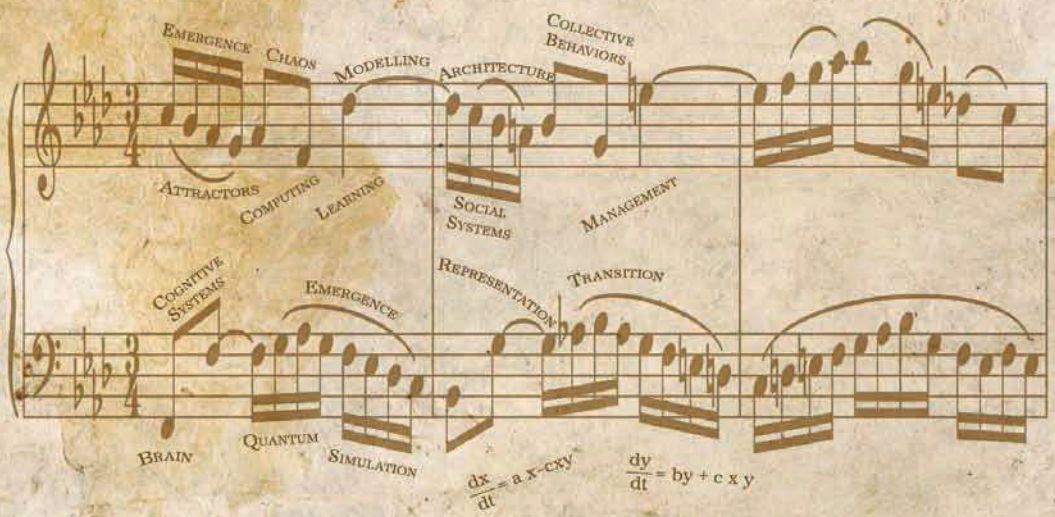
Systemics of Emergence

Research and Development



Edited by
Gianfranco Minati, Eliano Pessa
and Mario Abram

Springer

SYSTEMICS OF EMERGENCE: RESEARCH AND DEVELOPMENT

SYSTEMICS OF EMERGENCE: RESEARCH AND DEVELOPMENT

Edited by

Gianfranco Minati[1], Eliano Pessa[2] and Mario Abram[1]
[1]Italian Systems Society, Milano, Italy
[2]University of Pavia, Pavia, Italy

Springer

Gianfranco Minati
Italian Systems Society
Milan, Italy

Eliano Pessa
University of Pavia
Pavia, Italy

Mario Abram
Italian Systems Society
Milan, Italy

Library of Congress Cataloging-in-Publication Data

ISBN-10: 0-387-28899-6 (HB) ISBN-13: 978-0387-28899-4 (HB)
ISBN-10: 0-387-28898-8 (e-book) ISBN-13: 978-0387-28898-7 (e-book)

© 2006 by Springer Science+Business Media, Inc.
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science + Business Media, Inc., 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now know or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks and similar terms, even if the are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America. Printed on acid-free paper.

9 8 7 6 5 4 3 2 1 SPIN 11552154

springeronline.com

Contents

Program Committee xi

Contributing Authors xiii

Preface xv

Acknowledgments xix

OPENING LECTURE 1

Uncertainty and Information: Emergence of Vast New Territories 3
G. J. KLIR

APPLICATIONS 29

Complexity in Universe Dynamic Evolution. Part 1 - Present state and future evolution 31
U. DI CAPRIO

Complexity in Universe Dynamic Evolution. Part 2 - Preceding history 51
U. DI CAPRIO

Mistake Making Machines 67
G. MINATI AND G. VITIELLO

Explicit Velocity for Modelling Surface Complex Flows with Cellular Automata and Applications 79
M. V. AVOLIO, G. M. CRISCI, D. D'AMBROSIO, S. DI GREGORIO, G. IOVINE, V. LUPIANO, R. RONGO, W. SPATARO AND G. A. TRUNFIO

Analysis of Fingerprints Through a Reactive Agent 93
A. MONTESANTO, G. TASCINI, P. BALDASSARRI AND L. SANTINELLI

User Centered Portal Design: A Case Study in Web Usability 105
M. P. PENNA, V. STARA AND D. COSTENARO

BIOLOGY AND HUMAN CARE 115

Logic and Context in Schizophrenia 117
P. L. BANDINELLI, C. PALMA, M. P. PENNA AND E. PESSA

The "Hope Capacity" in the Care Process and the Patient-Physician Relationship 133
A. RICCIUTI

Puntonet 2003. A multidisciplinary and Systemic Approach in Training Disabled People Within the Experience of Villa S. Ignazio 147
D. FORTIN, V. DURINI AND M. NARDON

Intelligence and Complexity Management: From Physiology to Pathology. Experimental Evidences and Theoretical Models 155
P. L. MARCONI

Disablement, Assistive Technologies and Computer Accessibility: Hints of Analysis Through a Clinical Approach Based on the ICF Model 169
C. MASALA AND D. R. PETRETTO

Chaos and Cultural Fashions 179
S. BENVENUTO

COGNITIVE SCIENCE 191

Personality and Complex Systems. An Expanded View 193
M. MELEDDU AND L. F. SCALAS

Complexity and Paternalism 207
P. RAMAZZOTTI

A Computational Model of Face Perception 223
M. P. PENNA, V. STARA, M. BOI AND P. PULITI

The Neon Color Spreading and the Watercolor Illusion:
Phenomenal Links and Neural Mechanisms 235
B. PINNA

Usability and Man-Machine Interaction 255
M. P. PENNA AND R. RANI

Old Maps and the Watercolor Illusion: Cartography, Vision Science
and Figure-Ground Segregation Principles 261
B. PINNA AND G. MARIOTTI

EMERGENCE 279

Autopoiesis and Emergence 281
L. BICH

Typical Emergencies in Electric Power Systems 293
U. DI CAPRIO

Strategies of Adaptation of Man to his Environment:
Projection Outide the Human Body of Social Institutions 311
E. A. NUNEZ

Emergence of the Cooperation-Competition Between Two Robots 317
G. TASCINI AND A. MONTESANTO

Overcoming Computationalism in Cognitive Science 341
M. P. PENNA

Physical and Biological Emergence: Are They Different? 355
E. PESSA

GENERAL SYSTEMS 375

Interactions Between Systems 377
M. R. ABRAM

Towards a Systemic Approach to Architecture 391
V. DI BATTISTA

Music, Emergence and Pedagogical Process 399
E. PIETROCINI

Intrinsic Uncertainty in the Study of Complex Systems:
The Case of Choice of Academic Career 417
M. S. FERRETTI AND E. PESSA

A Model of Hypertextual Structure and Organization 427
M. P. PENNA, V. STARA, D. COSTENARO AND P. PULITI

LEARNING 435

Teachers in the Technological Age: A Comparison Between
Traditional and Hypertextual Instructional Strategies 437
M. P. PENNA, V. STARA AND D. COSTENARO

The Emergence of E.Learning 447
M. P. PENNA, V. STARA AND P. PULITI

Spatial Learning in Children 453
B. LAI, M. P. PENNA AND V. STARA

MANAGEMENT 461

Dynamics of Strategy: a Feedback Approach to
Corporate Strategy-Making 463
V. CODA AND E. MOLLONA

A Cognitive Approach to Organizational Complexity 495
G. FIORETTI AND B. VISSER

Normative Commitment to the Organization, Support
and Self Competence 515
A. BATTISTELLI, M. MARIANI AND B. BELLÒ

A Multivariate Contribution to the Study of Mobbing,
Using the QAM 1.5 Questionnaire 527
P. ARGENTERO AND N. S. BONFIGLIO

Representation in Psychometrics: Confirmatory Factor Models of Job Satisfaction in a Group of Professional Staff 535
M. S. FERRETTI AND P. ARGENTERO

SOCIAL SYSTEMS 549

The Impact of Email on System Identity and Autonomy: A Case Study in Self-Observation 551
L. BIGGIERO

Some Comments on Democracy and Manipulating Consent in Western Post-Democratic Societies 569
G. MINATI

Metasystem Transitions and Sustainability in Human Organizations. Part 1 - Towards Organizational Synergetics 585
G. TERENZI

Metasystem Transitions and Sustainability in Human Organizations. Part 2 - A Heuristics for Global Sustainability 601
G. TERENZI

SYSTEMIC APPROACH AND INFORMATION SCIENCE 613

Scale Free Graphs in Dynamic Knowledge Acquisition 615
I. LICATA, G. TASCINI, L. LELLA, A. MONTESANTO AND W. GIORDANO

Recent Results on Random Boolean Networks 625
R. SERRA AND M. VILLANI

Color-Oriented Content Based Image Retrieval 635
G. TASCINI, A. MONTESANTO AND P. PULITI

THEORETICAL ISSUES IN SYSTEMICS 651

Uncertainty and the Role of the Observer 653
G. BRUNO, G. MINATI AND A. TROTTA

Towards a Second Systemics 667
G. MINATI

Is Being Computational an Intrinsic Property of a Dynamical System? 683
M. GIUNTI

The Origin of Analogies in Physics 695
E. TONTI

Prisoner Dilemma: A Model Taking into Account Expectancies 707
N. S. BONFIGLIO AND E. PESSA

The Theory of Levels of Reality and the Difference Between Simple and Tangled Hierarchies 715
R. POLI

General System Theory, Like-Quantum Semantics and Fuzzy Sets 723
I. LICATA

About the Possibility of a Cartesian Theory Upon Systems, Information and Control 735
P. ROCCHI

Program Committee

G. Minati (chairman)
Italian Systems Society

E. Pessa (co-chairman)
University of Pavia

G. Bruno
University "La Sapienza", Rome

S. Di Gregorio
University of Calabria

M. P. Penna
University of Cagliari

R. Serra
University of Modena and Reggio Emilia

G. Tascini
University of Ancona

Contributing Authors

Abram M. R.	AIRS, Milano, Italy
Argentero P.	Università degli Studi di Pavia, Italy
Avolio M. V.	Università degli Studi di Calabria, Rende (CS), Italy
Baldassarri P.	Università Politecnica delle Marche, Ancona, Italy
Bandinelli P. L.	ASL Roma "E", Roma, Italy
Battistelli A.	Università degli Studi di Verona, Italy
Bellò B.	Università degli Studi di Verona, Italy
Benvenuto S.	CNR, Roma, Italy
Bich L.	Università degli Studi di Pavia, Italy
Biggiero L.	Università dell'Aquila, Roio Poggio (AQ), Italy
Boi M.	Università degli Studi di Cagliari, Italy
Bonfiglio N. S.	Università degli Studi di Pavia, Italy
Bruno G.	Università "La Sapienza", Roma, Italy
Coda V.	Università Commerciale "L. Bocconi", Milano, Italy
Costenaro D.	Università degli Studi di Cagliari, Italy
Crisci G. M.	Università degli Studi di Calabria, Rende (CS), Italy
D'Ambrosio D.	Università degli Studi di Calabria, Rende (CS), Italy
Di Battista V.	Politecnico di Milano, Italy
Di Caprio U.	Stability Analysis s.r.l., Milano, Italy
Di Gregorio S.	Università degli Studi di Calabria, Rende (CS), Italy
Durini V.	Villa S. Ignazio, Trento, Italy
Ferretti M. S.	Università degli Studi di Pavia, Italy
Fioretti G.	Università degli Studi di Bologna, Italy
Fortin D.	Villa S. Ignazio, Trento, Italy
Giordano W.	Università Politecnica delle Marche, Ancona, Italy
Giunti M.	Università degli Studi di Cagliari, Italy

Iovine G.	CNR-IRPI, Rende (CS), Italy
Klir G. J.	State University of New York, Binghamton, NY
Lai B.	Università degli Studi di Cagliari, Italy
Lella L.	Università Politecnica delle Marche, Ancona, Italy
Licata I.	ICNLSC, Marsala (TP), Italy
Lupiano V.	CNR-IRPI, Rende (CS), Italy
Marconi P. L.	ARTEMIS Neuropsichiatrica, Roma, Italy
Mariani M.	Università degli Studi di Bologna, Italy
Mariotti G.	Università degli Studi di Sassari, Italy
Masala C.	Università degli Studi di Cagliari, Italy
Meleddu M.	Università degli Studi di Cagliari, Italy
Minati G.	AIRS, Milano, Italy
Mollona E.	Università degli Studi di Bologna, Italy
Montesanto A.	Università Politecnica delle Marche, Ancona, Italy
Nardon M.	Villa S. Ignazio, Trento, Italy
Nunez E. A.	AFSCET, France
Palma C.	Istituto d'Istruzione Superiore, Roma, Italy
Penna M. P.	Università degli Studi di Cagliari, Italy
Pessa E.	Università degli Studi di Pavia, Italy
Petretto D. R.	Università degli Studi di Cagliari, Italy
Pietrocini E.	Accademia Angelica Costantiniana, Roma, Italy
Pinna B.	Università degli Studi di Sassari, Italy
Poli R.	Università degli Studi di Trento, Italy
Puliti P.	Università Politecnica delle Marche, Ancona, Italy
Ramazzotti P.	Università degli Studi di Macerata, Italy
Rani R.	Università degli Studi di Cagliari, Italy
Ricciuti A.	Attivecomeprima Onlus, Milano, Italy
Rocchi P.	IBM, Roma, Italy
Rongo R.	Università degli Studi di Calabria, Rende (CS), Italy
Santinelli L.	Università Politecnica delle Marche, Ancona, Italy
Scalas L. F.	Università degli Studi di Cagliari, Italy
Serra R.	CRSA Fenice, Marina di Ravenna, Italy
Spataro W.	Università degli Studi di Calabria, Rende (CS), Italy
Stara V.	Università Politecnica delle Marche, Ancona, Italy
Tascini G.	Università Politecnica delle Marche, Ancona, Italy
Terenzi G.	ATESS, Frosinone, Italy
Tonti E.	Università degli Studi di Trieste, Italy
Trotta A.	ITC "Emanuela Loi", Nettuno (RM), Italy
Trunfio G. A.	Università degli Studi di Calabria, Rende (CS), Italy
Villani M.	CRSA Fenice, Marina di Ravenna, Italy
Visser B.	Erasmus University, Rotterdam, The Netherland
Vitiello G.	Università di Salerno, Baronissi (SA), Italy

Preface

The systems movement is facing some new important scientific and cultural processes tacking place in disciplinary research and effecting Systems Research.

Current Systems Research is mainly carried out disciplinarily, through a disciplinary usage of the concept of system with no or little processes of generalization.

Generalization takes place by assuming:

- inter-disciplinary (when same systemic properties are considered in different disciplines), and
- trans-disciplinary (when considering systemic properties per se and relationships between them) approaches.

Because of the nature of the problems, of the research organization, and for effectiveness, research is carried out by using local inter-disciplinarily approaches, i.e. between adjacent disciplines using very similar languages and models.

General Systems Research, i.e. the process of globally inter- and trans-disciplinarizing, is usually lacking. General systems scientists are expected to perform disciplinary and locally inter-disciplinary research by, moreover, carrying out generalizations. The establishing of a dichotomy between research and generalizing is, in our view, the key problem of systems research today.

Research without processes of generalization produces duplications inducing besides fragmentations often used for establishing markets of details, symptomatic remedies, based on the concept of system, but with no or poor understanding of the global picture, that's without generalizing.

The novelty is that emergence is the context where generalization, that's globally inter- and trans-disciplinarizing is not a choice, but a necessity.

As it is well known emergence is, in short, the rising of coherence among interacting elements, detected by an observer equipped with a suitable cognitive model at a level of description different from the one used for elements. Well-known exempla are collective behaviors establishing phenomena such as laser effect, superconductivity, swarms, flocks and traffic.

It has been possible to reduce General Systems Theory (GST), by dissecting from it the dynamic process of establishment and holding of systems, that's by removing or ignoring the processes of emergence. By considering systems and sub-systems as elements it is still assumed the mechanistic view based on the Cartesian idea that the microscopic world is simpler than the macroscopic and that the macroscopic world may be explained through an infinite knowledge of the microscopic.

However, daily thinking is often still based on the idea that not only the macro level of reality may be explained through the micro level, but that the macro level may be effectively managed by acting on the micro level. Assumption of manageability of the emergent level through elements comes from linearizing the fact that it is possible to destroy the upper level by destroying the micro.

Studying systems in the context of emergence doesn't allow the dissection mentioned above because models relate to how new, emergent properties are established rather than properties themselves only.

In the GST approach it has been possible to focus on (emergent) systemic properties (such as open, adaptive, anticipatory and chaotic systems), by considering their specificity and not reducibility to the ones of components. GST allowed description and representation of systemic properties, and adoption of systemic methodologies. This reminds in some way initial (i.e. Aristotelian) approaches to physics when the problem was to describe characteristics, essences, more than evolution.

Emergence focuses on the processes of establishing of systemic properties. The study of processes of emergence implies the study of inter- and trans-disciplinarity. Research on emergence allows for modeling and simulating processes of emergence, by using the calculus of emergence. Examples are the studies of Self-Organization, Collective-Behaviors, and Artificial Life. In short, Emergence studies the engine of GST, while GST allowed focusing on results. Models of emergence relate, for instance to phase transitions, synergetic effects, dissipative structures, conceptually inducing inter- and trans- disciplinary research.

Because of its nature emergence is inter- and trans-disciplinary.

Paradoxically, this kind of research is currently made not by established systems societies, but by "new" systems institutions, like, just to mention a couple, the Santa Fe Institute (SFI), the New England Complex Systems Institute (NECSI), the Institute for the Study of Coherence and Emergence (ISCE), and in many conferences organized world-wide

General Systems Research is now research on emergence. As it is well known emergence refers to the core theoretical problems of the processes from which systems are established, as implicitly introduced in Von Bertalanffy's General Systems Theory by considering the crucial role of the observer, together with its cognitive system and cognitive models.

Emergence is not intended as a process taking place in the domain of any discipline, but as "trans-disciplinary modeling" meaningful for any discipline. We are now facing the process by which the General Systems Theory is more and more becoming a Theory of Emergence, seeking suitable models and formalizations of its fundamental bases. Correspondingly, we need to envisage and prepare for the establishment of a Second Systemics -a Systemics of Emergence- relating to new crucial issues such as, for instance:

- Collective Phenomena;
- Phase Transitions, such as in physics (e.g. transition from solid to liquid) and in learning processes;
- Dynamical Usage of Models (DYSAM);
- Multiple systems, emerging from identical components but simultaneously exhibiting different interactions among them;
- Uncertainty Principles;
- Modeling emergence;
- Systemic meaning of new theorizations such as Quantum Field Theory (QFT) and related applications (e.g. biology, brain, consciousness, dealing with long-range correlations).

We need to specify that in literature it is used, even if not rigorously defined, the term Systemics intended as a cultural generalization of the principles contained in the General Systems Theory. We may say that, in short, General Systems Theory refers to systemic properties considered in different disciplinary contexts (inter-disciplinarity) and per se in general (trans-disciplinarity); disciplinary applications; and theory of emergence.

More generally the term Systemic Approach refers to the general methodological aspects of GST. In base of that a problem is considered by identifying interactions, levels of description (micro, macro, and mesoscopic levels), processes of emergence and role of the observer (cognitive model).

At an even higher level of generalization, Systemics is intended as cultural extension, corpus of concepts, principles, applications and

methodology based on using concepts of interaction, system, emergence, inter- and trans-disciplinarity.

Because of what introduced above, Systemics should refer to the principles, approaches and models of emergence and complexity, by generalizing them and not popularizing or introducing metaphorical usages.

The Systems Community urges to collectively know and use such principles and approaches, by accepting to found inter- and trans-disciplinary activity on disciplinary knowledge. Trans-disciplinarity doesn't mean to leave aside disciplinary research, but apply systemic, general (disciplinary-independent) principles and approaches realized in disciplinary research. Focusing on systems is not anymore so innovative. Focusing on emergence is not anymore so new.

What's peculiar, specific of our community? In our view, trans- and global inter-disciplinary research, implemented as cultural values.

The third national conference of the Italian Systems Society (AIRS) focused on emergence as the key point of any systemic processes. The conference dealt up with this problem by different disciplinary approaches, very well indicated by the organization in sessions:

1. Applications.
2. Biology and human care.
3. Cognitive Science.
4. Emergence.
5. General Systems.
6. Learning .
7. Management.
8. Social systems.
9. Systemic approach and Information Science.
10. Theoretical issues in Systemics.

We conclude hoping that the systemic research will continuously accept the general challenge previously introduced and contained in the paper presented. This acceptance is a duty for the systems movement when reminding the works of the founding fathers.

The Italian Systems Society is trying to play a significant role in this process.

Gianfranco Minati, AIRS president

Eliano Pessa, Co-Editor

Mario Abram, Co-Editor

Acknowledgments

The third Italian Conference on Systemics has been possible thanks to the contributions of many people that have accompanied and supported the growth and development of AIRS during all the years since its establishment in 1985 and to the contribution of "new" energies. The term "new" refers both to the involvement of students and to the involvement and contribution of researchers realizing the systemic aspect of their activity.

We have been honoured by the presence of Professor George Klir and by his opening lecture for this conference.

We thank the Castel Ivano Association for hosting this conference and we particularly thank Professor Staudacher, a continuous reference point for the high level cultural activities in the area enlightened by his beautiful castle.

We thank the Provincia Autonoma of Trento for supporting the conference and the University of Trento, the Italian Association for Artificial Intelligence for culturally sponsoring the conference.

We thank all the authors who submitted papers for this conference and in particular the members of the program committee and the referees who have guaranteed the quality of the event.

We thank explicitly all the people that have contributed and will contribute during the conference, bringing ideas and stimuli to the cultural project of Systemics.

G. Minati, E. Pessa, M. Abram

OPENING LECTURE

UNCERTAINTY AND INFORMATION: EMERGENCE OF VAST NEW TERRITORIES

George J. Klir
Department of Systems Science & Industrial Engineering,
Thomas J. Watson School of Engineering and Applied Science,
State University of New York, Binghamton, New York 13902-6000, U.S.A.

Abstract: A research program whose objective is to study uncertainty and uncertainty-based information in all their manifestations was introduced in the early 1990's under the name "generalized information theory" (GIT). This research program, motivated primarily by some fundamental methodological issues emerging from the study of complex systems, is based on a two-dimensional expansion of classical, probability-based information theory. In one dimension, additive probability measures, which are inherent in classical information theory, are expanded to various types of nonadditive measures. In the other dimension, the formalized language of classical set theory, within which probability measures are formalized, is expanded to more expressive formalized languages that are based on fuzzy sets of various types. As in classical information theory, uncertainty is the primary concept in GIT and information is defined in terms of uncertainty reduction. This restricted interpretation of the concept of information is described in GIT by the qualified term "uncertainty-based information". Each uncertainty theory that is recognizable within the expanded framework is characterized by: (i) a particular formalized language (a theory of fuzzy sets of some particular type); and (ii) a generalized measure of some particular type (additive or nonadditive). The number of possible uncertainty theories is thus equal to the product of the number of recognized types of fuzzy sets and the number of recognized types of generalized measures. This number has been growing quite rapidly with the recent developments in both fuzzy set theory and the theory of generalized measures. In order to fully develop any of these theories of uncertainty requires that issues at each of the following four levels be adequately addressed: (i) the theory must be formalized in terms of appropriate axioms; (ii) a calculus of the theory must be developed by which the formalized uncertainty is manipulated within the theory; (iii) a justifiable way of measuring the amount of relevant uncertainty (predictive, diagnostic, etc.) in any situation formalizable in the theory must be found; and (iv) various methodological aspects of the theory must be developed. Among the many

uncertainty theories that are possible within the expanded conceptual framework, only a few theories have been sufficiently developed so far. By and large, these are theories based on various types of generalized measures, which are formalized in the language of classical set theory. Fuzzification of these theories, which can be done in different ways, has been explored only to some degree and only for standard fuzzy sets. One important result of research in the area of GIT is that the tremendous diversity of uncertainty theories made possible by the expanded framework is made tractable due to some key properties of these theories that are invariant across the whole spectrum or, at least, within broad classes of uncertainty theories. One important class of uncertainty theories consists of theories that are viewed as theories of imprecise probabilities. Some of these theories are based on Choquet capacities of various orders, especially capacities of order infinity (the well known theory of evidence), interval-valued probability distributions, and Sugeno λ-measures. While these theories are distinct in many respects, they share several common representations, such as representation by lower and upper probabilities, convex sets of probability distributions, and so-called Möbius representation. These representations are uniquely convertible to one another, and each may be used as needed. Another unifying feature of the various theories of imprecise probabilities is that two types of uncertainty coexist in each of them. These are usually referred to as nonspecificity and conflict. It is significant that well-justified measures of these two types of uncertainty are expressed by functionals of the same form in all the investigated theories of imprecise probabilities, even though these functionals are subject to different calculi in different theories. Moreover, equations that express relationship between marginal, joint, and conditional measures of uncertainty are invariant across the whole spectrum of theories of imprecise probabilities. The tremendous diversity of possible uncertainty theories is thus compensated by their many commonalities.

Key words: uncertainty theories; fuzzy sets; information theories; generalized measures; imprecise probabilities.

1. GENERALIZED INFORMATION THEORY

A research program whose objective is to study uncertainty and uncertainty-based information in all their manifestations was introduced in the early 1990's under the name "generalized information theory" (GIT) (Klir, 1991). This research program, motivated primarily by some fundamental methodological issues emerging from the study of complex systems, is based on a two-dimensional expansion of classical, probability-based information theory. In one dimension, additive probability measures, which are inherent in classical information theory, are expanded to various types of nonadditive measures. In the other dimension, the formalized language of classical set theory, within which probability measures are

formalized, is expanded to more expressive formalized languages that are based on fuzzy sets of various types. As in classical information theory, uncertainty is the primary concept in GIT and information is defined in terms of uncertainty reduction. This restricted interpretation of the concept of information is described in GIT by the qualified term "uncertainty-based information."

Each uncertainty theory that is recognizable within the expanded framework is characterized by: (i) a particular *formalized language* (a theory of fuzzy sets of some particular type); and (ii) *generalized measures* of some particular type (additive or nonadditive). The number of possible uncertainty theories is thus equal to the product of the number of recognized types of fuzzy sets and the number of recognized types of generalized measures. This number has been growing quite rapidly with the recent developments in both fuzzy set theory and the theory of generalized measures. In order to fully develop any of these theories of uncertainty requires that issues at each of the following four levels be adequately addressed: (a) the theory must be formalized in terms of appropriate *axioms*; (b) a *calculus* of the theory must be developed by which the formalized uncertainty is manipulated within the theory; (iii) a justifiable way of *measuring* the amount of relevant uncertainty (predictive, diagnostic, etc.) in any situation formalizable in the theory must be found; and (iv) various methodological aspects of the theory must be developed.

Among the many uncertainty theories that are possible within the expanded conceptual framework, only a few theories have been sufficiently developed so far. By and large, these are theories based on various types of generalized measures, which are formalized in the language of classical set theory. Fuzzification of these theories, which can be done in different ways, has been explored only to some degree and only for standard fuzzy sets. One important result of research in the area of GIT is that the tremendous diversity of uncertainty theories emerging from the expanded framework is made tractable due to some key properties of these theories that are invariant across the whole spectrum or, at least, within broad classes of uncertainty theories.

One important class of uncertainty theories consists of theories that are viewed as theories of imprecise probabilities. Some of these theories are based on Choquet capacities of various orders (Choquet, 1953-54), especially capacities of order infinity (the well known theory of evidence) (Shafer, 1976), interval-valued probability distributions (Pan and Klir, 1997), and Sugeno λ-measures (Wang and Klir, 1992). While these theories are distinct in many respects, they share several common representations, such as representations by lower and upper probabilities, convex sets of probability distributions, and so-called Möbius representation. All

representations in this class are uniquely convertible to one another, and each may be used as needed.

Another unifying feature of the various theories of imprecise probabilities is that two types of uncertainty coexist in each of them. These are usually referred to as *nonspecificity* and *conflict*. It is significant that well-justified measures of these two types of uncertainty are expressed by functionals of the same form in all the investigated theories of imprecise probabilities, even though these functionals are subject to different calculi in different theories. Moreover, equations that express relationship between marginal, joint, and conditional measures of uncertainty are invariant across the whole spectrum of theories of imprecise probabilities. The tremendous diversity of possible uncertainty theories is thus compensated by their many commonalities.

Uncertainty-based information does not capture the rich notion of information in human communication and cognition, but it is very useful in dealing with systems. Given a particular system, it is useful, for example, to measure the amount of information contained in the answer given by the system to a relevant question (concerning various predictions, retrodictions, diagnoses etc.). This can be done by taking the difference between the amount of uncertainty in the requested answer obtained within the experimental frame of the system (Klir, 2001a) in the face of total ignorance and the amount of uncertainty in the answer obtained by the system. This can be written concisely as

$$\begin{aligned}\text{Information}\,(A_S \mid S,Q) = {} & \text{Uncertainty}\,(A_{EF_S} \mid EF_S,Q) \\ & - \text{Uncertainty}\,(A_S \mid S,Q)\end{aligned}$$

where

- S denotes a given system
- EF_S denoted the experimental frame of system S
- Q denotes a given question
- A_{EF_S} denotes the answer to question Q obtained solely within the experimental frame EF_S
- A_S denotes the answer to question Q obtained by system S.

This allows us to compare information contents of systems constructed within the same experimental frame with respect to questions of our interest.

The purpose of this paper is to present a brief overview of GIT. A comprehensive presentation of GIT is given in a forthcoming book by Klir (2005).

2. CLASSICAL ROOTS OF GIT

There are two classical theories of uncertainty-based information, both formalized in terms of classical set theory. The older one, which is also simpler and more fundamental, is based on the notion of *possibility*. The newer one, which has been considerably more visible, is based on the notion of *probability*.

2.1 Classical Possibility-Based Uncertainty Theory

To describe the possibility-based uncertainty theory, let X denote a finite set of mutually exclusive alternatives that are of our concern (diagnoses, predictions, etc). This means that in any given situation only one of the alternatives is true. To identify the true alternative, we need to obtain relevant information (e.g. by conducting relevant diagnostic tests). The most elementary and, at the same time, the most fundamental kind of information is a demonstration (based, for example, on outcomes of the conducted diagnostic tests) that some of the alternatives in X are not possible. After excluding these alternatives from X, we obtain a subset E of X. This subset contains only alternatives that, according to the obtained information are *possible*. We may say that alternatives in E are supported by evidence.

To formalize evidence expressed in this form, the characteristic function of set E, r_E, is viewed as a *basic possibility function*. Clearly, for each $x \in X$, $r_E(x) = 1$ when x is possible and $r_E(x) = 0$ when x is not possible. Possibility function applicable to all subsets of X, Pos_E, is then defined by the formula

$$Pos_E(A) = \max_{x \in A} r_E(x) \tag{1}$$

for all $A \subseteq X$. It is indeed correct to say that it is possible that the true alternative is in set A when A contains at least one alternative that is also contained in set E.

Given a possibility function Pos_E on the power set of X, it is useful to define another function, Nec_E, to describe for each $A \subseteq X$ the necessity that the true alternative is in A. Clearly, the true alternative is necessarily in A if and only if it is not possible that it is in $\overline{A}$, the complement of A. Hence,

$$Nec_E(A) = 1 - Pos_E(\overline{A}) \tag{2}$$

for all $A \subseteq X$.

The question of how to measure the amount of uncertainty associated with a finite set E of possible alternatives was addressed by Hartley (1928).

He showed that the only meaningful way to measure this amount is to use a functional of the form

$$c \log_b \sum_{x \in X} r_E(x)$$

or, alternatively, $c \log_b |E|$ where $|E|$ denotes the cardinality of E, and b and c are positive constants. Each choice of values b and c determines the unit in which the uncertainty is measured. Requiring, for example, that

$$c \log_b 2 = 1,$$

which is the most common choice, uncertainty would be measured in *bits*. One bit of uncertainty is equivalent to uncertainty regarding the truth or falsity of one elementary proposition. Choosing conveniently $b = 2$ and $c = 1$ to satisfy the above equation, we obtain a unique functional, H, defined for any possibility function, Pos_E, by the formula

$$H(Pos_E) = \log_2 |E|. \tag{3}$$

This functional is usually called a *Hartley measure* of uncertainty. Its uniqueness was later proven on axiomatic grounds by Rènyi (1970).

Observe that the Hartley measure satisfies the inequalities

$$0 \leq H(Pos_E) \leq \log_2 |X|$$

for any $E \subseteq X$ and that the amount of information, $I(Pos_E)$, in evidence expressed by function Pos_E is given by the formula

$$I(Pos_E) = \log_2 |X| - \log_2 |E|. \tag{4}$$

It follows from the Hartley measure that uncertainty associated with sets of possible alternatives results from the lack of specificity. Large sets result in less specific predictions, diagnoses, etc., than their smaller counterparts. Full specificity is obtained when only one alternative is possible. This type of uncertainty is thus well characterized by the term *nonspecificity*.

Consider now two universal sets, X and Y, and assume that a relation $R \subseteq X \times Y$ describes a set of possible alternatives in some situation of interest. Consider further the sets

$$R_X = \{x \in X \mid (x,y) \in R \text{ for some } y \in Y\}$$
$$R_Y = \{y \in Y \mid (x,y) \in R \text{ for some } x \in X\}$$

which are usually referred to as *projections* of R on sets X, Y, respectively. Then three distinct Hartley measures are applicable, which are defined on the power sets of X, Y, and $X \times Y$. To identify clearly which universal set is involved in each case, it is useful (and a common practice) to write $H(X)$, $H(Y)$, $H(X,Y)$ instead of

$$H(Pos_{R_X}), \quad H(Pos_{R_Y}), \quad H(Pos_R)$$

respectively. Functionals $H(X)=\log_2 |R_X|$ and $H(Y)=\log_2 |R_Y|$ are called *simple* (or *marginal*) *Hartley measures*, while functional $H(X,Y)=\log_2|R|$ is called a *joint Hartley measure*.

Two additional functionals are defined,

$$H(X \mid Y) = \log_2 \frac{|R|}{|R_Y|} \quad \text{and} \quad H(Y \mid X) = \log_2 \frac{|R|}{|R_X|}, \tag{5}$$

which are called *conditional Hartley measures*. Observe that the ratio $|R|/|R_Y|$ in $H(X \mid Y)$ represents the average number of elements of X that are possible alternatives under the condition that a possible element of Y is known. This means that $H(X \mid Y)$ measures the average nonspecificity regarding possible choices from X for all possible choices from Y. Function $H(X \mid Y)$ has clearly a similar meaning with the roles of sets X and Y exchanged. Observe also that the conditional Hartley measures can be expressed in terms of the joint measures and the two simple Hartley measures:

$$H(X \mid Y) = H(X,Y) - H(Y) \text{ and } H(Y \mid X) = H(X,Y) - H(X). \tag{6}$$

If possible alternatives from X do not depend on selections from Y, and visa versa, then $R = X \times Y$ and the sets R_X and R_Y are called *noninteractive*. Then, clearly,

$$H(X \mid Y) = H(X) \quad \text{and} \quad H(Y \mid X) = H(Y), \tag{7}$$
$$H(X,Y) = H(X) + H(Y). \tag{8}$$

In all other cases, when sets R_X and R_Y are *interactive*, these equations become the inequalities

$$H(X \mid Y) < H(X) \quad \text{and} \quad H(Y \mid X) < H(Y), \tag{9}$$
$$H(X,Y) < H(X) + H(Y). \tag{10}$$

In addition, the functional

$$T_H(X,Y) = H(X) + H(Y) - H(X,Y), \tag{11}$$

which is usually called an *information transmission*, is a useful indicator of the strength of constraint between sets R_X and R_Y.

The Hartley measure is applicable only to finite sets. Its counterpart for subsets of the n-dimensional Euclidean space E^n ($n \geq 1$) was not available in the classical possibility-based uncertainty theory. It was eventually suggested (as a byproduct of research on GIT) by Klir and Yuan (1995b) in terms of the functional

$$HL(Pos_E) = \min_{t \in T} c \log_b \left[\prod_{i=1}^{n} (1 + \mu(E_{i_t})) + \mu(E) - \prod_{i=1}^{n} \mu(E_{i_t}) \right],$$

where E, T, E_{it}, and μ denote, respectively, a convex subset of E^n, the set of all isometric transformations from one orthogonal coordinate system to another, the i-th projection of E in coordinate system t, and the Lebesgue measure, and b and c are positive constants whose choice defines a measurement unit. This functional, which is usually referred to as *Hartley-like measure*, was proven to satisfy all mathematical properties that such a measure is expected to satisfy (Klir and Wierman, 1999; Ramer and Padet, 2001).

Let a measurement unit for the Hartley-like measure be defined by the requirement that $HL(Pos_E)=1$ when E is a closed interval of real numbers of length 1 in some assumed unit of length. That is, we require that

$$c \log_b 2 = 1.$$

It is convenient to choose $c = 1$ and $b = 2$ to satisfy this equation. Then,

$$HL(Pos_E) = \min_{t \in T} c \log_2 \left[\prod_{i=1}^{n} (1 + \mu(E_{i_t})) + \mu(E) - \prod_{i=1}^{n} \mu(E_{i_t}) \right]. \tag{12}$$

In these units, a unit square has uncertainty 2, a unit cube has uncertainty 3, etc. For any universal set X (a convex subset of E^n), a normalized Hartley-like measure, NHL is defined for each convex subset E of X by the formula

$$NHL(Pos_E) = \frac{HL(Pos_E)}{HL(Pos_X)}. \tag{13}$$

Clearly, NHL is independent of the chosen unit and $0 \leq NHL(Pos_E) \leq 1$.

2.2 Classical Probability-Based Uncertainty Theory

The second classical uncertainty theory is based on the notion of classical probability measure (Halmos, 1950). As is well known, the amount of uncertainty in evidence expressed by a probability distribution, p, on a finite set X of considered alternatives is measured (in bits) by the functional

$$S(p(x) \mid x \in X) = -\sum_{x \in X} p(x) \log_2 p(x). \tag{14}$$

This functional was introduced by Shannon (1948) and it is usually referred to as *Shannon entropy*. Its uniqueness has been proven in numerous ways (Klir and Wierman, 1999).

For probabilities on $X \times Y$, three types of Shannon entropies are recognized: joint, marginal, and conditional. A simplified notation to distinguish them is commonly used in the literature: $S(X)$ instead of $S(p(x)|x \in X)$, $S(X,Y)$ instead of $S(p(x,y)\,|x \in X, y \in Y)$, etc. Conditional Shannon entropies are defined in terms of weighted averages of local conditional entropies as

$$S(X \mid Y) = -\sum_{y \in Y} p_Y(y) \sum_{x \in X} p(x \mid y) \log_2 p(x \mid y), \tag{15}$$

$$S(Y \mid X) = -\sum_{x \in X} p_X(x) \sum_{y \in Y} p(y \mid x) \log_2 p(y \mid x). \tag{16}$$

As is well known (Klir and Wierman, 1999), equations and inequalities (2) - (6) for the Hartley measure have their exact counterparts for the Shannon entropy. For example, the counterparts of (2) are the equations

$$S(X \mid Y) = S(X,Y) - S(Y) \text{ and } S(Y \mid X) = S(X,Y) - S(X). \tag{17}$$

Moreover,

$$T_S(X,Y) = S(X) + S(Y) - S(X,Y) \tag{18}$$

is the *probabilistic information transmission* (the probabilistic counterpart of (7)).

It is obvious that the Shannon entropy is applicable only to finite sets of alternatives. At first sight, it seems suggestive to extend it to probability density functions, q, on E (or, more generally, on E^n, $n \geq 1$), by replacing in Eq.(14) p with q and the summation with integration. However, there are several reasons why the resulting functional does not qualify as a measure of uncertainty: (i) it may be negative; (ii) it may be infinitely large; (iii) it depends on the chosen coordinate system; and most importantly, (iv) the limit of the sequence of its increasingly more refined discrete approximations diverges (Klir and Wierman, 1999). These problems can be overcome by the modified functional

$$S'[q(x), q'(x) \mid x \in \mathsf{E}] = \int_{\mathsf{R}} q(x) \log_2 \frac{q(x)}{q'(x)} dx, \tag{19}$$

which involves two probability density functions, q and q'. Uncertainty is measured by S' in relative rather than absolute terms.

When q in (19) is a joint probability density function on E^2 and q' is the product of the two marginals of q, we obtain the information transmission

$$\begin{aligned} &T_S[q(x,y), q_X(x), q_Y(y) \mid x \in \mathsf{E}, y \in \mathsf{E}] \\ &= \int_{\mathsf{E}} \int_{\mathsf{E}} q(x,y) \log_2 \frac{q(x,y)}{q_X(x) \cdot q_Y(y)} dxdy. \end{aligned} \tag{20}$$

This means that (20) is a direct counterpart of (18).

2.3 Relationship Between the Classical Uncertainty Theories

It is fair to say that the probability-based uncertainty theory has been far more visible than the one based on possibility. Moreover, the distinction between the two classical theories was concealed in the vast literature on probability-based uncertainty theory, where the Hartley measure is almost routinely viewed as a special case of the Shannon entropy. This view, which is likely a result of the fact that the value of the Shannon entropy for the

uniform probability distribution on some set is equal to the value of the Hartley measure for the same set, is ill-conceived. Indeed, the Hartley measure is totally independent of any probabilistic assumptions. Furthermore, given evidence expressed by a possibility function Pos_E, any probability measure *Pro*, not only the one representing the uniform distribution, is consistent with Pos_E when

$$Pro(a) \leq Pos_E(A) \tag{21}$$

for all $A \subseteq X$. Function Pos_E captures thus a set of probability measures and we have no basis for choosing any one of them.

The fundamental distinction between the two classical theories, both conceptual and formal, was correctly recognized by Kolmogorov (1965), who refers to the possibility-based uncertainty theory as "the combinatorial approach to the quantitative definition of information" and makes a relevant remark: "Discussions of information theory do not usually go to this combinatorial approach at any length, but I consider it important to emphasize its logical independence of probabilistic assumptions."

In spite of Kolmogorov's clear discussion of the fundamental distinctions between the two classical uncertainty theories, many information theorists still continue in their writings to subsume the Hartley measure under the Shannon entropy or they dismiss it altogether. One rare exception is Rènyi (1970), who developed an axiomatic characterization of the Hartley measure and proved its uniqueness.

It is shown in Sec. 2.1 that the type of uncertainty measured by the Hartley functional is well described by the term *nonspecificity*. One way of getting insight into the type of uncertainty measured by the Shannon entropy is to rewrite Eq. (14) as

$$S(p(x) \mid x \in X) = -\sum_{x \in X} p(x) \log_2 \left[1 - \sum_{y \neq x} p(y) \right]. \tag{22}$$

The term

$$Con(x) = \sum_{y \neq x} p(y)$$

in Eq. (22) represents clearly the total evidential claim pertaining to elements that are distinct from x. That is, $Con(x)$ expresses the sum of evidential

claims that fully conflict with the one focusing on x. Clearly, $Con(x) \in [0,1]$ for each $x \in X$. The function

$$CON(x) = -\log_2[1 - Con(x)]$$

which is employed in Eq. (22), is monotonic increasing with $Con(x)$ and extends its range from $[0,1]$ to $[0,\infty]$. Hence, it represents the same conflict in a different scale. The choice of the logarithmic function is a result of axiomatic requirements for the Shannon entropy.

It follows from these facts and from the form of Eq. (22) that the Shannon entropy is the mean (expected) value of *conflict* among evidential claims within a given probability distribution function. This is another demonstration that the two classical uncertainty theories deal with distinct types of uncertainty. It is significant that in every generalization of the classical theories both uncertainty types coexist.

3. MATHEMATICAL FRAMEWORK FOR GIT

GIT is based on (i) generalizing classical measure theory (Halmos, 1950) by abandoning the requirement of *additivity*; and (ii) generalizing classical set theory (Stoll, 1961) by abandoning the requirement that sets have *sharp boundaries*. The former generalization results in the theory of *generalized measures*, which are also commonly called *fuzzy measures* for some historical reasons of little significance (Denneberg, 1994; Pap, 1995; Wang and Klir, 1992). The latter generalization results in the *theory of fuzzy sets* (Klir and Yuan, 1995a, 1996; Yager et al., 1987). The following is an overview of these two generalizations.

3.1 Generalized Measure Theory

Classical measures are generalized by abandoning the requirement of additivity. However, for utilizing the generalized measures to represent uncertainty, it is essential to only replace additivity with a weaker requirement of monotonicity. The following is a formal definition of monotone measures, which represent one dimension of the mathematical framework for GIT.

Given a universal set X and a non-empty family C of subsets of X that contains $\varnothing$ and X and has an appropriate algebraic structure (σ-algebra, ample field, power set, etc.), a *regular monotone measure*, g, on $\langle X, \mathsf{C}\rangle$ is a function

$$g : \mathsf{C} \to [0,1]$$

that satisfies the following requirements:
(g1) $g(\varnothing) = 0$ and $g(X) = 1$ (boundary conditions);
(g2) for all $A, B \in \mathsf{C}$, if $A \subseteq B$, then $g(A) \leq g(B)$ (monotonicity);
(g3) for any increasing sequence $A_1 \subseteq A_2 \subseteq \ldots$ of sets in C,

$$\text{if } \bigcup_{i=1}^{\infty} A_i \in \mathsf{C}, \text{ then } \lim_{i\to\infty} g(A_i) = g\left(\bigcup_{i=1}^{\infty} A_i\right) \quad \text{(continuity from below);}$$

(g4) for any decreasing sequence $A_1 \supseteq A_2 \supseteq \ldots$ of sets in C,

$$\text{if } \bigcap_{i=1}^{\infty} A_i \in \mathsf{C}, \text{ then } \lim_{i\to\infty} g(A_i) = g\left(\bigcap_{i=1}^{\infty} A_i\right) \quad \text{(continuity from above).}$$

Functions that satisfy requirements (g1), (g2), and either (g3) or (g4) are equally important in the theory of monotone measures. These functions are called *semicontinuous* from below or above, respectively. When the universal set X is finite, requirements (g3) and (g4) are trivially satisfied and may thus be disregarded. Observe also that requirement (g2) defines measures that are actually *monotone increasing*. By changing the inequality $g(A) \leq g(B)$ in (g2) to $g(A) \geq g(B)$, we can define measures that are *monotone decreasing*. Both types of monotone measures are useful, even though monotone increasing measures are more common in dealing with uncertainty. An excellent example of monotone decreasing measure is the measure of a potential surprise, which was developed by George Shackle in economics (Klir, 2002).

For any pair $A, B \in \mathsf{C}$ such that $A \cap B = \varnothing$, a monotone measure g is capable of capturing any of the following situations:

a) $g(A \cup B) \geq g(A) + g(B)$, called *superadditivity*, which expresses a cooperative action or synergy between A and B in terms of the measured property;
b) $g(A \cup B) = g(A) + g(B)$, called *additivity*, which expresses the fact that A and B are noninteractive with respect to the measured property;
c) $g(A \cup B) \leq g(A) + g(B)$, called *subadditivity* which expresses some sort of inhibitory effect or incompatibility between A and B as far as the measured property is concerned.

Observe that probability theory, which is based on classical measure theory, is capable of capturing only situation (b). This demonstrates that the theory of monotone measures provides us with a considerably broader framework

than probability theory for formalizing uncertainty. As a consequence, it allows us to capture types of uncertainty that are beyond the scope of probability theory.

Nonclassical uncertainty theories are based on various special classes of monotone measures. Some of the most visible theories are surveyed in Sec.4.

3.2 Generalized Set Theory

Classical sets are generalized by abandoning the requirement of sharp boundaries between sets. The generalized sets, which allow us to distinguish degrees (or grades) of membership, are called *fuzzy sets*. Each of these sets is fully characterized by a membership function of the form

$$A : D \rightarrow R . \tag{23}$$

Distinct types of fuzzy sets are primarily distinguished by distinct types of sets R (ranges) or, in some cases, by special sets D (domains) that are employed in their membership functions. For each $d \in D$, $A(d)$ is viewed as the degree of membership of object d in fuzzy set A or, alternatively, as the degree of compatibility of object d with the concept represented by fuzzy set A.

The following are perhaps the most visible of the recognized types of fuzzy sets:

- If $R = [0,1]$, then A is called a *standard fuzzy set* (Klir and Yuan, 1995a).
- If R is the set of closed subintervals of $[0,1]$, then A is called and *interval-valued fuzzy set* (Mendel, 2001).
- If R is the set of fuzzy subintervals of $[0,1]$, then A is called a *fuzzy set of type 2* (Mendel, 2001).
- If R is a lattice of qualitative labels of membership grades, then A is called an *L-fuzzy set* (Goguen, 1967).
- If D is a set of fuzzy sets, A is called a *fuzzy set of level 2* (Gottwald, 1979).

Additional important types of fuzzy sets include fuzzy sets of *higher types* and *higher levels*, and those known as *rough fuzzy sets* (Dubois and Prade, 1990) and *intuitionistic fuzzy sets* (Atanassov, 2000).

By abandoning the requirement of sharp boundaries between sets, the mathematical structure of a Boolean is abandoned as well in the area of fuzzy sets. Which structure replaces it depends not only on the type of fuzzy sets, but also on the chosen operations. This great diversity of formalized languages opened by the various types of fuzzy sets is a matter of future research in GIT. Thus far, only the standard fuzzy sets have been utilized for formalizing uncertainty.

4. UNCERTAINTY THEORIES EMERGING FROM THE GIT FRAMEWORK

4.1 Imprecise Probabilities

The theory of monotone measures has been instrumental in formalizing the notion of *imprecise probabilities*. Several theories of imprecise probabilities are now well developed. Among them, two theories, one developed by Walley (1991) and the other one pursued by Kyburg (1987), are currently the most generalized theories of imprecise probabilities. The former theory is formalized in terms of *lower and upper previsions*, the latter one in terms of *closed convex sets of probability distribution functions*. It is established that these two theories are formally equivalent.

All theories of imprecise probabilities that are based on classical set theory share some common characteristics. One of them is that evidence within each theory is fully described by a *lower uncertainty function*, $\underline{u}$, or alternatively, by and *upper uncertainty function*, $\overline{u}$. These functions are always regular monotone measures that are superadditive and subadditive, respectively, and

$$\sum_{x\in X}\underline{u}(\{x\})\le 1,\quad \sum_{x\in X}\overline{u}(\{x\})\ge 1. \tag{24}$$

In the various special theories of uncertainty, they possess additional properties.

When evidence is expressed (at the most general level) in terms of an arbitrary closed and convex set D of probability distribution functions p on a finite set X, functions $\underline{u}_{\mathsf{D}}$ and $\overline{u}_{\mathsf{D}}$ associated with D are determined for each $A \in \mathsf{P}(X)$ by the formulas

$$\underline{u}_{\mathsf{D}}(A)=\inf_{p\in\mathsf{D}}\sum_{x\in A}p(x),\quad \overline{u}_{\mathsf{D}}(A)=\sup_{p\in\mathsf{D}}\sum_{x\in A}p(x),$$

Since

$$\sum_{x\in A}p(x)+\sum_{x\notin A}p(x)=1,$$

for each $p \in \mathsf{D}$ and each $A \in \mathsf{P}(X)$, it follows that

$$\overline{u}_{\mathsf{D}}(A)=1-\underline{u}_{\mathsf{D}}(\overline{A}). \tag{25}$$

Due to this property, functions $\underline{u}_{\mathsf{D}}$ and $\overline{u}_{\mathsf{D}}$ are called *dual* (or *conjugate*). One of them is sufficient for capturing given evidence; the other one is uniquely determined by (25). It is common to use the lower probability function $\underline{u}_{\mathsf{D}}$ to capture the evidence.

As is well known (Grabisch, 1997), any given lower probability function $\overline{u}_{\mathsf{D}}$ is uniquely represented by a set-valued function m_{D} for which $m_{\mathsf{D}}(\varnothing)=0$ and

$$\sum_{A\in \mathsf{P}(X)} m_{\mathsf{D}}(A)=1. \tag{26}$$

Any set $A \in \mathsf{P}(X)$ for which $m_{\mathsf{D}}(A) \neq 0$ is often called a *focal element*, and the set of all focal elements with the values assigned to them by function m_{D} is called a *body of evidence*. Function m_{D} is called a *Möbius representation* of $\underline{u}_{\mathsf{D}}$ when it is obtained for all $A \in \mathsf{P}(X)$ via the *Möbius transform*

$$m_{\mathsf{D}}(A)=\sum_{B|B\subseteq A}(-1)^{|A-B|}\underline{u}_{\mathsf{D}}(B) \tag{27}$$

The inverse transform is defined for all $A \in \mathsf{P}(X)$ by the formula

$$\underline{u}_{\mathsf{D}}(A)=\sum_{B|B\subseteq A} m_{\mathsf{D}}(B). \tag{28}$$

Assume now that evidence is expressed in terms of a given lower probability function $\underline{u}$. Then, the set of probability distribution functions that are consistent with $\underline{u}$, $\mathsf{D}(\underline{u})$, which is always closed and convex, is defined as follows:

$$\begin{aligned}\mathsf{D}(\underline{u}) = \{p(x)|x\in X, p(x)\in[0,1], \sum_{x\in X} p(x)=1, \text{and}\\ \underline{u}(A)\leq \sum_{x\in A} p(x) \text{ for all } A\in \mathsf{P}(X)\}\end{aligned} \tag{29}$$

That is, each given function $\underline{u}$ is associated with a unique set D and vice-versa.

4.2 Some Special Uncertainty Theories

Uncertainty functions with various special properties have been studied in the literature, resulting in distinct theories of uncertainty. A few of these

special uncertainty functions are briefly introduced in this section to illustrate nonclassical uncertainty theories.

A simple uncertainty theory is based on the class of λ-measures, u_λ, which are regular monotone measures that satisfy the requirement

$$u_\lambda(A \cup B) = u_\lambda(A) + u_\lambda(B) - \lambda u_\lambda(A) u_\lambda(B) \tag{30}$$

for given pair of disjoint sets A, $B \subseteq X$, where $\lambda \in (-1,\infty)$ is a parameter by which individual measures in this class are distinguished. It is well-known that the value of λ is uniquely determined by values $u_\lambda(\{x\}) \in [0,1]$ for all $x \in X$ via the equation

$$1 + \lambda = \prod_{x \in X} [1 + \lambda u_\lambda(\{x\})] \tag{31}$$

When $\lambda > 0$, u_λ is a lower uncertainty function; when $\lambda < 0$, u_λ is an upper uncertainty function; when $\lambda = 0$, u_λ is clearly a classical probability measure. For further information, see (Wang and Klir, 1992).

In another, well developed uncertainty theory, lower and upper uncertainty functions, $\underline{u}$ and $\overline{u}$, are uniquely determined for all sets $A \subseteq X$ by their values $\underline{u}(\{x\})$ and $\overline{u}(\{x\})$ on singletons for all $x \in X$. It is required that $\underline{u}(\{x\}) \leq \overline{u}(\{x\})$ and that the inequalities (24) be satisfied. The primary reference dealing with this uncertainty theory is (Weicheselberger and Pöhlman, 1990).

A well defined category of theories of imprecise probabilities is based on *Choquet capacities of various orders* (Choquet, 1953-54). All these theories are generalizations of the theory based on Sugeno λ-measures, but they are not comparable with the theory based on feasible interval-valued probability distributions. The most general theory in this category is the theory based on *capacities of order 2*. Here, the lower and upper uncertainty functions, $\underline{u}$ and $\overline{u}$, are monotone measures for which

$$\underline{u}(A \cup B) \geq \underline{u}(A) + \underline{u}(B) - \underline{u}(A \cap B), \tag{32}$$

$$\overline{u}(A \cap B) \leq \overline{u}(A) + \overline{u}(B) - \overline{u}(A \cup B). \tag{33}$$

for all A, $B \subseteq X$. Less general uncertainty theories are then based on *capacities of order k*. For each $k > 2$, the lower and upper probabilities, $\underline{u}$ and $\overline{u}$, satisfy the inequalities

$$\underline{u}(\bigcup_{j=1}^{k} A_j) \geq \sum_{\substack{K \subseteq N_k \\ K \neq \emptyset}} (-1)^{|K|+1} \underline{u}(\bigcap_{j \in K} A_j), \tag{34}$$

$$\overline{u}(\bigcap_{j=1}^{k} A_j) \leq \sum_{\substack{K \subseteq N_k \\ K \neq \emptyset}} (-1)^{|K|+1} \overline{u}(\bigcup_{j \in K} A_j) \tag{35}$$

for all families of k subsets of X, where $N_k = \{1, 2, \ldots, k\}$. Clearly, if $k' > k$ then the theory based on capacities of order k' is less general than the one based on capacities of order k. The least general of all these theories is the one in which the inequalities are required to hold for all $k \geq 2$ (the underlying capacity is said to be of order ∞). This theory, which was extensively developed by Shafer (1976), is usually referred to as *evidence theory* or *Dempster-Shafer theory*. In this theory, lower and upper uncertainty functions are called *belief* and *plausibility* measures. An important feature of this theory is that the Möbius representation of evidence in this theory (usually called a *basic probability assignment function*) is a positive function ($m(A) \geq 0$ for all $A \in \mathsf{P}(X)$), and hence, it can be viewed as a probability distribution on the power set. This feature makes it possible to develop the theory via "non-traditional and sophisticated application of probability theory" as shown convincingly by Kramosil (2001). Another approach to evidence theory has been pursued by Smets (1988).

Other types of monotone measures have been introduced in the literature, but uncertainty theories based upon them have not been properly developed as yet. Among them, two broad classes of monotone measures seem to have a great potential utility: *decomposable measures* (Dubois and Prade, 1982) and *k-additive measures* (Grabisch, 1997).

For any pair of disjoint subsets A, B of X, a monotone measure, $u_{\otimes}$, that is *decomposable* with respect to a t-conorm $\otimes$ (Klir and Yuan, 1995a), is required to satisfy for every pair $A, B \in \mathsf{P}(X)$ the following requirement:

$$u_{\otimes}(A \cup B) = u_{\otimes}(A) \otimes u_{\otimes}(B). \tag{36}$$

The class of λ-measures is a particular class of decomposable monotone measures, in which t-conorms $\otimes_{\lambda}$ have the form

$$a \otimes_{\lambda} b = \min[a + b + \lambda ab, 1].$$

For classical probabilities, the t-conorm is the arithmetic sum.

The class of *k-additive measures*, which was introduced and investigated by Grabish (1997), is defined as follows: A monotone measure is said to be

k-additive ($k \geq 1$) if its Möbius representation satisfies $m(A) = 0$ for any A such that $|A| > k$, and there exists at least one subset B of X such that $|B| = k$ and $m(B) \neq 0$. The advantage of k-additive measures, especially for small values of k, is that they are characterized via the Möbius representation by a limited number of parameters (focal elements). They are thus suitable for approximating other monotone measures, as investigated by Grabisch (1997). Clearly, the 1-additive measures are classical probability measures.

4.3 Fuzzification of Uncertainty Theories

It is now well established that any area of classical mathematics can be extended to its fuzzy counterpart by the process of *fuzzification*. According to Goguen (1967), "Fuzzification is the process of imparting a fuzzy structure to a definition (concept), theorem, or even a whole theory." Fuzzification can be attained in three distinct ways:

1. via the α-cut representation of fuzzy sets;
2. via the extension principle for fuzzy sets;
3. via fuzzy morphisms and category theory.

The first two ways are well covered in fuzzy-set literature; the third one, which is technically more demanding, has a rather limited coverage (Rodabaugh et al., 1992).

Fuzzifying mathematical objects is not unique. For example, classical probability theory can be fuzzified by fuzzifying events or probabilities or both. Fuzzifying only events is perhaps the simplest (and oldest) way of fuzzifying uncertainty theories. Probabilities, $Pro(A)$, of fuzzy events A (Zadeh, 1968) are still classical probabilities, which are calculated either by the formula

$$\mathrm{Pro}(A) = \sum_{x \in X} A(x)p(x) \tag{37}$$

when X is finite and p is a probability distribution function, or by the formula

$$\mathrm{Pro}(A) = \int_X A(x)f(x)dx \tag{38}$$

when $X = \mathsf{E}^n$ for some $n \geq 1$ and f is a probability density function on X.

Efforts to fuzzify the various uncertainty theories have been rather limited so far. The following are some of the well developed fuzzifications:

1. Fuzzification of the classical (crisp) possibility theory by allowing grades of possibilities was suggested by Zadeh (1978). This generalized

possibility theory is covered quite extensively in the literature. Perhaps the most comprehensive overview of the theory is covered in a series of articles by De Cooman (1997). In Klir (1999), the issue of possibilistic normalization is addressed, which is largely neglected in the rest of the literature.

2. Fuzzification of the concept of a random variable is covered in the literature quite extensively. A Special Issue edited by Gil (2001) covers recent developments as well as references to previous work.
3. Perhaps the most definitive results have been obtained in fuzzifying the theory based on interval-valued probability distributions. In particular, Bayesian methodology developed for interval-values probability distributions by Pan and Klir (1997) was fuzzified by Pan and Yuan (1997). The fuzzification is based on extending relevant intervals of real numbers to fuzzy intervals via the α-cut representation and on using constrained fuzzy arithmetic (Klir, 1997) to perform required computations.
4. Interesting approaches to fuzzifying the Dempster-Shafer theory were proposed by Yen (1990) and Yang et al. (2003), but either of them has been adequately developed as yet and no comparative study has been conducted so far.

5. MEASURES OF UNCERTAINTY

A *measure of uncertainty* of some conceived type in a given uncertainty theory is a *functional* that assigns to each uncertainty function in the theory a nonnegative real number. This number is supposed to measure, in an intuitively meaningful way, the amount of uncertainty of the considered type that is embedded in the uncertainty function. Examples of measures of uncertainty are the Hartley measure and the Shannon entropy. Uncertainty functions that are directly involved in these measures are, respectively, possibility distribution functions and probability distribution functions.

To be acceptable as a measure of the amount of uncertainty of a given type in a particular uncertainty theory, a proposed functional must satisfy several intuitively essential axiomatic requirements. Specific mathematical formulation of each of the requirements depends on the uncertainty theory involved. However, the requirements can be described informally in a generic form, independent of the various uncertainty calculi.

The following axiomatic requirements, each expressed in a generic form, must be satisfied whenever applicable:

1. *Subadditivity* – the amount of uncertainty in a joint representation of evidence (defined on a Cartesian product) cannot be greater than the sum

of the amounts of uncertainty in the associated marginal representations of evidence.

2. *Additivity* – the two amounts of uncertainty considered under subadditivity become equal if and only if the marginal representations of evidence are noninteractive according to the rules of the uncertainty calculus involved.
3. *Range* – the range of uncertainty is $[0, M]$, where 0 must be assigned to the unique uncertainty function that describes full certainty and M depends on the size of the universal set involved and on the chosen unit of measurement.
4. *Continuity* – any measure of uncertainty must be a continuous functional.
5. *Expansibility* – expanding the universal set by alternatives that are not supported by evidence must not affect the amount of uncertainty.
6. *Branching/Consistency* – when uncertainty can be computed in more ways, all acceptable within the calculus of the uncertainty theory involved, the results must be the same (consistent).
7. *Monotonocity* – when evidence can be ordered in the uncertainty theory employed (as in possibility theory), the relevant uncertainty measure must preserve this ordering.
8. *Coordinate invariance* – when evidence is described within the n-dimensional Euclidean space ($n \geq 1$), the relevant uncertainty measure must not change under isometric transformations of coordinates.

When distinct types of uncertainty coexist in a given uncertainty theory, it is not necessary that these requirements be satisfied by each uncertainty type. However, they must be satisfied by an overall uncertainty measure, which appropriately aggregates measures of the individual uncertainty types.

The strongest justification of a functional as a meaningful measure of the amount of uncertainty of a considered type in a given uncertainty theory is obtained when we can prove that it is the only functional that satisfies the relevant requirements and measures the amount of uncertainty in some specific measurement units. A suitable measurement unit is uniquely defined by specifying what the amount of uncertainty should be for a particular (and usually very simple) uncertainty function.

Two types of uncertainty coexist in each nonclassical uncertainty theory: nonspecificity and conflict. They are measured, respectively, by appropriately generalized Hartley and Shannon measures.

5.1 Generalized Hartley Measures

The Hartley measure (introduced in Sec. 2.1) was first generalized to the theory of graded possibilities, then to the Dempster-Shafer theory and,

eventually, by Abellan and Moral (2002), to arbitrary closed convex sets of probability distributions and, hence, to all theories of imprecise probabilities. This generalized functional, *GH*, is defined by the formula

$$GH(m_{\mathsf{D}}) = \sum_{A \subseteq X} m_{\mathsf{D}}(A) \log_2 |A| , \tag{39}$$

where m_{D} denotes the Möbius representation of the lower uncertainty function associated with a given closed and convex set D of probability distributions on X.

Abellan and Moral (2002) showed that this functional possesses all the essential mathematical properties required for measures of uncertainty. In particular, they proved that the measure has the following properties:

1. It has the *proper range* $[0, \log_2 |X|]$ when measured in bits; 0 is obtained when D contains a single probability distribution; $\log_2 |X|$ is obtained when D contains all probability distributions on X and thus represents total ignorance.
2. It is *subadditive*: $GH(\mathsf{D}) \leq GH(\mathsf{D}_X) + GH(\mathsf{D}_Y)$, where

$$\mathsf{D}_X = \{p_X | p_X(x) = \sum_{y \in Y} p(x, y) \text{ for some } p \in \mathsf{D}\} ,$$

$$\mathsf{D}_Y = \{p_Y | p_Y(y) = \sum_{x \in X} p(x, y) \text{ for some } p \in \mathsf{D}\} .$$

3. It is *additive*: $GH(\mathsf{D}) = GH(\mathsf{D}_X) + GH(\mathsf{D}_Y)$ if and only if D_X and D_Y are not interactive, which means that for all $A \in \mathsf{P}(X)$ and all $B \in \mathsf{P}(Y)$,

$$m_{\mathsf{D}}(A \times B) = m_{\mathsf{D}_X}(A) \cdot m_{\mathsf{D}_Y}(B) \quad \text{and } \mathrm{mD(R)} = 0 \text{ for all } R \neq A \times B.$$

4. It is *monotonic*: if D and D' are closed convex sets of probability distributions on X such that $\mathsf{D} \subseteq \mathsf{D}'$, then $GH(\mathsf{D}) \leq GH(\mathsf{D}')$.

The generalized Hartley measure defined by Eq. (39) is thus a mathematically sound measure of nonspecificity. An overview of the history regarding this generalization prior to the work by Abellan and Moral (2002) is covered in (Klir and Wierman, 1999).

In the 1980s and early 1990s, several intuitively promising functionals were proposed as candidates for the generalized Shannon measure in evidence theory. However, each of them was found, upon closer scrutiny, to violate the essential requirement of subadditivity A historical summary of these unsuccessful efforts is covered in (Klir and Wierman, 1999)

Generalized Shannon measure, *GS*, was eventually defined indirectly, via an *aggregated uncertainty*, *AU*, covering both nonspecificity and conflict, and the well established generalized Hartley measure of nonspecificity, *GH* defined by Eq. (39). Since it must be that $GH + GS = AU$, the generalized Shannon measure can be defined as

$$GS = AU - GH. \tag{40}$$

Using this definition, the unsuccessful effort to find *GS* directly is replaced with the effort to find *AU* and define *GS* indirectly via Eq. (40). The latter effort was successful in the mid 1990s, when a functional *AU* satisfying all essential requirements was established in evidence theory (Klir and Wierman, 1999). However, this functional is applicable to all the other theories of imprecise probabilities as well, which follows from the common properties shared by these theories. Given any lower uncertainty function $\underline{u}$ associated with a closed convex set, D, of probability distributions (or vice versa), $AU(\underline{u})$ is defined by the formula

$$AU(\underline{u}) = \max_{p \in \mathrm{D}}[-\sum p(x)\log_2 p(x)]. \tag{41}$$

It is the maximum Shannon entropy within D. An efficient algorithm for computing this maximum, which was proven correct for belief functions of evidence theory (Klir and Wierman, 1999), is applicable without any change when belief functions are replaced with lower probabilities of any other kind.

Although functional *AU* is a well-justified measure of total uncertainty in the various theories of uncertainty, it is highly insensitive to changes in evidence due to its aggregated nature. It is an aggregate of the two coexisting types of uncertainty, nonspecificity and conflict. It is thus desirable to express the total uncertainty, *TU*, in a disaggregated form

$$TU = (GH, GS), \tag{42}$$

where *GH* is defined by (39) and *GS* is defined by (40) and (41).

6. GIT IN RETROSPECT AND PROSPECT

The principal aim of GIT have been threefold: (i) to liberate the notions of uncertainty and uncertainty-based information from the narrow confines of classical set theory and classical measure theory; (ii) to conceptualize a

broad framework within which any conceivable type of uncertainty can be characterized; and (iii) to develop theories for the various types of uncertainty that emerge from the framework at each of the four levels: formalization, calculus, measurement, and methodology. Undoubtedly, these are long-term aims, which may perhaps never be fully achieved. Nevertheless, they serve well as a blueprint for a challenging, large-scale research program.

The basic tenet of GIT, that uncertainty is a broader concept than the concept of classical probability theory, has been debated in the literature since the late 1980s (for an overview, see (Klir, 2001b)). As a result of this ongoing debate as well convincing advances in GIT, limitations of classical probability theory are increasingly recognized.

In addition to demonstrating limitations of classical uncertainty theories and conceptualizing a fairly comprehensive framework for studying the full scope of uncertainty, a respectable number of nonclassical uncertainty theories have already been well developed within GIT at the levels of representation, calculus, and measurement, and to some extent at the methodological level. Notwithstanding the significance of these developments, they represent only a tiny fraction of the whole area. Most prospective theories of uncertainty and uncertainty-base information are still undeveloped For example, none of the nonstandard fuzzy sets have been employed in formalizing uncertainty thus far.

The two-dimensional framework for formalizing uncertainty is quite instrumental in guiding future research in GIT. However, it should be emphasized that this broad framework may still not be sufficiently general to capture all conceivable types of uncertainty and the associated uncertainty-based information.

The role of information in human affairs has become so predominant that it is now quite common to refer to our society as information society. It is thus increasingly important for us to develop a good understanding of the broad concept of information. In the generalized theory of uncertainty-based information, information is viewed as a commodity whose value is its potential to reduce uncertainty pertaining to relevant situations, where the concept of uncertainty is understood in the broadest possible terms. The theory does not deal with the issues of how much uncertainty is actually reduced, if any, and how valuable this uncertainty reduction is to information users (cognitive agents) in the context of each given situation. However, the theory, when adequately developed, will be a broad base upon which it will be possible to erect a conceptual structure to capture semantic and pragmatic aspects relevant to information users under various situations of information flow. Only when this is adequately accomplished, a genuine science of information will be created.

When considering the various results emerging from GIT, as examined in this paper, GIT is now a well-established area of research. When considering future challenges of GIT, they are enormous. This huge gap between what has been accomplished and what is yet to be accomplished will undoubtedly make GIT a very active area of challenging and important research for many years to come.

REFERENCES

Abellan, J., and Moral, S., 2002, A non-specificity measure for convex sets of probability distributions, *Intern. J. of Uncertainty, Fuzziness and Knowledge-Based Systems* **8**(3):357-367.

Atanassov, K. T., 2000, *Intuitionistic Fuzzy Sets*, Springer-Verlag, New York.

Choquet, G., 1953-54, Theory of capacities, *Annales de L'Institut Fourier* **5**:131-295.

De Cooman, G., 1997, Possibility theory - I, II, III, *Intern. J. of General Systems* **25**,291-371.

Denneberg, D., 1994, *Non-additive Measure and Integral*, Kluwer, Boston.

Dubois, D., and Prade, H., 1982, A class of fuzzy measures based on triangular norms, *Intern. J. of General Systems* **8**(1):43-61.

Dubois, D., and Prade, H., 1990, Rough fuzzy sets and fuzzy rough sets, *Intern. J. of General Systems* **17**(2-3):191-209.

Gil, M. A., ed., 2001, Special Issue on Fuzzy Random Variables, *Information Sciences* **133**(1-2):1-100.

Goguen, J. A., 1967, L-fuzzy sets, *J. of Math. Analysis and Applications* **18**(1):145-174.

Gottwald, S., 1979, Set theory for fuzzy sets of higher level, *Fuzzy Sets and Systems* **2**(2):125-151.

Grabisch, M., 1997, k-order additive discrete fuzzy measures and their representation, *Fuzzy Sets and Systems* **92**(2):167-189.

Halmos, P. R., 1950, *Measure Theory*, D.Van Nostrand, Princeton, NJ.

Hartley, R.V.L., 1928, Transmission of information, *The Bell System Technical J.*, **7**:535-563.

Klir, G. J., 1991, Generalized information theory, *Fuzzy Sets and Systems* **40**(1):127-142.

Klir, G. J., 1997, Fuzzy arithmetic with requisite constraints, *Fuzzy Sets and Systems* **91**(2):165-175.

Klir, G. J., 1999, On fuzzy-set interpretation of possibility theory, *Fuzzy Sets and Systems* **108**(3):263-273.

Klir, G. J., 2001a, *Facets of Systems Science*, Plenum Press, New York.

Klir, G. J., 2001b, Foundations of fuzzy set theory and fuzzy logic: A Historical Overview, *Intern. J. of General Systems* **30**(2):91-132.

Klir, G. J., 2002, Uncertainty in Economics: The heritage of G.L.S. Shackle, *Fuzzy Economic Review* **VII**(2):3-21.

Klir, G. J., 2005, *Uncertainty and Information: Foundations and Applications of Generalized Information Theory* (in production).

Klir, G. J., and Wierman, M. J., 1999, *Uncertainty-Based Information: Elements of Generalized Information Theory*, Physica-Verlag/Springer-Verlag, Heidelberg and New York.

Klir, G. J., and Yuan, B., 1995a, *Fuzzy Sets and Fuzzy Logic: Theory and Applications* Prentice Hall, PTR, Upper Saddle River, NJ.

Klir, G. J., and Yuan, B., 1995b, On nonspecificity of fuzzy sets with continuous membership functions, in: *Proc. Intern. Conf. on Systems, Man, and Cybernetics,* Vancouver, pp.25-29.

Klir, G. J., and Yuan, B. (eds.), 1996, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh,* World Scientific, Singapore.

Kolmogorov, A. N., 1965, Three approaches to the quantitative definition of information, *Problems of Information Transmission* **1**:1-7.

Kramosil, I., 2001, *Probabilistic Analysis of Belief Functions*, Kluwer Academic/Plenum Publishers, New York.

Kyburg, H. E., 1987, Bayesian and non-Bayesian evidential updating, *Artificial Intelligence* **31**:271-293.

Mendel, J. M., 2001, *Uncertain Rule-Based Fuzzy Logic Systems*, Prentice Hall PTR, Upper Saddle River, NJ.

Pan, Y., and Klir, G. J., 1997, Bayesian inference based on interval probabilities, *J. of Intelligent and Fuzzy Systems* **5**(3):193-203.

Pan, Y., and Yuan, B., 1997, Bayesian inference of fuzzy probabilities, *Intern. J. of General Systems* **26**(1-2):73-90.

Pap, E., 1995, *Null-Additive Set Functions*, Kluwer, Boston.

Ramer, A., and Padet, C., 2001, Nonspecificity in P^n, *Intern. J. of General Systems* **30**(6):661-680.

Rènyi, A., 1970, *Probability Theory*, North-Holland, Amsterdam (Chapter IX, Introduction to Information Theory, pp.540-616).

Rodabaugh, S. E., Klement, E. P., and Höhle, Y., 1992, *Applications of Category Theory to Fuzzy Subsets*, Kluwer, Boston.

Shafer, G., 1976, *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ.

Shannon, C. E., 1948, The mathematical theory of communication, *The Bell System Technical J.* **27**:379-423, 623-656.

Smets, P., 1988, Belief functions, in: *Non-standard Logics for Automated Reasoning*, P. Smets et al., ed., Academic Press, San Diego, pp 253-286.

Stoll, R. R., 1961, *Set Theory and Logic*, W. H. Freeman, San Francisco and London.

Walley, P., 1991, *Statistical Reasoning With Imprecise Probabilities*, Chapman and Hall, London.

Wang, Z., and Klir, G. J., 1992, *Fuzzy Measure Theory*, Plenum Press, New York.

Weichselberger, K., and Pöhlmann, S., 1990, *A Methodology for Uncertainty in Knowledge-Based Systems*, Springer-Verlag, New York.

Yager, R. R., et al., eds., 1987, *Fuzzy Sets and Applications - Selected Papers by L. A. Zadeh*, John Wiley, New York.

Yang, M., Chen, T., and Wu, K., 2003, Generalized belief function, plausibility function, and Dempster's combination rule to fuzzy sets, *Intern. J. of Intelligent Systems* **18**(8):925-937.

Yen, J., 1990, Generalizing the Dempster-Shafer theory to fuzzy sets, *IEEE Transactions on Systems, Man, and Cybernetics* **20**(3):559-570.

Zadeh, L. A., 1968a, Probability measures and fuzzy events, *J. of Math. Analysis and Applications* **23**(2):421-427.

Zadeh, L. A., 1978, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* **1**(1):3-28.

1

APPLICATIONS

COMPLEXITY IN UNIVERSE DYNAMIC EVOLUTION. PART 1 - PRESENT STATE AND FUTURE EVOLUTION

Umberto Di Caprio
Stability Analysis s.r.l., Via Andrea Doria 48/A, 20124 Milano, Italy

Abstract: Recent experimental results gathered by spatial Telescope Hubble, from October 2003, deeply modify our knowledge of Universe. They give numerical estimates of fundamental quantities as "age of Universe", geometric form, radius, density of matter, expansion rate with time, birth of galaxies, Hubble constant. We frame these results in a general theory that explains present status and, in addition, forecasts future evolution, and extrapolates past structure from time zero on. We propose a simple solution of the problem of "missing mass" and point out existence of dark energy. Complexity enters the picture through formulation and testing of a two-body dynamic model in which visible Universe rotates around a central black-hole. Presentation is splitted in two parts, one dealing with present state and future evolution, the other with past evolution from time zero.

Key words: telescope Hubble; universe; missing mass; two-body; black-hole; cosmological model; SR; GR; stability; critical density; deceleration parameter.

LIST OF SYMBOLS

t_0, t_f	age of universe, time at the end of expansion;
$M_0 = M(t_0)$	mass, present value;
$M_f = M(t_f)$	mass, value at the end of expansion;
$R_0 = R(t_0)$	radius, present value;
$R_f = R(t_f)$	radius, value at the end of expansion;
ρ_0, ρ_f	density: present value, final value at the end of expansion;

ρ_c	critical density;
H_0, τ_H	Hubble constant, Hubble time;
s, s_0	deceleration parameter, present value;
M_B, M_{B0}	mass of central black-hole: present value, value at time t ;
$M_f = M(t_f)$	final mass of visible universe;
$R(t_f) = R_{ss}$	radius of stabilization round a black-hole;
R_s	minimum radius for stability round a black-hole;
c, $c(t)$	speed of light in vacuum: present value, value at time t ;
m_0, m_p	electron mass, proton mass;
q	unitary electric charge;
α	fine structure constant;
G, $G(t)$	gravitation constant: present value, value at time t ;
ε_0, $\varepsilon(t)$	permittivity in vacuum: present value, value at time t ;
μ_0, $\mu(t)$	permeability in vacuum: present value, value at time t ;
k, $k(t)$	Coulomb constant: present value, value at time t ;
Θ	temperature of visible universe;
K	expansion constant;
α_{mq}	electro-gravitational ratio;
r_B, R_L	Bhor radius, Lorentz radius;
k_B	Boltzman constant;
E_p, T	potential energy, kinetic energy;
$\gamma_r = 1.618$	golden ratio;
v_{rot}	rotational speed.
γ_{rot}	relativistic mass coefficient;

Remark: We denote s_0 the "deceleration parameter", rather than q_0 (standard notation) to avoid confusion with symbol q that, on the contrary, represents the unitary electric charge.

1. INTRODUCTION

Experimental information gathered by Hubble Space-Telescope in Autumn 2003, and successive months, shows precious and novel information about Universe structure and dynamic evolution. In particular:

1. Universe has a finite extension and closed spheroidal structure (fig.1) with radius $R_0 \approx 30\times10^9$ light-years.
2. Universe has a finite age $t_0 \approx 13.65\times10^9$ years $= 4.3\times10^{17}$ s.
3. Universe is in expansion.
4. In past time Universe had a smaller extension, smaller and smaller going back with time.
5. First galaxies have formed about 13.2 billion years ago.

6. Galaxies are rotating bodies around massive black-holes. Masses of galaxies and of their black-holes have been conveniently augmenting with time.

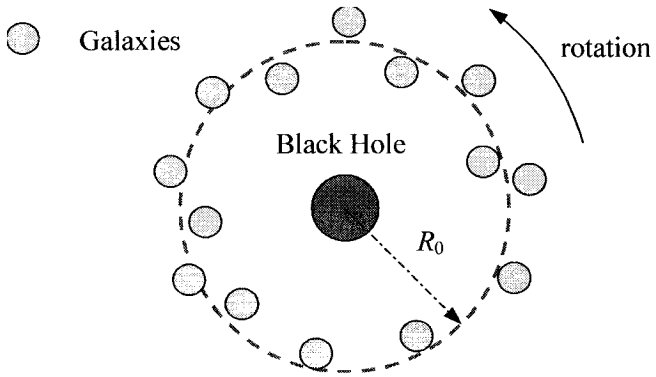


Figure 1. Universe Present Structure.

On the other hand it was already known that estimated mass of present universe is about 3% of the value indirectly determined from measurements of the "deceleration parameter" s_0

$$s_0 = s(t_0) \quad ; \quad s = -\frac{R\ddot{R}}{\dot{R}^2} \tag{1.1}$$

The problem is how to set together all above information.

We propose an original approach based on *complexity theory* and on the principle of *equivalence Potential Energy / mass*. (In *complexity theory* we include *stability theory*). In order to explain closed and spheroidal form we postulate a two-body dynamic model in which visible Universe rotates around a central black-hole and simultaneously undergoes radial expansion. We determine parameters of interest by postulating "similarity" with hydrogen and assuming that present value of deceleration parameter is $s_0 =$ 0.2355 as required to solve (as we shall see) the problem of "missing mass". To define the above similarity we introduce the adimensional quantity

$$\alpha_{mq} = \frac{Gm_p m_0}{kq^2} = 4.4067 \times 10^{-40}. \tag{1.3}$$

The present value of Hubble constant H_0 is determined by eq.

$$H_0 = \alpha_{mq}\left[\frac{c}{r_\Gamma}\right] \; ; \; r_\Gamma = \alpha^{4/3} r_B \tag{1.4}$$

in which α is the fine structure constant, r_Γ is the average radius of hydrogen atom r_B is the Bohr radius.
The corresponding Hubble time is $\tau_H = H_0^{-1} = 5.6696\times10^{17}$ s.
From (1.2), (1.5) and a noted formula derived by (Weinberg, 1972 pp 481-484, see Appendix 2) we determine the age of Universe:
$t_0 = 0.7592\ \tau_H = 4.30\times10^{17}$ s $= 13.65\times10^9$ years
Also, from (1.4) and (see Weinberg)

$$\rho_c = \frac{3}{8\pi G} H_0^2 \tag{1.7}$$

we determine the *critical density*, i.e. the value beyond which expansion is substituted by contraction , $\rho_c = 5.565\times10^{-27}$ kg/m^3.
Eq.s (1.2) (1.8) lead to determination of present time density
$\rho_0 = 2\, s_0\, \rho_c = 2.623\times10^{-27}$ kg/m^3.
Using for present Universe radius the similarity formula (with hydrogen)

$$R_0 = (\alpha\, r_B)\frac{1}{\alpha_{mq}}\frac{s_0}{\sqrt{1-2s_0}} \tag{1.10}$$

we find $R_0 = 2.844\times10^{26}$ m $= 30.06\times10^9$ light-years.
Eq.s (1.6) and (1.8) fully agree with experimental observations by Hubble space telescope. Furthermore from (1.9), (1.11) we can derive the present value of Visible Universe mass

$$M_0 = M(t_0) = \rho_0 \frac{4\pi}{3} R_0^3 = 2.54\times10^{53}\,\text{Kg}. \tag{1.12}$$

Summing up, the parameters of present Universe are

Table 1. Parameters of present Universe

Age 13.65×10^9 years	Radius 30.06×10^9 light-years
Mass 2.54×10^{53} Kg	Density 2.623×10^{-27} Kg/m^3
$H_0 =$ 54.3 Km/MegaParsec/s	Critical density 5.565×10^{-27} Kg/m^3
Deceleration parameter 0.2355	

1.1 Explicative comments

The fact that visible Universe has a finite extension and a spheroidal form undoubtedly means that Universe has a geometric center and a center of mass. The two properties must agree. Also, as known measurements show that the density has a finite value, then the mass of Universe must be finite. For reasons of "dynamical Equilibrium" the visible Universe must consist of a ring of matter in rotation round a central mass at rest. The Newton attraction force is counterbalanced by the centrifugal force. On the other hand a two body structure solves the known problem of missing mass, i.e. the discrepancy between estimates of density derived from measurements of luminosity of our galaxy and those derived from measurements of the red shift of cosmic radiation (owing to Universe total mass). In fact: in a two-body structure the coupling Potential Energy (of gravitational nature) is negative and since this Energy is equivalent to mass, visible mass is only a fraction of effective mass (*We postulate indeed that Potential Energy is equivalent to mass. This assumption is fully supported by other studies* (see Bibliography) *regarding the dynamic structure of the proton, the dynamic structure of the deuteron, the dynamic structure of the neutron, the dynamic structure of the photon*). This fraction grows in time, due to the fact that Potential Energy is inversely proportional to expansion radius and, by consequence, the absolute value of the energy in question decreases with time. A "stabilization" is got when expansion stops, i.e. when expansion speed becomes equal to zero.

A second factor contributing to the augmentation of effective mass with respect to visible mass is "rotation", since rotation produces well known relativistic effects (in the spirit of Special Relativity). At present time the relativistic "gamma factor" is about 1.5. At the beginning of expansion (i.e. 13.65 billion years ago) it was 1.618 (equal to the golden ratio).

A third reason is that total mass includes the contribution of central black-hole (*the three mentioned factors, altogether, are responsible for the seeming missing of mass*) (see Sec. 4). In Universe future evolution, mass of black-hole keeps constant while mass of visible Universe grows up to a final value 12.5×10^{53} Kg (see Sec. 3).

We emphasize that a black-hole is not a monster, nor practically nor conceptually. Stability theory tells us that any massive body behaves like a black-hole if the physical radius is smaller than the radius of instability around the body (see Sec.2). Such critical radius is half the Scharzschild radius of General Relativity. However we do not use GR whilst we use stability theory and SR for showing that the Scharzschild radius is but the stabilization radius around a black-hole. Expansion of Universe will stop when radius of visible universe will reach its stabilization value. In such

condition the density of universe will be equal to the critical density defined by equation (1.7).

This result might mean that the final density of Universe equals the critical density defined by classical GR models (Einstein, Friedmann, Weinberg, De Sitter) in which visible Universe does not rotate (around any body) and expansion (or contraction) is not contrasted (nor favored) by any dynamical force. In other words we find $\rho_f = \rho_c$ (*this result should be considered extraordinary*) with

$$\rho_f = \frac{M_f}{(4\pi/3)R_f^3}$$

Alternative models in literature (De Witt, Dicke, Milne) don't either consider rotation nor assume a two-body structure. None of them puts into evidence a negative coupling energy (which, in our model, is the gravitational potential energy). Not only this: we assume "absolute time" and "absolute space" in the spirit of Newton theory. In this respect our approach is "post-newtonian" and vaguely resembles Milne's approach. The likeness ends here, since we assume that potential energy has an inertia ("mass") and this allows us to avoid any cosmological correction and to leave aside the cosmological principle. Further note: part of the success of our formulation is related to the definition of the proper value of the Hubble constant in eq. (1.7) (i.e. in the eq. that determines the critical density). The Hubble constant is usually defined as a quantity proportional to the ratio

$$\frac{\dot{R}(t_0)}{R(t_0)}$$

and its determination is assigned to experiments. The latter, however, do not lead to univocal nor consistent estimates. We use an analytic and drastically innovative formula (eq. (1.4)) which proves to work fine. The inverse of the Hubble constant (i.e. the "Hubble time") is proportional to the time employed by light to cross an atom of Hydrogen; the proportionality factor represents the ratio between electric force and gravitation force in Hydrogen. This criterion appeals to similitude between the two-body structure of Universe and the two-body structure of Hydrogen (electron + proton); in particular, it invokes correspondence between gravitation force (in Universe) and electric force (in Hydrogen). An identical correspondence is at the basis of eq. (1.10) that binds together the radius of Universe to the Bohr radius (note that αr_B is the Compton radius).

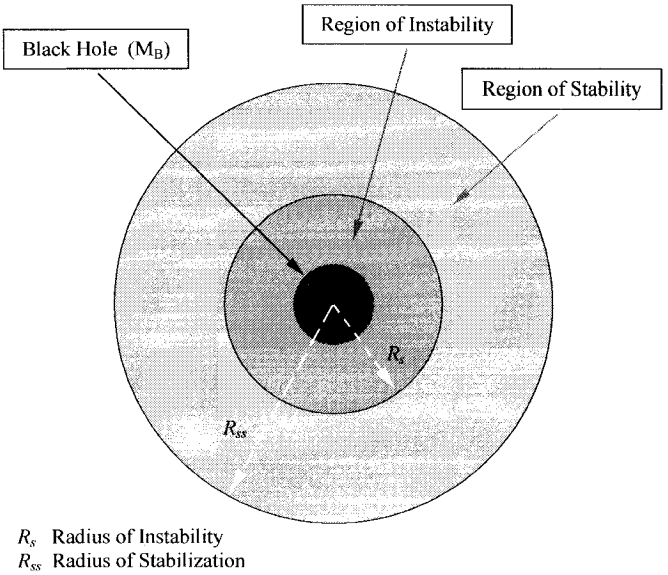


Figure 2. Regions of influence of a black hole.

Eq. (1.10) includes a correction factor that depends on the present value of deceleration parameter s_0. We shall afterwards comment on this. However, we soon note that formula (1.10) establishes a tie between s_0 and R_0 so that we can determine s_0 from the experimental value of R_0. If $R_0 \approx 30$ billion years, as indicated by Hubble telescope, the deceleration parameter must about be equal to 0.235 (formula (1.2)). From (1.2), (1.7) and (1.9) we go back to present value of density ρ_0. Finally, from density and volume we determine mass of visible Universe. Regarding the age of universe we use the classic formula of Weinberg (1972, p. 484) for a matter dominated universe in expansion. The formula is reported in Appendix 2 and gives the age as a function of s_0 and H_0. We find a value in perfect agreement with the experimental estimate from Hubble telescope. Any one familiar with classic GR could recognize the advancement provided by our theory.

2. STABILIZING EFFECT OF CENTRAL BLACK-HOLE

We have postulated that visible Universe rotates around a central black-hole. This is necessary in order that Universe had a spheroidal form and could evolve towards a stable structure. We preliminary recall general

properties of black-holes deriving from SR, Equivalence between Potential Energy and mass, and stability theory (Di Caprio, 2001).

2.1 Relativistic equations of a black-hole

By using Special Relativity (SR), the principle of equivalence Potential Energy/mass and Stability Theory, Di Caprio showed that a necessary condition for stability (of any body) around a black-hole is that the rotation radius be lager than a minimum quantity R_s (Section 5 about stability)

$$R_s = \frac{GM_B}{c^2}. \tag{2.1}$$

Also, a sufficient condition for stability is that rotation radius be equal to R_{ss}, with

$$R_{ss} = 2R_s = 2\frac{GM_B}{c^2}. \tag{2.2}$$

2.2 Black-hole in our cosmological model

In Part 2 we point out that mass of central black-hole has been growing in time up to present value $M_{B0} = M_0 = 2.54\times10^{53}$ Kg. However, contribution of black-hole to present rate of expansion is determined by quantity $M_B(t_0) = 0.558\ M_0 = 1.417\times10^{53}$ Kg(see sec 3.2), while at the end of expansion (i.e. at time $t_f = 14.39\times10^6$ years) it will be $M_B(t_f) = M_0 = 2.54\times10^{53}$ Kg. The difference between $M_B(t_0)$ and M_{B0} is due to (special) relativistic effects. Using formulas (2.1) and (2.2) we find
$R_s(t_f) = 1.8855\times10^{26}$ m ; $R_{ss}(t_f) = 3.771\times10^{26}$ m.

As (see 1.10) the present value of R_0 is 2.844×10^{26} m, it is $R_s(t_f) < R_0 < R_{ss}(t_f)$ i.e., roughly speaking 20×10^9 light-years < 30×10^9 light-years < 40×10^9 light-years. Therefore: the present radius of Universe is outside the stability region pertaining to present mass of central black-hole. Conversely it is inside the stabilization region pertaining to final mass of central black-hole. Expansion will continue until radius reaches the value $R(t) = R_{ss}(t_f) = 40\times10^9$ light-years. In Appendix 5 we show that $R(t_f) = 0$ while $R(t_0) > 0$.

In the following analysis we set
$R_f = R_{ss}(t_f) = 3.771\times10^{26}$ m $\approx 40\times10^9$ light-years.

2.3 The way in which stabilization occurs

Using the Principle of equivalence Potential Energy /mass, we show in the following Section that the mass of visible universe increases with time up to a value $M_f = 4.92\ M_0 = 12.5\times10^{53}$ Kg.

The growth results in growth of density, in spite of the parallel increase of volume. The final effect is that density reaches the critical value ρ_c (see Sec. 1). Then expansion ceases and a stable equilibrium is reached.

Central black-hole has a fundamental role in stabilization of Universe.

3. THE ROLE OF COMPLEXITY

3.1 Dark Energy

The existence of a central black-hole is at the roots of Complexity. Two masses M_1 and M_2, at a mutual distance R, are gravitationally bonded by a negative Potential Energy

$$E_p = -\frac{G M_1 M_2}{R}.$$

An analogous relation holds with regard to a ring of matter M_2 in rotation round a central body M_1 (fig.1). Postulating that this Energy is equivalent to mass and using the Einstein conversion factor, we find that the effective inertia of the rotating ring is a function of R:

$$M_2(R) = M_2 + \frac{E_p}{c^2} = M_2 - \frac{GM_1M_2}{Rc^2}$$

with c speed of light in vacuum. In particular if $R = R(t)$ then

$$M_2(t) = M_2 - \frac{GM_1M_2}{R(t)c^2}. \tag{3.1}$$

In our cosmological model we precisely assume that visible Universe is a ring of matter rotating round a central black-hole. Rotation is accompanied with radial expansion. Consequently mass is varying with time according to a law of type (3.1) and, in particular, it grows with expansion. The increase

ceases when expansion ceases. The following eq. gives us the relation between the present mass M_0 and the final mass M_f:

$$M_0 = M_f - \frac{GM^2{}_f}{c^2}\left(\frac{1}{R_0} - \frac{1}{R_f}\right). \tag{3.2}$$

Since we know the present mass, and we know R_0 (the present radius) and R_f (the final radius), we can use (3.4) for re-ascending from M_0 to M_f = 12.5×10^{53} Kg = 4.922 M_0.

Of course M_f should be considered the effective value of Universe mass, while M_0 is the corresponding value at present time, which turns out smaller because of the "darkening" effects of Potential Energy. Indeed in the interval ($t_0 < t < t_f$) we have

$$M(t) = M_f - \frac{GM_fM_f}{c^2}\left[\frac{1}{R(t)} - \frac{1}{R_f}\right] = M_f\left\{1 - \frac{GM_f}{c^2}\left[\frac{1}{R(t)} - \frac{1}{R_f}\right]\right\}$$

while the mass of central black hole is

$$M_{B0} = M_f - \frac{GM_fM_f}{c^2}\left[\frac{1}{R_0} - \frac{1}{R_f}\right] = M_0.$$

Dark Energy: *We call Dark Energy the coupling Potential Energy between rotating visible Universe and central black-hole.*

This energy is responsible for the fact that the present mass of visible Universe is smaller than the final value at the end of expansion, and density is increasing with expansion (contrary to traditional intuition).

Since we are dealing with a two-body problem we use the following differential equation of future expansion

$$\frac{M(t)M_B(t)}{M(t)+M_B(t)}\dot{R}^2 - \frac{GM(t)M_B(t)}{R} = \text{constant} \tag{3.3}$$

in which $M(t)$ is the (time varying) mass of visible Universe and $M_B(t)$ is the relativistic mass of central black-hole as defined by (Di Caprio, 2001)

$$M_B(t) = \left(1 - \frac{\dot{R}^2}{c^2}\right) M_{B0} \quad \rightarrow \quad M_B(t) = \left(1 - \frac{\dot{R}^2}{c^2}\right) M_0 .$$

Eq. (3.3) is equivalent to

$$\dot{R}^2 + K(t) = \frac{GM(t)}{R(t)}\left[1 + \frac{M_B(t)}{M(t)}\right] \tag{3.4}$$

which entails

$$\dot{R}^2 + K(t) = \frac{4\pi}{3} G\rho(t) R^2 \left[1 + \frac{M_B(t)}{M(t)}\right]$$

We can rewrite (3.4) as

$$\dot{R}^2 + K(t) = \frac{8\pi}{3} G\rho^* R^2 ; \qquad \rho^*(t) = \rho(t)\frac{1}{2}\left(1 + \frac{M_B(t)}{M(t)}\right) \tag{3.5}$$

which is recognized to be of the classical Einstein-Weinberg form

$$\dot{R}^2 + K = \frac{8\pi}{3} G\rho R^2 ; \quad K = (R_0 H_0)^2 (1 - 2s_0) \tag{3.6}$$

where (K/R^2) is "spatial curvature" according to GR (General Relativity).

Although this is only a formal analogy (we are not using GR but two-body Newtonian Dynamics and the Principle of equivalence Potential Energy / mass), it allows us to use noted results from GR and, in addition, to particularize such results in a definite numerical form. So, in order to determine the age of Universe we use Weinberg's eq. in Appendix 2. In order to determine the present time expansion speed we use eq.s (3.5), (3.6), in which we assume $K(t_0) = K$.

3.2 Present time expansion speed

In Appendix 3 we explicitly determine the value of K and find:

$$K = -\frac{s_0^2}{\alpha^{2/3}} = -1.325\times 10^{17}(m/s)^2 \quad \rightarrow \quad K(t_0) = K \tag{3.7}$$

$$\frac{\dot{R}^2(t_0)}{c^2} = -\frac{K(t_0)}{c^2} - \varphi_0 - \varphi_{B0}\,; \qquad \varphi_0 = \frac{GM_0}{R_0 c^2} = 0.6628$$

$$\varphi_{B0} = \frac{GM_0[1-\dot{R}^2(t_0)/c^2]}{R_0 c^2} = 0.6628\left[1-\frac{\dot{R}^2(t_0)}{c^2}\right]$$

Consequently

$$\frac{\dot{R}^2(t_0)}{c^2} = 1.474 - 0.6628 - 0.6638 + 0.6628\frac{\dot{R}^2(t_0)}{c^2}$$

$$\rightarrow \quad \frac{\dot{R}^2(t_0)}{c^2} = 0.6648$$

This yields the present value of expansion speed. Note that eq. (3.7) points out that K is independent from R_0 and H_0 though it satisfies eq. (3.6). Indeed from (3.7) and from (1.4) we can reascend to eq. (1.10) that gives the value of radius R_0. Such formulation is original and might be considered a peculiar fruit of our study. Moreover, in general, K must be substituted by $K(t)$ with $K(t) \neq K$.

3.3 Present time rotation speed

In Appendix 4 we derive the value of present time rotation speed:

$$\frac{v_{rot}(t_0)}{c} = 0.65 \quad ; \qquad \frac{v_{rot}^2(t_0)}{c^2} = 0.423\ .$$

Such speed is lower then the maximum value allowed by stability which, as shown by (Di Caprio, 2001) is v_{rot}/c = 0.78615. The relativistic mass coefficient due to rotation is

$$\gamma_{rot}(t_0) = \frac{1}{\sqrt{1-[v^2{}_{rot}(t_0)/c^2]}} = 1.317\,.$$

This quantity contributes to the explanation of the problem of missing mass.

4. THE PROBLEM OF MISSING MASS

Our formulation gives a satisfactory solution of the very well known problem of missing mass, which consists in this: experimental measurements of density derived from measurements of Luminosity in our Galaxy give about only 3% of the theoretical value obtained from observed cosmological red-shift. We explain the discrepancy by four separate effects:

1. The deceleration parameter is 0.2355 rather than 1;
2. The effective mass of rotating Universe is 4.922 the apparent mass at present time;
3. Rotation introduces a relativistic mass factor $\gamma_{rot}(t_0) = 1.5041$;
4. In addition to the mass of visible universe we must consider the mass of central black-hole (which is equal to the mass of visible Universe).

We then have $M_{tot} = (1{,}317\times 4.922 + 1)\, M_0 = 7.48\, M_0$ and consequently

$$\frac{s_0}{(M_{tot}/M_0)} = \frac{0.2355}{7.48} = 0.031 = 3.1\% \tag{4.2}$$

in agreement with experimental data. The final density of Universe turns out to be equal to the critical density:

$$\rho_f = \frac{M_f}{(4\pi/3)R^3{}_f} = 5.565\times 10^{-27}\, Kg/m^3 = \rho_c \tag{4.3}$$

The fundamental properties turn out to be verified:

1. radius R_f is the final "stabilization" radius, then expansion ceases when $R(t) \to R_f$;
2. when $R = R_f$ it is $M = M_f$ and universe density equals the critical density.

What is left is the determination of time t_f and the law of variation $R=R(t)$ in the interval $(t_0 < t < t_f)$. We show in Appendix 5 that $t_f = 1.0538\, t_0 = 14.39\times 10^9$ years, while at intermediate times

$$\begin{aligned} R(t) &= R_0 + 2.8\left[\left(\frac{t_f}{t_0} + \ln t_0\right)^2 - \left(\frac{t_f}{t} + \ln t\right)^2\right] R_0 \\ \dot{R}(t) &= 2\times 2.8\left(\frac{t_f}{t} - 1\right)\frac{R_0}{ct} \end{aligned} \tag{4.4}$$

with

$$R(t_f) = R_0 + 0.32592R_0 = R_f = 3.771\times10^{26}\,m;$$

$$\dot{R}(t_f) = 0; \quad \frac{\dot{R}(t_0)}{c} = 0.6648$$

Note that, as the density is bonded to critical density by $\rho = 2s\rho_c$ and $\rho_f = \rho_c$, the final value of deceleration parameter is $s_f = 0.5$. Also, as

$$s_f = -(R_f\ddot{R}_f)/\dot{R}_f^2 \quad \text{and} \quad \dot{R}_f = 0$$

it must be

$$\ddot{R}_f = 0 \quad \rightarrow \quad GM_{B0}/R_f = \gamma_{rot}v_{rot}^2(t_f),$$

which gives the final rotation speed of visible Universe as the solution of equation

$$\gamma_{rot}\frac{v_{rot}^2(t_f)}{c^2(t_f)} = 0{,}5 \quad ; \quad \frac{v_{rot}(t_f)}{c(t_f)} \approx 0.625.$$

In conclusion, when the expansion ceases, rotation goes on and has a permanent Equilibrium value. Note that our relations allow for the computation of density at any time in the interval $(t_0 < t < t_f)$:

$$\rho(t) = \frac{M(t)}{(4\pi/3)R^3(t)}$$

where $R(t)$ is defined from equation (4.4) and $M(t)$ from equations (3.3), (3.5), (3.6). The density is increasing in time (up to the critical density).

5. WHAT IS STABILITY ?

Stability is a basic property of the relative motion of a two-body system that kept mutual coupling in spite of disturbances. By "disturbance" we mean a perturbed "initial condition". The simplest way to introduce stability concepts in our context is to make reference to the following classical concepts. The total energy of a body in rotation round a second body

(gravitationally or electrically coupled) is defined by $E = T + E_p$ with T being the kinetic energy and E_p the potential energy. A necessary condition for stability is $T + E_p \leq 0$.

A sufficient condition is that the potential energy had a minimum at the reference Equilibrium (point that identifies the dynamical equilibrium). On the other hand in equilibrium $E_n = -m v^2 = -\gamma\, m_0 v^2$ while, on the other hand, T satisfies the relativistic eq. $T = (\gamma - 1)\, m_0\, c^2$. Thereby the above mentioned conditions can be formulated as conditions on the rotation speed and, in the final analysis, on the rotation radius. Taking account of this and, furthermore, replacing the mass m_0 with

$$\hat{m}_0 = m_0 + \frac{E_p}{c^2} = m_0\left(1 - \gamma\frac{v^2}{c^2}\right)$$

one shows (Di Caprio, 2001) that the following conditions are necessary and sufficient for relativistic stability:

$$\gamma\frac{v^2}{c^2} \leq 1 \quad \text{(necessary condition);} \qquad \gamma\frac{v^2}{c^2} = \frac{1}{2} \quad \text{(sufficient condition).}$$

The necessary condition identifies the maximum admissible value of the rotation speed and then the minimum admissible radius for stability. This radius is given by eq. (2.1). The region inside the circle of radius R_{min} represents a *forbidden* region: any body crossing the border of such region falls on the central body ("black-hole"). We call this region "region of influence of black-hole".

The sufficient condition identifies a specific rotation speed so that the circular trajectory covered by the external body is stable. Such speed corresponds to a stable rotation radius equal to twice R_{min} (see eq. (2.1)). Stability implicates that if, following a disturbance, the rotating body is moved outwards or inwards, the forces in action bring back the body to initial equilibrium.

6. SIMILARITY WITH THE HYDROGEN ATOM

The formula of the Hubble constant (eq. 1.4) and the formula of the Universe present radius (eq. 1.10) point out that the two-body structure of the Universe is similar to the two-body structure of the hydrogen atom. The scale factor is the adimensional quantity α_{mq} that represents the ratio between

gravitational and electric forces. This must be considered a result, more than an a-priori and unjustified assumption.

What is the deep physical meaning of such result? While deferring to successive studies an exhaustive analysis, here we underline the following analogy: inasmuch as electric quantization in hydrogen determines the Bohr radius (and the related radii αr_B and $r_\Gamma = \alpha^{4/3} r_B$), in a similar way the gravitational quantization determines the radius R_0 (apart from a correction depending on the deceleration parameter s_0 that accounts for the peculiarity of the cosmological model, i.e. the "expansion"). Also inasmuch as electric quantization can be related to rotation of the three quarks constituting the proton, so gravitational quantization should be related to rotation of central black hole. An implied inference is that a moving mass produces a gravi-magnetic field which is the correspondent of a classical electromagnetic field.

7. CONCLUSION

Using a two-body model we have been able to fully specify the present parameters of Universe as measured by Space Telescope Hubble, and much more: spheroidal form, radius, age, mass of visible universe, mass of central black-hole, density, critical density, Hubble constant, deceleration parameter, expansion speed, rotation speed. We completely solved the problem of missing mass and of dark energy. Universe is in expansion, but expansion is slowing down, up to cease completely in next 800 Millions years. The present mass is only a fraction of the final mass: the difference is due to the inertial effects of the coupling Potential Energy, which is negative. At any time such energy is inversely proportional to the expansion radius and, at any time, it determines subtraction, from final mass, of a decreasing amount of inertia. The present mass is about 1/5 of the final mass. In future its value and the resulting density will be larger. A final Equilibrium state will be reached in 800 Million years: afterwards expansion speed will reduce to zero and density will stop at a maximum value which is equal to the critical density of standard cosmological models. The deceleration parameter will be growing from the present value 0.2355 to 0.5 (final value). Thereafter Universe will continue to rotate, in a stable way, round the central black-hole. While the present radius is about 30 billions light-years, the final radius will be equal to 40 billions light-years. The two classical extremes, i.e. big crunch and indefinite expansion become excluded by our theory. From a methodological point our approach completely differs from classical GR (Einstein, De Sitter, Weinberg, Bondi, Friedmann) and from modified GR (e.g. Dicke, De Witt). We have assumed decoupled

"absolute space" and "absolute time" in the spirit of Newton and, in this respect, one could speak of "post-newtonian" formulation (which vaguely resembles Milne's approach). In reality our distinguishing feature is the use of complexity theory and of relativistic (in the sense of SR) stability theory, altogether with the Principle of Equivalence Potential Energy/Mass. They allowed us to point out a negative coupling energy between visible Universe and central black-hole, that explains the "missing mass" and the fact that, contrary to traditional intuition, the density is growing with expansion (up to a future value equal to the "critical density" of classical GR). From a formal point of view our eq.s look like a generalization of those presented by (Weinberg, 1972) for a matter dominated Universe. This motivates our preference for Weinberg's results. Of course Weinberg and others do not account for rotation. On the contrary we have shown the necessity of that and pointed out the existence of a final dynamical Equilibrium. Last but not least, as outlined in Sec. 5, *our theory puts together stability theory and Special Relativity* (*in a generalized version that implicates space-time decoupling and application to curvilinear motions*). By this way we were able to determine the relativistic mass factor γ_{rot} pertaining to rotation (which contributes to the difference between observed mass and effective mass) and, in addition, we proved that the stabilization radius around a black-hole is but the Schwarzschild radius of classical GR.
In Part 2 we show that accurate knowledge of present state and present Energy allow us to extrapolate past evolution from time zero, when expansion begun.

REFERENCES

See Part 2.

Appendix 1. Hubble constant

Equation (1.4), which gives the Hubble constant, is equivalent to

$$H_0^{-1} = \frac{1}{\alpha_{mq}} \frac{r_\Gamma}{c} \quad ; \quad r_\Gamma = \alpha^{4/3} r_B = \alpha^{1/3} \left(\frac{h/2\pi}{m_0 c} \right) \tag{A1.1}$$

As $m_0 c^2 r_\Gamma^3 = \alpha\,[h/2\pi]^3$ radius r_Γ is the average radius of electromagnetic coupling in Hydrogen and (c/r_Γ) is the average time of crossing of said atom. Eq. (A1.1) tells us that Hubble time is proportional to average time in question by a cosmological factor equal to the electro-gravitational ratio.

Appendix 2. Age of Universe

The classic Weinberg's formula (1972, pp 481-484) giving the age of Universe in expansion in a matter dominated era is

$$t_0 = H_0^{-1} = \left[(1-s_0)^{-1} - s_0(1-s_0)^{-3/2} \cosh^{-1}\left(\frac{1}{s_0} - 1 \right) \right]$$

Using our values of H_0 and s_0 we find $t_0 = 13.65\times10^9$ years.

Appendix 3. Derivation of the expansion constant $K = K(t_0)$

The following formula is quoted in (Weinberg, 1972, page 481) $K = (H_0 R_0)^2(2s_0 - 1)$. With reference to eq. (3.4) we set $K(t_0) = K$ and, since we have explicit formulas for H_0 and R_0 we find

$$K = -\frac{s_0}{\alpha^{2/3}} c^2 .$$

Appendix 4. Present time rotation speed

As the deceleration parameter is defined by equation (1.1) and

$$\ddot{R}(t_0) = \gamma_{rot} \frac{v_{rot}^2}{R_0} - \frac{GM_{B0}}{R_0^2} \quad \rightarrow \quad \dot{R}(t_0)s_0 = \frac{GM_{B0}}{R_0} - \gamma_{rot} v_{rot}^2$$

from which we obtain

$$\gamma_{rot} \frac{v_{rot}^2(t_0)}{R_0} = \frac{GM_{B0}}{R_0^2 c^2} - \frac{\dot{R}^2(t_0)}{c^2} s_0$$

Using our values of s_0, R_0, M_{B0}, $R(t_0)$ we find $v_{rot}(t_0)/c = 0.65$; $v^2_{rot}/c^2 = 0.423$, $\gamma_{rot} = 1.317$.

Appendix 5. Computation of final time for the end of expansion

Denoting t_f the time at which R reaches its stabilization value R_f, we determine t_f from R_f (and from t_0, R_0) by solving the following non-linear problem: in the interval (t_0, t_f) it must be

$$\dot{R}(t) = b\left(\frac{t_f}{t} + \ln t_0\right)\frac{R_0}{t}$$

$$R(t) = R_0 + \frac{b}{2}\left[\left(\frac{t_f}{t_0} + \ln t_0\right)^2 - \left(\frac{t_f}{t} + \ln t_0\right)^2\right]R_0$$

with b and t_f such that $R(t_f) = \kappa_f$. As $R_f = 1.3259\ R_0$, the problem is solved with $b = 2.8$, $t_f = 1.0538\ t_0$; $(t_f/t_0 - 1) = 0.0538$. We can check that

$$\frac{b}{2}\left[\left(\frac{t_f}{t_0} + \ln t_0\right)^2 - (1 + \ln t_0)^2\right] = 0.3259$$

$$\rightarrow \quad R(t_f) = R_0 + 0.3259 R_0 = R_f$$

while, on the other hand

$$\frac{\dot{R}(t_0)}{c} = 2b\left(\frac{t_f}{t_0} - 1\right)\frac{R_0}{t_0 c} = 0.6648\,.$$

COMPLEXITY IN UNIVERSE DYNAMIC EVOLUTION.
PART 2 - PRECEDING HISTORY

Umberto Di Caprio
Stability Analysis s.r.l., Via Andrea Doria 48/A, 20124 Milano, Italy

Abstract: We define the initial structure of Universe and the disturbance that destroys the dynamical equilibrium relating to such structure. We reconstruct subsequent evolution and explain how expansion led to a matter dominated era. Further on we illustrate the birth of galaxies and the growth up to present time. According to our model Universe was transparent up to the beginning of the matter dominated era and afterwards, for a certain time, was obscured by the presence of central black-hole. A conjecture about ether is set forward.

Key words: Universe evolution; time zero; universe energy; two-body model; birth of galaxies; cosmic radiation; universe temperature; ether.

1. INTRODUCTION

In Part 1 we defined a two-body dynamical model of Universe, explaining present state (age, radius, mass, density, critical density, expansion and rotation speeds, energy, missing mass) and next foreseeable evolution towards a final state of dynamical equilibrium. Such evolution is implied in present state since the latter does not identify a stable equilibrium and, on the other hand, it is characterized by an expansion speed greater than zero. Here we go back in time and show that present state is the point of arrival of a continuous process of expansion, started 13.65 billions years ago. During such process the masses of visible Universe and of central black-hole were not at all constant but, on the contrary, grew-up with convenient dynamical laws. We define the "initial state" and determine the sequence of "intermediate states" up to now. We keep some strong holds.

1. The past structure was similar to present structure (i.e. a two-body system formed by a ring of matter in rotation round a central body).
2. The Universe total Energy is constant and equal to present value. Then

$$E_u(t) = E_u(t_0)\,; \qquad E_u(t_0) = [M_0 + M_{B0}]c^2 = 3.55\times 10^{70}\,\text{Joule}$$
$$E_u(t) = [M(t) + M_B(t)]\,c^2(t); \tag{1.1}$$
$$M_0 = 2.54\times 10^{53}\,\text{Kg}; \qquad M_{B0} = 0.558 M_0$$

3. The observations by Hubble telescope and others show that past values of radius were smaller, and even much smaller going back in time. Consequently the past values of dark energy (that we identified with the negative gravitational energy that couples the two bodies) were larger, and even much larger (in absolute terms). Then the past value of observable masses were smaller than nowadays. We shall see that the initial value of the mass of visible Universe was the electron mass m_0 and that of the mass of central body was the proton mass m_p. In conclusion, the universe evolution from time zero to nowadays was marked by a simultaneous increase of radius and masses. As a consequence of (2) and (3) the speed of light in vacuum has been varying with time:

$$c(t_0) = c = 3\times 10^8\,\text{m/s}\,; \qquad c(t) > c \quad \text{for} \quad t < t_0 \tag{1.2}$$

4. The birth of Galaxies can be dated at 13.25 Millions years ago. We identify this time with the time when (according to our dynamical model) visible Universe went out the region of instability pertaining to central black-hole (see Part 1).
5. The matter dominated era begun at $t_m \approx 380\times 10^3$ years $\approx 1.2\times 10^{13}$ s, when, according to Weinberg (1972, p. 540) the temperature fell down 4000 °K. Before such time the radiation was dominant. We show that, in our model, t_m represents the time when the radius of the region of influence of (growing) central black-hole became equal to the radius of visible Universe. So, for $t < t_m$ the Universe was "transparent" while, for $t \geq t_m$ it was "obscured". Such black-out lasted until birth of galaxies (see preceding point (4)).
6. We postulate that inasmuch the speed of light in vacuum varies with time, so does the gravitational constant and, moreover

$$\frac{G(t)}{c^2(t)} = \frac{G}{c^2} = 7.42 \frac{\text{Joule} \cdot \text{m}}{\text{Kg}^2 (\text{m/s})^2} \tag{1.3}$$

The above Postulate gives consistency among the various equations that involve time-varying masses, time-varying speed of light, growing radius of visible Universe, decreasing dark energy and time variation of the region of influence of central black–hole. The quantity $\mu_G = (G/c^2)$ defines the "gravitational permeability" which is the gravitational correspondent of electromagnetic permeability μ_0.

7. Since the speed of light is a function of time we define an instantaneous permittivity $\varepsilon(t)$ and an instantaneous magnetic permeability $\mu(t)$ so that

$$\varepsilon(t) \cdot \mu(t) = 1/c^2(t) \tag{1.4}$$

8. Before the "time zero" Universe already had a two-body dynamical structure, in Equilibrium by the action of the gravitational force only plus the centrifugal force. Its inertia was equal to zero and its energy was a pure binding energy (nor kinetic nor potential).

2. EVOLUTION FROM TIME ZERO

We call "time zero" the time when expansion begun and start our analysis by subdividing $(0,t_0)$ in three parts. The first $(0,t_1 = 1.103$ s) where t_1 is the time when temperature was 10^9 K°, the second (t_1=1.103 s, t_2=1.103×10^{13} s) where t_2 is the time when temperature fell down to about 4000 K°, the third (t_1,t_0) with $t_0 = 13.65\times10^9$ years (present time). The second interval is pointed out by Weinberg in his book "Gravitation and Cosmology" (1972, pp 481-490; p. 540). In particular t_2 coincides with the time of beginning of matter-dominated era. In the interval in question the ratio (R/R_0) and the temperature Θ are known from thermodynamics.

We assume that the energy of Universe satisfies equations (1.1) (namely Universe total energy is finite, constant and equal to present value). Consequently the speed of light must be a function of time.

2.1 Insertion of our results (Part 1) in Weinberg's results

Weinberg's analysis of Universe evolution in interval (t_1, t_2) can be made plainer inserting in it our value of radius $R_0 \approx 30\times10^9$ light-years (such value agrees with the experimental one given by STH). So

$$R(t_1) = R(1.103) = 1.9\times10^{-10} R_0 = 5.4\times10^{16}\,\mathrm{m}$$
$$R(t_2) = R(1.2\times10^{13} s) = 6.3\times10^{-4} R_0 = 1.79\times10^{23}\,\mathrm{m} \tag{2.1}$$

From this we extrapolate dynamic behaviour in "early age" and, in parallel, subsequent evolution in (t_2, t_0).

To complete the picture we correlate Weinberg's values of temperature (for different ratios R/R_0) with a formula that involves the energy of visible Universe (Sec.6).

In order to establish the expansion law before 1.103 s we need defining an initial condition. This condition must be consistent with preceding equations and must explain dark energy effects in terms of masses that vary with time.

2.2 The Initial Condition

Reasons of theoretical consistency impose that the initial condition could not be other than an Equilibrium condition referring to a two-body system (fig.1). We call Big-Bang the sudden destruction of this Equilibrium and the consequent start of radial expansion. We formulate a precise conjecture and validate the conjecture through its effectiveness in explaining all subsequent evolution.

Before expansion, Universe was consisting of a Graviton, formed by a ring of matter with mass equal to the electron mass m_0, in stable rotation round a concentrated mass equal to the proton mass m_p (fig. 1). *Such assumption about initial value of the two masses turns out fully validated by following analysis. In fact growth of the masses in question by reason of expansion univocally explains the value of the time when the matter-dominated era began* (Sec. 4). The energy of this two-body system was equal to $E_u = 3.55\times10^{70}$ Joule $= 3.957\times10^{53}$ Kg c^2. The "equivalent mass", however, was equal to zero. The coupling was gravitational only. Equations were

$$\frac{m_p m_0}{m_p + m_0} + \frac{E_p(0^-)}{c^2(0^-)} = 0\,; \quad E_p(0^-) = \frac{G(0^-) m_p m_0}{R(0^-)}$$
$$(m_p + m_0)c^2(0^-) = E_{Tot}; \quad \frac{G(0^-)}{c^2(0^-)} = \frac{G}{c^2} \tag{2.2}$$

with $R(0^-)$, $c(0^-)$, $G(0^-)$ initial values of radius, speed of light, gravitation constant; and with $E_p(0^-)$ initial value of Potential energy.

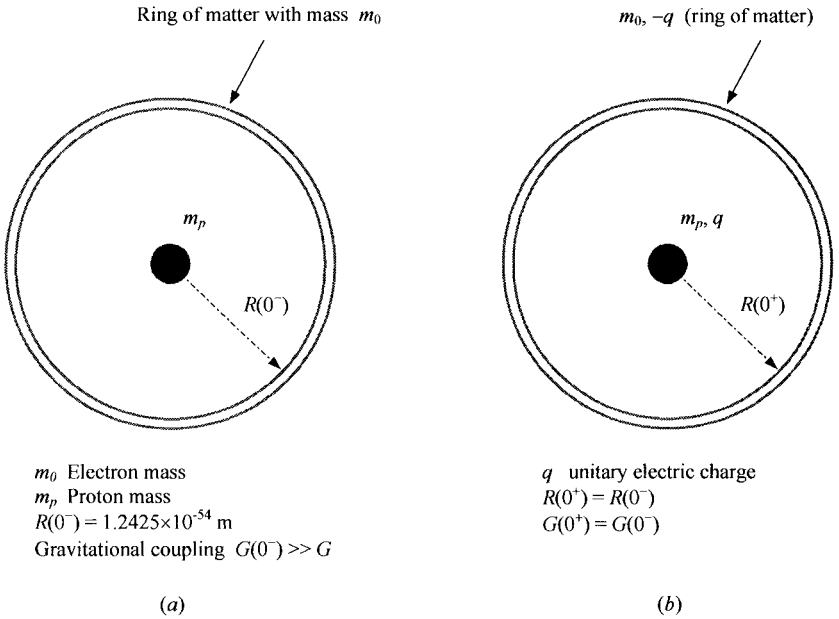


Figure 1. Universe initial structure.

Eq. (2.2) tells us the postulated equivalence Potential energy/mass and tells us that the "equivalent mass" of the two-body system is equal to zero and then the system is a graviton. From above equations we get

$$R(0^-) = \frac{G}{c^2}(m_p + m_0) = 1.2425\times10^{54}\,\mathrm{m}$$

$$c^2(0^-) = \frac{E_{Tot}(0^-)}{m_p + m_0} = 2.12\times10^{97}\,\mathrm{m}^2/\mathrm{s}^2 \tag{2.3}$$

$$\frac{G(0^-)}{c^2} = 2.35\times10^{80}; \quad G(0^-) = G\frac{c^2(0^-)}{c^2} = 1.575\times10^{70}\,\frac{\mathrm{Joule}\cdot\mathrm{m}}{\mathrm{Kg}^2}$$

The above data completely define the initial condition i.e. the initial dynamic equilibrium. After assuming that rotation speed was determined by the relativistic stability criterion proposed by (Di Caprio, 2001b) we also derive the pre-disturbance rotation time T_0

$$\frac{v_{rot}}{c} = 0.624; \qquad T_0 = \frac{2\pi R(0^-)}{0.624\,c(0^-)} = 2.734\times10^{-102}\,\mathrm{s}.$$

2.3 Rupture of initial Equilibrium

What perturbation did destroy the aforesaid dynamical Equilibrium? We propose a conjecture that works fine for practical purposes, as it satisfactorily allows us to reconstruct the Universe dynamics up to Weinberg's time t_1=1.103 s and, afterwards, up to time t_2=1.2×10^{13} s when matter-dominated era begun. (Further on, stability theory and conservation of energy will lead us to understanding evolution from time t_2 to present time t_0). The perturbation is the sudden appearance of an electric polarization, so that m_p and m_0 abruptly acquired a positive and a negative unitary electric charge, respectively (fig. 2). This determined the birth of an electric coupling that summed up to the original gravitational coupling, causing a modification of energy balance. In principle this could have determined the violation of conservation law. Conservation however was kept thanks to the instantaneous raise of a (positive) Kinetic energy that counterbalanced the negative Potential energy of electric origin. Such Kinetic energy was the expansion energy. Expansion determined the growth of radius and, because of the constraint of conservation of total energy, a slowing down of expansion itself. Dynamical evolution can be rebuilt through the following steps.

3. FROM TIME 0 TO WEINBERG'S TIME 1.103 SECONDS (VERY EARLY UNIVERSE; EARLY UNIVERSE; BIRTH OF ETHER)

We assume that expansion was accompanied by rotation and that rotation speed was equal to the maximum value allowed by relativistic stability (i.e. 0.78615c, see Di Caprio, 2001). We subdivide the interval $(0 < t < 1.103$ s) in two parts:

1. The interval $(0< t < t_L)$ with t_L=1.096×10^{-61} s that identifies "very early" Universe;
2. The interval $(t_L < t < 1.103$ s) that identifies "early" Universe.

In interval (1) the radius grows proportionally to time and the expansion speed is constant. It is

$$R(t) = R(0^-)\cdot\frac{t}{T_0}; \qquad \dot{R}(t) = \frac{R(0^-)}{T_0} \tag{3.1}$$

$$R(t_L) = 1.475\times10^{-14}\,\mathrm{m};$$

$$\dot{R}(t_L) = 0.458\times10^{-48}\,\mathrm{m/s}; \qquad t_L = 1.096\times10^{-61}\,\mathrm{s}$$

In interval (2) radius grows proportionally to $t^{-1/2}$ (as usually assumed for a radiation dominated Universe). One shows (see Appendices) that

$$R(t) = R(t_L + \Delta t_L)\sqrt{\frac{t}{1.488\times 10^{-68}\,\mathrm{s}}}; \qquad R(t_L + \Delta t_L) = 1.993\times 10^{-14}\,\mathrm{m}$$

$$R(1.103) = 5.4\times 10^{16}\,\mathrm{m}$$

in agreement with Weinberg's model and with our estimate of present radius R_0. The transition from "very early" to "early" Universe is marked by a negative jump of the expansion speed. In early Universe

$$\dot{R}(t_L^+) = \frac{1}{2}\frac{R(t_L)}{t_L} = \frac{1}{2}\cdot\frac{1.475\times 10^{-14}}{1.096\times 10^{-61}} = 0.067\times 10^{48}\,\mathrm{m/s}$$

Such value is smaller than value referring to very-early Universe. In fact we have

$$\dot{R}(t_L^-) = 0.458\times 10^{48}\,\mathrm{m/s}; \qquad \dot{R}(t_L^+) < \dot{R}(t_L^-)$$

The jump points out a sudden loss of kinetic energy due to some event. We postulate that the event is the production of photons, i.e. generation of ether, when magnetic permeability falls down the nowadays value $\mu_0 = 4\pi\times 10^{-7}$ and stabilizes itself. Stabilization is precisely due to generation of photons and occurs at time $t = t_L = 1{,}096\times 10^{-61}$. The total energy of ether is

$$E_\Phi = \frac{1}{2}m_0(t_L)[R^2(t_L^-) - R^2(t_L^+)]$$

$$= \frac{1}{2}425 m_0[(0.458\times 10^{48})^2 - (0.067\times 10^{48})^2] = 3.97\times 10^{67}\,\mathrm{Joule}$$

It satisfies the relation

$$E_\Phi = N_\Phi \cdot m_0 c^2 \qquad N_\Phi = 4.85\times 10^{80} \text{ (number of photons).}$$

The value of $m_0(t_L)$ has been determined from the conservation equation

$$[m_p(t_L) + m_0(t_L)]c^2(t_L) = E_{Tot} = (m_p + m_0)c^2(0^-)$$

in which

$$c^2(t_L) = \gamma_r^2 \dot{R}^2(t_L) = 2.618(0.458\times10^{48})^2 = 5.48\times10^{95}(\text{m/s})^2$$

In conclusion the ether was born at time $t = t_L = 1{,}096\times10^{-61}$ s and consisted of N_Φ photons. (N_Φ is about equal to the number of protons forming visible Universe.) We emphasize that this conjecture about ether needs to be validated by experimental data, of the type collected by Hubble telescope.

In the above formula of the speed of light (and in subsequent ones) $\gamma_r = 1.618$ is the relativistic mass coefficient corresponding to a rotation speed $(v_r / c) = 0.78615$ that, as said at the beginning of this Section is the assumed rotation speed of m_0 around m_p. It is noted that γ_r equals the famous "golden ratio" and this represents an important finding of our study. Further note that the sum of the rotation speed and the expansion speed must not exceed $c(t)$ and, for this reason, the expansion speed is smaller than the speed of light.

Before time t_L the magnetic permeability was varying according to

$$\mu(t) = 4\pi \frac{m_0}{q^2}\frac{R(t)}{\gamma_r^2} \quad ; \qquad \mu(t_L) = 4\pi \frac{m_0}{q^2}\frac{R(t_L)}{\gamma_r^2} = 4\pi\cdot10^7 = \mu_0$$

$$\frac{R(t_L)}{\gamma_r^2} = \frac{1.475\times10^{-14}\,\text{m}}{2.618} = 2.8179\times10^{-15}\,\text{m} = \alpha^2 r_B$$

where α is the fine structure constant and r_B the Bhor radius.
On the other hand μ_0 satisfies the equation $\mu_0 = 4\pi(m_0/q^2)(\alpha^2 r_B)$ and the elementary photon is made up of an electron and a positron rotating round their center of mass at a distance equal to $\alpha^2 r_B$ (Di Caprio, 2003).
The above equations implicate that the Coulomb's constant and the speed of light vary according to

$$k(t) = \frac{m_0}{q^2} R(t)\dot{R}^2(t); \qquad c^2(t) = \frac{k(t)}{(\mu(t)/4\pi)} \rightarrow c^2(t) = \frac{1}{\varepsilon(t)\mu(t)}$$

In interval (2), from $t_L = 1.096\times10^{-61}$ s to $t_1 = 1.103$ s, it is

$$\mu(t) = \mu_0 = 4\pi\times10^{-7} = \text{const}\,; \qquad k(t) = \frac{\mu_0}{4\pi}c^2(t) = \frac{\mu_0}{4\pi}\gamma_r^2\dot{R}^2(t)$$

with

$$\dot{R}(t) = 1.993\times10^{-14}\frac{1}{1.488\times10^{-61}}\frac{1}{\sqrt{t/1.488\times10^{-61}}} = \frac{5.166\times10^{16}}{\sqrt{t}} .$$

4. UNIVERSE WAS TRANSPARENT UP TO BEGINNING OF MATTER-DOMINATED ERA

By extending our equations up to time $t_m = t_2 = 1.2\times10^{13}$ s that marks the beginning of the matter-dominated era, we find

$$R(t_m) = 1.79\times10^{23}\,\mathrm{m}\,; \qquad \dot{R}(t_m) = 49.75\,c \tag{4.1}$$

Eq. (4.1) allows us show that t_m is but the time when the region of influence of the central black hole embraced visible Universe (figures 3, 4, 5, 6). More precisely, for $t < t_m$, such region was smaller than the radius of visible Universe and ,consequently, Universe was transparent. On the contrary, at $t=t_m$, Universe was instantaneously darkened and this determined an abrupt variation of the expansion law: matter-dominated era begun. To see this, note that shrinkage of radial speed from $t = 0$ to $t_m = 1.2\times10^{13}$ s caused an increase of both mass of visible Universe and of black-hole. However, the conservation of energy imposed

$$[m_p(t) + m_0(t)]\,c^2(t) = (m_p + m_0)c^2(0^-) = (m_p + m_0)c^2(0^-)$$

$$[m_p(t) + m_0(t)] = \frac{c^2(0^-)}{c^2(t)}(m_p + m_0) = c^2(0^-)(m_p + m_0)/2.618\dot{R}^2(t)$$

$$m_p(t_m) = \frac{c^2(0^-)}{2.618\dot{R}^2(t_m)}m_p = \frac{2.12\times10^{97}}{2.618(24.99c)^2}m_p = 2.418\times10^{50}\,\mathrm{Kg}$$

$$m_0(t_m) = m_p(t_m)/1836 = 1.314\times10^{47}\,\mathrm{Kg}$$

This means (see our preceding formulation) that at $t = t_m$ masses of visible Universe and of central black hole were respectively given by $M(t_m) = m_0(t_m) = 1.314\times10^{47}$ Kg and $M_B(t_m) = m_p(t_m) = 2.41\times10^{50}$ Kg. As the radius of the instability region pertaining to a black-hole is defined by (Di Caprio, 2001, also see Part 1) $R_s=(G/c^2)M$ we get $R_s(t_m)=(G/c^2)M_B(t_m)$ $=1.79\times10^{23}$ m. By comparison with (4.1) it is inferred that $R_s(t_m) = R(t_m)$ (whilst $R_s(t) < R(t_m)$ or $t < t_m$).

5. FROM BEGINNING OF MATTER-DOMINATED ERA TO NOW

We divide the interval (t_m,t_o) in (t_m,t_s) and (t_s,t_o) where t_s is the time when growth of central black hole came to an end and visible Universe went out again the region of instability around the black-hole.

5.1 Time t_s when visible Universe came out of the region of influence of central black hole

From time t_m to time $t_s = 0.4\times10^9$ years the visible Universe was darkened by central black-hole and the mass of central black-hole grew up linearly according to

$$M_B(t) = (t/t_m)M_B(t_m)\,; \qquad M_B(t_s) = M_{B0} = 2.54\times10^{53}\,\mathrm{Kg}$$

the radius of the corresponding region of influence reached the value $R_s = 1.8855\times10^{26}$ m:

$$R_s = \frac{G}{c^2}M_{B0}$$

Meanwhile the radius of visible universe grew up linearly as well

$$R(t) = (t/t_m)\cdot R(t_m)$$
$$R(t_s) = \frac{0.4\times10^9\,\mathrm{years}}{1.2\times10^{13}\,\mathrm{s}}\cdot 1.79\times10^{23}\,\mathrm{m} = 1.8855\times10^{26}\,\mathrm{m} = R_s(t_s) \tag{5.1}$$

In the final analysis, it was $R(t) = R_s(t)$ in the whole interval in question, which means that visible Universe was continually darkened.
Growth of black-hole definitively ceased, however, at time t_s (namely it was $R_s(t) = \text{const} = R_s$ for $t > t_s$) whilst expansion of visible Universe went on. Consequently the radius of Universe satisfied the relation $R(t) > R_s(t)$ for $t > t_s$ (that indicates the emission of visible Universe from "forbidden region").

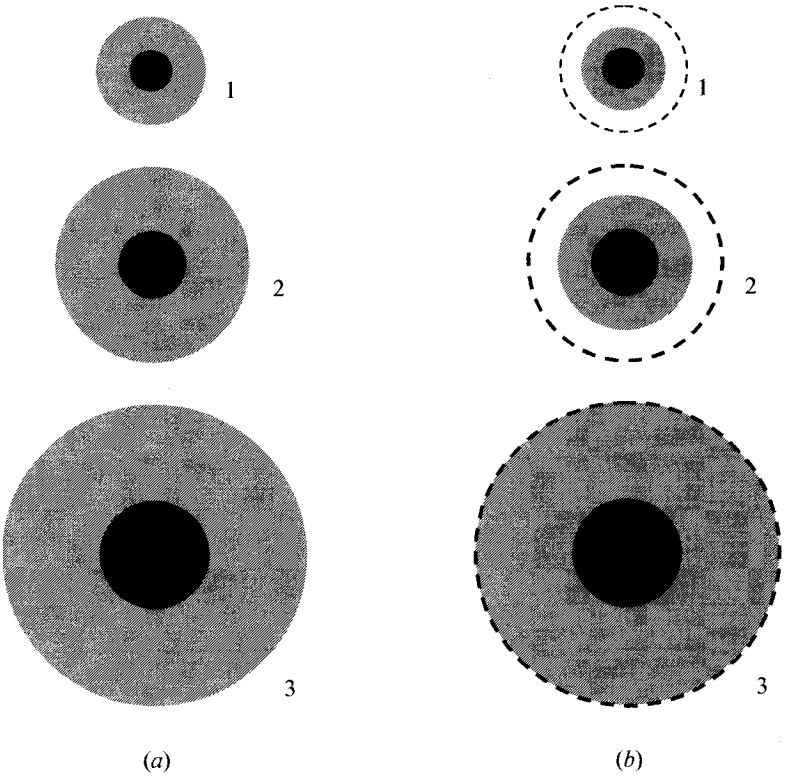


Figure 2. (*a*) Growing black hole and growing region of instability. (*b*) Early Universe: growing region of instability (gray) and growing radius of Visible Universe

Note: it is $(t_0 - t_s) = 13.25\times10^9$ years and this time interval corresponds to the age of first galaxies, as observed by Space telescope Hubble.

So according to our model first the galaxies were born when visible Universe out-came the instability region around central black-hole. We underline that, as a consequence of eq. (5.1) the expansion speed was constant during the black-out (cfr. with eq. (4.1))

$$\dot{R}(t) = \text{const} = \frac{1}{2}\frac{R_m}{t_m} = 49.75\,c \qquad \text{in} \quad (t_m, t_s)$$

Therefore the mass of visible Universe remained constant in (t_m, t_s). while at time t_s it abruptly acquired a value about 780 times bigger, i.e. about 10^{50} Kg $\approx M_0/49.75^2$. Then expansion abruptly slowed down (see next paragraph).

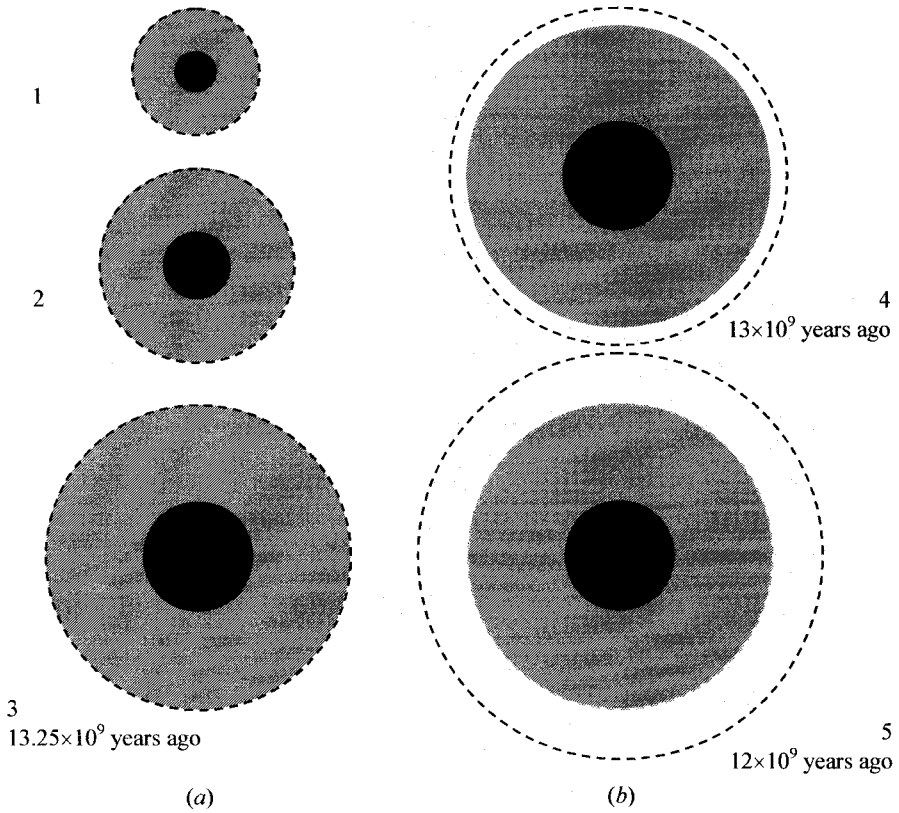


Figure 3. (*a*) Darkening interval from t_m =380×10^5 years to t_G = 425×10^6 years; in this interval the Visible Universe was obscured by black hole. (*b*) Birth of galaxies: at time t_G = 425×10^6 years the visible universe exceeds the border of instability region

5.2 From time t_s to time t_0 (from birth of galaxies to present time)

The following law rules the evolution of Universe from time t_s to present time:

$$R(t) = R_s + \dot{R}(t_s^+)\left(\frac{t-t_s}{t_s}\right) - R_s\left(\frac{t-t_s}{t_s}\right)^{0.426} \quad ; \qquad \dot{R}(t_s^+) = 5.463c$$

$$\dot{R}(t) = \dot{R}(t_s^+) - \frac{R_s}{t_s} 0.426 \left(\frac{t - t_s}{t_s} \right)^{0.426-1}$$

$$= 5.463c - 1.505c\, 0.426 \left(\frac{t - t_s}{t_s} \right)^{0.426-1}$$

The required edge conditions are verified:

$$R(t_0) = 2.844 \times 10^{26}\,\mathrm{m} = R_0; \qquad \dot{R}(t_0) = 0.6648c\,.$$

Note that as

$$\dot{R}(t_s^-) = \dot{R}(t_m) = 49.75\ c \quad \text{it is} \quad \dot{R}(t_s^+) << \dot{R}(t_s^-)\,,$$

the birth of galaxies strongly slowed down the expansion.

6. TEMPERATURE AND BACK-GROUND COSMIC RADIATION

We use the following formula for the determination of temperature of visible universe during the "matter dominated era" (from t_m to t_0):

$$\frac{\Theta(t)}{1K^0} = \frac{(64/81)\alpha^2 r_B}{R(t)} \left(\frac{M_0 c^2}{k_B} \cdot \frac{1}{1\,°\mathrm{K}} \right)^{4/9} \tag{6.1}$$

where k_B is the Boltzmann constant. The formula is correlated to the entropy formula and allows one to prove that Universe entropy has a local minimum at the final Equilibrium point. Here ,however, we do not afford this issue. We note that (6.1) is numerically equivalent to

$$\frac{\Theta(t)}{1°\mathrm{K}} = \frac{7.158 \times 10^{26}\,\mathrm{m}}{R(t)}$$

and then it gives

$$\Theta(t_m) = \Theta(1.2\times10^{13}\text{s}) = \frac{7.158\times10^{26}}{1.79\times10^{23}}\ °\text{K} = 4000\ °\text{K}$$

in agreement with Weinberg thermo-dynamical model. In addition we can determine the Universe temperature at time t_s, when radius was 1.8855×10^{26} m, and at present time (radius = 2.844×10^{26} m).
We find $\Theta_s = 3.797$ °K ; $\Theta_0 = 2.511$ °K ; $(\Theta_s + \Theta_0)/2 = 3.15$ °K.
Θ_s can be identified with the temperature of fossil radiation discovered by Penzias and Wilson (Weinberg, 1972). So our formulation tells us that fossil radiation is bonded to birth of galaxies.

7. CONCLUSION

The accurate knowledge of present state and of present energy (Part 1) allowed us to extrapolate past evolution from time zero, when expansion began. The Universe initial state was defined by a dynamical Equilibrium involving the gravitational force and a rotating ring of matter (with mass equal to the electron mass) round a central body with mass equal to the proton mass. Destruction of this initial Equilibrium, owing to sudden appearance of an electric polarization, gave start to expansion and determined a well defined series of events, among which the birth of ether, the reach of the matter dominated-era (about 385 thousands years from beginning) and the subsequent evolution up to now. Using stability theory we have dated the birth of galaxies at time 0.4 billion years, i.e. about 13.25 billion years ago, which agrees with STH observations. Also, we have correlated such event with the microwave back-ground cosmic radiation. As regards the start of the matter dominated era we have seen that, in precedence, Universe was transparent but the mass of central body (black-hole) was growing faster than radius of visible Universe and finally its "region of influence" darkened the visible Universe. The black out ceased when such region stopped its increase: then galaxies were born. The early stages of Universe life have been subdivided in two parts: very early Universe, when radius grew up linearly with time, and early Universe when growth was proportional to square root of time (like usually assumed). The transition from one another determines an energy jump that we attributed to the birth of ether (in the form of convenient photons).

From the point of view of theory we have ignored General Relativity and have based our study on "generalized" Special Relativity, Stability theory and the equivalence Potential Energy/mass. In particular we have assumed flat Euclidean space and absolute time, totally decoupled from space. Like in

Part 1 we have kept a two body structure and this is an essential part of our study. Another key point is the postulate of conservation of energy which allowed us to maintain a stronghold in the whole Universe evolution up to final equilibrium. Conversely Universe physical parameters (speed of light, Gravitation "constant", electric permittivity, electromagnetic permeability) have been assumed to be continually varying. This is a novelty with respect to any other formulation in literature.

Appendix 1. Early Universe and very-early Universe

We set

$$R(0^+) = R(0^-); \qquad c(0^+) = c(0^-); \qquad R(0^+) = c(0^+)/\gamma_r$$

$$\frac{1}{2} m_0 \dot{R}^2(0^+) = \frac{k(0^+)q^2}{R(0^+)} = \frac{k(0^+)q^2}{R(0^-)}$$

where $k(0^+)$ is the initial value of Coulomb's constant. The above equations implicate

$$k(0^+) = \frac{m_0}{q^2} R(0^+) \frac{c^2(0^+)}{2\gamma_r^2} = \frac{m_0}{q^2} R(0^-) \frac{c^2(0^-)}{2\gamma_r^2} = 1.78 \times 10^{50} \frac{\text{Joule} \cdot \text{m}}{\text{Coulomb}}$$

and, as $k(t_0) = 8.854 \times 10^{-12}$ it is $k(0^+) >> k(t_0)$. This means that we are proposing a dynamic model with time varying parameters.
We subdivide the interval $(0 < t < 1.103 \text{ s})$ in two parts:

1. The interval $(0 < t < t_L)$ with $t_L = 8.54 \times 10^{-62}$ s that identifies very early Universe;
2. The interval $(t_L < t < 1.103 \text{ s})$ that identifies early Universe.

In the interval (1) the radius grows proportionally to time and the expansion speed is constant

$$R(t) = R(0^+) \frac{t}{T_0} = R(0^-) \frac{t}{T_0}; \qquad T_0 = 2.73 \times 10^{-102} \text{s}$$

$$\dot{R}(t) = \frac{R(0^-)}{T_0} = \text{const} = 0.458 \times 10^{48} \text{m/s}$$

$$R(t_L) = 1.475 \times 10^{14} \text{m}; \quad \dot{R}(t_L) = 0.458 \times 10^{48} \text{m/s}; \quad t_L = 1.096 \times 10^{-61} \text{s}$$

In the interval (2) the radius grows proportionally to $t^{1/2}$. This interval is decoupled in two sub-intervals (2*a*) and (2*b*) which are:

the sub-interval (2*a*): ($t_L < t < t_L+\Delta t_L$); $t_L=1.096\times10^{-61}$ s; $\Delta t_L=0.392\times10^{-61}$ s; $t_L+\Delta t_L = (1.096 + 0.392)\times10^{-61}$ s = 1.488×10^{-61} s

$$R(t) = R(t_L)\sqrt{\frac{t}{t_L}}\,; \qquad \dot{R}(t) = \frac{1}{2}\frac{R(t_L)}{t_L}\frac{1}{\sqrt{t/t_L}}$$

The sub-interval (2*b*)

$$R(t) = R(t_L + \Delta t_L)\sqrt{\frac{1}{1.488\times10^{61}\mathrm{s}}}\,; \qquad R(t_L + \Delta t_L) = 1.993\times10^{-14}\mathrm{m}$$

The above equations give $R(1.103) = 5.4\times10^{16}$. It is

$$R(t_L + \Delta t_L) = R(t_L) + \dot{R}(t_L)\Delta t_L\,.$$

REFERENCES

Arcidiacono, G., 1973, *Relatività e Cosmologia*, Veschi, Roma.

Bondi, H., 1961, *Cosmology*, Cambridge University Press, Cambridge, UK.

Di Caprio, U., and Spavieri, G., 1999, A new formula for the computation of the spectra of complex atoms, *Hadronic J.* (December).

Di Caprio, U., 2000, The effects of the proton's magnetic field upon quantization, *Hadronic J.* (December).

Di Caprio, U., 2001a, The dynamic structure of the proton, *Supplement to Hadronic J.* (July).

Di Caprio, U., 2001b, *Relativistic Stability*, Rep. S.A.-1.

Di Caprio, U., 2003, *Photon Dynamic Structure*, Rep. S.A.-5.

Dicke, R. H., 1964, Experimental relativity, in: *Relativity, Groups and Topology*, C. De Witt and B. De Witt, eds., Gordon and Breach, New York.

Dolgov, A. D., and Zeldovich, Ya. B., 1991, Cosmology and elementary particles, *Rev. of Modern Physics* **53**.

Macvittie, G. C., 1965, *General Relativity and Cosmology*, Chapman & Hall, London, UK.

Milne, E. A., 1935, *Relativity, Gravitation and World Structure*, Clarendon Press, Oxford, UK.

North, J. D., 1952, *The measure of the Universe*, Oxford University Press, Oxford, UK.

Robertson, H. P., and Noonan, T. W., 1968, *Relativity and Cosmology*, Saunders, Philadephia.

Ryan, M. P., and Shepley, L. C., 1975, *Homogeneous Relativistic Cosmologies*, Princeton University Press, Princeton, NJ.

Sciama, D. W., 1971, *Modern Cosmology*, Cambridge University Press, Cambridge, UK.

Schramm, D. N., and Steigman, G., 1988, Particle accelerators test cosmological theory, *Scientific American* (June).

Universe today; http://www.universetoday.com.

Weinberg, S., 1972, *Gravitation and Cosmology*, Wiley, New York.

MISTAKE MAKING MACHINES

Gianfranco Minati[1] and Giuseppe Vitiello[2]
[1]*Italian Systems Society, Via P. Rossi 42, 20161 Milano, Italy*
www.airs.it, Tel./Fax: +39-2-66202417, Email: gianfranco.minati@AIRS.it,
http://www.geocities.com/lminati/gminati/index.html
[2]*Dipartimento di Fisica "E.R.Caianiello", INFN and INFM,*
Università di Salerno, Via S. Allende, 84081 Baronissi (Salerno), Italy
Email: vitiello@sa.infn.it, http://www.sa.infn.it/giuseppe.vitiello/vitiello/

Abstract: Classic approaches consider errors and mistakes related to inadequate usage or functioning of physical or logical devices. They are usually considered problems to be fixed, like in engineering the ones related to reliability and availability. Mistake making processes or machines are assumed to be repaired. Another phase has been established when considering the role of the observer and the introduction of uncertainty principles. It is then possible to consider processes, at a certain level of description, as observer-related mistake making machines. We discuss the topic related to the possibility to design a mistake making device as a problem having correspondences with designing emergence. Emergence may be considered as a possible error appearing in Mistake Making Processes. We introduce the possibility to design an intrinsically (non observer-related) mistake making device, which has been proposed to be named Spartacus. This project is proposed with reference to the dissipative quantum model of brain. Another approach may be the one related to chaotic neural network designing, introduced in literature as Creativity Machine.

Key words: error; mistake; dissipation; emergence; model; observer; brain; quantum fields.

1. INTRODUCTION

There are different possible approaches to model systems by using different theories. One possibility is to adopt the brain and its functioning as a model for artificial systems. One of the striking features of the brain behavior is the autonomy which it exhibits in its functioning. The brain

appears to be capable of producing novelties, namely the brain behavior sometimes appears not rigidly determined by some functional or logical chain of steps in its functioning. Sometimes the brain acts as making mistakes with respect to some otherwise defined normal behavior. Remarkably, the brain may exhibit a behavior which is erratic not necessarily with respect to the expectations of the observer, but, rather, its being erratic is intrinsic to the brain dynamics. It is then tempting to explore the possibility to model systems in such a way to allow also for them this peculiar capability (such a privilege!) of making mistakes.

In order to clarify the general frame we propose for such a system modeling, we consider in this paper both the cases, the one of observer-related erratic devices and the one of intrinsically (non observer-related) erratic devices.

We introduce in Section 2 some descriptions of concepts helpful in the discussion of mistake making devices:

a) Observer,
b) Error and mistake making,
c) Working,
d) Reliability and availability
e) Machines and systems.

In Section 3, with reference to the observer-related erratic devices, we introduce the concept of trivial and non-trivial mistake making device. We observe that such a "machine" cannot be an algorithm. In Section 4, by resorting to some specific features of the dissipative quantum model of the brain, we put forward the question of whether it is possible to design intrinsically (non observer-related) erratic devices.

We observe that in referring to the erratic system we use the word device, but also the word machine. However, one must consider that, generally speaking, machines are by definition and construction not allowed making mistakes; in the case they make mistakes, they are considered to be wrong or out of order (see below). The discussion about the possibility to design a mistake making device may be related to the problem of designing emergence. Emergence may be indeed considered as a possible error appearing in Mistake Making Processes.

2. INTRODUCTORY DESCRIPTIONS

2.1 Observer

The subject has been crucial in many approaches and disciplines like second order cybernetics, cognitive science and the studies about consciousness (von Foerster et. al., 1974, von Foerster, 1981, 2003).

The reference to the observer usually relates to its cognitive processing and its role in the processes of emergence.

In the evaluation of "right and wrong", "true and false", the problem is not so much in the identification of a source of judgment or selection (i.e. relative to an observer), but rather it is in adopting the proper process of generating and using models by the observer, the Dynamic Usage of Models (Minati 2001; Minati and Pessa, in progress), the Logical Systemic Openness (Minati et al., 1996; 1998).

It is generally assumed that the observer is capable of awareness emerging from the global network of long range correlations in the universe.

It is interesting to note that in physics the role of the observer, due to his possible interference with the phenomenon under observation, is of crucial importance in measurement processes. The observer interference can be indeed the source of uncertainties and errors. This is not typical of quantum physics, although there it may occur with high probability. It also happens in classical physics, where indeed the theory of the errors was originally introduced (see below). The observer interference with the observed phenomenon has been sometimes considered to be the source of the Heisenberg uncertainty principle in Quantum Mechanics. However, it must be stressed that the uncertainty principle has its roots in the quantum fluctuations of the vacuum state, which is an intrinsic property of Nature, fully independent of the observer presence or actions.

2.2 Error and mistake making

In general, errors or mistakes are considered in relation to some observer action or evaluation. In the following we also consider a second possibility, namely errors as intrinsic to the system dynamics, occurring independently from any observer action.

As well known, the theory of the errors, introduced by Karl Friedrich Gauss (1777-1855), deals mostly with the first option, namely with the evaluation and the manipulation of the errors made in measuring a quantity by an observer.

In this context, reference is to errors in measuring, decision making, statistical analyses, and foreseeing. One of the purposes is to get

measurements having a level of error smaller of a determinate threshold, making the measurement acceptable in a specific context.

Here, an error is considered as a deviation from accuracy (Murphy 1961; Youden 1961) or correctness, taking place

- in applying rules (computing, logical fallacies),
- in modeling,
- in judging,
- in speaking and writing (linguistic errors), and so on.

An error is thus considered to be 'negative', as 'negation of correct', as lack of something missed or that cannot be completely reached, forcing acceptance of approximations.

The way to deal with an error is to fix it, reduce its negative influences, try to avoid it in the future and learning is often assumed to be a process making us able to do not repeat errors and to teach how to avoid them. In this sense, learning is a rather conservative concept, it is mostly consolidating the knowledge of how to perform correctly (i.e. without errors) well determined procedures or to preserve specific notions. Changes, extensions of the knowledge, inventions, novelties are excluded.

By changing level of description, evaluation rules or observer it is always possible to make all actions correct or incorrect.

Mistake making relates to real processes in a context. It must be related to a spatial and temporal context. It may be related to a level of description and/or to an observer. However, it may be also an intrinsic feature of the system behavior, independent of any description or action by an external (to the system) observer.

In the first case, mistake making takes place when the reaction or behavior of an entity, like a machine, a system, or a living being is different from the designed or expected behavior. This difference may be detected, evaluated, measured by an observer comparing the real behavior with the designed or expected one. In this case the observer is supposed to have such a cognitive model to be able to know which behavior is the right one. It is in this sense that the 'mistake' or error is not intrinsic to the system behavior. It is rather a system-observer relational feature.

When different observers have different expectances about the behavior, then some of them have a wrong cognitive model (i.e. wrong expectances) or they assume incorrect operational conditions. It is interesting to observe that the same observer has in this context the role of a (measuring) device or machine which may or may not make mistakes (according to the evaluation of other observers; the conventional or social (ethical) aspects of the 'mistakes' or 'deviances' here appear).

It is possible to look for errors made in previous processes having the purpose to improve a machine, a theory, a procedure (i.e. to avoid it in the

future). Usually the goal of a designer is to avoid mistakes by the designed machine.

Sometimes mistakes are useful to introduce or observe unexpected behaviors or results. Testing a newly designed machine has in general the meaning of detecting erratic behaviors to be avoided in an improved design of the tested machine.

When mistakes are not rejected, they constitute additions to or extensions of the observer knowledge. Examples are in production processes and discoveries made by chance. In these cases, the term by chance means indeed by mistake (with respect to what was expected). In some sense, the term dis-covery is equivalent to the term mistake (the discovery is always by chance, otherwise it is not a discovery).

In this connection, the problem with computers is that they do not make mistakes. Therefore they are useless for making discoveries.

There are scientific experiments which properly speaking are tests of validity of a theoretical model. Another category of experiments is the one of exploring experiments made in the hope to add knowledge of territories not covered by the tested theories. Sometimes it is very hard to distinguish between these two categories of experiments. Sometimes some experiments belong to both of them.

The error theory in studying ethics is the theory of those statements within a certain area which are reputed to be false. In general, any statement may be made false by changing the context in which it is considered. This approach has been introduced in dealing with moral statements, like by John Mackie (Mackie, 1977).

The correspondence between error and being false takes place when it is considered the result of a logic inference: an error may be considered equivalent to the attribution of the quality to be false when referred to a process of reasoning, to the result of a theory, to the comparison with experimental data.

There is a crucial difference between making mistake with respect to the expectance of the designer or of the observer at its level of description and making mistake as an intrinsic property of the dynamical behavior of the system. We refer to such a second possibility as to the unpredictability. Like the quantum fluctuations mentioned above, an unpredictable behavior is "gratuitous", extraneous to any logical or procedural "necessity", it cannot be expected or unexpected in any given context. It is not "derivable". Inside a given context, the unpredictable behavior is not a "negation", is not a "deviance" with respect to any possible behavior. It is a novelty. The unpredictable behavior gets to be known only after it is fully realized, only afterward. In this sense, when applied to this second case, the meaning of the

words 'mistake' or 'error' excludes any expectation content attached to them by any observer.

2.3 Working

Working always relates to an observer having ways of using in mind. A machine or a device is considered to be working when producing the desired effects.

Working is always a process of emergence, making the initial configuration of elements (the machine "off") to become something having new properties (the machine "on", having elements interacting thanks, for instance, to power). Anyway, working in this machine case is a process coming from the organized interactions between elements, like in an electronic card. This kind of emergence may be designed and controlled as it is different from intrinsic emergence (Cruchtfield, 1994).

Working may also relates to the effectiveness of an approach, a methodology, and a rule. Working relates to the fitness of something to a design, to expectances more than to objectivistic behaviors intended as absolute characteristics. Characteristics are ways of behaving and of using considered standard in a local context, by the observer. In this case usages correspond to functions, to linearly use them, and not how to use, to invent usages of functions. A machine may display errors in behaving and not inappropriateness in usage.

Ways of not working for a machine may be to do not make emergent, for some reasons, expected properties when powered or to do not properly make the job the observer had in mind.

An error takes place in the working of a machine when effects are not the ones desired by the observer. This may happen because of change in the working context, interferences, inappropriate supplied input or wrong usage.

A non-working machine may be considered different from an improperly working machine. It depends on how the observer considers a machine non-working. The property of being not working may refer to a constant, stable state of a machine, or may refer to the displaying of instabilities, irregularities in making available functions.

2.4 Reliability and availability

The German mathematician Eric Pieruschka collaborated with Von Braun about the problem of reliability for rockets. At those times it was assumed that the rocket would be as reliable as the least reliable part. Pieruschka showed that the rocket's reliability R would be equal to the

product of the reliability of its components. This result formed the basis for what later became know as Lusser's law: $R = R_1 \times R_2 \times \ldots \times R_n$

In engineering reliability is the probability that a device, process, or system will perform the expected job without failure for a given time when operated correctly in a specified environment: $R = e^{-t/\check{T}}$, where $\check{T}$ is the mean time between failures, t is the time over which the device is expected to operate. The failure rate is $\lambda = 1/\check{T}$ and $N = \lambda\, t$ is the number of failures during the activity.

The index of availability A is given by the ratio between the time in which a device, a system or a process has been effectively available (i.e. operational by respecting some pre-defined standards) to the user and the total time of observation (Villemeur, 1992):

$$A = Uptime \,/\, (Uptime + Downtime)$$

In engineering the parameters usually considered are:

- MTBF - *Mean Time Between Failures*,
- MTTR - *Mean Time To Repair*.

2.5 Machines and systems

In literature, in short, a system is intended to be emergent from interacting elements components, where any change on a component affects all other components, and making emergent a new entity (i.e. the system) having characteristics different from the ones of the components (von Bertalanffy, 1968).

The subject relates to the Systems Theory and the theory of emergence (Minati and Pessa, in progress).

The concept of machine has been introduced from different approaches in engineering, physics, cybernetics, chemistry, mathematics, information science and so on.

In physics a machine is intended as a device for doing work.

In the course of the 18th century the view that an organism is a mechanical machine, gradually turned into the idea that it is a chemical machine, a new engine exploiting the chemical reactions of combustion to produce mechanical movements. The steam engine brought together the two sciences: mechanists and vitalists realized that a chemical machine is not a contradiction in terms.

In systemic terms a machine may be intended as a state determined system, that is any system in which the specification of a state determines the subsequent state, having a behavior fully determined by knowing its initial state and the transformations.

A Markovian machine is a probabilistic machine, i.e. any system in which the probability of any given state determines the probability of the subsequent state. The Turing machine, introduced by Alan Turing to give a mathematically definition of algorithm, has opened a new era in considering computation and mental processing (Boolos, 1980; Putnam, 1975).

We just mention that a Turing machine is an abstract computing device. consisting of a read/write head able to bi-directionally scan a one-dimensional, infinite tape divided into boxes, each of which is inscribed with a 0 or 1. Computation takes place with the machine, in a given state, scanning a box. The machine behavior is completely determined by: (a) the state, or functional states, the machine is in, (b) the number, 0 or 1, on the box it is scanning, and (c) a table of instructions specifying, for each, finite number of state and possible input (0 or 1), what the machine should write (erase in some versions), which direction it should move in, and which state it should take. The concept of right or wrong particularly applies to machines because of their state determined behavior. Mistake making relates to steps of a machine working.

3. OBSERVER-RELATED MISTAKE MAKING MACHINES

A mistake making machine cannot be a Turing machine, namely an algorithm generating mistakes. Indeed, it is not possible to design an algorithm doing nothing but the expected result, even if such a result is defined to be wrong with respect to certain criteria.

Even by introducing probabilistic behavior the machine may be intended to be right as far as its behavior fulfills the one expected by the observer.

With reference to the definition introduced at the point b) in Section 2, a machine is assumed mistake making if it contradicts the expectance of the designer, of the observer in its context, at its level of description. The crucial point is the role of the observer, not the computational aspects.

Let us consider the activity of a machine examined by an observer. It is not possible to exclude the existence of a level of description in which the machine is not mistake making as stated by the Theorem of Chaitin (Chaitin, 1974a; 1974b; 1992).

A trivial mistake making machine is a machine doing other than what expected by a single observer.

A non-trivial mistake making machine is a machine doing other than what expected by any observer. A machine is always making mistake when its behavior contradicts expectances of any observer of the universe (this definition is time independent).

If a machine has a finite number n of possible behaviors and there are n observers each of them expecting one of the n possible behaviors, then the machine cannot be mistake making for all observer.

It is possible to have an always mistake mistaking machine, that is mistake making for any observer, if the number of observers is less than the number of the possible behaviors assumable by the machine.

The problem of the existence of a never mistake making machine occurs when its behavior fulfills expectances of any observer of the universe (this definition is time independent). Because the observers may have different, contradictory expectances, then this situation cannot occur.

Necessary condition for the existence of a non-trivial mistake making machine is that the universe of possible behaviors for the machine is greater than the one of the possible observers. In case of infinite number of behaviors and infinite number of observers the cardinality of the set of the behaviors must be greater than the cardinality of the set of the observers.

An always mistake making machine must have the possibility to assume a number of states having cardinality greater than the one of the set of possible observers. Sufficient condition for the existence of a non-trivial always mistake making machine is the occurring of the necessary condition and, simultaneously, the access to information about the universe of all the observers in order to be able to assume a behavior considered wrong by any and all possible observers.

4. NON OBSERVER-RELATED MISTAKE MAKING MACHINES

In the dissipative quantum model of brain (Vitiello, 1995; 2001) it is possible to describe the time evolution of a given memory (long term and short term memories) as a trajectory in the space of the memories. This is the collection (the space) of the Hilbert spaces each one associated to a different code (order parameter) representing a different memory. It has been recently shown (Pessa and Vitiello, 2004a; 2004b) that these trajectories are classical chaotic ones, which means that each trajectory is a bounded trajectory in the parameter space and does not intersect itself (i.e. it correspond to a non periodic deterministic evolution); trajectories corresponding to different initial conditions (different values of the code at initial time and thus different memories) do not intersect each other; trajectories specified by different initial conditions are diverging trajectories (even for slightly different initial conditions). These chaos features of the trajectories hold in the infinite volume limit.

It has been also shown (Pessa and Vitiello, 2004a; 2004b) that the dissipative character of the brain dynamics (arising from the unavoidable fact that the brain is a system permanently open onto the external world) implies quantum noise in the fluctuating random force in the brain-environment coupling. Such a noisy fluctuating background interferes with the strict determination of the trajectories initial conditions in a fully unpredictable way. It has been conjectured (Vitiello, 2001; 2004) that the "secret flavor of subjectivity" may find its roots in this "doubtful predictability" in this "precious unfaithfulness" which depicts the subject dynamical identity, even his consciousness mechanisms.

In the mentioned chaotic regime, the fuzziness of the initial conditions due to quantum fluctuations means that the same stimulus for the same subject is never associated to the same trajectory in the memory space. Which indeed describes the quite familiar experience of the uniqueness of the now. The dense sequence of 'nows', all of them unique and therefore 'distinct' among themselves, "constitute the multi-time dimensions of the self, its own time space", that "spring of time-lines through which the self can move "freely"" (Vitiello, 2004).

Due to boundary effects or to inhomogeneities the chaoticity may be smoothed out (Pessa and Vitiello, 2004a; 2004b), although not completely eliminated, so that intersections of two or more trajectories may occur, and starting from such a intersection point any of the trajectories departing from there may be then followed. This may result in the confusion of memories or also in the association of memories. These intersection points are switching points and which one will be the trajectory which will happen to be followed is fully unpredictable since the 'initial conditions' associated to that crossing point are not sufficient to uniquely determine it among those from there departing (loss of determinism). Confusion of memories, association of memories and weakening or loss of determinism are thus associated to unpredictable paths in the memory space: erratic behavior thus appears in the brain activity.

In addition to that, it should be considered the corruption of the code specifying a given memory. This corruption occurs in the course of the time evolution as an effect of the dissipative dynamics. In the absence of an action by the subject aimed to refresh the memory (brushing up the memory), it manifests as changes, deformations, blurring in the originally recorded memory. Such a memory corruption process may also contribute to the uncertainty (fuzziness) characterizing the initial conditions associated to the trajectory crossing point. This point then may be the origin of a completely new, unforeseen trajectory, not belonging to the set of the trajectories converging at that point. This mis-taken trajectory appears as the

brain error, intrinsic in its quasi-chaotic dissipative dynamics. It characterizes the brain autonomy (Vitiello, 2004).

The uncertainty associated to the initial conditions out of which the erratic trajectory arises has been associated to the doubts characterizing the continual self-questioning of consciousness modes (Vitiello, 2004). On the other hand, the "internal freedom" from the strict deterministic evolution of the memory paths has been associated to the possibility of the brain "active response" to the world stimuli and to the "aesthetical dimension" (Vitiello, 2004). In this sense, the brain error or mistake might be associated to what commonly is referred to as responsibility (the ethical dimension) and as creativity, respectively. Perhaps, the note the composer is going to write after another note, the color or the word the painter or the poet are respectively going to use in their painting or poetry is not known a priori also by them, it is a surprise also to them. They get awareness of that note, color or word only afterward. Perhaps this is why the literary tradition wants that Omer, the poet, was blind.

The challenge is now: is it possible to design an artificial device capable of making such a kind of mistakes? In view of the positive content of these mistakes and of the fact that they are privileges of the natural brain, the name which has been proposed for such a distinguished (hypothetical) mistake making machine is Spartacus (Vitiello, 2004).

Work aimed to pursue such a program is in progress.

5. CONCLUSIONS

We have considered the possibility to model systems in such a way to allow for them the capability of making mistakes. We have considered both the cases, the one of observer-related erratic devices and the one of intrinsically (non observer-related) erratic devices.

With reference to the observer-related erratic devices, we have introduced the concept of trivial and non-trivial mistake making device. We have observed that such a "machine" cannot be an algorithm.

By resorting to some specific features of the dissipative quantum model of the brain, we have put forward the question whether it is possible to design an intrinsically (non observer-related) erratic device, which has been proposed to be named Spartacus. In this connection, we remark that since 1996 S. L. Thaler is pursuing an interesting and promising research program along the line of chaotic neural network designing (Thaler, 1996a; 1996b). It is interesting to investigate about possible relations between Thaler's Creativity Machine and the features exhibited by the dissipative quantum model of brain in connection with the Spartacus designing program.

REFERENCES

Boolos, G., and Jeffrey, R., 1980, *Computability and Logic*, 2nd ed., Cambridge University Press, Cambridge.

Chaitin, G. J., 1974a, Information-theoretic computational complexity, *IEEE Transactions on Information Theory* **20**:10-15.

Chaitin, G. J., 1974b, Information-theoretic limitations of formal systems, *Journal of the ACM*, **21**:403-424.

Chaitin, G. J., 1992, Information-theoretic incompleteness, *Applied Mathematics and Computation* **52**:83-101.

Cruchtfield, J. P., 1994, The calculi of emergence: computation, dynamics and induction, *Physica D* **75**:11-54.

Mackie, J. L., 1977, *Ethics: Inventing Right and Wrong*, Penguin.

Minati, G., 2001, *Esseri Collettivi*, Apogeo, Milano, Italy.

Minati, G., and Pessa, E., (in progress), *Collective Beings*, Kluwer Academic/Plenum Publishers, New York.

Minati, G., Penna, M. P., and Pessa, E., 1996, Towards a general theory of logically open systems, in: *Proceedings of the 3rd European Systems Conference*, Edizioni Kappa, Rome.

Minati, G., Penna, M. P., and Pessa, E., 1998, Thermodynamic and logical openness in general systems, *Systems Research and Behavioral Science* **15**(3):131-145.

Murphy, R. B., 1961, On the meaning of precision and accuracy, *Mater. Res. Stand.* **1**:264-267.

Pessa, E., and Vitiello, G., 2004a, Quantum noise, entanglement and chaos in the quantum field theory of Mind/Brain states, *Mind and Matter* **1**:59-79.

Pessa, E., and Vitiello, G., 2004b, Quantum noise induced entanglement and chaos in the dissipative quantum model of brain, *Int. J. Mod. Phys. B* (in print).

Putnam, H., 1975, The nature of mental states, in: *Mind, Language and Reality: Philosophical Papers II*, H. Putnam, ed., Cambridge University Press, Cambridge.

Thaler, S. L., 1996a, Neural nets that create and discover, *PC AI*, May/June, pp. 16-21.

Thaler, S. L., 1996b, Is neuronal chaos the source of stream of consciousness?, *World Congress on Neural Networks, (WCNN'96)*, Lawrence Erlbaum, Mawahm, NJ.

Villemeur, A., 1992, *Reliability, Availability, Maintainability and Safety Assessment: Volume 1 - Methods and Techniques*, Wiley and Sons.

Vitiello, G., 1995, Dissipation and memory capacity in the quantum brain model, *Int. J. Mod. Phys. B* **9**:973-989.

Vitiello, G., 2001, *My Double Unveiled*, John Benjamins, Amsterdam, The Netherlands.

Vitiello, G., 2004, The dissipative brain, in: *Brain and Being*, G. Gordon, K. H. Pribram and G. Vitiello, eds., John Benjamins, Amsterdam.

von Bertalanffy, L., 1968, *General Systems Theory*, George Braziller, New York.

von Foerster, et. al., eds., 1974, Cybernetics of cybernetics. Biological computer laboratory, in: *Cybernetics of Cybernetics*, K. Krippendorf, ed., 1979, Gordon and Breach, New York.

von Foerster, H., 1981, Observing Systems, in: *Selected Papers of Heinz von Foerster*, Intersystems Publications, Seaside, CA.

von Foerster, H., 2003, *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer-Verlag, New York.

Youden, W. J., 1961, How to evaluate accuracy, *Mater. Res. Stand* **1**:268-271.

EXPLICIT VELOCITY FOR MODELLING SURFACE COMPLEX FLOWS WITH CELLULAR AUTOMATA AND APPLICATIONS

M. V. Avolio[1,3], G. M. Crisci[2,3], D. D'Ambrosio[1,3], S. Di Gregorio[1,3], G. Iovine[4,3], V. Lupiano[4], R. Rongo[2,3], W. Spataro[1,3] and G. A. Trunfio[3]
[1]*Dept. of Mathematics,* [2]*Dept. of Earth Sciences,* [3]*Center of High-Performance Computing, University of Calabria, Arcavacata, 87036 Rende, CS, Italy;*
[4]*CNR-IRPI, Via Cavour, 87030 Rende, CS, Italy*

Abstract: Fluid-dynamics is an important field of Cellular Automata applications, that give rise also to specialised models as lattice gas automata and lattice Boltzmann models. These models come up against difficulties for applications to large scale (kilometres) phenomena. Our research group *Empedocles* faces up to macroscopic phenomena concerning surface flows, developing an alternative strategy, that was significantly improved, introducing explicit velocities in own Cellular Automata models. This paper illustrates the methodology with applications to lava flows, debris flows and pyroclastic flows.

Key words: complex systems; cellular automata; computational fluid-dynamics; lava flow; debris flow; pyroclastic flows.

1. INTRODUCTION

Cellular Automata (*CA*) are a paradigm of parallel computing; they are good candidates for modelling and simulating complex dynamical systems, whose evolution depends exclusively on the local interactions of their constituent parts (Di Gregorio and Serra, 1999) and represent sometime an alternative approach for that type of phenomena.

A *CA* involves a regular division of the space in cells, each one embedding a finite automaton (*fa*). The *fa* state changes according to a transition function that depends on the states of neighbouring cells and of the

cell itself; the transition function and the neighbourhood are invariant in time and space. At the time $t = 0$, cells are in states, describing initial conditions, and the *CA* evolves changing the state of all the cells simultaneously at discrete times, according to the transition function.

An equivalent description could be given in terms of a regular lattice, whose sites correspond to the cell centres and neighbourhood conditions are given by vectors joining a site with its neighbours.

Many complex fluid-dynamical phenomena were modelled by *CA*: lattice gas automata models (Frisch et al.,1990) simulate fluid dynamical properties in terms of motion and collision of "particles" on a grid, leading to the Navier-Stokes equations at the continuum limit for a defined class of conditions, while lattice Boltzmann models (Chopard and Luthi, 1999) can simulate a larger range of fluid-dynamical cases, considering the density of fluid particles by more numerous states.

Complex macroscopic fluid-dynamical phenomena like surface flows seem difficult to be modelled in these *CA* frames, when they take place on a large space scale and need practically a macroscopic level of description that involves the management of a large amount of data, e.g., the morphological data.

An empirical method was developed in order to overcome these limits (Di Gregorio and Serra, 1999):

- a nearly unlimited (but finite) number of states is permitted; the state is composed of substates, each substate describes a feature of the space portion related to the own cell (e.g. the substate "temperature");
- the transition function is split in several parts, applied sequentially, each one corresponds to an "elementary" process of the macroscopic phenomenon (e.g. lava solidification);
- substates of type "outflow" are used in order to account for quantities moving from a cell toward another one in the neighbouring.

This approach didn't overcome the problem of the constant "velocity": a fluid amount moves from a cell to another one in a *CA* step (which is considered a constant time), it implies a constant "velocity" in the *CA* context of discrete space/time. Nevertheless, velocities can be deduced as emergent properties by analyzing the global behavior of the system in time and space. In such models, the flow velocity can be deduced by averaging on the space (i.e. considering clusters of cells) or by averaging on the time (e.g. considering the average velocity of the advancing flow front in a sequence of *CA* steps).

Constant velocity could be a limit for modelling finely macroscopic phenomena, because it is difficult to introduce physical considerations in the modelling at local level.

A solution is here proposed: fluids, e.g., the lava, are characterised also by substates specifying the mass centre position and velocity inside the cell.

The second section considers the improved *CA* methodological approach for modelling macroscopic surface flows; the third section presents the applications to lava, debris and pyroclastic flows; some conclusions are reported at the end.

2. AN EXTENDED DEFINITION OF *CA*

CA were conceived in the 1950's by J. von Neumann (1966) in order to investigate self-reproduction were one of the first Parallel Computing models. The most widely used CA were developed adopting a microscopic approach with a small number of states (from two states to no more than hundred states).

The complexity of macroscopic natural phenomena demands an extension of the original computational paradigm for many cases: phenomena with a macroscopic approach could involve as far as billion of states with a proportionally complex transition function. Furthermore additional specifications need for permitting a correspondence between the system with its evolution in the physical space/time, on the one hand, and the model with the simulations in the cellular space/time, on the other hand. The following considerations about *CA* and macroscopic systems introduce an extended definition of *CA* for surface flows, where the explicit velocity notion is included.

2.1 Global parameters

Primarily, the dimension of the cell (e.g. specified by the cell side p_s) and the time correspondence to a *CA* step p_t must be fixed. These are defined as "global parameters", as their values are equal for all the cellular space. They constitute the set P together with other global parameters, which are commonly necessary for simulation purposes.

2.2 Substates

The state of the cell must account for all the characteristics, relative to the space portion corresponding to the cell, which are assumed to be relevant to the evolution of the system. Each characteristic corresponds to a substate; permitted values for a substate must form a finite set.

When a characteristic (e.g. a physical quantity) is expressed as a continuous variable, a finite, but sufficient, number of significant digits are utilised, so that the set of permitted values is large but finite.

The substate value is considered always constant inside the cell (e.g. the substate altitude). Then the cell size must be chosen small enough so that the approximation to consider a single value for all the cell extension may be adequate to the features of the phenomenon.

The set S of the possible values of state of a cell is given by the Cartesian product of the sets $S_1, S_2, \ldots S_n$ of the values of substates: $S = S_1 \times S_2 \times \ldots \times S_n$; the set Q is also defined: $Q = \{S_1, S_2, \ldots, S_n\}$.

The cellular space is two dimensional because quantities concerning the third dimension (the height) may be included among the substates of the cell in a phenomenon concerning the earth surface

2.3 Elementary processes

The state transition function τ must account for all the processes (physical, chemical, etc.), which are assumed to be relevant to the system evolution, which is specified in terms of changes in the states values of the *CA* space. As well as the state of the cell can be decomposed in substates, the transition function τ may be split into "elementary" processes, defined by the functions $\sigma_1, \sigma_2, \ldots \sigma_p$ with p being the number of the elementary processes.

The elementary processes are applied sequentially according a defined order. Different elementary processes may involve different neighbourhoods; the *CA* neighbourhood is given by the union of all the neighbourhoods associated to each processes.

In the empirical approach of Di Gregorio and Serra (1999), an elementary process is individuated by: σ: $Q_a^m \rightarrow Q_b$, where Q_a and Q_b are Cartesian products of the elements of subsets of Q, m is the number of cells of the neighbourhood, involved in the elementary process; Q_a individuates the substates in the neighbourhood that effect the substate value change and Q_b individuate the cell substates that change their value. Furthermore the movement of a certain amount of fluid from a cell toward another cell is described introducing substates of type "outflows", that specify the involved fluid quantities to be moved in the neighbourhood.

2.4 External influences

Sometimes, a kind of input from the "external world" to the cells of the *CA* must be considered; it accounts for describing an external influence which cannot be described in terms of local rules (e.g. the lava alimentation at the vents) or for some kind of probabilistic approach to the phenomenon.

Of course special and/or additional functions must be given for that type of cells.

2.5 Dimensions of the cell size and clock

The choice of the value of the parameters cell size and clock is dependent on the elementary processes. They could be inhomogeneous in space and/or time: the opportune dimension of a cell can vary for different elementary processes; furthermore very fast local interactions need a step corresponding to short times on the same cell size; the appropriate neighbourhoods for different local interactions could be different. An obvious solution to these problems is the following: the smallest dimension of a cell must be chosen among the permitted dimensions of all the local interactions. Then it is possible to define for each local interaction an appropriate range of time values in correspondence of a *CA* step; the shortest time necessary to the local interactions must correspond to a step of the *CA*. It is possible, when the cell dimension and the *CA* step are fixed, to assign an appropriate neighbourhood to each local interaction; the union of the neighbourhoods of all the local interactions must be adopted as the *CA* neighbourhood.

2.6 Extended *CA* formal definition

Considering these premises, the following formal definition of two dimensional square or hexagonal *CA* is given:

$<\boldsymbol{R}, \boldsymbol{G}, \boldsymbol{S}, \boldsymbol{X}, \boldsymbol{P}, \tau, \gamma>$

- $\boldsymbol{R} = \{(x, y) | \; x, y \in \mathfrak{I}, -l_x, \leq x \leq l_x, -l_y \leq y \leq l_y\}$ is the set of points with integer co-ordinates in the finite region, where the phenomenon evolves. Each point identifies a square or hexagonal cell.
- $\boldsymbol{G} = \{G_1 \cup G_2 \cup \ldots \cup G_n\}$ is the set of cells, which undergo to the influences of the "external world"; n external influences are here considered, each one defines a subregion G_i $1 \leq i \leq n$ of the cellular space, where the influence is active. Note that $G \subseteq R$.
- $\boldsymbol{X}$, the neighbourhood index, is a finite set of two-dimensional vectors, which defines the set $N(X,i)$ of neighbours of cell $i = <i_x, i_y>$ as follows: let $X = \{\xi_0, \xi_1, \ldots, \xi_{m-1}\}$ with $m = \#X$, then $N(X,i) = \{i+\xi_0, i+\xi_1, \ldots i+\xi_{m-1}\}$; ξ_0 is always the null vector. Note that $\boldsymbol{X} = \{X_1 \cup X_2 \cup \ldots \cup X_p\}$, where X_j $1 \leq j \leq p$ is correlated to the j-th elementary process.
- $\boldsymbol{S} = S_1 \times S_2 \times \ldots \times S_s$ is the set of the state values; it is specified by the Cartesian product of the finite sets of the values of the s substates $S_1, S_2, \ldots S_s$.

- $\boldsymbol{P}$ is the finite set of global parameters, which effect the transition function.
- $\tau : S^m \rightarrow S$ is the deterministic state transition. It is specified by the ordered set of elementary processes in order of application: $<\sigma_1, \sigma_2, \ldots \sigma_p>$.
- $\gamma : N \times G \rightarrow S$ expresses the external influences to cells of G in the cellular space; it determines the variation of the state S for the cells in G. N, the set of natural numbers, is here referred to the steps of CA. γ is specified by the sequential applications of the n functions $\gamma_1 : N \times G_1 \rightarrow Q_1$, $\gamma_2 : N \times G_2 \rightarrow Q_1$, $\gamma_n : N \times G_n \rightarrow Q_n$ where Q_1, Q_2, ... Q_n are Cartesian products of the elements of subsets of Q ($Q = \{S_1, S_2, \ldots S_n\}$).

2.7 Explicit Velocity

The quantity of fluid inside the cell is described in general by the following substates: fluid thickness, mass centre co-ordinates and velocity components. The x and y co-ordinates are related to the cell centre; initial conditions involve value 0 for x, y co-ordinates and velocity components.

Outflows from the cell to its adjacent cells are computed in three steps:

- The first step involves the computation of the size of the minimising flows, i.e. the flows that minimise the differences of certain quantities in the neighbourhood (Di Gregorio and Serra, 1999) or between the cell and its adjacent cell (Crisci et al., 2004).
- A velocity is computed for each minimising flow according to the following equations (deduced in sequence and similar to the Stokes equations) where F is the force, m is the mass of the fluid inside the cell, v is its velocity, t is the time, v_0 is the initial velocity, θ is the angle of the slope between the two cells, α is the friction parameter.

$$F = mg \sin\theta - \alpha m v$$
$$dv/dt = g \sin\theta - \alpha v$$
$$v = (v_0 - g \sin\theta / \alpha)\, e^{-\alpha t} + (g \sin\theta / \alpha)$$

  These equations describe a motion, which is depending on the gravity force and is opposed by friction forces. An asymptotic velocity limit is considered because the effect of the friction forces increases as the velocity increases.
- The mass centre of the outflow is shifted according to its velocity and the time corresponding to a step of the CA. If the shift is sufficient, a part (or all the outflow) will be transferred to the adjacent cell.
- New co-ordinates of the mass centre and components of its velocity are obtained considering outflows and inflows of the cell.

3. APPLICATIONS

The methodological approach of the previous section for modelling surface flows by *CA* was applied to two different phenomena, which were modelled by our research group *Empedocles* without explicit velocity: lava flows (SCIARA (Barca et al., 1994; Crisci et al., 2004a)), debris flows (SCIDDICA (Di Gregorio et al., 1999; Avolio et al., 2000, D'Ambrosio et al., 2002, 2003a, 2003b; Iovine et al. 2003)). Such models were partially modified introducing the explicit velocity and are illustrated in the following subsections together with the model for pyroclastic flows (PYR (Crisci et al., 2004b)), that was developed directly in terms of explicit velocity.

3.1 The model SCIARA-γ2

SCIARA, *Simulation by* **C***ellular* **I***nteractive* **A***utomata of the* **R***heology of* **A***etnean lava flows* (sciara means the solidified lava path in Sicilian) is a family of models for simulating lava flows (Barca et al., 1994; Crisci et al., 2004a); the version γ2 of SCIARA (submitted to Computers and Geosciences) adopts the explicit velocity.

SCIARA-γ2 = $\langle R, V, X, S, P, \tau, \gamma \rangle$:

- $R = \{(x, y) | x, y \in \Im, -l_x, \leq x \leq l_x, -l_y \leq y \leq l_y\}$ identifies the set of regular hexagons covering the finite region, where the phenomenon evolves.
- V is the set of cells, corresponding to the vents.
- $X = \{(0,0), (0,1), (0,-1), (1,0), (-1,0), (1,1), (-1,-1)\}$.
- S is the set of *fa* states, i.e. the Cartesian product of the values sets of the following substates:
  S_a is the cell altitude;
  S_{th} is the thickness of lava inside the cell;
  S_x, S_y are the co-ordinates of mass centre of lava inside the cell;
  S_{Vx}, S_{Vy} are the velocity components of mass centre of lava inside the cell;
  S_T is the lava temperature.
- P is the set of the global parameters:
  p_a is the apothem of the cell;
  p_{step} is the time correspondence of a step;
  p_{TV} is the lava temperature at the vent;
  p_{TS} is the lava solidification temperature;
  p_{adhV} is the adherence (i.e. unmovable thickness of lava at the emission temperature at the vents;
  p_{adhS} is the adherence value at the solidification temperature;
  p_{cool} is the cooling parameter;

p_{vl} is the "limit of velocity" for lava flows.

- $\tau : S^7 \rightarrow S$ is the deterministic transition function, composed by the following "elementary" processes: determination of the lava flows by application of minimisation algorithm (Di Gregorio and Serra, 1999); determination of the lava flows shift by application of velocity formulae; mixing of inflows and remaining lava inside the cell (determines new thickness and temperature); lava cooling by radiation effect and solidification;
- $\gamma : Q_{th} \times N \rightarrow Q_{th} \times Q_T$ specifies the emitted lava from the source (vents) cells at the CA step $t \in N$.

3.2 Simulations with SCIARA-γ2

A first application of SCIARA-γ2 concerns the crisis in the autumn of 2002 at Mount Etna (Sicily). The eruption started October 24 on the NE flank of the volcano, with lava generated by a fracture between 2500 m a.s.l and 2350 m a.s.l., pointing towards the town of Linguaglossa. After 8 days, the flow rate diminished drastically, stopping the lava front towards the inhabited areas.

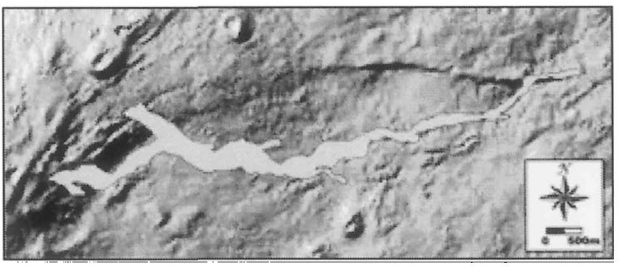


Figure 1. The maximum extension of the 2002 Etnean lava flow of NE flank, menacing the town of Linguaglossa.

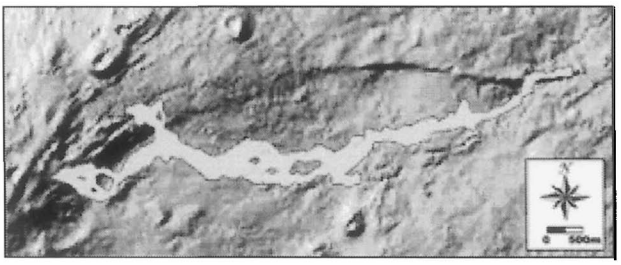


Figure 2. The simulation of 2002 Etnean lava flow of NE flank toward the town of Linguaglossa.

The Fig.1 shows the real lava flow at the maximum extension, Fig. 2 shows the corresponding simulation. Comparison between real and simulated event is satisfying, if we compare the involved area and lava thickness.

3.3 The model SCIDDICA S5a

SCIDDICA, *Simulation through* **C***omputational* **I***nnovative methods for the* **D***etection of* **D***ebris flow path using* **I***nteractive* **C***ellular* **A***utomata* (sciddica means "it slides" in Sicilian) is a family of models (Di Gregorio et al., 1999; Avolio et al., 2000, D'Ambrosio et al., 2002, 2003a, 2003b; Iovine et al. 2003) for simulating landslides of type debris/mud flows; the version S5a of SCIDDICA (presented at EGU 2004, 25-30 April 2004, Nice, France) adopts the explicit velocity.

SCIDDICA S5a = <*R, X, S, P,* τ> :

- $\boldsymbol{R} = \{(x, y) | x, y \in \mathfrak{I}, -l_x, \leq x \leq l_x, -l_y \leq y \leq l_y\}$ identifies the set of regular hexagons covering the finite region, where the phenomenon evolves.
- $\boldsymbol{X} = \{(0,0), (0,1), (0,-1), (1,0), (-1,0), (1,1), (-1,-1)\}$
- $\boldsymbol{S}$ is the set of *fa* states, i.e. the Cartesian product of the values sets of the following substates:
  S_a is the cell altitude
  S_{th} is the thickness of debris inside the cell
  S_e is the energy of landslide debris
  S_x, S_y are the co-ordinates of mass centre of debris inside the cell;
  S_{Vx}, S_{Vy} are the velocity components of mass centre of debris inside the cell;
  S_d is the depth of erodable soil cover.
- $\boldsymbol{P}$ is the set of the global parameters:
  p_a is the apothem of the cell;
  p_{step} is the time correspondence of a step;
  p_{fa} is the height threshold (related to friction angle) for debris flows;
  p_{rl} is the run-up loss (at each step), due to frictional effects;
  p_{adh} is the adhesion (i.e. the unmovable amount of landslide debris);
  p_{mt} is the activation threshold for mobilisation of the soil cover;
  p_{pe} is the parameter of progressive erosion of the soil cover;
- $\tau : S^7 \rightarrow S$ is the deterministic transition function, composed by the following "elementary" processes:
  debris flows determination by application of minimisation algorithm (Di Gregorio and Serra, 1999) with run up (D'Ambrosio et al., 2002, 2003a, 2003b);

mixing of inflows and remaining debris inside the cell (determines new thickness and energy);
mobilisation triggering and effect;
energy loss by friction.

3.4 Simulations with SCIDDICA S5a

SCIDDICA S5a was applied to the Chiappe di Sarno (Italy) debris flows, triggered on 5-6 May 1998 by heavy rains.

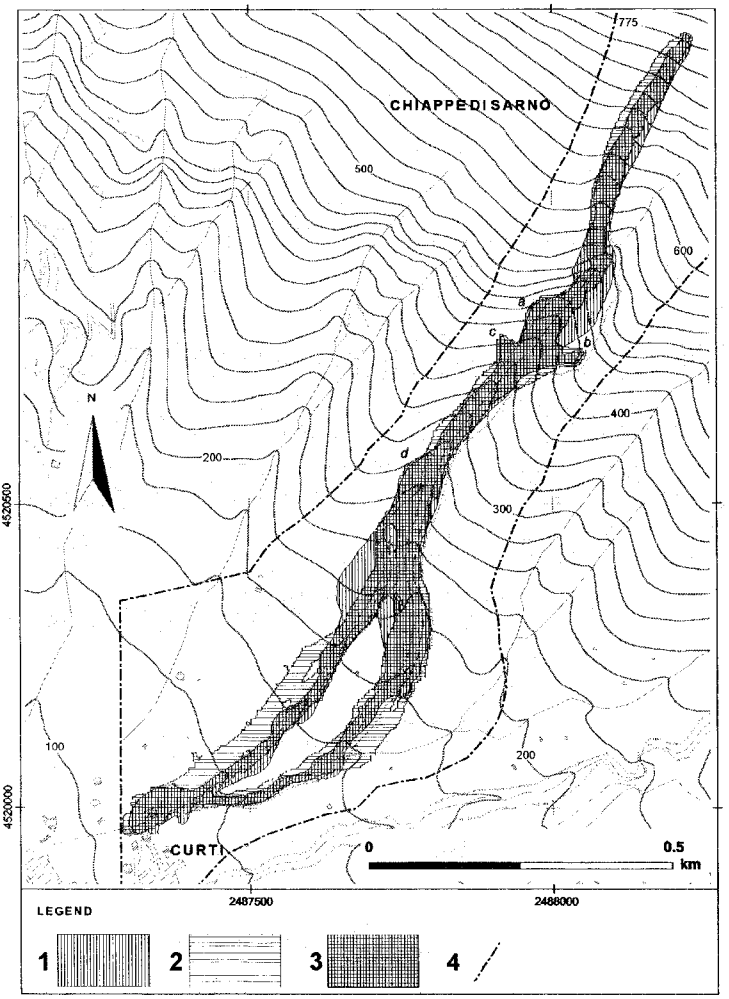


Figure 3. The Chiappe di Sarno landslide: quantitative comparison between real and simulated cases. Key: area affected by (1) real landslide, (2) simulated landslide, (3) both cases; (4) border of the area considered in the GIS analysis.

Debris slides were originated in the soil mantle, and transformed into rapid/extremely rapid debris flows, deeply eroding the soil cover along their path (avalanche effect). Landslides caused serious damage and numerous victims.

Fig. 3 shows the superposition of a real and simulated event. Comparison between real and simulated event is satisfying, if we compare the involved area and debris thickness.

3.5 The model PYR

Pyroclastic flows are fluidized masses of rock fragments and gases that move rapidly in response to gravity. They can form following the collapse of an eruption column. The flow contains water and gas from the eruption, water vapour and air can be enclosed as it moves down slope; pyroclastic flows can move at rates of 200 m/s.

Pyroclastic flows represent a greatest volcanic hazard; they can incinerate, burn, and asphyxiate people. More people have died due to these hazards than any other volcanic hazard.

The model PYR is a first attempt of *CA* model for simulating pyroclastic flows, the explicit velocity method was conceived the first time for this type of application.

PYR = $\langle \boldsymbol{R}, \boldsymbol{G}, \boldsymbol{X}, \boldsymbol{S}, \boldsymbol{P}, \tau, \gamma \rangle$:

- $\boldsymbol{R} = \{(x, y) | x, y \in \mathfrak{I}, -l_x, \leq x \leq l_x, -l_y \leq y \leq l_y\}$ identifies the set of squares covering the finite region, where the phenomenon evolves.
- $\boldsymbol{G}$ is the set of cells, corresponding to the area, where the volcanic column begins to collapse and to generate the pyroclastic flows.
- $\boldsymbol{X} = \{(0,0), (0,1), (0,-1), (1,0), (-1,0)\}$.
- $\boldsymbol{S}$ is the set of *fa* states, i.e. the Cartesian product of the values sets of the following substates:
  S_a is the cell altitude;
  S_{th} is the thickness of solid particles deposit;
  S_x, S_y, S_z are the co-ordinates of the mass centre of the pyroclastic column inside the cell;
  S_E is the elevation of the pyroclastic column;
  S_V is velocity of the pyroclastic column;
  S_{RT} is the residual time of the *CA* step during which a flow may leave the cell.
- $\boldsymbol{P}$ is the set of the global parameters:
  p_s is the side of the cell;
  p_{step} is the time correspondence of a step;

p_{SP}, p_G is the solid particles and gas content of pyroclastic column (in percent: $p_{SP} + p_G = 100$);
p_{erl} is the degassing - particles deposition relaxation rate (elevation loss rate);
p_α is a parameter ruling the friction effect;

- $\tau : S^7 \rightarrow S$ is the deterministic transition function, composed by the following "elementary" processes:
  degassing and particles deposition;
  internal shift of the pyroclastic column;
  outflowa determination;
  outflows and inflows composition.
- $\gamma : S_E \times N \rightarrow S_E$ specifies the feeding of pyroclastic matter, $t \in N$, the natural numbers set.

3.6 Simulations with PYR

PYR was applied to pyroclastic flows which occurred during the 1991 eruption of Mount Pinatubo in the island of Luzon, the Philippines (Crisci et al., 2004b).

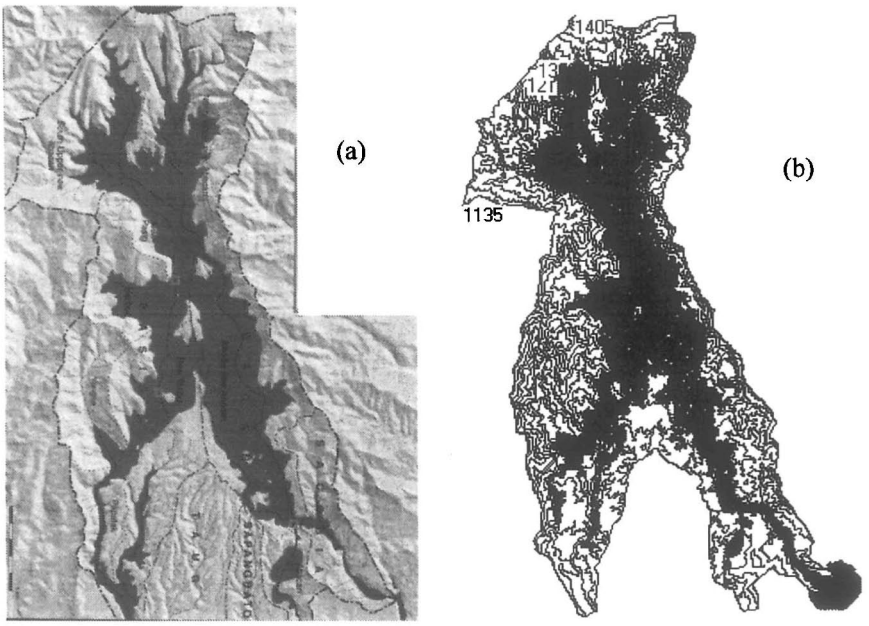


Figure 4. The pyroclastic deposits in Sacobia river valley after the 1991 eruption of the Pinatubo are represented in darker grey in (a) for the real event and in black on surface curves in (b) for the simulated event.

The flows that were generated in the 1991 eruption and their ash clouds travelled for about 15 km from the main vent, covering an area of ca 400 Km^2.

Our simulations refer exclusively to the Sacobia area (Fig. 4,a) because of the good pre and post eruption available geo-morphological data.

The simulation results seems satisfying enough, if the comparison between real (Fig. 4,a) and simulated (Fig. 4,b) event is performed, considering the pyroclastic flow paths and the area involved in the event.

4. CONCLUSION

Our interdisciplinary research group *Empedocles* refined the empirical approach for modelling complex macroscopic phenomena with *CA*, introducing the explicit velocity for simulation applications to surface flows, as lava, debris and pyroclastic flows, that are very difficult to be managed with differential equation systems.

The a-centric (local interactions based) world-view which characterises *CA* models involves a different viewpoint, with respect to partial differential equations, in treating complex macroscopic phenomena. Therefore, physics laws of conservation have to be rewritten (at a given approximation level) in a very different context of space-time discretisation.

The explicit velocity formula, applied for the three models is perhaps too much simple and would be substituted by a more complex one or by different formulae for each phenomenon.

Anyway, SCIARA simulations were definitely improved in comparison with the ones of the previous releases; a good correspondence is evidenced also in the event evolution. Time scalability was successful, because similar, but more precise, simulations were obtained reducing the time corresponding to a SCIARA-$\gamma 2$ step.

SCIDDICA S5a is more problematical: simulations (better than ones of the previous SCIDDICA releases in terms of extension of landslide) involve a number of steps, whose correspondence in time is larger than the estimated duration of the event. An attempt to improve the explicit velocity formula is in progress.

PYR needs deep verifications with simulations of other events.

We have the feeling that an important way for the simulation of complex surface flows is being covered.

REFERENCES

Avolio, M. V., Di Gregorio, S., Mantovani, F., Pasuto, A., Rongo, R., Silvano, S., and Spataro W., 2000, Simulation of the 1992 Tessina landslide by a cellular automata model and future hazard scenarios, *International Journal of Applied Earth Observation and Geoinformatics* **2**(1):41-50.

Barca, D., Crisci, G. M., Di Gregorio, S., and Nicoletta, F., 1994, Cellular Automata for simulating lava flows: a method and examples of the Etnean eruptions, *Transport Theory and Statistical Physics* **23**(1-3):195-232.

Chopard, B., and Luthi, P. O., 1999, Lattice Boltzmann Computations and Application to Physics, *Theoretical Computer Science* **217**:115-130.

Crisci, G. M., Di Gregorio, S., Rongo, R., and Spataro, W,. 2004a, SCIARA, a model for simulation of Etnean lava flows: real case applications, *Journal of Vulcanogy and Geothermal Research* **132**(2-3):253-267.

Crisci, G. M., Di Gregorio, S., Rongo, R., and Spataro, W., 2004b, PYR: a Cellular Automata Model for Pyroclastic Flows and Application to the 1991 Mt. Pinatubo eruption, *Future Generation Computer Systems* (in press).

D'Ambrosio, D., Di Gregorio, S., Iovine, G., Lupiano, Merenda, L., V., Rongo, R., and Spataro, W, 2002, Simulating the Curti-Sarno Debris Flow through Cellular Automata: the model SCIDDICA (release S2), *Physics and Chemistry of the Earth* **27**:1577-1585.

D'Ambrosio, D., Di Gregorio, S., Iovine, G., Lupiano, V., Rongo, R., and Spataro, W, 2003a, First simulations of the Sarno debris flows through cellular automata modelling, *Geomorphology* **54**(1-2):91-117.

D'Ambrosio, D., Di Gregorio, S., Iovine, G., 2003b, Simulating debris flows through a hexagonal cellular automata model: SCIDDICA $S_{3\text{-hex}}$, *Natural Hazards and Earth System Sciences* **3**:545-559.

Di Gregorio, S., and Serra, R., 1999, An empirical method for modelling and simulating some complex macroscopic phenomena by cellular automata, *Future Generation Computer Systems* **16**:259-271.

Di Gregorio, S., Rongo, R., Siciliano, C., Sorriso-Valvo, M., and Spataro, W, 1999, Mount Ontake landslide simulation by the cellular automata model SCIDDICA-3, *Physics and Chemistry of the Earth (A)* **24**(2):97-100.

Frisch, U., D'Humieres, D., Hasslacher, B., Lallemand, P., Pomeau, Y., and Rivet, J. P., 1990, Lattice gas hydrodynamics in two and three dimensions, *Complex Systems* **1**:6-49.

Iovine, G., Di Gregorio, S., and Lupiano, V., 2003, Debris-flow susceptibility assessment through cellular automata modeling: an example from 15–16 December 1999 disaster at Cervinara and San Martino Valle Caudina (Campania, southern Italy), *Natural Hazards and Earth System Sciences* **3**:457-468.

von Neumann, J., 1966, *Theory of self reproducing automata*, University of Illinois Press, Urbana.

ANALYSIS OF FINGERPRINTS THROUGH A REACTIVE AGENT

Anna Montesanto, Guido Tascini, Paola Baldassarri and Luca Santinelli
Dipartimento di Elettronica, Intelligenza Artificiale e Telecomunicazioni
Università Politecnica delle Marche, 60131 Ancona, Italy

Abstract: The aim of this job is to study the process of self-organisation of the knowledge in a reactive autonomous agent that navigates throughout a fingerprint image. This fingerprint has been recorded using a low cost sensor, so it has with her a lot of noise. In this particular situation the usual methods of analysis of the minutiae fail or need a strong pre-processing of the image. Our system is a reactive agent that acts independently from the noise in the image because the process of self-organising of the knowledge carries to the emergency of the concept of "run toward the minutiae" through a categorisation of the sensorial input and a generalisation of the situation "state-action". The system is based on hybrid architecture for the configuration recognition and the knowledge codifies.

Key words: fingerprints; reactive agent; navigation; planning.

1. INTRODUCTION

The problem of the navigation could be defined like the problem of reaching a final position in a space, departing from an initial position and avoiding the obstacles in the environment. The ability of an agent of move in semi-structured or not structured environments is determined from the quality of its perceptions. They depend on the reliability and on the wealth of the input deriving from the sensors of which the agent is endowed. Particularly, in the reactive systems the knowledge of the environment is neither modelled neither memorized but is extracted through sensorial information. Therefore, the behaviours of the agent are revealed as reaction to the information perceived from the environment.

As regard the reactive systems the sensorial processes (perception) and the motor processes (action or behaviour) develop together and also they are inseparable, mutually informative and structured in a way to increase the cognitive system (Jain et al., 1997). The intelligence of a reactive agent is considered as a biological feature (Varela et al., 1991). The intelligence should be considered as the capacity of an agent to interacting with its environment rather than represent it internally (Maturana and Varela, 1987). The concept of "behaviour-based systems" in robotics was introduced for the first time by R. Brooks (Brooks, 1986). According to Brooks the introduction of these systems imposed a turning point to the process of research of characteristics necessaries to the models of intelligent agent (Brooks, 1991). He highlighted the necessity of creating agents those they physically interact with the real world and they act in order to solve complex problems of navigation. This because the models exclusively based on the representation of the environment and on the planning of the movements, are not able to give efficient performances in real and dynamic environments.

In the present work we realized a navigation system of a reactive agent in a fingerprint. The agent receives information on the environment through the sensors of distance of which is endowed. The sensors are distributed uniformly on all the perimeter of the same agent and they calculate the distance from an obstacle along their direction of competence. The agent has in input only the "range images" derived by the sensors of distance.

The purpose of the navigation concerns the movement of an agent in the environment avoiding the obstacles, which are represented by the ridge line, recognizing and reaching a goal that is represented by the so-called minutiae (termination and bifurcation) present in the fingerprint. This task is rather complex because the environment consisted of very narrow corridors the so-called valleys and delimited by the walls that are represented by the crests. In order to facilitate the task of the agent and therefore to improve the navigation we realized a "hybrid" system that adds a planning to a reactive behaviour (Montesanto et al, 1999). Particularly, the system has a further task that it concerns the planning of a path toward a goal, using as reference some prototypic paths to follow, in a very elastic manner. The term "prototypes" presupposes that the system has the concept of "path toward the goal". The system has not this a priori knowledge, but it has to extract it, categorizing the sensorial input and generalizing the situations "state-action". In this way the system obtain a local knowledge. This local knowledge is used to build a global knowledge through an abstraction of sensorial data.

2. ANATOMY OF A FINGERPRINT

A fingerprint consisted of ridge lines, also called crests, that they could intersect or end, forming a particular design known as ridge pattern. The points in which the crests finish or fork are said minutiae and they constitute an important factor of the discrimination of the fingerprints. In this work, we consider only the bifurcations and the terminations (Bolle et al., 2000) among all the different classifiable minutiae.

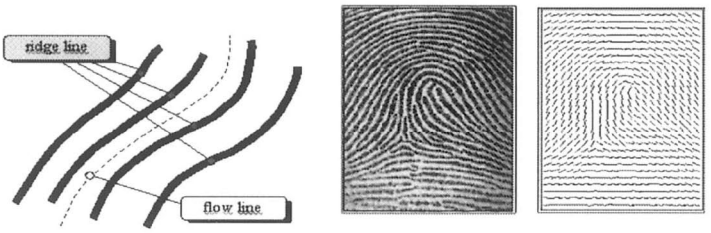


Figure 1. Representation of Ridge Line, Flow Line, Image of a Fingerprint, and Directional Field.

An algorithm of recognition of digital fingerprints has to give reliable answers in a few times. The environment of navigation is rather complex, disturbed, narrow but static. In fact, the crests and valleys form a maze, in which the walls are very close among them. In order to compare two digital fingerprints it is necessary to consider some particular characteristics. These characteristics must be able to establish univocally the corresponding fingerprint from which they are previously extracted (Jain et al., 1997, 2002). The information about the minutiae refers to their position and to their orientation in the structure of the crests of the fingerprint. In order to extract a vector of minutiae it is necessary to find particular points in a complex environment, therefore it is necessary to move in this environment until an efficient position is not found. The research of minutiae can not be performed on all the surface of the image. This because in the edges the contact between the sensor and the finger is latent and the deformation of the skin is excessive. For these reasons the edges of the image are noisy and then the structure of crests and valleys is completely lost. So the area of research of minutiae concerns the central part of the image and therefore the structure of the crests is clearer and then the singleness is more real and likely. The main concentration of minutiae is in the central zone around the singular points. An excessive area of research could reduce the precision performance of the algorithm of survey.

The fingerprints matching (Jain et al., 2001, 2002) must have an extraction system of characteristics that receives in input an image of a digitalized fingerprint, and that gives in output a whole of characteristics. The characteristics will be compare with those in a database belonging to fingerprints of subjects previously recorded. So, the matching system receives in input two whole of characteristics, one related to the fingerprint that must be identified and one extracted from the database. They have compared through an appropriate method, one of these is the object of the present work. This comparison procedure must be repeated until the whole database is not controlled in order to give an answer relative to the identity of the subject. As we said the minutiae are the characteristics of the fingerprints that we considered. Since the vectors of minutiae associated with each fingerprint are calculated, there is the comparison procedure among the characteristics. The recognition of the fingerprints must be invariant respect to possible translations, rotations or variation of dimensions of the same fingerprint. For this purpose we used the calculus of moments, particularly the regular invariant moments (RMI). The moments are expensive from the point of view of the calculus, therefore we stopped the calculus to moments of the third order. This is sufficient to reconstruct the image and to recognize the pattern. The moments are evaluated for all of minutiae. Then the real comparison is realized by a simple procedure that adds the mean square errors between the RMI concerning the template image and the RMI concerning the sample in the database. If the value is lesser to a threshold, arbitrarily chosen, the two fingerprints coincide, and so they belong to the same individual. In order to obtain efficient performances of recognition is necessary to choose a right threshold.

3. EXPERIMENTAL SET-UP

The environment in which the agent navigates is similar to a maze, consisted of ridge lines (walls) that they represent the obstacles. In the environment there was not "landmarks", or points of reference, those they can be useful to the orientation and to the movement of the agent. So, the agent has not information about its position in the environment and about the position of the goal. Besides the agent has not a global vision of the environment, but only local information.

The information relative to the environment derives from the range sensors, simulated "sonar", those they give the distance between the system and the obstacles met along the direction of survey. Particularly the sonar gives the distance in pixel from the valley or from any other obstacle along the direction in which the sensor is oriented. In this case, there are 8

directions of survey useful to check (8 sensorial profiles), so that they form an angle of 45° between the previous and the successive sensor. The digital image represents a matrix, and for each pixel we considered an 8-dimensional metrics (diagonal distance). This kind of distribution contemplates all the information near the point, but moving away from the centre some information could be lost. Therefore, in order to have a better precision we could consider a bigger number of range sensors. In this way the information about the obstacles could be more accurate also away from the centre. In this particular case a high depth of exploration is not necessary because the conditions of the environment change notably from a position to another. It is important to have more precise information near the agent, guaranteeing an immediate answer to avoid the obstacles during the navigation.

4. THE SYSTEM ARCHITECTURE

In this work we propose a hybrid complex architecture of neural networks for the classification of the sensorial inputs and for the navigation. This architecture is composed by two modules. The first module consisted of two Self-Organizing Map (SOM) networks of Kohonen (Kohonen, 1982, 1990, 1995) and of one Multilayer Perceptron (MLP) based on the Error-Backpropagation rule (Minsky and Papert, 1969; Rosenblatt, 1958). The first module elaborates the sensorial inputs giving a classification of them in output. The output of the first module represents the input of a second module consisted of an only MLP, that it manages the phase of navigation.

4.1 Categorization and association sensorial profiles-prototypic paths

Initially the user constructs some paths in the environment, each of these is considered like a succession of positions appropriately sampled. For each position belonging to a path is associated a sensorial profile. This profile is constructed through a largely used in literature method. For each sonar sensor of which the agent is endowed, the x distance from the first obstacle along the direction of the sensor is calculated, and then the potential P associated to that sensor is determined using the following equation (1):

$$P = k_1(d_m - x) + k_2(d_m - x)^2 + k_3(d_m - x)^3 \qquad (1)$$

where d_m is the maximum distance determined by the sensor, k_1, k_2, k_3 are opportune parameters.

The sensorial profile is the whole of P values, one for each sensor ordered from the parallel sensor to the direction of the actual movement and then to the next sensors considered in hourly sense.

The normalized sensorial profiles between [0,1] are the input of a feedforward network with two layers and a tri-dimensional layer of categorization subjected to an unsupervised learning based on the SOM algorithm of Kohonen. Then, after this first phase the values of the weights of the categorization network are extracted. For each position of the agent in the environment, we obtain the sensorial profile associated to the actual position through the values of the weights of the adapted network. This phase is relative to the categorization of sensorial profiles.

Successively there is the phase regarding the categorization of paths. Each path (except the starting point) is represented like a succession of couples: direction of movement and length of the path. This kind representation is obtained extracting the characteristic points from the whole path that are changing points of direction: direction and length. Then, the extracted numerical values are normalized between [0,1]. The individual paths are the input of a two layers feedforward network, with 3-dimensional categorization layer, learned by the SOM algorithm. After the training, the values of the updated weights of the network associate to each path its category of belonging.

Once that the paths and the sensorial profiles are categorized, there is the procedure of association profiles-paths, in order to obtain the information concerning the path for each profile. For each path we considered all the significant points and one at a time they are categorized through categorizator of sensorial profiles. Then the whole path is categorized through categorizator of the paths. The coordinates of the category of the profile and the coordinates of the category of the path which have the same profile are memorized in a file of profile-path association. Since it could happen that a same profile belongs to more paths, we valued the individual paths through three categories: 1) path that directly conducts to the goal, 2) path that conducts to the goal with a long way 3) path that does not conduct to the goal.

In order to identify the prototype path to follow the recognition of local sights is applied (image of range recorded from the sensors of distance). Such prototype path is learned during the first phase when the system is driven toward the goal: in this way the system built a set of representative paths made up of "sights" concatenated. Because of the aspect of the environment (a maze) the categories of the sensorial inputs are relatively

few. Therefore such input are classified in categories and so the paths are characterized by chain of categories more or less recurrent.

4.2 Reaction and planning

Once that the association profiles-paths file is created, there is the learning phase of the second module. The three layers MLP receive in input the categories of the sensorial profile and the categories of the path associated to the position in which the agent is situated. The output of the network represents the direction toward which the agent must move and the information concerning the presence or less of a minutia. The following description analyzes the implemented procedure to identify the input and the desired output. The inputs are determined beginning from the coordinates of a point of the environment occupied by the system. The sensorial profile associated to the point is calculated and through the categorizator of profiles is categorized. The recorded path until this moment is categorized to determine the coordinates of the category of the same path. If there is not any associated category, the coordinates have value equal to zero. The coordinates represent the input of the MLP. For each path we consider the profiles that they compose it.

The desired output are constructed beginning from the point, in which the agent is situated, in other words the associated sensorial profile is determined and through the categorizator of profiles is categorized. Consequently the category of the path until this time is determined and it is associated to the actual profile. The vector of the weights corresponding to the category of path represents the prototype of the path belonging to this category. This vector is decoded obtaining the coordinates of the points belonging to the same path and the order in which are situated in the path, through an inverse procedure to that used for create the file of management of the learning of the categorizator of the paths. For each of these points the sensorial profile is determined and is looked for the point that has the sensorial profile nearer (according to the Euclidian metric) to the sensorial profile of the point in which the agent is located. Since this point the successive immediately point in the path-prototype is identified and the direction of the line between the two points is determined. This direction represents the desired output.

The supervised learning of MLP network is implemented using a file containing all the necessary information. For each path the profiles that compose each path are extracted, and the association of the category of the profile and of the category of the path is the input of the neural network. The direction corresponds to the direction of the tract of path successive to the

characteristic point of the particular profile, in the prototype of path to which it belongs.

Concerning the information on the kind of profile identified from the sensors of the agent, the knowledge about the type of prototypic path is exploited. The paths those they lead to a particular minutia (identified with the category 1) have a profile relative to a minutia. The profile is not relative to a minutia for all the other typologies of path. Regarding a minutia, the direction of the shift is not more important and it has a value equal to 0, while the others two output of the network identify the type of minutia: [0 1] for the termination and [1 0] the bifurcation.

4.3 Navigation

At the beginning of the navigation we have to define the characteristics of the environment in which the agent moves, the starting point, the initial direction and the condition why the navigation ends. The navigation proceeds in the following way. The sensorial profile associated to the point relative to the position of the agent is calculated and this profile is categorized, obtaining the coordinates of category. On the basis of the information concerning the path completed by the system we look for the category of such path, that is the category associated to category of the actual profile. Then the coordinates relative both to the category of the profile and to the category of path, are encoded through a binary representation and they represent the input of perceptron. The output of perceptron is the desired direction and the typology of image (if there is a minutia or a crest). The angular direction of moving is determined decoding the direction of the output. The x, y coordinates of the new point reached by the agent are obtained using the following equation:

$$\begin{aligned} x &= x_0 + v_0 \cos(\theta_n) \\ y &= y_0 + v_0 \sin(\theta_n) \end{aligned} \tag{2}$$

where x_0, y_0 are the coordinates of the point relative to the previous position of the agent, v_0 is the intrinsic speed (constant for hypothesis). The angle θ_n is evaluated through the (3) equation:

$$\theta_n = \theta + [k_p(\theta_f - \theta) + \varepsilon\, Obst]\Delta t \tag{3}$$

θ is the direction of the agent in the point with coordinates x_0, y_0, k_p and ε are opportune parameters, Δt is the width of the temporal interval

corresponding to one step, *Obst* is determined by the following equation (4) and it identify the potential of the obstacle along the direction:

$$Obst = \begin{cases} h_1(r_0 - r_{\min}) + h_2(r_0 - r_{\min})^2 + h_3(r_0 - r_{\min})^3 & r_{\min} \le r_0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

In the last equation h_1, h_2, h_3 are parameters, r_0 is an opportune value of threshold (the distance for which the obstacles does not constitute a danger). Besides $r_{\min}$ represents the distance to which there is the first obstacle along the direction identified by the angle θ_f. The next equation could be considered for the calculus of the potential of the obstacle or repulsive.

$$Obst(K) = A\frac{e^{-\alpha K}}{K} \quad (5)$$

where α and K are opportune parameters.

Then the new x, y coordinates are replaced to x_0, y_0, and the process restarts. The equation (3) means the avoiding obstacles function and it deserve a brief explanation. The direction of the shift is influenced by the previous direction, by the direction θ_f returned from the system of control of the agent (MLP), for the presence of the obstacles along the same direction. The direction θ_n changes if there are obstacles (potential > 0). The changing of direction is greater if the obstacles are closer to agent (potential with a value near to 1). This term determines the variation of the delta angle $\theta_f - \theta$. Therefore this value is added to θ in order to give the actual direction.

During the navigation the system has only local information regarding the category of the sensorial input. This information is associated to the information relative to the category of the path performed until that moment and kept in memory by the agent. In case this profile belongs to a prototypic path, the system is implicitly reinforced to follow this particular path. This means that the systems have to perform the encoded action, step by step. In the following state the system could not find again the category of the sensorial input contained in the path that it was following. So, the system looks for another path that it contains the category of the actual sensorial input. If this path is found, this last new path is reinforced whilst the previous is neglected. This procedure is an implicit reinforcement of the state-action couple, that it is realized by the existence of a prototype path to follow. The system reacts against the sensorial input, but this reaction is totally untied by the context and particularly by the specific position of the same system. Therefore it reacts to the category of input.

5. EXPERIMENTAL RESULTS

In this section the obtained results implementing the proposed architecture are related, supposed that there were 20 agents in the environment of navigation. The agents scattered along the surface succeed to explore a wide surface of the same image, and so this makes easier the identification of a bigger number of minutiae. Because of the low quality of the original image of the fingerprint, experimental results demonstrated that the agent are not been able to identify a certain number of bifurcation and termination minutiae. The image of the fingerprint is acquired by a capacitive sensor and is originally a colour image. Successively the image is converted in a binary format, assigning the 0 value to the valleys and the 1 value to the crests.

The experiments are relative to a sample consisted of 11 individuals and so with a sample of 11 fingerprints. In the first experiment when we considered the sample of fingerprints, we did not make any selection, in order to consider only the better images, but we chose the first acquired images for each individual, like a normal system of acquisition. The first fingerprint is memorized like sample in the database and it assumes an identity, whilst the following fingerprints are used for the matching procedure. The images of testing from which we identified the minutiae and we compared with those presents in the database are 31 (random taken by the 11 individuals). So totally the fingerprints acquired by the same sensor are 42.

Considering the total number of images (31) to identify we obtain a rate of recognition rather low with a value equal 48%, as shown in the table 1. During the testing phase, whenever we considered only the best images (19 images on 31) the value of recognition drastically increase, obtaining a rate of recognition equal to of the 79%. The results are shown in the following table 1.

Table 1. Results of identification of fingerprints.

	Identified	Rejected	False identification	Total
All the images	15	0	16	31
Best images	15	0	4	19

Substantially the 12 rejected images, those they are visually worse than the others, represent all the images those the matching algorithm does not succeed to identify. On the basis of the obtained results is clear that the quality of the images is an important and decisive factor for the recognition. The images of low quality notably worsen the performances of the system

determining a false identification. When the images increase their quality, results clearly better are obtained.

6. CONCLUSIONS

The architecture that we realized intends to resolve the problem of recognition of fingerprints making use of the concept of navigation of a reactive agent in a complex environment. Particularly the image of a fingerprint could be comparable to a maze. The developed method is an interesting and above all innovative application for this kind of work and environment. This approach originates like a system of control for a reactive agent that moves in a particular environment in order to reach a goal on basis of local information obtained during the exploration phase. In case of an environment identified by a fingerprint, there is not any landmark; there are not a priori information about the distance from the goal represented by the minutia. Besides the environment is very narrow, there are, in particular, narrow corridors, relatively rapid curves, and sudden obstacles to avoid.

Considering these problems we have used a "hybrid" system. In other words, the system makes use of local information reacting in real time to sudden stimuli but it is also based on a previous procedure of planning of the path, through the prototypic paths. The experimental results have shown that the rate of recognition notably depends on the quality of the image that they must be identified. In fact, we have experimentally verified that the testing procedure guarantees a good level of recognition on the images of a rather high quality, result that decreases if also images of worse quality are considered. In order to overcome this problem we could introduce a previous phase of pre-processing of the image that reduces the noise, bettering the structure of the crests and identifying more easily and with better precision the minutiae. Besides the arrangements made in phase of development, that is to have a local localization, to codify the information valley and then obstacle in the memory of the agent, they could open a passage to the applicability of the system in any kind of environment.

REFERENCES

Bolle, R. M., Senior, A. W., Ratha, N. K., and Pankanti, S., 2000, Fingerprint minutiae: a constructive definition, *Lecture Notes in Computer Science*.

Brooks, R. A., 1986, A robust layered control system for a mobile robot, *IEEE Journal of Robotics and Automation* **2**:14–23.

Brooks, R. A., 1991, Intelligence without representation, *Artificial Intelligence* **47**:139–159.

Jain, A. K., Hong, L., Pankanti, S., and Bolle, R., 1997, An identity-authentication system using fingerprints, *Proc. IEEE* **85**:1365-1388.

Jain, A. K., Prabhakar, S., and Pankanti, S., 2001, Matching and classification: a case study in fingerprint domain, *Proc. of Indian National Science Academy (PINSA)* **67**:223-241, (Special Issue on Image Processing, Vision and Pattern Recognition).

Jain, A. K., Prabhakar, S., and Pankanti, S., 2002, On the individuality of fingerprints, *IEEE Trans. On Analysis Machine Intelligence* **24**(8).

Kohonen, T., 1982, Self-organized formation of topologically correct feature maps, *Biol. Cybern.* **49**:59-69.

Kohonen, T., 1990, Self-organizing map, *Proc. IEEE* **78**:1464-1480.

Kohonen, T., 1995, *Self-Organizing Maps*, Springer Verlag.

Maturana, H., and Varela, F. J., 1987, *The Tree of Knowledge: the Biological Roots of Human Understanding*, New Science Library, Boston, MA.

Minsky, M., and Papert, S., 1969, *Perceptrons*, MIT Press, Cambridge, MA.

Montesanto, A., Penna, M. P., Pessa, E., and Tascini, G., 1999, A hybrid architecture for autonomous robots undertaking goal-directed navigation, *Workshop Apprendimento e Percezione nei Sistemi Robotici*, Parma.

Rosenblatt, F., 1958, The perceptrons: a probabilistic model for information storage and organization in the brain, *Psychological Review* **65**:386-408.

Varela, F. J., Thompson, E., and Rosch, E., 1991, *The Embodied Mind: Cognitive Science and Human Experience*, MIT Press, Cambridge, MA.

USER CENTERED PORTAL DESIGN: A CASE STUDY IN WEB USABILITY

Maria Pietronilla Penna[1], Vera Stara[2] and Daniele Costenaro[1]
[1]*Università degli Studi di Cagliari, Facoltà di Scienze della Formazione, Dipartimento di Psicologia; E-mail: maria.pietronilla@unica.it, costenaro@email.it;*
[2]*Università Politecnica delle Marche, Facoltà di Ingegneria, DEIT-Email: v.stara@univpm.it*

Abstract: The chance to share information, receive feedbacks, tips and suggestions, in a "one-to-many" way, represented the strength of the WWW in recent years, becoming the most important channel of information delivery of all times. However, the "virtually unlimited" growth of the web finds now an actual limit: the desire of employing new technologies when developing internet sites has pushed web designers to pay more attention to tools first of all, almost as in a closed system, rather than to process of human-computer interaction. Starting from a real case study, this contribution proposes a total analysis of the multiple aspects that make usability the true problem of the web.

Key words: web usability; user centered design.

1. INTRODUCTION

The "Web Usability" seems to represent the perfect bridging point between cognitive science and web design and proposes itself as a cross-section discipline adopting a systemic point of view, that leads to complexity analysis of relationships between typical elements of human-computer interaction, on one side, and on the other side as an approach in which the sum of single disciplinary units represents more than the single parts, both qualitatively and quantitatively. In fact web sites, as cognitive artifacts equipped with graphical interfaces would fulfill criteria of effectiveness, efficiency and satisfaction in the specific contexts of use (ISO 9241). Usability is not only a technological factor, and could represent a fixed and reliable point in web *mare-magnum*, and, moreover, a challenge for e-business success.

Enterprise portals, being mirrors of represented organizations, can be viewed as tipical cases of e-commerce services in which this aspect can have important consequences in the relationship between the enterprise itself and its users. The portal has two fundamental functions: it combines and organizes huge amounts of information, in particular data generated within the organization, and it can retrieve and show any of the contained intems in a user-interface that should be simple and easily usable.

However, many e-commerce sites in the web do not respect usability criteria and risk to be not appreciated by internet users, because of common mistakes both in design and communication fields, so that an e-commerce site is not of daily use if its interface will not be simple as walking down the street, entering a real shop and purchasing a product.

The costs/benfits ratio of usability engeneering ranges form 1/10 up to 1/100 for each dollar spent for usability implementation, and this is surely a benefit both for user and for the e-business company.

A positive user experience of portal contents could produce as a consequence the increase both in profit and in sale volume, while lowering costs for customer care. An increase in parameters SPR (Site Penetration Rate) and CCR (Consumer Conversion Rate) is equivalent to decrease of online customer acquisition costs with consequent increase in satisfaction (Michelini, 2002): "customer looks at the site and then pays" (Nielsen, 2000).

Starting from a real case study, this contribution proposes a total analysis of the multiple aspects that make usability the true problem of the web, developed in two stages:

- a phase of qualitative analysis of the web site according to the model EtnoTeam Lab (200x);
- a phase of quantitative analysis to measure:
  a) the modalities of interaction between specific web site and customers in a experimental setting based on assisted navigation;
  b) the satisfaction of customers after navigation.

2. 1ST STEP

The chosen model subdivides the site's quality into five dimensions:

1. *Communication*: a dimension which is a summary of the features used to catch up communication objects established with respect to target customers;
2. *Functionality*: dimension along which it is possible to measure the interaction modalities between the portal and its customers with respect to web site's aim;

3. *Contents*: it makes reference to relevance, completeness, reliability of the contained information;
4. *Accessibility*: it measures the ease in accessing the site, with respect to the platform used at the client side and to its connectivity;
5. *Management*: it makes reference to aspects related to control of navigation in the site.

2.1 Communication

Trenitalia is a large organization and it is not necessary to dwell on the brand: it clearly communicates the web site's information aim. We attribute greater prominence to the headline that camps in the cover of the site and in the Home Page, because its content not corresponds to the actual situation.

Visual communication: the chosen colours for the site are few, but not well chosen: colours belonging to opposite ends of the visual phantom are often used, like yellow and violet.

Figure 1. Synthesis of the chromatic scale.

Another misleading element is given by colours of some active icons that breaks the chromatic scale chosen for the rest of site, often creating distorted perception of the consequences of correlated actions.

Scegli	Partenza	Arrivo	Treni	Durata	Acquista
1.	13:30 ROMA TE	18:00 MI C.LE	ES*	04:30	
2.	14:30 ROMA TE	19:00 MI C.LE	ES*	04:30	
3.	15:30 ROMA TE	20:00 MI C.LE	ES*	04:30	
4.	16:10 ROMA TE	22:15 MI C.LE	IC	06:05	
5.	16:30 ROMA TE	21:00 MI C.LE	ES*	04:30	

Figure 2. The meaning of basket icon, now widely shared from the community web, can suggest purchase of the shown product, the red colours of the icon create a perception of danger and the possibility that a click could start the process of purchase in a not reversible way. Actually this icon, in Trenitalia, represents the only way to know the price of the selected ticket.

The use of the font maintains the web site in the average: the Verdana 8 character is used for the body of the text, peculiar to many of the last generation sites; it is not present any application useful to modify the font size, and this creates a problem of accessibility: because of the font; namely at high resolutions the site can difficultly be usable by people having problems at eyes, and moreover this situation cannot be modified even when acting on the preferences of the browser (testing on IE 6, Mozilla e Opera 7).

The explanation of links is completely absent, being based on the useless repetition of the link label itself. It is immediate to wonder about the reasons of this choice: the huge amount of data of this site requires a better explanation of every link. Considering not useful to send additional explanation of the link, it is a waste of bandwidth to send redundant information in labels.

Links are highlighted with various colours during navigation; the majority of them are textual, and some are icon link.

2.1.1 Logical subdivision of the information in homepage

The most clearly noticeable subdivision is in the home page, between the dynamic zone of interaction with the customer (on the left), and the other links in the central and right columns.

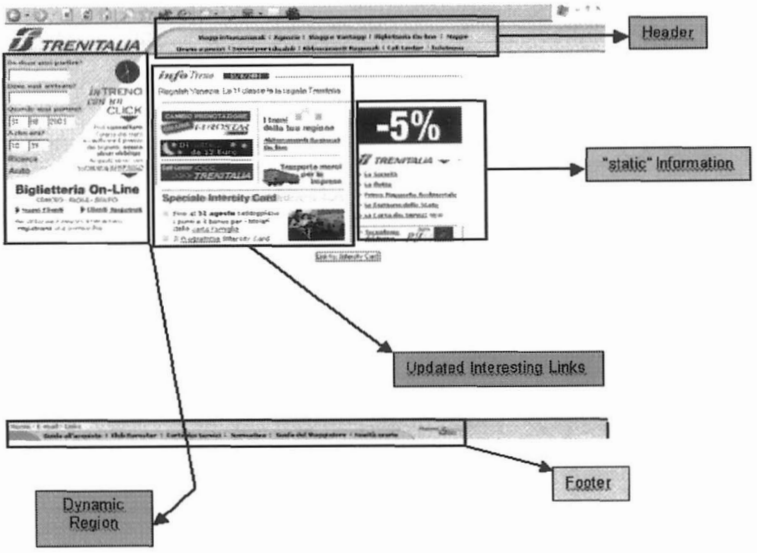


Figure 3. Logical subdivision of the information in homepage.

The central column is striking because it is concentrated and too much rich in colours; but the zone that is first perceived from the eye and captures the attention (also for a precise one choice of order of page download) is the dynamic region on the left.

Header and footer of the page contain links that can be found also in the body of the same homepage.

This choice can be discussed and it makes navigation less easy, confuses the customer and decreases usability occupying additional bandwidth. The layouts of the other pages are different from the homepage, and also between the various pages structural (and rather obvious) differences exist; these differences can be connected to the whole subdivision of pages according to their argument.

The reading direction is the classical, from left to right and from top to bottom. Texts are well chosen and short, and this makes reading "on the screen" easy and not heavy.

A negative aspect is the excessive usage of popup windows, in order to communicate with the customer. They are used both for short texts and for loading entire pages of the site, making navigation heavier.

The graphical layout is not invasive, but rich; self-explaining photos are used to present business activity of the organization; in the cover of the site an animation realized with Macromedia Flash™ is placed, and in the homepage a small application that reproduces an analogical clock is present, realized with the same technology. The two graphical elements slow down navigation.

2.2 Functionality

The most important interactive function of Trenitalia site is searching for timetables of trains and the online ticket purchase. It is possible to access these functions directly from the dynamic areas of the homepage. The search engine and the links force the customer to paths which are not in a perfect accordance with the idea "in train with a click". Clicking on the link "hours and prices" in the header of the homepage it is possible to get access to the same search page.

We show here the complete path in order to search and purchase a railway-ticket.

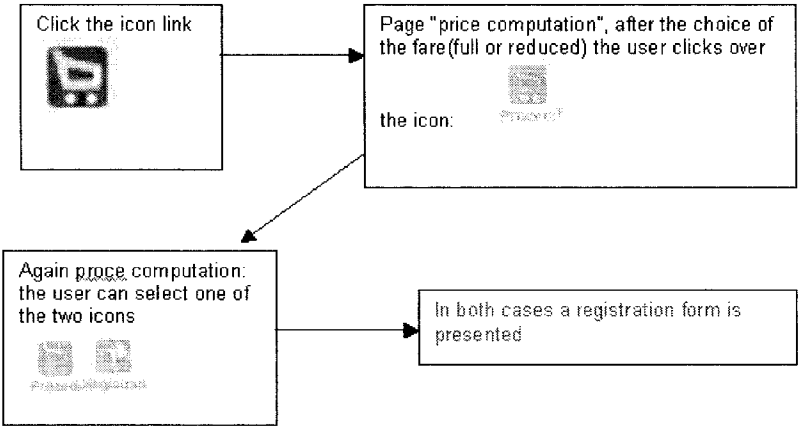


Figure 4. Interactive function.

The registration page is moreover available clicking in homepage on link "the ticket office on-line" and "in train with a click". The same link "ticket office on-line" is reported in the header of the site. We found no reason for the creation of such ambiguous and misleading paths and link redundancies that give the same results. Customers experience confusion while using a service that, for its nature and the declared aim, would be fastly understandable and easy to use, instead of becoming extremely complex.

The customer registration to Trenitalia services is made up by three fast steps; they do not require the insertion of too many data. The lack of fields to fill up for market researches is a remarkable merit. Unfortunately we must signal that the password cannot be chosen from the customer, but sent by Trenitalia via e-mail, and delivery times are variable. That suggests a clear bad working of server side's site. The search engine for the timetables and railway-tickets is precise and detailed, but confirms problems at server side, since in 10 days of testing (from the 21/08/03 to 31/08/03) only once (29/08/03 hours 19:07) it calculated the price and produced the ticket on-line purchase.

It would be opportune to pay more attention to control (or better, modernization) of the server. An interesting and well-built tool is the service "maps" (constructed with JAVA™ technology), that allows the customer to get much interactivity and access to various information, but it is not put in good evidence inside Trenitalia site; the link is in fact present only in the header of the homepage.

2.3 Content

Information is introduced in a clear way, with titles that are coherent with the proposed content, even if in some cases the lack of comments to the link can generate some misunderstandings on the content of the visited page, but these case are rare and they do not disturb navigation.

The access to the information is constantly guaranteed from the structure of the pages constituted of three frames; even if it is not designed in compliance with the most modern tendencies (frames can give several problems of management of the pages, and makes difficult the indexing of site on search engines, thus reducing accessibility); in this case the choice is functionality in order to always maintain in sight the top and bottom menus.

The external links are very well organized and coherent with the economic-social destination of the site.

2.4 Accessibility

Trenitalia is accessible from three addresses and is indexed on the most important world-wide search engines. From the point of view of accessibility, intended as portability and accessibility to the customers with disabilities, the site shows the previously mentioned problems: the impossibility to modify the font size has been remarked, together with a mistaken use of the chromatic phantom; the use of colours does not take into account the chromatic alterations provoked from the colours-blindness; the homepage examined with the Vis-Check software showed that in colours-blindness presence a large part of the contrasts gets lost.

The service is accessible in five languages, not limiting to per-word translation but offering a service differentiated for every country (Italy, England, Spain, Germany, France).

2.5 Management

A greater control would be surely necessary at server side since in the ten days of testing we found some links not working; however they have been restored in 24h.

3. 2ND STEP

To a sample of 130 university students (age between 18 and 35), we assigned a navigation task: know a railway-ticket's cost.

Every performance has been monitored through a desktop recording software that recorded every single navigation.

The user-experiences have respected the times and the subjective interaction ways with the site.

In order to estimate how much the site is usable we constructed an evaluation grid that contains some situations described in the following taxonomy inspired to the Cognitive Walkthrough:

- The customer attains the goal: the subject can decide to continue the interaction or to abandon the navigation ("goal" variable);
- The customer has the necessary knowledge to attain the goal: the system does not preview specific competences for its fruition;
- The sample includes expert and inexpert people ("expertise" variable);
- The association between correct and attainment action is perceived from the customer: in the system are implemented explicit and recognizable marks that guide the customer to attend the goal.

3.1 Hardware and Software Requirement

We used a Personal Computer Intel Pentium IV CPU 2.02 GHz and 256 MB of RAM, whit monitor Plug and Play KEIMAT ™ on RADEON 7000/RADEON VE ™ 17 thumb, resolution 1024x768 pixel; it was connected to a LAN 256 Kbit/s; we used a Microsoft Internet Explorer 5.0 Italian version browser.

3.2 Data analysis

77 subjects attained the goal instead of 53 subjects that did not attain the goal. An analysis of data, done through ANOVA evidenced how who attained the goal used a lower time of navigation than who did not attempt the goal ($F = 29{,}128$; $df = 1$; $p < .000$). ANOVA didn't evidence significant differences in direction of "expertise" and "goal" variables.

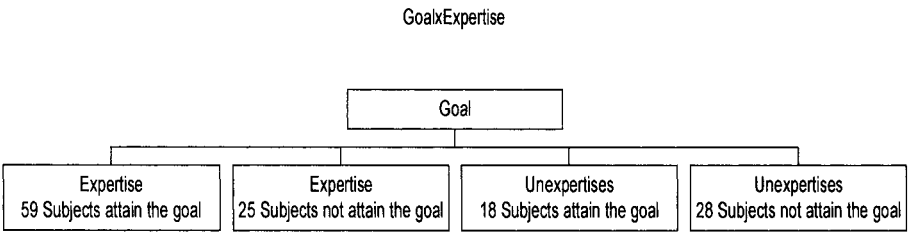


Figure 5. Goal vs Expertise.

These outcomes evidence that a previous knowledge of the used iconology should concretely support subjects in site usability and it should favours who have a medium navigation expertise: a second analysis, done through a Log Linear method, evidenced how to attend the goal it is fundamental a good experience (Y2 = 11.9 with $df = 1$ and $\alpha = .001$), a fairly metaphor recognition (Y2 = 37.36 with $df = 1$ and $\alpha = .001$).

These data evidence how Trenitalia is so far from a User Friendly interface and from User Centered Design (UCD).

UCD is not only a philosophy that places the user at the centre, but it is also a systemic process that improves the usability focusing on cognitive aspects, whereas, in this study case, icons and links are misleading for meaning, scope, position inside the page and the diagram; they suggest misleading paths and do not increase the times of navigation but also prevent the attainment of task unless the interface is known. We see, also, how users are not impressed by complexity that seems gratuitous, especially for those users who may be depending on the site for timely and accurate work-related information. The interface metaphors should be simple, familiar, and logical, and for this reason, we assumed that Trenitalia satisfies only the first requirement of the Cognitive Walkthrough.

4. CONCLUSIONS

From an institutional site of the dimension of Trenitalia, we would expect a greater attention to usability and accessibility: the insufficient organization of the informative architecture (we refer in particular to the ambiguous organization of menus and links and to the somewhat hard paths of the dynamic area of the site) is not exactly what a service of social usefulness would offer. From first qualitative analyses and consequent quantitative analysis we suggest that navigation is made difficult by these defects and by a not sufficient monitoring of the server side. The site is penalized by a lack of accessibility from part of customers with eye problems. However, the not invasive graphical layout and a good web-writing make pages quick and readable. Trenitalia's interface invites the customer to interaction. However when accessing it in a deeper way customers perceives a sure loss (lack of correct affordance). Icons and links used do not aid the perception of correlation between action and feedback of the system, due to incorrect use of metaphors highly shared (incorrect use of natural mapping). This does not supply aids, but hinders the attainment of task in terms of time and link used. This investigation reveals once again the necessity of user centered design and underlines the importance of integration between technical standards defined and features of human cognitive processing, like perception,

attention, learning and memory in the web-design domain (Penna and Pessa, 1996).

REFERENCES

Bottazzini, P., Fietta, R., Fonti, D., Mandelli, S., and Ponasso, L., 2003, *Web Usability*, Tecniche Nuove, Milano.

Bussolon, B., and Valentini, E., 2003, *Progettazione orientata all'usabilità*; http://www.hyperlabs.net/ergonomia/dispensa/02.html .

Etnoteam; http://www.etnoteam.it .

ISO 9241, 1994, *Ergonomic Requirements for Office Work Visual Display Terminals, Part 10, Dialogue principles.*

ISO 9241, 1996, *Ergonomic Requirements for Office Work Visual Display Terminals, Part 11, Guidance on specifying and measuring usability.*

Nielsen, J., 2000, *Web Usability*, Apogeo, Milano.

Michelini L., 2002, *Perché progettare siti usabili*, http://www.prohtml.it/print_articolo.asp/id_140/stampe.html.

Penna, M. P., and Pessa, E., 1996, *Le interfacce Uomo-Macchina*, Di Renzo, Roma.

Postai, S., 2003, *Il mestiere del web*, HOPS, Milano.

Shneiderman, B., 1997, *Designing User The User Interface: Strategies for Effective Human-Computer Interaction*, Addison Wesley, Reading, MA.

Stara, V., Costenaro, D., Merenda, E., Murgia, S., Piroddi, M. A. L., and Farci, E., 2003, La costruzione di interfacce efficaci: un'indagine sperimentale sul portale di Trenitalia, *Atti del Congresso AIP*, Bari, pp. 88-90.

Usability; http://www.usability.gov .

2

BIOLOGY AND HUMAN CARE

LOGIC AND CONTEXT IN SCHIZOPHRENIA

Pier Luca Bandinelli[1], Carlo Palma[2], Maria Pietronilla Penna[3] and Eliano Pessa[4]

[1]*Dipartimento di Salute Mentale ASL Roma "E". Servizio Psichiatrico di Diagnosi e Cura c/o Azienda Complesso Ospedaliero S. Filippo Neri Via G. Martinotti 20, 00135 Roma, Italy; pluc.band@libero.it*

[2]*Istituto d'Istruzione Superiore, Via Cesare Lombroso 120, 00168 Roma, Italy; briciola.02@tiscali.it*

[3]*Università degli Studi di Cagliari, Facoltà di Scienze della Formazione Località sa Duchessa, Via Is Mirrionis 1, 09123, Cagliari, Italy; maria.pietronilla@unica.it*

[4]*Università degli Studi di Pavia, Dipartimento di Filosofia e Psicologia Piazza Botta 6 (Palazzo S. Felice), 27100 Pavia, Italy; eliano.pessa@unipv.it*

Abstract: In this paper the authors analyze the pattern of reasoning in schizophrenia, according to proof theory. In particular they consider the clinical form of "organized" (paranoid subtype) and "disorganized" schizophrenia. In the first form they focusing on the conservation and an "excess" of the use of standard inference rules that formalize certain logical modes of reasoning, but also the incorrect use of premises not context sensitive. The authors also suggest that in disorganized subtype the inference rules are not derived from a tautological proposition, but the patient use non-standard inference rules like assonance, analogy and metaphor, relative to a particular focalized and pervasive mental state. In these case the premises and conclusion of reasoning are represented by a formalized logic expression

Key words: logical reasoning; schizophrenia; common sense.

1. LOGICAL REASONING IN SCHIZOPHRENIA

Patients with schizophrenia exhibit an exceedingly wide range of symptoms from a variety of domains. The cardinal features are abnormal ideas (such as delusions); abnormal perceptions (such as hallucinations); formal thought disorder (as evidenced by disorganized speech); motor,

volitional, and behavioural disorders; and emotional disorders (such as affective flattening or inappropriateness). Delusional ideation produces much of the social alienation, lack of treatment compliance, and poor functioning associated with these diseases.

In addition to these diverse, and sometimes bizarre symptoms, it has become increasingly apparent that the disorder is, to variable degrees, accompanied by a broad spectrum of cognitive impairments.

These dysfunction may occur in the perceptual, thought, and/or linguistic processes of the individual diagnosed as schizophrenic.

The general clinical goal concerning the apparently disordered cognition characteristic of schizophrenia, has been the return of the schizophrenic's cognition to ordered functioning. To this purpose, in the 1950s, Bateson, Jackson, Haley, and Weakland proposed that schizophrenic communication is characterized, and induced, by double-bind, an apparently self-contradictory form of communication, but "coherent" inside familiar system.

Von Domarus (1944) and Arieti (1955) have theorized that schizophrenics reason by different rules of logic from others and that this tendency may explain their irrational thought and speech. In particular, Von Domarus believed that schizophrenic reasoning is dominated by the principle that two objects are identical when they share a common attribute. Arieti espoused a similar view, but believed inappropriate logic in schizophrenics was limited to subject material with emotional content. The Von Domarus hypothesis and Arieti's revision, were focussed on one feature of deductive reasoning: the principle of identity and one of its underlying process, the inclusion relationship. Because the syllogism, in Aristotelian logic, was considered to be the prototype of deduction, most experiments have relied on this type of paradigm.

Only a limited number of studies have been run to test Von Domarus and Arieti's hypotheses.

Nims (1950), Williams (1964), Wyatt (1965), and Jacobs (1969) reported predominantly non-significant differences between schizophrenics and controls on their syllogism tests, but Chapman and Chapman (1973) have suggested that their findings may be artifacts that resulted from the use of atypically intact schizophrenics. Gruber (1965) found that male schizophrenics produced more Von Domarus (similarity implies identity) errors when the shared predicate utilized a preferred-meaning response than when it did not, but was unable to replicate his findings with female samples.

Gottesman and Chapman (1969) compared the frequencies of Von Domarus and Non-Von Domarus syllogistic reasoning errors among schizophrenics and normals. Their schizophrenics performed at a lower level than their normals, but the two groups did not differ on the percentage of errors that coincided with the Von Domarus prediction, which suggest that

schizophrenic thinking may be characterized by reasoning errors of various sorts. However, their schizophrenics' mean IQ was 18 points lower than that of their controls, and it is possible that the poor performance of the former reflected general intellectual weakness rather than disease-specific deficit.

Watson, Wold and Kucala (1976) compared schizophrenics and psychiatric controls on four abstraction measures and a logic test after the groups had been matched closely for IQ. They found no difference on the abstract ability test, but their schizophrenics were inferior to controls on the logical reasoning test. Their results suggested that inability to abstract is not a disorder-specific schizophrenic deficit, but that inability to reason logically may be.

The results of a more recent work of Watson and Wold (1981) indicate that, inability to use syllogistic reasoning properly, is probably not the root cause of schizophrenic thought disorder, and added no credence to the already challenged view that the use of peculiar rules of logic accounts for schizophrenic thought disorder.

2. SCHIZOPHRENIA AND MATHEMATICAL LOGIC

In a natural and common way we can think the logic as the analysis of the rules of inference and the study of the notion of proof. The study of logic is interested to the syntactical form of the topics. The basic concept is if the truth of hypotheses imply the truth of conclusion. The deepest study of the mathematical proof is the aim the mathematical logic known as proof theory.

All avowedly logic-based theory assume that reasoning is a process of applying inference rules, and typically take their lead from proof-theoretical systems such as natural deduction.

Proof theory is a way of approaching validity, but it is not the only way, and recently we assist to a gradual decline of the logistic paradigm. The reason of new emergent point of view is based on the concept about the controversy over the relation between logic and reasoning. In this case the opinions lining up along a gradient ranging from the view that nothing could be less relevant to psychology than logic, on the one hand, to the view that logic is central not only to reasoning but to cognition at large, on the other. In this sense, alternatively non-logical frameworks to proof theory are Johnson-Laird at al.'s theory of mental models (1983, 1992, 1995), and Chater and Oaksford's probabilistic approach (1999, 2001).

3. MATHEMATICAL LOGIC

A proposition in the English language is a statement which can be said to be true or false. From proposition A, B we can form more complex propositions by using logical symbols or connectives:

1. negation: **not** ($\neg$);
2. conjunction: **and** ($\wedge$);
3. disjunction: **or** ($\vee$);
4. exclusive disjunction: **aut** or **xor** ($\oplus$) (i.e.: *Hic vincendum aut moriendum est.* Livio)

Let be □ a connective. We have the following property for 2), 3) and 4):

- idempotency [only for 2) and 3)]: A □ A = A;
- commutativity: A □ B = B □ A;
- associativity: (A □ B) □ C = A □ (B □ C).

For 1) we have that the negation of the negation of a proposition is equivalent to the proposition itself (i.e.: $\neg\neg A \equiv A$).

First/second De Morgan law:

The negation of the conjunction/disjunction of two propositions is equivalent to the disjunction/conjunction of the negations of the two proposition ones.

4. THE IMPLICATION

The logical symbol or connective $\rightarrow$ is called the implication sign. We can consider an interpretation of cause and effect for implication sign. The two propositions **a** and **b**, in **a** $\rightarrow$ **b**, are called *antecedent* and *consequent*.

We now consider an interpretation of propositional calculus and we define the operations $\wedge$, $\vee$, $\rightarrow$, $\neg$ by means of the following Table 1:

Table 1.

x	y	$x \wedge y$	$x \vee y$	$x \rightarrow y$	$\neg x$
0	0	0	0	1	1
0	1	0	1	1	1
1	0	0	1	0	0
1	1	1	1	1	0

Sometimes instead of 0, 1 the word *false* and *true* are used. Then the Table 2 will indicate rules of assigning truth values to the connectives used.

Table 2.

x	y	$x \wedge y$	$x \vee y$	$x \rightarrow y$	$\neg x$
false	false	false	false	true	true
false	true	false	true	true	true
true	false	false	true	false	false
true	true	true	true	true	false

Implication connective can assume the following four forms

$x \rightarrow y$ direct
$\neg x \rightarrow \neg y$ contrary
$y \rightarrow x$ inverse
$\neg y \rightarrow \neg x$ counternominal

Table 3.

x	y	$x \rightarrow y$	$\neg x \rightarrow \neg y$	$y \rightarrow x$	$\neg y \rightarrow \neg x$
false	false	true	true	true	true
false	true	true	false	false	true
true	false	false	true	true	false
true	true	true	true	true	true

The direct implication and counternominal implication are logically equivalent.

4.1 Rule of inference

The logical symbol $\Rightarrow$ is called the turnstile or yield sign and meaning a *sequents*. Sequents are forms of statements, theorem in which we may clearly distinguish conditions (hypotheses) and conclusion. We can, now, consider the following rules of inferences:

- Modus ponens or detachment: $(a \rightarrow b) \wedge a \Rightarrow b$
- Modus tollens: $(a \rightarrow b) \wedge \neg b \Rightarrow \neg a$
- Tertium non datur: $a \vee \neg a \Rightarrow t$ where t is always true
- Sillogism: $[(a \rightarrow b) \wedge (b \rightarrow c)] \Rightarrow (a \rightarrow c)$
- Contrapposition law: $(a \rightarrow b) \Rightarrow (\neg b \rightarrow \neg a)$
- Reductio ad absurdum: $(\neg a \rightarrow f) \Rightarrow a$ where f is always false
- Analysis of (two) possible case:
  $\{ [(\Gamma \wedge \Phi) \Rightarrow \Psi] \wedge [(\Gamma \wedge X) \Rightarrow \Psi] \wedge [\Gamma \Rightarrow (\Phi \vee X)] \} \Rightarrow (\Gamma \Rightarrow \Psi)$

In the brain, the left parietal regions (especially the left superior parietal lobule), are activated in both modus ponens and modus tollens problems (Goel, Buchel at al., 2000; Goel and Dolan, 2003).

4.2 Gödel's incompleteness theorem and schizophrenic "world"

In logic, the foundation of a logical system is a set of fundamental proposition or statements and rules to manipulate these fundamental propositions so that other proposition can be derived. A logical system is consistent when it is no possible to derive a proposition and its negation in the system self, or in another way we can say that a calculus is consistent if not all formulas that calculus are provable in it. Importantly, a proposition derived in a consistent logical system is presumed to be true (*i.e.* it corresponds to "the way things are"). Similarly, ordered thought may be termed consistent when it does not allow a deduced thought and its negation to both be reasoned from a particular set of premises. A deduced thought in an ordered system of thought that is consistent is considered true, meaning that it corresponds to reality. In logic, a logical system is complete when all truths can be represented by theorems of the system. This system is incomplete when there exists at least one truth, indicating an element of "the way things are", for which a corresponding theorem does not exist in the system. Applied to cognition, the term incomplete means that the logical system that can be used to characterise mature human thoughts is incapable of generating statements corresponding to all truths concerning those areas with which this thought is concerned. Further, this logical system characterising thought is incomplete in a particular way, namely trough a statement that comments on itself and, indeed, denies its own probability. It is Gödel who, in logic, provided the means for this self-referential possibility of a statement commenting on itself. This incompleteness is particularly intriguing because when a statement comments on itself, the notion of "the way things are" can be applied to itself.

In schizophrenic patients delusions and hallucinations are a needed element in a psychological theory incorporating an absolute view of world, a view in which the world is one way and exists independently of the experiencing individual. In more traditional terms, a delusion is generally considered a belief that does not correspond to "the way things are", and hallucinations is a perception that does not correspond to "the way things are". Delusions and hallucinations often have very sophisticated structures which can be called "well thought out" and yet do not correspond to what "normal" individuals maintain are "the way things are". For "normal" individuals, this lack of correspondence is considered a reflection of the schizophrenic's disordered cognition.

5. THE ROLE OF CONTEXT AND INFERENCE RULES IN PARANOID ORGANIZED SCHIZOPHRENIA

Logic is understood in the general sense of "the way cognitive elements are linked together".

Since logic was first developed to formalize rationality, it make sense that it would serve as a useful tools in modelling aberration of reason. If individuals infer conclusion from a set of premises by applying a pre-established category of rules of reasoning, then false conclusions may be arrived at by either starting from false premises or by invalid inferences.

A key feature of deductive arguments is that conclusions are contained within the premises and are independent of the content of the sentences. They can be evaluated for validity, a relationship between premises and conclusions involving the claim that the premises provide absolute grounds for accepting the conclusion. In inductive reasoning the conclusion reaches beyond the original set of premises, allowing for the possibility of crating new knowledge.

The vast majority of literature on schizophrenia and logic, address the possibility that the patients use of invalid inferences, beginning with Von Domarus' idea that patients with schizophrenia consistently use a specific fallacious inference (Von Domarus, 1944). More modern studies have tested patients' abilities to use standard logical inferences (Ho, 1974; Kemp et al., 1997; Watson and Wold, 1981), and the correlation between delusion thought and a peculiar style of reasoning in which patients "jump to conclusion" (Huk, Garety and Hemsley, 1988).

Little attention has been paid to the other possibility by which false conclusion may be reached: inappropriate choice of premises. This absence is all the more striking because modern empirical studies of normal cognition suggest a paradigm of reasoning, like mental models, that is radically at odds with that presupposed by standard tests of logic (Johnson-Laird, 1995).

The mental model theory assumes that reasoning is based on the truth conditions of a given statement. According to this account, a reasoner forms a mental model based on the premises of an argument, arrives at a putative conclusion by discovering a new, informative proposition that is true in the model, and then looks for counter-examples. In the event no such counterexamples are found, then the conclusion is stated as a valid consequence of the premises. This approach originally assumed (given the spatial nature of models) that reasoning is a right hemisphere activity.

The most obvious difference of the two models (proof theory and mental models) lies at the level of premises. Tests of deductive logic provide pieces

of information that are explicitly described as the material from which conclusions ought to be derived. In the real world, however, our premises are seldom laid out so neatly before us. Instead, a large portion of our mental work must go towards discriminating between relevant and irrelevant information, choosing that from which we will later derive conclusions. Since available premise-groups are usually incomplete, most "conclusions" are actually closer to be being *hypotheses*, over-inclusive sets that are then restricted by confrontation with new evidence.

Johnson-Laird's experiments suggest that the capacity for recognizing counter-examples to our provisionary models, and these models' subsequent revision, are just as critical in the formation of belief systems as the inferences that initially give rise to the models (Oakhill and Johnson-Laird, 1985). Since perhaps the most characteristic feature of delusions is not the strangeness of the conclusions reached, but of their perseverance in the face of systemic evidence to the contrary (Jones and Watson, 1997) one would expect the recognition and application of counter-examples to be a cognitive ability that is seriously impaired in patients with delusions. Responsible of delusions is the failure to sort premises: distinguishing relevant from irrelevant information and, in particular, the recognition and application of counter-examples.

It is possible that such a failure may be the result of a normal prioritization of neural resources during periods of emotional stress (fear, anxiety, mania, psychotic depression), inappropriately activated in patients. While it may be natural for healthy individuals to initially from false or partially false models, these models are normally revised in the face of contradictory evidence. In the presence of anxiety or fear, this self-correcting mechanism may be temporarily disabled in order to devote full mental resources to avoiding the cause of threat.

5.1 Clinical Case 1

One of our patients, in order to prove that she was the Virgin Mary to other people who did not believe her, walked for almost one hour up and down the carriageway's central line along a busy thoroughfare.

After being admitted to our ward, she denied being aware of her state of illness, claiming with conviction that she was the Virgin Mary, and that the inconfutable evidence was that she had not been run down by cars. This patient suffers from a schizoaffective disorder, and has psychotic symptoms (delusions of mystic type) during the stages in which her mood shows clear signs of turning high. During the intercritical periods she is well integrated in her everyday's world, yet showing a basic personality structure in which the narcissistic and the paranoid dimensions are entwined.

The increasing of mood level can modify the subjectivist probability of one (or more) events. How many times must be executed an action that is dangerous for our life because we can feel our selves supernatural? (One hour!; Run with the car at 120 M/h, and so on). The increasing of mood level can diminish the subjectivist probability of the awareness of the state of the illness. During the state of the illness, schizoaffective disorder with psychotics symptoms (mystic type delusions), the abduction reasoning schemes are transformed into deterministic schemes, so that lead to sure conclusions and no to probable conclusions.

Theorem: *I'm the Virgin Mary.*
Proof. It is need prove that: Whoever do dangerous actions and haven't injuries is supernatural. In an other form:

To do dangerous actions; Don't have injuries $\Rightarrow$ To be supernatural.

Walking for one hour with the risk to be run down lead to the sure of to be supernatural. The subjectivist probability to be not run down is transformed in the sure of to be supernatural. The reasoning scheme used is the induction scheme. The run of the first car haven't caused the run down, then the same for the second car, and so on, therefore the conclusion is that I'm supernatural. On this way we have the first truth:

- I'm supernatural
  (The abduction is used and the transition to the induction happen by the parameter regulator of the manic affective state, so we are out of the common sense)

Axiom: to do dangerous actions cause injuries or can lead to die. Is used the tertiun non datur excluding the possibility of to be undamaged (again out of the common sense).

Modus ponens:

- First premise: **if** I'm supernatural **then** I'm *not* run down
- Second premise: I'm supernatural
- Conclusion: I'm *not* run down

Scheme of the final reasoning

The scheme of the final reasoning is based on the following tautology $[(a \rightarrow b) \wedge \neg b] \rightarrow \neg a$, that lead to the modus tollens reasoning $(a \rightarrow b) \wedge \neg b \Rightarrow \neg a$ with

- First premise: $(a \rightarrow b)$
- Second premise: $\neg b$ and the principle of the tertium non datur so that is possible to have the following
- Conclusion: $\neg a$

And now
Let be the following propositions:
a ≡ I'm the Madonna
b ≡ I'm *not* run down

- First premise: **if** I'm *not* the Madonna **then** I'm run down
- Second premise: I'm *not* run down

using the modus tollens reasoning scheme we obtain the
- conclusion: *not* I'm *not* the Madonna ≡ I'm the Madonna. Q.E.D.

In this case is evident how fundamental affective states (emotions, feeling, moods) are continuously and inseparably linked to all cognitive functioning (or "thinking" and "logic" in a broad sense), and that affects have essential organizing and integrating effects on cognition (Ciompi, 1997).

5.2 Clinical Case 2

One of our patients, a university researcher suffering from paranoid schizophrenia, was admitted to our ward, because in receiving his wage, the bank-clerk asked him to sign the receipt. The patient objected that he could not sign the receipt, because this act could not coincide exactly with the very moment he was handed the money, but it would have been delayed, although for a few seconds only. This would have entailed either to sign before having been handed the sum, in which case the clerk could have been able to steal the money without the opportunity for him to prove that he had not taken it (having already signed), or to sign after having been given the money, in which case he could have taken it from the bank unduly, as the clerk would have not been able to prove that he had signed. When the clerk replied to the patient that the delay between signing the receipt and being given the money was absolutely insignificant, the patient reacted attacking him, and this caused the man to be taken to the psychiatric emergency department.

The problem of process synchronization arise from the need to share some resources in a system. This sharing requires coordination and cooperation to ensure correct operation.

Operating in a context of cooperating sequential process can be viewed as consisting of a linear sequence of actions executed for obtain a goal. In a multi-acting environment the operations may be activated by several users simultaneously. This will lead to inconsistencies if no synchronization conditions are observed. The effect of synchronization conditions will be that, depending on the state of the object, some operations will be delayed somewhere during the course of their execution, until an appropriate state of the object is reached by the effect of operations performed by other users. It is not always possible to specify the necessary synchronization in term of the

externally operation alone. The user of the object is not aware of, and need not be interested in, the details of the synchronization. It is the implementation's responsibility to maintaining consistency by observing suitable constraints. In a system the state transition can occur only if a primitive operation is initiated. Depending on the current state, a transition to another state may be allowed or forbidden. The synchronization conditions warrantee that no illegal state transitions can occur. A process may be locked for different reasons:

- It may be inside a wait-order operation.
- The process may wait for an answer to an order it has sent to another process. If the called process may send orders (not only answer) to the caller then the system get into a circular wait state. This is a possibility of deadlock.

If the answer to any order is computed in a finite time and if every process that is no blocked proceeds with no zero speed, the answer to every order will arrive in a finite interval of time, i.e. there is no deadlock. We can say that a pathologic state consist in a process that send an order and an answer consequent to the previous order to the process it self (i.e. clinic case 2). In this case we can consider as:

- shared resources the *receipt* and the *salary*
- process the *employee* and the *university researcher* (the *patient*)
- the jobs consist of sign the receipt (for the university researcher) and pay the salary to the university researcher (for the employee).

The parallelism of the two jobs is impossible, in abstract sense, but in the common sense for a delay of a little few of seconds we can assume that the two events are happened simultaneously. In other words the interval of time is considered an indivisible region of acting. The loose of the common sense and premises context sensitive, cause a pathological behaviour. The transition is in a state of deadlock with consequent broil, in this case the university researcher assail the employee. At this point the system is in a state of deadlock and, the external psychiatric quick intervention, have removed the deadlock by changing the university researcher in patient.

5.3 Clinical Case 3

One of our patients, who works as a parking lot attendant, began to develop the perception that one of her shoulders no longer belonged to her (showing it to me, and making me touch it, as if this had been absolutely evident). As an explanation, she told me that she had certainly been kidnapped by a criminal organization, that had drugged her and had replaced her shoulder, as a retaliation for the fact of having "understood" that some criminal organizations are involved with the international trade of human organs, but she could not tell me more than

this, being under the risk of a further revenge. As I remarked the absurdity of this story, the patient started showing an aggressive attitude towards me, saying that I was but another one who did not believe her, and opposing an absolute firmness to my criticism of her delusional state. (From a clinical point of view, this situation may be interpreted as a case of partial somatic depersonalization, on which a delusional t interpretation, or a delusional perception of somesthesic kind is developed).

In this case there are two elements that seems very strange to common sense: the wrong perception of non-self of her shoulder, and the explanation of this phenomenon in delusional terms, as the results of a transplantation occurred for her damage. All these phenomenon are both cause and effect one for the other, but the common element is the prevalent mental state of the patient, that is represented by the sensation to be damaged in her physic integrity.

We can say that the common sense is determined by some rules that aren't write but are accepted and observed by many peoples (hyper-objective rules). The common sense is also the context within it is possible to construct the reasoning, in other words the environment in which sort out the premises for deductions and/or inductions and/or abductions reasoning. The premises aren't only a collection: we establish, by an empirical hierarchical structure derived from the common sense itself, the importance of the premises that will be used. The knowledge from the external common sense (the reality), after an elaboration and some change caused by our perception, is transferred into our mind and become, sometime with some distortions out of the standard common sense, the our realty. We know the not all can be demonstrate and some things can became axioms, so that it is not necessary a demonstration. In this way we increase the number of axioms and can happen that we don't understand if a proposition is a theorem or an axiom, all become true. All is true! Therefore the persons that don't believe our statement have a strange behaviour. In this way we haven't need to demonstrate ours propositions because every proposition is a truth (an axiom). In other words we can say that our reality is out of the common sense.

6. TAUTOLOGY: THE LOOSE OF CERTAINTY IN DISORGANIZED SCHIZOPHRENIA

These clinical cases are not so characteristics like the former, and patient affected by disorganized schizophrenia show thoughts deeply disjointed until a "salad of words" in which is not more possible to estimate the principal elements of the speech, that become incoherent, not finalistic, with

associative relation not preserved. This kind of speech is without a shared sense, except that some fragments in relation with the history or prevalent mental state of the patient. The patients seem to resemble each other in this confused magma of mental disorganization.

6.1 Clinical Case 4

This is a fragment of disorganized speech spoken by a patient suffering from chronic schizophrenia (S. Piro, 1967):

"Death that does not come. Death is the food of the Gods, because the Gods do not love man, as man is finite. Man is the saddest among the humanized beasts, because his brain cannot see beyond the ray of a pomegranate, which yet has a thousand suns, and a thousand grains of health, that no human mind can conceive, and whose purpose is to restore the soul thirsty of knowledge and delight. This short note acts as a teaching for the women in childbed and for scholars of knowledge, to feed their life in the case of need. With a pomegranate one dines without water, and drinks without wine; and one can even stay without food for several days, by having together with the pomegranate a cup of vegetable soup that can be usefully drunk also without a spoon, provided that it is warmed by the light of an electric bulb, or by the lens of a levelling-rod or by the fire of a match. The way back to Rome. The chicory broth made from salty water helps the digestion. For this reason Caesar in Gaul before the battle with Vergingetorix poured on the ground a spoonful of Leporino's onion soup, to wish misfortune."

What mean premises? A truth? Or more simply a word? The answer is a word that is *citizen* of the dictionary. With the words it possible to create sentences syntactically and semantically rights. Therefore became possible a salad of phrases. May be necessary to use the reasoning schemes? No, there is another way for creating new reasoning schemes. We can jump sentence to sentence if the intersection from two different sentence is a word. But sometimes to go to a new phrase can require an approach based on a behaviour turned to the aim, that realise a kind of reversed chronological order: in this sequence the future or final result of an action precedes or anticipates the action itself. In this case we can named the phrase non-contextual sentence. For example: (*the way back to Rome*). In this case there isn't intersection both previous and next phrases. But the sentence is needed for introducing *Caesar*. A salad of phrases can be created choosing, randomly, the phrases with a word for intersection and/or sometimes using one non-contextual sentence.

The principle of random choose of the words in the phrases, may be dominated by a prevalent mental state that in this case is related to grandiose, immortality, and omnipotence themes (the Gods, Caesar, Rome). Every single fragment therefore, even if not respect standard inference rules,

may considered to be coherent with the others, if referred to a prevalent mental state.

In this case, the meddle term of conjunction, loss the cause-effect link present in tautology of transitive property of material implication (that is the basis of syllogistic reasoning). In fact, in the example above, there are others criteria that rule the terms of conjunction: the *assonance* (death ... death; gods ... gods; man ... man), the *metaphor* (the pomegranate like symbol of copiousness and nourishment), and the *analogy* (alimentation like food and alimentation like knowing) that are all coherent with the prevalent mental state, devoid of reference with context.

7. CONCLUSIONS

Schizophrenic patient is like someone that is constructing an artificial private language (in the manner of Chomsky theories). In organized paranoid subtype, inference rules are strongly correct, but the premises are closed related with his particular emotional states, without a shared sense (that assume merely auto-referential characteristics). The premises are processed without to be set in a context at every single situation.

In the *clinical case 1* there is a strong presence of the standard inference rules that are correctly used. In this case the reasoning scheme, in the first part of the proof, is made by induction and abduction, but the premises aren't rights because they are probabilistic (the probability of to be not run down for an interval of time). The induction always need of true premises. In the abduction a reasoning adopt, as starting point a collection of data D; the hypothesis H, if true, could explain D; no one else may explain better then H; hence the H is probably true (Minati, 1998). In our clinic case 1 the patient reach a certain conclusion: I'm supernatural.

In the *clinical case 2* the inference rules are rights, but the premise that define that some actions can't be broken off is not respected. In this way the critical region between the sign of the receive and the payment of the salary is broken off, and therefore the premise is broken.

In the *clinical case 3* aren't existing evident inference rules, but the mainframe is dominated by the prevalence of the altered perceptive state. The last one state involve an exasperated no critic use of premises that are non contextual to the situation. The extreme use of premises (exist the Mafia, exists the transplantation of organs, exists the menaces), involve that the use of the inference rules is quasi needless in demonstrating that the shoulder has been transplanted per wish of the Mafia and therefore this one conclusion assume the appearance of an axiom. The patient seems to ask (perhaps unconsciously): how many axiom I need for proving the theorems?

From Gödel's incompleteness theorem and Church's thesis we see that there is even no effective way of giving the axioms. The last one consideration can be used for an extreme increasing of the premises so that the patient have no need of proving the theorem of transplantation that in this way become an axiom.

In the case of disorganized schizophrenia (*clinical case 4*) seems to be accepted the existence of a principle, a theorem or a statement, that can be both acceptable and refutable. In this way is broken the non contradiction principle. The meddle term, in the syllogism inference rule, has characteristics of assonance, metaphor and analogy, coherent to a prevalent mental state and non context sensitive. In other words the meddle term is not used as in the standard inference rule derived from the tautology of the transitivity of the connective of implication (the syllogism reasoning).

Is it, now, correct that the paranoid organized schizophrenia and the disorganized schizophrenia are again under the same common "hat" of the schizophrenia, or instead, from the above discussion they are two deeply different troubles?

REFERENCES

Arieti, S., 1955, *Interpretation of Schizophrenia*, Brunner, New York.

Bateson, G., Jackson, D. D., Haley, J., and Weakland, J., 1956, Toward a theory of schizophrenia, *Behavioral Science* **1**:251-64.

Chapman, L. J., and Chapman, J. P., 1973, *Disorder Thought in Schizophrenia*, Appleton-Century-Crofts, New York.

Chater, N., and Oaksford, M., 1999, The probability heuristics model of syllogistic reasoning, *Cognitive Psychology* **38**:191-258.

Ciompi, L., 1997, The concept of affect logic: an integrative psycho-socio-biological approach to understanding and treatment of schizophrenia, *Psychiatry* **60**:158-170.

Goel, V., Buchel, C., Frith, C., and Dolan, R. J., 2000, Dissociation of mechanisms underlying syllogistic reasoning, *Neuroimage* **12**:504-514.

Goel, V., and Dolan, R. J., 2003, Explaining modulation of reasoning by belief, *Cognition* **87**:B11-22.

Gottesman, L., and Chapman, L. J., 1969, Syllogistic reasoning errors in schizophrenia, *Journal of Consulting Psychology* **24**:250-255.

Gruber, J., 1965, *The Von Domarus Principle in the Reasoning of Schizophrenics*, Unpublished doctoral dissertation, Southern Illinois University.

Ho, D. Y. F., 1974, Modern logic and schizophrenic thinking, *Genetic Psychology Monographs* **89**(1):145-165.

Huk, S. F., Garety, P. A., and Hemsley, D. R., 1988, Probabilistic judgements in deluded and non-deluded subjects, *Quarterly Journal of Experimental Psychology* **40A**:801-812.

Jacobs, M. R., 1969, *The effect of interpersonal content on the logical performance of schizophrenics*, Doctoral dissertation, Case Western Reserve University.

Johnson-Laird, P. N., 1983, *Mental Models: Toward a Cognitive Science of Language, Inference, and Consciousness*, Cambridge University Press, Cambridge.

Johnson-Laird, P. N., 1995, Mental models, deductive reasoning, and the brain, in: *The Cognitive Neuroscience*, M. S. Gazzaniga, ed., MIT Press, Cambridge.

Johnson-Laird, P. N., Byrne, R. M. J., and Schaeken, W., 1992, Propositional reasoning by model, *Psychological Review* **99**:418-439.

Jones, E., and Watson, J. P., 1997, Delusion, the overvalued idea and religious beliefs: a comparative analysis of their characteristics, *British Journal of Psychiatry* **170**:381-386.

Kemp, R., Chua, S., McKenna, P., and David, A., 1997, Reasoning and delusions, *British Journal of Psychiatry* **170**:398-405.

Minati, G., 1998, *Sistemica*, Apogeo, Milano.

Nims, J. P., 1959, *Logical Reasoning in Schizophrenia: The Von Domarus Principle*, Unpublished doctoral dissertation, University of Southern California.

Oakhill, J. V., and Johnson-Laird, P. N., 1985, Rationality, memory and the search for counter-examples, *Cognition* **29**:79-94.

Oaksford, M., and Chater, N., 2001, The probabilistic approach to human reasoning, *Trends Cognitive Science*, **1;5(8)**:349-357.

Piro, S., 1967, *Il Linguaggio Schizofrenico*, Feltrinelli, Milano.

Von Domarus, E., 1944, The specific laws of logic in schizophrenia, in: *Language and Thought in Schizophrenia*, J. S. Kasanin, ed., University of California Press, Berkeley.

Watson, C. G., and Wold, J., 1981, Logical reasoning deficits in schizophrenia and brain damage, *Journal of Clinical Psychology* **3**(3):466-471.

Watson, C. G., Wold, J., and Kucala, T., 1976, A comparison of abstractive and nonabstractive deficits in schizophrenics and psychiatric controls, *Journal of Nervous and Mental Disease* **163**:193-199.

Williams, E. B., 1964, Deductive reasoning in schizophrenia, *Journal of Abnormal and Social Psychology* **69**:47-61.

Wyatt, L. D., 1965, *The Significance of Emotional Content in the Logical Reasoning Ability of Schizophrenics*, Unpublished doctoral dissertation, Purdue University.

THE "HOPE CAPACITY" IN THE CARE PROCESS AND THE PATIENT-PHYSICIAN RELATIONSHIP

Alberto Ricciuti
AIRS - Associazione Italiana per la Ricerca sui Sistemi, http://www.airs.it
Attivecomeprima-Onlus (Breast Cancer Association), Via Livigno, 3 – 20158 Milano, Italy
http://www.attivecomeprima.org
email: alberto.ricciuti@fastwebnet.it

Abstract: Especially in serious pathologies, in which there is a real risk of dying, the capacity of the sick person of keeping alive the will of living and participating actively in the care process, is intimately linked to what Fornari calls "hope capacity". Such capacity is often reduced and inhibited by the kind of arguments developed in the patient-physician communication due to a misunderstanding or a wrong use of the concept of "probability". In the context of the actual biomedical model, inspired on a narrow conception of the living beings, we currently use, in the patient-physician communication, the statistical and probabilistic evaluations referred to clinical/epidemiological data as predictive of the possible evolution of the pathology in the single person. When that happens – for a misunderstanding of the concept of "probability" or a semantic one – the "hope capacity" of the sick person fades away until it definitely disappears. This work shows how, in a systemic conception of health problems – where the barycentre of our attention is shifted from the illness to the sick person – new and fertile spaces for the discussion in the patient-physician communication about his/her possible futures can be opened, referring on one hand to the logistic concept of probability developed in the XX century by Wittgenstein and Keynes and on the concept of subjective probability developed more recently by De Finetti, on the other hand to the objective interpretation of the probability theory proposed by Popper, which defines probability as "propensity", that is as a physical reality analogous to forces or forces field.

Key words: hope; patient-physician communication; probability; systemics.

1. INTRODUCTION

Every physician who lives every day close to sick people knows for experience that in the moments of deepest crisis for a person, when it's extremely difficult even to imagine a future, the possibility to see the "door of hope" always open is the only thing (besides the physician's competence) that can give to the sick person the strength to take therapies on – sometimes very heavy, as the oncologic ones – and that can keep the patient-physician relationship always lit up. This hope is not necessarily the hope to recover – since in some situations it could be seriously difficult- but the hope for the sick person to live his/her life at his/her best. More than the fear of death, what really scares the person it's the fear of suffering, of living painfully. This pain is not only physical – that medicine can quite completely defeat today – but also moral, linked with the absence of hope. A woman told me recently: "Mentioning those statistical results when I was in hospital, they killed my hope ... and now it's more difficult for me to do everything ... even to accept a cure".
One of the most important recommendations for a physician in order to have a correct behaviour – according to a certain orthodoxy– is not to give false hopes to the sick person, because it sounds like a prediction of his/her illness' outcome.

But hope can't be false: it can be or not be.

Perhaps who talks about "false hope" means the hope that an event considered totally improbable – by our scientific knowledge and statistical results – will happen. But anyway, the cases that evolve in a radically different way from our expectations exist among the patients of every physician, and are extremely numerous in literature (Simonton, Matthews-Simonton and Creighton, 1978; Siegel, 1990; Hirshberg and Barasch, 1995).

But the "hope capacity" depends on what? Is it linked somehow with a personal Faith in a Mystery that helps us? Or is it also our reaction to a reasonable hint of possibility of living, that we can glimpse in the depths of our soul? How much our capacity to light on a "hope process" is linked with our *cultural imprinting*? Is it possible to imagine that, using a different *paradigm* (as Kuhn means, 1962), it's possible to individuate new reasons for lighting on the "hope capacity" (both in the physician and in the patient) in order to make the quality of the cure better (and perhaps its effects ...)?

The Cartesian paradigm that informs our cultural imprinting places every consideration about the theme of hope in the sphere of the *res cogitans*, domain of the human sciences, in opposition to the *res extensa*, domain of natural sciences. Therefore, the hope could live and act in a psycho-affective dimension that medicine has to take into account, perhaps for its traditional

spirit of charity. But the *principle of separation* of the Cartesian paradigm retrogrades the hope as it was a sort of rhetoric of feelings without any real contact with our concrete being. Our cultural paradigm, that separates the subject from the object, the sick person from the illness, separates also the hope from the corporeity giving them to two different spheres of phenomena completely separated the one from the other, one psychic, the other physic. Moreover, due to the *reduction* of the human to the natural in the Cartesian paradigm – also called for this reason "paradigm of simplification" (Morin, 1982) -, the hope, when it is in contrast with the expectation of the statistic results of the scientific knowledge, is degraded to "false hope", and who supports it (as the physician for example) is seen as someone whose behaviour is ethically incorrect.

The process of the "paradigm of simplification" that reduce the human to the biologic separates – for epistemological statute– the person from his/her life context. Both from the outer sphere – since it doesn't consider the relationships of the person with his/her social and affective context – and from the inner sphere, that is the representation of the world around him/her and of him/herself, that helps the person to decide his/her actions and gives them a "sense". According to the cognitive stereotype of this cultural paradigm, even the illness is separated from the body, that is its "container". Therefore, who is "affected" by the illness has the moral duty to "fight against" it, and during this period of "fight", he/she is virtually (and sometimes physically ...) excluded from the social context of "normal" people, since he/she is "unproductive". Consequently this, for the person and his/her usual way of "perceiving the illness", leads to a precise feeling of a regression in the social status. The person tends to protect him/herself from this regression – more or less consciously – with a behaviour that denies the illness. For example, the person tends to hide the illness with the colleagues or with other people more or less close to him/her. This contributes to increase his/her loneliness, and to create an inner feeling much more unfavourable to the "hope capacity".

With his brilliant and rigorous methodology of psychoanalytic research, Franco Fornari – particularly sensitive to these arguments because of his degree in medicine - affirms that the "hope capacity" is an *innate* characteristic of the human being (Fornari, 1985). It's a resource, a sort of survival hint that lights up in the moments of deepest crisis and bewilderment of a person, helping the person to re-design his/her life, besides the physical condition identified by the medicine as "illness".

Starting from this brilliant and fertile theory by Fornari, our purpose is to lead medicine to begin a reflection on the issue of the "hope capacity" , taking into account its structural value in order to benefit the person, and

consequently taking the human being again in the middle of the physician's activity. In other words, we have to set up another medical knowledge around the concept of the "hope capacity"; this new approach should consider the "hope capacity" a precious project resource for the person and should light it on as "fertilizer" of the therapeutic project, to evolve it in *care project*, that is addressed to the person in his/her entirety.

2. PROBABILITY AND HOPE

Even admitting for many people the existence of a "hope capacity" strictly linked to a religious faith, everybody tends to hope much more in the events he/she considers more probable. It is exactly using this concept incorrectly that medicine kills sick people's hope.

The concept of probability is used in the patient-physician communication at least for two kinds of argumentations: the first – typically technical – in order to motivate to the patient the decisions taken and the evaluations concerning the definition of the therapeutic program and the timing of the clinical controls. The second one – typically dialogic – in order to communicate with the patient about the prediction of the evolution of his/her pathology: for example predictions of the possibility to recover, of eventual recurrences, of the unfavourable evolution of his/her pathology within a certain period of time.

Obviously in different kinds of argumentations the word "probability" has very different meanings, leading to an ambiguity within the communication between the sick person and the person who takes care of him/her. In fact - as Mario Ageno well describes (1987) – "usually it happens that, during the conversation, we shift from a concept to another, more or less consciously". This leads to a serious confusion and misunderstanding, that can blow out the flame of a hope that the sick person hardly lighted on, and also the wish to live.

The *statistical* concept of probability used in the patient-physician communication derives from a mathematical theory formulated within the "paradigm of simplification", and the ambiguities it introduces in the patient-physician communication are due to two reasons:

1. it refers to the illness, and not to the ill person. Therefore it introduces a gap in the patient-physician communication, unless its meaning is absolutely clarified.
2. it consists of a datum referred to a combination of events (e.g. the percentage of survival in 5 years' time within a *group of patients* affected by lung cancer) as useful to predict a single event (the estimated survival in 5 years' time for *that single patient*).

Since these misunderstandings are very frequent within the patient-physician communication, I cite (Ricciuti, 1998), as paradigmatic, what a patient told me after her operation for an intestine cancer (in its initial stage, therefore with a very optimistic prognosis): "...then when I asked if I could hope, they told me: "we will know it in three years' time". Why everybody told me I was lucky, and then gave me three years of life?".

It is clear, for the professionals in the field, that the person who told her that referred to statistical-probabilistic evaluations, that means to data taken from the scientific literature about the evolution of that illness in the totality of examined patients. But that patient, who evaluates the situation from *her* point of view, taking into account *her* whole personal history, gives to that sentence a meaning of a three-years-of-life sentence. Or at least she looks at those three years as a period of intolerable and uncertain wait, that can leads to the risk of making this person living as a "dead person", while perhaps she will never die for that illness.

The semantic misunderstanding - much less banal than it appears at first sight – consists in the fact that in our common language, when we use the word probability we refer to something that perhaps *will happen*, that means something that has a certain possibility to happen in our uncertain future. On the contrary, when we talk about probability within the mathematic theory – that is what we use in medicine for technical evaluations and in the communication with the sick person -, we refer to something that is *already happened*, since in this case the term probability is defined as the ratio between the number of favourable events, and the number of all the events we observed. The problem emerges when we make predictions of life or death for *that single patient*, using, as predictive for *his/her* future, data describing only the number of lives and deaths within a sample of patient observed in a statistical research. Consequently, it is completely senseless to use those statistical data to make a prediction of what will happen in the future of *that single patient*. For that single patient the risk of dying is 0% or 100%, and not 30% or 60% ...

Unfortunately, this kind of misunderstanding is really frequent in the common language and mentality, even among physicians. This statistical datum, however, is expression of a valuable knowledge within the field of the biological and epidemiological research; therefore it must take part of that inheritance of data and technological-scientific knowledge the physician, as professional in the field, uses in order to make better technical evaluations, as the kind of therapy to propose to the patient and the follow-up planning. On the contrary, the use of the statistical concept of probability in the patient-physician communication in order to talk about *his/her*

possibility of life or death is completely illogical and, sometimes, very harmful.

In order to go over this kind of problems, trying to assign a degree of possibility to a single event, there are several models giving different interpretations of probability. The more fertile and useful for our purpose draw a path starting from the *logistic concept* of probability – pointed out by G. Boole (1854) and F. H. Bradley (1883) and than recovered by J. M. Keynes (1921-1962) – and ending in the concept of *subjective probability* pointed out by Ramsey (1931) and developed to its end by De Finetti (1970).

Avoiding specific theoretical details, the guiding thread linking these Authors' studies is a concept of probability as the *degree of rational belief* on a certain hypothesis. In Keynes' view, the probability that an event will happen is a logic relation between propositions, and it depends on our whole knowledge when we formulate the question.

In our case, the question to be answered, in order to help the patient to light on his/her hope again, is not "Which is the probability that this kind of cancer will present again in a three years' time?", but is, for example: "Which is the probability that my sentence "in the next three years I will be healthy" will be true?".

Therefore, it is a question concerning no more the event "illness", but our considerations about that event, that means something concerning, finally, the ego as a subject.

In other words, this approach permits to shift our attention to the sick person with his/her personal history and experiences, habits, behaviours, expectations, plans and affections. So the sick person can perceive him/herself non more like an anonymous biologic device, being a prey to a breakdown that condemns him/her to the uncertainty and anxiety about a passive and intolerable wait, but like a subject responsible for him/herself and his/her choices, capable to activate his/her resources and reorganize his/her hope in order to make a new *plan* of life. Moreover, a plan is a dynamic process evolving within a certain period of time and the probability of realizing it, as Finetti says, depends on the *degree of trust* expressed by the person to the reachness of his/her objective. This *degree of trust* – that is function of the available evidences – can be kept only if the subject is willing to modify it in relation to the growing information and coherently with the whole system of his/her expectations.

Therefore, in order to help the patient to answer that question, the physician has not only to take into account technical-scientific elements concerning the illness (among which there are also the statistical data referred to that illness), but also to orient the attention on the patient, looking for "important" information concerning him/her life very closely, since

"concretely, the probability in this sense is the degree of rational belief on a certain hypothesis and we can give to the same proposition a higher degree of evidence or probability, according to the related knowledge we have" (Ageno, 1987).

Surely, it is particularly difficult to quantify this probability, but it is not impossible. But are we sure that this can add efficacy to our argumentations within the dialogue with the sick person about *his/her* possibility to manage? On the contrary, isn't it true that we have a great resistance to drop, even if just for one second, the "paradigm of simplification" and the reduction of the human to the biologic, "theoretically" prescribed by that paradigm, because we are afraid of being technically disarmed and professionally more defenceless or less credible?

Therefore, the first step to leave this ambiguity that mortgage and stifle the sick person's hope capacity is to be completely aware of it. The second step is to begin by considering the possibility to use, with appropriate flexibility, different concepts of probability and methods of evaluation according to our argumentations within different fields of knowledge and different ups and downs of life. In other words, we have to clarify the distinction between different contexts of the medical/epidemiological research (where the statistical concept of probability is surely a useful instrument) and the clinical contest relating to the single patient (where the concept of probability can be used with different meanings).

3. THE REASONS FOR HOPE

Our considerations lead us to conclude that the hope, at least when it can be easily lighted on and nourished by the events we consider more probable, is linked to our *knowledge*, that means the paradigms that form our thoughts. If our *knowledge* concerns the elaboration of the information we consider "important", and if the evaluation of the probability depends on our knowledge, we can easily comprehend how much our evaluation about the evolution of the clinical history of a sick person can change according to the kind and the complexity of the our argumentations about the evaluation of the "case".

But this is just a little step. We must recognise that, if we remained attached to the patterns of thought of the Cartesian paradigm, anyhow we talk about hope, this will be surely restricted into the domain of the *res cogitans*, that has lost its continuity with the biological processes that belong to the *res extensa* domain.

On the contrary, if we take into account the larger cultural horizon of the "paradigm of complexity", considering ourselves as "autopoietic" beings

(Maturana and Varela, 1980) and considering life – whatever is the organizational stage – as a "cognitive process" (Maturana, 1980), the continuity between *res cogitans* and *res extensa* is not only built, but also it become a one-dimensional simultaneity in which the human being consists, giving him/her again the status of person.

It is difficult not to be seduced by the charm of this way of argumentation about ourselves, but it is not just a pure aesthetic pleasure. We must recognize that the *systemic thought* is now the most fertile and advanced cultural approach which offers the possibility to theoretically reunify the two terms of the Cartesian dualism and – linking again the biologic with the biographic - to give back to our person his/her unitary structure.

The "hope capacity", in this conceptual horizon, is an *emergent property*, that means a property that gives to the human being his/her specific consistence and defines his/her degree of complexity. This capacity, besides the dichotomy mind-body, has to be considered a fundamental part of the whole combination of the autoregulative capacities of the biological system, that are those autopoietic processes maintaining unvaried its organizational scheme and guaranteeing its autonomy, unity and individuality.

Hoping that an event will happen because it is probable means *recruiting* that event among those we evaluated possible *for us*. That means to introduce in our system a piece of information that has the role of an autoregolative dynamic contributing to address and support our autopoietic processes, that are the fundamentals of our biological organization.

It is extremely significant that Karl Popper, maybe the most important philosopher of science of the last century, whose work gives the soundest theoretical justification to the modern empirical method, attended to these topics until he proposed an interpretation of probability as *propensity*. While illustrious scientists as Heisenberg and Einstein – as Popper underlined in a lecture given the 24 August 1988 to the World Congress of Philosophy in Brigthon (Popper, 1990) – said the probability theory was mainly due to *lack of knowledge*, therefore regarding the *subjective* status of our mind, Popper affirms that "an objective interpretation of the probability theory" must comprehend also the whole combination of different possibilities with different "weights", that have *more than one single statistical probability* to happen, as the famous example of loaded dice. The different "weight" of these different probabilities determines a *propensity* to realize that makes the different probabilities non-equiprobable and strictly *depending on the context*. Therefore, the propensity – and this is the fantastic assumption made by Popper – is not a sort of abstract concept, but it has exactly the consistence of a physical reality, as what we call in physics *force* or *forces field*. Probability can be considered as "a vector in the *space of possibilities*"

(Popper, 1957). Moreover, the introduction of the concept of propensity means to generalize and extend exactly the concept of force.

As noted by Abbagnano (1971), "this interpretation tends to reduce the distance between the two basic concepts of probability", the statistical one which considers classes of events, and the logic-subjective one which considers single events.

As Popper (1990) pointed out, "in our concrete and changeable world, the situation and, with it, the possibilities, so the propensities, continuously change. They surely change if we (or any other organism) *prefer* a possibility rather than another one; or if we *discover* a possibility where we didn't see anything before. It is our comprehension of the world that modifies the conditions of this changing world, expectancies, dreams, fantasies, hypothesis and theories".

It is clear that the "extension" of this way of thinking to the possible futures of hour horizon is really different from the narrow tunnel where often some estimated predictions about *our* future conduct us, predictions that come from statistical-probabilistic evaluations referred to someone else's illness, and not to ourselves. At most, these evaluations represent a propensity, among many others existing in our life, that we can influence when – as Popper said- we *prefer* a possibility rather than another one, or when we *discover* a possibility where we didn't see anything before.

4. TOWARDS A BIOLOGY OF HOPE

"The concrete is not a step toward something else:
it is how we arrive and where we are"
(Francisco J. Varela, 1992)

When we are willing to bet than an event we desire will really happen, it means that we succeeded in finding out at least one reason, within our uncertain future, for thinking that this event has more than one possibility to happen.

When we are aware that we passed from the *desirable* domain to the *possible* domain, it means that we introduced in our system a piece of information consisting in the lighting on of a psycho-neuro-immuno-endocrinous process that can modify our biology. In the context of the *paradigm of complexity*, it doesn't make sense asking ourselves if an activity of thought can modify *in the concrete reality* a biological process. First of all, because the activity of thought is a biological process in itself, and it is absolutely linked with the biological processes leading to it and defining ourselves as autopoietic system. Secondly, because there isn't any *concrete*

reality besides that cognitive process defining life. The most important question is not *if* what we want will be realized, but *how*: that means asking ourselves which are the processes of which the *propensity*, that guides us toward the realization of one of our possible future, consists.

We can't face now the complexity of this problem, but we can take some aspects into consideration, that can be useful for our purpose. We already know that the network of relations between processes concerning our psycho-neuro-immuno-endocrinous system's activation is exactly the same whether due to a cognitive or a non-cognitive stimulus (Blalock, 1992) (Figure 1) and that, therefore, there isn't any difference in the electroencephalographic response to real or imaginary stimulus (Weinberg, Grey Walter and Crow, 1970).

On the other hand, the richness of interpretative and explicative hypotheses from the most recent researches go over these concepts and show us very interesting sceneries. The description of the brain as a computer, that is still very common, is definitely misleading and is simply in contrast with its nature. "Cognitive life is not a continuous flow, but it is punctuated with behavioural structures arising and dissolving in very small time fractions" (Varela, 1992), coherently with the *organizational network* of our nervous system, and, more generally, our whole autopoietic system. In human beings about 10^{11} interneurons connect about 10^{6} moto-neurons with 10^{7} sensorial neurons distributed on the body's receptive surfaces. The proportion is of 100.000:1:10 interneurons mediating the matching of the sensorial and motorial surfaces.

The *motorial* is resonant with the *sensorial* in rapid and spontaneous appearances and disappearances of emerging configurations of neuronal connections and activities, created by a background of incoherent and chaotic activity that, with rapid fluctuations, proceeds until it reaches a global electric configuration concerning the whole cerebral cortex, and then it dissolves again in the chaotic background (Freeman, 1985).

As Varela writes (1992), "these very rapid dynamics concern all the sub-networks that give the stimulus to the quickness of action in the very following moment, and they concern not only the sensorial interpretation and motorial action, but also the whole range of cognitive expectations and emotional tones that are fundamental in forging a micro-world; on their basis a neuronal group (a cognitive sub-network) prevails and becomes *the behavioural modality of the next cognitive moment, a micro-world*".

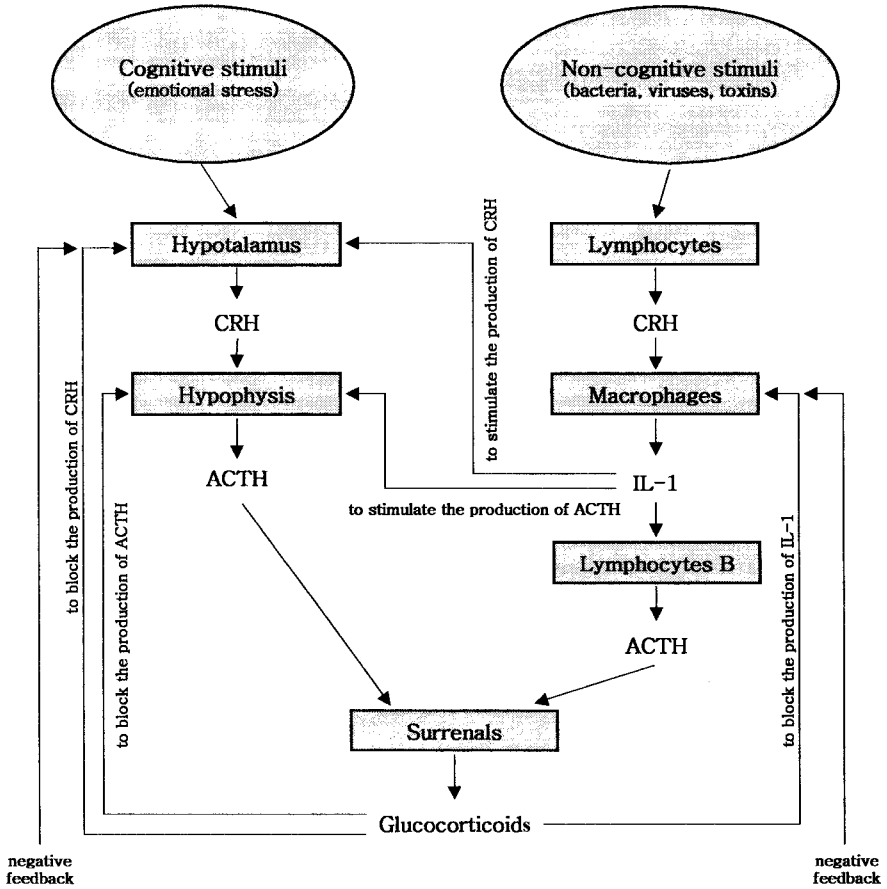


Figure 1.

In the same way, the biological processes illustrated in Figure1 can be described as *configurations of biological processes' networks* that identify as many *micro-identities*. It is from their continuously fluctuations that our life derives every instant.

5. CONCLUSIONS

Therefore, if we take into account the larger cultural horizon of the "paradigm of complexity", the possible futures for every person are absolutely unpredictable, since one of the most important characteristic of complex systems, as living beings as we are, is the non-linearity of their responses to different stimuli. We regenerate ourselves continuously using the context around us and we have the extraordinary capacity to modify the rules of the game and to evaluate their effects, to plan, to program and activate new tendencies to the action, new behaviours that can orient our future – sometimes in a very different way from the "statistical-probabilistic" expectations.

Our scientific method and the probability theory are surely a compass that helps us to sail the sea of our possibilities about our future, but we must take into account – especially in the patient-physician communication- that a compass is sensitive only to the field of possibilities for which it has been designed and programmed. In other words, it can't "weigh" our virtues and weak points, our fears and our "hope capacity", that means the *propensities* that our mind and heart can produce to change the path of our life and that, as Popper says, have exactly the consistence of forces inside and around us. These propensities make our possible futures absolutely non-equiprobable. It is not easy to choose which future. First of all, because we need at least one good reason to think that we can manage and light on a *hope process* having effectively a positive biologic relapse. Secondly, because often, with no awareness of it, we involve dynamics (ways of thinking, convictions, behaviours, so propensities ...) that work against us.

It is exactly in this moment that an *allied* physician can be precious: he/she can help us to individuate from time to time the most reasonable possibility to follow, helping us to walk toward it. Our strength will be his/her capacity to show us that we can manage our difficult situation. Concretely, this means "keep the door of hope always open". Nobody has the right to close that door, in the name of a theory – even if precious and effective when correctly used as a work-tool – that always represents an incomplete way of describing ourselves.

Someone says, as we wrote before, that "we can't give false hopes to the sick person"... But hope can't be "false"; it can be realized or not. Anyway, in very difficult moments of his/her life, the human being has always had the need to believe that there is, somewhere, a place where something really extraordinary for him/her can realize, also because the future – as Popper writes- is *objectively not fixed* and *objectively* open. Our inner hope capacity, even with its taste and sacred charm of a divine gift, can find its good reasons for lighting up with our complete awareness of living in the field of

possible ... and can find a support in a physician who uses the scientific theories, but who is not used by them. Because - as Jaspers says (1986) - "the physician has not only the responsibility of the precision of his/her statements, but also the responsibility of their effect on the patient".

REFERENCES

Abbagnano, N., 1971, *Dizionario di Filosofia*, Unione Tipografico-Editrice Torinese, Torino

Ageno, M., 1987, *La Biofisica*, Laterza, Bari.

Blalock, J. E., 1992, *Neuroimmunoendocrinology*, Kargen, Basilea.

Boole, G., 1854, *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probability*, London.

Bradley, F. H., 1883, *The Principles of Logic.*

De Finetti, B., 1970, *Teoria della Probabilità. Sintesi Introduttiva e Appendice Critica*, Einaudi, Torino.

Fornari, F., 1985, *Affetti e Cancro*, Raffaello Cortina, Milano.

Freeman, W., and Skarda, Sh., 1985, Spatial EEG patterns, nonlinear dynamics and perception: the neo-sherringtonian view, *Brain Research Review* **10**.

Hirshberg, C., and Barasch, M. I., 1995, *Remarkable Recovery*, Riverhead Books, New York, (it. tr.: *Guarigioni Straordinarie*, Mondadori, Milano 1995).

Jaspers, K., 1986, *Der Arzt im Technischen Zaitalter*, R. Piper, Monaco, (it. tr.: *Il Medico nell'Età della Tecnica*, Raffaello Cortina, Milano, 1991).

Keynes, J. M., 1962, *A Treatise on Probability*, London, 1921, New-York.

Kuhn, T. S., 1962, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, (it. tr.: *La struttura delle rivoluzioni scientifiche*, Einaudi, Torino, 1969).

Maturana, H. R., 1980, *Biology of Cognition, in Autopoiesis and Cognition. The Realization of the Living*, D. Reidel Publishing Company, Dordrecht, Holland, (it. tr.: Biologia della cognizione, in: *Autopoiesi e Cognizione*, Marsilio, Venezia, 1985).

Maturana, H. R., and Varela, F. J., 1980, Autopoiesis. The Organization of living, in *Autopoiesis and Cognition. The Realization of the Living*, D. Reidel Publishing Company, Dordrecht, Holland, (it. tr.: Autopoiesi. L'organizzazione del vivente, in: *Autopoiesi e Cognizione*, Marsilio, Venezia, 1985).

Morin, E., 1982, *Science avec Conscience*, Fayard, Paris, (it. tr.: *Scienza con Coscienza*, Franco Angeli, Milano, 1984).

Popper, K. R., 1957, The propensity interpretation of the calculus of probability, and the quantum theory, in: *Observation and Interpretation, A Symposium of Philosophers and Physicists*, Körner.

Popper, K. R., 1990, *A World of Propensities*, Thoemmes Antiquarium Books, Bristol, (it. tr.: *Un Universo di Propensioni*, Vallecchi, Firenze, 1991).

Ramsey, F. P., 1931, Truth and probability, in: *The Foundation of Mathematics and Other Logical Essays*, F. P. Ramsey, ed., London, (it. tr.: *I Fondamenti della Matematica e altri Scritti*, Feltrinelli, Milano, 1964).

Ricciuti, A., 1998, Le speranze, le preoccupazioni e la relazione terapeutica del medico personale, in: *...e poi cambia la vita*, Attivecomeprima Onlus, ed., Franco Angeli, Milano.

Siegel, B. S., 1990, *Love, Medicine and Miracles: Lessons Learned about Self-Healing from a Surgeon's Experience with Exceptional Patients*, Quill, (it. tr.: *Amore, Medicina e Miracoli*, Sperling, Milano, 1990).

Simonton, O. C., Matthews-Simonton, S., and Creighton, J. L., 1978, *Getting Well Again: a Step-by-step, Self-help Guide to Overcoming Cancer for Patients and their Families*, Bantam Books, Toronto, (it. tr.: *Stare bene nuovamente*, Edizioni Nord-Ovest, Milano, 1981).

Varela, F. J., 1992, *Un Know-how per l'Etica*, Laterza, Bari.

Weinberg, H., Grey Walter W., and Crow H. J., 1970, Intracerebral events in humans related to real and imaginary stimuli, *Electroenchephal. Clin. Neurophysiol.* **29**.

Wittgenstein, L., 1922, *Tractatus Logico-Philosophicus*, Routledge and Kegan Paul, London, (orig. german ed. 1921; it. tr: Bocca, Milano-Roma, 1954).

PUNTONET 2003. A MULTIDISCIPLINARY AND SYSTEMIC APPROACH IN TRAINING DISABLED PEOPLE WITHIN THE EXPERIENCE OF VILLA S. IGNAZIO

Dario Fortin, Viola Durini and Marianna Nardon
Villa S. Ignazio, Cooperativa di Solidarietà Sociale, Via alle Laste 22, 38100 Trento
e-mail: vsi@vsi.it

Abstract: In this paper we will present Puntonet 2003, a about 900 hours' course intended for disabled people and co-financed by the European Social Fund. Our approach in developing the course structure itself was focused in taking into account both the future employment of the participants and the personal and social reinforcement. The organizing and teaching team is itself multidisciplinary, combining engineers and scientific professionals and professionals with social, educational and psychological skills. The course Puntonet 2003 aims the inclusion of disabled people in the Information Society, improving professional skills but also reinforcing knowledge and integration in the social network.

Key words: information society; disability; social integration; employment.

1. INTRODUCTION

In this paper we will present the course Puntonet, designed and realized within the experience of the cooperative enterprise Villa S. Ignazio, engaged for 30 years in preventing social exclusion and in promoting different training activities. The activities of Villa S. Ignazio respond to specific social needs, raising from the local community.

As noted in the paper "*E-Inclusion: The Information Society's potential for social inclusion in Europe*"(2001), the more the Information Society advances, the more social and economic opportunities depend on ICT usage.

Digital exclusion increasingly becomes a barrier for individuals, also in our territory, not only as far as employment is concerned, but also for the social inclusion. The core intention of the Puntonet course is then to equip disabled people with ITC skills and encourage their participation in the Information Society. In this sense, the epistemological approach of the Puntonet course is systemic, because it emphasizes cooperation and exchange among the different systems, in that the individual is involved.

The Puntonet project would like to contribute to the E-inclusion purposes, promoted by the European Union through the Europe Action. The main purpose of the Puntonet course is in fact to provide ICT literacy skills to people with physical, psychological or sensory disabilities, so that they can do simple office works and manage job research strategies using the ICT. Information and Communication Technologies have the potential to overcome traditional barriers to mobility and geographic distances, and to distribute more equally knowledge resources. They can generate new services and networks that support and encourage disabled people in a flexible and pro-active way, also offering new job opportunities.

On the other hand, new risks of digital exclusion need to be prevented. In an economy increasingly dominated by the usage of ICT across all sectors, Internet access and digital literacy are a must for maintaining employability and adaptability, and for taking economic and social advantage of on line content and services.

The Puntonet project takes into account the main concrete measures, that the ESDIS (High Level Group "*Employment and Social Dimension of the Information Society*") proposes in order to fight against digital exclusion. First, to realize the Information Society's potential for disadvantaged people working for new ICT job opportunities; then to remove barriers by raising awareness of ICT opportunities and promoting digital literacy; last but not least, to encourage the networking of Social Partners and civil society organizations at a local level, supporting innovative forms of stakeholder partnerships.

The experience of the past Puntonet courses shows in fact how important ICT skills are, not only for professional training and employment, but also, from a systemic point of view, as a mean to get Information and services, to search for a job and to manage formal and informal relationships, thus supporting personal and social reinforcement.

Moreover, the Italian research "*Information Society and disabled people*" underlines that problems for disabled people in the working field are particularly concerned with the relational area, more than with the professional performances.

Competence to manage leadership relations and interpersonal interactions within a professional context is often lacking; as a result relational tuition

and training has a very important role in the Puntonet course. We think in fact that relational skills can support disabled people sustaining motivation for an independent management of professional and personal relationships. As a consequence, we work in a network with the Social and Employment Services, in that the pupil is called to build up an active role.

2. A BRIEF HISTORY OF VILLA S. IGNAZIO

Villa S. Ignazio in Trento is a cooperative enterprise, that has been involved in activities of personal and vocational training since the end of the 60's. Since then, Villa S. Ignazio has been working also on:

- activities of reception, training and prevention for young adults with economic, working, relational, psychological and learning difficulties;
- information and promotion activities about social justice and solidarity;
- courses for professionals in the areas of social troubles, employment orientation, training, voluntary service and multimedia.

Villa S. Ignazio since 1996 plans and manages Projects co-financed by the European Social Fund through an internal organization called VSI Progetti.

3. WHAT IS PUNTONET 2003

Within VSI Progetti a team works on issues related to disability focusing on training projects promoting ITC literacy. Since 1998 different courses have been organized and about 30 people (disabled people and professionals dealing with social issues) were trained. Since 2002 the training courses for disabled people have become one-year-courses and are called Puntonet.

Puntonet 2003 course is articulated into 6 units alternating different training methods: frontal lessons, one by one training, *Computer Assisted Distance Education* and Internship.

During the first months most of the lessons are *Basic Computer Science* and *MS Office* frontal lessons; these are followed by the *Computer Assisted Distance Education* in which the pupils begin to work in a semi-autonomous way simulating some easy office tasks, such as writing down commercial letters, sending e-mails, printing addresses on labels etc. Individual tutors help the students giving hints when necessary but also supporting their self-organization and learning.

The unit called *Communication's Techniques* guides the pupils during the whole course and its aim is to reinforce relational competence and encourage both the individual and professional communication skills. During this unit

the pupils should also strengthen their skills in managing their relationship with their social and professional network.

During two one-by-one units (The *Project Work* unit and the *Active Internship Search* unit) a personal tutor takes care of each pupil. During the PW unit each pupil defines her own professional project work by analyzing her interests, attitudes and resources also by paying attention to critic elements due to the specific disability. Moreover, the unit aims to make the pupil aware of the social network she can count on. On the other hand, during the AIS module each pupil builds up her competency in job searching by planning an internship (via Internet, newspapers, Job Agencies etc.), writing the CV, preparing a cover letter, simulating a job interview. Each student plays in fact an active role in defining her Internship, thus simulating the decision-making process, that will then be necessary in looking for a job.

The last part of the course is the Internship, in that each pupil works in a firm for about two months putting in practice what learnt. During the Internship a personal tutor will support the pupil when necessary.

During this unit the pupils try out how to enter different systems, integrating the competence acquired in the course to the labor market, keeping in contact with the firm, with the course's tutor, keeping the Social Service up-to-date.

The whole course is intended to promote in the trainees the independent management of professional and personal relationship, during each phase a psychologist is available to talk about the difficulties step by step encountered, and regular meetings with the Social Service and the Public Employment Office are organized, thus supporting both the personal and social reinforcement.

4. THE CONTEST AND THE INTERNAL ORGANISATION

The contest in which the Puntonet course is intended and developed comprehends different and complex systems inter-connected: the labor market in Trento Province, the Public Employment Office and the Law 68/99, the Social Services, Villa S. Ignazio, the disabled pupils, their families and their social network.

Starting from the *labor market*, in Trento Province, the statistics of the local Employment Service underline the interest in people with office and basic ICT competence and the course aims to fill the lack of computer-skilled people requested for office works.

The *Law 68/99* (Law for disabled people employment rights) aims to integrate people in the market labor with disability by providing them

support services and focused job brooking services. The *Public Employment Office* is devoted to connect the law 68 and its application. Each disabled person who decides to look for a job has got a contact person inside the Employment Office helping him going back into work. The person responsible for law 68/99 from the EO works taking into account both the Labor Market requests and the competence, limits and constraints of the single person with disability. This officer works with the Social Service, the Health Service and the other Social Actors that take care of the disabled person looking for a job.

Villa S. Ignazio and *Puntonet* belong to the context of the social actors dealing with disadvantaged people; The staff is permanently in cooperation with the Health and Social services, as well as with the *Public Employment Office*.

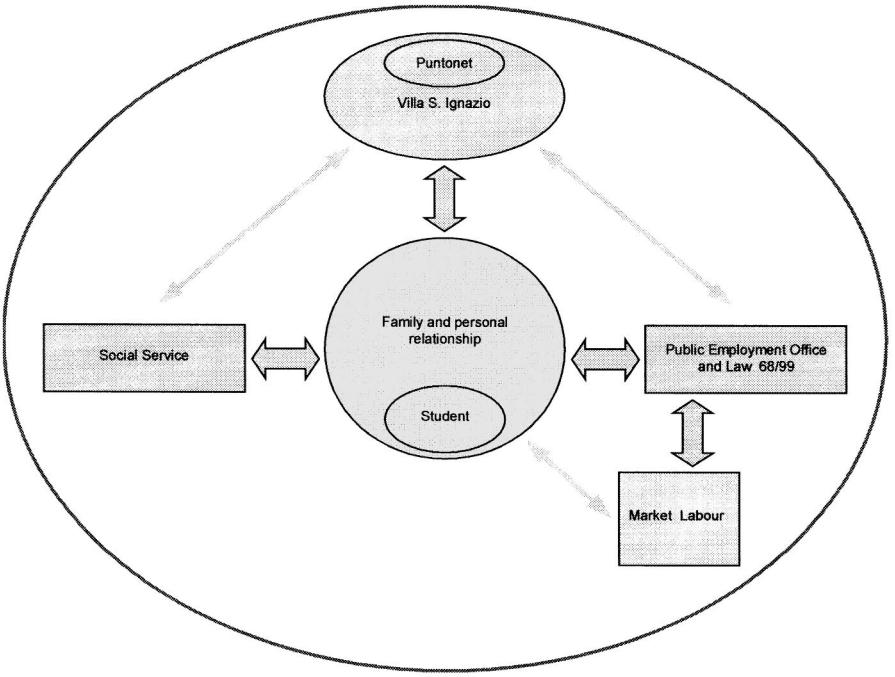


Figure 1. The context.

The *person* with her family and social network is the focus system from which we start to integrate all these systems mentioned above for both job recruitment and social integration reinforcement.

The Puntonet's organizing and teaching team faces the context's complexity itself by a *multi-disciplinarily composition* and frequent *updating meetings*.

A *sociologist* looks after the organizational aspects of the course and maintains the contacts with both the *external* social network (Public Employment Office and Local Government Offices) and the *internal* Villa S. Ignazio network; a *psychologist* looks after the psycho-pedagogical aspects dialoguing with the teachers/ trainers staff and supporting the pupils and, when useful and possible, their families, two *computer scientists* teach Information science and care about course's program, didactic equipment and learning evaluations while the trainers are *pedagogues* that help the pupils individually during the internship and the one-by-one units.

Thanks to the regular meetings, the staff discusses the different issues related to the pupils and plans how to build up the relationships with the external systems involved (market labor, Social Service ...).

5. RESULTS FROM THE EXPERIENCE OF PAST PUNTONET COURSES AT VILLA S. IGNAZIO

The statistics of the local Employment Service underline the interest of the local market labor in people with office and basic ICT competence and the Puntonet course would like to respond to these needs. In fact, the past course editions had important employment results, as emerges from the follow up realized in June 2004.

8 out of 9 former students (2001 and 2002 courses) have been interviewed. 7 have a job at the moment, where they use their ICT competence, and they declare, the course was very useful to get the job. Moreover, 100% say, the course was an important experience from a relational and personal point of view, to know other people and change the usual way to get in relationship.

As far as the 6 pupils of edition 2003 is concerned (the course finished in July 2004), one of them passed an examination for a public job as secretary with ICT skills and another is employed in the firm, where she did her internship during the course.

These employment opportunities have been possible thanks to the competence acquired by the students, but also thanks to the cooperation with the Social and Employment Services, and with the firms, where the students have done their internship.

The relationships with these social actors are getting permanent, thus simplifying the job insertion process. In effect, we notice that local firms and public organizations are getting more and more open to training experiences with disabled people.

6. FUTURE PROJECTS AND PERSPECTIVES

We believe that the results have been achieved thanks to the cooperation and the dialogue between us and the other social actors involved.

In next edition, we would like to go forward in the systemic epistemology taking into account the different meanings, given by the different social actors to the same issues: social and occupational integration of disabled people. We realized, that ICT skills as cross-competence for the professional and personal area are essential, as well as the internship designed by the student himself, in cooperation with the teachers' team and her specific social network.

Moreover, in the future, we would like to promote a counseling and information point, dedicated to former pupils, disabled people interested in job insertion and professionals involved in this topic, thus allowing the dialogue between disabled people and social actor to continue.

REFERENCES

AA.VV., 1998, *Accoglienza Sociale, Ospitalità, Inserimento lavorativo*, Volume realizzato nell'ambito del progetto co-finanziato FSE e PAT "ADEGUA".

AA.VV., 2003, *Il Lavoro Accessibile*, Provincia Autonoma di Bolzano.

Albini, P., Crespi, M., and di Seri, E., 2000, *Il Nuovo Diritto al Lavoro dei Disabili*, Cedam, Padova.

Commissione Interministeriale sullo Sviluppo e L'Impiego delle Tecnologie Della Informazione per le Categorie Deboli, 2003, *Tecnologie per la disabilità: una società senza esclusi*, Libro Bianco, Roma; http://www.innovazione.gov.it .

European Commission, 2001, *E-Inclusion – The Information Society's potential for Social Inclusion in Europe*, Commission Staff Working Paper; http://europa.eu.int/comm/employment_social/knowledge_society/eincl_en.pdf.

Garito, M. A., ed., 1996, *La Multimedialità nell'Insegnamento a Distanza*, Garamond, Roma.

Giordani, M. G., ed., 1995, *Disabili, Tecnologie e Mercato del Lavoro*, Etaslibri, Milano.

Ianes, D., Celi, F., and Cramerotti, S., 2003, *Il piano Educativo Individualizzato, Progetto di Vita*, Edizioni Erickson, Trento.

Minati, G., 1996, *Introduzione alla Sistemica*, Edizioni Kappa, Roma.

Osservatorio del Mercato del Lavoro Trentino, 2004, XIX Rapporto sull'occupazione in provincia di Trento, *Bollettino di documentazione sulle politiche del lavoro*, Provincia Autonoma di Trento.

Ridolfi, P., ed., 2002, *I Disabili nella Società dell'Informazione*, Franco Angeli, Milano.

Scialdone, A., Checcucci, P., and Deriu, F., eds, 2003, *Società dell'informazione e persone disabili: dal rischio di esclusione ai percorsi di integrazione*, Guerini e Associati, Milano.

Zanobini, M., Manetti, M., and Usai, M. C., 2002, *La Famiglia di Fronte alla Disabilità*, Edizioni Erickson, Trento.

INTELLIGENCE AND COMPLEXITY MANAGEMENT: FROM PHYSIOLOGY TO PATHOLOGY. EXPERIMENTAL EVIDENCES AND THEORETICAL MODELS

Pier Luigi Marconi
ARTEMIS Neuropsichiatrica, Roma - Italy

Abstract: Intelligence is the most evident of the emergent properties of the evolution of life on the hearth. The thought process of humans is still not very understood as background information processing. Most of the observations come from clinical practice where an impairment of the thought process is believed as the back ground phenomena of behavioural disfunctions. Data from clinical observation, patients self reports and antypsychotic treatment efficacy are the main source of present models of thought process. Other modeling arise from experimental psychology and cognitive sciences. Just in the last 20 years new data are available by pooling together neuropsychological reasults with clinica observations, and self reports. In present work the statistical structure of such pooling of data is presented from observations performed in normal, psychiatric patients and people with mental retardation. A model of thought process is presented taking into account both this statistic structure and clinical observations. Two main component are suggested as main modules of thought process: the "Rule Inference Processor" and the "Input Organizing Processor". Impairment of one of both processor can explain the formal thought disorders observed in clinical diagnostic group of patients.

Key words: thought process modeling; formal thought disorders; neuropsychology; psychopathology; rating scales; factorial analysis; discriminant analysis; complexity management.

1. BACKGROUND

Intelligence is the most evident of the emergent properties of the evolution of life on the hearth. Its evolutionary goal is to give more probability to the living species to survive with the best quality in spite of an increasing variety of environmental states. So we can think to the intelligence as a property which can sustain problem solving in front of a wide range of inputs. If we think the complexity of a system as the number of its components or states, we can thing about the intelligence as an evolutionary property of life to manage complexity. Actually intelligence is structured linked to other 3 properties: the consciousness, the thought process, and the sociality. All together these four properties are the characteristic of the humankind as the most evolved specie on the hearth. The sociality is perhaps is so important for the intelligent management of complexity that we can speak at present time of "social or pluri-individual" kinds of intelligence. The study of intelligence was concerned with many medical and psychological approach: experimental psychology, neurology, neuropsychology, psychiatry and cognitive science and artificial intelligence. In clinical psychiatry the study of intelligence is strictly linked with the study of consciousness and thought process. This clinical approach is the object of psychopathology which at present time is not only performed by clinical observation and psychological comprehension of symptoms (what the patient describe of inner experience) and signs (what we can observe the behavior of the patient), but it is performed also with the use of assessment instruments. These ones are constructed on the basis of the clinical knowledge of the psychopathological syndromes (collections of symptoms and signs linked to common clinical course and outcomes and/or common clinical response to treatments) using methods developed in experimental psychology (psychometrics). So the instruments of "conventional" or "evidence-based" psychopathology are classical psychometric instruments, neuropsychological test and clinical scale just developed to objective the clinical observations. Using this new approach the study of intelligence was performed as a cofactor influencing the thought process.

2. OBJECTIVE OF THE STUDY

Our study was primarily aimed to look for a statistical structure in the relation of the neuropsychological assessments of executive functions, visual memory, input organization and intellective level with the clinical assessments of man psychopathological dimensions and of mental status.

Secondly the study was to define a theatrical model of the intelligent though process on the basis of such statistical structure.

3. METHODS

49 subjects were clinically observed and assessed by psychometric test (tab. 1): 15 non affected people, 10 affective disorder patients, 10 schizophrenic patients and 14 mental retarded patients. The patients were private out patient, who gave they informed consent to use the clinical date for research purposes (tab.2).

Table 1. Included Subject (global).

	N	%	M	F	N (PNP)	N (PANSS)	N (PANSS + PNP)
Controls	15	27,3	10	5	15	13	13
Affective	12	21,8	7	5	10	12	10
Schizophrenics	12	21,8	7	5	10	12	10
Minus	16	29,1	13	3	14	-	-

Table 2. Included Subjects (by provenience).

	Inpatients	Ambulatoriali	On Rehabilitation	Voluntary
Controls				15
Affective	2	10		
Schizophrenics	8	4		
Minus			16	

The non affected people were university students, screened by means of clinical structured interview (MINI Interview), self assessment questionnaires, and clinical interview. The clinical dimensions were assessed with PANSS Scale of Kay, implemented with 6 more items form BPRS-24 of Ventura to reach a complete comparability either with PANSS and BPRS data. The clinical evaluation of Mental Status was performed with the Mini Mental State Examination. The neuropsychological assessment was done by Wisconsin Card Sort Test, Progressive Matrices (PM38) and an ad hoc built Visual Memory Test (VMT). The VMT was made by a sequence of three slide with a blue rectangle displayed in a black background. The screen was divided in 9 equal areas and the rectangle in each slide was appearing randomly in one of the nine different position. The subject was asked to remember the position of the rectangle in the first, second or third slide. Which slide was randomly defined, but equal to each subject. The VMT was performed with 30 sample of such sequence. The psychopathological dimensions were computed taking into account the factorial structure

computed in the GISES Study performed on more then 800 psychiatric patients. The dimensions were 6, 3 affective (depression, anxiety and mania) and 3 psychotic (reality distortion, cognitive disorganization, and psychomotor poverty syndrome).

The Statistical Analysis was performed on the neuropsychological scores, extracting a factorial structure using the OBLIMIN rotation. Then a discriminant analysis on such factor was performed to validate the ability to contain a sufficient descriptive power of such factor, in a continuum of individuals with different disturbance at the intelligent thought process level. A second level factorial Analysis was performed to find the latent statistical structure of the relation between psychopathological dimensions and neuropsyhological factors. Statistical procedures were performed by Statistical Package SPSS.

4. RESULTS

On table 3 are reported the mean values of the scores resulted on each test by each group of people.

Table 3. Mean Score for each group on each test.

	Control People		Affective Patients		Schizo-phrenics		Mental Retardation	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Categories #	6,0	,0	2,6	2,3	,5	1,0	1,1	,8
Exposed Cards	75,5	5,9	114,9	22,7	128,0	,0	89,2	19,9
Correct Responses	66,0	5,1	55,7	20,4	43,2	11,9	47,1	20,3
Total Errors #	4,4	2,8	54,6	28,4	84,7	11,9	40,5	6,1
Perseverant Errors #	2,2	2,0	26,0	16,1	34,3	22,5	27,9	14,2
Perseverant errors %	2,8	2,6	21,0	11,5	26,7	17,5	30,8	18,8
Attempts to 1st Category	12,8	2,7	23,6	27,8	12,2	27,6	25,5	16,3
Conceptual Level Resp. %	85,8	4,3	42,9	27,9	15,1	11,0	30,6	15,1
Failure Maintain Test	,26	,45	1,18	1,16	1,00	1,26	1,31	1,53
Perseverant Responses #	4,9	1,7	28,9	19,4	40,1	31,2	33,1	18,7
VMT Correct %	99,1	1,5	78,61	29,5	67,7	33,7	16,0	9,4
PM38 Correct %	88,0	95,3	68,0	21,0	44,0	23,4	18,6	5,5
MMSE Total Score	27,7	1,8	27,0	3,2	24,0	5,7	13,0	5,8
Cognitive Disorganiz.	1,34	,28	1,24	1,40	2,54	,99		
Psychomotor Poverty	+1,43	,21	+2,77	,94	+2,96	,88		
Anxiety	2,04	,79	3,82	1,02	2,94	1,19		
Reality Distortion	+,97	,15	+2,54	1,51	+3,18	1,22		
Depression	1,06	,33	1,72	1,36	1,38	1,30		
Mania	+1,10	,44	+1,81	1,63	+1,26	,81		
Age	27,4	9,5	34,1	14,5	38,8	15,7	43,9	9,4

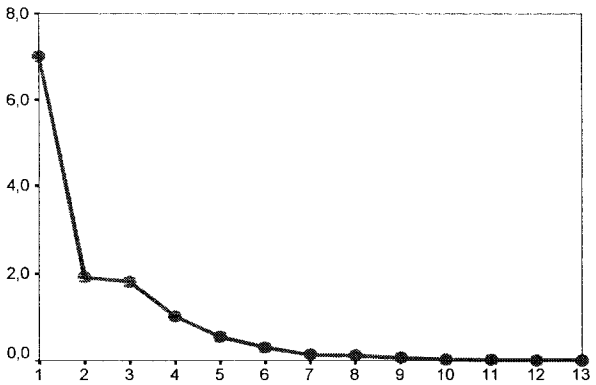


Figure 1. Eigen Values Scree-Plot of neuropsychological data Factorial Analysis.

The evaluations performed on the whole group of people were converging in a 4 factorial model (tab.4), extracted with the Eigen value > 1 and the screeplot criteria (fig. 1), and rotated with a OBLIMIN procedure.

The first factor (tab.4) was that one able to "explain" the most of variance in data scores. It is mainly linked to data concerning "perseverance", such as the ability of subject to change "rule" when the criteria change in ordering cards on the WCST. This factor was called here as "Perseveration Factor". The second factor instead was mainly linked to the PM38, MMSE, and Visual Memory Test performance, being sensitive to the common functions tested by all: the memory performance and the ability to spatially organize the data. For such interpretation of the factor it was called "Memory and ordering factor". The third factor was linked to the failure of the subject to maintain a working hypothesis, although able to give correct responses, with the consequence to have a trend toward an high number of attempts before completing the first category at the WCST. The factor was called the "schemata lability factor". The fourth factor, finally, was the factor linked to the main performance at the WCST, such the ability to complete all the categories at the WCST, with few errors and consequently with a low number of exposed cards. This factor is interpreted as expression of the ability of the subject to make correct "rule inference" by data sequence processing, and was called "Rule Inference factor".

After performing Discriminant Analysis, only 2 of 4 factors were able to distinguish between the 4 groups of people with an high level of statistical significance (tab. 4). In fig. 2 we can see the clear cut psychotics, mental retarded and controls.

Table 4. Neuropsychological Factors Extracted after OBLIMIN Rotation.

Parameter	Test	Fact.#1	Fact.#2	Fact.#3	Fact.#4
		Perseve-ration	Memory & Ordering	Schemata Lability	Rule Inference
	Variance %	29,1	20,1	13,6	11,2
Perseverant Responses #	WCST	-,932			
Perseverant Errors #	WCST	-,864			
Perseverant errors %	WCST	-,846			
Total Score	MMSE		,930		
Correct %	VMT		,901		
Correct %	PM38		,883		
Attempts to 1st Category	WCST			,865	
Correct Responses	WCST			,686	,361
failure maintain test	WCST		-,355	,633	-,434
Exposed Cards	WCST				-,942
Total Errors #	WCST				-,896
Categories #	WCST		,359		,697
Conceptual Level Resp.%	WCST				,695

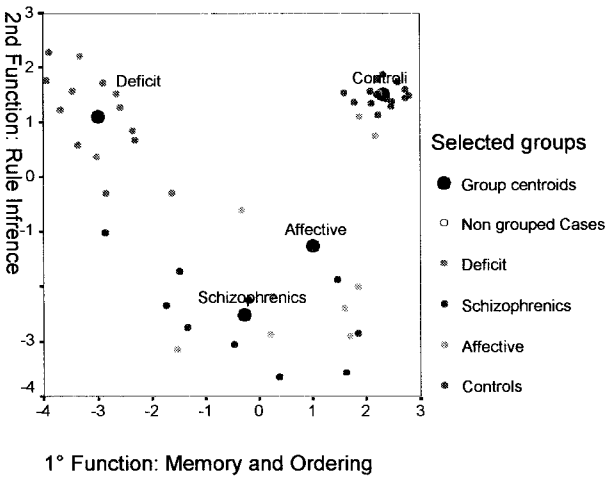


Figure 2. Discriminant Functions graphic form the Discriminant Analysis performed on the Neuropsychological Factors.

These two factor were the "Memory and Ordering" factor, and the "rule inference" factor. The first one was altered in mental retarded people, but with impairments also in schizophrenics; the second one instead was altered

in schizophrenics and at a lower level in affective patients, where was found similar in controls and retarded people (tab. 5).

Table 5. Discriminant Analysis on Neuropsicological Factors.

	1st Function p<.001	2nd Function p<.001
Memory and Ordering	,956	-,361
Rule Inference	,158	1,009
Controls (N=15)	2,312	1,498
Affective (N=10)	1,002	-1,275
Schizophrenics (N=10)	-,279	-2,515
Mental Retarded (N=14)	-2,994	1,102

Merging the neuropsyhcological factors with the six psychopathological factors found in previous researches (3 psychotic factors: reality distortion, psychomotor poverty, and cognitive disorganization; 3 affective factors: anxiety, mania, depression) we found again four factor extracted with the Eigen Value>1 criteria, and after a OBLIMIN rotation (fig.3). The data were concerning all the subject but the mental retarded people, since it was considered reliable all the data of the PANSS, being related also to what it is referred subjectively by people, and most of the subject were severely impaired.

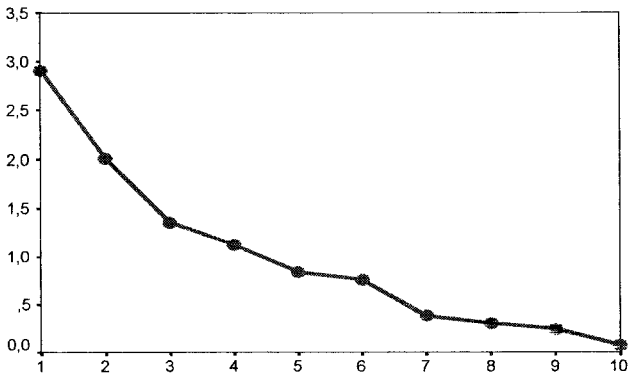


Figure 3. Eigen Values Scree-Plot of neuropsychological and psychopathological factors second level Factorial Analysis.

In table 6 we can see as the two affective bipolar factors (mania and depression) are converging together in the "mood state factor", failing to

converge with any of the neuropsychological factors. Conversely, the "schemata lability" factor of neuropsychological factors remains alone also, failing to converge with any one of the psychopathological factors; it was interpreted as expression of the subject ability to focus and maintain attention on a specific inference schemata with which preprocess data inputs sequence ("Attention and concentration factor").

The other two factor instead are expression of a convergence between neuropsychological observations and clinical psychopathological ones. The main factor explains quite 30% of the whole data variance and is "built" by the convergence of reality distortion, psychomotor poverty, rule inference, anxiety, and perseveration factors. It explains neuropsychologically the core process of the psychotic thought disturbance, as a defect in rule inference process in presence of an high trend in perseverating on unvalidated hypothesis.

Table 6. Second Level Factors Extracted after OBLIMIN Rotation pooling together neuropsyhological and psychopathological factors.

		Fact.#1	Fact.#2	Fact.#3	Fact.#4
	Observation Level	Interference Management & Rule Inference	Mood State	Input Organization and Memory (low)	Attention and Concen-trating Abilities
Variance %		29.1	20.1	13.6	11.2
Reality Distortion	Psycho-pathology	-,839			
Psychomotor Poverty Syndrome	Psycho-pathology	-,745			
Rule Inference	Neuro-psychology	,664		-,498	
Anxiety	Psycho-pathology	-,571			
Perseveration (Low)	Neuro-psychology	,509			
Mania	Psycho-pathology		-,900		
Depression	Psycho-pathology		,792		
Cognitive Disorganization	Psycho-pathology			,945	
Memory and Ordering	Neuro-psychology			-,878	
Schemata Lability	Neuro-psychology				-,897

Psychopathologically it is correlated either to the core psychotic symptoms (reality distortion) and to the Psychomotor Poverty Syndrome, i.e. social withdrawal, apathy, inactivity, affective bluntening. This factor also it is linked to an "affective" factor such as anxiety, giving an important experimental evidence to the feeling of fear associated to the decrease of ability to "understand" properly the reality. This factor was called "Rule Inference factor" but also "interference management factors" being the rule inference strictly connected with a "clean sequence of data", in a "input stream" made either by external complex data sets, and internal emotional inputs. The last factors was made by the convergence of the "Memory and Ordering factor" form neuropsychology and the "Cognitive Disorganization Factor" from psychopathology. This psychopathological factor is linked mainly to the formal thought disturbances, as failures in logics, in making plans, disorientation, attention, difficulty in abstract thinking, bizarre behavior. Giving importance in naming this second level factor to the neuropsychological correlates, we called it as "Input Organization and Memory".

Performing the Statistical Analysis to evaluate the discriminating power of these 4 factors in distinguishing the schizophrenic patients by affective patient and controls, we found a statistical significance only in two factors (tab. 7): the main factor was able to discriminate psychotics by non psychotics, and was made by the convergence of the ability to organize and memorize inputs associated to a good interference management ability and a good capability to make inferences form data sequences. The controls were performing high as the psychotics very low in such a factor. The second discriminating factor was made mainly by an impairment in input organization and memory, which was severely impaired only in schizophrenics, were the best performance is instead observed in Affective Patients.

Table 7. Discriminating factors form Second Level Factors extracted pooling together neuropsyhological and psychopathological factors.

	1st Function p<.001	2nd Function p=.026
Rule Inference & Interference Management	1,164	,293
Input Organization and Memory	-,889	,807
Controls (N=13)	3,197	,124
Affective Patients (N=10)	-1,247	-,590
Schizophrenics (N=10)	-2,909	,430

5. DISCUSSION AND MODELING OF RESULTS

These experimental finding give a particular relevance in psychopathology to 2 events in data processing at a neural level:

1. the **input organizing process**, in which the working memory properties to remember items is linked to the ability to organize them in sequences by multiple ordering criteria
2. the **rule inference process**, in which the rule that links the antecedent to consequent is extracted from the ordered sequences, in a context of input filtering, in which significant inputs are gained by non significant one (noise or chaos).

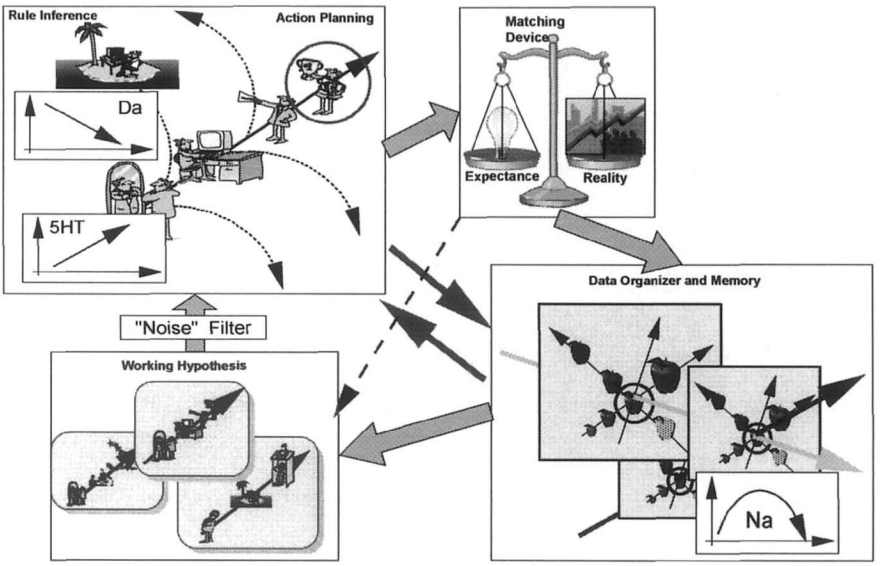


Figure 4. The Thought Process.

So making a correct rule inference, it seems necessary to focus the attention on the significant items. This subprocess, however, results to be statistically independent from the rule inference and to be not significant in discriminating schizophrenic patients by non affected people and affective disorder patients. This focusing process may allow a correct induction of rules, when the interference of non significant stimuli may reduce the quality of the inference and impair the process at all.

The input ordering process give a "structure" to data, allowing to discriminate objects and collocating the data among a "knowledge map" made by cognitive coordinates; this structure allow to "recognize" objects, giving them meanings, and putting them among a frame of rules, class, etc.. This process extract form the chaos the information to be oriented.

The "interference" is an additional input which is not cut off and is included to the structured data input stream, making this structure low contrasted. The presence of an interference may represent by itself an increase of complexity of the stimuli, so is stressing the organizing module. Either because of the presence of a "complex environment" or because of the maintenance of a "complex processing data set" or because of the presence of an high interference, the organizing module may fail the goal to structure the data in semantic frames, which represent the precursor of the focusing filtering schemata.

After the structuring process and the focusing of input filter on the relevant inputs, the thought process is extracting the association between the data sets, ordered in time sequence; extracted a regularity in such a sequence, the "rule" become" a schemata in the further sequence of data inputs. The process goes over with a circularity between, data structuring process, sequence ordering, rule inference, expectance validation, and so on. When the circularity of the process is convergent toward a valid theory by which the expectance is confirmed by the evidence from experience, the system is relaxing toward the minima, and the allert-anxiety state is released.

However when the process is non convergent, an attempt to reduce complexity of stimuli is done; it can be performed be "cutting off" noising signals, as inner state signals or outside world signal, or just reducing the complexity of external world at all. This last reaction may induce the "poverty syndrome" seen either in chronic schizophrenic patients and in chronic obsessive patients. Both can have a failure in the "theory validation" process, and may attempt to "project" in the external world the schemata arisen from the inner affective inputs.

However in the obsessive patient the orienting and memory system is not on failure as with the schizophrenic patient; this partial failure allows to the obsessive patient to maintain the correct orientation in interpreting the "interfering thoughts" just as inner interference. However when the orienting module fails together with the rule inference one, the thought disturbance explode in a psychotic buffet in which any organizing capability is lost as it is observed in the schizophrenic psychotic acute episode.

Looking to the mental retarded people, the most common reaction is anxiety as physical complaints or as agitated behavior when the complexity of external environment changes (i.e. the daily routines changes, the steps toward problem solution becomes too high, or the concurrent stimuli

increase too much). In this case we know that the weak module is the organizing and orienting module, whose inability to produce a valid schemata does not allow to the process to reduce activation; the most autonomous attempt to react to this overload is controlling the environment imposing their expectances or routines; however, when, in spite such an adaptive reaction fail to reduce the complexity of stimuli, a "psychotic reaction" may occur.

In patient with mood disorder the depressive state may reduce the capability to pay attention to external stimuli, with a decrease of attention and concentrating capability seen in these patient; but the ability to organize and memorize inputs is not impaired, but it seems conversely just increase in spit to control people also. The psychotic state "mood congruent", may be not an actual psychosis by interference, but just a decrease of outside word signal, linked to the depressive state, and the increase of inner world one, with the failure of the "theory validation" module. These setting lead to a perseveration on the thoughts just derived from inner drives.The psychotic state mood incongruent (interpreted by the present module as an increase of inner interference also) may arise only if an anxiety or hypomanic activation is concurrent, or a mixed state (concurrent presence of agitation, iperactivity and depression, feeling of guilt, and difficulty to concentrate) arises. In this case an actual interference between signal can lead to misinterpretation of outside world stimuli.

Confirmations from actual clinical practice are taken from clinical monitoring data. Preliminary results that are confirming the neuropsychological factor involved in psychosis are show in tab. 9-11.

Table 8. Discriminating Psychotics from Affective patients by clinical and neuropsychological parameters (N= 59 outpatients).

	Factor 1 p < .000	Factor 2 p < .004
Behavioural Disfuntion	-1,274	,579
Affective Complaints	1,304	,506
Diagnostic Group		
None	-0,029	-1,271
Affective	1,108	,273
Psychotics	-1,277	-,207
Both	-,859	1,233

Table 9. Discriminant Power of Discriminant Function computed pooling together clinical and neuropsychological parameters.

	Affective	Psychotics	Both
None	5,598 ,013	3,183 ,066	7,319 ,005
Affective		9,013 ,002	6,281 ,009
Psychotics			2,366 ns

Table 10. Discriminating Psychotics from Affective patients by clinical and neuropsychological parameters (N= 59 outpatients).

N= 59 (OutPatients)	Factor P < .002
Rule Inference	1,000
Diagnostic Group	
None	1,176
Affective	,489
Psychotics	-,339
Both	-,987

Table 11. Discriminant Power of Discriminant Function computed pooling together clinical and neuropsychological parameters.

	Affective	Psychotics	Both
None	1,620 ns	6,687 ,015	15,038 ,001
Affective		2,931 ns	10,791 ,003
Psychotics			1,657 ns

In tab.9-11 results from discriminant analysis by data from clinical setting is presented. The impairment of **rule inference process** is confirmed as the main neuropsychological factor which discrimantes "psychotics" by non psychotics and by affective patients without psychotic feature. The data has few schizophrenics among the psychotic population so the discriminating factor between schozophrenics and affective psychotics may be not detected.

6. CONCLUSION

Taking into account the present model and the experimental evidence, two processes are basic in thought disorders patophysiology, who manage complexity of environmental states, and complexity of rules linking one state to the consequent one. The first one is a "Input Organizing Processor" (IOP), the second one is a "Rule Inference Processor" (RIP). The hypothesis presented here can be useful not only in interpreting psychopathological phenomena., abut also planning therapeutical, rehabilitating and preventing interventions.

REFERENCES

Berdia, S, and Metz, J. T., 1998, An artificial neural network simulating performance of normal subjects and schizophrenics on the Wisconsin Card Sorting Test, *Artif. Intell. Med.* **13**(1-2):123-38.

Pancheri, P., Marconi, P. L. Brugnoli, R., and Carilli, L., 1999, Pensiero, in: *Trattato Italiano di Psichiatria*, P. Pancheri and G. B. Cassano ,eds., Masson, Milano.

Amos, A., 2000, A computational model of information processing in the frontal cortex and basal ganglia, *J. Cogn. Neurosci.* **12**(3):505-19.

Callicott, J. H., Mattay, V. S., Verchinski, B. A., Marenco, S., Egan, M. F., and Weinberger, D. R., 2003, Complexity of prefrontal cortical dysfunction in schizophrenia: more than up or down, *Am. J. Psychiatry* **160**(12):2209-15.

Savage, R. M., Jackson, W. T., and Sourathathone, C. M., 2003, A brief neuropsychological testing battery for evaluating patients with schizophrenia, *Community Ment. Health J.* **39**(3):253-62

Marconi, P. L., 2004, L'approccio dimensionale in psicopatologia, in: *Manuale di Psicofarmacologia Dimensionale*, T. Cantelmi, ed., Adelfi, Roma.

Peled, A., 2004, From plasticity to complexity: a new diagnostic method for psychiatry, *Med. Hypotheses* **63**(1):110-4.

DISABLEMENT, ASSISTIVE TECHNOLOGIES AND COMPUTER ACCESSIBILITY: HINTS OF ANALYSIS THROUGH A CLINICAL APPROACH BASED ON THE ICF MODEL

Carmelo Masala and Donatella Rita Petretto
Department of Psychology, University of Cagliari

Abstract: Aim of the present paper is to introduce some hints of analysis about the problems encountered using computer science instruments in school to improve the participation and the learning skills in people who are in a situation of disability. The ICF model of the World Health Organization has been explained as a proposal to classify the health conditions according to a bio-psycho-social approach. Classification tools of the computer Assistive Technology are described in the field of the Assistive Technologies and then hints of analysis are proposed to describe the various factor influencing the disability situations and then, starting from this, a definition is given of the principles of individualization to create a customized setup of the computer aids for people in a disability situation. Such approach has been experimentally tested in the project M@rte-Handicap of MIUR and Regione Autonoma della Sardegna, and is available on line at the site: www.portalem@rte.it.

Key words: disability; computer assistive technology; ICF model.

1. THE NEW CULTURE OF DISABILITY

Over twenty years have gone by since the World Health Organization published a Classification of Impairments, of Disabilities and Handicaps called I.C.I.D.H. (International Classification of Impairment, Disabilities and Handicaps) (WHO, 1980). All these years have not been enough to delete the expression "bearer of an handicap" ("portatore di handicap") not only from the common language, but even from the technical language of the

experts. It's likely that if you ask, even to a field expert, to describe a certain handicap situation, he will describe the disability or the dysfunction. In other words, he will say that the handicap is the amputation or the paraparesis and he will not talk about the quality of life or the limitations coming from the society itself to the impaired person. The person will be seen as someone who "bears a handicap", a "paraparetic" and the handicap will never be seen as a situation that can evolve into better or worst, also influenced from context factors. People still tend to use the word "handicap" as something a person bears, an attribution, and not some kind of dynamical situation, depending on the interaction between person and environment, as the World Health Organization proposed. This tendency not to modify the meaning of the term leads to a very important implication: it's hard to identify what really causes the handicap, and the cause will still be seen as lying in the limit of the person itself; as if the responsibility for the handicap only lies on the person's impairment and not in the whole society. In the time from 1980 to today, the World Health Organization has tried to spread the I.C.D.I.H. model, influencing people working in the health field and legislators, and has also started a revision of the model called I.C.D.I.H.-2, recently renamed ICF (WHO, 2001). It is a new classification tool of a person's health conditions and it takes part in family of international classifications of the World Health Organization, published in November, 15, 2001 and tested in 65 countries from 1994 to 2001. The ICF is now acknowledged by 191 countries as an international reference to describe and measure health and disablement process.

According to the W.H.O., such tools gets over the old ideas about health and disablement. The person is not classified as healthy or sick anymore (much less as someone who bears a handicap or a disease), but as a person operating in a more or less suitable way in the various situations which the challenge of life has in store. The ICF, since it keeps into account also the environmental factors, allows to classify and quantify components of life enabling or disabling, and components that make participation easier or harder, and it suggests different ways to find the tools to give a chance to keep on living an active life in the family, in the working place, in the community. The intention to focus the attention on social aspects of handicap is further strengthened in the ICF with an operating proposal which consists in giving a tool to determine the impact of the social and physical environment on the "functions", on the physiology of a person. This intention can be clarified by reading one of the examples made to introduce the ICF: when a person with a severe impairment finds it hard to work in a building with no access ramps or lift operators, it can be observed that every intervention to be worked can be classified as: intervention on the person and intervention on his life environment. Thus it will be possible to find out

if it is necessary to remove the architectural barriers or to set up tools to get over them, and so the model will point out not only the impairments of body functions, the impairments of body structures and the activity limitations, but also permit an analysis of other personal and environmental components which "enable" o "disable" the person, such as the need to set up some aids instead of rejecting the person from his working place or from his living place.

The ICF points out the idea of "participation". It tries to give a "historical" evaluation, good for that moment and not for another, a "hic et nunc" indicating the suitable interventions to work to get over that moment of difficulty or to keep a given health condition and/or some positive aspects in the condition of a person. The ICF is also a universal model, integrating, not inclusive for a minority. Any person can have a disablement, but this must be considered in his dynamicity and in his multidimensionality.

As the development of a new international classification of health's conditions, the ICF model is the goal of a new cultural evolution representing the change from a medical model centered on pathology to a bio-psycho-social model, widening to consider the meeting between the person and its social and physical environment (AAATE, 2003).

During the same time in which the WHO introduced ICIDH model and then ICF model, in scientific and market field there has been a great development in the tools of assistive technology for people who are in a situation of disablement. Such development did not only consist of those tools allowing to reduce the impairments of a person, i.e. an articular prosthesis, but also consisted of trying to satisfy the needs which go beyond the mere functions, to reach the psychological needs and the social participation needs. In this evolution the new computer technologies must be considered, even though it must be said that only in the last decade the scientific research started to take care of the concept of Computer Accessibility, of the environmental accessibility, and even of Internet and Multimedia tools (for example see Vanderheiden, 1997, 1998; Hardwick et al., 1998; Hayes, 1998; Alberti, 1999; Slatin, 2001; Evans and Blenkhorn, 2003). The birth in an international level of new guidelines like the ones created by the World Wide WEB Consortium (W3C), about the computer accessibility, rises from the interest of software developers to the so-called "handicapped people". In Italy a new law has just been made about WEB and computer accessibility in the Public Administration (Legge Stanca, 4 del 9 gennaio 2004, "Disposizioni per favorire l'accesso dei soggetti disabili agli strumenti informatici"; formerly the directives of W3C had been received in the Circolare Ministeriale del 13.03.2001).

2. COMPUTER ACCESSIBILITY AND ENVIRONMENTAL ACCESSIBILITY

The common meaning for "computer accessibility" is being able to take advantage from contents and information (through the tools of multimediality) by people in a situation of disability. Some aspects of computer accessibility are related with the wider aspects of environmental accessibility, so that it appears useful to distinguish three different fields:

- physical (environmental) accessibility which means accessing to rooms and workstations;
- accessibility to WEB and software;
- accessibility to the computer.

The first field is structural and environmental in nature and can initially leave apart clinical knowledge being almost closer to engineering and architectural knowledge. Yet it represent the first step to reach computer accessibility. Environmental accessibility means being able to reach a building and all its units, to get in easily and to take advantage of all spaces and equipments in a safe and autonomous way. The presence or the absence of any architectural barriers or any environmental aids are one of the first things to be considered when evaluating the accessibility of the computer tools. This is not always considered in all its complexity, because the concept itself of barrier (drawback of environmental accessibility) is a very complex concept and changing in its nature from the characteristics of disability of the single. The meaning of architectural barrier is:

- any element present, absent, lacking, unsuitable which for its shape, color or position hinders mobility, use or orientation of anyone and particularly people with reduced or hindered motor ability;
- any obstacle keeping from or limiting the correct and safe usage of parts of equipment to anyone;
- missing warnings or technical solutions.

The important thing is that different barriers can exist for different disabilities, to the extreme case in which an environmental element can be an aid for a person in a certain situation of disability and a barrier for another, i.e. a phone booth at the same level of the wheelchair may help someone in a wheelchair but may hinder a low-seer or a blind person using a signalling tip (AAATE, 2003).

The other two meanings of the concept of accessibility are about the use of a computer (or interfacing with the computer) and accessibility in different contents, WEB sites or software.

3. ASSISTIVE TECHNOLOGIES AND COMPUTER AIDS

The two last meanings of accessibility are about the use of aids, because often being able to use a computer or to access the contents of a software requires using tools of Assistive Technology. The concept of Assistive technology includes that of computer aid, but not only. In fact the person in a situation of disablement might need an expanded keyboard to access the computer, but also a system of posture and head control. They are all examples of Assistive Technology, but only the first one is a computer aid[1]. Such distinction is important to make customized definitions of aids.

A classification proposed for the computer aids is the one distinguishing between hardware and software aids for the accessibility and for learning aids. The first fields includes all those programs or input peripherals with the aim to improve the access of the person in a situation of disability to the computer as a tool; the second group includes all the software with learning contents, no matter if they are developed especially for learning reasons, or if they are commercial programs in which some learning content can be seen. The idea of Assistive Technology is clearer for what concerns the accessibility aids and less clear for what concerns the learning aids, so that according to some authors the learning aids should not be considered as aids (AAATE, 2003). In our view, even didactic software should be considered as aids, since they are compensating tools in some cases and training and rehabilitation tools in others for people in a situation of disability in the learning field (Masala and Petretto, 2004).

As we have just said, as in ICF model disablement should intended as a dynamical process (no more an attribution to the person but a situation in which any person can be any time there is a difference between the individual capacities and the environmental factors) the Assistive Technologies take part of the environmental factors category as aids (WHO, 2001). A chance to understand the dynamics of such process is given by the

[1] Even in the text of the Stanca's Law there is a hint to both concepts and their relationship. The text reads: (1) for "Accessibility" the ability of computer systems, in the forms and the limits allowed by technological knowledge, to give services and usable information, without discriminations, even from the ones who because of disability need Assistive Technologies or particular setups; (2) for "Assistive Technologies" those tools or technical solutions, in the specific field of computer Assistive Technologies referring to: a hardware and software tools, allowing the disabled person to access information and services given by computer services, getting over or reducing the disadvantage condition; (3) for "Computer Aid" any special computer environment or hardware or software system that allows to interact with the computer environment and that allows or makes easier certain activities by compensating functional limitations.

distinction between capacity and performance. Capacity is what a person is able to accomplish in an aidless environment, performance is what he can achieve when aids and Assistive Technologies are present. A good example is the person whose leg has been amputated, showing limitations in walking and running, but with a prosthesis (aid) such limitation are reduced and can even disappear, as long as the prosthesis is present. A better example for our field of study is that of a person with limited motor abilities as a consequence of a tetraparesis, who cannot use a mouse to operate a computer (capacity). Using a mouse emulator that takes advantage of a residual movement in the eyes allows controlling the cursor and thus allows the autonomous control of the computer (performance). In this examples the mouse emulator stands between capacity and performance, as an aiding environmental factor, and reduces the disability condition when using a computer. Another example can concern a person showing a sight limitation as a consequence of a severe hypermetropy, with nystagm and visual field reduction. Such a person can not use a computer autonomously because he cannot see the cursor, he cannot read what is on the screen and his eyes get easily tired (capacity). Using a magnified cursor, a magnifying software and a vocal synthesis allows him to use the computer autonomously and to prevent sight stress (performance). Also in this case the tools of assistive technology stand in the gap between capacity and performance and reduce the condition of disability in which the person is.

Taking a look at the state of the art of computer aids, there are two remarks we can make (AAATE, 2003; WATI, 2002). An optimistic one according to which it can be "virtually" thought that any person with dysfunctions such that he cannot use a computer, might use it in an effective way if only he had the suitable aid, which can be any environmental feature, or any tool of Assistive Technology which cooperates with his capacity and improves his performance, thus reducing the situation of disability in using a computer. The second and more realistic remark links clinical and application needs to economic factors of the aids market: "a clinically and methodologically based process is necessary to identify the computer tool useful for the reduction of the disability situation in that person". The second remark is not to reduce the optimism of the first one but to mark the need of a study approach of the single case, meant to customize the aid, as to avoid situation in which the benefit/cost ratio is not always benefit-oriented and in which buying high cost tools doesn't always reduce the disability situation, because such aids have not been customized enough. We are talking about all situations in which giving a computer aid, as a magnifying software, (or a non computer aid, as a wheelchair) happened without a suitable customizing and without evaluating the former skills and the satisfaction of the user. Such cases have often turned in a too early rejection of the aid by

the user who found it useless, and also made the user reluctant to new proposals and made the authorities and clinics who should had given new aids unable to provide them.

It is remarkable that there can be a continuity between the user with no disability and the user with disability about the accessibility aids. In an impressive description of the future of Assistive Technologies for WEB, Vanderheiden hinted a comparison between the needs of a daily user of computer tools (using them in all the different channels and situations and interfaces) and people in a disability situation (Vanderheiden, 1997, 1998). The first comparison is the usage of an interface in a noisy environment or deaf or low hearing persons: both situations need hearing aid. The second is the use of an interface while driving a car and blind or low seeing persons: in both cases you cannot and you must not trust your sight too much. The third example is the need of a few keys or ergonomic keyboards for people with some kind of dysfunction in moving: in both cases the keyboard needs to be adapted. The fourth kind is about the use of an interface by a person with busy hands and the case of a person who cannot use hands or does not have hands at all: in both situations tools are needed that can be used with no hands. The fifth kind is about the comparison between people working in chaotic environments and people who cannot concentrate: the presence of too many stimulation negatively influences performance. Such examples are useful because they recall the concept of universal design which is useful for everyone and so it is useful for people in a disability situation, too, and besides recall the concept of flexibility of interfaces, which was outlined by the same author in the same description (Vanderheiden, 1997). This last concept is defined in literature as the need of creating mode-dependent interfaces and mode-redundant interfaces. The first feature is about introducing the same concept through different modes (seeing, hearing, tactile, and so on) and the second one is about the fact that the same content is represented at the same time in all the possible modes (i.e. a software with a soundtrack, pictures and text showing the same content and allowing full use of the same content by people with different dysfunctions or users in different situations or interfaces). Such comparisons recalls the idea that basic principles are needed to design and create accessible user/machine interfaces and to improve the participation to the WEB (Vanderheiden, 1998, Hayes, 1998). Such principles should be flexible, general but not generic and with a remarkable and widely customizable according to needs and aims that the user is willing to reach. But a general fundamental principle must concern, most of all, the avoiding of harming which means keeping from limiting further the skills of the person (Vanderheiden, 1998). As an example we recall what was found out during a test. Trying to improve learning skills, reading skills and computer using skills of a low seeing person, a project was

defined whose intervention was to magnified the fonts of the word processor. As a result: the person used the word processor only to read or write single words, under dictation with the reading limited to a single word because of the dimension of the character and the fact that every single word took an entire page. Trying to improve the accessibility of a low seer he had been turned into a severe dyslexic! The story told clearly shows the need to outline principles and guidelines in defining tools and customized setups for the accessibility improvements.

Thus it becomes important to define general principles related to computer accessibility for people in a situation of disability starting from the available literature and the tools available on the national and international market and in the research field. The research filed Is useful because there are some aids, often prototypes, set up ad hoc for single cases or for test situations which hardly find a place in the commercial market but are useful to outline some guiding principles.

In the M@rte-Handicap project of MIUR and Regione Autonoma della Sardegna (Moduli di Apprendimento con Reti Tecno-Educative) (Learning Modules with Techno-Educational Networks), a project who involves 500 Sardinian schools in an extensive use of computer science tools for didactic purposes, we dedicated particular attention to the topic of accessibility and to the use of computer aids with person in situation of disability (Masala and Petretto, 2004).

All the previous considerations led us to use ICF model as reference tool for classifying the situations of disabilities. Such classification refers to the functional needs and the contextual problems related to accessibility, rather than to the pathology or the impairment condition: the only reference to pathology or to impairment cannot predict the best computer aids or the best solution for accessibility.

As ICF model refers to a dynamic approach, how can we solve the problem of customizing aids solution? A descriptive approach based to a reduced number of classes cannot solve the problem of single case peculiarity, even if it can be useful for a general approach. It is necessary an approach based on a collaboration between a clinician, an engineer, other rehabilitation and educational agents and the subject itself. For our opinion and experience, a clinical approach is the way to solve the needs of fitting an aids solution to the features of any single person in situation of disability.

REFERENCES

AA.VV., 1998, Access to multimedia technology by people with sensory disabilities, *Proceedings of National Council on Disability*, March.

AAATE, Association for the Advancement of Assistive Technology in Europe, 2003, *Tecnologie e Disabilità, scenario 2003, il punto di vista dell'AAATE*, in: www.aaate.net.

Alberti, P. W., 1999, Accessibility for the hearing impaired, *International Journal of Pediatric Otorhinolaryngology* **49**(1):55-58.

Evans, G., and Blenkhorn, P., 2003, Architectures of assistive software applications for Windows-based computers, *Journal of Network and Computer Application* **26**: 213-228.

Hayes, M. G., 1998, Individuals with disabilities using the Internet: a tool for information and communication, *Technology and Disability* **8**:153-158.

Masala, C., and Petretto, D. R., 2004, *Progetto M.A.R.T.E. – Report Studio Pilota e Protocollo di Valutazione degli Ausili Informatici*, PO E8, in: www.portalem@rte.it .

Slatin, J. M., 2001, The art of ALT: toward a more accessible WEB, *Computers and Composition* **18**:73-81.

Vanderheiden, G. C., 1997, Anywhere, anytime (+ anyone) access to the next-generation WWW, *Computer Networks and Systems* **29**:1439-1446.

Vanderheiden, G. C., 1998, Cross-modal access to current and next-generation Internet – fundamental and advanced topics in Internet accessibility, *Technology and Disability* **8**:115-126.

WATI, Wisconsin Assistive Technology Initiative; http://www.wati.org .

WHO, World Health Organization, 1980, *International Classification of Impairments, Disabilities and Handicaps*, World Health Organization, Geneva.

WHO, World Health Organization, 2001, *ICF Classification of Funtioning*, World Health Organization, Geneva, (Tr. it. Organizzazione Mondiale della Sanità, *Classificazione Internazionale del Funzionamento, della Disabilità e della Salute*, Erickson, Trento).

CHAOS AND CULTURAL FASHIONS

Sergio Benvenuto
CNR, Viale Marx, 15 - 00137 Roma, Italy

Abstract: The author approaches the old sociological riddle of cultural fashions--that is, of the ephemeral success of a cultural trait--making use of the theories of Chaos and Complexity. Reconsidering a classic of the sociological literature--Georg Simmel's interpretation of fashions--the author approaches the dynamics of fashions as the result of two contradictory tensions, one pushing people to imitate others perceived as "up", and another pushing each individual to distinguish her- or himself from the others.

Key words: cultural fashions; urn problem; Simmel's paradigm; strange attractors; systems dependent on initial conditions.

1. INTRODUCTION

I have always gone to see many movies in my life. But the reasons for the success of certain films and the failure of others largely eluded me: if one took two films of the same genre, and of a more or less equivalent quality, one was a box-office smash and the other quickly disappeared from the scene. Why was this?

Even some famous cinema experts had asked themselves the same thing, they find the reasons for the success or the failure of a film often incomprehensible. The cinema distribution companies are well aware that it is not enough to have a cast of stars or a famous director for their film: there is always the risk of it being a flop.

De Vany and Walls (1996, 2003) analysed the box-office takings of the 50 most popular films in the United States during a nine-month period in 1995. They found that the top four films on the list had taken over 20% of all the tickets sold, while the last four on the list took only 0,001%. The first

film on the list had takings of 49 million dollars and the last under 5.000 dollars--a relationship of 10.000:1!

But this enormous disparity in viewing figures can be observed in almost all cultural sectors, philosophy and religion included. In other words, cultural processes tend to function *like fashions*: some traits or products become popular while others--despite their true value or qualities--remain ignored. But how is this avalanche effect produced, according to which we all go to see the same few films, read the same few books, and always profess the same few philosophies or religions?

Thanks to the research inspired by the theories of chaos and complexity, we have finally become aware of the fact that this "iniquity" is not specific to human beings. It is the effect of processes which mathematicians call non-linear and chaotic[2].

2. THE ANTS AND THE BALLS

In the '80s various entomologists repeated the following experiment: they put two deposits of food for ants at the same distance from the anthill[3]. When the ants had discovered the existence of these deposits and had started taking the food, the experimenters kept supplying the two sites with food so that they were always rigorously equivalent. This then was an application of the famous dilemma of Buridano's donkey to the world of ants.

But the results of this experiment surprised the researchers. They had predicted that, following an initial period of oscillation, with time the choices of the ants would have become stabilized in an equal way: 50% of the ants would have gone to site A, and the other 50% to site B. But that did not happen. In fact the choices of the ants were subject to chaotic oscillations, which never became stabilized, as can be seen from figure 1.

On the Y-axis we have the percentage of the ants that visited site A, while on the X-axis we have the time elapsed. Over time we have continuous variations, not a stable equilibrium. We can say that in certain periods it is "fashionable" to go to site A, in other periods it is "fashionable" to go to site B, and in other periods still the popularity of the two sites is equal. It is said that the system of the anthill-site relationship is *far from equilibrium*: it undergoes continual variations.

[2] Not all non-linear processes are chaotic. As we shall see, chaos does not mean pure chance.

[3] See Kirman, A. (1997). Cf. Ormerod, P. (1998).

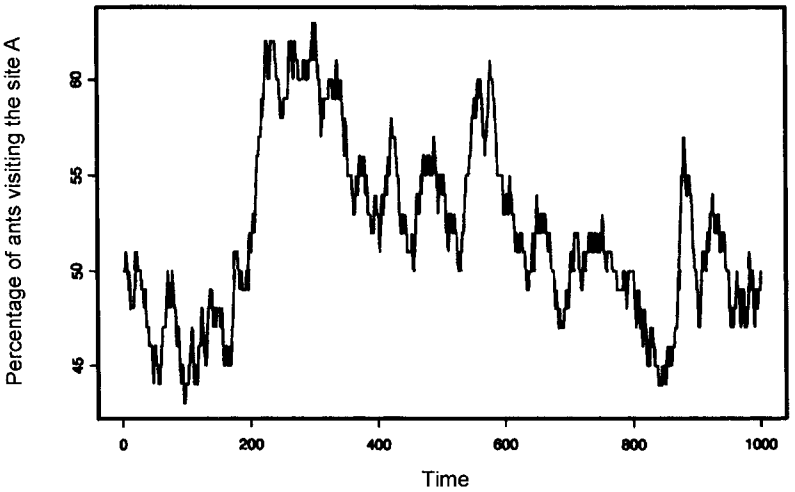


Figure 1. Typical Solution for the Ants Model.

We notice that a chaotic series of the same type describes processes as different and varied as, for example, the annual variations of the GNP of an industrialised country or the annual variations of the numbers of sun spots.

And yet after a sufficiently long period of observation, it turns out that *on average* the ants pass more or less the same amount of time in A as they do in B: or—as we say in "chaos-ology"--the series has *a strange attractor*. In the long term then, order appears; it is said that this system is locally unpredictable and globally stable. But how can we explain this chaotic unpredictability?

Ants, like human beings, are imitative animals: if an ant goes first to site A, others will follow it to the same site, which will attract still more ants and so on. It is probable that some ants are more able to lead others, more "charismatic" than others: this would explain why at a certain moment most of the ants go to A and not to B, or vice versa. This avalanche effect is called a *self-reinforcing process* and it was formulated mathematically by Arthur, Ermoliev and Kaniovski (1983) as the "urn problem".

Let us imagine that there are some black and white balls, equal numbers of each, in an urn and imagine that someone takes out a certain number of balls randomly. The rule for putting them back into the urn is "if by chance more than half of the balls taken out are white, then you must put them all back in, but with an extra white ball; the same thing applies if more than half of the balls taken out are black". If this operation is repeated, following the rule every time, it can happen that a distinctive tendency towards white balls,

for example, develops. Since the possibility of taking more white balls increases each time a handful is taken out (because a white ball is added every time, there are more white balls in a handful), after a while the urn could become filled with an overwhelmingly greater number of white balls. In other words, at each "turn" the probability of getting a majority of white or black balls does not stay the same, because at every turn the probabilities of the following turn are modified.

This mathematical game represents what happens with the ants: if by chance, at the beginning, more ants go to a certain site, this behaviour will influence that of the other ants, and therefore an "avalanche" effect is created in favour of one of the two sites. The future evolution of a process therefore depends on the initial processes, even if they have a minuscule effect. The scientists and experts say that the system is *notably dependent on the initial conditions*.

3. THE MARKET, BLINDFOLDED GODDESS

Some economists have observed that the same dynamics happens in commodities markets. In many cases, when two products are in competition, the objectively better one does not prevail: instead the winner is the one which, for whatever reason, is chosen by an initially superior number of people, often for totally casual reasons, or in any case inscrutable ones. It is now obligatory to cite the example of the QWERTY keyboard for typewriters and computers, the one which we all use to this day (Gould 1991). This is a particularly irrational system, and yet it is the one which has prevailed. The industries of the sector have proposed easier to use and more efficient types of keyboard many times, but every time this has been a failure.

When the first video recorders came onto the market two companies were in competition, each of which offered a different model: Betamax and VHS. In the end VHS won and Betamax disappeared. And yet, from a technical point of view the Betamax video system was, in many ways, superior. The point is that at the time, since they were both new products, the consumers were not experts. Buyers--also of books, political ideas, artistic tendencies, religions, etc.--are *interactive agents*; they influence each other reciprocally, and are not isolated atoms. Like the ants, also the consumers of video recorders have ended up imitating each other and fortune has rewarded the VHS. All those who preach the thaumaturgic virtues of the market, according to which the mechanisms of the market always reward the best products (also cultural and political ones) should meditate on cases of this kind.

In general, unjustifiable disparities very often arise between two or more cultural objects of analogous quality. In fact very few products (today they are usually American) take up alone the greater part of our cultural consumption, and all the other products have to be content with the crumbs that are left. This happens because of the urn effect described above: a "virtuous circle" of success is created starting from an initial positive impact. In other words, the market *hyperbolises objective differences*--or creates them wherever they do not exist or almost exist. In other words, a large proportion of cultural processes have the characteristics of what we could define as *fashions*. This was the view of a classic author like Veblen (1899), who unfortunately had only a slight influence on XXth century sociology.

So, do the dynamics of reciprocal imitation alone explain the production of fashions in human cultures?

4. SIMMEL'S PARADIGM

The only truly important general theory of fashion is an 1895 essay by Georg Simmel, *Die Mode* (Simmel, 1904, 1905, 1957). Simmel says that every fashion is a process which is always unstable and which depends on two contrasting and interacting forces. One force is the impulse of every human being *to imitate* someone else--usually a person who is considered "up", for some reason, as superior or in any case worthy of being imitated. The second force is the impulse everyone has *to distinguish him/herself* from his/her fellows--above all from those perceived, for some reason, as "down", inferior. The relation between the tendency to imitate and the tendency to distinguish oneself varies from one human being to another, but it is rare for one of the two to be totally absent in an individual.

Let us consider a feature of recent fashions, for example wearing short shirts revealing the female navel. This was certainly first done by young women of the higher classes and the most influential cultural sectors, living in big Western metropolises. Then, gradually, women of an increasingly inferior social condition and those who live in less and less central areas imitated this exhibition of the navel. But as this trait propagated and imposed itself--becoming ever more fashionable--it became less and less distinctive: this is why the ladies of a higher and more influential world or the trend-setters who have launched a fashion tend for this very reason to abandon it and pass on to something else. This explains the perennial instability of fashion: its triumph with the masses is tantamount to the digging of its grave. If one says of anything that it is "obligatory fashion", this means that it is already on the wane and due to disappear.

Raymond Boudon (1979) has applied this model also to the field of sophisticated intellectual fashions. A new philosophical idea, let us, for the sake of argument, call it Heraclitism, springs up in one of the prestigious cultural breeding grounds--in the American campuses of the ivy league, in Paris, Oxford or Cambridge, in some German university, or any of the few other truly influential centres of intellectual production in the world. Then, gradually, Heraclitism spreads to the outer provinces of Western culture, followed by Oriental culture. But as it becomes diffused, Heraclitism is appropriated by intellectuals and professors of lesser calibre, less and less brilliant, ever more pedestrian and conformist--and so it becomes the obligatory paradigm taught even in universities of marginal importance. Thus after a few decades the new philosophical elite, instructed in the rules of Heraclitean obedience in one of the above-mentioned great philosophical centres, precisely in order to distinguish themselves from the mass of their colleagues, opt for a rival but less successful theory. Since by now Heraclitism has become commonplace, a way of thinking which is taken for granted and therefore lazy, it is not too difficult for these young lions of the intellect to upset the status quo and promote the alternative philosophy. And so the cycle begins again.

But a point remains which the theory of Simmel and Boudon does not deal with: why is it precisely *that trait*--why exactly the navel *en plein air*--which is imitated and not another? Why is Heraclitism adopted and not another philosophy having its own justificatory arguments and supporters? In effect the foremost avant-garde stylists usually base their work on a stock of ideas: they hope that at least *one* proposal will be imitated and become fashionable. Every ambitious philosopher tries to launch a new and original way of thinking, but in the end only very few of these become hegemonic schools. So, what qualities must a cultural trait possess in order to be a success, even if only an ephemeral one?

5. CULTURAL EVOLUTION HAS NO MEANING

The theory of chaos suggests the following idea to us: it is not necessary for a fashionable trait to have any particular qualities; instead it is enough for it to obey "the dynamics of the ants". Of course it is necessary for some facilitating conditions to be satisfied: that the trait should first of all distinguish the persons who "make the fashion", in other words that they should be in the prestigious position which makes them *elegantiarum arbitri*--or *cogitationum arbitri* (arbiters of concepts). It is also necessary for certain influential media to start acting as a sufficient sounding board. Given these conditions, a cultural product will be able to impose itself, while

another analogous one will quickly disappear. We may consider the case of the philosophical book *Empire* by Negri and Hardt (2000), which seems to give a new lymph to the radical and alternative social tendencies of the sixties: it became a best seller in Italy because it was already a blockbuster in the guiding market of America. This previous success abroad convinced the Italian reviewers to take it seriously, and thus "the process of the urn" was triggered off.

The unpredictability of the fortunes of a product is due to the fact that various negligible and minimal differences at the moment the process begins--the fact that from the beginning a book or a film had a good review, for example--can lead to spectacular differences when it becomes fully developed. We have already seen that the system shows a sensitive dependence on the initial conditions, better known as the *butterfly effect*: "the flap of a butterfly's wings in Brazil can cause a tornado in texas" (see Lorenz, 1979a, 1979b). Modern cultures too, like the weather, are particularly unstable systems: minimal variations can produce dramatic results; while in more stable systems (such as certain archaic societies) even enormous impacts do not manage to disturb the basic equilibrium.

The classical sociologists of culture usually maintain that cultural fashions have deeply rooted sociological motivations. For example, it is said that young people today tattoo themselves in order to overturn a dominant conception which exalts the reversibility of any choice, and the unlimited ability and tendency to change; thus in this way they are polemically affirming their preference for irreversible acts. Take as an example the great vogue for the thought of Popper in Italy in recent years: it is seen as indicating the decline of the totalising theories (such as Marxism) and the growing power of science. By following this line of argument it has even been declared that the periods in which women's dresses become shorter coincide with periods of irrepressible female emancipation! But this is simply not true.

It is all very well to look for the deeper meanings of fashions, both in futile as well as serious fields. But one should ask oneself whether these fashions prevail because they express certain deep tendencies in the social way of being, or if they only seem to express deep tendencies because they happened to prevail in a certain specific period. The relationship between signifiers (the fashionable trait) and signified is much more complex than the classic sociology of culture would lead us to believe.

6. INTELLIGIBLE UNPREDICTABILITY

And so cultural processes are often non-linear and chaotic. A linear process is a classic chain of causes and effects: if I heat some water I know that it will boil at 100° centigrade (212°F). But human culture is a non-linear system and the single changes are thus unpredictable. An order can instead be mapped out over a long period. In other words, the socio-cultural facts show an aleatory and unpredictable tendency when seen in detail: no one knows what women will be wearing in a couple of years. But with time stability is revealed which, far from denying chaotic processes, is like their precipitate. An order disguised as disorder emerges.

However, the claim of being able to predict cultural phenomena in a precise way comes up against some decisive limits. Today various fashion industries devote substantial financing to sophisticated research by social psychologists, hoping to discover a theory which will make fashion in some way predictable. It would be like manna from heaven: these companies could launch a sure-fire winning product every year. But they are just throwing their money away, because the increase of intelligibility that a theory of fashion can give us does not necessarily, *ipso facto*, lead to more predictability. Luckily, that which human beings tend to prefer from one moment to another--whether it be a type of trousers or a philosophical or aesthetical conception--remains mostly unpredictable. Non-linearity is the ontological expression of the liberty of nature, and therefore also of those natural beings who are human beings.

In fact no sociologist has ever truly been able to predict any social macro-phenomenon--and when he has actually got it right, it has almost always been by chance. Thirty years ago who could have predicted the re-awakening of Islam and the jihad, which worries us so much today? Who could have foreseen even in 1965 the explosion of the radical protest movements of only two or three years later? And even as recently as four years ago who would have predicted the no-global vogue? Which American intellectual of the '60s would ever have predicted the hegemony of the deconstructionist tendencies in the American humanistic faculties? And we could continue this list indefinitely.

There is no need to speak at length of economic predictions, which almost always turn out to be wrong. It is a good habit not to read economic predictions even for the coming year and even if they come from the most prestigious sources.

As Ormerod (1994) underlines, even the immense sums paid to the famous financial consultants employed by private clients and companies are simply thrown away: the oscillations of the financial markets are not even chaotic, they are often simply random. This is because the development of human culture (and also of the economy) is indeterminate: at any given moment there is not just one track that a culture (or an economy) could follow, but many.

It would however be a mistake to conclude that this failure of the social sciences to predict the future evolution of such developments is the effect of the scientific backwardness of these sciences. In reality other much "harder" sciences are unable to do much better.

For example, the Darwinian theory of evolution--which is the dominant paradigm in biology--does not allow us to predict the future changes and developments of animal species. Eighty million years ago, in the Cretaceous era, no biologist would have been able to predict the advent of a mammal of medium size called homo sapiens, around 78 million years later. Like the evolution of culture, the evolution of life too is to a large extent unpredictable (on this point, see Pievani, 2003). And as Edward Lorenz has demonstrated, not even meteorology is ever able to predict with much precision. This should be enough to make us very mistrustful of futurologists of whatever tendency or school they may belong to, even if they are paid vast amounts: there is something rotten in the pretension of being able to predict the future, in many fields.

Chaos theory in fact tells us that certain unpredictability is not the effect of our inability to accumulate a mass of information which would allow us to make the prediction, but is an integral part of the non-linear structure of many natural processes.

But fortunately not everything is unpredictable. One should not confuse chaos with pure randomness. Not even the best meteorologist can say if next summer will be hotter or cooler than this summer, but we can all safely bet that the next summer will be hotter than the next winter.

The theory of chaos also shows how order emerges within natural and cultural processes and that there is therefore something predictable: but it underlines the fact that order is basically a form of stable chaos.

The fact that for centuries only women have worn dresses, for example, is probably the effect of stable chaos in the field of clothing habits, and perhaps even the fact that in Western societies the Christian faith continues to prevail is a form of stable chaos.

For example, certain fashions in clothing turn out to be more or less periodical, and therefore they manifest a sort of order. It can be noticed that there is a cyclic variation in the length and the breadth of skirts (Kroeber

even calculated the period involved[4]). But even if we have periods of vertiginously short skirts and periods of extraordinarily long skirts, we can recognise an attractor, which is to say a sort of medium length of skirts in the West, for example in the last 200 years. Moreover, we can suppose with a certain margin of certainty that for the whole 21st century in the West women will continue to wear skirts, while--except for a few Scots--men will not wear them (this is however only a fairly safe bet, not a certainty). This stable order is always provisional and threatened by complexity. We should finally start thinking that we all live *on the edge of chaos*.

For this reason, if they were truly digested, the theories of complexity and chaos could change our way of seeing what happens in our cultures. They lead us to mistrust all the totalising and totalitarian conceptions which have the pretension of telling us with certainty what the world will be like and which therefore supply us with the instruments to dominate as we may please--or to help us submit to those who, in their opinion, will dominate us. Living on the edge of chaos is also an aesthetic choice: the acceptance of living joyously with the unpredictable, the new and the unknown. Rather than being simply the humiliation of our arrogance, this way of thinking gives up the imaginary "regular income" of determinism and the transformation of our uncertainties into a genuine wealth to help us to survive.

7. THE NEED FOR LINEARITY

Today many intellectuals avail themselves of every opportunity to refer to the theories of complexity and chaos. Also the chaotic theory of fashions will soon become, one can safely suppose, a fashion itself. And yet this approach has hitherto been and still remains--also among open-minded and well-informed intellectuals--a dead letter. For example, the overwhelming majority of people, even those with a certain degree of culture, continue to think of political processes in linear terms. I cannot exclude the idea that modern democracies function only on the basis of a *linearist illusion* according to which it is necessary for us to think that policies are either good or bad in absolute terms, and that certain actions of the government or the

4 Kroeber and Richardson (1940). They noticed that the rhythm of change of women's evening gowns "not only is regular (the amplitude was of around half a century, the complete oscillation of a century), but it also tends towards the alternation of the forms according to a rational order: for example the width of the skirt and the width of the waist are always in an inverse relationship: when one is narrow the other is wide". See also Barthes (1970).

Central Bank are the real causes of a certain economic disaster or on the contrary of an economic boom. The supposition that the relationship between the input (political or economical measures) and the output is linear seems to many people a necessary condition for being able to judge political actions. But the theories of chaos teach us that the causal relationships in societies are not linear--that a wise and far-sighted policy in certain contexts can lead to dreadful results, or vice versa. Who could have foreseen, for example, that the application of the free-market doctrine of opening the markets fostered by the International Monetary Fund would have led to excellent results in some countries, while it is leading other countries, especially in Latin-America, towards bankruptcy?

In political and social life we always need to identify someone to be held responsible--or indeed a scapegoat--while in reality there is no linear relationship between certain inputs and the final output. If the chaotic conception of society and politics truly became a part of our mentality, a significant portion of the old categories which still condition our thinking--for example, the fundamental political opposition between the left and the right, or between innovation and conservation--would lose the greater part of their meaning. But the irrepressible need to simplify reality will surely prevail over the disenchanted acceptance of complexity.

REFERENCES

Arthur, B., Ermoliev, Y., and Kaniovski, Y., 1983, A generalised urn problem and its applications, *Kibernetica*.

Barthes, R., 1970, *Sistema della Moda*, Einaudi, Turin, pp. 299-300.

Boudon, R., 1979, *La logique du Social*, Hachette, Paris.

De Vany, A., and Walls, W. D., 1996, Bose-Einstein dynamics and adaptive contracting in the motion-picture industry, *Economic Journal* (November).

De Vany, A., and Walls, W. D., 2003, Quality evaluations and the breakdown of statistical herding in the dynamics of box-office revenue, *Presented at the Annual Meeting of the American Economic Association.* January 2003, Washington DC, USA.

Gould, S. J., 1991, The Technology's Panda's Thumb, in: *Bully for Brontosaurus. Reflections in Natural History*, W.W. Norton & Co., London-New York.

Hardt, M., and Negri, A., 2000, *Empire*, Harvard Univ. Press.

Kirman, A., 1997, in: *The Economy as an Evolving Complex System II*, Arthur, Durlauf and Lane, Santa Fe Institute, Addison-Wesley.

Kroeber, A. L., and Richardson, J., 1940, *Three Centuries of Women's Dress Fashion*, Univ. of California, Berkeley and Los Angeles.

Lorenz, E., 1979a, *Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?*, American Association for the Advancement of Science.

Lorenz, E., 1979b, On the prevalence of a periodicity in simple systems, in: *Global Analysis*, Mgremela and Marsden, eds., Springer, New York.

Ormerod, P., 1994, *The Death of Economics*, Faber and Faber, London.

Pievani, T., 2003, The Contingent Subject. For a Radical Theory of Emergence in Developmental Processes, *Journal of European psychoanalysis* **17**(Summer-Winter).
Ormerod, P., 1998, *Butterfly Economics*, Pantheon Books, London.
Simmel, G., 1904, Fashion, *International Quarterly* **10**(1):130-155,(Eng. Tr.).
Simmel, G., 1905, *Philosophie der Mode*, Pan-Verlag, Berlin.
Simmel, G., 1957, Fashion, *American Journal of Sociology* **62**.
Veblen, T., 1899, *The Theory of the Leisure Class*, (Penguin Books, 1994).

3

COGNITIVE SCIENCE

PERSONALITY AND COMPLEX SYSTEMS. AN EXPANDED VIEW

Mauro Meleddu and Laura Francesca Scalas
Department of Psychology, University of Cagliari

Abstract: Nowadays, dynamic S-P-R models, developed within personality research, seem to offer the chance for the development of a unitary research framework, one including individual and environmental factors, from both structural and dynamic levels of personality. Complexity concept is a fundamental element of convergence of the different existing theoretical and methodological approaches. From this expanded view, personality appears as an "hyper-complex" system of intra-individual circular interactions, and of recursive relations between internal and external factors, that develops and grows self-organizing. This approach takes into account, within a wide interpretative framework, the contributions from various humanistic and scientific disciplines.

Key words: personality; complex systems; chaos.

1. INTRODUCTION

The study of personality is a research field that aspires to recompose, in a unitary framework, individual, affective, cognitive, behavioural, biological and environmental components (Hall and Lindzey, 1978; Mischel, 1993; Pervin and John, 1997). The extent of the subject has led to a wide theoretical and methodological debate (cf. Bem, 1983; Caprara and Van Heck, 1992a; Endler, 1983; Mischel 1968; Pervin, 1990a). Furthermore, the high number of approaches have raised problems concerning its validity and the relation with other scientific disciplines.

For a long time, there was a strong contrast between two major perspectives: one oriented to *internal factors*, the other oriented to *situations*. The first includes the trait-types approaches (e.g. Allport, 1965; Cattell, 1950; Eysenck, 1953; Guilford, 1959; McCrae and Costa, 1996; Murray,

1938; Vernon, 1950) and psychoanalysis. The second perspective is represented by the developments of behaviourism within the interactionism of Mead and social learning approach (e.g. Bandura, 1971, 1977a, 1986, Mischel, 1973, 1977, 1990). The trait-type theory provides a valid description of the intra-individual structure, but neglects dynamics (Epstein, 1994); whereas, psychoanalytic perspective examines deep dynamics, but neglects situations. Situationism tends to focus on processes, but ignores the stability of individual factors and their predictive value (Funder, 1991, 1994).

Recent findings demonstrating the stability of personality basic components, achieved also with the contribution of neuroscience's research, has paved the way to the reconciliation between the trait (dispositional) theory and social learning approach (cf. Mischel and Shoda, 1998). In particular, the transition from the neo-behaviourist formula *S-O-R* to the *S-P-R* models has contributed to this development. In *S-P-R* models the relationship between the situation *S* and the response *R* is mediated by personality *P*, instead of organism *O* (cf. Fraisse, 1968-1969). *P* represents an element of the behavioural organization of the person, which includes traits, biological bases, and neuropsychological processes and their cognitive, emotional and motivational components. Moreover, in the last few decades, the common reference to *complexity concept* (Caprara and Van Heck, 1992b; Pervin, 1990b) has contributed to give more coherence to the field.

Circular dynamic interactions have a central role in advanced *S-P-R* models. Such kind of relations do not permit the usual separation between dependent and independent variables (Raush, 1977), in contrast to the traditional research formula $y = f(x)$. Circular dynamic interactions can determine different solutions, may become non-linear, and can activate growing developmental processes in terms of complex systems. That trend can be described by sequences of calculi according to the general recursive formula: $x_{n+1} = f_c(x_n)$. This requires the introduction of a research paradigm that exceeds the deterministic idea at the basis of the classical science and experimental model.

2. FLUCTUATIONS AND DETERMINISTIC CHAOS

Fluctuation theory allows the examination of non-linear phenomena and the description of the macroscopic order development, as a consequence of microscopic, disordered and chaotic fluctuations, inside open processes in conditions far from equilibrium (Prigogine et al., 1977; Nicolis and Prigogine, 1977). This theoretical approach has provided an essential

contribution to knowledge unification, and exceeds the deterministic conceptions of classical science (Prigogine and Stengers, 1979).

Nowadays, the study of chaos considers nature as a global complex system that develops itself alternating deterministic phases (conditions of relative macro-structural equilibrium) and stochastic periods (conditions of high micro-structural instability). Overall, the model includes physical, chemical, genetic, biological, social and mental factors (Eigen,1987, 1988; Gleick, 1987; Haken, 1981). From this point of view, fluctuation theory can be coherently applied to a wide number of problems, ranging from physical, biological and social phenomena (Prigogine, 1972; Prigogine et al., 1972) to psychological aspects (cf. Masterpasqua and Perna, 1997). Later, we will examine the possibility to extend this approach to complex *S-P-R* dynamic models.

3. THE S-P-R INTERACTIVE DYNAMIC MODELS

The development of the *S-P-R* paradigm into *dynamic models* gives to individual factors an active, finalistic and constructive role that can be represented as in fig. 1.

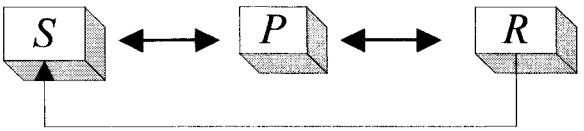


Figure 1. S-P-R dynamic model.

The double arrows represent the mutual interaction of personality with both external stimuli and responses. Furthermore, the direct link between *R* and *S* involves a causal relationship between behavioural responses and stimuli. Thus, the schema shows that individual processes can also determine modifications in the physical environment, and activate a process of reciprocal transformation between the person and the environmental stimuli.

On the same line, the personality *interactive model*, developed by Endler (Endler and Magnusson, 1976; Magnusson and Endler, 1977), considers the circular process that takes place in the person-situation interactive relationship. The dynamic model includes interpersonal relations, where each personal response constitutes a recursive stimulus for other persons (cf. Endler, 1983).

The *Cognitive-Affective Personality System (CAPS)* (Mischel and Shoda, 1998) extends the circular interactions to the different domains of *P*, that is to the reciprocal influence between the organic and psychological levels, and between the different parts of this latter level (conscious, unconscious, cognitive, affective, motivational). Traits and different processes, such as cognitive and affective, can be considered inside a unitary interpretative framework (Mischel and Shoda, 1995). The whole system has an *active-proactive* nature and includes *self-perceptions* and *expectancies*. These activate internal feedback loops that work as schemas or dispositions for behavioural direction. The model considers self-processes, such as "*self-efficacy*" (Bandura, 1977b), and other cognitive constructs such as "*expectancy*" (Mischel, 1973) and "*locus of control*" (Rotter, 1975). Furthermore, the model considers traits as directly connected to affective-cognitive processes and to their biological bases. This approach allows to explain more adequately behaviour as a function of traits, self-perception and situational variability (Chiu et al., 1995). In fig. 2 is represented the circular interconnection among the different parts of the system.

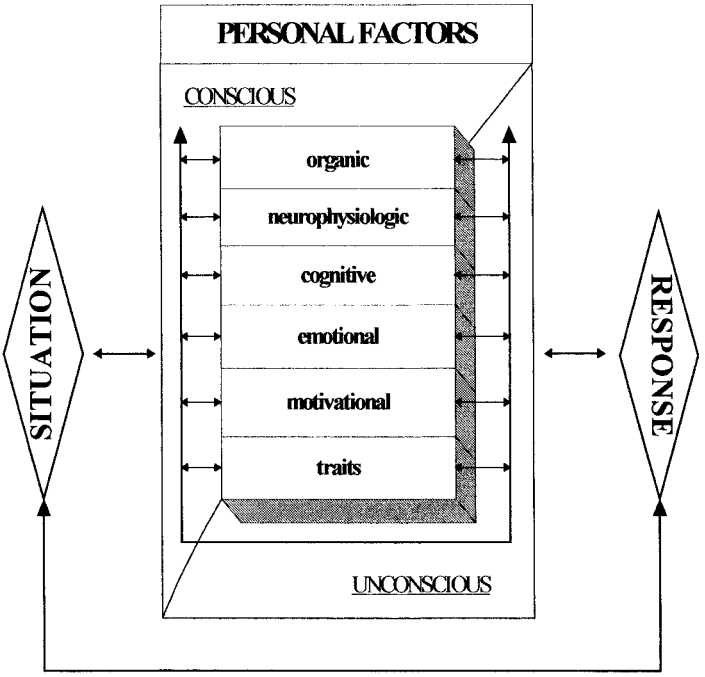


Figure 2. The *CAPS* model.

In general, the *CAPS* model is a further dynamic extension of the *S-P-R* model. It permits to go beyond the person-situation contrast and the prevailing conception of the *cognitivist revolution* of the '60s and '70s (Miller et al., 1960; Neisser, 1967).

The examination of the circular interactions, within fluctuations theory, allows another extension of the *S-P-R* models, and makes it possible to explain the relationship between stability and change in personality. From this point of view, personality organizes itself, according to the same modalities of complex systems, through the improvement and stabilization of casual juxtapositions results.

4. THE "STATE SPACE" MODEL

The "*state space*" model (cf. Lewis, 1995; Lewis and Junk, 1997) is a recent example of the fluctuations theory application to emotional and cognitive processes of human development. The *state space* refers to all possible "states" that a system can attain: stable, unstable and transitory. The model considers human behaviour as stochastic, unpredictable and determined by multiple causes. In particular, the model involves uncertainty and predictability, in fact, behaviour is thought to be caused by chaotic fluctuations and can alternate unpredictable and predictable phases.

Fundamental elements of the system are "attractors" and "repellors". Attractors represent states of self-organization and relative stability toward which the system moves from one or more other states. Repellors are unstable and transient states from which the system tends to move away.

In personality, attractors take the place of traits properties, but depend on individual goals and on the whole situational factors. Attractors are constellations of cognitive, emotional and behavioural elements. These constellations may have a global (e.g. global anxiety) or content-specific (e.g. fear for a specific situation) nature. Repellors, on the other hand, are constellations of rare or transient behaviours and internal states. They include states that the system tend to avoid, such as distress.

Attractors and repellors determine change and stability conditions with the contribution of chaos. Change and stability alternate. There are two types of changes: "micro-developments" and "macro-developments" (Lewis, 1995). They relate respectively to the "weakly" and "strongly" self-organizing trajectories that characterize complex systems (Haken, 1977; Prigogine and Stengers, 1984). The first concerns momentary adaptive responses. Moreover, micro-developments represent the movement from one condition to another accessible at the same time in the system; such as the transition between two attractors represented, for example, by two ideas or

two different emotional states. Macro-developments regard the formation of new and more stable conditions or phases. They produce transitions between cognitive stages or personality shifts.

Personality stability and change depend on cognitive-affective interactions, according to recursive development of chaotic fluctuations, that characterize the activity of attractors and repellors. During the initial phases of both micro-developments and macro-developments, the system becomes sensitive to chaotic fluctuations that may transform the habitual configuration and produce a new organization.

The explanatory model of psychological self-organization processes, developed by Lewis (1995, 1996, 1997), focuses on two sets of constituents: *cognitive elements* (e.g. concepts, scripts, beliefs, images) and *emotions* (e.g. anxiety, shame, sadness, anger). Inside the *S-P-R* models, it is possible to consider the relation between these two components as a recursive, self-enhancing and self-regulating interaction within the person (cf. Lewis, 1995). The process can be represented as in fig. 3.

Cognitive interpretation, or "*appraisal*", of events takes place continuously and immediately. The process mostly involves the unconscious level (cf. LeDoux, 1996). Cognitive appraisal of a situation is automatically associated to an emotion (cf. Beck, 1976; Scherer, 1984; Stein and Tabasso, 1992). Coupling determines a cognitive adjustment, that origins a new emotional interpretation of the situation. Between cognitive and emotional components takes place a feedback loop that produces a global interpretative process of the events. Thus, during this process, a recursive dynamic develops.

This dynamic enhances and stabilizes itself, and gives rise to the consistent and habitual patterns that constitute personality. Self-organization takes place rapidly in micro-development through the coupling process. The consolidation of these connections, in terms of tendencies or habits, self-organizes in macro-development.

This recursive dynamic approach considers behaviour as a special case of cognitive adjustment following from emotional evaluation (Lewis and Junk, 1997). In emotional relevant situations, cognitive appraisals give rise to specific emotional reactions. These reactions activate cognitive attention on situational features and determine adaptive behavioural responses. Cognitive changes following from behaviour influence recursively the circular relation with emotions. The loop self-enhances until the system stabilizes itself. The stabilization consists in coherence between the micro-development elements such as images, concepts, associations and memories, that give a meaning to the situation. At the same time, the recursive interaction between micro-development and macro-development elements is already active.

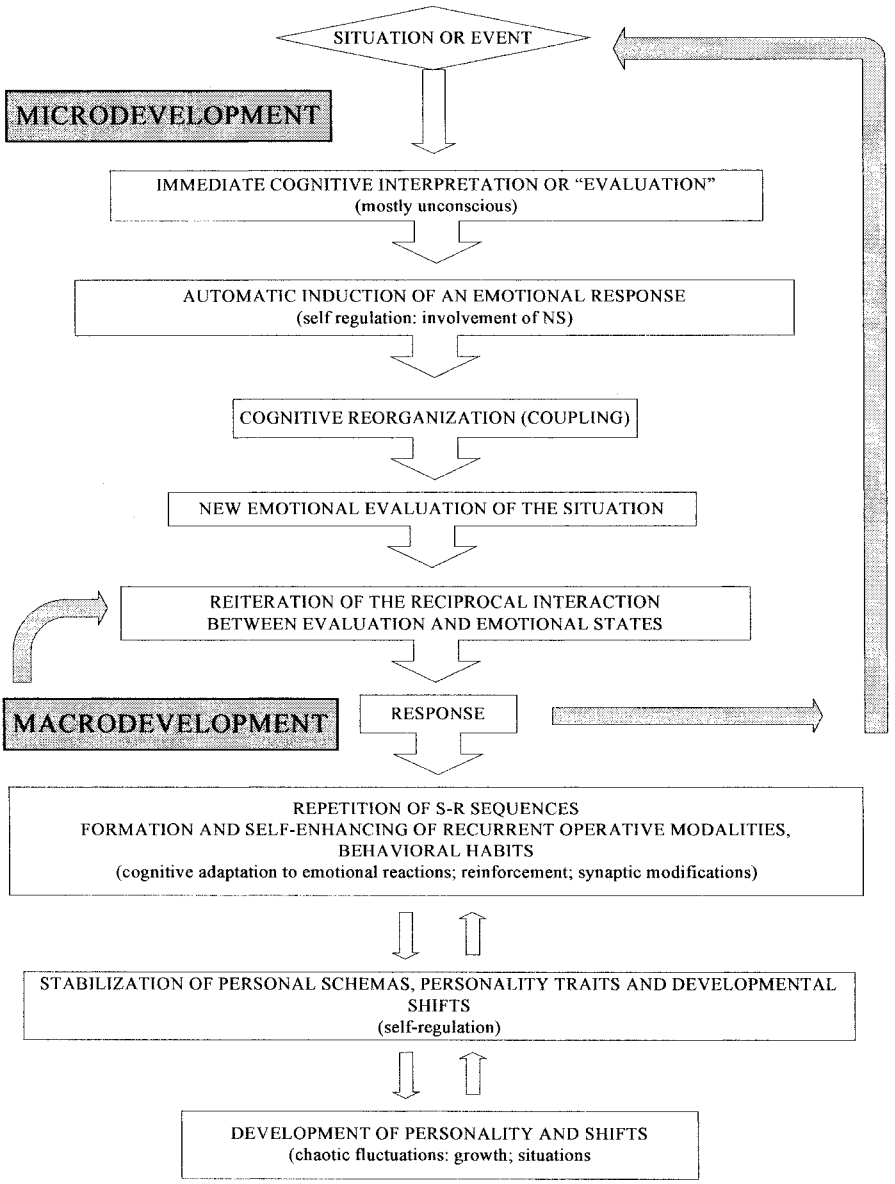


Figure 3. The *state space* model.

The reciprocal adjustment between cognitive elements, coupled in micro-developments, supports the stability of the global interpretation and the circular activation of emotions. The emotional feedback favours cognitive interaction reinforcement, consolidates the global interpretation and, at the same time, maintains the emotional activation. Thus, cognitive evaluations and emotions activate and fuel each other in a recursive, self-correcting and self-stabilizing relationship.

Emotional interpretations self-organize in real time. Moreover, personality self-organization can be considered as a process in which attractors development crystallize over the life span in macro-development. Attractors consolidation specifies the repertoire of habitual responses in social interactions. During the development of the interpretative process, the coupling between cognitive and emotional elements produces microscopic changes. These modifications influence the coupling process on the subsequent interpretative phase, and give rise to changes in relations and structures. At the biological level, these variations involve synapses.

In personality development, the consolidation of cognitive-emotional attractors does not proceed in a linear way, but is marked by periods of rapid changes and reorganizations. According to the self-organizing shifts of complex systems, transitional phases of personality growth follow a branching path (Lewis, 1995). Chaotic fluctuations, linked to maturational and environmental changes, influence the circular interaction between cognitive and emotional elements. The self-amplification of fluctuations has the potential to produce new couplings between cognitive and emotional activities. This process gives rise to developmental bifurcations, that correspond to personality changes. Every transitional phase is characterized by a recursive loop, and is affected by the previous stage, under the influence of recurrent emotions. Consequently, during development, orderliness prevails and gives continuity to personality. Orderliness, however, is dynamic and unpredictable on the basis of situational context, in fact the development of new schemas is linked to casual fluctuations.

5. BIOLOGICAL BASES AND COMMUNICATION PROCESSES

In general, the *CAPS* and *State space* models are in line with different studies that highlight the link between traits and their dynamic affective-cognitive processes (cf. Bolger and Zuckerman, 1995; Mischel and Shoda, 1995). These processes have specific neural bases (Bates and Wachs, 1994; LeDoux, 1994, 1996; Zuckerman, 1991) and interact inside the brain through connections more complex than it was thought in the past. The psychological

level is not completely referable to neural properties (Gazzaniga, 1992; Sperry, 1982). However, it is possible that, through these connections, circuits take place and, according to the models here on examination, give rise to the formation and development of behavioural responses and different, mostly stable, mental schemas. In the *state space* model, as we have seen before, traits stability prevails because recursive fluctuations of transitions are influenced by the previous organization of the system. With reference to biological bases, it is important to consider that the various parts of NS have diverse hereditary foundations: for example, the sub cortical connections, associated to a more rigid genetic programming, are tendentially more stable in respect of cortical connections. Thus, it is possible to assume that the different plasticity of encephalic structures, involved in behavioural responses, influences the person stability. From this point of view, it is possible to justify the structural properties of general dimensions of personality such as *extroversion* and *neuroticism*. In Eysenck's model (1953, 1967), these dimensions have different biological basis: respectively the *ARAS* and the *limbic system*. Anyway the development and stabilization of specific responses (e.g. extraverted or neurotic) is influenced by the interaction with learning processes.

For the approaches here in consideration, dynamic models imply communication processes, internal and external to the individual.

On the biological level, Eigen (1987, 1988) considers evolution as the result of casual mutations determined by errors during the replication process of nucleic acids. Natural selection influences the propagation of informational reading errors. The process presents phase drops, which develop in conditions far from equilibrium characterized by microscopic fluctuations. Interpersonal relations, individual and socio-cultural development, build on a shared communication system. Social interaction has a symbolic nature (cf. Mead, 1934). Thus, within the social system, responses involve the meaning attributed to the stimuli, rather than their physical properties. The attribution of meaning to events and the creation of links between contexts, give rise to a self-enhancing process that assumes the characteristics of "knowledge of knowledge" (Morin, 1986). The individual evaluates situations through the *event-self-context interaction*. From this point of view, the self-concept can be considered as the higher level of integration of biological, cognitive, affective, behavioural, symbolic and social processes. These processes influence each other recursively at different levels (cf. Harter, 1996; Le Doux, 2002; Mead, 1925; Morin, 1991).

6. CONCLUSIONS

The recent development of *S-P-R* models allows to consider a wide range of dynamic and multi-factorial relations, including linear, non-linear and circular interactions. Recursive relations can produce developments, in which deterministic phases alternate to stochastic periods, and can produce results not easy to predict and not referable to a single cause. Moreover, the reference to complexity concept permits to reconcile, in a wider dimension, determinism and teleology, as well as other contrasts such as that between randomness and necessity, and between life sciences and nature sciences. From this point of view, personality can be seen as a complex open system that refers to *instability*, *disorder*, *dynamic relations between systems*, *equifinality*, *equipotentiality*, *non linearity*, *recursivity*, *finalism* and *self-organization* concepts. This approach includes communication processes and the circular interaction with the context, that highlight the observer functions of "interference" and creation of meanings. The self-concept takes the integration role of the various individual and environmental processes that influence each other circularly at different levels. This underlines the complex interrelations existing among different parts of the same system, and between the individual and other systems. Thus, personality can be seen as an "hyper-complex" system that grows self-organizing, according to the interactive processes between its internal factors and the environment.

This approach considers contribution from different psychological currents and various disciplines such as mathematics, physics, chemistry, biology, genetic, neurophysiology, epistemology, and social sciences. The integration of the different perspectives has a fundamental importance within personality research (Eysenck, 1997; Houts et al., 1986; Pervin, 1990b), and requires an adequate methodological framework inside the scientific research canons, that justifies the accumulation of knowledge and the debate between different approaches and disciplines.

This openness does not involve an anarchic (cf. Feyerabend, 1975) or eclectic conception (Caprara and Cervone, 2000), but refers to the epistemological tradition of self-corrective science (Kuhn, 1962; Lakatos, 1978; Laudan, 1977; Popper, 1959) and of critic realism (Bhaskar, 1975; Manicas and Secord, 1983). It is possible to frame this openness into a pluralistic complementary approach that takes into account the interference between the observer and the observed (Heisenberg, 1930), and the constructive function of the knower (Morin, 1986, 1991). From this point of view, it is possible to consider the different disciplines and their approaches as open knowledge systems, characterized by internal coherence, that grow up recursively and permit information exchange on the basis of shared hypothetical-deductive criteria.

REFERENCES

Allport, G. W., 1965, *Pattern and Growth in Personality*, Holt, Rinehart and Winston, New York.

Bandura, A., ed., 1971, *Psychological Modeling: Conflicting Theories*, Aldine-Atherton, New York.

Bandura, A., 1977a, *Social Learning Theory*, Prentice-Hall, Englewood Cliffs, NJ.

Bandura, A., 1977b, Self-efficacy: Toward a unifying theory of behavioral change, *Psychological Review* **84**:191-215.

Bandura, A., 1986, *Social Foundations of Thought and Action*, Prentice Hall, Englewood Cliffs, NJ.

Bates, J. E., and Wachs, T. D., ed., 1994, *Temperament: Individual Differences at the Interface of Biology and Behavior*, American Psychological Association, Washington, DC.

Beck, A. T., 1976, *Cognitive Therapy and Emotional Disorders*, International University Press, New York.

Bem, D. J., 1983, Constructing a theory of the triple typology: Some (second) thoughts on nomothetic and idiographic approaches to personality, *Journal of Personality* **51**:566-577.

Bhaskar, R., 1975, *A Realistic Theory of Science*, Leeds Books, Leeds.

Bolger, N., and Zuckerman, A., 1995, A framework for studying personality in the stress process, *Journal of Personality and Social Psychology* **69**:890-902.

Caprara, G. V., and Cervone, D., 2000, *Personality: Determinants, Dynamics, and Potentials*, Cambridge University Press, Cambridge UK.

Caprara, G. V., and Van Heck, G. L., 1992a, Personality psychology: some epistemological assertions and historical considerations, in: *Modern Personality Psychology. Critical Review and New Directions*, G. V. Caprara and G. L. Van Heck, eds., Harvester Wheatsheaf, London.

Caprara, G. V., and Van Heck, G. L., 1992b, Future prospects, in: *Modern Personality Psychology. Critical Review and New Directions*. G. V. Caprara and G. L. Van Heck, eds., Harvester Wheatsheaf, London.

Cattell, R. B., 1950, *Personality*, McGraw-Hill, New York.

Chiu, C., Hong, Y., Mischel, W., and Shoda, Y., 1995, Discriminative facility in social competence: Conditional versus dispositional encoding and monitoring-blunting of information, *Social Cognition* **13**:49-70.

Eigen, M., 1987, *Stufen zum Leben*, Piper GMBH, München.

Eigen, M., 1988, *Perspectiven der Wissenshcaft*, Deutsche Verlags-Anstalt GMBH, Stuttgart.

Endler, N. S., 1983, Interactionism: a personality model, but not yet a theory, in: *Nebraska Symposium on Motivation, 1982: Personality - Current Theory and Research*, R. A. Dienstbier and M. M. Page, eds., University of Nebraska Press, Lincoln, NE.

Endler, N. S., and Magnusson, D., 1976, Toward an interactional psychology of personality, *Psychological Bulletin* **83**:956-974.

Epstein, S., 1994, Trait theory as personality theory: Can a part be as great as the whole?, *Psychological Inquiry* **5**:120-122.

Eysenck, H. J., 1953, *The Structure of Human Personality*, Methuen, Wiley, London-New York.

Eysenck, H. J., 1967, *The Biological Basis of Personality*, C. C. Thomas, Springfield, IL.

Eysenck, H. J., 1997, Personality and experimental psychology: The unification of psychology and the possibility of a paradigm, *Journal of Personality and Social Psychology* **73**:1224-1237.

Feyerabend, P., 1975, *Against Method: Outline of an Anarchist Theory of Knowledge*, New Left Book, , London.

Fraisse, P., 1968-1969, Modèles pour une histoire de la psychologie, *Bulletin de psychologie*, **276**:9-13.

Funder, D. C., 1991, Global traits: a neo-allportian approach to personality, *Psychological Science* **2**:31-39.

Funder, D. C., 1994, Explaining traits, *Psychological Inquiry* **5**:125-127.

Gazzaniga, M. S., 1992, *Nature's Mind: The Biological Roots of Thinking, Emotions, Sexuality, Language and Intelligence*, Basic Books, New York.

Gleick, J., 1987, *Chaos*, Viking Penguin, New York.

Guilford, J. P., 1959, *Personality*, McGrow-Hill, New York.

Haken, H., 1977, *Synergetics. An Introduction: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry and Biology*, Springer-Verlag, Berlin.

Hall, C. S., and Lindzey, G., 1978, *Theories of Personality*, Wiley, New York.

Harter, S., 1996, Historical roots of contemporary issues involving self-concept, in: *Handbook of Self-Concept*, B. Bracken, ed., Wiley, New York.

Heisenberg, W., 1930, *The physical principles of the quantum theories*, University of Chicago Press, Chicago.

Houts, A., Cook, T., and Shadish, W., 1986, The person-situation debate: A critical multiplist perspective, *Journal of Personality* **54**:52-105.

Kuhn, T. S., 1962,*The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

Lakatos, I., 1978, *The Methodology of Scientific Research Programs*, Cambridge University Press, Cambridge, MA.

Laudan, I,. 1977, *Progress and Its Problems. Towards a Theory of Scientific Growth*, University of California Press, Berkeley, CA.

LeDoux, J., 1994, Cognitive-emotional interactions in the brain, in: *The Nature of Emotions: Fundamental Questions*, P. Eckman, and R. J. Davidson, eds., Oxford University Press, New York.

LeDoux, J., 1996, *The Emotional Brain. The Mysterious Underpinning of Emotional Life*, Simon & Shuster, New York.

LeDoux, J., 2002, *Synaptic Self: How Our Brains Become Who We Are*, Viking Penguins, New York.

Lewis, M. D., 1995, Cognition-emotion feedback and the self-organization of developmental paths, *Human Development* **38**:71-102.

Lewis, M. D., 1996, Self-organising cognitive appraisals, *Cognition and Emotion* **10**:1-25.

Lewis, M. D., 1997, Personality self-organization: Cascading constraints on cognition-emotion interaction, in: *Dynamics and Indeterminism in Developmental and Social Processes*, A. Forgel, M. C. Lyra, and J. Valsiner, eds., Erlbaum, Mahwah, NJ.

Lewis, M. D., and Junk, N., 1997, The self-organization of psychological defenses, in: *The Psychological Meaning of Chaos*, F. Masterpasqua, and P. A. Perna, eds., American Psychological Association, Washington, DC.

Magnusson, D., and Endler, N. S., 1977, Interactional psychology: Present status and future prospects, in: *Personality at The Crossroads: Current Issues in Interactional Psychology*, D.Magnusson, and N. S. Endler, eds., Erlbaum, Hillsdale, NJ.

Manicas, P. T., and Secord, P. F., 1983, Implications for psychology of the new philosophy of science, *American Psychology* **4**:399-413.

Masterpasqua, F., and Perna, P. A., eds., 1997, *The Psychological Meaning of Chaos*, American Psychological Association, Washington, DC.

McCrae, R. R., and Costa, P. T., 1996, Toward a new generation of personality theories: Theoretical contexts for the five-factors model, in: *The Five-Factor Model of Personality. Theoretical Perspectives*, J. S. Wiggins, ed., Guilford, New York.

Mead, G.H., 1925, The genesis of the self and social control, *International Journal of Ethics* **35**:251-73.

Mead, G. H., 1934, *Mind, Self and Society, From the Standpoint of a Social Behaviourist*, University of Chicago Press, Chicago.

Miller, G. A., Galanter, E., and Pribram, K. H., 1960, *Plans and the Structure of Behavior*, Holt, New York.

Mischel, W., 1968, *Personality and Assessment*, Wiley, New York.

Mischel, W., 1973, Toward a cognitive social learning reconceptualization of personality, *Psychological Review* **80**:252-283.

Mischel, W., 1977, The interaction of person and situation, in: *Personality at The Crossroad: Current Issues in Interactional Psychology*, D. Magnusson, and N. S. Endler, eds., Erlbaum, Hillsdale, NJ.

Mischel, W., 1990, Personality dispositions revisited and revised: A view after three decades, in: *Handbook of Personality: Theory and Research*, L. A. Pervin, ed., Guilford, New York.

Mischel, 1993, *Introduction to Personality*, Holt, Rinehart and Winston, New York.

Mischel, W., and Shoda, Y., 1995, A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure, *Psychological Review* **102**:246-268.

Mischel, W., and Shoda ,Y., 1998, Reconciling processing dynamics and personality dispositions, *Annual Review of Psychology* **49**:229-258.

Morin, E., 1986, *La Méthode. Vol. III. La Connaissance de la Connaissance*, Seuil, Paris.

Morin, E., 1991, *La Méthode. Vol. IV. Les Idées. Leur Habitat, leur Vie, leur Moeurs, leur Organisation*, Seuil, Paris.

Murray, H. A., 1938, *Explorations in Personality*, Oxford University Press, New York.

Neisser, U., 1967, *Cognitive Psychology*, Appleton-Century-Crofts, New York.

Nicolis, G., and Prigogine, I., 1977, *Self-Organization in Non-Equilibrium Systems*, Wiley, New York.

Pervin, L. A., 1990a, A brief history of modern personality theory, in: *Handbook of Personality: Theory and Research*, L. A. Pervin, ed., Guilford, New York.

Pervin, L. A., 1990b, Personality theory and research: Prospects for the future, in: *Handbook of Personality: Theory and Research*, L. A. Pervin, ed., Guilford, New York.

Pervin, L. A., and John, O. P., 1997, *Personality. Theory and Research*, Wiley, New York.

Popper, K. R., 1959, *The Logic of Scientific Discovery*, Hutchinson, London.

Prigogine, I., 1972, La thermodynamique de la vie, *La Recherche* **3**:547-562.

Prigogine, I., Allen, P. M., and Herman, R., 1977, The evolution of complexity and the laws of nature, in: *Goals in a Global Community, a Report to the Club of Rome. Volume I Studies on the Conceptual Foundations*, E. Laszlo, and J. Bierman, eds., Pergamon Press, New York.

Prigogine, I., Nicolis, G., and Babloyantz, A., 1972, Thermodynamics of evolution (I). *Physics Today* **25**(11):23-28.

Prigogine, I., Nicolis, G., and Babloyantz, A., 1972, Thermodynamics of evolution (II). *Physics Today* **25**(12):38-44.

Prigogine, I., and Stengers, I., 1979, *La Nouvelle Alliance. Métamorphose de la Science*, Gallimard, Paris.

Prigogine, I., and Stengers, I., 1984, Order Out of Chaos, Bantam, New York.

Raush, H. L., 1977, Paradox, levels, and junctures in person-situation systems, in: *Personality at The Crossroads: Current Issues in Interactional Psychology*, D. Magnusson, and N. S. Endler, eds., Erlbaum, Hillsdale, NJ.

Rotter, J. B., 1975, Some problems and misconceptions related to the construct of internal versus external control of reinforcement, *Journal of Consulting and Clinical Psychology* **43**:56-67.

Scherer, K. R., 1984, On the nature and function of emotions: A component process approach, in: *Approaches to Emotions*, K. R Scherer , and P. Ekman, eds., Erlbaum, Hillsdale, NJ.

Sperry, R. W., 1982, Some effects of disconnecting the cerebral hemispheres, *Science* **217**:1223-1226.

Stein, N. L., and Tabasso, T., 1992, The organization of emotional experience: Creating links among emotion, thinking, language, and intentional action, *Cognition and Emotion* **6**:225-244.

Vernon, P. E., 1950, The Structure of Human Abilities, Methuen, London.

Zuckerman, M., 1991, *Psychobiology of Personality*, Cambridge University Press, New York.

COMPLEXITY AND PATERNALISM

Paolo Ramazzotti
Dipartimento di Istituzioni Economiche e Finanziarie, Università di Macerata
Via Crescimbeni 20, 62100 Macerata, Italy
email: ramazzotti@unimc.it

Abstract: The aim of the paper is to assess the features of public policy in a complex environment. The point of departure is provided by a number of recent papers by David Colander where he argues that progress in mathematics and computational technology allows scholars and policymakers to grasp features of economic reality that, up to some time ago, were beyond their reach. Since the technical difficulties associated to these new tools hardly allow single individuals to use them, Colander suggests that there is scope for public intervention. This intervention need not preclude individual freedom. He refers to it as "libertarian paternalism". The paper argues that Colander focuses on first order complexity, which is associated to economic dynamics, but neglects second order complexity, which relates to cognitive processes. Cognition implies that actors can formulate their choices only by learning, i.e. by constructing appropriate knowledge contexts. This requires appropriate public action in order to prevent the establishment of restrictive knowledge contexts. In turn, this implies a "democratic paternalism" that is markedly different from the paternalism Colander refers to.

Key words: complexity, knowledge, paternalism, public policy, choice

1. INTRODUCTION

The aim of the paper is to discuss paternalistic economic policy in relation to complexity. It attempts to do so by examining how complexity affects economic inquiry as a whole. The point of departure is a number of papers written by David Colander, where he contends that progress in mathematics and computational technology provides a new outlook on economic policy, making it reasonable to advocate "libertarian paternalism".

Colander focuses on complexity in economic dynamics. I argue, however, that he neglects a range of issues associated to complexity in cognition. My contention is that both individuals and policy makers do not only need to make choices; they also need to choose how to choose. The discretionary nature of these choices reasserts the value-laden nature not only of economic policy but of the economic inquiry that underlies it. When these issues are taken into account, Colander's view of public policy and paternalism turns out to be too simple and a different notion of paternalism is required.

The paper is structured as follows. Following a brief outline of Colander's views I discuss a few features of cognitive complexity in order to point to the shortcomings of libertarian paternalism. I then introduce the notion of a knowledge context and argue that it can be affected by the purposive action of vested interests. Finally I outline the characteristics of two possible forms of paternalistic policy.

2. COLANDER ON COMPLEXITY

In a range of fairly recent papers David Colander (Brock and Colander, 2000, Colander 2003a, Colander 2003b, Colander, Holt and Rosser, 2003) has been arguing that the "Complexity Revolution" is leading to an overall different view of the economy and of public policy. This change is not easy to see although it is extremely important: "Policy economists, and sophisticated economic theorists, quickly learn that the Walrasian general equilibrium worldview is, at best, a first step to developing a useful worldview of the economy, (...). Recognizing this, they develop a more sophisticated worldview incorporating real-world insights and assumptions as well as modifications of general equilibrium theory, game theory and mechanism design theory. Unfortunately, that more sophisticated worldview is often ambiguous and undeveloped, since developing it formally is an enormous task." (Brock and Colander, 2000: 76-77).

According to Colander, neoclassical textbook theory is still based on what he defines the holy trinity: rationality, greed and equilibrium. His view is that advanced theory based on these assumptions has led us to a dead end in terms of its policy implications. Microeconomic theory is totally disjointed from heuristically based policy suggestions. Macroeconomic theory is consistent with microfoundations but completely inadequate to cope with real world problems. Consequently, young economists are switching to a more flexible, behaviorally grounded, trinity: purposeful behavior, enlightened self interest and sustainability.

The new assumptions do not lead to straightforward conclusions but they do allow simulation. Thanks to computing technology, researchers can do away with analytical constraints such as equilibrium requirements while taking into account a range of possible non-linear patterns such as chaotic ones. Thus, the whole approach to policy is changed. Rather than using theory, one only has to model a situation and figure out what the outcome is going to be. Since the results generally are not analytically neat, in that they rely on trial and error rather than assuming consistency beforehand, Colander refers to this as a "muddling through" approach to policy.

Although it is not rewarding in terms of deductive theory, Colander argues that "muddling through" provides for more relevant policy analysis. The switch in the assumptions and in the techniques is consistent with a broader view of the economy, which "will evolve from its previous vision of highly complex, 'simple system' to a highly complex 'complex system'" (Colander, 2003b: 7). Note that this switch does not extend the scope of previously existing theory: it actually changes its premises. First, simulations are not designed to solve equations but to "gain insight into the likelyhood of certain outcomes, and of the self-organized patterns that emerge from the model." (Colander, 2003a: 11). "[E]quations describing the aggregate movement of the economy" are dispensed with: "one simply defines the range and decision processes of the individual actors" (*ibid.*). Second, in so far as complex systems involve emergent properties, the total is not just the sum of its parts. Thus, outcomes cannot be merely deduced from microfoundations. The latter may exist but they "can only be understood in reference to the existing system" (*ibid.*: 7).

This overall shift in the approach to policy analysis leads Colander to argue in favor of paternalism. Indeed, since agents are only boundedly rational, they are likely not to be fully aware of what they want. Furthermore, owing to emergent properties, agents may be unable to perceive the direct – let alone the indirect - consequences of their actions. Finally, since the techniques required to outline future scenarios are rather sophisticated, even if agents had that information, they might be unable to compute it. Governments may be better equipped to figure out what the future could be and to envisage what appropriate conducts should be.

Colander's view points out how changes in the available techniques are providing a unique path towards economic policy. The latter involves value judgements but the economics underlying it are just a technical issue[5]. Different views of the world are irrelevant. All you need to do is model a

[5] "A [...] change brought about by complexity is that it adds a theoretical neutrality to the abstract debate about policy." (Brock and Colander, 2000: 79).

problem. In the following section I will point to a few features of problem solving. I will argue that there is more than technical issues involved.

3. FEATURES OF COGNITIVE COMPLEXITY

3.1 Decision making

Let us return to the terms of the 'holy trinity'. Colander believes that the notion of 'purposeful behavior' will eventually substitute that of 'rationality'. The latter was originally criticized by Simon who rejected the notion of substantive rationality in favor of the notion of procedural rationality. The reason lies in the absence of the mental ability to take account of all the possible future moves. Under these circumstances a player is forced gradually to identify a strategy that she deems appropriate. In order to do so she will often resort to heuristics, i.e. problem solving techniques based on previous experience. Her rationality is procedural in that it relates to the process – as opposed to the product - of choice (Simon, 1978).

Let us consider a problem solving process in greater detail. It consists in identifying an algorithm that eventually provides a solution. For any single problem, however, a multiplicity of algorithms may exist (Egidi, 1992). In order to identify which one is more appropriate a second order algorithm would be required. The identification of the best second order algorithm would require a third order algorithm, and so on in an infinite regression. The decision when to stop the search process may be informed but it is ultimately based on an aspiration level.

The satisficing – thus discretionary - nature of the steps that the agent takes during her problem solving process emerges also in relation to information. She needs whatever information is relevant. Her problem, however, is that she does not know whether some information that she lacks is relevant to her decision. As long as she does not have it, she cannot say whether it is convenient to collect it or not. We are, therefore, back to the same type of infinite regress we referred to with regard to algorithms. Whatever decision the agent eventually takes is based on the information that she decided is adequate to decide.

The above issues characterize the problem solving activity, thus the decision making process, of private agents. Making sense of their behavior and of the outcomes that may ensue is what Colander is concerned with. The above issues, however, also relate to the problem solving activity and decision making of policy makers. They involve aspiration levels, i.e. discretionary decisions not only in choosing what policy is appropriate but also in choosing how far to go in the search for a solution. The distinction

between the neutral problem solving techniques of the economist and the value judgements of the policy maker is, therefore, inappropriate. Different views among economists are likely to persist in terms of how far they have to go in their search for an appropriate solution to their problems.

The key issue, here, is that there are two types of complexity. One relates to the economy, the other relates to the participant observer. They are what Delorme (1998) refers to as first order and second order complexity.

3.2 Decision contexts

The above discussion implicitly assumed that the goal underlying the problem solving process was fairly well identified. Consider chess: your problem – choosing the nth move - may be extremely difficult to solve but you know perfectly well why you have to solve it, i.e. what goal you are pursuing. Indeed, the goal is determined by the rules of the game. These, together with the information on the previous n-1 moves, determine your decision (or choice) context.

There are instances where your goal is fairly straightforward but the context is not as clear. Suppose you want to buy a pair of shoes. The environment you need to search in is not circumscribed as in the case of the chessboard. Similarly, the information you need is not completely available as in the game of chess and the rules of the game are not as binding. Consequently you may not be really sure about what pair of shoes you actually want to buy. Your goal and your decision context will become clear as you proceed in your search.

Similar considerations apply to business decisions. While profit is the typical goal for a firm, it remains to be seen whether it is long term or short term profit, whether it is associated to production of some sort (real profit) or it includes quasi-rents, speculative earnings and the like (money profit). When a decision relates to "making a profit", it involves the prior identification of the relevant context. The problem, here, is precisely to understand what "relevant" is in terms of time horizon, geographical space, political space, rules of conduct, etc..

The identification of a goal is more complicated when we consider individuals rather than business. The reason is that individuals are more likely to have multiple ranges of activity, each one with its goals and priorities. Thus, on strictly economic grounds an individual may want to maximize income, possibly because that allows her to buy as many goods as possible. The means to achieve this goal may include activities she deems immoral – e.g. corruption, blackmail, etc. – provided she knows she can get away with them. These actions, however, may clash with her (non-economic) values.

A conventional economist would view this as a typical trade off: the extra income is the price of honesty. According to this view, honesty and income are on the same ground and can be measured in terms of the same metric. An alternative view is that the two values (call them economic and moral) are on different grounds because they arise out of, and belong to, different contexts. The reason why there is no unique decision context, where everything can be sorted out, is that individuals are boundedly rational. They have to make sense of a range of different issues and situations but they cannot take everything into account when they do so. They have to simplify matters by drawing boundaries (Georgescu-Roegen, 1976), i.e. defining different contexts, each one with its own features, e.g. its rules.

As long as decision contexts are kept apart no problem arises. If they are mutually independent, no clash occurs: a strict "Give to Caesar what is Caesar's and to God what is God's" rule holds. Sometimes this is not the case, some overlap occurs and a clash ensues (Hirschman, 1984). The reason for this is that drawing boundaries is a typical problem solving activity, subject to the above outlined difficulties. Only substantively rational agents would be able to achieve that task in an optimal way. Independently of rationality, overlaps occur also because systems hardly are fully decomposable. Near decomposability (Simon, 1981) is a more plausible assumption, which suggests that the economy is best viewed as an open system (Kapp, 1976).

4. LIBERTARIAN PATERNALISM

These considerations are rather important if we consider the scope for paternalism. A paternalistic policy may be required, as we mentioned above, when bounded rationality precludes individuals from ascertaining what they actually want. It is under these circumstances that inertia often plays a dominant role. Since people are not aware, the libertarian policy maker[6] herself cannot easily understand what they really want: "What people choose often depends on the starting point, and hence the starting point cannot be selected by asking what people choose." (Thaler and Sunstein, 2003: 178).

Thaler and Sunstein consider the case of the director of a company cafeteria who realizes that people will choose what to eat according to how the courses are arranged. Whether the director likes it or not, her choice affects other people's choices. So the problem she faces is how to choose. The authors suggest three possible methods. The first one is to choose "what

[6] A libertarian policy maker is one who does not want to use coercion (Thaler and Sunstein, 2003).

the majority would choose if explicit choices were required and revealed" (*ibid.*: 178). Thus, one might identify an optimal equilibrium - or a desirable dynamic outcome - that agents would achieve, if they only knew what they want, and direct agents towards that equilibrium.

This method is relatively easy to follow when the decision context is so narrow that a "single-exit solution" occurs (Latsis, 1976). If actual economies are open systems, however, agents can choose within a wide range of possible decision contexts that may or may not include relative prices, bargaining power, ethical and/or religious values, etc.. It is not possible to identify the choice that agents would make if they were substantially rational precisely because they are not: they are procedurally rational. They define decision contexts in a different way. They would feel coerced if they were told that decisions are taken on their behalf according to *what they are supposed to choose* rather than according to what they would actually be willing to choose.

This leads us to the second method, which consists in choosing what "would force people to make their choices explicit" (Thaler and Sunstein, 2003: 178). A priori this is not impossible. Note, however, that the choice context need not be defined in advance. In the example of the cafeteria, a choice context might consist in (unhealthy) desserts placed before – or after - (healthy) fruit. Another choice context might provide a greater variety of fruit, in order to make fruit more appealing than dessert. Since this alternative might clash with budget constraints, the choice context might be extended to include the budget. The range of possible choice contexts is practically without limits. Under these circumstances, it is not clear who is supposed to choose the choice context. It could be people themselves but, provided they were capable and willing to do so, nothing would ensure that individual choices would be mutually consistent. Alternatively, it could be the policy maker but, then, people would be forced to choose in relation to a choice context that they might not accept.

The third method raises the same kind of problems. It consists in choosing the starting point so as to minimize the number of people who choose not to accept that starting point. The number of opt-outs, however, may turn out to be very low only because the choice context is not made explicit. Agents would not really be aware of what they are supposed to choose.

The implicit assumption in Thaler and Sunstein's (and Colander's) discussion of paternalism is that it is possible to envisage choice sets, i.e. given bundles of goods that agents, with given preferences, are supposed to choose from. What I argued above is that those bundles of goods are not given: agents need to choose which goods should be included and which goods should be kept out. They need to do so on the basis of what they

know, which may not be much. Furthermore, they must know what they want: they need to make out what their preferences are and they must deal with whatever inconsistencies arise between their preferences and their moral values. This is not a merely technical matter, however, as the example about income maximization through "immoral" expedients highlights. Finally, each good in a choice set is supposed to have its price tag, so that money can be a general metric[7]. But, owing to the boundary issue, it is far from clear that everything can be assessed in terms of a unique metric.

The broader notion of choice context was introduced precisely to take these issues into account. Its implications for social welfare – the underlying goal of all public policy – are worth emphasizing. According to the conventional approach to economics, social welfare can be measured in terms of money income (North, 1990), the key assumption being that only market transactions matter. The boundaries of social welfare may be extended in order to include externalities or some social costs but this only requires a reassessment of the money value of income (Coase, 1960). This implies that agents – whose preferences are assumed to be given – need choose only in relation to relative (market) prices. Truly, some information may be lacking when they choose, but the related transaction costs may nonetheless be assessed.

If social welfare generally cannot be restricted to income, however measured, and if income and non-income welfare are not mutually independent, an entirely different criterion is required. This leads Sen (1999) to suggest that social welfare be assessed in terms of capabilities, i.e. in terms of the command that individuals may have over their lives.

The upshot of the above discussion is that in some instances the choice context is commonly acknowledged and the goals are intuitive so that choices may be fairly easy to make and the policy maker can resort to libertarian paternalism. In general, however, decision contexts and goals are not straightforward. In the absence of perfect knowledge, a policy maker must formulate value judgements regarding the appropriate decision context. It is therefore inappropriate to believe that identifying the policy to be followed is a technical matter alone.

5. KNOWLEDGE CONTEXTS

The identification of a choice context or of a goal is a way to conceive of reality and act on it. Three situations may be pointed out, based on M. Polanyi (1962). In the first one – finding the fountain pen you lost – the

[7] Money is obviously assumed not to affect relative prices.

object of the search process is clearly known even though the search context – a special case of choice context - has to be defined as the process develops. Success occurs when the outcome of your search process matches your pre-defined goal, quite independently of whether you know how actually to achieve it.

In the second one – finding a word in a crossword puzzle – the object is unknown but the search context is strictly defined: the crossword puzzle is a closed system where only single-exit solutions apply. Success, here, depends on whether the outcome is achieved according to the rules defined within the search context.

In the third type – research in mathematics – both the object and the context are not known in advance. They are defined as the research – thus, the learning process - goes on. There are instances where the context is defined so that a theorem can be proved, much like in the second situation above. There are other instances where a goal suggests a redefinition of the search context.

The first and the second situations are fairly straightforward. Anyone can assess whether the search process has been successful or not. This is not the case with the third situation. Owing to its openness, it may involve the pursuit of a range of goals, each one requiring the identification of its corresponding search context. Since each goal bounds the search process, the objects of the search will depend on those bounds. Different search paths may therefore be followed. They depend on the idiosyncratic nature of the agent who is searching. They may change as the search proceeds. Under these circumstances success is difficult to assess or even define.

A further distinction among goals is appropriate. Consider the crossword puzzle. It is fairly obvious that we may refer to success when we find the right word. On the other hand, that success is usually only a proximate goal, which is functional to a more far-reaching goal, such as enjoying some relax. Success, therefore, turns out to be a subtle concept even in relation to the first two situations.

A range of intermediate goals is generally possible between the proximate and the most far-reaching one. Thus, finding a pen may be functional to carrying out one's task in a company, which in turn may be functional to the firm's profitability or to one's career, etc.. I suggest that the issues pertaining to ultimate goals – those that provide guidelines to one's life - are of the third type depicted above: they resemble much more research in mathematics than finding a pen or a word in a crossword puzzle. They relate to open-ended processes where contexts are likely to have broad boundaries and, owing to bounded rationality, loose internal connections as well as internal inconsistencies. This is particularly so if we consider that the people involved are not necessarily scholars, who are generally specialized

in rigorously pursuing connections, but normal people who try to make sense of their lives. I also suggest that problems that pertain to immediate goals or contexts are more similar to the first two types: searching for a pen or solving a crossword puzzle[8]. Either goals or contexts are easier to identify. They appear to be more "practical".

Owing to different views among (groups of) individuals, disagreement may relate to either proximate or far-reaching goals. In the first case the existence of a clearly identified goal or decision context makes it easy to distinguish technical issues from value judgements. Furthermore, a common ground is generally provided by shared far-reaching goals. This is where the most appropriate problem solving strategy is to rely on 'persuasion' (March and Simon, 1958).

In the second case, a common ground is more difficult to find. Owing to the extension of the decision context and the relative vagueness of the related goals, it is more difficult to distinguish technical issues from value judgements. When disagreement relates to the latter, there may be no higher tier goals that the parties share. This situation can be dealt with only through 'bargaining' and 'politics' (*ibid.*). Bargaining consists in reaching a compromise between alternative views. It may be achieved if a common solution to proximate problems may be found, independently of disagreement on the far reaching ones. Politics is required when disagreement involves all issues. It consists in creating a common ground for subsequent agreement by providing a shared goal or a shared decision context. Based on this common ground, a strategy of persuasion or bargaining can then be followed.

What all this leads to is that public policy can deal with major (i.e. political) problems only by fostering a shared view of what the relevant issues are. In turn, this involves providing an outlook of the relevant reality, i.e. tracing boundaries, that most people will be willing to accept. This entails a paternalism which is quite different from the one discussed above. The aim of the policy maker is not to comply with what people would choose but to provide the framework within which people learn and eventually choose, i.e. a decision context to choose how and what to learn. This framework I refer to as a "knowledge context". Two questions arise, here. How far is this paternalism possible? How far is it desirable? They are discussed in the following section.

[8] Note that the cases outlined are rather extreme. Thus, the crossword puzzle can be solved because it involves a single exit solution. Other cases may occur where a range of possible outcomes may result from the same set of rules.

6. INTEREST GROUPS

Up to this point my discussion referred to agents and policy makers. A more accurate look at the economy suggests that there are a range of distinct interest groups and stakeholders – above all employers, workers and consumers - who struggle to get a higher share of income relative to others. In some instances their mutual conflicts may relate to specific issues. In other instances, a far-reaching conflict may occur. In order to deal with these conflicts, these parties need to act in the same way outlined for policy makers: they must resort to persuasion, bargaining and politics. While persuasion and bargaining are fairly intuitive, politics requires some discussion.

Politics was defined above as creating the conditions for subsequent persuasion or bargaining. It consists in providing a shared knowledge context, i.e. a common ground where goals and search contexts may be defined. This is exactly what collective actors do. Firms advertise their products. In so doing they do try to persuade but they also provide a general view of what is supposed to improve the quality of life. Labor unions and consumer action groups act in much the same way. When unions defend worker rights or claim a wage hike, they are putting forth a view of social welfare – thus of the quality of life - which is not based on the level, or the rate of growth, of income but involves at the very least distribution. Similarly, when consumer groups argue that some product is too expensive or that it does not meet some requirement (safety, pollution, etc.), they are providing a view of social welfare which differs from money income or the amount of goods bought. The upshot is that the above interest groups pursue sectional interests but they also provide their view of the general interest.

These considerations allow us to reassess the features of public policy. For any given cultural heritage, what policy makers do, other collective actors do as well. Each one tries to direct learning processes in a way that fits their goals. The knowledge context of a community results from the joint action of a range of actors. From this perspective, 'public paternalism' interacts with a range of 'private paternalisms'. As a result, overlaps, as well as inconsistencies, affect the final outcome which in no way need be compact: despite her efforts, no single individual has a fully consistent worldview; despite commonalities, no two individuals share the same knowledge and the same worldview.

The question, now, is whether public paternalism is desirable. It is reasonable to believe that the sectional and general interest views provided by each group are internally consistent but may be incompatible with the views provided by other groups. This depends not only on bounded rationality but on the existence of inconsistencies and conflicting interests

within the economy. It is also reasonable to believe that the general interest each group upholds is conceived so as to be consistent with its own sectional interests. Should an inconsistency arise, the general interest view would be the one to be reassessed, not the sectional interests. Under these circumstances the scope for public policy should be to reassert a general interest view. This would involve creating a shared view of the common good and identifying a common ground to make sectional interests mutually consistent or, at least, compatible.

This scope for public policy is subject to two qualifications. First, a shared view of the common good need not imply that the latter actually exists. As should be clear by now, there is no unique way to look at the world we live in: whatever view policy makers were to suggest would be as discretionary as any other. Its success would depend not on the goodness of the society it upholds but on its acceptance by social actors. Basically, this is what underlies the term "social cohesion".

Second, precisely because no unique view exists, one has to be chosen. Democracy is supposed to be the means to choose among these views. Note, however, that which view is preferred depends on the existing knowledge context. Actors who pursue sectional interests, however, (purposefully) act so as to affect that knowledge context. Public policy cannot diregard this issue. It must pursue a knowledge context that is consistent with the view of society that it wishes to enact. How it can do this is the subject of the section that follows.

7. DEMOCRATIC PATERNALISM

One of the key tenets of libertarian paternalism is that individuals should choose. If they do not, whatever decision is taken on their behalf should respect what they (would) want. This is a reasonable claim as long as we assume that individuals know what they want or that they could figure it out if only they thought about it. What individuals know, however, may lie beyond their control: their decision context may be restricted by the purposive action of other actors who want to direct their choices. These knowledge asymmetries feed back on themselves. The less individuals know, the more they are forced to rely on external information and knowledge. Thus, although learning is always a social process, it may nonetheless be either self- or hetero-directed (Ramazzotti, forthcoming): individuals may choose what and why they are learning or they may lose control of what they learn because someone else is choosing for them.

Policy may attempt to correct those asymmetries by allowing individuals to extend their knowledge. It can do so by enhancing the circulation of

information as well as the variety of its sources. It can also allow that information to be appreciated by fostering discussion and confrontation of ideas. In most instances, however, individuals pursue knowledge only in so far as it may be put to use: what point is there in discussing about change or trying to find out how best to change things if change is deemed impossible[9]? What this leads to is that policy should affect knowledge by enhancing conditions that require knowledge and learning, i.e. conditions that enable people to choose how they wish to conduct their lives. In the absence of such public action, individuals would have to rely on knowledge contexts that are created by agents who protect their vested interests.

The policy I refer to may be termed democratic paternalism. It consists in enhancing the insurgence of knowledge contexts that allow people to actually control what policy makers do, including how they enhance knowledge contexts. A key question, here, is whether such a policy is possible. Much like the individual I discussed above, who must choose whether to give priority to her economic or to her moral values, a policy maker may have to choose between her personal goals and the social values she fosters: corruption, just like honesty, is always possible. Furthermore, although she is likely to be more aware than common citizens of the vested interests that are at stake, her choices may nonetheless be influenced by the knowledge contexts that are associated to those interests: aside from immoral behavior, a biased outlook is also possible. Thus democratic paternalism may be hindered or even precluded by the establishment of a paternalistic democracy, one where policy makers are just another manifestation of vested interests and knowledge contexts reflect these interests to the point that people cannot identify alternative viewpoints.

Further inquiry will be necessary to identify the circumstances that may favor the former, rather than the latter, outcome. What is sure is that the libertarian view is too simple to be of any help in terms of public policy.

8. CONCLUDING REMARKS

The general discussion stressed the discretionary nature of choices concerning what and how much information is required to make a decision, what algorithm is appropriate, what boundaries should be traced - thus also what decision contexts should be chosen – and, finally, what goals should be pursued. It emphasized that the economy is an open system and that this

[9] Sen (1999) points out that someone could be happy despite her dismal living conditions simply because she cannot envisage any other way to conduct her life. I am suggesting that this holds for learning too.

must be the point of departure for any policy analysis. It stressed that discretion applies to individuals but it applies also to the policy maker: complexity has to do not only with how agents and the economy behave but also with how agents, including the policy maker, observe them.

From this perspective, even though mathematical and computational progress may help to model specific situations, it cannot substitute the strictly qualitative features of explanation and value judgement in economic theory.

ACKNOWLEDGEMENTS

I wish to thank Stefano Solari for his comments on a previous version of this paper.

REFERENCES

Brock, W. A., Colander, D., 2000, *Complexity and Policy, The Complexity Vision and the Teaching of Economics*, Elgar, Cheltenham.

Coase, R. H., 1988, The problem of social cost, in: *The Firm, the Market, and the Law*, Chicago University Press, Chicago.

Colander, D., 2003a, *Muddling Through and Policy Analysis*, Middlebury College Economics Discussion Paper, No. 03-17.

Colander, D., 2003b, *The Complexity Revolution and the Future of Economics*, Middlebury College Economics Discussion Paper, No. 03-19.

Colander, D., Holt, R., and Rosser, B., 2003, *The Changing Face of Mainstream Economics*, Middlebury College Economics Discussion Paper, No. 03-27.

Delorme, R., 1998, *From First Order to Second Order Complexity in Economic Theorising*, Mimeo.

Egidi, M., 1992, Organizational learning, problem solving and the division of labour, in: *Economics, Bounded Rationality and the Cognitive Revolution*, M. Egidi and R. Marris, eds., Elgar, Aldershot.

Georgescu-Roegen, N., 1976, Process in farming versus process in manufacturing: a problem of balanced development, in: *Energy and Economic Myths. Institutional and Analytical Economic Essays*, Pergamon Press, New York.

Hirschman, A. O., 1984, Against parsimony: three easy ways of complicating some categories of economic discourse, *The American Economic Review* **74**:89-97.

Kapp, K. W., 1976, The open-system character of the economy and its implications, in: *Economics in the Future: Towards a New Paradigm*, K. Dopfer, ed.,Macmillan, London.

Latsis, S. J., 1976, A research program in economics, in: *Method and Appraisal in Economics*, S. J. Latsis, ed., Cambridge University Press, Cambridge.

March, J. G., and Simon, H. A., 1958, *Organizations*, John Wiley and Sons, New York.

North, D. C., 1990, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge.

Polanyi, M., 1962, *Personal Knowledge. Towards a Post-Critical Philosophy*, Routledge, London.

Ramazzotti, P., (forthcoming), Constitutive rules and strategic behavior in: *Institutions in Economics and Sociology: Variety, Dialogue and Future Challenges*, K. Nielsen and C.A. Koch, eds., Elgar, Cheltenham.

Rosser, J. B, Jr., 1999, On the complexities of complex economic dynamics, *Journal of Economic Perspectives* **13**:169-192.

Sen, A., 1999, *Development as Freedom*, Alfred Knopf, New York.

Simon, H. A., 1978, Rationality as process and as product of thought, in: *Decision Making. Descriptive, Normative and Prescriptive Interactions*, D. E. Bell, H. Raiffa, and A. Tversky, eds., Cambridge University Press, Cambridge.

Simon, H. A., 1981, The architecture of complexity, in: The *Sciences of the Artificial*, MIT Press, Cambridge, MA.

Thaler, R. H., and Sunstein, C. R., 2003, Behavioral economics, public policy, and paternalism, *The American Economic Review* (May):175-179.

A COMPUTATIONAL MODEL OF FACE PERCEPTION

Maria Pietronilla Penna[1], Vera Stara[2], Marco Boi[1] and Paolo Puliti[2]
[1]*Università degli Studi di Cagliari, Facoltà di Scienze della Formazione, Dipartimento di Psicologia - Email: maria.pietronilla@unica.it, marcoboi@tiscali.it*
[2]*Università Politecnica delle Marche, Facoltà di Ingegneria, DEIT Email: v.stara@univpm.it, p.puliti@univpm.it*

Abstract: In this paper we are interested in the perceptual aspect of face recognition that is the process that categorizes the visually perceived face into a perceptive space made by as many categories as the possible discriminations are. The question we want to answer to is if it is possible to model some aspect of face perception using a neural network architecture and if this model is able to provide any useful information about that conditions such as apperceptive prosopagnosia in which face perception appears to be impaired. We will propose an answer to these question using a computational model. The research was dived into two experiments: the first one aimed to test the ability of the network to discriminate between different faces and to generalize between similar faces and the other one aimed to investigate the behaviour of the system when noise is added to the normal operation of the network.

Key words: face perception; neural network; prosopagnosia.

1. INTRODUCTION

Interest in visual face recognition has been increasingly growing in recent times due to the need for reliable automatic face recognition systems and to recent evidences that show that face recognition is somehow separable from common object recognition in terms of cognitive processes and neural correlates.

As far as concern the cognitive aspect, face recognition is usually claimed to imply a holistic processing of image opposed to simple part decomposition processing used in object recognition: visual face

representation is thus probably processed and stored as a whole rather than a sum of parts (Young, Hallaway and Hay 1987; Tanaka and Farah 1993). Moreover this face specific processing seems to be elicited as a response to the presence of a face gestalt in the visual field as suggests the "face inversion effect" that consists in a worst performance in a face recognition task for faces presented upside-down than for face presented right side up.

Another difference between object and face recognition lies on the visual information these process seems to rely on. Whereas for the former appear to be prominent edge based information, the latter seems to rely more on shadow and shading information. Biederman and Kalocsai (1997) believe that face recognition utilizes a visual information mapped directly from the early visual processing stages.

The different effects of negation and changing in lighting direction on objects and faces probably depend on such early visual information processing difference as this transformation affects face recognition much more than object recognition (Bruce and Langton, 1994; Subramaniam and Biederman, 1997).

Following Biederman and Kalocsai we suppose that face representation implied in face recognition could imply a more direct mapping of early visual processing areas (V1, V2, V4) neurons activations. This could account for the configurational representation implied in face inversion and for the sensitivity to lighting and shadowing change.

Probably this process has its neural substrate in the brain areas recruited during face recognition task as shown by fMRI studies (Kanwisher et al., 1997; Hasson et al., 2001). Face selective regions have been found also in macacus (Perret et al., 1982; Young and Yamane, 1992) by single unit recording but probably the most impressive evidence of the existence of an area specialized in face perception comes from prosopagnosia. This is a clinical condition characterized by the an impaired ability to recognize faces in presence of a relatively preserved capacity to recognize objects, that generally occurs after a right occipital-temporal lesion. Prosopagnosia has often been considered as a consequence of the malfunctioning of an area dedicated to face recognition (Farah et al., 1995) even though there are some alternative explanations (Tarr and Cheng, 2003).

Furthermore some authors (De Renzi et al., 1991) operate a distinction between an apperceptive and an associative form of prosopagnosia: the former is characterized by a perceptual deficit (individuals cannot distinguish between two simultaneously presented faces) whereas the latter seems more related to a mnemonic deficit (individuals are not able to retrieve the semantic information associated to the perceived face). Michelon and Biederman (2003) showed also a preserved imagery for faces in presence of an apperceptive prosopagnosia.

Similarly we believe that face perception and recognition should be considered as two different stages of the process that links a visual stimulus to the semantic information associated, by first producing a categorization of a face visual stimulus, processed by early visual, areas and then using this percept to retrieve the information associated to the individual the face belongs to.

2. THE MODEL

We used a neural network architecture in order to account for some features of visual information implied in human face processing. It is remarkable that our model doesn't reproduce neither the processes underlying the individuation and the extraction of face representation from the visual field nor the processes leading from face representation to the semantic information relative to individuals which the face belongs to. It just focuses on the perceptive aspect of face recognition that is the mechanism that allows the discrimination of faces according to visual information they convey and the experience acquired during face perception.

Like in the human brain face area, where a visual representation coming from early visual processing areas is processed and categorized according to a given perceptive space, our model system receives a visual representation of a face and returns its categorization in a perceptive space consisting of a given number of categories. A node of the output layer fires every time a face with suitable features is presented to the input layer. The characteristics every node is most likely to respond are determined by network self-organization, occurring during a period named "learning phase" in which the network output layer nodes compete for activation, tuning their response with the occurrence of suitable stimulus features.

We used a Kohonen self-organizing neural network, including two-dimensional input and output layers and receiving as input an image coded as a grey level matrix. The input layer is composed of a number of elements (nodes) equal to the number of matrix pixels. Every node in the input layer translates pixel luminance value into an activation value.

The network operation can be subdivided into two different phases, the learning and the test one. During the former the network self-organizes in order to discriminate between the input stimuli, while in the latter it extends the categorization ability acquired during the previous phase to new stimuli not presented before.

The number of the output layer nodes (16 in our simulations) is specified according to the number of expected categories. The output layer nodes are connected each other with fixed weights that give rise to a competitive

dynamics allowing just one output node to fire for every presentation of input pattern. The output layer is connected to the input one by a set of weighted connections that change during the learning phase according to the learning rule.

$$w_{ij}(t+1) = \begin{cases} w_{ij}(t+1)\alpha(t)[S_j^{(k)} - w_{ij}(t)] & \text{if } i \in B \\ w_{ij} & \text{if } i \notin B \end{cases}$$

Here B is the set of nodes that are involved in weight change, w_{ij} is the weight of the connection between the input and the output node, $\alpha(t)$ is a learning parameter that changes with time, $S_j^{(k)}$ is the incoming pattern.

The weight changes affects all output layer nodes that are within the radius of a circle centered on the winner node whose value is specified at the start of the learning phase and decreases during it.

The activation of every output node is defined by Kohonen shortcut algorithm:

$$y_i = \sum_{j=1}^{M} (x_j \cdot w_{ij})$$

Here y_i is the activation of output layer node, x_i is the activation of input node, w_{ij} is the weight of the connection between input and output node and M the total number of output nodes.

We trained the network with a set of different images displaying different faces with different lighting conditions and after we tested the network presenting images of the same faces taken in different pose and lighting conditions. We checked whether network was able to cluster the images according to the face they represent, that is to respond with the activation of a specific node to the presentation of a specific face.

As the network gave rise to categories based on statistical features of incoming stimuli, we expected the network to cluster together the images belonging to the same face, even with different pose and lighting conditions, by firing with the same node and to discriminate between images of different faces by activating different nodes.

3. EXPERIMENT 1: CATEGORIZATION

We trained the network with a set of 45 different images representing 15 faces in 3 different poses and we expected it to be able not only to

discriminate between different faces and generalize between different images of the same face but also to extend this classification to new images displaying the same faces in different pose and lighting conditions. The images were taken from the Yale database (Yale). In particular we expected the network to fire with a different output layer neuron for images of different faces but not for different images of the same face.

The training session lasted for 52500 epochs with the following parameters and time variation rules:

$$\eta_0 = 0.10; \quad \beta = 0.0001; \quad R_0 = 3.9; \quad b_0 = 0.0001$$

$$R(t) = R_0 \exp(-b_0 t) \qquad \text{radius of activity bubble}$$

$$\alpha(t) = \eta_0 \exp(-\beta t) \qquad \text{learning parameter.}$$

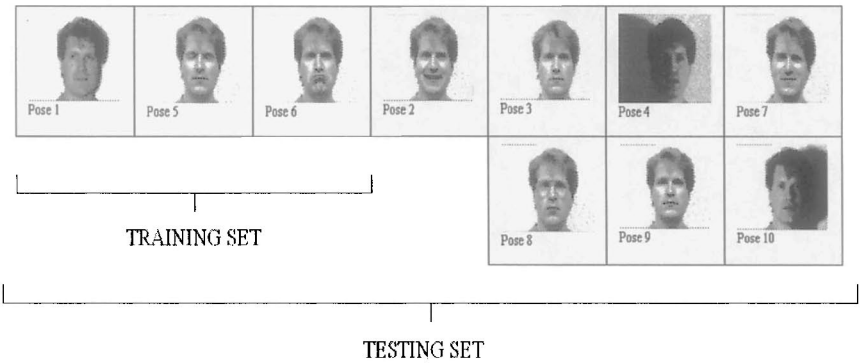


Figure 1. The poses used in the training and testing set. Every face used in the experiment is presented in the displayed pose.

We computed the standardized residuals for the frequency tables displaying the activated node and the displayed face for every image. The results are presented in table 1. Every cell reports the shift of the observed score from the expected score (computed assuming the independence between node and face) expressed in terms of standard deviations. The result suggests an association between node and face although some nodes seems to be associated with two or even three faces.

Table 1. Standardized residuals table for the face-node table. The values in the cells represent the measure of the difference between observed and expected frequencies in terms of standard deviations. A high positive value represents a high relative frequency while a high negative value represents a low relative frequency for that cell. For every given face we highlighted the

cell corresponding to the node that has the highest activation frequency related to the presentation of an image displaying that specific face. The highest frequencies cells are displaced on 11 nodes.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
0,0	-1	1,3	-1	5,3	-1	-1	-1	0,3	-1	-1	-1	1,3	-1	-1	-1
0,1	-0	-0	-0	2,5	-0	-0	-0	-0	-0	3,5	-0	-0	-0	-0	-0
0,2	-1	-1	0,1	0,1	-1	-1	1,1	0,1	-1	5,1	-1	-1	-1	-1	2,1
0,3	-1	-1	-1	-1	-1	-1	-1	5,3	-1	-1	-1	-1	-1	-1	4,3
1,0	8,3	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0,3	-1	-1	-1
1,1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	9,3	-1	-1	-1	0,3
1,2	-0	-0	-0	-0	-0	-0	-0	0,8	0,8	-0	-0	-0	-0	-0	-0
1,3	-2	-2	4,7	-2	-2	-2	5,8	-2	6,8	-2	-2	-2	-0	-2	-2
2,0	-0	-0	-0	-0	-0	6,5	-0	-0	-0	-0	-0	-0	-0	-0	-0
2,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2,2	-1	-1	0,8	-1	7	-1	-0	-0	-0	-1	-1	2,9	-0	-1	-1
2,3	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	7,4	-1	-1
3,0	0,7	-0	0,7	-0	0,7	-0	-0	-0	-0	-0	-0	-0	-0	0,7	0,7
3,1	-1	-1	-1	-1	-1	1,3	-1	-1	-1	-1	-1	-1	-1	7,3	-1
3,2	-0	-0	-0	-0	-0	0,8	-0	-0	-0	-0	-0	-0	-0	0,8	-0
3,3	-1	7,3	-1	-1	0,2	-1	-1	-1	-1	-1	-1	2,2	-1	-1	-1

The table shows the specific nodes that every face appears to be preferentially associated to. In fact if we consider for every face the node that is mainly associated with the presentation of a given face, every face has a node that is more likely to fire as a response to the presentation of an image of that face. That means that for every face there is a class the image of a given face is most likely to be classified in. There are 11 such different classes because the preferred node is the same for some faces (e.g. faces 3, 7 and 9). So a set of 15 faces is categorized preferentially in 11 classes.

The preferred activation node is different for 8 of 15 faces (faces 1, 2, 4, 6, 10, 11, 13, 14 in nodes 1;0 , 3;3, 0;0, 2;0 , 0;2 , 1;1 , 2;3 , 3;1) while 7 faces share some preferential nodes (faces 3, 7, 9 in node 1;3, faces 8 and 15 in node 0;3, face 5 and 12 in node 2;2). Thus it seems the network tends to distinguish 15 faces in 11 classes with consequent partial overlapping of some faces on the same node. Unfortunately the sample size is not large enough to allow a probability test on the standardized residuals but a descriptive analysis of the response of the network to the presentation of faces suggest that it is able to selectively activate a node as a response to the face displayed in the presented image even if there are some faces that are represented in the same node.

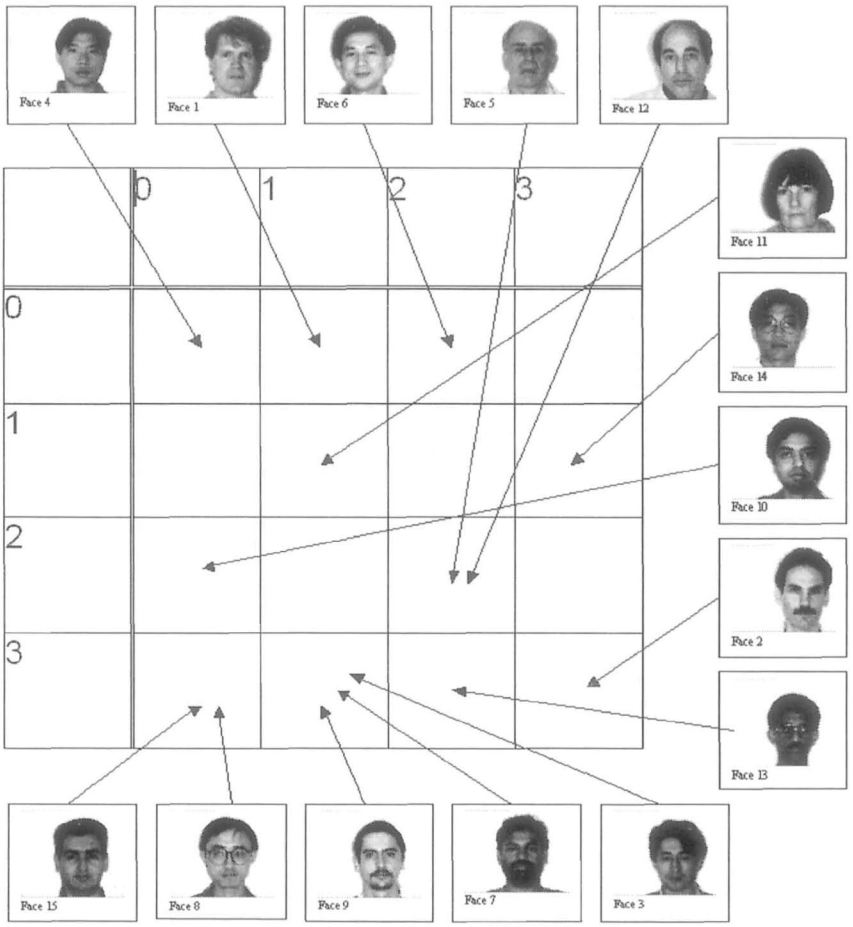


Figure 2. The set of faces used in the experiment. The arrow links the face with the node that show the highest standardized residual for its presentation.

4. EXPERIMENT 2: APPERCEPTIVE PROSOPAGNOSIA

In this experiment we tested the effect of a noise component on the discrimination of different faces and investigated how this alteration influences the network operation and thus the categorization of faces.

The noise consists in the substitution of the weight value between a given input and output node, with a new weight given by the weighted mean

between the old weight and a random value sorted between the minimum and maximum value a weight can assume for that specific output node.

$$y_i = \sum_{j=1}^{M} x_j [w_{ij}(1-p) + m_i r_{ij} p]$$

Here y_i is the activation of the output layer node, x_i is the activation of the input node, w_{ij} is the weight of the connection between the input and the output node, p is a value set by the experimenter, m_i is the highest absolute value of a weight w_{ij} and r_{ij} is a matrix of random values chosen within the interval from -1 to $+1$.

We tested the network for different values of p varying from 0 to 1 and we recorded the categorization produced. We wonder what could be the effect of the introduction of different amounts of noise on network operation.

The probability associated to the chi-square test of independence is 2.17 E-84 when p is 0 and gradually decreases as p tends to 1 and became 0.999 when p is 1. So the first result is that the association between node and face tends to decrease with growing noise.

But it is interesting to note that the change that occurs in network operation cannot be limited to this phenomenon. It is worth of note that for some faces the most responsive node change as the noise value grows. For example face 2 preferential node changes from 3;3 to 0;0 when noise reaches 60% and face 4 changes from 0;0 to 0;1 at 70%.

Table 2. Standardized residuals for face 1 for variable proportion of noise.

Face1	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
0;0	-0,86	0,144	0,144	0,144	-0,86	-0,9	-0,1	1,018	0,268	1,049	0
0;1	-0,68	-0,73	-0,73	-0,77	-0,68	-0,68	0,415	-0,9	-0,62	-0,68	-0,57
0;2	-0,98	-0,98	-0,98	-0,94	-0,94	-0,98	-0,9	-0,9	-0,9	-0,73	-0,68
0;3	-0,86	-0,86	-0,86	-0,9	-0,82	-0,77	-1,02	-1,13	0,268	0,184	-0,73
1;0	8,184	7,184	7,184	6,203	8,164	7,169	6,11	5,092	0,034	-0,86	-0,68
1;1	-0,86	-0,86	-0,86	-0,86	-0,82	-0,82	-0,86	-0,77	-0,62	-0,62	0,07
1;2	-0,36	-0,36	-0,36	-0,25	-0,25	-0,44	-0,44	-0,36	-0,73	0,034	0,225
1;3	-1,27	-1,27	-1,27	-1,3	-1,37	-1,23	-0,98	-0,77	-0,98	0,106	-0,77
2;0	-0,68	-0,68	-0,68	-0,68	-0,68	-0,73	-0,73	-0,68	-0,51	-0,68	2,151
2;1	0	0	0	0	0	-0,25	-0,57	0,415	0,474	-1,02	-0,94
2;2	-1,13	-1,09	-1,09	-1,09	-1,13	-1,13	-0,94	-0,9	1,938	0,225	-0,57
2;3	-0,73	-0,73	-0,77	-0,73	-0,73	-0,73	-0,62	-0,57	-0,68	0,415	1,221
3;0	0,415	0,415	0,415	0,474	0,415	0,415	-0,68	-0,36	-0,77	-0,57	0,07
3;1	-0,82	-0,82	-0,82	-0,82	-0,82	-0,77	0,268	0,034	0,07	-0,44	-0,68
3;2	-0,36	-0,36	-0,25	-0,36	-0,44	-0,36	-0,57	-0,77	0,314	0,144	-0,68
3;3	-0,9	-0,9	-0,9	0,07	-0,82	0,106	-0,57	-0,51	0,144	1,081	0,106

Table 3. Standardized residuals for face 2 for variable proportion of noise.

Face2	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
0;0	1,148	1,148	1,148	2,151	0,144	2,121	5,051	5,092	0,268	2,064	2,037
0;1	-0,68	-0,73	-0,73	-0,77	-0,68	-0,68	-0,57	0,106	0,362	-0,68	-0,57
0;2	-0,98	-0,98	-0,98	-0,94	-0,94	-0,98	-0,9	-0,9	-0,9	-0,73	-0,68
0;3	-0,86	-0,86	-0,86	-0,9	-0,82	-0,77	-1,02	-1,13	-0,73	-0,82	-0,73
1;0	-0,82	-0,82	-0,82	-0,77	-0,9	0,144	-1,02	0	1,049	1,148	-0,68
1;1	-0,86	-0,86	-0,86	-0,86	-0,82	-0,82	-0,86	-0,77	-0,62	-0,62	0,07
1;2	-0,36	-0,36	-0,36	-0,25	-0,25	-0,44	-0,44	-0,36	-0,73	-0,98	0,225
1;3	-1,27	-1,27	-1,27	-1,3	-1,37	-1,23	-0,98	-0,77	-0,98	-0,9	-0,77
2;0	-0,68	-0,68	-0,68	-0,68	-0,68	-0,73	-0,73	-0,68	-0,51	-0,68	0,144
2;1	0	0	0	0	0	-0,25	-0,57	-0,57	-0,51	0	-0,94
2;2	-1,13	-1,09	-1,09	-1,09	-0,1	-1,13	-0,94	-0,9	-0,13	0,225	-0,57
2;3	-0,73	-0,73	-0,77	-0,73	-0,73	-0,73	-0,62	-0,57	-0,68	-0,57	-0,77
3;0	-0,57	-0,57	-0,57	-0,51	-0,57	-0,57	-0,68	-0,36	-0,77	-0,57	1,081
3;1	-0,82	-0,82	-0,82	-0,82	-0,82	-0,77	-0,73	-0,98	-0,94	-0,44	0,314
3;2	-0,36	-0,36	-0,25	-0,36	-0,44	-0,36	-0,57	-0,77	-0,68	-0,86	-0,68
3;3	7,157	7,157	7,157	6,135	7,184	5,142	3,363	1,453	4,159	2,092	0,106

Table 4. Standardized residuals for face 3 for variable proportion of noise.

Face3	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
0;0	5,162	5,162	5,162	4,159	5,162	4,135	4,021	0	1,261	2,064	1,018
0;1	2,292	2,254	2,254	3,214	2,292	1,303	0,415	4,135	-0,62	0,314	-0,57
0;2	0,034	0,034	0,034	0,07	-0,94	-0,98	0,106	1,114	-0,9	0,268	-0,68
0;3	-0,86	-0,86	-0,86	-0,9	0,184	-0,77	-1,02	-1,13	-0,73	-0,82	-0,73
1;0	-0,82	-0,82	-0,82	-0,77	-0,9	0,144	0	0	3,078	-0,86	0,314
1;1	-0,86	-0,86	-0,86	-0,86	-0,82	-0,82	-0,86	-0,77	-0,62	0,362	-0,94
1;2	-0,36	-0,36	-0,36	-0,25	-0,25	-0,44	-0,44	-0,36	-0,73	-0,98	-0,77
1;3	-1,27	-1,27	-1,27	-1,3	-1,37	-1,23	-0,98	-0,77	-0,98	-0,9	0,225
2;0	-0,68	-0,68	-0,68	-0,68	-0,68	-0,73	-0,73	-0,68	-0,51	-0,68	1,148
2;1	0	0	0	0	0	0,719	-0,57	-0,57	-0,51	0	-0,94
2;2	-1,13	-1,09	-1,09	-1,09	-1,13	-1,13	1,081	-0,9	-0,13	1,221	-0,57
2;3	-0,73	-0,73	-0,77	-0,73	-0,73	-0,73	-0,62	-0,57	0,314	-0,57	0,225
3;0	-0,57	-0,57	-0,57	-0,51	-0,57	-0,57	-0,68	-0,36	-0,77	-0,57	-0,94
3;1	-0,82	-0,82	-0,82	-0,82	-0,82	-0,77	-0,73	-0,98	-0,94	-0,44	0,314
3;2	-0,36	-0,36	-0,25	-0,36	-0,44	-0,36	-0,57	-0,77	-0,68	-0,86	0,314
3;3	-0,9	-0,9	-0,9	-0,94	-0,82	0,106	-0,57	0,474	1,148	0,07	0,106

We interpret this result as a form of internal reorganization of the network in order to cope with the noise introduced. As a consequence of this process the network changes some of its coding properties and produces a different categorization of the stimuli.

We wonder if this phenomenon have some correspondence with what occurs in the prosopagnosic brain, that is if the damage that occurs in prosopagnosics produce a simple exclusion of the face specific device or

cause it to function in a different fashion producing a categorization in which certain characteristics of the perceptual space are altered and the individual is no more able to extract from a given percept the semantic information associated not because the percept is no more available but because that percept is no more the same and produces a wrong recognition.

Unfortunately it seems that studies on prosopagnosia have not yet investigated the possibility of a different pattern of categorization but simply have focused on the accuracy of the discrimination showing that prosopagnosics have more difficulty in recognizing face that is in associating a face to its identity.

Interestingly some studies show effects that seems to imply the role of a malfunctioning face specific device. For example Farah (1995) has shown that while normal subjects have a worse performance with inverted than with upright faces, some prosopagnosics show an opposite pattern with a better performance when face is presented inverted than when it is presented upright. Boutsen and Humpreys (2002) showed that a prosopagnosics, when asked to say whether two simultaneously presented faces are the same or different, tends to answer "same" more than different.

We believe that, like in our artificial system the malfunctioning was simulated by a random noise in the value of the weight, in the human brain area dedicated to face perception a lesion could result in a restructuring of the perceptive space.

5. CONCLUSION

In this paper we have considered different aspects that distinguish face from object recognition (Young, Hallaway and Hay, 1987; Tanaka and Farah, 1993). On the basis of some nenuropsychological data (De Renzi et al., 1991; Michelon et al., 2003) we distinguished in this process a perceptive component from a mnemonic one and tried to model some aspects of the face perception system using a neural network architecture.

The model was able to respond with the activation of a specific node to the presentation of a given face although this association was not straightforward for every node in the network.

After that we wondered if a malfunctioning of the network could model some aspects of apperceptive prosopagnosia.

We simulated a lesion in the network by introducing a noise component in network operation and we observed a gradual decrease of the association between node and face as the noise grew.

Moreover in correspondence to suitable noise values we observed a change of the preferentially activated node. In our opinion this effect may

arise from a change in perceptive space occurred as a consequence of a restructuring of the organization of the network. We believe that as the noise get higher, the network change its organization, but this new organization results in a different operation of the network that is in a different categorization.

We have shown that in an artificial neural network able to discrimanate between different faces, suitable alterations of its normal operation could result not only in a simple decrease in the efficiency of discrimination but also in a change of the classification produced by the network. We wonder if apperceptive prosopagnosia, that is thought to be a consequence of a malfunctioning of the face perception area, would consist in the absence of face specific processing or in the change of properties of perceptive space and consequently of the classification operated by the face perception system. For example it could be interesting to investigate whether even in those individuals with impaired face perception is still preserved some kind of discrimination, also if not fine enough to discriminate between different individuals.

REFERENCES

Biederman, I., and Kalocsai, P., 1997, Neurocomputational bases of object and face recognition, *Philosophical Transactions of the Royal Society B* **352**:1203–1219.

Boutse, L., and Humpreys, G.W., 2002, Face context interferes with local part processing in prosopagnosic patients, *Neuropsychologia* **40**:2305-2313.

Bruce, V., and Langton, S., 1994, The use of pigmentation and shading information in recognizing the sex and identities of faces, *Perception* **23**:803-822.

De Renzi, E., Faglioni, P., Grossi, D., and Nichelli, P., 1991, Apperceptive and associative forms of prosopagnosia, *Cortex* **27**, 212-221.

Farah, M. J., Wilson, K. D., Drain, H. M., and Tanaka, J. R., 1995, The inverted face inversion effect in prosopagnosia: evidence for mandatory, face specific perceptual mechanism, *Vision Research* **14**:2089-2093.

Hasson, U., Hendler, T., Ben-Bashat, D., and Malach, R., 2001, Face or vase? A neural correlate of shape selective groping process in human brain, *Journal of Cognitive Neuroscience* **13**:744-753.

Kanwisher, N., Mc Dermat, J., and Chun, M. M., 1997, The fusiform face area: a module in extrastriate cortex speciaized in face perception, *Journal of Neuroscience* **17**:4302-4311.

Michelon, P., and Biederman, I., 2003, Less impairment in face imagery than face perception in early prosoppgnosia, *Neuropsychologia* **41**:421–441.

Perret, D. I., Rolls, E. T., and Caan, W., 1982, Visual neurones responsive to faces in the monkey temporal cortex, *Experiemtal Brain Researh* **47**:329-342.

Subramaniam, S., and Biederman, I., 1997, Does contrast reversal affect object identification?, *Investigative Opthalmology and Visual Science* **38**(998).

Tanaka, J. W., and Farah, M. J., 1993, Parts and wholes in face recognition, *Quarteerly Journal of Experimental Psychology* **46A**:225-245.

Tarr, M. J., and Cheng, Y. D., 2003, Learning to see faces and objects, *Trends in Cognitive Sciences* **1**:23-30.

Yale database; http://cvc.yale.edu/projects/yalefaces/yalefaces.html .

Young, A. W., Hallaway, D., and Hay, D. C., 1987, Configurational information in face perception, *Perception* **16**:747-759.

Young, M. P., and Yamane, S., 1992, Sparse population coding of face in the inferotemporal cortex, *Science* **256**:1327-1331.

THE NEON COLOR SPREADING AND THE WATERCOLOR ILLUSION: PHENOMENAL LINKS AND NEURAL MECHANISMS

Baingio Pinna
Facoltà di Lingue e Letterature Straniere, University of Sassari, Via Roma 151, I-07100 Sassari, Italy, e-mail: baingio@uniss.it

Abstract: This work explores the interactions between the cortical boundary and coloration and figural properties of two illusions: the neon color spreading and the watercolor effect. Through psychophysical and phenomenal observations the neon color spreading has been compared with the watercolor illusion. The results showed that the phenomenal qualities of both effects can be reduced to a basic common limiting case that can explain the perceptual differences between the two illusions. Finally, the article proposes a unified explanation of the properties of the two illusions in terms of the FACADE neural model of biological vision (Grossberg, 1994). The model clarifies how local properties, such as spatial competition, can control some properties of both illusions, and how more global figural properties, determining the shape and strength of contours, can explain differences between the two illusions.

Key words: Neon color spreading; watercolor illusion; Gestalt principle of grouping; border ownership; figure-ground segregation; FACADE model.

1. NEON COLOR SPREADING

In 1971 Varin reported a "chromatic diffusion" effect obtained by using four sets of concentric black circumferences arranged in a virtual cross and partially composed of blue arcs creating a virtual large central blue circle (see Figure 1). Under these conditions the central virtual circle appears as a ghostly transparent veil of bluish tint extending among the boundaries of the blue arcs. The chromatic translucent diffusion fills the entire illusory circle induced by the terminations of the black arcs (see Bressan et al., 1997, for a review).

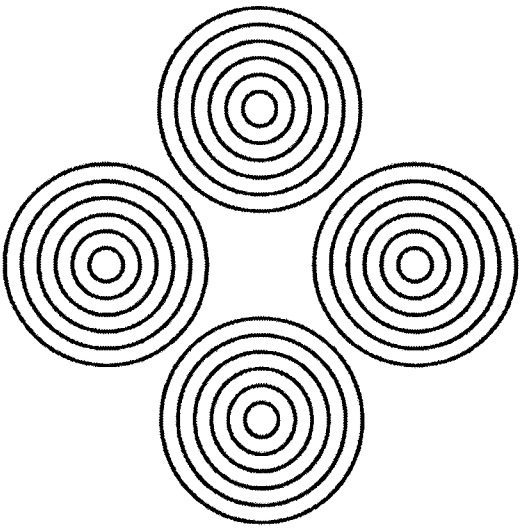


Figure 1. The neon color spreading.

The "chromatic diffusion" effect was independently rediscovered in 1975 by van Tuijl (see also van Tuijl and de Weert, 1979), who named it "neon-like color spreading". Van Tuijl used a lattice of horizontal and vertical black lines, where segments creating an inset virtual diamond shape had a different color (i.e. blue). The perceptual result is a delicately tinted transparent diamond-like veil above the lattice.

A common geometrical property of all the known cases of the neon color spreading concerns the continuation of one line in a second line differently colored or, in other words, a single continuous line varying at a certain point from one color to another.

The neon color spreading manifests two basic phenomenal properties: coloration and figural effects.

1.1 Coloration effect in the neon color spreading

The phenomenology of the coloration effect peculiar to the neon color spreading reveals the following perceptual qualities: i) the color appears as a diffusion of a little amount of pigment of the embedded chromatic segments; ii) the coloration is transparent like a light, a shadow, or a fog; iii) the way of appearance (*Erscheinungweise*, Katz, 1911, 1930) of the color is diaphanous and can appear in some (not all) cases similarly to a veil that glows like a neon upon the background, like a transparent layer, or (under achromatic conditions) like a dirty, shadowy, foggy or muddy filmy blanket; iv) if the

inset virtual figure is achromatic and the surrounding inducing elements chromatic, the illusory veil appears tinted not in the achromatic color of the embedded elements, as expected, but in the complementary color of the surrounding elements, i.e. the gray components appear spreading reddish or yellowish color when the surrounding components are respectively green or blue (van Tuijl, 1975).

1.2 Figural effect in the neon color spreading

The previous ii)-iii) qualities refer not only to the coloration effect of the neon color spreading but also to its figural effect. Phenomenally, i) the illusory "thing", produced according to the coloration property, has a depth stratification: it can appear in front of or behind the component elements; ii) by reversing the relative contrast of embedded vs. surrounding components, the depth stratification reverses as well, i.e. when the surrounding elements have less contrast than the embedded ones, the inset components appear as a background rather than as a foreground (Bressan, 1993); iii) in both perceptual conditions the illusory "thing" is perceived as a transparent film; iv) the illusory "thing" may assume different roles or may become different phenomenal objects: a "light", a "veil", a "shadow" or a "fog"; v) when the transparent film, usually perceived in front of the stimulus components, is pitted against depth stratification (for example by using stereograms, Nakayama et al., 1990, or flicker-induced depth, Meyer and Doughrty, 1987) the neon color spreading is lost; vi) the neon color spreading reveals the "phenomenal scission" (*Spaltung*, Koffka, 1935; Metger, 1954) of an elevated transparent colored veil and underneath components that appear to amodally continue without changing in color: the physical variation of color of the inset elements is charged to the transparent film, while the variation of color of the surrounding components is phenomenally discharged, so they appear as having the same color.

2. WATERCOLOR ILLUSION

The "watercolor illusion" is a long-range assimilative spread of color sending out from a thin colored line running parallel and contiguous to a darker chromatic contour and imparting a strong figural effect across large areas (Pinna, 1987; Pinna, Brelstaff and Spillmann 2001; Pinna, Werner and Spillmann, 2003; Spillmann, Pinna and Werner, 2004, Pinna, in press; Pinna and Grossberg, submitted).

Geometrically, while the neon color spreading is elicited by the continuation of one segment with a different color, the watercolor illusion occurs through the juxtaposition of parallel lines.

In Figure 2, purple wiggly contours flanked by orange edges are perceived as rows of undefined shapes (polygons, and flower-like shapes different in each row) evenly colored by a light veil of orange tint spreading from the orange edges.

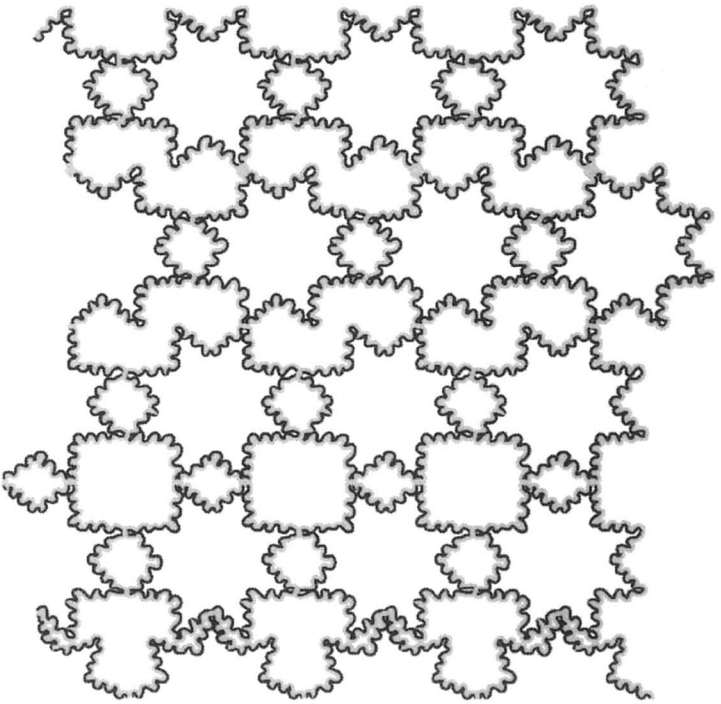


Figure 2. The watercolor illusion: Rows of undefined shapes appear evenly colored by a light veil of orange tint spreading from the orange edges.

In Figure 3 rows of stars are now perceived evenly colored of the same illusory faint orange as in Figure 2. The different coloration and figural results of Figure 2 and 3 are obtained although both figures have the same geometrical structure, and depend on the inversion of the purple and orange lines: the purple/orange wiggly lines in Figure 2 become orange/purple in Figure 3. This reversion affects both the coloration and figural effects of the watercolor illusion: what in Figure 2 appears as illusory tinted and segregated as a figure, in Figure 3 appears as an empty space without a clear

coloration and without a delimited shape (only the figure has a shape but not the background); what is visible in Figure 2 is imperceptible in Figure 3.

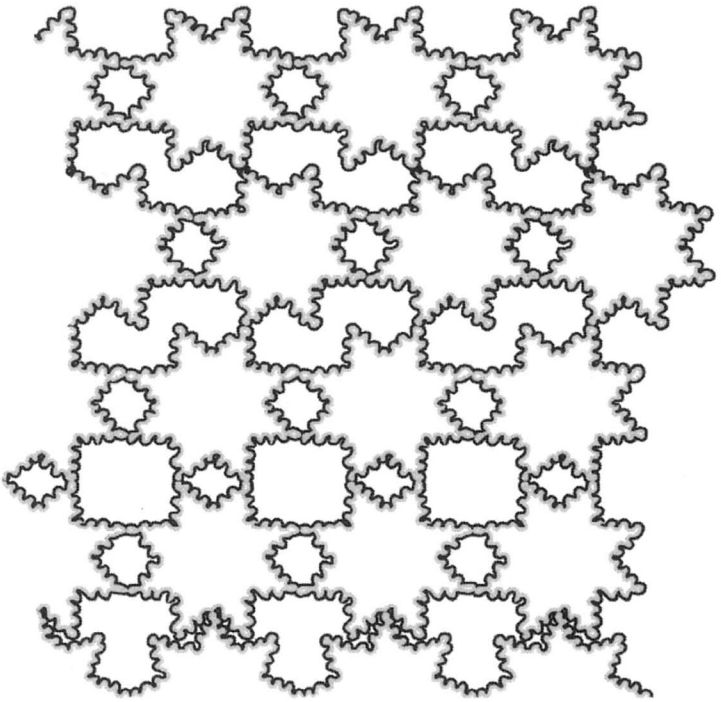


Figure 3. Rows of stars are perceived evenly colored by a light veil of orange tint.

Similarly to the neon color spreading, the watercolor illusion shows both coloration and figural effects.

2.1 Coloration effect in the watercolor illusion

Phenomenally, some coloration qualities analogous to the neon color spreading can be highlighted within the watercolor illusion: i) the illusory color appears as a spreading of some amount of tint belonging to the orange fringe; ii) the coloration does not appear transparent as in the neon color spreading but solid and impenetrable; iii) the illusory color appears epiphanous and as a solid surface color (Katz, 1930).

Compared to the neon color spreading, it has not been demonstrated yet if the watercolor illusion induces a complementary color like the neon color spreading when one of the two juxtaposed lines is achromatic and the other

chromatic. This is the argument of the next experiment. Furthermore, the watercolor illusion differs in the way of appearance of the coloration: transparent vs. solid and impenetrable, and diaphanous vs. epiphanous. The aim of Section 4 is to demonstrate that the watercolor illusion presents ways of appearance of the coloration similar to those of the neon color spreading.

2.2 Figural effect in the watercolor illusion

The watercolor illusion not only determines a long range coloration effect, it also induces a unique figural effect that can compete with the classical Gestalt principles of grouping and figure-ground segregation (Wertheimer, 1923; Rubin, 1915, 1921). Pinna et al. (2001) and Pinna and Grossberg (submitted) demonstrated that, all else being equal, the watercolor illusion determines figure-ground segregation more strongly than the well-known Gestalt principles: proximity, good continuation, prägnanz, closure, symmetry, convexity, past experience, similarity. It was shown (Pinna, in press) that within the watercolor illusion a new principle of figure-ground segregation is contained: "asymmetric luminance contrast principle", stating that, all else being equal, given an asymmetric luminance contrast on both sides of a boundary, the region, whose luminance gradient is less abrupt, is perceived as a figure relative to the complementary more abrupt region perceived as a background. This phenomenal and physical asymmetry across the boundaries makes the figural effect due to the watercolor illusion stronger than in the classical figure-ground conditions, and prevents reversibility of figure-ground segregation. The asymmetric luminance contrast principle strengthens Rubin's notion of unilateral belongingness of the boundaries (Rubin, 1915): The boundaries belong only to the figure and not to the background, which appears as an empty space without a shape. This notion has been also called "border ownership" (Nakayama and Shimojo, 1990).

The main figural qualities of the watercolor illusion are: i) the illusory figure has a univocal (poorly reversible) depth segregation similar to a rounded surface with a bulging and volumetric effect; ii) the resulting surface appears thick, solid, opaque and dense; iii) as shown in Figures 2 and 3, by reversing the colors of the two lines running parallel, figure-ground segregation reverses as well; in these two figures, the border ownership is reversed, i.e. the boundaries belong unilaterally only to one region and not to the other; iv) the figural effect of the watercolor illusion may be perceived in terms of phenomenal scission, in fact it looks like that obtained through the painting technique of *chiaroscuro* (the modeling of volume by depicting light and shade): A highlight marks the point where the light is most directly (orthogonally) reflected, moving away from this highlight, light hits the

object less directly and therefore has a darker value of gray. The scission is between a homogenous or uniformly colored object and a reflected light on a rounded object. This scission, by sculpturing the coloration, sculptures even the shape that appears as having a volumetric 3D pictorial shape.

The figural effect of the neon color spreading compared with that of the watercolor illusion shows again some differences: respectively, transparency vs. opaque and dense appearance; respectively, appearance as a "light", a "veil", a "shadow" or a "fog" vs. rounded thick and opaque surface bulging from the background.

Despite the differences between the two illusions, and particularly despite the different perceptual roles assumed by illusory things, the two effects are very similar structurally in their strong color spreading and clear depth segregation. These similarities can suggest common basic mechanisms to explain both illusions. Furthermore, the different coloration and figural roles and the two kinds of phenomenal scissions, as previously described, can be attributed to the geometrical differences between the two illusions, respectively, continuation of a segment in a different color and juxtaposition of at least two lines.

The questions to be answered are: can the watercolor illusion assume figural properties like the neon color spreading under geometrical conditions different from those used in Figure 2 and 3? This is the main topic of the phenomenal new cases presented in Section 4.

Summing up, i) the aim of the next experiment (Section 3) is to complete the coloration comparisons between the two phenomena by testing whether the watercolor illusion can induce complementary color under conditions similar to those of the neon color spreading; ii) the aim of the new cases presented in Section 4 is to demonstrate that, under different geometrical-figural conditions, the watercolor illusion manifests figural properties analogous to those of the neon color spreading. The results are discussed in the light of a limiting case (Section 5) common to both illusions that can explain similarities and differences between the two phenomena. Finally, two parallel and independent processes as proposed within the FACADE model (Grossberg, 1994, 1997) are suggested to account for the coloration and figural effects in both neon color spreading and watercolor illusion.

3. EXPERIMENT: WATERCOLOR ILLUSION AND COMPLEMENTARY COLOR INDUCTION

It is well known (van Tuijl, 1975) that in the neon color spreading, when inset elements are achromatic and surrounding ones are chromatic, the illusory color spreading, occurring within the inset elements, appears in the

complementary color of the surrounding inducing components (see Figure 4).

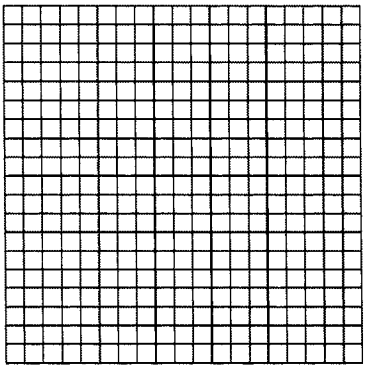


Figure 4. When inset elements are black and the surrounding ones are blue, the illusory color spreading within the black elements appears yellow.

The demonstration of this effect also in the watercolor illusion (see Figure 5) strengthens the links and similarities between the two illusions and suggests a common mechanism for the coloration effect in both phenomena.

Figure 5. The jagged annulus appears evenly colored by a light veil of yellowish tint complementary to the blue outer edges.

3.1 Subjects

Different groups of fourteen naive subjects participated to each experimental condition. All had normal or corrected-to-normal vision. The stimuli were presented in a different random sequence for each subject.

3.2 Stimuli

The stimuli were composed by two conditions - neon color spreading and watercolor illusion - with four stimuli each, where the color of the surrounding components were blue, green, yellow, and red. The CIE x, y chromaticity coordinates for the stimuli were: blue (0.201, 0.277), green (0.3, 0.5), yellow (0.46, 0.42), and red (0.54, 0.33).

Stimuli were hand-drawn chromatic/achromatic contours, in continuation for the neon color spreading condition and running parallel for the watercolor condition, on a white background. The stroke width of the magic marker was approx 6 arcmin.

Figure 4 was the basic stimulus for the neon color spreading condition, while Figure 5 was the one for the watercolor condition.

The overall size of the stimuli was about 21X15 deg of visual angle. Luminance contrast for stimulus components (Lx) was defined by the ratio (Lwhite background - Lx) / Lwhite background. The luminance of the white (background) paper was 80.1 cd/m^2. Black lines had a luminance contrast of 0.97. Stimuli were presented under Osram Daylight fluorescent light (250 lux, 5600° K) and were observed binocularly from a distance of 50 cm with freely moving eyes.

3.3 Procedure

The task of the subjects was to report the perceived color within the black region by naming it, say yellowish, reddish, etc. There was a training period preceding each experiment to familiarize subjects with the color spreading in the neon color spreading, in the watercolor illusion and with the task. During practice, subjects viewed some examples of neon color spreading and watercolor illusion different from the stimuli to familiarize them with both coloration and figural effects. Observation time was unlimited.

3.4 Results

In Figure 6, the number of subjects perceiving and naming the color within and between the black lines is plotted as a function of the four chromatic colors of the surrounding elements for the two conditions: neon color spreading and watercolor illusion. The results clearly showed that for a significant number of subjects the perceived coloration within the black elements was of the color complementary to the chromatic color of the surrounding lines. The results for both conditions did not differ qualitatively. No statistical comparison was done, because of the different stimulus patterns used.

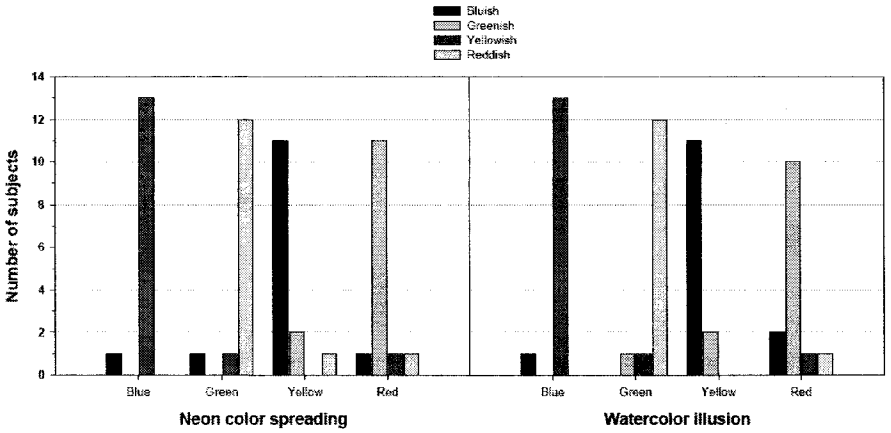


Figure 6. Number of subjects perceiving and naming the color within and between the black lines plotted as a function of the four chromatic colors of the surrounding elements for the two conditions: neon color spreading and watercolor illusion.

It is interesting to notice that 10 subjects spontaneously described the stimuli of the watercolor condition as if the complementary color within the black region (see Figure 5) was like a surface and of opaque coloration peculiar to a figure, while the surrounding blue coloration was like a halo of blue light or a blue illumination spreading everywhere along the background and coming from behind the figure (back-lighting). This result demonstrates that the coloration effect of the watercolor illusion may appear as having different phenomenal ways of appearance, hence making the figural differences between neon color spreading and watercolor illusion more contest-dependent than absolute differences. This hypothesis is the topic of the following Section.

4. WATERCOLOR ILLUSION AND WAYS OF APPEARANCE OF COLORATION AND FIGURAL EFFECTS

By changing the geometrical conditions, the watercolor illusion manifests different coloration and figural effects. When the figure is segregated independently from the presence of a colored fringe, that appears now turned toward the background, the resulting color spreading of the fringe does not

assume surface color properties, but properties belonging to the background:

the illusory coloration is perceived diaphanous like a foggy coloration diffusing everywhere in the background. Figure 7 shows one example of the new background properties belonging to the watercolor illusion (see also Pinna and Grossberg, submitted).

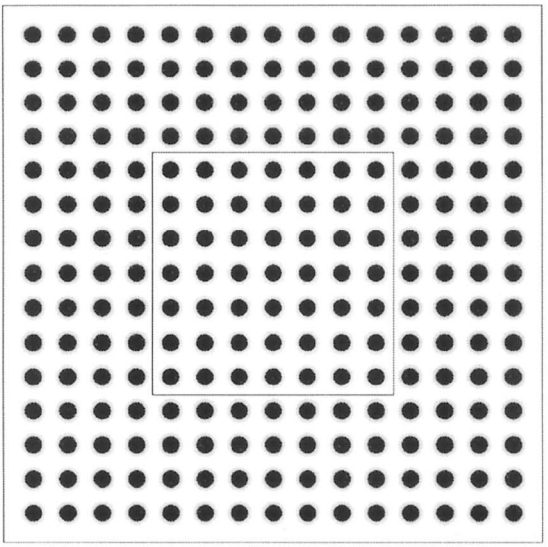


Figure 7. A light blue coloration spreading from the inset square of elements is surrounded by a red spreading. Both coloration effects are not followed by a figural effect with a volumetric property, but the coloration appears as a diaphanous color like a foggy veil of color.

Another example is illustrated in Figure 8. Here the illusory coloration gives to the illusory star a fuzzy luminous quality. While in Figure 7 the coloration is part of the background, in Figure 8 it is a property of the figure (the star), however it differs from the strong surface and volumetric appearance peculiar to Figures 2 and 3. Its inner surface appears brighter and yellowish, but foggy, soft and smooth.

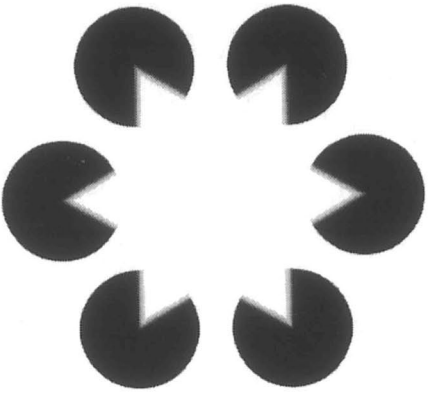


Figure 8. The illusory coloration of the star appears fuzzy and luminous.

A similar effect but more volumetric than that of Figure 7, resembling to the chiaroscuro, is illustrated in Figure 9 under different conditions (see also Pinna and Grossberg, submitted).

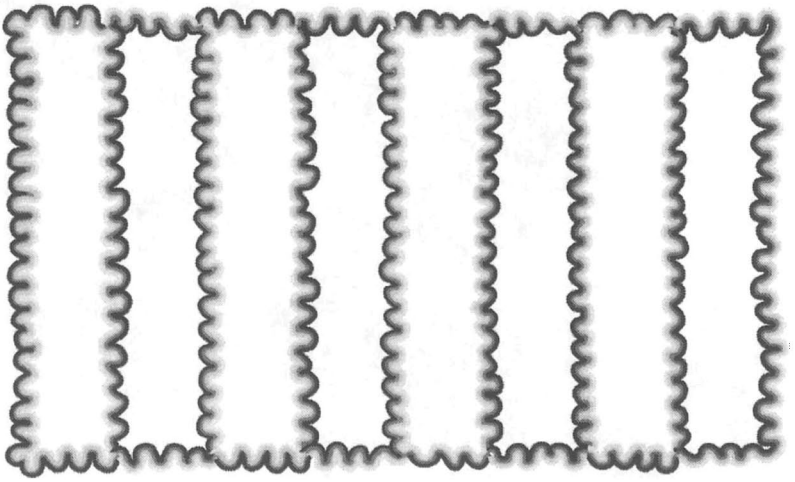


Figure 9. The illusory coloration of the columns appears volumetric.

Figure 10 shows another figural role that the watercolor illusion can assume: transparency (see Pinna and Grossberg, submitted).

At first sight, the two halves of the figure appear alike, but in the two halves the purple and orange colors are in cross-continuation (purple with

orange and orange with purple). Under these conditions, because of the reversed contrast of the two halves, if within one half, the frame appears slightly orange, within the second, it appears bluish. The opposite is true in the small inner square. Nevertheless, both halves manifest the same figural effects: The interspace between the two square-like shapes is perceived as a transparent figure or as a transparent frame, despite the differences in the color fringes and in the inner coloration effect. Summing up, in Figure 10 the same figural transparent effects and different chromatic colorations are seen in the two halves, even if no clear and immediate contradiction is perceived. These phenomenal results are certainly in agreement with the watercolor illusion as a grouping or figure-ground segregation principle but in disagreement with the similarity principle: the two halves are dissimilar; therefore they should not be grouped.

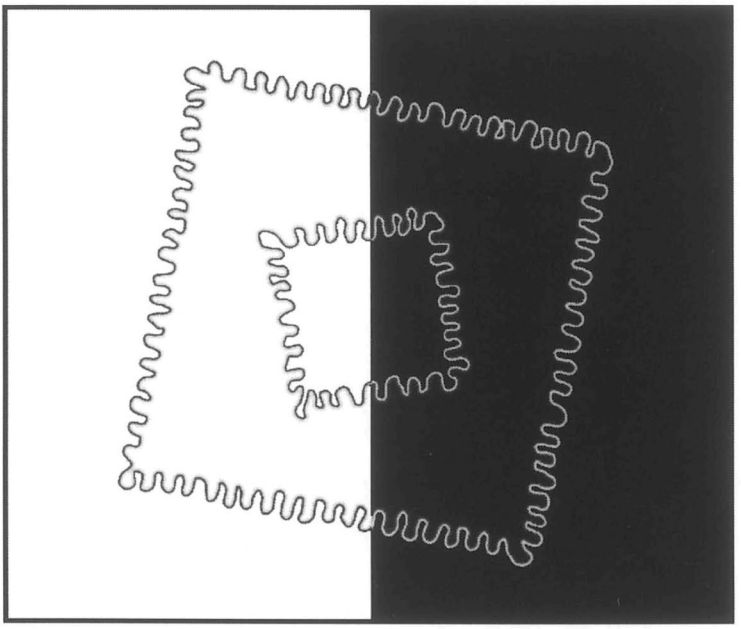


Figure 10. A transparent watercolored surface.

An interesting case is illustrated in Figure 11, where a direct comparison between a quasi equiluminant condition and a high contrast difference between the two juxtaposed color lines, induces different coloration and figural effects: near the quasi equiluminant conditions the coloration appears not as a surface color but as an ethereal soft coloration without any figural or background properties; near the high contrast difference the figural effect and the surface color properties are restored.

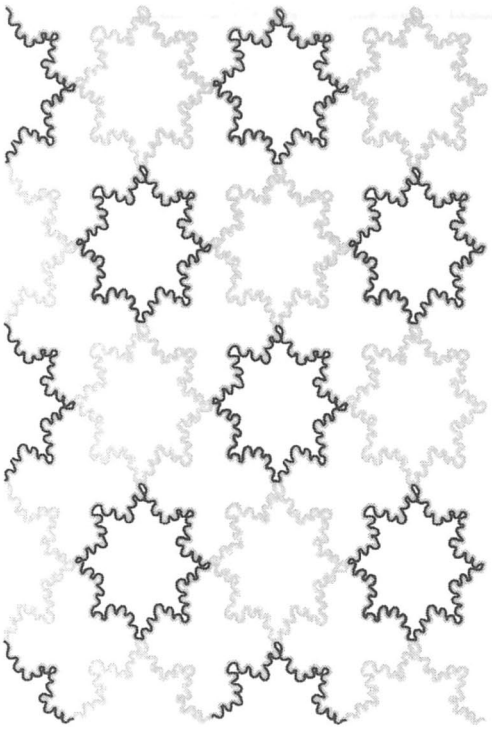


Figure 11. The quasi equiluminant adjacent lines (gray and red) show an ethereal soft coloration without any figural or background properties; the high contrast adjacent lines (black and red) show a clear figural effect and a surface color property.

5. NEON COLOR SPREADING AND WATERCOLOR ILLUSION REDUCED TO A LIMITING CASE

The phenomenal results obtained through these watercolor conditions plus the results of the previous experiment suggest some hypotheses useful to draw a bridge between the two illusions.

First of all, the variety of coloration and figural effects within the watercolor illusion are mainly accompanied to geometrical variations that influence the figural effect. These variations suggest that the ways of appearance of the coloration is strongly linked to the figural properties. If this is true the switch between the neon color spreading and the watercolor illusion may depend from different geometrical properties (continuation or juxtaposition of lines) inducing different figural and, as a consequence,

different coloration effects. This hypothesis is supported by Figures 7, 8 and 9 that are geometrically and phenomenally in between neon color spreading and watercolor illusion. The geometrical difference may activate neural dynamics that deeply interact to create a whole phenomenal result where the two effects are synergistically reinforcing each other, as defined previously in terms of phenomenal scission.

Second, because the variety of ways of appearance of the watercolor illusion cannot be obtained in the neon color spreading, the watercolor illusion can be considered as a more general case including the more specific neon color spreading condition.

Third, given this variety of appearances the two illusions can be reduced to a more simple geometrical condition that may be considered as a limiting case that can explain similarities and dissimilarities between neon color spreading and watercolor illusion. Because neon color spreading is defined by the continuation of lines (see Figure 12), while the watercolor illusion is defined by juxtaposition, the two illusions can be firstly combined as illustrated in Figure 13 and secondly as in Figure 14.

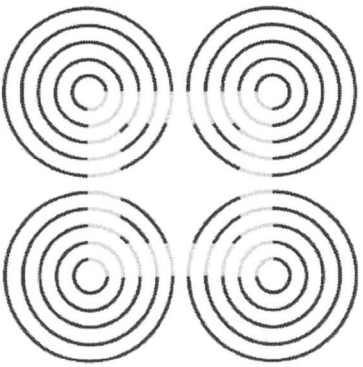


Figure 12. The neon color spreading defined by the continuation of lines.

In Figure 12, the continuation of purple surrounding arcs in orange arcs, so creating a square annulus, induce a clear neon color spreading, whose coloration and figural properties appear not to glow as in Figure 1 (possibly due to the high contrast between the two colors) but more as a transparent orange veil.

Now, if the orange inset arcs are reduced to very small arcs or dots, as illustrated in Figure 13, a condition in between neon color spreading and watercolor illusion is created: the inducing elements are lines that continue in dots and at the same time the termination of each inducing arc present a juxtaposed dot. So the same figure can be read in two ways: from the neon

color spreading point of view and from the watercolor illusion perspective. The phenomenal result shows a clear coloration effect not weaker than that of Figure 12, and a figural effect that differs from the known neon color conditions and more similar to Figures 8 and 9: a fuzzy illusory square annulus yellowish and brighter than the background.

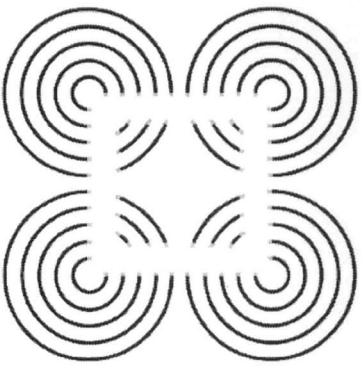


Figure 13. A condition in between neon color spreading and watercolor illusion.

By reducing the surrounding purple arcs to dots (as illustrated in Figure 14), the geometrical conditions are more similar to those of the watercolor illusion. It has been already shown that the watercolor illusion also occurs not only by using juxtaposed lines but also by using juxtaposed chains of dots (see Pinna et al., 2001). Under these conditions both coloration and figural effects become weaker and weaker as the density of the dots becomes sparser and sparser. Unlike neon color spreading, there is no transparency effect.

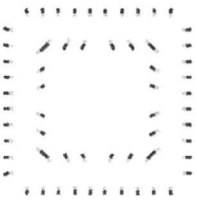


Figure 14. The two-dots limiting case can be considered as the basis for a common neural model to account for the neon color spreading and the watercolor illusion.

Figure 14 may represent the two-dots limiting case useful to find mechanisms for the coloration and figural effects common to all the

coloration and figural variations considered (see also Pinna and Grossberg, submitted).

On the basis of these results, it can be said that (i) the neon color spreading and the watercolor illusion have a common limiting case: The two-dots juxtaposition; (ii) this limiting case can be considered as the basis for a common neural model to account for both illusions; (iii) the perceived coloration and figural differences between the two illusions depend on the geometrical differences that elicit different local color interactions and different figural organizations; (iv) coloration and figural effects may derive from parallel processes, i.e. at a feature processing stage, the small interaction area around and between the two dots produces the color spreading common to both illusions, and at a parallel boundary processing stage, the different geometrical structures in both illusions produce the different figural effects.

Color spreading may arise in two steps: First, weakening the contour by lateral inhibition between differentially activated edge cells (local diffusion); and second, flowing the color onto the enclosed area (color diffusion). Edge polarity neurons in areas V2 and V4 of the monkeys responding to a luminance step in one, but not the other, direction may be responsible for border ownership.

The next section proposes how the FACADE neural model of 3D vision and figure-ground separation can explain these effects.

6. FACADE EXPLANATION OF NEON COLOR SPREADING AND THE WATERCOLOR ILLUSION

The separation between the coloration and figural effects suggests different mechanisms for color spreading and the figural effect. The FACADE model (Grossberg, 1994, 1997) assumes that parallel boundary grouping and surface filling-in processes are respectively defined by the Boundary Contour System (BCS) and Feature Contour System (FCS) (Cohen and Grossberg, 1984; Grossberg and Mingolla, 1985a, 1985b; Grossberg and Todorovic, 1988). The two processes are realized by the cortical interblob and blob streams within cortical areas V1 through V4. These boundary and surface processes show complementary properties (Grossberg, 2000). Boundaries are oriented and are insensitive to contrast polarity or, in other words, boundaries pool contrast information at each position from opposite contrast polarities. Surfaces fill-in outwardly from individual lightness or color inducers in an unoriented way using a process

that is sensitive to contrast polarity. The two systems can explain neon color spreading (Grossberg, 1987, 1994, 2000; Grossberg and Mingolla, 1985a; Grossberg and Swaminathan, 2004; Grossberg and Yazdanbakhsh, 2004).

The watercolor illusion can be explained by a process of spatial competition where stronger inputs to the boundary system occur at the edges of higher contrast colored lines than at lower contrast ones. Thus the layer 6-to-4 spatial competition is stronger from the boundaries of higher contrast edges to those of lower contrast edges than conversely. The boundaries of the lower contrast edges are thus weakened more by competition than the boundaries of the higher contrast edges. Hence more color can spread across these boundaries than conversely.

A similar idea has been used to explain why neon color spreading is sensitive to the relative contrasts of the edges at which neon color is released (Grossberg and Mingolla, 1985a). For a wider and more exhaustive discussion on the neon color spreading and the watercolor illusion see Pinna and Grossberg (submitted).

FACADE theory also proposes how two-dimensional monocular properties of the BCS and FCS may be naturally embedded into a more comprehensive theory of 3-D vision and figure-ground separation (Grossberg, 1987, 1994, 1997, 2004) that is the best candidate to explain the different figural roles assumed by the coloration effect in both neon color spreading and watercolor illusion. This idea has been developed in a series of quantitative studies to explain several different types of perceptual and neural data linked to the 3-D vision (Grossberg and Howe, 2003; Grossberg and Kelly, 1999; Grossberg and McLoughlin, 1997; Grossberg and Pessoa, 1998; Grossberg and Swaminathan, 2004; Grossberg and Yazdanbakhsh, 2004; Kelly and Grossberg, 2000; McLoughlin and Grossberg, 1998), however it needs to be further developed to fully explain the qualitative ways of appearance of the watercolor illusion showed.

ACKNOWLEDGMENTS

This research was supported by Fondazione Banco di Sardegna, the Alexander von Humboldt Foundation, ERSU, Banca di Sassari and Fondo d'Ateneo (ex 60%).

I thank Stephen Grossberg for precious comments and suggestions and Massimo Dasara for assistance in testing the subjects.

REFERENCES

Bressan, P., 1993, Neon color spreading with and without its figural prerequisites, *Perception* **22**:353-361.

Bressan, P., Mingolla, E., Spillmann, L., and Watanabe, T., 1997, Neon colour spreading: a review, *Perception* **26**:1353-1366.

Cohen, M., and Grossberg, S., 1984, Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance, *Perception & Psychophysics* **36**:428-456.

Grossberg, S., 1987, Cortical dynamics of three-dimensional form, color, and brightness perception II: Binocular theory, *Perception & Psychophysics* **41**:117-158.

Grossberg, S., 1994, 3-D vision and figure-ground separation by visual cortex, *Perception & Psychophysics* **55**:48-120.

Grossberg, S., 1997, Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures, *Psychological Review* **104**:618-658.

Grossberg, S., 2000, The complementary brain: Unifying brain dynamics and modularity, *Trends in Cognitive Sciences* **4**:233-245.

Grossberg, S., 2004, How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex, *Behavioral and Cognitive Neuroscience Reviews*, (in press).

Grossberg, S. and Howe, P., 2003, A laminar cortical model of stereopsis and three-dimensional surface perception, *Vision Research* **43**:801-829.

Grossberg, S., and Kelly, F., 1999, Neural dynamics of binocular brightness perception, *Vision Research* **39**:3796-3816.

Grossberg, S., and McLoughlin, N. P., 1997, Cortical dynamics of three-dimensional surface perception: Binocular and half-occluded scenic images, *Neural Networks* **10**:1583-1605.

Grossberg, S., and Mingolla, E., 1985a, Neural dynamics of form perception. Boundary completion, illusory figures and neon color spreading, *Psychological Review* **92**:173-211.

Grossberg, S., and Mingolla, E., 1985b, Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations, *Perception & Psychophysics* **38**:141-171.

Grossberg, S., and Pessoa, L., 1998, Texture segregation, surface representation and figure-ground separation, *Vision Research* **38**:1657-1684.

Grossberg, S., and Swaminathan, G., 2004, A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: Development, attention, and bistability, *Vision Research*, (in press).

Grossberg, S., and Todorovic, D., 1988, Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena, *Perception & Psychophysics* **43**:241-277.

Grossberg, S., and Yazdanbakhsh, A., 2004, Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading, (in press).

Katz, D., 1911, Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung, *Zeitschrift für Psychologie* 7:6–31,(Leipzig: Barth).

Katz, D., 1930, *Die Erscheinungsweisen der Farben*, 2nd edition, (Translation into English: MacLeod R. .B, and Fox C. W., 1935, *The World of Color*, Kegan Paul, London).

Kelly, F., and Grossberg, S., 2000, Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception, *Perception & Psychophysics* **62**:1596-1618.

Koffka, K., 1935, *Principles of Gestalt Psychology*, Harcourt Brace, New York.

McLoughlin, N. P., and Grossberg, S., 1998, Cortical computation of stereo disparity, *Vision Research* **38**:91-99.

Metzger, W., 1954, *Psychologie, Die Entwicklung ihrer Grundannahmen seit der Einführung des Experimentes*, Zweite Auflage (Darmstadt: Steinkopff).

Meyer, G. E., and Dougherty, T., 1987, Effects of flicker-induced depth on chromatic subjective contours, *Journal of Experimental Psychology: Human Perception and Performance* **13**:353-360.

Nakayama, K., Shimojo, S., and Ramachandran, V. S., 1990, Transparency: Relation to depth, subjective contours, luminance, and neon color spreading, *Perception* **19**:497-513.

Nakayama, K., and Shimojo, S., 1990, Towards a neural understanding of visual surface representation, *Cold Spring Harbor Symposia on Quantitative Biology* **40**:911–924.

Pinna, B., 1987, Un effetto di colorazione, in: *Il laboratorio e la città, XXI Congresso degli Psicologi Italiani*, V. Majer, M. Maeran and M. Santinello, eds., 158.

Pinna, B., (in press), The role of the Gestalt principle of similarity in the watercolor illusion, *Spatial Vision.*

Pinna, B., Brelstaff, G., and Spillmann, L., 2001, Surface color from boundaries: A new 'watercolor' illusion, *Vision Research* **41**:2669–2676.

Pinna, B., Werner J. S., and Spillmann, L., 2003, The watercolor effect: A new principle of grouping and figure-ground organization, *Vision Research* **43**:43–52.

Pinna, B, and Grossberg, S., The watercolor illusion: new properties and neural mechanisms, *Journal of Vision*, (submitted).

Rubin, E., 1915, *Synsoplevede Figurer*, Kobenhavn: Glydendalske.

Rubin, E., 1921, *Visuell Wahrgenommene Figuren*, Kobenhavn: Gyldendalske Boghandel.

Spillmann, L., Pinna, B., and Werner, J. S., 2004, (in press), Form-from-watercolour in perception and old maps, in: *Seeing Spatial Form*, M. R. M., Jenkin and, L. R., Harris, Eds., Oxford University Press.

van Tuijl, H. F. J. M., 1975, A new visual illusion: neon-like color spreading and complementary color induction between subjective contours, *Acta Psychologica* **39**:441-445.

van Tuijl, H. F. J. M., and de Weert, C. M. M., 1979, Sensory conditions for the occurrence of the neon spreading illusion, *Perception* **8**:211-215.

Varin, D., 1971, Fenomeni di contrasto e diffusione cromatica nell'organizzazione spaziale del campo percettivo, *Rivista di Psicologia* **65**:101–128.

Wertheimer, M., 1923, Untersuchungen zur Lehre von der Gestalt II, *Psychologische Forschung* **4**:301–350.

USABILITY AND MAN-MACHINE INTERACTION

Maria Pietronilla Penna and Roberta Rani
Dipartimento di Psicologia, Università di Cagliari, Via Is Mirrionis, 09100 Cagliari, Italy

Abstract: This paper deals with the issue of usability of computer software, mainly as regards the one supporting web navigation. It is evidenced how the concept of usability cannot be defined in an objective way, bur results from an interaction between user characteristics and environmental demands. It is proposed that the study of such a question should start from a monitoring of user expertise. In this regard the paper introduces a new model for designing web search engines, called Guest Resource Model. It states that each expertise level should trigger the operation of a different kind of search engine. A possible practical implementation of this model is discussed, having in mind the goal of protecting children accessing Internet from the dangers of web navigation.

Key words: software usability; user characteristics; user expertise; web navigation; search engines.

1. INTRODUCTION

The recent revolution in personal computer technology and falling hardware prices are making personal computers available to ever broader groups of users, for a more and more larger variety of tasks. Such a situation has given rise to the emergence of new needs, not existing in the era in which computers were used only by a small number of people. Amongst these needs one of the most important is related to the search for a satisfactory definition of software usability. This question has a strong systemic valence. Namely it is impossible to define the concept of software usability per se, that is only on the basis of objective, measurable features, whose occurrence can be detected by every observer in a whatever context. On the contrary, software usability (see, for instance, Nielsen, 1993) is a construct emergent from the interaction between particular users and

computers within a complex system including human beings, endowed with a cognitive and an emotional system, having goals, beliefs, fears, mental schemata, interacting with a continuously changing environment, reacting to their actions and source of new needs, questions, situations. In this paper we will therefore deal with the problem of defining software usability within a systemic framework, focussing mainly on the role played by subject expertise in interacting with computer programs.

2. THE CONCEPT OF SOFTWARE USABILITY

What does usability mean? This concept has multiple components (see Preece, 1994; Landauer, 1996; Vicente, 1999) and is traditionally associated to five usability attributes: learnability, efficiency, memorability, error propensity, satisfaction. They can be roughly defined as follows:

- Learnability is related to the fact that the use of a software should be easy to learn so that the user can rapidly start getting some works through the software itself.
- Efficiency is related to software performance, so that, once the user has learned to use it, a high level of productivity is possible.
- Memorability is related to the fact that the instructions for use should be easy to remember, so that the casual user be able to use again the software after some time of inactivity, without having to learn everything all over again.
- Error propensity is related to the fact that software use is characterized by a low user error rate, besides, when the user makes errors, he can easily recover from them; further, catastrophic errors must not occur.
- Satisfaction is related to the fact that software should be pleasant to use, so that users are subjectively satisfied when using it.

The first step in studying software usability is to analyze the intended users and the goals to be reached through its use. Individual user characteristics and variability in tasks are the two factors with the largest impact on usability, so they need to be studied carefully. Of course the concept itself of user should be defined in such a way as to include everybody whose work is affected by the software under consideration in some way.

The individual user characteristics let us know the class of people who will use the system. In many cases this knowledge is very easy to obtain, since it is possible to identify these users with concrete individuals. An account of users' work experience, age, educational level, previous computer experience, and so on, makes possible to anticipate their learning difficulties

to some extent and to better set appropriate limits for the complexity of the user interface (see Helander et al., 1997). For instance the latter must be designed in the simplest way if users are expected to use it with minimum training.

The most basic advice with respect to interface evaluation is simply to do it, and especially to perform some user testing (see, in this regard, Lindgaard, 1994; Gray and Salzman, 1998). The benefits of employing some reasonable usability engineering methods to evaluate a user interface rather than releasing it without evaluation are much larger than the incremental benefits of using exactly the right methods for a given project.

3. THE ROLE OF USER EXPERTISE

The user expertise is probably one of most important factors in assessing software usability. Namely an inexperienced user can lower the efficiency of most software programs, despite the efforts of programmers and of experts in software usability. In this regard, the most suited example is given by Internet, which, at first sight, appears as the definitive solution to give full power to user, but whose efficiency is counteracted by the fact that not all "users" are able to correctly navigate in the World Wide Web.

Unfortunately all interface design strategies have been for long time oriented to support only scientific communication. As this form of communication is typical of very specific kinds of users, endowed with high expertise, such a circumstance induced to neglect the important role played by the inexperience of most users.

This problem is particularly serious when the users belong to special categories, such as, for example, children. They haven't enough experience and their interaction with Internet gives rise to the need for their protection against the dangers implicit in every navigation activity. In this regard we propose a new interface design model that we called Resource Guest model (RGM). Within it the users are divided in three types, that is:

- Inexpert user
- Middle user
- Expert user

As regards Internet the RGM implies that the research engine used must change with a change of user type. This can be obtained through two different steps. In the first step a new password is created. The latter serves as a sort of right to use when approaching Internet. The user is endowed with a card, containing his/her data and the associated password, but this "bond" isn't a limit for user, because he/she will be also protected against navigation dangers.

A second step consists in preventing from the access to some websites when in presence of not suitable user features, as occurring, for example, when the users are children. All this could be possible with the utilization of a password (the one introduced in the first step) that the user puts in before starting the search of the wanted items. This password can be a method for allowing the possibility of access only to expert users.

Such a strategy could help in increasing web usability, owing to the fact that each user has his own password, obtained through a suitable registration phase. In this way, in agreement with the principles of RGM, we can implement an association between single user characteristics and web resources available to him/her. Of course the research engines should be modified in order to conform to RGM. In this way we could reduce informatics crimes without renouncing to user freedom of choice.

In the Figure 1 we describe through a picture how the second password could appear on the page of an Internet connection.

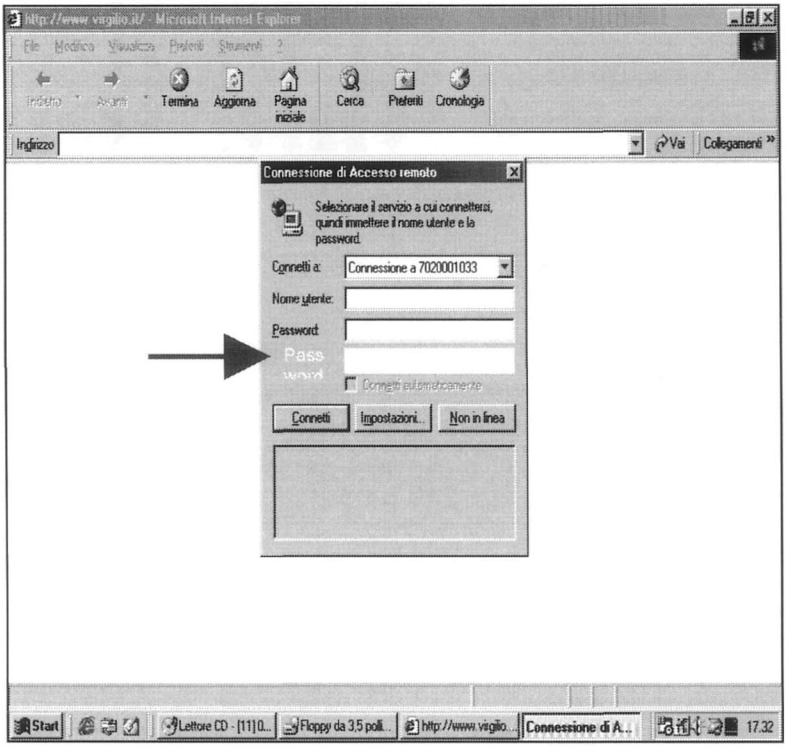


Figure 1. A hypothetical Internet connection window with the second password.

4. CONCLUSIONS

As we discussed before, the concept of software usability cannot be defined without taking into account the whole system of user-environment relationships. However the formidable difficulties of the enterprise aiming to capture the essentials of this concept will never be surmounted if we don't begin to deal with specific problems, such as the one of user experience, on which we focussed within this paper. In this regard we proposed the RGM, giving suitable criteria, based on a knowledge of subject's expertise level, for designing interfaces allowing the user for an easy and productive use of the resources effectively available to him/her. In such a way we can simultaneously protect the user against the dangers deriving from an access to resources requiring a high level of emotional and cognitive maturation.

We proposed a possible implementation of these criteria in the case of Internet access, through the introduction of a second password associated to a list of user features. So far, this solution appears as the best one for granting a right compromise between freedom of navigation and protection against informatics crimes. We feel, however, that a validation of this proposal cannot come only from theoretical analyses of interface design or of human-machine interactions, but mostly from a experimental activity, performed according to the rules of scientific inquiry, about this new form of interfacing. Such an activity will require, of course, a strong transdisciplinary effort.

REFERENCES

Gray, W. D. and Salzman, M. C., 1998, Damaged merchandise? A review of experiments that compare usability evaluation methods. *Human-Computer Interaction* **13**:203-261.

Helander, M., Landauer, T. K. and Prabhu, P. V., Eds., 1997, *Handbook of Human-Computer Interaction,* 2nd edition, North Holland, Amsterdam.

Landauer, T. K., 1996, *The Trouble with Computers: Usefulness, Usability, and Productivity*, MIT Press, Cambridge, MA.

Lindgaard, G., 1994, *Usability Testing and System Evaluation: A Guide for Designing Useful Computer Systems*, Chapman and Hall, London.

Nielsen, J., 1993, *Usability Engineering*, Academic Press, Boston, MA.

Preece, J., 1994, *Human Computer Interaction*, Addison-Wesley, Reading, MA.

Vicente, K. J., 1999, *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*, Lawrence Erlbaum, Hillsdale, NJ.

OLD MAPS AND THE WATERCOLOR ILLUSION: CARTOGRAPHY, VISION SCIENCE AND FIGURE-GROUND SEGREGATION PRINCIPLES

Baingio Pinna and Gavino Mariotti
Facoltà di Lingue e Letterature Straniere, University of Sassari,

Abstract: Long-range color spreading is studied in two experiments in which a purple contour is flanked by an orange contour. The color faintly but uniformly fills in the surface from the orange contour inducing a strong figural effect and is therefore known as the Watercolor illusion. The figural effect of the watercolor illusion is compared with classical Gestalt factors (past experience, similarity and symmetry). The results of the experiments reveal a more effective role of the watercolor illusion in form perception and in figure-ground segregation than the one of the past experience principle and, in addition, corroborate the perceptual effects in distinguishing and demarcating geographical regions obtained by early cartography using the outline-color technique. The results are explained in terms of Grossberg's FACADE neural model of biological vision.

Key words: historical cartography, watercolor illusion, grouping principles, unilateral belongingness of boundaries, figure-ground organization, neural model.

1. COLORATION AND FIGURAL EFFECTS IN EARLY CARTOGRAPHY

Different colors are used by most of contemporary cartographers to distinguish, for example, different political regions, as illustrated in Figure 1.

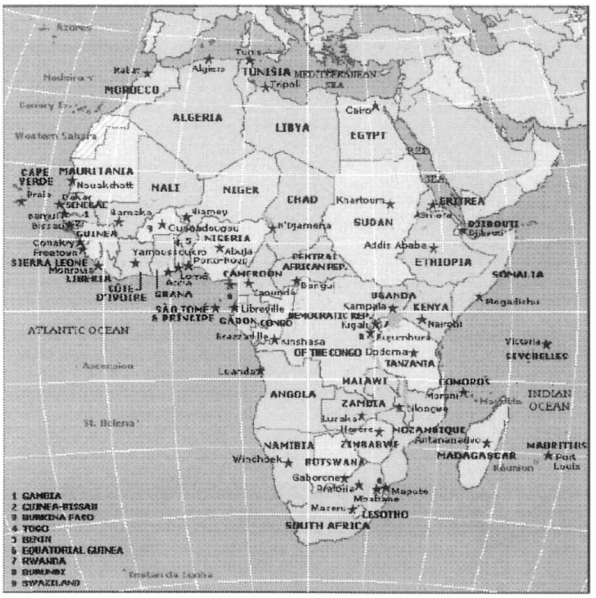


Figure 1. Different colors are used to distinguish different African political regions.

Under these conditions each state is segregated from the neighbor ones in accordance with the similarity principle (Wertheimer, 1923), stating that, all else being equal, the most similar elements (in color, brightness, size, shapes, etc.) are grouped together and appear segregated from dissimilar elements.

In Figure 1, the same principle distinguishes lands from the sea that appears uniformly colored with a blue tint. No strong figural properties are here involved. The resulting perceptual effect shows flat regions without any clear object demarcation, whose basic properties are: (i) the presence of a boundary defining the shape of the object; (ii) a brightness and color determining the appearance of the object; (iii) a segregation in depth with respect to the background. These properties can be better understood in the light of figure-ground segregation. Figure and background have opposite qualities: (i) a figure has a surface color that appears compact and opaque; the background, on the other hand, has a diaphanous color and appears empty; (ii) figure and ground are both linked together through the presence of a border; the figure is defined by a border and appears to lie above the background; on the contrary, the background is unlimited and continuous underneath the figure; (iii) the border surrounding the figure is unidirectional, i.e. it belongs to the figure, not to the background.

Cartographers may have known these principles a long time ago pushed possibly by the need to show strong and clear differentiations and figural demarcations between regions (see Pinna, Brelstaff and Spillmann, 2001; Wollschläger, Rodriguez and Hoffman, 2002). Thus coloration alone as defined by the similarity principle might not be enough. Figural effects to strengthen the differentiation were probably required. In figure 1 the figural properties are not as strong as the ones illustrated in Figure 2, where a XVI century Map drawn by Joan Blaeu (1663; see also Donkersloot de, 1973; Schilder, 1993; Wollschläger, Rodriguez and Hoffman, 2002), using the outline-color technique (about the use of this technique by early cartographers see Bagrow and Skelton, 1985), shows how different colors can be used as fringes to enhance the different parts of a geographical region. This technique reveals how the surface color of a figure is defined by two flanked borders surrounding it. Under these conditions, not only coloration but also figural properties are strengthened: the whole Africa (from Le Grand Atlas ou Cosmographie Blaviane, Vol.10: Africa, 1663. British Library Maps C.5.b.1) appears more strongly segregated in depth and with enhanced figural properties compared with the Africa illustrated in Figure 1.

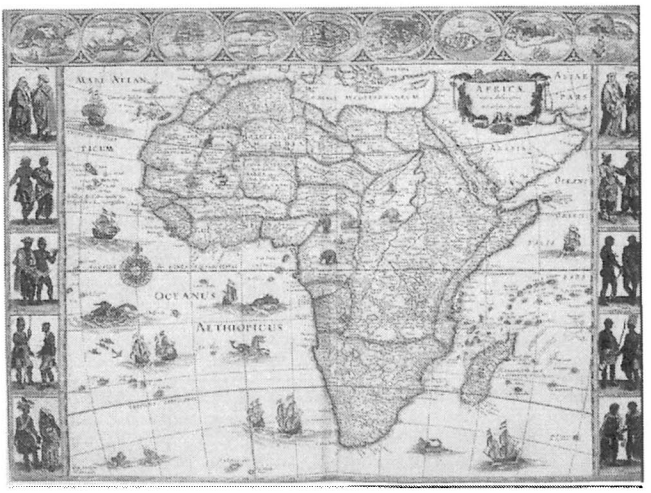


Figure 2. 16th century Map where the outline-color technique enhances different African regions.

The clear figural segregation occurs not only among different regions of the map but most significantly between lands and sea as it is illustrated in Figures 3 (from Le Grand Atlas by Joan Blaeu, Vol.XI: Asia, 1667. British Library Maps F.1010038) and 4 (from Le Grand Atlas by Joan Blaeu,

Vol.XII: America, 1667. British Library Maps F.1000002), where the figure-ground segregation is further strengthened.

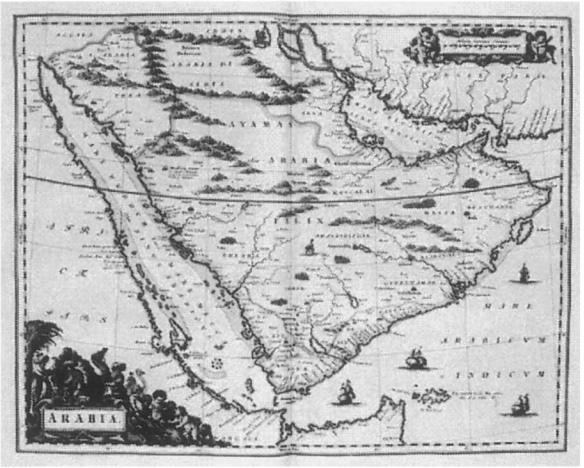


Figure 3. A clear figural segregation between lands and sea is obtained by using the outline-color technique.

Figure 4. The outline-color technique shows how different colors can be used as fringes to enhance figure-ground segregation.

Although the idea of cartographers was shown just through illustrations and not scientifically defined in terms of grouping or figure-ground segregation, it phenomenologically anticipated Rubin's (1915, 1921) unilateral belongingness of the margin (*Zusammmengehörigkeit*) rule: the

contour belongs to the figure (ownership), not to the background. These parallel and independent findings represent only one of the many interesting examples of identical ideas, developed in different sciences, pushed by different needs but eliciting similar discoveries.

Physical coloration and figural effects resembling to those illustrated by cartographers were independently discovered by Pinna (1987) with the *watercolor illusion*.

2. COLORATION AND FIGURAL EFFECTS IN THE WATERCOLOR ILLUSION

The watercolor illusion is a long-range color spreading (coloration effect) imparting a figure-ground segregation (figural effect) across large regions (Pinna, 1987; Pinna, Brelstaff and Spillmann, 2001; Pinna, Werner and Spillmann, 2003; Spillmann, Pinna and Werner, 2004).

In Figure 5, purple wiggly contours flanked by orange edges are perceived as two asymmetrical peninsulas (see also Pinna and Grossberg, submitted), one half of each pointed and the other curved, leftward oriented (i.e. going from right to left) and slightly tinted by an orange veil of surface color (*Erscheinungweise*, Katz, 1911, 1930).

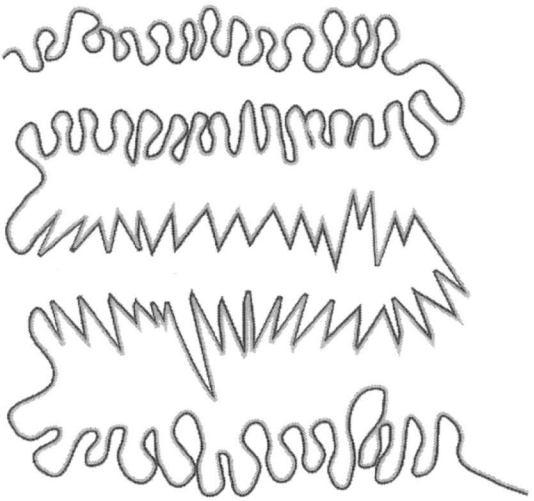


Figure 5. The watercolor illusion: purple wiggly contours flanked by orange edges are perceived as two asymmetrical peninsulas slightly tinted by an orange veil of surface color.

The peninsulas manifest a strong figural appearance and a univocal (very poorly reversible) figure-ground segregation that give to them the perceptual property of a rounded surface segregated in depth and a volumetric effect extending out from the flat surface. On the other hand, the complementary regions appear as holes or empty space around them.

As the purple and orange lines of Figure 5 are reversed, the main directions of coloration and figural effects are also reversed (see Figure 6).

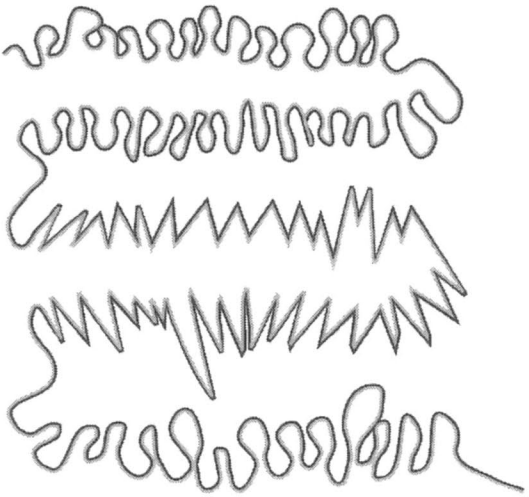


Figure 6. By reversing the purple and orange lines the main directions of coloration and figural effects are also reversed.

Although the two figures are geometrically the same, they appear as two groups of different symmetrical peninsulas hardly recognized as derived by the same figure with reversed colored lines: in Figure 5 two peninsulas appear asymmetrical and leftward oriented, while in Figure 6 two peninsulas, different from those in Figure 5, appear symmetrical and rightward oriented (i.e. going from left to right). What appears as a "bulging object" in Figure 5 appears as an "empty space" in Figure 6 and vice versa. The two opposite percepts show clearly Rubin's notion of unilateral belongingness of the boundaries (Rubin, 1915). This notion was also termed "border ownership" (Nakayama and Shimojo, 1990). In Figures 5 and 6 the border ownership is reversed: the boundaries belong unilaterally to only one (asymmetrical or symmetrical) region and not to the other. These perceptual results demonstrate that the watercolor illusion plays respectively against or in favor of the Gestalt principle of similarity of the edge shape (pointed vs.

curved) and symmetry (see also Morinaga's principle, 1942, *Ebenbreite*). According to these principles, all else being equal, more similar elements in the edge shape tends to be grouped together, and all else being equal, symmetrical contours are grouped together. For a demonstration of the effectiveness of these principles alone without the watercolor illusion, see Figure 7.

Figure 7. Similarity and symmetry principles induce two peninsulas from left to right.

The perceptual result tends to be more similar to the one of Figure 6 than the one of Figure 5, even if the figural organization is not so strong: similar and symmetrical elements are mostly grouped together but some reversibility can be perceived showing dissimilar and asymmetrical peninsulas segregated as figures as shown through the result of Figure 5.

It is not clear if early cartographers were aware of the coloration and figural effects of using two adjacent lines differently colored as illustrated in the watercolor illusion. As far as we know there are no references describing the two effects individually or jointly depending on the outline-color technique and in the same terms of the watercolor illusion. However, by comparing the two maps in Figure 3 and 4 with the watercolor illusion in Figures 5 and 6, the geometrical similarity (two juxtaposed differently colored lines) between the perceptual effect due to the cartographic technique and the watercolor illusion is self-evident. Nevertheless, all the known cases of old maps (see Figures 3 and 4) did not show a coloration and a figural effect as strong as the watercolor illusion: this was because: (i)

cartographers used to add a thick colored line to the inside of a black boundary, thus weakening the coloration effect of the watercolor illusion that is stronger when the two lines have the same width (Spillmann, Pinna and Werner, 2004 in press); (ii) the external boundary was always achromatic, while the watercolor illusion prefers chromatic edges; (iii) cartographers inserted place names within the watercolored region further weakening both coloration and figural effect.

In this work our purpose is not to distinguish the perceptual results obtained by the cartographer technique and the watercolor illusion but to demonstrate how strong and unique is the coloration and figural effects produced when two different color lines run parallel (juxtaposed), winning against previous knowledge and past experience and creating irregular geographic-like edges as those illustrated in Figure 8.

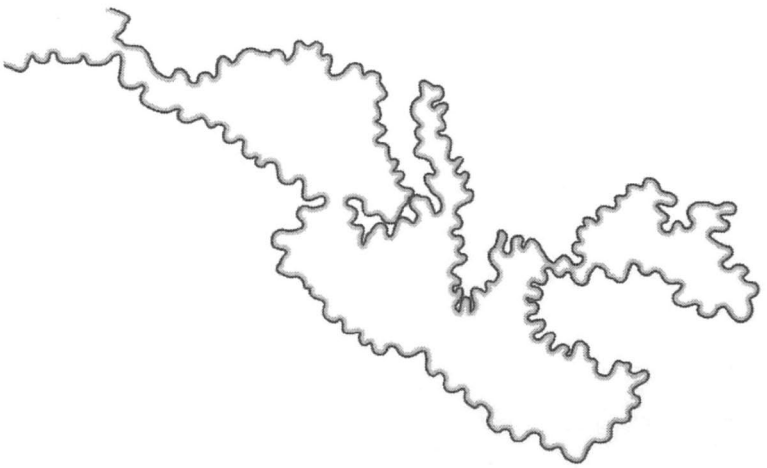


Figure 8. A watercolored unknown peninsula.

The percept of a land similar to an unknown peninsula emerges immediately. It takes some time to discover and recognize the hidden Mediterranean Sea. On the contrary, it immediately emerges when the orange fringes are added to the outside edge of the purple line (see Figure 9). These two Figures are identical in the outline, but the figure-ground organization is reversed by the watercolor illusion. The hue of the orange contour uniformly spreads to fill-in complementary surfaces and thereby defines the figural appearance of land masses and the sea in reciprocal regions.

The watercolor illusion seems to play against or in favor of the past experience principle of grouping and figure-ground organization (see also Pinna, Werner and Spillmann, 2003; Spillmann, Pinna and Werner, 2004 in press). This principle is well known within the Gestalt literature even if it has been considered as a principle on its own, i.e. not concerning geometrical properties of the stimulus configuration itself. It states that elements previously grouped in preceding viewing or regions experienced or familiar as figure tend to appear grouped or segregated similarly in spite of other principles playing against it.

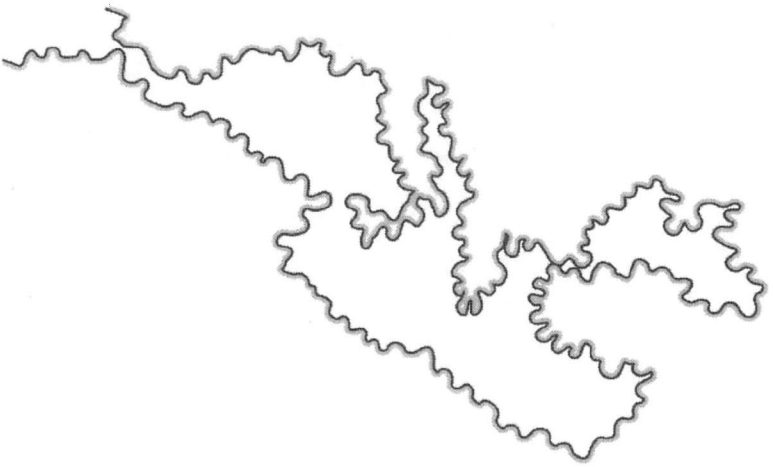


Figure 9. The Mediterranean Sea.

Pinna, Werner and Spillmann (2003), Spillmann, Pinna and Werner (2004 in press) and Pinna (in press) demonstrated that the watercolor illusion can be considered as a new principle of figure-ground segregation by studying it pitted against and combined with (i) the classical Gestalt grouping principles of proximity, good continuation, closure, convexity, symmetry, similarity, *prägnanz* and past experience (Wertheimer, 1923); and against (ii) the figure-ground segregation principles of surroundedness, size, orientation, symmetry, convexity, and parallelism (Rubin, 1915, 1921). In each of these cases the resulting effect due to the watercolor illusion on grouping and figure-ground organization is stronger than for each of the gestalt principles alone.

The aim of this work is to extend previous results on past experience to well known geographical regions presented under different conditions. In

two experiments the Mediterranean Sea, as illustrated in Figures 8 and 9, was studied statically in different tilted static positions (Experiment 1) and dynamically by rotating them (Experiment 2).

3. EXPERIMENT 1: STATIC WATERCOLORED MEDITERRANEAN SEA IN DIFFERENT ORIENTATIONS

Kanizsa (1969, 1980) showed a shaded map of Europe turned upside down so that the sea appears as a bas-relief while the land appears hollow. Only less than 15% of subjects recognized Europe. This result, weakening the role played by past experience on visual perception, was interpreted in terms of relief vs. hollow reversion (Lauenstein, 1938; Fieandt von, 1938; Metzger, 1975) and unfamiliarity (Rock and DiVita, 1987) when a figure is turned upside down. Figure 8 is related to Kanizsa's experiment. However it includes additional elements: (i) the orientation is not turned upside down; (iii) the perceptual effect induced by the watercolor illusion does not appear like a relief vs. hollow object but like a bulging figure vs. empty space; (ii) by turning the figure upside down the bulging figure vs. empty space effect does not change as the relief vs. hollow effect does.

The aim of the experiment was to measure the figural effect of the watercolor illusion in terms of percentage of recognition of the Mediterranean Sea under conditions like those illustrated in Figure 8 and 9, i.e. orange fringe respectively in the inside or outside edge of the purple line, when the stimulus is statically shown in different orientations. This was done to compare the figural effect of the watercolor illusion with past experience (familiarity) and with unusual orientations that change the retinal shape and the expressive qualities of the object (Kanizsa, 1969, 1980). The questions were: Does the perceptual differences between the two maps in Figures 8 and 9 in terms of perceptual recognition and figure-ground segregation change by changing the stimulus orientation? How does the perceptual results of Figure 8 and 9 change by changing the stimulus orientation despite what we know about geographical maps and *ipso facto* despite past experience, unusual orientation and changes in the retinal image? Does the watercolor illusion weaken or strengthen the effect due to past experience, to the retinal shape variations and to the changes in the expressive qualities of the object under different orientations of the stimulus pattern?

3.1 Subjects

Separate groups of 10 naive subjects participated to each experimental condition where only one stimulus was presented on a specific orientation. All observers had normal or corrected-to-normal vision.

3.2 Stimuli

The stimuli were obtained by 8 different rotation angles (from 0 to 360 deg with increments of 45 deg) of the geographical map of the Mediterranean Sea for the two conditions illustrated in Figures 7 and 8, where orange fringe is in the inside or outside edge of the purple line, i.e. respectively, inside the Mediterranean Sea or inside the land, and for a condition where only the purple line (purple-boundary-only condition) was used without the orange fringe, as illustrated in Figure 10.

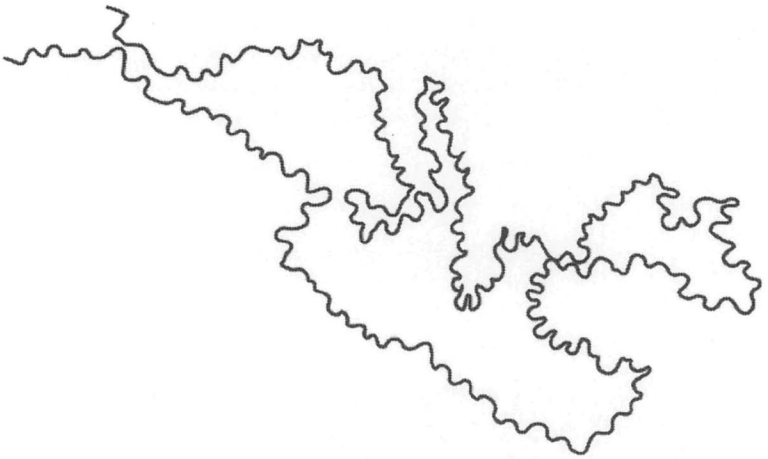


Figure 10. The Mediterranean Sea not watercolored.

The overall size of the stimuli was 18.8 x 10.6 deg of visual angle. The luminance of the white background under our test conditions was 80.1 cd/m^2. The CIE x,y chromaticity coordinates of the chromatic components of the patterns were: (orange) 0.57, 0.42; (purple) 0.27, 0.302.

The stroke width of the lines was approx 6 arcmin. Stimuli were presented on a computer screen under Osram daylight fluorescent light (250

lux, 5600° K) and were observed binocularly from a distance of 50 cm with freely moving eyes.

3.3 Procedure

The subject's task was to describe and report what they perceived by naming it (e.g. a peninsula, the Mediterranean Sea, etc.). In addition, subjects rated the relative strength (in percent) of a given percept being perceived as a figure, when reversible results were reported.

There was a training period preceding the experiment. During practice, subjects viewed some known figures from the Gestalt literature (*e.g.*, face-vase, etc.) to familiarize them with concepts of figure and ground. They practiced scaling the relative strength or salience of each reversible percept using percentages. Observation time was unlimited. Only at the end of the experiment all subjects were tested to recognize the South Europe, the Mediterranean Sea and North African map in a condition where only the purple line outlined these regions and under normal experienced rotation of them (0 deg). All the observers answered correctly and quite promptly.

3.4 Results

In Figure 11, the number of subjects perceiving and naming the Mediterranean Sea (Figure 11left) and the mean rating (in percent) of the Mediterranean Sea perceived as a figure (Figure 11right) are plotted as a function of the eight rotation angles for the three conditions: orange fringe in the inside edge of the Mediterranean Sea, inside South Europe and North Africa lands and purple-boundary-only map.

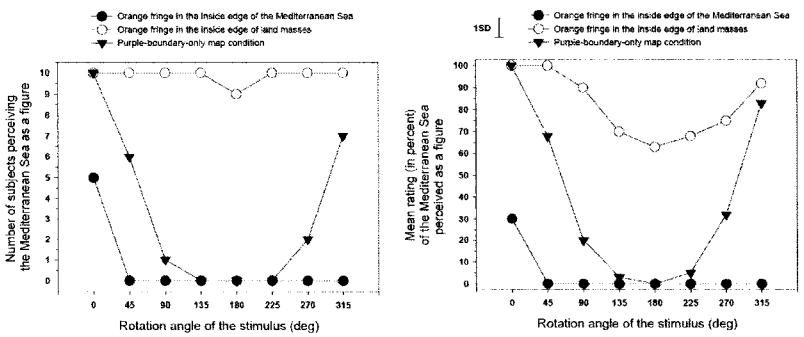


Figure 11. The number of subjects perceiving and naming the Mediterranean Sea (left) and the mean rating (in percent) of the Mediterranean Sea perceived as a figure (right) plotted as a function of the eight rotation angles for three conditions.

All subjects did not recognize the Mediterranean Sea when the orange fringe was within its edges (Figure 8). Only in the experienced rotation angle (0 deg) and after explicit suggestions by the experimenter, 5 out of 10 subjects perceived the Mediterranean Sea (see Figure 11left), even if this percept was not compelling (30% as illustrated in Figure 11right).

The opposite perceptual result was obtained when the orange fringe was within South Europe and North Africa land masses: all the subjects (except one when the map was 180 deg rotated) quite promptly recognize the Mediterranean Sea and the lands all around it. However, as the rotation angle of the map changes from the 0 deg experienced orientation, the strength of the known geographical map perceived as a figure decreases and reaches its minimum (a rating of about 60%, see Figure 11right) when the map is upside down (180 deg rotation) then increases again. Nevertheless, the figural effect induced by the watercolor illusion combined with past experience is much stronger than the figural effect induced by the past experience principle working alone in the purple-boundary-only condition, where the number of subjects perceiving the Mediterranean Sea decreased abruptly reaching a minimum when the map was almost turned upside down (from 135 to 225 deg) and then increased again. The same trend described the strength of the known geographical map perceived as a figure. It decreases quickly and reaches its minimum in the rotation interval from 135 to 225 deg (see Figure 11right). This implies that the known geographical Mediterranean Sea was hardly recognized by the subjects immediately after that the ±45 deg rotation angle was exceeded.

A two-way ANOVA revealed that the mean rating (in percent) of the Mediterranean Sea perceived as a figure varied significantly by varying the rotation angle of the stimulus ($F_{7,216}$=798.721, $p<0.001$) as well as through the three conditions ($F_{2,216}$=999.834, $p<0.001$). The interaction between the two factors was also significant ($F_{14,216}$=867.572, $p<0.001$).

The result concerning the purple-boundary-only condition is in agreement with unfamiliarity hypothesis (Kanizsa, 1969, 1980; Rock and DiVita, 1987) when a figure is turned upside down, because this unusual rotation angle changes the retinal shape and the expressive qualities of the object. However, the same hypothesis is very weak when the figural effect induced by the watercolor illusion is in favor of past experience. This result demonstrates that under watercolored conditions a boundary belongs strongly only to the side where the orange fringe flows out of its confines inducing poorly reversible figure-ground segregation.

4. EXPERIMENT 2: WATERCOLORED MEDITERRANEAN SEA DYNAMICALLY ROTATED

The results of the previous experiment consider each stimulus as a condition in itself independent from the others and phenomenally independent from the perception of a rotation. The rotation is not perceived at all. To perceive it, it is, first, necessary to recognize the map as a rotated map and, second, mentally rotate or compare the rotated map so as to match it with the one not rotated (0 deg). These operations require more than a mere primary perceptual process. The aim of the following experiment is to show directly the map and its rotation dynamically and not as a single photogram. This gives to the entire rotation sequence a whole object attribute that favor the recognition process and that may change the results of the previous experiment. The questions to answer were: By making the recognition process easier and stronger through the dynamic rotation of the stimulus, is the strong figural effect induced by the watercolor illusion weakened or annulled?

4.1 Subjects

Three new and separate groups of 10 naive subjects participated to each experimental condition. All observers had normal or corrected-to-normal vision.

4.2 Stimuli

Three dynamic stimuli were obtained by rotating each of the three conditions illustrated in Figures 8, 9 and 10. The rotation speed was 40 fps with a total of 1200 frames for a complete rotation. The rotation (clockwise or anticlockwise) was varied randomly for each subject. Before reporting their answer, subjects perceived two complete rotations of the stimulus. To make the recognition process and the role of past experience stronger, each rotation started from the experienced rotation angle at 0 deg. Illumination and observation distance were the same as in the first experiment. The procedure was the same as in the previous experiment except that subjects reported their evaluation of the relative strength (in percent) of a given percept being perceived as a figure as they hear a sound when one of the eight rotation angles with increments of 45 deg was reached. The subjects repeated their rating 10 times.

4.3 Results

In Figure 12, the mean rating (in percent) of the Mediterranean Sea perceived as a figure is plotted as a function of the eight rotation angles for the three conditions: orange fringe in the inside edge of the Mediterranean Sea, inside the South Europe and North Africa lands and purple-boundary-only map. Differently from the first experiment the measure of the number of subjects is not needed.

Like in the first experiment no subject recognizes the Mediterranean Sea when the orange fringe was within its edges (see Figure 8). However, after explicit suggestions by the experimenter all the subjects perceived the Mediterranean Sea within the range between the experienced rotation angle (0 deg), that obtained the maximum rating value (about 45%), and between ±90 deg, where little by little the Mediterranean Sea became weaker and weaker (a rating of about 5%), then it almost disappears letting the complementary perceptual result, the unknown peninsula, be perceived at a rating of almost 100%.

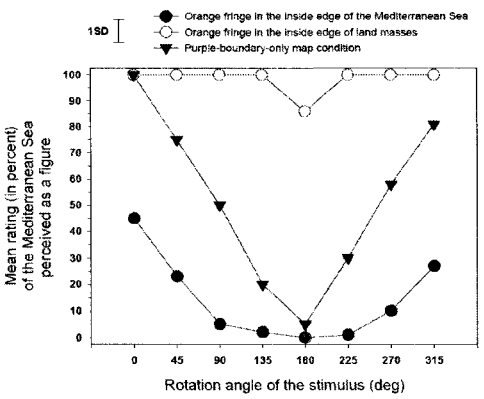


Figure 12. The mean rating (in percent) of the Mediterranean Sea perceived as a figure plotted as a function of the eight rotation angles for three conditions.

When the orange fringe was within South Europe and North Africa land masses (Figure 9), all subjects quite promptly recognize the Mediterranean Sea and the lands all around at a rating of 100% for all rotations but at a rating of about 85% for the180 deg rotation, i.e. when the map was upside down.

In the purple-boundary-only condition (Figure 10), the strength of the Mediterranean Sea perceived as a figure decreases similarly but less abruptly

than in the first experiment from 0 deg (a rating of 100%) to about 180 deg when it reached the minimum rating value.

A two-way ANOVA showed an overall effect of the rotation angle of the stimulus ($F_{7,216}$=983.442, $p<0.001$) as well as of the three conditions ($F_{2,216}$=886.475, $p<0.001$). The interaction between the two factors was also significant ($F_{14,216}$=957.776, $p<0.001$).

Summing up, the results showed that even by increasing the strength of the whole object property with the dynamic rotation of the stimuli, the figural induction due to the watercolor illusion is not annulled but slightly weakened: the recognition of the map is improved but not enough to annul the figural effect due to the watercolor illusion. Under these conditions the watercolor illusion determines figure-ground segregation more strongly than past experience and dynamic rotation. It is worth remembering that the starting rotation was chosen at 0 deg. When the starting orientation was chosen randomly (data of this experiment are not reported here), the results become more similar to those of the first experiment. A final important result is that when the retinal views are changed by showing the stimulus at different rotation angles, subjects' recognition rates drop abruptly. This is in agreement with previous results obtained by Rock and DiVita (1987) and Rock, DiVita and Barbeito (1981). However, our results show poor recognition even when the stimulus is dynamically rotated preserving his shape constancy and uniqueness. The map appears to change and to reverse its figure-ground organization through the rotation, even if it does not change its shape. This is an important point that deserves further experimental attentions, and that was first investigated using different stimuli in unpublished experiments by Pinna and Rock.

5. CONCLUSIONS

The FACADE model (Grossberg, 1994, 1997) is the best candidate to explain the coloration and figural effects of the watercolor illusion by assuming parallel boundary grouping and surface filling-in processes respectively processed by the Boundary Contour System (BCS) and Feature Contour System (FCS) (Cohen and Grossberg, 1984; Grossberg and Mingolla, 1985a, 1985b; Grossberg and Todorovic, 1988). The two Systems obey to complementary rules (Grossberg, 2000) that can explain the watercolor illusion through a process of spatial competition between stronger inputs occurring at the edges of higher contrast colored lines than at lower contrast ones. (Grossberg, 1987, 1994, 2000; Grossberg and Mingolla, 1985a; Grossberg and Swaminathan, 2004; Grossberg and Yazdanbakhsh, 2004). The boundaries of the lower contrast edges are thus weakened more

by competition than by the boundaries of the highest contrast edges. Hence more color can spread across these boundaries than conversely. Briefly, the coloration effect can be summarized in two steps: First, the contour is weakened by lateral inhibition between differentially activated edge cells (local diffusion); and, second, the color onto the enclosed area flows out (color diffusion). Edge polarity neurons in areas V2 and V4 of the monkeys responding to a luminance step in only one, but not the other, direction may be responsible for border ownership and, as a consequence, of the figural effect.

ACKNOWLEDGMENTS

This research was supported by Fondazione Banco di Sardegna, the Alexander von Humboldt Foundation, ERSU, Banca di Sassari and Fondo d'Ateneo (ex 60%).

We thank Massimo Dasara for assistance in preparing the stimuli and testing the subjects.

REFERENCES

Bagrow, L., and Skelton, R. A., 1985, *History of Cartography,* Precedent Publ, Chicago

Cohen, M., and Grossberg, S., 1984, Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance, *Perception & Psychophysics* **36**:428-456.

Donkersloot de, M. V., 1973, *Willem Jansz. Blaeu: A Biography and History of His Work as a Cartographer and Publisher.*

Fieandt von, K., 1938, *Über das Sehen von Tiefengebilden bei Wechselnder Beleuchtungsrichtung*, Psychologisches Institut Universität Helsinki, Helsinki.

Grossberg, S., 1987, Cortical dynamics of three-dimensional form, color, and brightness perception II: Binocular theory, *Perception & Psychophysics*, **41**:117-158.

Grossberg, S., 1994, 3-D vision and figure-ground separation by visual cortex, *Perception & Psychophysics* **55**:48-120.

Grossberg, S., 1997, Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures, *Psychological Review* **104**:618-658.

Grossberg, S., 2000, The complementary brain: Unifying brain dynamics and modularity, *Trends in Cognitive Sciences* **4**:233-245.

Grossberg, S., and Mingolla, E., 1985a, Neural dynamics of form perception. Boundary completion, illusory figures and neon color spreading, *Psychological Review* **92**:173-211.

Grossberg, S., and Mingolla, E., 1985b, Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations, *Perception & Psychophysics* **38**:141–171.

Grossberg, S., and Swaminathan, G., 2004, A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: Development, attention, and bistability, *Vision Research*, (in press).

Grossberg, S., and Todorovic, D., 1988, Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena, *Perception & Psychophysics* **43**:241-277.

Grossberg, S., and Yazdanbakhsh, A., 2004, Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading, (in press).

Kanizsa, G., 1969, Perception, past experience, and the "impossible experiment", *Acta Psychologica* **21**:66-96.

Kanizsa, G., 1980, *La grammatica del vedere*, Il Mulino, Bologna.

Katz, D., 1911, Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung, *Zeitschrift für Psychologie,* Ergänzungsband 7 (Leipzig: Barth), 6-31.

Katz, D., 1930, *Die Erscheinungsweisen der Farben*, 2nd edition [Translation into English: MacLeod R. .B, and Fox C.W., 1935 *The World of Color* (London: Kegan Paul)].

Lauenstein, L., 1938, Über räumliche Wirkungen von Licht und Schatten, *Psychologische Forschung* **19**:267-319.

Metzger, W., 1975, *Gesetze des Sehens*, Kramer, Frankfurt-am-Main.

Morinaga, S., 1942, Beobachtungen über Grundlagen und Wirkungen anschaulich gleichmäßiger Breite, *Archiv für die gesamte Psychologie* **110**:309-348.

Nakayama, K., and Shimojo, S., 1990, Towards a neural understanding of visual surface representation, *Cold Spring Harbor Symposia on Quantitative Biology* **40**:911-924.

Pinna, B., 1987, Un effetto di colorazione, in: *Il Laboratorio e la Città. XXI Congresso degli Psicologi Italiani*, V. Majer, M. Maeran and M. Santinello, eds., 158.

Pinna, B., in press, The role of the Gestalt principle of similarity in the watercolor illusion, *Spatial Vision.*

Pinna, B., Brelstaff, G., and Spillmann, L., 2001, Surface color from boundaries: A new 'watercolor' illusion, *Vision Research* **41**:2669-2676.

Pinna, B., Werner J.S., and Spillmann, L., 2003, The watercolor effect: A new principle of grouping and figure-ground organization, *Vision Research* **43**:43-52.

Pinna, B., and Grossberg, S., submitted, The watercolor illusion: new properties and neural mechanisms, *Journal of Vision.*

Rock, I., and DiVita, J., 1987, A case of viewer-centered object perception, *Cognitive Psychology* **19**(2):280-293.

Rock, I., DiVita, J., and Barbeito, R., 1981, The effect on form perception of change of orientation in the third dimension, *Journal of Experimental psychology: Human Perception & Performance* **7**(4):719-732.

Rubin, E., 1915, *Synsoplevede Figurer*, Kobenhavn: Glydendalske.

Rubin, E., 1921, *Visuell Wahrgenommene Figuren*, Kobenhavn: Gyldendalske Boghandel.

Schilder, G., 1993, *The State of Research into the Cartographical Work of Willem Jansz. Blaeu*, Monumenta Cartographica Neerlandica.

Spillmann, L., Pinna, B. and Werner, J. S., 2004, Form-from-watercolour in perception and old maps, in: *Seeing Spatial Form*, M. R. M. Jenkin and L. R. Harris, eds., Oxford University Press, (in press).

Wertheimer, M., 1923, Untersuchungen zur Lehre von der Gestalt II, *Psychologische Forschung* **4**:301–350.

Wollschläger, D., Rodriguez, A. M., and Hoffman, D. D., 2002, Flank transparency: The effects of gaps, line spacing, and apparent motion, *Perception* **31**:1073-1092.

4

EMERGENCE

AUTOPOIESIS AND EMERGENCE

Leonardo Bich
Dipartimento di Psicologia, Università di Pavia, Piazza Botta, 6 - 27100 Pavia, Italy,
e-mail: leobich@libero.it

Abstract: Autopoietic theory is more than a mere characterization of the living, as it can be applied to a wider class of systems and involves both organizational and epistemological aspects. In this paper we assert the necessity of considering the relation between autopoiesis and emergence, focusing on the crucial importance of the observer's activity and demonstrating that autopoietic systems can be considered intrinsically emergent processes. From the attempts to conceptualize emergence, especially Rosen's, autopoiesis stands out for its attention to the unitary character of systems and to emergent levels, both inseparable from the observer's operations. These aspects are the basis of Varela's approach to multiple level relationships, considered as descriptive complementarities.

Key words: autopoiesis; emergence; observer; descriptive complementarity; Robert Rosen; Francisco Varela.

1. INTRODUCTION

After about thirty years since the first formulation of autopoiesis, it is opportune to consider the possible developments of Maturana's and Varela's insights. It is necessary then not to stop at the statement that living systems are autopoietic, but to go further, taking into consideration the whole framework of the theory.

The first fundamental aspect to be considered is the role of organization, defined as the set of invariant relations that determine the identity of a system, apart from its effective material realization, specifically from the attributes of its concrete components. With respect to this, autopoietic systems together with nervous, immunitary and social ones, also belong to the wider class of autonomous systems.

The problem of the identity and the unitary character of a system, as well as that of organizational levels, is tightly connected to the observer's activity. His role is the crucial aspect of the theory, that is characterized by the rejection of an objective reality independent from observative experience, and by the importance given to relations between different points of view on the system described.

Because of the inseparability of organizational and epistemological aspects in autopoietic theory, we can revise it in relation to emergence, a fundamental problem for Systemics. Few years after the loss of Robert Rosen and Francisco Varela, two of the main thinkers in the study of the nature of organisms and emergent processes, it is necessary to give new impulse to their researches. It is therefore opportune not to consider autopoietic theory as complete and concluded, confining ourselves to its identifications with the living, but rather to develop it further, starting from its analysis as a suitable point of view on emergent phenomena.

2. AUTOPOIETIC THEORY

In the first place autopoiesis is a definition of the living. Belonging to systemic tradition, it focuses attention on the characterization of organisms as autonomous unities, different from simple aggregates of inorganic components. It identifies the mechanism common to all organisms, that generates the identity of the living, independently from its different material realizations. In order to do this, it is necessary to distinguish between organization and structure: the former is the set of relations that defines the identity of the system; the latter its effective realization. Organization is a subset of all actual relations between components and for this reason it doesn't specify their concrete nature, but only the kind of relations they must satisfy. So it can be implemented by different structures.

Autopoietic organization is defined as a unity by a net of processes of production, transformation and destruction of components, that:

1. through their interactions and transformations recursively realize and regenerate the same complex that produces them;
2. constitute the system as a concrete unity in the space in which they exist, by establishing its boundary and thus specifying its topological domain (Varela et al., 1974; Maturana and Varela, 1980).

To analyze autopoiesis it is necessary to consider the observer's activity both by recognizing his role in the process of distinction of a unity, and by considering him as a one of these entities. At the same time this dual nature belongs to the observed system.

The two basic concepts with respect to this, are *structural determinism* and *distinction.* The former is tightly connected to the tradition of biological structuralism, that takes into consideration the active nature of organisms in their relation with the environment (von Bertalanffy, 1949). It focuses attention on the internal structure of the system instead of on the role of its medium. The system interacts and changes in time in a way totally determined by its structure, that specifies the set of all possible changes and effective perturbations. And these therefore don't define, but only trigger structural changes. As a consequence, these systems cannot be studied through input-output relations, because a same stimulus can cause different alterations. Closely connected to structural determinism, the concept of *structural coupling* describes interactions between two systems of this kind as a series of mutual selective perturbations, that can lead to the arousal of reciprocal behavioral coherences: there is no transmission of information.

The importance of the role of the observer results from him being a structural determined system too. He doesn't record an objective reality but only describes those structural changes that perturbations trigger in him. For this reason, concepts like representation and reality become meaningless and there isn't anymore an objective and independent basis for knowledge. The starting point is the observer's experience as a constitutive condition and a primary unavoidable fact (Maturana, 1988).

The primitive observational activity is distinction, the procedure that takes a unity apart from its background, specifying its conditions and domain of existence. In addition, according to points of view, attention can be focused on either the inner structure of the system or its interaction with the environment. In autopoietic theory this is the basis of the difference between a simple and a composite unity. The former is distinguished from the medium it interacts with, and it is characterized by a set of *given* properties. The latter instead is the consequence of a distinction where the point of view is internal to the system: it is made of components, that are recognized in respect to the entity they integrate. Organization then, is the set of relations that integrate components into a composite unity of some kind, defining its identity and properties when it is distinguished as a whole. As we can see, this identity, is not an objective property of reality, but depends on the operation of distinction performed by the observer.

Autonomous systems, whose paradigm is autopoiesis, are a special kind of structural determined systems, characterized as composite unities by *organizational closure*. They act not only upon their internal variables, as homeostatic machines do, but upon their very organization: it is the product of their operations and their behavior is secondary to its maintenance. They self-distinguish from a background, as unities, through their very operations of production of components and of a physical boundary, defining at the

same time their organization: activity and identity coincide. The system is then a unitary whole, it is producer and product of itself. It is an *all-or-nothing* unity, not describable as an aggregate through a succession of intermediary steps, but only as already realized. This nature is expressed by a circular relation between the dynamics of components and the membrane: they are inseparable, because the former realize the latter, which makes them possible.

The identification of these systems by the observer depends on an operation of distinction that defines their identity in the same domain where they specify it through their operations. Otherwise on the same structure another kind of system is distinguished, whose identity depends on the recognition of a different organization from an autopoietic one.

3. EMERGENCE

At an intuitive level, the concept of emergence can be considered synonymous with qualitative novelty in nature, unpredictable or undeducible on the basis of pre-existing features, both concerning the different behavior of components in isolation or in the whole, and the properties of the system as a unity.

Following Crutchfield (1994), we can distinguish three kinds of definitions of emergence. The intuitive concept belongs to the first, that doesn't investigate the nature of novelty and the system that exhibits it. And most of all it doesn't take into consideration the role of the observer, as in British Emergentism (Alexander, 1920; Morgan, 1923; Broad, 1926). The second, called "pattern formation" and characterized by self-organizing processes, had been identified with emergence until the beginning of the eighties. A pattern is considered emergent when it is an unpredictable consequence of non-linear dynamics: different possibilities emerge, without being able to predict which one will be realized. This kind of emergence depends on the observer's predictive limitations, due to inaccuracy in the knowledge of initial conditions. The many possibilities are only apparent, as the result is totally determined and deducible by laws that describe system dynamics, without the need of adopting a new model.

The third kind of emergence, called "intrinsic", concerns a change in the structure of the observed system, that leads to a new behavior, undeducible on the basis of the adopted model, so that a new description is necessary: the paradigmatic example of this kind of emergence, also formally, is *Spontaneous Symmetry Breaking* (Anderson and Stein, 1985; Pessa, 1998). In this context the role of the observer doesn't concern his predictive limitations, but the inability of any of his models to capture, even in

principle, all features of the world. Attention is shifted from objective properties of reality to relations between descriptions, pointing out the partiality of a single point of view on a same system so that manifold models are needed. The distinction between epistemological and ontological emergence, as the one expressed by Silberstein and McGeever (1999), loses its meaning because both unavoidably depend on the observer's activity, although in different ways.

The difference between the two last kinds of emergence is expressed in its epistemological nature by Rosen (1978), through the concepts of "error" and "emergence relative to a model". He describes modelling as establishing a relation between two entailment structures: the causal one of natural system and the formal one of the model that encodes it. The existence of emergent phenomena depends on the limitations of the encoding process, the measurement of certain aspects of the natural system through its interaction with a measuring device, or meter. Every meter m defines and observable f that associates every state s, distinguished in a system S, with a real number. However every meter, and so every observable, is partial because it can distinguish only certain states of the system. For this reason, what a model can describe depends on the choice of meter.

The concepts of error and emergence follow from the incompleteness of observables in relation to the effective behavior of the system. The former has its origin in the limits of the resolution power of meters, which cannot discriminate between values beyond a certain threshold of accuracy. Consequently, some changes cannot be caught and after some temporal gap the behavior of the system becomes unpredictable.

Emergence instead doesn't concern the accuracy of observables, but the relationships between sets of them, that characterize different models. It is defined as the bifurcation between the effective behavior of a system and that predicted by a model, due to the partiality of a single set of observables in describing the evolution of the system, even in an ideal situation of perfect knowledge of initial conditions. As a result, the model defined by that set has to be substituted, because there are new features invisible before, like new relations between observables, that need a new description, not consistent with the previous one. Instead in Newtonian reductionistic conception, different models can always be reduced to a fundamental one, the motion of basic components.

In Rosen (1978, 1991, 2000) the concept of emergence is strictly connected to complexity. This is opposed to the simplicity of aggregative systems, which are describable on the basis of the behavior of their isolate parts, and it is recognized through the appearance of new properties, not deducible from the adopted model. It can so be defined as the inability of any formalism alone to capture all the properties of a system, and the

necessity of making use of alternative descriptions, neither reducible to each other nor to a more comprehensive one.

4. AUTOPOIESIS AS AN INTRINSICALLY EMERGENT PROCESS

Autopoietic theory shows many emergentist aspects, but unlike some approaches only focused on their relational character (Ruiz-Mirazo and Moreno, 1998), it is necessary to consider them with regard to their inseparable connection with observative activity, consistently with the structure of the theory and with the concept of emergence analyzed above.

The organizational and processual identity of these systems, enables us to consider the origin of levels and their relations, instead of a mere covariation of their properties. Also, the unitary, distributed and *all-or-nothing* characters of autopoietic systems, leave out the possibility of applying a reductionistic approach that identifies the unity with a sum of parts or functions: the whole depends on the relations between its components, but it is not reducible to the properties of them in isolation. Components in fact depend on the very network that specifies them. For this reason they don't coincide with material parts, as they don't exist apart from their relations.

The difference between a component and an isolated part depends on the observer. They are in fact epistemological notions, the result of two different operations of distinction. The identification of material parts is independent of relations they can take part in, and consequently it is impossible to restore the system from them. A component instead, is distinguished from a background constituted by the unity it integrates through its interactions and cannot be considered in isolation. In a similar way, according to Rosen (1991) in complex systems operations of analysis and synthesis don't correspond.

The notion of level has an epistemological nature too. For instance an autopoietic unity can be considered as a component of a wider system, not on the basis of its identity as such but for its role in the whole that it is a part of. Its nature depends on the observer's purposes and activities: there isn't a fundamental or preferential level.

The main emergentist aspect in autopoietic theory concerns levels of a system, as inseparable from operations of distinction. They depend in fact on the observer's point of view and on the distinctions he performs, which specify a unity and the domain in which it exists and interacts through its properties. Consequently, unities identified by different operations of distinction, differ in their domain of existence (Maturana and Varela, 1980). Referring to the difference between a simple and a composite unity, we can

perform two distinctions on a same system, so specifying two kinds of domain. As a composite unity, the system exists in the space specified by its components, because it is distinguished through their properties. Instead, considered as a simple unity, it exists in the domain of its properties as an unanalyzable whole, in interaction with its medium.

Also the difference between properties of whole and parts is therefore connected to different operations of the observer, according to whether the point of view is external or internal to the system. This exists in two domains: although its behavior as simple unity is determined by its organization as a composite one, i.e. by relations between the components that realize it, it belongs to another space, defined by another distinction. These domains, even though they depend on each other, cannot be described by the same kind of relations, because they are specified by different distinctions, which define different spaces and properties.

These aspects are typical of emergent phenomena: dynamics of components generate something different and irreducible to them. This is due to the fact that the distinction of an autopoietic system as a composite unity determines the emergence of a new phenomenological domain, as simple unity, which belongs to a different descriptive level. As a consequence of this view, two different domains cannot be reduced to each other, as they take into consideration distinct kinds of interactions: the whole cannot interact as a component at lower level (Maturana, 1980). As the two domains don't intersect, their relation can be considered only in a metadomain established by the observer, who considers their correlations likewise structural coupling: levels trigger changes on each other. Putting them on the same level would be a methodological mistake, as it wouldn't consider the crucial role of the observer, who distinguishes the system as a unity of one or the other kind. These are the reasons why downward causation is so puzzling.

Autopoiesis is therefore an example of intrinsic emergence, giving an explanation of it in terms of observative activity. From a synchronic point of view, the emerging of a level requires a new descriptive framework. This is the case of the difference between composite and simple unity, and their domains. The irreducibility of levels is the consequence of the difference between the observative operations, which define two distinct domains that cannot interact on the same plan. From a diachronic point of view, both considering the level of components and that of the unity as a whole, rules are not defined in advance, but are recursively generated by the system. In fact only the organization is kept invariant, unlike structural processes. Consequently with structural change, the possible states of the system vary and cannot be deduced on the basis of original rules any more. In addition autonomous systems specify interactions of their medium, that cannot be

fixed in a transfer function of stimuli into responses, but they go on changing because of structural modifications that occur in both participants and are determined by their structures. Rules of interaction are continually redefined through the interplay of reciprocal perturbations and compensations that generate an open-ended process, without converging on an optimal state. In both synchronic and diachronic cases, a single model is unable to describe the system.

Varela makes a further step, trying to express not only the irreducibility of emergent levels but also their reciprocal relations, considered as mutual constraints instead of mere covariations. At the same time he keeps the centrality of the role of the observer, avoiding the limits of a "realistic" analysis of downward causation. In every description of these systems, according to the assumed point of view, we see the opposition between the observed object and the process it is generated by. Varela considers these two observative levels of the same unity as complementary ways of description (Varela, 1979), in the sense that they are distinct and irreducible to each other, but at the same time they are in a relation of reciprocal definition. It is a particular dialectic, where there is neither synthesis nor contradiction, where terms are "not one, not two". In fact by enlarging the point of view to a metadomain, these two poles can be considered as intertwined, although belonging to two different descriptions. In such a wider view, each term emerges from the other in a reciprocal but asymmetrical relation that extends itself through levels. Also, we must consider that no observative pole can be conceived either as privileged or self-established, as a substance independent from its underlining processes, but terms are always co-dependent. A founding principle cannot be found, as it always appeals to another one and vice versa, giving rise to a vicious circle. This apparent paradox can be avoided only considering the wider metadomain where both terms are poles of a generative relation (Varela, 1979; Varela and Dupuy, 1992).

With respect to autopoietic systems, three levels must be considered: components, whole and environment. These give rise to a double dialectic, characterized on the one hand by the mechanism of realization of the identity of the system, and on the other by the production of significance from a indistinct background. Their common logic is that of paradox, identifiable in the self-referential and co-dependent circularity between unitary identity and processes of production, and between behavioral autonomy and coupling with the medium. These two dialects are to be considered as emergent processes (Varela, 1997), tightly connected to each other and characterized by the complementarity of two interacting levels. Also, this double emergence can be seen in a metadomain where the identity of the system is inseparable from dynamics that realize it and from the interaction with its

background, in respect to which it is distinguished as a unity. According to the observative level new complementary pairs are generated.

The first step of this approach consists in considering the emergence of the system's identity. The mechanism which generates a self-defined and irreducible unity is the circular, self-referential relationship between whole and parts, that are distinct, because on two different levels, but at the same time interdependent: there is in fact an endless co-definition between local rules of interaction (the flux of components) and global properties (the unitary and constant pattern). But the unity so generated is never completely defined. On the one hand it looks coherent, durable and steady, on the other it is only virtual, as it is the result of an endless dynamic. Thus it is not localized, or identifiable as a substance: it is a "selfless self" (Varela, 1992), never completely defined but nevertheless crucial for the level of the interaction of the whole as simple unity. The system is in fact the node of intersection between its realization and its interactions with the medium. Structural and interactional processes are both crucial.

It is therefore necessary to consider also the second emergence, where a dialectic of complementarity is instantiated between the production of a world of significance and the structural coupling that makes it possible. This is the basis of Varela's *enactive* approach (Varela et al., 1991), that consists in linking emergent process and structural coupling. The autonomous system, through its activity in the medium produces a domain of distinctions from a random and uniform background: it selects certain environmental regularities as meaningful.

The key concept in Varela's approach is the lack of foundations. They cannot be found in an objective reality but neither in the subject, that is never completely defined and substantial. It depends in fact on its underlining processes and its interactions with the environment: subject and object arise simultaneously, as each one cannot exist without the other. The meaning of the double dialectic of emergence that characterizes Varela's approach, consists then in the reciprocal relations between levels with an endless interplay of complementary points of view, without a privileged one.

5. CONCLUSIONS

The main difference between autopoietic and Rosen's approaches to emergence concern their epistemological frameworks. In both theories, the idea of a mere recording of an objective reality is rejected, and attention is focused on the role of the observer. In Rosen however, this aspect is connected with the problem of limitations of models and the impossibility of capturing all the characteristics of reality through a same description.

Maturana's and Varela's approach analyzes these aspects but goes deeper, as it points out the observer's activity in constituting a reality. Attention is focused in particular on the identity and the unitary character of the system, considered with regard to its background through the concept of distinction. This difference can also be seen comparing autopoiesis and Rosen's M/R-systems (Rosen, 1972; 1991; Nomura, 1997). The latter fail to grasp the unitary character of the living and for this reason they can express only the first part of Maturana's and Varela's definition. The prospects of an autopoietic approach to emergence are thus wider, as they take into consideration the origin and relations between levels and the identity of the system. But they lack a consistent formalism: the used ones are limited to pattern formation (Varela et al., 1974; Varela, 1979; Thompson and Varela, 2001).

A possible line of research can be to identify the correspondences between autopoietic emergence and relational holism of quantum mechanics, that asserts that the emergent properties of the system don't exhaustively supervene on components' intrinsic ones (Teller, 1986). The search for such a correspondence must extend to the inquiry about emergence to the insights coming from the analysis of autopoietic theory, by keeping the centrality of observational and processual aspects, beyond a mere comparison of properties of parallel levels.

REFERENCES

Alexander, S., 1920, *Space, Time and Deity*, Macmillan, London.

Andersen P. B., Emmeche, C., Finnemann, N. O., and Christiansen, P. V., eds., 2000, *Downward Causation*, Arthus University Press, Arthus.

Anderson, P. W., and Stein, D. L., 1985, Broken symmetry, emergent properties, dissipative structures, life: are they related?, in: *Self-Organizing Systems: The Emergence of Order*, E. F. Yates, ed., Plenum Press, New York, pp. 445-458.

Baas, N. A., and Emmeche, C., 1997, On emergence and explanation, *Intellectica* **25**(2):67-83.

Beckermann, A., Flohr, H., and Kim, J., eds., 1992, *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, de Gruyter, Berlin.

Birckhard, M. H., 1998, A process model of emergence and representation, in: *Emergence, Complexity, Hierarchy, Organization*, G. L. Farre and T. Oksala, eds., Finnish Academy of Technology, Espoo, pp. 263-270.

Broad, C. D., 1926, *The Mind and Its Place in Nature*, Routledge and Kegan Paul Ltd., London.

Castellani, E., 2002, Reductionism, emergence, and effective field theories, *Studies in History and Philosophy of Modern Physics* **33**:251-267.

Crutchfield, J. P., 1994, The calculi of emergence: computations, dynamics, and induction, *Physica D* **75**:11-54.

Kim, J., 1993, *Supervenience and Mind, Selected Philosophical Essays*, Cambridge University Press, Cambridge, MA.

Kim, J., 1997, Explanation, prediction and reduction in emergentism, *Intellectica* **25**(2):45-57.

Maturana, H. R., 1980, Autopoiesis: reproduction, heredity, and evolution, in: *Autopoiesis, Dissipative Structures, and Spontaneous Social Orders*, M. Zeleny, ed., Westwiew, Boulder, pp. 45-79.

Maturana, H. R., 1988, Reality: the search for objectivity or the quest for a compelling argument, *Irish J. Psych.* **9**(1):25-85.

Maturana, H. R., and Varela, F. J., 1980, *Autopoiesis and Cognition. The Realization of the Living*, Reidel Publishing, Dordrecht.

Maturana, H. R., and Varela, F. J., 1984, *El árbol del conocimiento*, Editorial Universitaria, Santiago del Chile.

Mikulecky, D. C., 2000, Robert Rosen: the well posed question and its answer-Why are organisms different from machines?, *Syst. Res.* **17**(5):419-432.

Minati, G., Penna, M. P., and Pessa, E., 1998, Thermodynamical and logical openness in general systems, *Syst. Res.*, **15**(2):131-145.

Morgan, C. L., 1923, *Emergent Evolution*, Williams and Norgate, London.

Nomura, T., 1997, An attempt for description of quasi-autopoietic systems using metabolism-repair systems, in: *Fourth European Conference on Artificial Life*, P. Husbands and I. Harvey, eds., MIT Press, Cambridge, MA, pp. 48-56.

Pessa, E., 1998, Emergence, self-organization, and quantum theory, in: *First Italian Conference on Systemics*, G. Minati, ed., Apogeo, Milano, pp. 59-80.

Rosen, R., 1972, Some relational cell models: the Metabolism-Repair Systems, in: *Foundations of Mathematical Biology*, R. Rosen, ed., Academic Press, New York, vol. II, pp. 217-253.

Rosen, R., 1978, *Fundamentals of Measurement and Representation of Natural Systems*, North-Holland, New York.

Rosen, R., 1991, *Life Itself: a Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*, Columbia University Press, New York.

Rosen, R., 2000, *Essays on Life Itself*, Columbia University Press, New York.

Ruiz-Mirazo, K., and Moreno, A., 1998, Autonomy and emergence: how systems become agents through the generation of functional constraints, *Acta Polytechnica Scandinavica* **91**:273-282.

Silberstein, M., and McGeever, J., 1999, The search for ontological emergence, *The Philosophical Quaterly* **49**(195):182-200.

Stephan, A., 2002, Emergentism, irreducibility, and downward causation, *Grazer Philosophische Studien* **65**:77-93.

Teller, P., 1986, Relational holism and quantum mechanics, *Brit. J. Phil. Sci.* **37**:71-81.

Thompson, E., and Varela, F. J., 2001, Radical embodiment: neural dynamics and consciousness, *TRENDS in Cognitive Sciences* **5**(10):418-425.

Varela, F. J., 1979, *Principles of Biological Autonomy*, North-Holland, New York.

Varela, F. J., 1992, Autopoiesis and a biology of intentionality, in: *Autopoiesis & Perception*, B. McMullin and N. Murphy, eds., Dublin City University, Dublin, pp. 4-14.

Varela, F. J., 1997, Patterns of life: intertwining identity and cognition, *Brain and Cognition* **34**:72-87.

Varela, F. J., and Dupuy, J.-P., 1992, Understanding origins: an introduction, in: *Understanding Origins*, F. J. Varela and J.-P. Dupuy, eds., Kluwer Academic Publisher, Dordrecht, pp. 1-26.

Varela, F. J., Maturana H. R., and Uribe, R. B., 1974, Autopoiesis: the organization of living systems, its characterization and a model, *Biosystems* **5**:187-196.

Varela, F. J., Thompson, E., and Rosch, E., 1991, *The Embodied Mind. Cognitive Science and Human Experience*, MIT Press, Cambridge, MA.

Von Bertalanffy, L., 1949, *Das Biologische Weltbild: Die Stellung des Lebens in Natur und Wissenshaft*, Francke, Bern.

TYPICAL EMERGENCIES IN ELECTRIC POWER SYSTEMS

Umberto Di Caprio
Stability Analysis s.r.l., Via A. Doria 48/A, 20124 Milano, Italy

Abstract: Typical emergencies, capable to produce extended black-outs in large interconnected power systems, are set into evidence. The most appropriate control actions are thoroughly discussed. A classification is proposed, in which one considers medium, heavy, extremely severe disturbances. The role of stability is pointed out, automatic load-shedding plans that expressly account for stability are suggested.

Key words: black-out; load-shedding; emergency control; multimachine; stability; power system.

1. INTRODUCTION

1.1 Generalities

Recent serious black-outs in Europe and in USA awakened interest about theoretical mechanisms explaining raise and evolution of emergencies in multimachine electric power systems. The following ingredients look determinant: 1) Nonlinearities; 2) Multiplicity of Equilibria; 3) Stability; 4) Complexity.

The nature of the first three can be explained and understood in a relatively simple way, even with reference to the elementary case of a single machine connected to an "infinite bus" (namely to an ideal network that keeps frequency and voltage). The fourth element ("complexity") comes into evidence in large interconnected systems, like, e.g. the western european power system. Non linearities appear to be the principal reason for raise and spreading of black-outs. Essentially: as the electric power generated by a single machine is a nonlinear function of the machine "electric rotor angle",

it varies non-linearly with such angle. So in transient condition generated power oscillates with time, whilst input mechanic power keeps constant. The unbalance between electric and mechanic powers determines an electric acceleration that can be oscillatory or, even, exponentially unstable. Electric acceleration is defined as the second derivative of the electric rotor angle with respect to time. In parallel, the first derivative gives the electric speed. The latter is proportional to the mechanic rotation speed of the machine. The non linearity binding together power and rotor angle is sinusoidal and, for this reason, two distinct equilibrium points (in a two-machine system) exist, one of which is stable whilst the other is unstable. Not only this: in emergency conditions it can well happen that no equilibrium point exists at all. This looks evident, e.g., from the analysis of the black-out of September 2003 in the italian system.

Furthermore, even when at least one stable equilibrium exists, the "region of stability in the large" is bounded. Consequently severe disturbances (i.e. power unbalances) lead the system state out of this region and then result in instability. This sets severe limits on the maximum allowable steady-state flows of power among subsystems ("areas") of a large system.

"Complexity" introduces further problems, both practical and conceptual. Above all the existence of multiple unstable equilibria, only one of which determines the size of the effective stability region. In addition one has a multiplicity of the basic "oscillation modes " (e.g. in a 5-machine system one finds 4 oscillation modes), and "nonconservation" of the system Energy.

Recent theoretical developments (Di Caprio, 2002) help in study of non-linear stability, by the use of a convenient Energy function that represents the "conservation part of the energy".

1.2 Structure of an electric power system

We make reference to a classic system for production, transmission and distribution of electric energy. The system contains a certain number of *generators*, a certain number of *transmission lines*, a certain number of *electric loads*. As electric generators are but synchronous machines, it is standard convention to use the name *n-machine electric power system*.

We emphasize. that complexity begins with $n = 3$. Namely a three-machine system must be considered a complex system. Fig. 1 illustrates a one-machine infinite bus, fig. 2 a two-machine system, fig. 3 a three machine system.

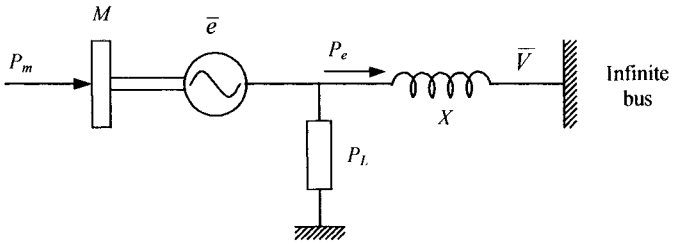


Figure 1. One-machine infinite-bus system.

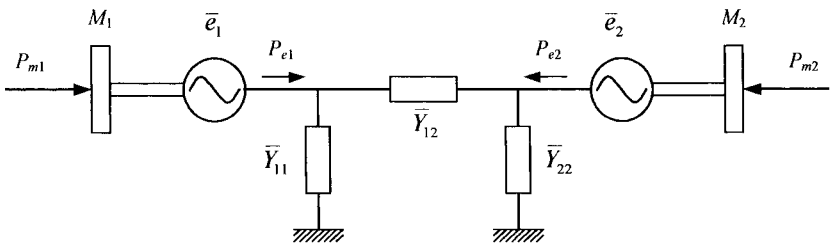


Figure 2. Two-machine system.

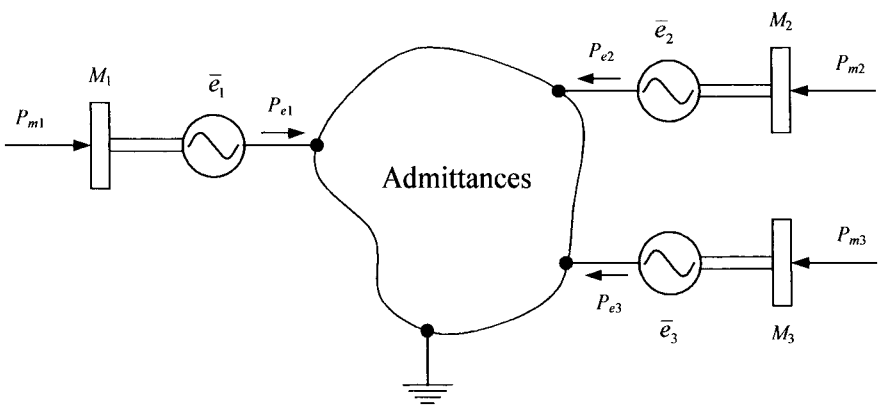


Figure 3. Three-machine system.

The meaning of symbols is as follows: P_m mechanic input power; M inertia constant; P_e electric output power; P_L load power; $\bar{e}$ internal e.m.f; X reactance; $\bar{V}$ infinite bus voltage; δ rotor angle, Y_{ij} transfer-admittance between busses i and j, Y_{ii} self-admittance of bus i.

Besides:

$$\ddot{\delta} = \frac{(P_m - P_L) - P_e}{M} \tag{1.1}$$

$$P_e = \frac{eV}{X}\sin\delta \tag{1.2}$$

In a two-machine system

$$\ddot{\delta}_1 = \frac{(P_{m1} - P_{L1}) - P_e}{M_1} \quad ; \quad \ddot{\delta}_2 = \frac{(P_{m2} - P_{L2}) + P_e}{M_2} \tag{1.3}$$

$$P_e = k\sin(\delta_1 - \delta_2) \quad ; \quad k = \text{const.} > 0 \tag{1.4}$$

$$\ddot{\delta}_1 - \ddot{\delta}_2 = \left[\frac{(P_{m1} - P_{L1})}{M_1} - \frac{(P_{m2} - P_{L2})}{M_2}\right] - \left(\frac{1}{M_1} + \frac{1}{M_2}\right)P_e \tag{1.5}$$

Eq. (1.5) is equivalent to

$$\frac{M_1 M_2}{M_1 + M_2}(\ddot{\delta}_1 - \ddot{\delta}_2) = \frac{M_2(P_{m1} - P_{L1}) - M_1(P_{m2} - P_{L2})}{M_1 + M_2} - P_e \tag{1.6}$$

i.e.

$$\frac{M_1 M_2}{M_1 + M_2}\ddot{\delta}_{12} = P_0 - k\sin\delta_{12} \tag{1.7}$$

with

$$\delta_{12} = \delta_1 - \delta_2 \tag{1.8}$$

$$P_0 = \frac{M_2(P_{m1} - P_{L1}) - M_1(P_{m2} - P_{L2})}{M_1 + M_2} \tag{1.9}$$

In a three-machine system one can check that the dynamic equations are of the following type:

$$M_1\ddot{\delta}_1 = P_{m1} - P_{e1} \; ; \quad M_2\ddot{\delta}_2 = P_{m2} - P_{e2} \; ; \quad M_3\ddot{\delta}_3 = P_{m3} - P_{e3} \tag{1.10}$$

$$
\begin{aligned}
P_{e1} &= e_1^2 g_{11} + e_1 e_2 [b_{12}\sin(\delta_1 - \delta_2) + g_{12}\cos(\delta_1 - \delta_2)] + \\
&\quad e_1 e_3 [b_{13}\sin(\delta_1 - \delta_3) + g_{13}\cos(\delta_1 - \delta_3)] \\
P_{e2} &= e_2^2 g_{22} + e_2 e_1 [b_{12}\sin(\delta_2 - \delta_1) + g_{12}\cos(\delta_2 - \delta_1)] + \\
&\quad e_2 e_3 [b_{23}\sin(\delta_2 - \delta_3) + g_{23}\cos(\delta_2 - \delta_3)] \\
P_{e3} &= e_3^2 g_{33} + e_3 e_1 [b_{31}\sin(\delta_3 - \delta_1) + g_{31}\cos(\delta_3 - \delta_1)] + \\
&\quad e_3 e_2 [b_{32}\sin(\delta_3 - \delta_2) + g_{23}\cos(\delta_3 - \delta_2)]
\end{aligned}
\tag{1.11}
$$

with b_{ij} transfer-susceptance; g_{ij} transfer-conductance. The state variables are the relative angles $(\delta_1 - \delta_3)$ and $(\delta_2 - \delta_3)$ as well as the relative speeds

$(\dot{\delta}_1 - \dot{\delta}_3)$ and $(\dot{\delta}_2 - \dot{\delta}_3)$.

2. EMERGENCY CONDITIONS

2.1 The concept of Area (or subsystem)

A large power system, like e.g. the italian power system or the european power system, can be usually described as an interconnection of a restricted number of subsystems, called "areas" which exchange with each other convenient fluxes of power and energy. For our didascalic purposes we consider a three-area system (fig. 4), in which a large area (3) injects power into a system formed in turn by two interconnected areas (1 and 2). We want to analyze the consequences on Areas 1 and 2 of a sudden interruption of the flow from area 3. We assume that both areas 1 and 2 are schematized by a synchronous machine and a given electric load (power consumption).

Two typical emergencies are considered:

1. The frequency drop is higher than allowed by admissible physical limitations;
2. Area 2 is unstable with respect to Area 1 and, then, a separation occurs.

2.2 Control apparatus

Voltage regulators and speed governors are standard control apparatus in normal operation. They keep frequency and voltage in view of assuring high quality service.

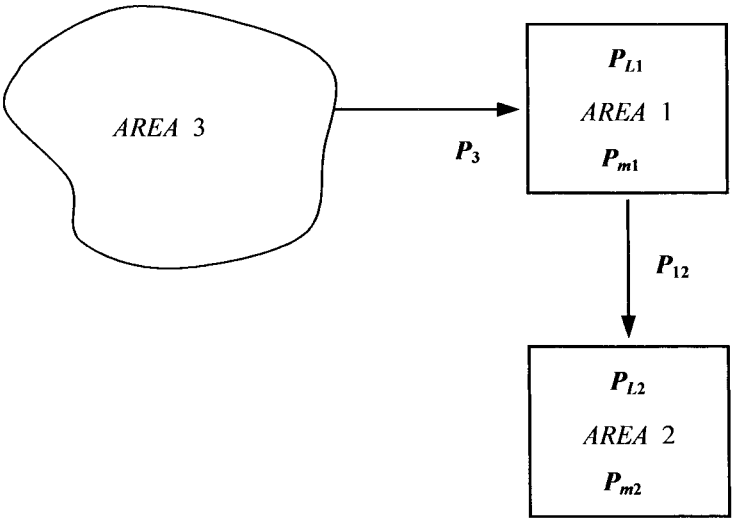


Figure 4. Three-Area system.

However their contribution to network security in emergency conditions can be considered negligible. By far more important are "load-shedders", "distance protections" on transmission lines, minimum frequency or maximum frequency breakers on power plants (electric machines).

Automatic load shedding foresees load rejection (of convenient and prescheduled entity) when frequency drops below 49.5 Hz, or 49Hz, or 48.5 Hz, in a network with nominal frequency 50 Hz. This kind of drops are due to unbalance between generated power and consumed power.

If, however, frequency falls down a critical threshold about 48 Hz, minimum frequency breakers automatically disconnect generation plants, thus giving origin to a total black-out.

"Distance" protections on transmission lines are "minimum impedance" protections, namely "switches" that open the line when the ratio voltage/ current (i.e. the "impedance") falls below preestablished levels. Typically this kind of event occurs due to unstability between connected Areas (e.g. Areas 1 and 2 in our example), and results in disconnection of Areas. Indeed unstability determines large oscillations in voltage and current and then in "impedance". A strong drop of impedance i.e. occurs when voltage reaches a minimum and simultaneously current reaches a maximum. The ultimate effect is separation among Areas, so that each Area must face emergency with its own control resources, without support from others. This represents a very critical situation that eventually degenerates into a total black-out.

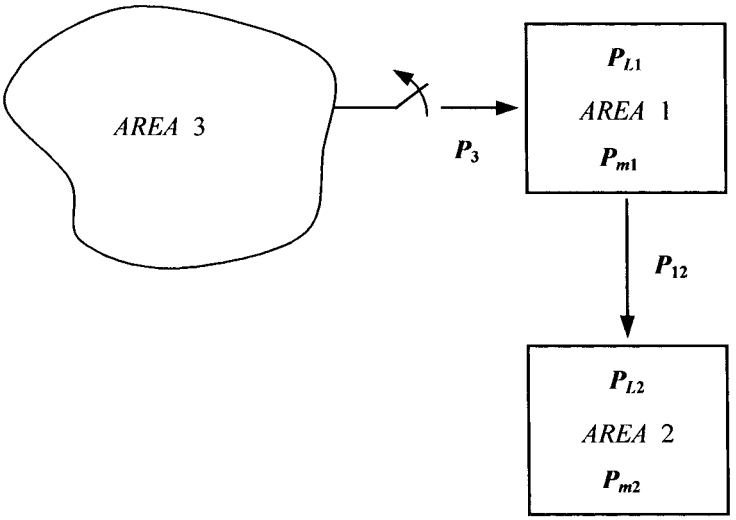


Figure 5. Disconnection of Area 3.

2.3 Preemergency control

In some circumstances, i.e. when the value of the imported power is particularly high, it can be safe to make a scheduled and preventive load shedding so that the power unbalance in case of (automatic and unwanted) trip of the importing transmission lines be smaller.

2.4 The role of stability

If interconnection between exporting Area 3 and the remainder of the system is abruptly interrupted (e.g. due to an unforeseen current overload), the isolated subsystem formed by Areas 1 and 2 behaves like an excited two-machine system with initial kinetic energy different from zero and with final equilibrium different from preexisting equilibrium. The stability problem is to see whether the initial kinetic energy is lower than the maximum value compatible with "stability in the large" of the final equilibrium. We use the "equal area" criterion (fig. 5) to determine such value: a one-machine infinite bus system is stable if the Lyapunov-function defined by

$$V(\delta,\dot{\delta}) = \frac{1}{2}M\dot{\delta}^2 + \int_{\delta_0}^{\delta}[P_e - P_m]\cos\delta$$

satisfies the condition

$$V(t=0^+) < V_{Lim}\ ; \qquad V_{Lim} = \int_{\delta_0}^{\delta_m} [P_e - P_m]\, d\delta\ ; \qquad \delta_m = \pi - \delta_0$$

In particular the initial kinetic energy must be lower than V_{Lim} :

$$\frac{1}{2} M \dot{\delta}^2 (0^+) < V_{Lim}$$

the quantity V_{Lim} is called *Limit value* of the V-function. Also, V-function is tied to Energy by the relation $V = \omega_s E$, where ω_s is synchronous speed.

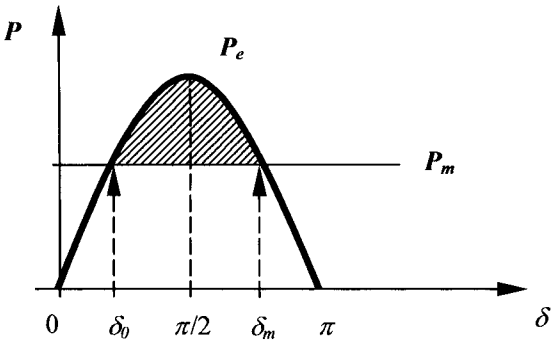


Figure 6. Equal-area stability criterion.

3. CASE STUDIES

In order to strongly simplify numerical analysis we make reference to round ciphers, which are not meaningful by themselves, but in terms of ratios between couples of values. They can be regarded as quantities in p.u. (per unit): i.e. a Power 70 referred to a nominal power 1000 MVA, corresponds to 70000 MW.

Area 1 has an electric load $P_{L1} = 70$, an inertia constant $M_1 = 1$.
Area 2 has an electric load $P_{L2} = 30$, an inertia constant $M_2 = (1/3)$.
The total load of Area 1 + Area 2 is: $P_L = P_{L1} + P_{L2} = 100$. Thus

$$\frac{P_{L1}}{P_L} = 0{,}7\,; \qquad \frac{P_{L2}}{P_L} = 0{,}3 \tag{3.1}$$

$$M_2 = \frac{M_1}{3}\,; \qquad M_1 + M_2 = \frac{4}{3}\,; \qquad \frac{M_1 M_2}{M_1 + M_2} = \frac{1}{4} \tag{3.1}$$

The electric power flowing from Area 1 to Area 2 is assumed to satisfy the equation

$$P_e = 20 \sin \delta_{12} \tag{3.2}$$

The electric power P_3 flowing from Area 3 to Area 1 equals a given percentage of the total load of Areas 1 and 2, while the mechanical input powers on Areas 1 and 2 are such that the total power balance is satisfied:

$$P_3 + P_{m1} + P_{m2} = P_L = 100$$

We consider three basic initial conditions regarding power P_3, and analyze the consequences of a sudden and unforeseen interruption of flow P_3 (see table 1).

Table 1. Case Studies 1, 2, 3. [(*a*) before load-shedding; (*b*) after load-shedding]

	P_{L1}	P_{M1}	P_{L2}	P_{M2}	P_3	ΔP_L	P_{12}
1 (*a*)	70.0	68.8	30.0	27.0	5.0 %	5.0	3.0
(*b*)	66.5	68.0	28.5	27.0	0.0	0.0	1.5
2 (*a*)	70.0	64.0	30.0	26.0	10.0 %	10.0	4.0
(*b*)	63.0	64.0	27.0	26.0	0.0	0.0	1.0
3 (*a*)	70.0	55.0	30.0	20.0	25.0 %	25.0	10.0
(*b*)	52.5	55.0	22.5	20.0	0.0	0.0	2.5

3.1 Case Study 1

We assume (see fig. 4) $P_3 = 5$; $P_{m1} = 68$; $P_{m3} = 27$; $P_{12} = 3$. Then

$$P_e = 3\,; \quad 20 \sin \delta_{12}^0 = 3\,; \quad \sin \delta_{12}^0 = \frac{2}{20}\,; \quad \delta_{12}^0 = 0.15056 \text{ rad} = 8.626°$$

The above values define the predisturbance Equilibrium. After disconnection of flow P_3, the system formed by Areas 1 and 2 has a power deficit of 5% and consequently undergoes a frequency drop

$$-\frac{\Delta f}{50\,\text{Hz}} = 5\% \quad \rightarrow \quad \Delta f = -2.5\ \text{Hz}$$

Such drop cannot be tolerated and, hence, it is counterbalanced by a convenient automatic load shedding, initiated by the drop of frequency.

What happens if the entity of shedding is exactly 5%, equitably distributed between Area 1 and Area 2? Clearly the system reaches a final equilibrium, but this equilibrium is different from the original one (see fig. 5).

Denoting with * the quantities after load shedding we have

$$P_{L1}^* = P_{L1} - 5\% P_{L1} = (70 - 3{,}5) = 66{,}5$$
$$P_{L2}^* = P_{L2} - 5\% P_{L2} = (30 - 1{,}5) = 28{,}5$$
$$P_{m1}^* = P_1 = 68\,; \qquad P_{m2}^* = P_2 = 27; \qquad P_e^* = 1.5 \neq P_e\,;$$
$$20 \sin \delta_{12}^* = 1.5\,; \qquad \delta_{12}^* = 0.0750\,\text{rad} = 4.3°$$

Is the transition (from the original equilibrium to the final equilibrium) stable? Using the equal-area stability criterion we find

$$V_{Lim} = \int_{\delta_{12}^*}^{\tau - \delta_{12}^*} [20 \sin \delta_{12} - 1.5]\, d\delta_{12} = 40.399 - 4.487 = 35.912$$

while

$$V(0^+) = V_{kin}(0^+) + \int_{\delta_{12}^*}^{\delta_{12}^0} [20 \sin \delta_{12} - 1.5]\, d\delta = V_{kin}(0^+) + 0.0043$$

where

$$V_{kin}(0^+) = \frac{M_1 M_2}{M_1 + M_2} \ddot{\delta}_{12}(0^+) = \frac{1}{4}\left[\ddot{\delta}_1(0^+) - \ddot{\delta}_2(0^+)\right]$$
$$\ddot{\delta}_1(0^+) = \frac{P_{m1} - P_{L1}^* - P_e}{M_1} = \frac{68 - 66.5 - 3}{1} = -1.5$$
$$\ddot{\delta}_2(0^+) = \frac{P_{m2} - P_{L2}^* + P_e}{M_2} = \frac{27 - 28.5 + 3}{(1/3)} = 4.5$$

One then has

$$V_{kin}(0^+) = \frac{1}{4}[-1.5 - 4.5]^2 = 9\,; \quad V(0^+) = 9 + 0.0043 = 9.0043$$

and, as $V_{Lim} = 39.9$, one finds $V(0^+) < V_{Lim}$ which means that the transition is stable.

In conclusion: *an automatic load shedding of 5% both in Areas 1 and 2 satisfactory counteracts a power loss of 5%, both in equilibrium and in transient condition. The perturbed system behaviour turns out to be stable.*

3.2 Case Study 2

The initial value of power flow from Area 3 is 10 MW and the initial values of mechanic input powers on Areas 1 and 2 are respectively equal to 64 and 26 (see fig. 4): $P_3 = 10$; $P_{m1} = 64$; $P_{m3} = 26$.

$$P_3 + P_{m1} + P_{m2} = 100 = P_{L1} + P_{L2}$$

Hence abrupt tripping of the interconnection line from Area 3 determines a power perturbation equal to 10%. Automatic load shedding 10% is necessary to stop frequency drop.

The initial value of power flow from Area 1 to Area 2 was $P_{12} = 4$ and the corresponding angle was the solution of eq.

$$20 \sin \delta_{12}^0 = 4 \quad \rightarrow \quad \delta_{12}^0 = 0.3041 \text{ rad} = 17.45°$$

After interruption of the flow from Area 3 (10%) and automatic load shedding (10%) equitably distributed between Areas 1 and 2 we have (fig. 5)

$$P_{L1}^* = P_{L1} - (10\%)P_{L1} = 63; \qquad P_{L2}^* = P_{L2} - (10\%)P_{L2} = 27$$

$$P_{m1} = 64\,; \qquad P_{m2} = 26$$

$$\ddot{\delta}_1 = \frac{64 - 63 - P_e^*}{M_1} = 1 - P_e^*\,; \qquad \ddot{\delta}_2 = \frac{26 - 27 + P_e^*}{(1/3)} = 3 + 3P_e^*$$

$$\ddot{\delta}_1 = \ddot{\delta}_2 \quad \rightarrow \quad 1 - P_e^* = 3 + 3P_{e\,2}^* \quad \rightarrow \quad 4P_e^* = 4 \quad \rightarrow \quad P_e^* = 1$$

$$20 \sin \delta_{12}^* = 1\,; \qquad \delta_{12}^* = 0.05 \text{ rad} = 2.86°$$

The limit value of the Lyapunov function referring to the equal-area criterion is given by

$$V_{Lim} = \int_{\delta_{12}^*}^{\tau-\delta_{12}^*} [20\sin\delta_{12} - 1]d\delta_{12} = 39.95 - 3.04 = 36.91$$

The initial kinetic energy is

$$V_{kin}(0^+) = \frac{1}{4}\left[\ddot{\delta}_1(0^+) - \ddot{\delta}_2(0^+)\right]$$

$$\ddot{\delta}_1(0^+) = \frac{P_{m1} - P_{L1}^* - P_e(0^-)}{M_1} = \frac{68-63-4}{1} = -3$$

$$\ddot{\delta}_2(0^+) = \frac{P_{m2} - P_{L2}^* + P_e(0^-)}{M_2} = \frac{26-27+4}{(1/3)} = 9$$

$$\ddot{\delta}_1(0^+) - \ddot{\delta}_2(0^+) = -3 - 9 = -12$$

and hence

$$V_{kin}(0^+) = \frac{1}{4}(-12)^2 = 36$$

Furthermore

$$\int_{\delta_{12}^*}^{\delta_{12}^0} [20\sin\delta_{12} - 1]d\delta_{12} = 0.638$$

and then $V(0^+) = V_{kin}(0^+) + 0.638 = 36.638$.
As $V_{Lim} = 36.91$, the above equations implicate $V(0^+) < V_{Lim}$ and then the transition is stable. However $V(0^+) \approx V_{Lim}$ so that stability margin is practically equal to zero.

In conclusion, *a large perturbation (10%) on the isolated system formed by Areas 1 and 2, is efficaciously counteracted by a 10% automatic load-shedding equitably distributed between Areas 1 and 2, without rupture of the interconnection between such Areas. However the stability margin is scanty.*

Things are substantially worse if perturbation ("disturbance") is extremely severe, e.g. 25% as postulated in case study 3.

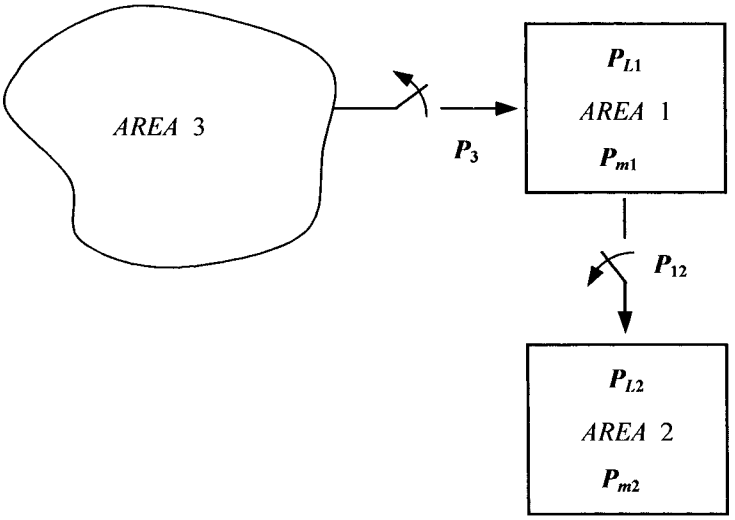


Figure 7. Double separation.

4. EXTREMELY SEVERE DISTURBANCES

4.1 Inadequacy of standard control strategies

If the initial value of the flow from Area 3 to Areas 1 and 2 was very high, e.g. 25, interruption of such flow would cause 25% loss of power, i.e. an extremely severe disturbance. In this case an automatic load-shedding of 25%, equitably distributed between Areas 1 and 2, would not only be problematic, but inadequate to keep stability and then connection. The final result would be a total black-out in both Areas.

In fact, let us assume e.g. that the predisturbance equilibrium is defined (fig. 4) by the following values:

$$P_3 = 25; \qquad P_{L1} = 70; \qquad P_{m1} = 55; \qquad P_{L2} = 30; \qquad P_{m2} = 20$$

$$P_{12} = 10; \qquad 20 \sin \delta_{12}^0 = 10\,; \qquad \delta_{12}^0 = 0.3041\,\text{rad} = 0.5235°$$

After interruption of power flow P_3, and 25% load shedding equitably distributed between Areas 1 and 2, we have a new final Equilibrium (fig. 5) with

$$P_{L1}^* = 52.5; \qquad P_{m1} = 55; \qquad P_{L2}^* = 22.25; \qquad P_{m2} = 20$$
$$P_{12} = 2.5; \qquad 20\sin\delta_{12}^* = 2.5\,; \qquad \delta_{12}^* = 0.1253\,\text{rad}$$

The limit value of the Lyapunov function is

$$\begin{aligned} V_{Lim} &= -20[\cos(\pi - \delta_{12}^*) - \cos\delta_{12}^*] - 2.5(\pi - 2\delta_{12}^*) \\ &= 39.676 - 7.227 = 32.45 \end{aligned}$$
$$\ddot{\delta}_1(0^+) = \frac{55 - 52.5 - 10}{1} = -7.5; \qquad \ddot{\delta}_2(0^+) = \frac{20 - 22.5 + 10}{(1/3)} = 22.5$$
$$\ddot{\delta}_1(0^+) - \ddot{\delta}_2(0^+) = -30; \qquad V_{kin}(0^+) = \frac{1}{4}(-30)^2 = 225 >> V_{Lim}$$

Then, as

$$V(0^+) = V_{kin}(0^+) + \int_{\delta_{12}^*}^{\delta_{12}^0} [20\sin\delta_{12} - 2.5]d\delta_{12} > V_{kin}(0^+)$$

it is $V(0^+) >> V_{Lim}$. *This means that transition from old to new Equilibrium is unstable.*

Isolated system formed by Areas 1 and 2 splits itself in two parts (fig. 7) in each of which frequency takes critical values (Δf = +2.25 Hz in Area 1 and Δf = –6.25 Hz in Area 2) leading to automatic exclusion of generation plants, and then to total black-out.

4.2 Alternative strategies

A radically alternative strategy is to shed load in a selective way, without pursuing an equitable distribution between Area 1 and Area 2. We show that the best strategy is to concentrate automatic load shedding mostly in Area 1. Let us first consider an intermediate strategy in which we shed 31.4% in Area 1 and 10% in Area 2. Then

$$P_{L1}^* = 70 - (31.4\%)70 = 48; \qquad P_{L2}^* = 30 - (10\%)30 = 27$$
$$P_{m1} = 55\,; \qquad P_{m2} = 20\,; \qquad P_{L1}^* + P_{L2}^* = 75 = P_{m1} + P_{m2}$$
$$P_{12} = 7\,; \qquad 20\sin\delta_{12}^* = 7\,; \qquad \delta_{12}^* = 0.6228\,\text{rad}$$

$$V_{Lim} = \int_{\delta_{12}^*}^{\pi-\delta_{12}^*} [20\sin\delta_{12} - 7]d\delta_{12} = 32.489 - 13.27 = 19.21$$

$$\ddot{\delta}_1(0^+) = \frac{(55-48)-10}{1} = -3; \qquad \ddot{\delta}_2(0^+) = \frac{(20-27)+10}{(1/3)} = 9$$

$$\ddot{\delta}_{12}(0^+) = \ddot{\delta}_1(0^+) - \ddot{\delta}_2(0^+) = -9; \quad V_{kin}(0^+) = \frac{1}{4}(-9)^2 = 20.25 > V_{Lim}$$

The transition is *unstable*.

On the contrary *if we make 5.5% shedding in Area 2 and 33% shedding in Area 1*, we get

$$P_{L1}^* = 70 - (33\%)70 = 46.66; \qquad P_{L2}^* = 30 - (5.55\%)30 = 28.33$$

$$P_{m1} = 55; \qquad P_{m2} = 20; \qquad P_{L1}^* + P_{L2}^* = 75 = P_{m1} + P_{m2}$$

$$P_{12}^* = 8.33; \qquad \delta_{12}^* = 0.429\,\text{rad}$$

$$V_{Lim} = \int_{0.429}^{\pi-0.429} [20\sin\delta_{12} - 10]d\delta_{12} = 36.3 - 19.022 = 17.012$$

$$\ddot{\delta}_1(0^+) = \frac{(55-46.66)-10}{1} = -1.66$$

$$\ddot{\delta}_2(0^+) = \frac{(20-28.33-0)+10}{(1/3)} = 5$$

$$\ddot{\delta}_{12}(0^+) = -6.66; \qquad V_{kin}(0^+) = \frac{1}{4}(-6.66)^2 = 11.1$$

$$V(0^+) = V_{kin}(0^+) + \int_{\delta_{12}^*}^{\delta_{12}^0} [20\sin\delta_{12} - 10]d\delta_{12} = 11.02$$

The transition is *stable*.

The proposed strategy is exceptional and should be based on prescheduled sectioning schemes on Area 1, so as to separate an entire large subsystem (e.g. load of large cities) when frequency time derivative exceeds preset and exceptional values. The sampling period should be about 0.1 s, which is compatible with performance of electronic frequency meters. This strategy fully preserves Area 2 and, at one time, gives to Area 1 concrete chances of restoration in a few minutes time interval. Note that *preservation of Area 2 is in interest of Area 1. Excessive shedding in Area 2 would cause instability and, then, ruinous black-outs both in Area2 and in Area 1.*

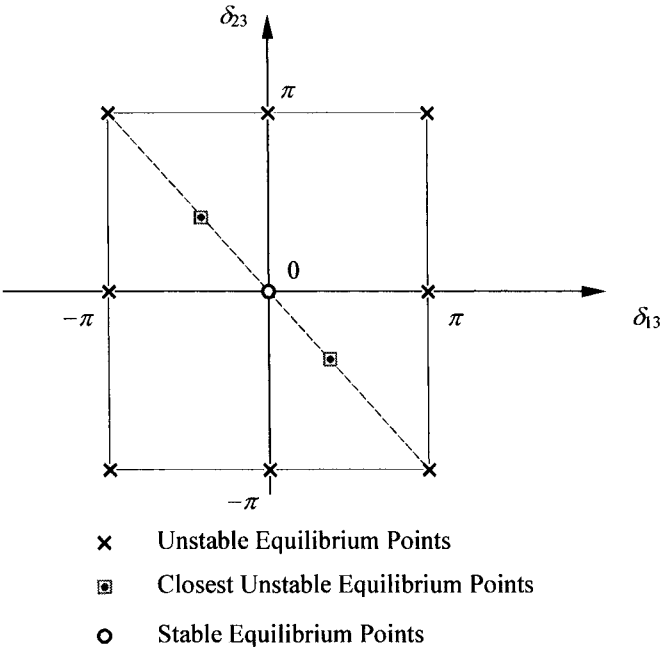


Figure 8. Equilibrium Points in a three-machine system.

5. OTHER SIDES OF COMPLEXITY

We have seen complexity from one of its more important aspects, i.e. mutual interactions among areas of a large interconnected system during emergencies. Another side worth mentioning is non-linearity and its interlacing with complexity in view of analytic formulation of conditions for stability "in the large" (according to nonlinear stability theory) vs. large disturbances. We have various points:

1. multiplicity of equilibria;
2. multiplicity of oscillation modes;
3. non-conservation of energy.

A satisfactory analysis and discussion is given in quoted references. Here we confine ourselves to a very simple though illuminating matter, i.e. multiplicity of equilibria. As an example, in a three-machine system with negligible transfer-conductances, we have a variety of Equilibrium points (fig. 8), only one of which is stable. The size of the *Stability Region* turns out determined by the *closest unstable Equilibrium*, i.e. the one in which the Lyapunov function takes on its minimum positive value. In order to determine such point we must use sophisticated numerical methods, e.g. by

optimizing of convenient performance indexes. Complexity strongly increases with the number of synchronous machines. Preliminary machine grouping based upon coherency-based *dynamic equivalents* greatly helps. In addition we need general criteria for defining the energy of non-conservative systems with n degrees of freedom.

6. CONCLUSION

We have shown typical emergencies in a large power system formed by three interconnected Areas. We have shown that in order to avoid extended black-outs we must simultaneously control frequency and interareas oscillations (stability). Both tasks can be substantially achieved via automatic load-shedding but the important point is that, in case of extremely severe disturbances, one must carefully distribute the shedded load among various areas. Otherwise unstable oscillations arise which finally lead to total black outs, in spite of load shedding.

REFERENCES

Byerly, R. T., and Kimbark, E.W., 1974, *Stability of Large Electric Power Systems*. IEEE Press, New York, N. Y.

Di Caprio, U., and Saccomanno, F, 1970, Nonlinear stability analysis of multimachine power systems, *Ricerche di Automatica* **1**.

Di Caprio, U., 1972, An approach to the on-line evaluation of stability margins in multi-area systems, *IV PSCC*, Grenoble.

Di Caprio, U., 1979, Problemi di sicurezza dinamica in una rete elettrica - (Dynamic security problems in power systems), *ENEL Rassegna Tecnica di Problemi dell' Energia Elettrica* **27**(5).

Di Caprio, U., 1981, Conditions for theoretical coherency in mulimachine power systems, *Int. Jour. of Automatica* (September).

Di Caprio, U., 1981, Controllo e dinamica dei sistemi elettici - (Control and dynamics of electric power systems), *ENEL Rassegna Tecnica di Problemi dell' Energia Elettrica*, fasc. 4, (July).

Di Caprio, U., 1982a, Use of Lyapunov and Energy methods for stability analysis of multimachine power systems, in: *Proc. of the Int. Symposium on Circuits and Systems*, Rome, May, p. 581.

Di Caprio, U., 1982b, Emergency Control, *Int. Jour. of EPES* **4**(1).

Di Caprio, U., 1982c, Theoretical and practical dynamic equivalents in multimachine power systems - Part I, *Int. Jour. of EPES* (October).

Di Caprio, U., 1983, Theoretical and practical dynamic equivalents in multimachine power systems - Part II, *Int. Journ. of EPES* (January).

Di Caprio, U., 1984, Status of power system research at ENEL, *Int. Jour. of EPES* (April).

Di Caprio, U., 1985, Practical and structural coherency in multimachine power systems, *Int. Jour. of EPES* (July).

Di Caprio, U., 1986, Lyapunov stability analysis of a synchronous machine with damping fluxes - Part I, *Int. Jour. of EPES* (January).

Di Caprio, U., 1987, Accounting for transfer conductance effects in Lyapunov stability analysis of multimachine power systems, *Int. J. of EPES* (July).

Di Caprio, U., 2002, The effect of friction forces upon stability in the large, *Int. J. of EPES* (May).

Di Caprio, U., 2002, The role of stability theory in the great theories of the XX centuy, In: *Emergence in Complex, Cognitive, Social and Biological Systems*, G. Minati and E. Pessa, eds., Kluwer Academic/Plenunm Publishers, New York, pp.127-140.

Di Caprio, U., Barretta L., and Marconato, R., 1976, The application of simplified dynamic models for the analysis of the parallel operation between the yugoslav and the italian power systems, and the evaluation of stabilizing signals, *I.E.E. Int. Conf. on On-line Operation and Optimization of Transmission and Distibution Systems*, London, June.

Di Caprio, U., Bruschi, G., and Marchese, V., 1981, Experience of use of the RIPES system for the detection of electromechanic disturbances in the ENEL network, *CIGRE Study Committee 32*, Rio de Janeiro, September.

Di Caprio, U., Clerici, E., Faro Ribeiro, L. F., and Nakamura, U., 1976, Digital and hybrid simulation studies to improve intersystem oscillation damping, *IEE PES Summer Meeting*, Portland, USA, July.

Di Caprio, U., Humphreys, P., and Pioger, G., 1982, The techniques and application of power system dynamic equivalents at CEGB, EDF and ENEL, *L'Energia Elettrica* **59**(12).

Di Caprio, U., and Marchese, V., 1982, Il sistema RIPES per la rivelazione e registrazione in tempo reale dei disservizi in una rete elettrica (The RIPES system for detection and real time recording of disturbances on power systems), *ENEL Rassegna Tecnica dei Problemi dell' Energia Elettrica*, fasc. 4, (July).

Di Caprio, U., and Marconato, R., 1975, A novel criterion for the development of multi-areas simplified models oriented to the on-line evaluation of power system dynamic security, *5th PSCC*, Cambridge, U.K., September.

Di Caprio, U., and Marconato, R., 1979, Automatic load-shedding in multimachine elastic power systems, *Int. Jour. of EPES* **1**(1).

Di Caprio, U., Marconato, R., and Mariani, E., 1974, Studio di alcuni piani per il controllo in emergenza di una rete elettrica a mezzo di alleggerimento automatico del carico (Emergency control plans by means of automatic load-shedding in an electric power system), *LXXV Riunione AEI*, A.89,Rome, September.

Di Caprio, U., Mariani, E., Ricci, P., and Venturini D., 1974, Simulation of power system behaviour under severe disturbances causing sequential trips of transmission lines or heavy power swings, *CIGRE Session*, 32-15, Paris, August.

Di Caprio, U., and Prandoni, W., 1988, Lyapunov stability analysis of a synchronous machine with damping fluxes - Part II, *Int. Jour. of EPES* (January).

Hahn, W., 1963, *Theory and application of Lyapunov's direct method*, Prenctice-Hall,

Hahn, W., 1967, *Stability of Motion*, Springer Verlag,

Huseyin, K., 1975, *Non linear theory of elastic stability*, Noordhoff Int. Publ., Leyden.

STRATEGIES OF ADAPTATION OF MAN TO HIS ENVIRONMENT: PROJECTION OUTIDE THE HUMAN BODY OF SOCIAL INSTITUTIONS

Emmanuel A. Nunez
AFSCET, Association Française de Science des Systèmes
1 rue de l'Echiquier, 78760 Jouars-Ponchartrain, France
Email: emmanuel.nunez@wanadoo.fr

Abstract: We present an hypothesis of the existence of analogies between the bio-psycho-cognitive living organism, working as a model, and social institutions. These institutions are created to protect man against stress and changes. This hypothesis is supported by: 1) the analogies which exist between an enterprise and living organism. 2) the existence of "out of body experiences" observed in some natural conditions and electrophysiological manipulations. Furthermore the possibility to project out of the subject a virtual object is one of the elements contributing to human identity and consciousness. A trinitrian situation is realized between the subject, the out of body object and the outside observer. This observer (mirror of the subject) is classically recognized as one of the essential factors needed for the subject identity construction which constitute one of the defense factors of a living organism or social institution. So, a "trinitrian intelligent loop" exist, allowing the emergence of the consciousness of the conscience.

Key words: bio-psycho-cognitive living organism; out of body experience; observer; social institutions.

1. INTRODUCTION

The reaction of a living organism to the action of a stressor must be compatible with life, avoiding detrimental unbalanced or irreversible attitudes.

When challenged by externally or internally (endogenous pathological situations) aggressive factors (biological, psychological or social), the organism develops an appropriate reaction by proceeding in phases (Nunez, 1995) (see figure below) The objective being to obtain temporization gaining time in order to recognize the identity of the stressor and build new weapons to neutralize, accept or incorporate the stressor.

The system may obtain this temporization using the first line of defense found in the stable "external identity" (e.g. skin, psycho-social personality) and the unstable, adaptive "internal identity" (homeostasis, immune system). The second phase can develop in two possible ways, both of which have as objective to escape from the reception of the importunate signal. One is to revert to a lower level of organization. We call this procedure "**retrogression**" (e.g. depression, ideologies ...). This mechanism can explain the expression of violence which appear in many circumstances characterized by cortex inhibition, with activation of the reptilian brain, induced by ideological (nationalism, integrism ...) or double bind (François, 1997) situations.

The second is to create temporarily a higher level of organization in the psychocognitive domain or in the immune network. We call this phenomenon "**supragression**" (e.g. activation of creativity, synthesis of new antibodies, divinities, angels, god ...).

Once these preparatory steps have been followed, the organism is then able to act by creating either new emergent concepts or new biological or artifactual procedures (e.g. vaccination, social institutions) respectively inside or outside the body. We call «**extracession**» the creation, outside the body, of artifacts or systems able to optimize the reaction to a stressor.

Thus, a biological or psychocognitive level of organization is converted, translated into an artifact which reproduces, with artificial constituents, the biological or the psychocognitive function. These artifacts can evolve outside the body, under human creative control, into more sophisticated systems. This projection will free or serve an organism function which requires a great deal of energy and therefore depletes the organism's energy capital. The resulting economy allows the organism to consecrate the conserved energy to the functioning of another already existing or emergent level. An illustration of this process is the example of the washing-machine liberating the house-keeper and thereby allowing her or him to perform other functions. Other extracessions can be the projection of a higher level of organization in technocognitive (e.g. computers), technological or social domains (e.g. enterprise) are created to protect man (figure 1).

An especially significant example of extracession-retrogression is represented by procreation. As death can be considered as an extreme form of retrogression, we can consider that one of the most remarkable biological

adaptation-protection to death is procreation. In this case we observe a process of extracession which stems from a retrogression. Thus, gametes may be considered as an archaic unicellular form of life with the potential to undergo, after fertilization, a certain number of phases of development in a controlled environment different from that prevailing during the development of their genitors.

The new being which results will develop, encountering new psychocognitive, sociocultural, and technological conditions which will enable it to create new emergent strategies of existence in order to achieve a better adaptation to its environment (e.g. easier adaptation of young persons to computation).

An other example of extracession which can be specified as an imitative-extracession is given by the construction of technological artifacts (boat, aircraft ...) taking objects (fishes, birds or floating-tree ...), observed in the nature, as models (Quilici-Pacaud, 1992).

2. STRUCTURAL AND FUNCTIONAL ANALOGIES BETWEEN A LIVING ORGANISM AND A SOCIAL INSTITUTION, THE INTERPRISE

The existence of an extracession mechanism from the body, used as a bio-psycho-cognitive model, to a social institution, appears to us as a rational explanation of the structural and functional analogies (subsystems associated and auto-regulated by an intertwined central hierarchic and peripheral networks of information) observed when we studied the enterprise by comparison with a living human organism. It is clear from our observations (Nunez, 1995) and other authors (Landier,1987, 1991; Fustec and Fradin, 2001; Foulard, 1998; Beer, 1972) that an enterprise can be considered as a living organism. Many analogies can be observed between an enterprise and a living organism. Both are non trivial systems having numerous intertwined subsystems or organs which activities are devoted to various complementary functions. These organs and functions communicate and are regulated by similar transferring integrative and regulatory information systems (top-down centralized transfer of information, feed-back regulation, peripheral information network regulated by ago-antagonistic systems; (Bernard-Weil, 1988).

It is possible to mention many other similar properties of both domains: birth, evolution, death, similar defense strategies, symbiotic associations etc.

3. POSSIBLE MECANISM OF EXTRACESSION, A PROPERTY OF LIVING ORGANISMS PROVIDING BOTH CREATION OF ARTIFACTS AND PROCREATION

Living and thinking organisms may create artifacts outside the body. The evolutionary goal of these objects is to become elements of defense, of survival for the individual and the species using a process which we call "extracession" as described and illustrated in the figure below. We describe this phenomenon in terms of a traduction-transduction from a biological factor to a psycho, technological, social factor. For example, a biological function performed by the hand may be replicated in the form of a prosthesis whose mechanical elements provide the same function. Artificial kidneys have likewise been developed having a blood purifying function.

Recent works (Blanke et al., 2002) sustained experimentally the existence of the hypothesis of extracession, showing that "out of body experiences", described as a personal feeling to be out of our body, looking it as an object, can be reproduced by brain electrophysiological stimulation. So, it can be possible to envisage that human brain is able to project out, part or totality of his body structure(s) or function(s). This filiation is somewhat hidden, owing to the fact that extracessions are realized from one domain (e.g. living matter) to another (e.g. prosthesis, social institutions) whose material structure can be very different but whose function(s) is (are) similar.

In the figure we introduce a new strategy of defense which can be used, by living organisms or not, to cope with aggressive factors. Thus, the use of fractal geometry (Mandelbrot, 1995) can be considered as a strategy that attenuate directly or indirectly the effects of a stressor, e.g. erosion of a coast (Sapoval, 2004; Sapoval et al., 2004; Baldassari, 2004).

The figure also shows the positive or negative control which exist between the extracessed features and the aggressive factor or stressor. In other word, the extraceded feature can directly or indirectly be aggressive or inhibitory (Simondon, 1989).

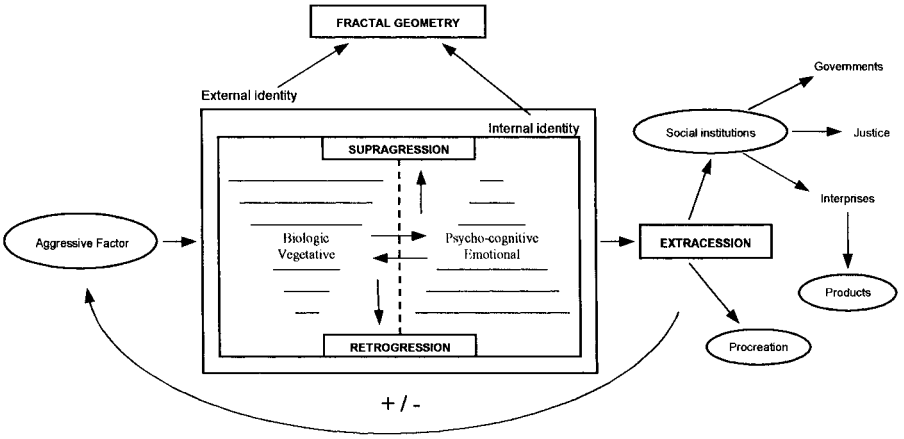


Figure 1. Representation of the varying intra- and extrabody strategies enabling a living organism to respond to stress.

This approach to the analysis of relationships existing between different domains also lays the groundwork for explanations of the motivation and sources of human creativity, sought in the study of the history and evolution of science and technology. We will develop elsewhere this subject.

In addition, the possibility to project out of the body a virtual object representing this body constitute one of the factors which contribute to human identity and consciousness.

In these conditions, a trinitrian situation (Morin, 1991; Donnadieu, 1992) is realized between the subject, the out of body projected subject, becoming so a virtual object, and the outside observer. This outside observer is classically recognized as an essential factor (mirror of the subject) needed for the construction of the identity of the subject. Identity, as seen before, being an important factor of defense.

So, a «trinitrian intelligent loop» is realized, allowing the emergence of the consciousness of the conscience.

BIBLIOGRAPHY

Baldassari, A., 2004, La percolotion en gradient: Des fronts de diffusion au fronts de mer, Abstracts of the *International congress on Benoit Mandelbrot 80th anniversary «Fractals en Progrès»*.

Beer, S., 1972, *Neurologie de l'Entreprise*, Presses Universitaires de France, Paris.

Bernard-Weil, E., 1988, *Précis de Systèmique Ago-antagoniste*, Interdisciplinaire, Limonest.

Blanke, O., Ortigue, S., Landis, T., and Seek, M., 2002, Stimulating illusory own body perception, *Nature* **419**:269-270.

Donnadieu, G., 1992, De quelques illustrations de la trialectique. A propos des interactions à l'œuvre dans les systèmes complexes, in: *Proceedings of the 5th EUSS*, *Acta systemica* (on line): http://www.afscet.asso.fr/res systemica .

Foulard, C., 1998, *L'entreprise Communicante*, Hermes, Paris.

François, C., 1997, Double bind, in: *International Encyclopedia of Systems and Cybernetics*, K. G. Saur, München.

Fustec, A., and Frandin, J., 2001, *L'entreprise Neuronale. Comment Maîtriser les Émotions et les Automatismes pour une Entreprise plus Performante*, Edition d'Organisation, Paris.

Mandelbrot, B., 1995, *Les Objets Fractals, Forme, Hasard et Dimension*, Flammarion, Paris.

Morin, E., 1991, *Introduction à la Pensée Complexe*, ESF, Paris.

Nunez, E. A., 1995, Analogies structurelles, fonctionnelles et évolutives des systèmes biologiques, psychocognitifs, sociaux et technologiques, in: *Proceedings AFCET Congress*, Toulouse, pp. 383-392.

Quilici-Pacaud, J. F., 1992, De la technologie comme modèle de représentation pour la conception (d'artefacts, outils) et de cognition en général, in: *Proceedings of the Second European School of System Science*, AFCET, Paris, pp. 281-282.

Simondon, G., 1998, *Du Mode d'Existence des Objets Techniques*, Aubier, Paris.

Sapoval, B., 2004, Résonateurs fractals et murs anti-bruits, Abstracts of the *International congress of Benoit Mandelbrot 80th anniversary «Fractals en Progrès»*.

Sapoval, B., Baldassari, A., and Gabrielli, A., 2004, Self stabilized fractality of seacoats through damped erosion, *Physic review letter* **9**:27-93.

EMERGENCE OF THE COOPERATION-COMPETITION BETWEEN TWO ROBOTS

Guido Tascini and Anna Montesanto
Dipartimento di Elettronica, Intelligenza Artificiale e Telecomunicazioni
Università Politecnica delle Marche, Ancona
E-mail: g.tascini@univpm.it

Abstract: The work studies the trajectories, the emergent strategies and the effects due to the interaction of two robots, in a simulated environment. The robots have the same task: crossing some fixed zones of the environment. The study is focused on the emergence of collaborative or competitive behaviour, which is valued by taking into account the interaction area locations and the impact of the interaction on the behaviour. The results of the research show emergent behaviours with a strong analogy with those of dominance in nature, in which animals organize itself in groups that follow particular geometries. The experiments highlight that the resulting interaction geometries depend on the agent evolution degree and on the interaction area locations, while the relationship between these two factors appears as reciprocal.

Key words: cooperation; competition; simulated agent; evolutionary learning.

1. INTRODUCTION

The interest on the emergent properties in robotic planning, is related to the fact that complex behaviours may evolve from simple assignments, given to a varying number of robot. In this vision the robots know only some aspects of the phenomenon that they will go to create and they do not need a global vision of the objective to achieve. This causes a reduction in the hardware costs, being in this case rather simple. The concept of emergency and the emergent theory of evolution was firstly introduced by Morgan in the book "Emergent Evolution" of 1923. In the same period the philosopher C. D. Broad (1925) argued about emergent properties with different levels of complexity.

The emergency, during many years, was conceived as relevant for the Biology. In fact, in biological evolution, is often possible to observe unpredictable characteristics on the base of the previously existing one. So we use the "emergent" attribute for pointing out something of "new" and "unpredictable". Afterward in different disciplines, with initial predominance of the physics, it was understood that the emergency conception was implicit in the general theory of the systems, proposed by Von Bertalanffy (1969): from a whole of interactive elements it could emerge behaviours and properties that are unpredictable by considering the simple features of the elements.

Normally the science, for studying complex behaviours, uses a reductionistic approach that tries to divide a complex "object" in simple slots that are singularly analyzed. Despite this method had a big success, it has a lot of limitations in: in fact it is often impossible forecasting the global behaviour of a dynamic system by using the acquired information from only constituent components. What escapes in this type of approach is definite emergent behaviour.

The emergent properties are features that born from this type of system and they bear from the interaction both between constituent elements and between these ones and environment. The more interesting and fascinating aspect is that this type of behaviour is not a priori defined.

Another interesting aspect is related to the partial acquaintance of the active constituent elements that often is limited to the phenomenon at microscopic level and to local information. In nature we can see different cases in which emergent behaviours are compared, like for instance the cloth made by a spider (Krink and, Vollrath, 1998) or the run of a group of gazelles. For instance an ant alone would not be able to plan, communicate with the others and build an ant hill, but a numerous group of ants could build sophisticate structures without need of any supervision. The examples that follow illustrate the concept of emergent properties in a different number of systems. We will show how they are developed agent-based solutions that supply emergent behaviours similar to those one in nature.

The economist Thomas Shelling (1969) formulated a sociological model in which affirmed that the varied forms of segregation that we could meet in nature, in the animals and in the man, like the grouping around the dominant animal or the birth of ghettos between men, they seem to be more rigid as regards the desires of the single individuals. His model consists of a grid-world constituted from two types of individuals each of which preferred to be surrounded from a certain percentage of individuals of the same type. The minority caused the migration of the individuals toward other subgroups containing a lot of elements of the same type giving origin to a positive feedback in which the micro pattern are amplified in macro pattern. The

departure of an individual from a subgroup induced the departure of individuals of the same type. Vice versa, their arrival in other subgroups pushed the individuals of the other type to take off. In this way we have a limit situation in which were delineated for the most part groups of the same type. The interesting appearance of this model was that the convergence on this type of structure was not deliberate and not genetically inherited. Much kind of animals tend to assemble itself in groups or crowds, by following a spatial structure in which the dominant individuals get to the centre and those subordinates get to the outskirts. Hamilton (1971), with his theory on the "selfish-herd", explains the reason of this disposition by affirming that this configuration have some advantages, the more evident being the defence. In substance the members in the centre, profit of a better protection from the raiders: this evolutionary form of behaviour is named "centripetal-instinct". A secondary advantage derives from this disposition, with some exceptions: it provokes a kind of visual trouble in the raider that is not able to well focus the prey having a whole group in movement.

Hemelrijk (1999), together with other collaborators, developed an agent-based model, named "Dom World", in which it is reproduced the competitive behaviours of a group of agents that attempt to conquest a hierarchical positions. From this search in simulated environment, the following emergent properties have been underlined:

- mutual help of the members of the group in the struggle
- reduction of the aggressions when the agents have well known each other.
- evidence of the phenomenon of spatial-centrality of the dominant agent.

The artificial creatures that populate this world have only two types of behaviour: grouping and having interactions of dominance. The interactions reflect the competition between the individuals for the division of the resources. When a member of the group invades the hierarchical space reserved to another one, bears a "dominance interaction" for the conquest of such space. If the attacker wins, it takes the place of the adversary, while if it loses is forced to take off. The combined effects of the positive feedback from the victory and of the strong hierarchical disposition, allow the system to converge on an emergent spatial configuration, that involves stratification in rings of dominance and presents the same spatial structure exposed in the theory of Hamilton without centripetal-instinct.

In the spatial-centrality, by gradually going from the centre toward the outside, we can find individuals that occupe hierarchical positions more and more low (weaker). Always Hemelrijk (2000), by departing from an artificial model similar to the previous one, have shown that by simply varying the intensity of the aggressions, it was possible to transform an arbitrary society, in which the benefits are concentred in individuals with

hierarchically elevated positions, into an equality society (Vehrencamp, 1983), with a more equitable distribution of the benefits.

2. MODEL OF THE ROBOT

The model of robot reproduced in the simulation is called Khepera and is product from the K-Team. The experiments that will follow are developed using a simulator. It is a software that reproduces the real robot (Khepera), the interactions of this one with the environment and those with other robot.

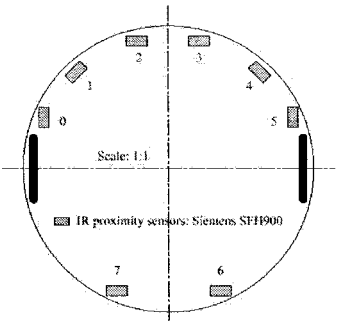


Figure 1. Spatial disposition of the IR sensors in the Khepera robot.

The object sampling is realised in 3D, but the simulation in practice evolves in a plane and the robot has only two freedom degree for translation. Besides the gravity is not taken into account.

The YAKS system allows simulating the real sensors of Khepera (fig. 1), as well as ad hoc sensors. It may simulate the following sensors: frontal IR of proximity ; back IR sensors ; array of light sensors , gripper, ground (for robot parking), compass (rotation angle); energy (parking on energy zone); rod (for recognition from another robot); rod sensor: gives 1 if a rod is detected in its vision field of 36 degrees.

2.1 The simulated environment

The simulator offers a virtual environment (world) constituted by walls and objects, that could be fixed or not, and gives the possibility of put in this world the artificial robots. All the actions performed by the robots and the structural characteristics of the elements that compose the world, are entirely managed from the core of the simulator. The simulator is called YAKS and

is written from Johan Carlsson of the Center for Learning System of the university of Scövde (Sweden).

YAKS presents these characteristics: the possibility of parallel evolution of more robot, the use of neural nets for the robot control, the use of genetic algorithms for the neural networks evolution, high use of parameter, easy expansion (bottom-up realization) and physics of the simulated world.

Figure 2 shows the graphic interface and the neural activities of YAKS.

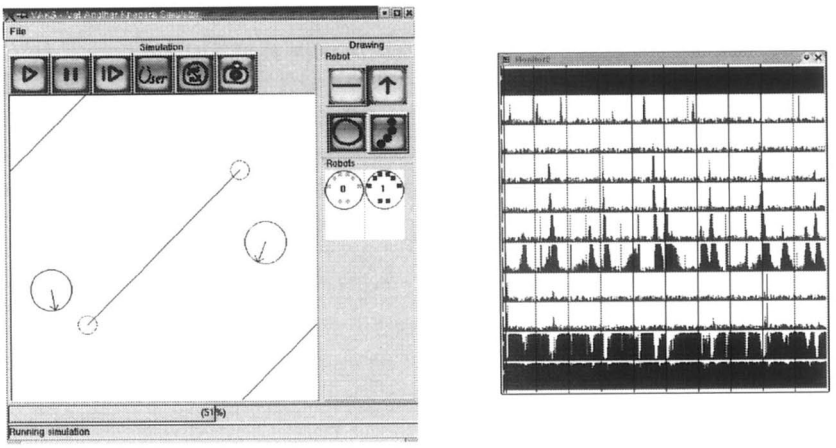


Figure 2. YAKS in execution in graphic mode and monitoring of the neural activities in YAKS.

The environments (worlds) is described by an ASCII file containing the list of objects that we want to insert; the coordinates are express in millimeters. Here we show an example of "world".

```
# The walls that constitute the external walls of the
simulated environment
wall 0.000000 0.000000 1000.000000 0.000000
wall 1000.000000 0.000000 1000.000000 1000.000000
wall 0.000000 1000.000000 1000.000000 1000.000000
wall 0.000000 0.000000 0.000000 1000.000000
# The (possible) departure position of the robot
start 640.0 440.0 90.0
```

The used operating system is the GNU/Linux, free variant and open source of Unix. The programming language is the C++.

2.2 The control system of the robot

The controller of the simulated robot is realised in evolutionary form. An artificial neural network (ANN) constitutes the control system. The input to the controller is constituted by the values registered from the sensors, the output by the amplitude of the motor command.

2.3 Models of ANN

The ANN used in the experiments is constituted by a neuron for each sensor in input and always from two neurons for the output, without hidden layers. The outputs were used to check the motors of the Khepera. The activation function of the neurons is a sigmoid. Each input neuron has a link with the two output neurons, with varying weight between -10 and +10 encoded with 8 bit. There are not recurrent connections; the structure of the ANN is therefore feed-forward. In figure 3 it is shown a three-dimensional representation of the structure of control when it has six frontal IR sensors and two back sensors.

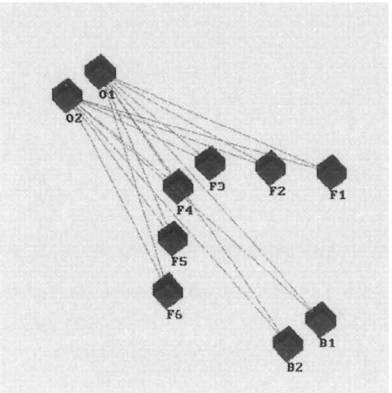


Figure 3. Neural structure relative to a Khepera endowed with six frontal sensors and two back sensors.

The red cubes represent the neurons, the green lines the links. The F are the neurons related to the frontal IR sensors, the B is those related to the back IR sensors, the O those relative to the exit, and therefore to the control of the motors.

2.4 Models of GA

The genetic algorithm present in YAKS offers different options relative to the simulation. Also in this case is used an ASCII line in which it is possible to define all the necessary parameters.

```
# Parameters of GA
GENERATIONS 100
START_GENERATION -1
INDIVIDS 100
EPOCHS 2
NR_OF_STEPS 100
PARENTS 20
OFFSPRINGS 5
SELECTION_METHOD 0  # default 0
BIT_MUTATION 1
FITNESS_FUNCTION 1
TEST_SAME_START_POSITION 0
```

3. EXPERIMENTAL PHASE

In this experiment we try to get behaviour of exploration of the robots in an open environment. In practice the environment was a box having dimension of 1000×1000 mm, besides there was not walls, corridors and obstacles. In the environment of square form, was insert 5 circular zones to explore. They were positioned along one of his edges. To detect these zones, has been necessary add a new sensor to the Khepera carrying the total number of his input neuron to 9. The fitness function in this experiment is very elementary being equal to the sum of the number of zones visited from the robot during the four epochs. We effect 20 simulations for every experiment, making to evolve a population of 100 individuals in 100 generations and 4 epochs with 1000 time steps of stay of the simulated robot in the environment. The interesting trajectories are harvests of evolution of the population from the 100 at 101 generation. We have to notice that the excursion of a single generation has involved the creation of almost 13000 trajectories, which are analyzed from a suitable program able to find those redundant, eliminate them and gather the remaining one in a directory.

3.1 Exploration of the environment and emergent properties

From the analysis of the trajectories birth during the simulations, we can see two kinds of emergent pattern, all runs in frontal movement:

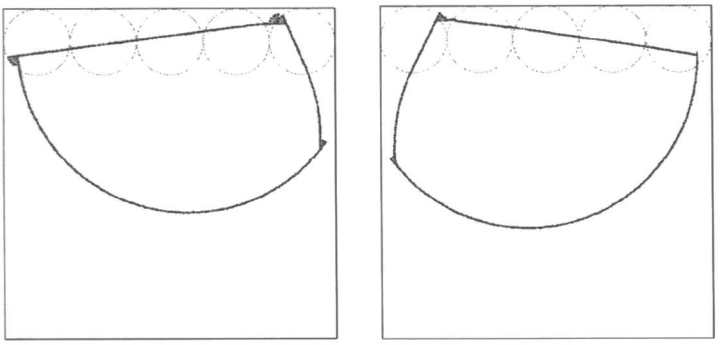


Figure 4. Curvilinear trajectories.

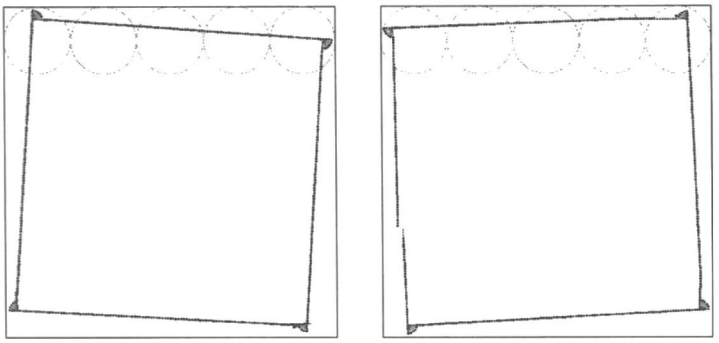


Figure 5. Rectilinear trajectories.

The rectangular pattern (fig. 5) and the semicircular one (fig. 4) differ between them for the sense of run, that are respectively clockwise and counter clockwise. Any characteristics of interest emerged for curvilinear pattern; they are the followings: the robot, from any point depart except from the zones, describes trajectories curvilinear; in the moment in which it enter in one of them, the angle of shift tends to settle, this means that it change typology of run from curvilinear to rectilinear. The zones are circles. So, While it cross the zones, it could happen that a part of his body escapes partially from those zones and the angle of shift suffers of the variations that tend to settle in the moment in which a possible and complete return is had

to the inside of the following zone. If the robot never meets the zones, the trajectory that describes is purely curvilinear. In the case of rectangular pattern it is not present an alternation of geometric typology of run (rectilinear, curvilinear) and the angle of crossing tends to sustain constant also if the body of the robot escapes partially from the zones.

4. THE INTERACTION BETWEEN ROBOTS

For the study of the interactions, we have grouped in range (by dividing them per robot) the trajectories that presented the same emergent pattern. The points of corresponding interaction are divided in groups and analyzed, showing in which way they condition the fitness function of the 2 robot.

4.1 Experiment 024-R1

In this first range the emergent pattern are represented from a degenerate semicircle for the blue robot (run in frontal movement and counter clockwise direction) (fig. 6) and from a rectangle for the red (run always in frontal movement, but the sense is counter clock) (fig. 7).

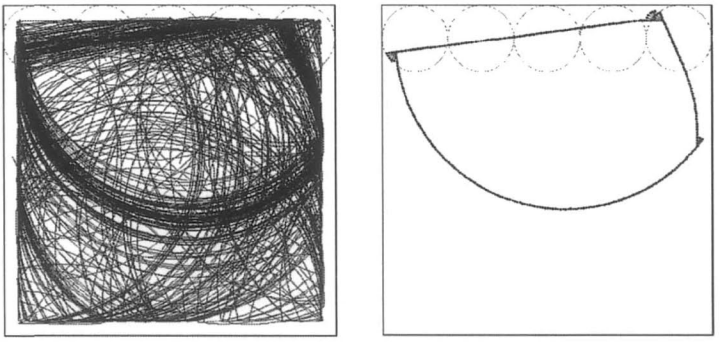


Figure 6. Superimposed trajectories and spatial emergent pattern (blue robot).

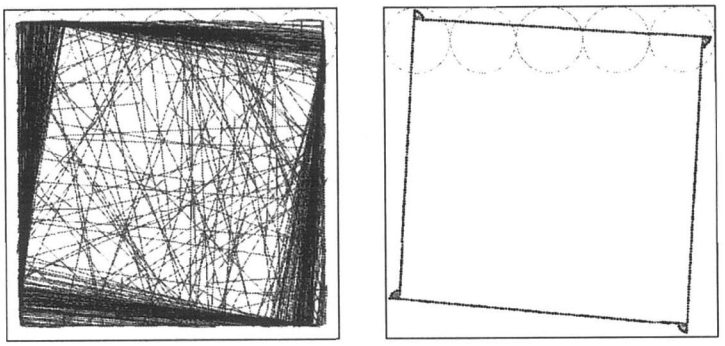


Figure 7. Superimposed trajectories and spatial emergent pattern (red robot).

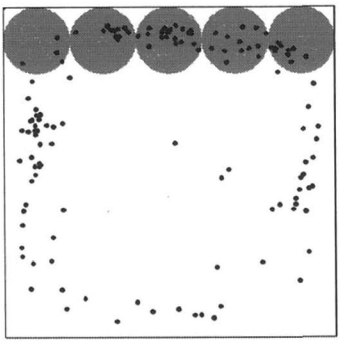


Figure 8. Spatial disposition of the points of interaction.

Figure 8 represents the spatial disposition of the points of interaction birth from the two robots; we can note immediately like these tend to be disposed prevalently along the perimeter of the environment, this is justified from the geometries of the 2 emergent pattern we have seen before that do not have points of intersection in the central zone of the environment. The present points in this zone are interactions that are verified before that the 2 robot begin to follow the direction of the proper emergent pattern. The more evident accumulations are along the left side, right side and in proximity of the central zones of the environment. We will show the variations of trajectory caused from these points. Particularly, we consider only those that carry to meaningful variations of fitness function in the 2 robot.

4.1.1 Accumulations along the left wall of the environment

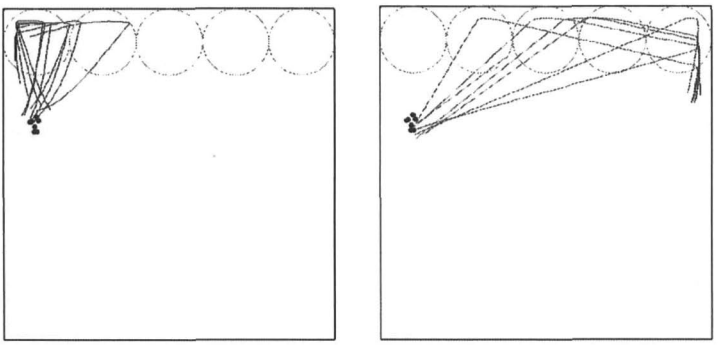


Figure 9. Trajectories post-interaction in the 2 robot (accumulations left side).

The points located in this zone (fig. 9) determine the deviations of trajectory, that correspond to a disadvantage (in terms of fitness) for the red robot and in an advantage for the blue one. We see in fact that the interaction pushes the blue again in the zones (from which originated the dominant pattern) while the red "loses" always the zone to the extreme left.

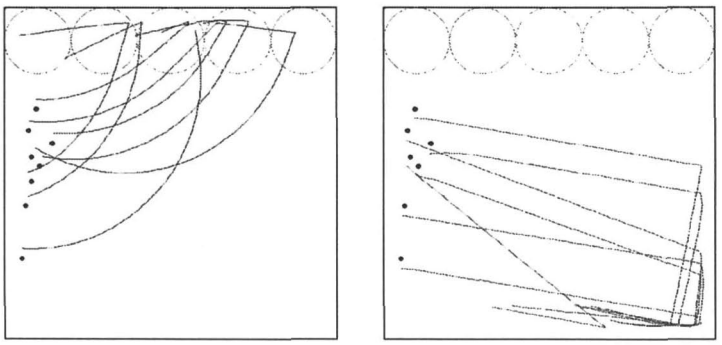


Figure 10. Trajectories post-interaction in the 2 robot (accumulations left side).

In the case of fig. 10 the blue robot profits in predominant manner from the interactive effects so his trajectories cover all the 5 zones. The red instead strongly loses: no trajectory of exit licks up any zone and it must spend a lot of time steps before it could try an entry in the zones again.

4.1.2 Accumulations along the zones

In this case (fig. 11), since both the trajectories of exit tend to carry the robots out of the zones, is not immediate verify the impact that this has on the fitness.

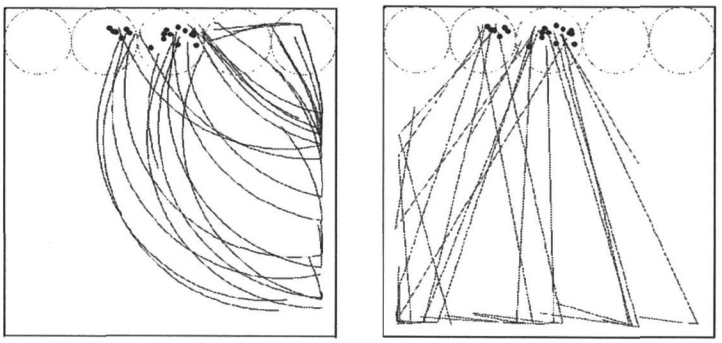


Figure 11. Trajectories post-interaction in the 2 robot (accumulations side zones).

We have compiled a series of charts in which the disposed values on the columns have the following meaning:

Table 1.

Trajectory	number of trajectory that is being considered
Gain (Blue, Red)	number of zones crosses because of the interaction (profit)
Lose (Blue, Red)	number of zones that the robots would have crossed one has not interacted between them (loss)

Zone A, B, C, D points out specific zones of the environment (Fig. 12) the figures contained between parenthesis they represent the relative averages to the values of each column.

Table 2. Zone A: the red robot lose 3 zones and the blu one lose 1 zone.

Trajectory	Gain Blue	Gain Red	Lose Blue	Lose Red
56	0	0	1	3
71	1	0	1	3
112	0	0	1	3
135	0	0	2	3
148	0	0	1	3
154	0	0	1	3
Total	1 (0.17)	0 (0)	7 (1.17)	18 (3)

Table 3. Zone B: in this case, both the robots forgive 2 zones, but the blue loses less because it has trajectories post-interaction that cross a zone.

Trajectory	Gain Blue	Gain Red	Lose Blue	Lose Red
56	0	0	1	3
71	1	0	1	3
112	0	0	1	3
135	0	0	2	3
148	0	0	1	3
154	0	0	1	3
Total	1 (0.17)	0 (0)	7 (1.17)	18 (3)

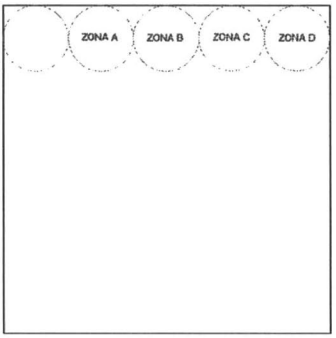


Figure 12. Subdivision in zones of interest.

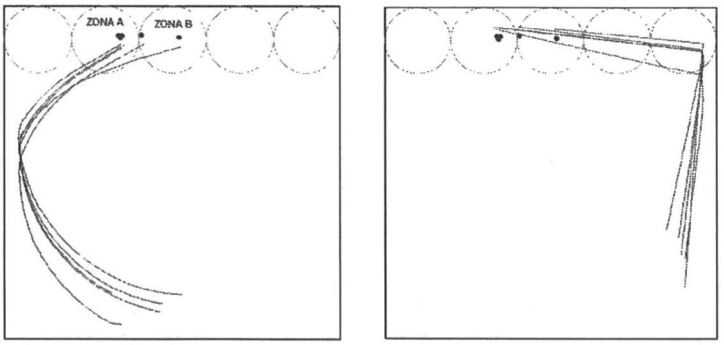


Figure 13. Interactive effects provoked from an agglomeration of points in Zone A.

This whole of points (fig. 13) is the only one, of those disposed on the side zones, that doesn't provoke a loss (and not even a profit) for the red robot: we see in fact that despite the interactions, the resultant trajectory crosses all the zones. In this case the blue loses a zone.

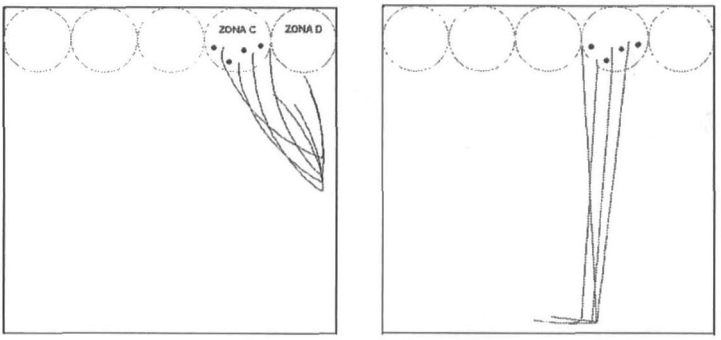


Figure 14. Interactive effects provoked from an agglomeration of points in C Zone.

Table 4. Zone C.

Trajectory	Gain Blue	Gain Red	Lose Blue	Lose Red
96	0	0	3	1
130	0	0	3	1
140	0	0	3	1
174	0	0	3	1
Total	0 (0)	0 (0)	12 (3)	4 (1)

The blue robot, in the interaction, curtains to lose many more zones than the red but his sense of run (counter clockwise) and the typology of the trajectory of exit (curvilinear) allows to him a fast recovery. A profit deriving from the re-enter in the C Zones and D zone too. The red instead, also losing less, comes hijacked on the inferior side of the environment and it doesn't ever cross the D Zone (fig. 15).

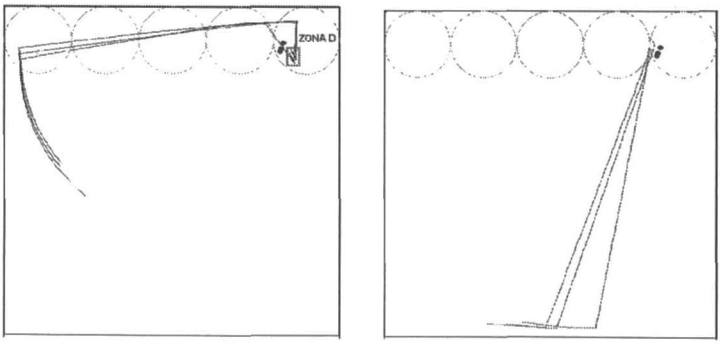


Figure 15. Interactive effects provoked from an agglomeration of points in D Zone.

The interactions modify the angle of shift of the blue robot that succeeds slightly to cross without problems all the 5 zones. The red is as usual inclined toward the lower part, but in this time does not lose the D Zone.

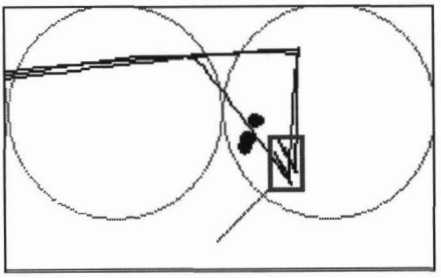


Figure 16. Reverse movement.

An interesting appearance (fig. 16) has given from the formality of reaction of the blue robot to the interaction: an inversion of run from frontal to reverse that remain few time steps, just for the time for make divert and remove the red. Then the blue takes back the direction of frontal run and begins to cross the zones.

4.1.3 Accumulations along the right wall

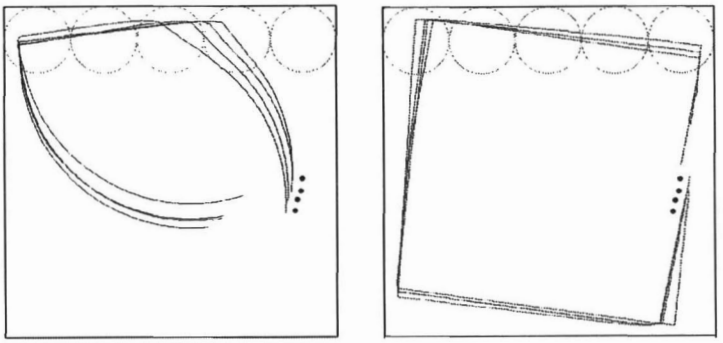


Figure 17. Trajectories post-interaction in the 2 robot (accumulations along the right wall).

In figure 17 the blue robot suffers a variation of bending that carries it to lose the first zone aloft to the right; also the red has a small variation of the angle of shift but this don't prevent him from reaching and cross all the 5 zones.

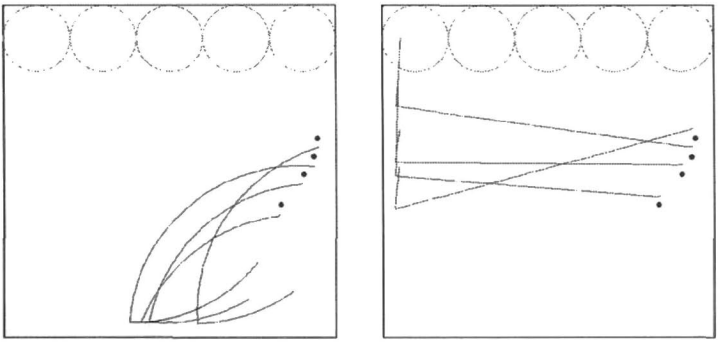


Figure 18. Trajectories post-interaction in the 2 robot (accumulations along the right wall).

In figure 18 the interaction pushes the blue robot in direction of the inferior wall of the environment, so it lose the 5 zones (the robot will reach them in the following time steps); the red has some benefited, in fact reaches and cross all the zones in briefer once as regards what it would get if it doesn't suffer interaction and remain coherent to his emergent pattern.

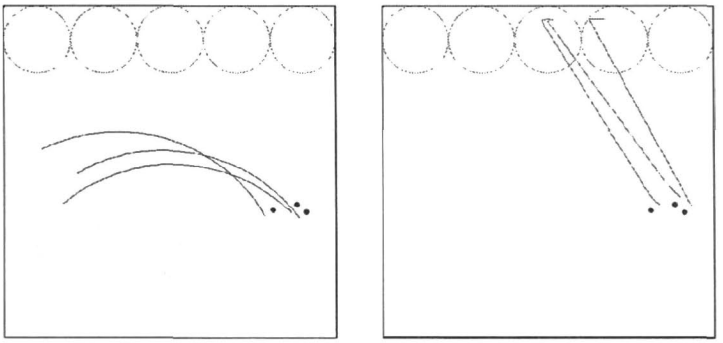


Figure 19. Trajectories post-interaction in the 2 robot; (accumulations along the right wall).

Also in this case to the blue is denied the possibility of reach the zones; the red instead "return back" and earn from 2 to 3 zones. The accumulations that are on the right side tend to not promote the blue, but how we have seen the more substantial accumulation is on the left side and along the zones and in these the red curtains to lose in manner rather marked.

4.1.4 Accumulations along the inferior wall

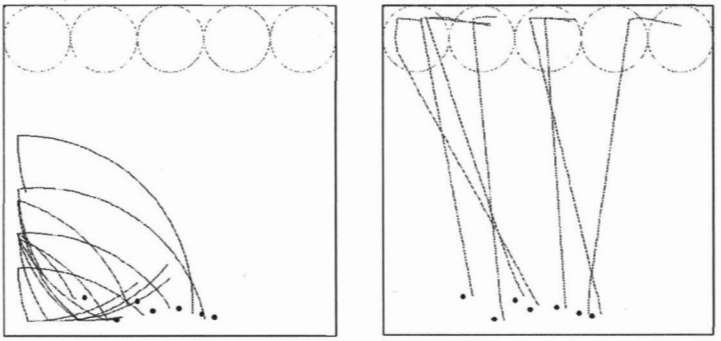


Figure 20. Trajectories post-interaction in the 2 robot (accumulations along the inferior wall).

We have seen that the most substantial accumulations of points are in correspondence of the intersections between emergent patterns of the 2 robots, this should not surprise us, and in fact in these points the probability of interaction is higher.

In the actual case (fig. 20), the points that are on the inferior wall of the environment, do not origin from this phenomenon. They derive mainly from two factors: interactions that have been verified before the blue robot has begun to follow their own emergent pattern and variations of bending in the trajectory of the blue robot caused from preceding interactions that they have pushed it in this zone (multiple interactions).

4.2 Experiment 024-R2

In this range of trajectories they emerge patterns that have the characteristic of present both pulls curvilinear and direction of frontal run, the only difference is the sense of run: counter clockwise for the red robot (fig. 21) and counter clock for the blue one (fig. 22).

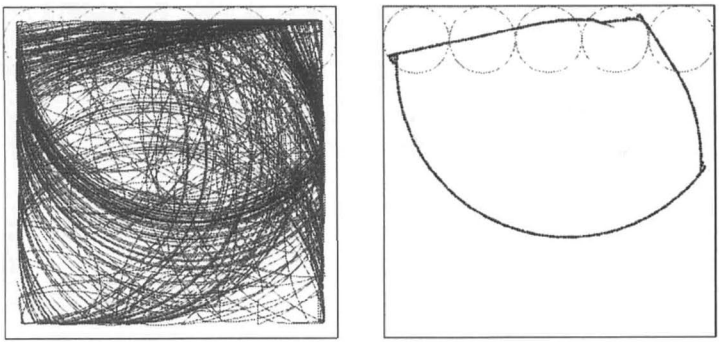


Figure 21. Superimposed trajectories and spatial emergent pattern (red robot).

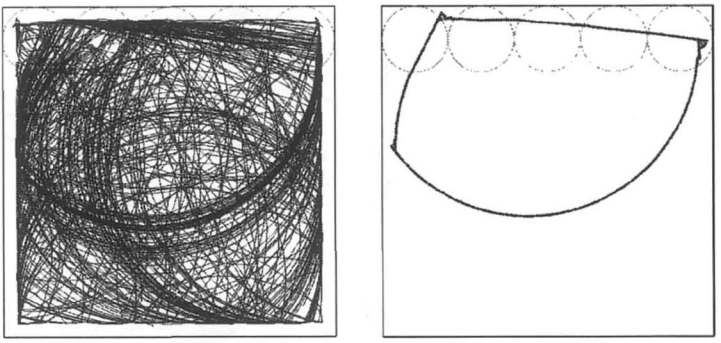


Figure 22. Superimposed Trajectories and spatial emergent pattern (blue robot).

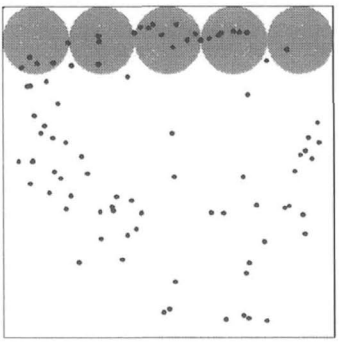


Figure 23. Spatial Disposition of the points of interaction Exp024-R2.

The structure of the emergent pattern (fig. 23) have not arranged, the points of interaction, along the contour of the environment but have spread into his inside. The accumulations that have caused more interesting results

in terms of fitness are those presents along the zones, in the left and right areas of the environment.

4.2.1 Accumulations along the zones

The trajectories of exit of the red robot tend to create a crushed version of his dominant pattern: this is surely a positive appearance in as it use less time steps for the crossing of the 5 zones and the part in excess could be used for following re-crosses (fig. 24).

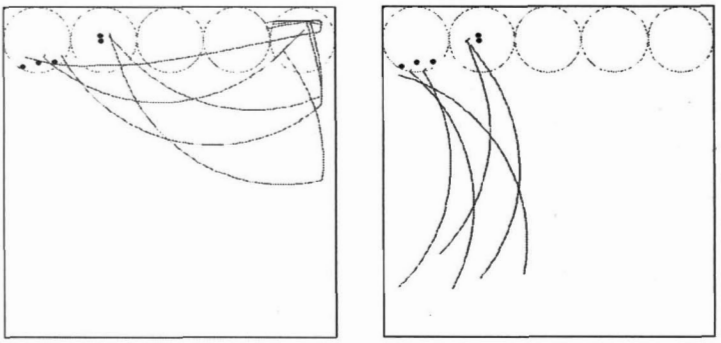


Figure 24. Partial Trajectories post-interaction: side zones.

The blue suffers damage instead, we see in fact that it meet just a pair of zones then it had an inversion of tendency that carries his trajectories in opposed direction to the zones.

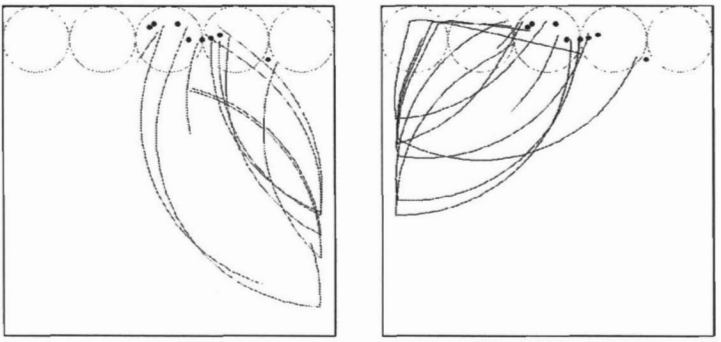


Figure 25. Partial Trajectories post-interaction: side zones.

Table 5.

Trajectory	Gain Blue	Gain Red	Lose Blue	Lose Red
259	0	0	1	4
266	1	0	1	3
283	1	0	2	2
303	0	0	2	3
310	0	0	2	2
349	1	0	2	2
351	0	0	1	3
369	0	0	2	2
Total	3 (0.37)	0 (0)	13 (1.62)	21 (2.62)

The red robot hears again in manner rather evident of the effects of the interaction (fig. 25). It does not have trajectories post-interaction that they cross any zone; the blue presents trajectories more squeezing than the red that allow him to recover the missing crossing of zone in small time.

4.2.2 Accumulations along left area of the environment

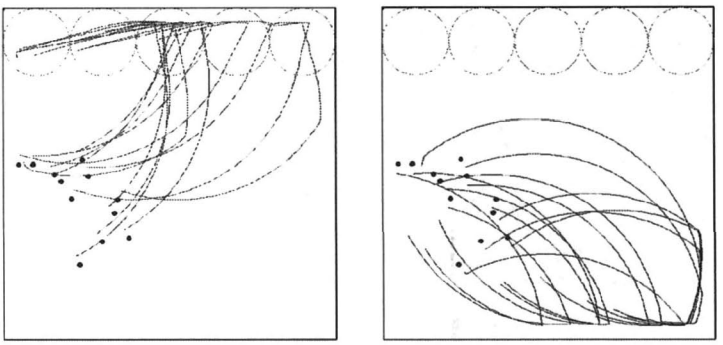


Figure 26. Partial trajectories post-interaction: area left environment.

In this case (fig. 26), also if the interactions are numerous, it is not difficulty appraise the effects that these has on the behaviour of the 2 robot. We can see clearly like the blue robot has a damage losing (temporarily) the possibility to cross the 5 zones. The red is instead "re-inserted" in the zones from which it originated (in accord with his emergent pattern and the direction of run) having an advantage.

4.2.3 Accumulations along the right wall.

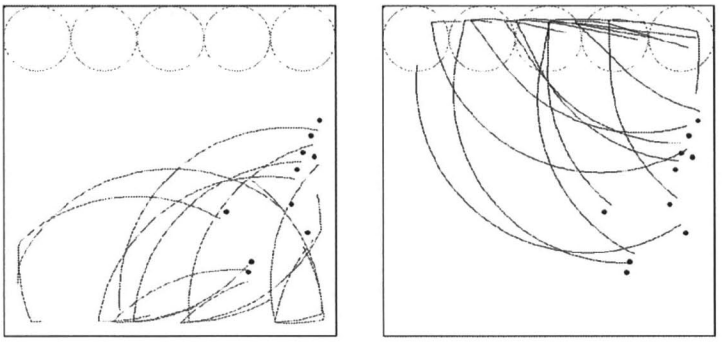


Figure 27. Partial Trajectories post-interaction: right area of the environment.

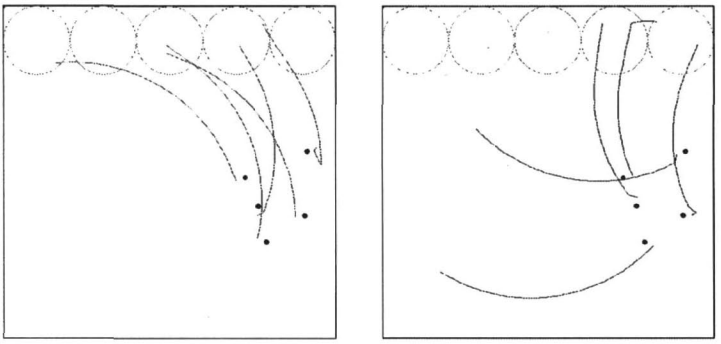


Figure 28. Partial Trajectories post-interaction: right area of the environment.

The interactions in that area (Fig. 27) tend to not promote the red robot, by pushing it in direction of the inferior wall of the environment; the blue, in the same way as for the red, is introduced in the zones from which it originated, getting an increase of his fitness and in conclusion an advantage. The loss suffered from the red is limited from any trajectories of escape in direction of the zones (Fig. 28).

5. RESULTS

In the course of the experiment 024 we have studied the effects that the interaction causes in the behaviour of the 2 robot, particularly we have shown like this could influence the value of the fitness. By comparing the results emerged in the 2 range of analyzed trajectories, we could affirm that,

while in the first case (rectangular, semicircle- range 1) the looses tend to be located mainly towards a single robot, in the other (semicircle, semicircle-range 2) a removal is gotten less evident between gains and looses. Like ulterior confirmation than above exposed, we have shown graphically the trend of the fitness in the two ranges (R1 and R2) above analyzed:

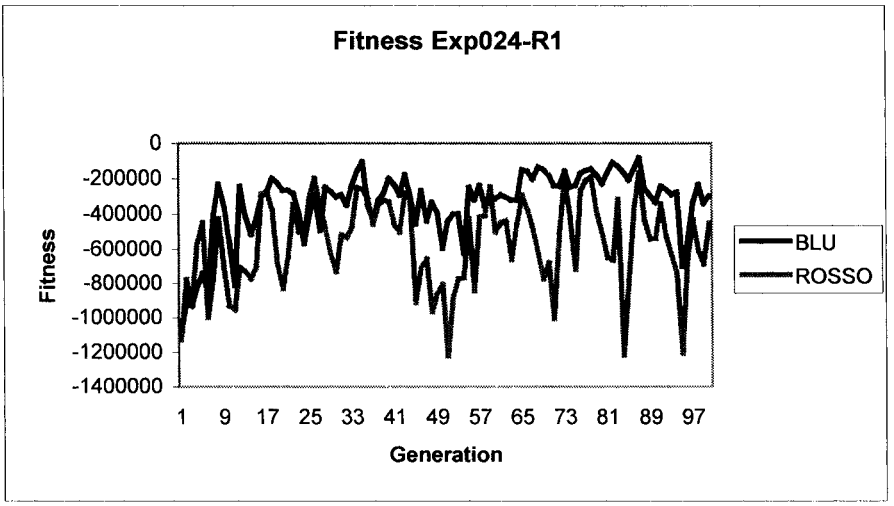


Figure 29. Trend of the fitness in the range of trajectories (R1) that has originated a curvilinear pattern for the blue robot and rectilinear for the red one.

From this chart (Fig. 29) we can note like already from the first generations, the red robot, and curtain to have a value of fitness inferior as regards that of the blue. Also when we have decrements for both the robots, the red robot is always that one that loses more.

The analysis developed in this range of trajectories (R1) let emerge a main point appearance deriving from the interaction between robot that has something in analogy with the competitive and dominance behaviours that are in nature between the animals. They gather in groups following particular geometries, in which the stronger elements (dominant) are arranged to the centre and those subordinates in periphery.

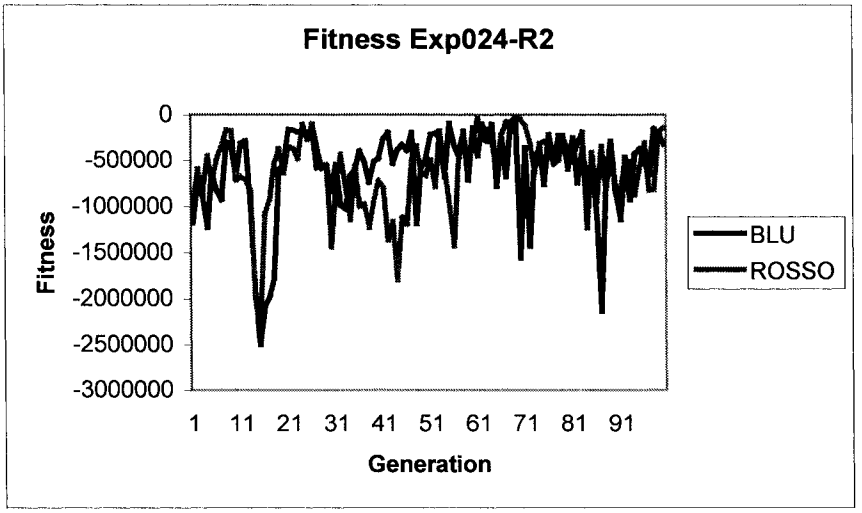


Figure 30. Trend of the fitness in the range of trajectories (R2) that have curvilinear pattern for both the robots.

The analysis turn for the range R2 find confirmation in this graph, where we can note like the gains and the losses are distributed in manner more uniform as regards the preceding case (range R1). This result seems to depend from the fact that the two agents are of similar rank; in fact they have pattern of trajectories symmetrical in mirror.

6. CONCLUSIONS

In this job we have developed experiments of genetic evolution of individuals to the inside of an open environment (without obstacles) in which the agents must complete a specific assignment: crossing of given zones of the environment. This is a typical assignment in the case, for instance, of supplying resources. In this situation it is reasonable that more agents interact in a conflictual way and they spring behaviours of cooperative-competitive type.

From the evolution of the individuals are sprung two typology of pattern of trajectories to move to the inside of the environment trying to optimize the fitness function, that is to cross the more possible zones. The two patterns are: a) semicircle and b)a kind of wall-following in which the pattern is rectangular, similar to the structure of the environment.

From an evolutionary point of view, we can deduce that a spatial pattern, like a semicircle, is more favourable for enhancing the fitness function. This

pattern time allows cutting the run. Particularly this pattern is more complex conceptually in as requires a better spatial representation of the environment that is missing in a pattern like wall-following. We could tell that the pattern like a semicircle is evolutionarily more advanced.

Another concept that is analysed in this work, concerns the environmental areas in which they are gathered the interactions. We could say that the points of interaction depend tightly from the typology of emergent pattern of the single agent: in fact the overlap of pattern determines the probability of interaction. We can also adfirm that the relationship between the areas of interaction and the emergent pattern is biunivocal, in fact the areas determine a profit or a loss depending on the typology of pattern (curvilinear or linear) and on the toward of the run (counter clock or counter clockwise). Therefore the areas are an additive element that determines an increase of the positive sides of a pattern and the diminution of those negative, like pits a kind of lens of enlargement.

REFERENCES

Broad, C. D., 1925, *The Mind and Its Place in Nature*, 1st edition, Routledge & Kegan, London.

Brooks, R. A., 1992, Artificial life and real robots, in: *Towards a Practice of Autonomous Systems: Proceedings of The First European Conference on Artificial Life*, MIT Press/Bradford Books.

Hamilton, W. D., 1971, Geometry for the selfish herd, *J. Theor. Biol.* **31**:295-311.

Hemelrijk, C. K., 1999, An individual-oriented model on the emergence of despotic and egalitarian societies, *Proceedings of the Royal Society London B: Biological Sciences* **266**:361-369.

Hemelrijk, C. K., 2000, Social phenomena emerging by self-organization in a competitive, virtual world ("Dom World"), in: *Learning to behave, Workshop II: Internalising Knowledge*, K. Jokinen, D. Heylen and A. Nijholt, eds.,. Ieper, Belgium, July, 2000, Venice, Italy, pp. 11-19.

Krink, T. and Vollrath, F., 1998, Emergent properties in the behaviour of a virtual spider robot, *Proc. Royal Society London* **265**:2051-2055.

Minati, G., 1996, *Introduzione alla Sistemica*, Edizioni Kappa, Roma.

Morgan, C. L., 1923, *Emergent Evolution*, Williams and Norgate, London.

Schelling, T. C., 1969, Models of Segregation, *American Economic Review, Papers and Proceedings* **59**(2):488-493.

Vehrencamp, S. L., 1983, A model for the evolution of despotic versus egalitarian societies, *Anim. Behav.* **31**:667-682.

von Bertalanffy, L., 1969, *General Systems Theory*, George Braziller, New York.

OVERCOMING COMPUTATIONALISM IN COGNITIVE SCIENCE

Maria Pietronilla Penna
Dipartimento di Psicologia, Università di Cagliari
Via Is Mirrionis, 09100 Cagliari, Italy

Abstract: This paper analyzes the role of computationalism in Cognitive Science in order to highligth its shortcomings. The main thesis is that, rather than eliminating computationalism from Cognitive Science, we would better reconsider the distinction between computable and uncomputable. Whereas such a distinction is useful to stress the intrinsic limitations of a mechanistic view of cognitive processing, it is useless when dealing with the main problem of post-computational Cognitive Science, namely the one of understanding the emergence of cognitive abilities from biological stuff.

Key words: computazionalism; decomposition method; cognitive science; emergence;

1. INTRODUCTION

Since the beginning of Cognitive revolution, the computationalist attitude dominated the development of Cognitive Psychology and Artificial Intelligence. The essence of the computationalist stance is synthetically expressed by the Physical Symbol System Hypothesis of Newell and Simon (see, e.g., Newell and Simon, 1976):

1. cognitive abilities and whence, in a broad sense, "intelligence" are possible only in presence of a symbolic representation of events and situations, external as well as internal, and of the ability to manipulate the symbols constituting the representations themselves;
2. all cognitive systems share a common set of basic symbol processing abilities;
3. every model of a given cognitive processing can always be cast in the form of a program, written in a suitable symbolic language, which, once

implemented on a computer, produces exactly the same behavior observed in the human beings in which we suppose the same cognitive processing be acting.

It is to be remarked that 1), 2) and 3) imply that the computer, on which every model of cognitive processing can be implemented , must be a digital computer, as an analogic computer is not suited to manipulate discretized symbols (even if it could perform such a task in an approximate way, provided suitable conventions would be adopted). A second remark is that computationalism implies that 1), 2) and 3) characterize every kind of cognitive processing and not only the one of a scientist resorting to symbol manipulation to draw a conclusion from a computational model of cognitive processing. Then, if this scientist uses a computational model of a non-symbolic cognitive processing (such models, for instance, are very common in Physics or in Computational Neuroscience), he/she cannot be considered as adhering to a computationalist view.

The long history of Cognitive Psychology and of Artificial Intelligence, as well as of Philosophy of Mind, evidenced how the adoption of a computationalist stance entails a number of advantages (see Fodor, 1975; Pylyshyn, 1984), such as:

a) avoiding a number of philosophical problems, such as the mind-brain dichotomy and the intervention of the *homunculus*;
b) stimulating the introduction of a number of computational models of cognitive activity, suited to be implemented on a computer and tested in laboratory experiments.

We could therefore say that the diffusion of computationalist view is directly responsible for the development of a scientific Cognitive Psychology out of the fog generated by nineteenth century philosophical discussions. The strange fact, however, is that this happened without *any experimental proof* of the computational nature of mental processing. Of course, we cannot deny that most high-level cognitive behaviors seem intuitively to be endowed with a genuine computational nature, for instance when we perform a mathematical operation, a logical inference, or we use concepts and language. However for other kinds of cognitive processing, such as perceptual ones, such an intuitive evidence is somewhat lacking. Besides, many experimental studies by cognitive psychologists have shown that the symbol manipulation program underlying concept manipulations by human beings is difficult to discover and that, provided it exists, its form would definitely be very different from the one expected on the basis of usual symbol manipulation rules employed in logic, mathematics or computer science.

Notwithstanding these troubles, we could continue to adopt the computationalist view, if it were able, at least in principle, to account for all observed features of human behaviors (eventually limiting ourselves only to "cognitive" ones), as they concretely appear both in laboratory experiments and in everyday life. However this paper will argue that the computationalist view is *in principle* unable to meet this requirement. This circumstance raises the issue of finding a viable alternative to computationalism. In this connection, we will try to sketch a possible way to continue doing scientific research in Psychology without adopting the computationalist view but, at the same time, without giving up the advantages of the latter.

2. UNSOLVABLE PROBLEMS FOR THE COMPUTATIONALIST APPROACH

There are three problems which, in principle, the computationalist approach cannot solve. They can be named as *complexity problem*, *implementation problem*, and *decomposition problem*. We will give in the following a detailed description of each one of them.

2.1 The complexity problem

The *complexity problem* stems from the fact that observed cognitive behaviors appear so complex as to make it impossible to describe (and reproduce) them through a computer program written in any standard programming language, and implemented on a digital computer. Here the attribute "complex" can have many different meanings, according to one of the different definitions of *complexity* proposed so far; notwithstanding the one we choose, it will nevertheless apply to cognitive behaviors. For instance, if the complexity is defined as the number of components (that is, of information elements) underlying a concrete every day cognitive behaviors, then such behaviors can surely be qualified as "complex". In this regard, evidence is provided by the failure, already apparent in the Sixties, of the methods based on "general principles" in Artificial Intelligence, the best representative case being the General Problem Solver of Newell and Simon. At that time it was recognized how the cognitive ability shown by human beings in solving different kinds of problems was impossible to reproduce by resorting only to programs based on general heuristics and a sufficient amount of stored information, mainly because every specific knowledge domain was characterized by specific rules and specific heuristics. Even with the advent of expert systems, every effort to build a system of this kind,

endowed with an amount of knowledge, in many different domains, comparable to the one currently utilized by the average human being, ran into failure. The most typical case was CYC (Lenat and Guha, 1990). This research program was intended to build an expert system endowed with the whole knowledge of an average subject, living in a Western country, with a regular high school education. After more than fifteen years of research, CYC has been only able to deal with very limited portions of such knowledge, whereas the most part of it is still unexplored.

There is, however, some reason to conjecture that cognitive behavior be "complex" even in a computational sense. Let us, consider, for instance, the linguistic productions of a human subject. As it is well known, they rely on the knowledge of a suitable lexicon, which, for an average subject, can contain some 40,000-50,000 words, or more. Let us now estimate the possible sentences that the subject can build by resorting to this lexicon; to simplify our computation, let us suppose that the maximum sentence length be finite, and equal to k words. Besides, let us denote by N the number of words known by this individual (the extension of his/her lexicon). As for every word in each sentence there are N different possibilities, the total number of possible different sentences is $S = N^k$ (in practical cases this is a huge number: with $N = 50{,}000$ and $k = 5$ we have that S is close to 3×10^{23}, about half the number of molecules contained in a gas mole). An obvious objection to this estimate is that it overlooks that the production of correct sentences is based on a finite number of grammatical rules; therefore, a computer program that simulates the linguistic competence of a human subject could correctly work only on the basis of these rules. However, this objection is not correct, for two reasons. First, observed linguistic behavior doesn't include only correct sentences (therefore, the set of all possible sentences should be taken into account) ; second, knowledge of grammatical rules cannot fully explain why, in speaking or writing, we choose certain words rather than others. We can, now, characterize each possible kind of linguistic behavior through a suitable Boolean function on the set of all possible different sentences; this function returns value 1 for the sentences that can be produced by the particular linguistic behavior under consideration, and value zero for the sentences that cannot be produced. It is immediate to see that the total number of different Boolean functions of this kind is $B = 2^S$. This number is thus identical to the number of the different possible linguistic behaviors, and we can conjecture that a symbolic program simulating the human linguistic behavior be endowed with a number of instructions and operations which, roughly, scales with B. According to standard theory of computational complexity, we could say that such a program solves a NP-hard problem, as B depends upon N, the number of information elements, in an exponential way. Of course, these rough

arguments cannot pretend to give a proof of the fact that the simulation of the human linguistic behavior is a NP-hard problem, but we recall that, if this were the case, every hope of simulating this behavior through a computer program would be vain (see, for a review, Hochbaum, 1997).

Moreover, even if things were arranged in such a way that the simulation of human cognitive behavior would, in principle, be feasible by resorting to a computer program of reasonable complexity, then we would still face the impossibility of discovering the operations performed by such a program through experiments on human subjects. For, in such experiments, it is impossible to know what other independent variables, besides the ones chosen in an explicit way by the experimenter, could influence the observed performance. And such a circumstance makes it impossible to detect the existence of input-output relationships in a reliable way.

Of course, these shortcomings affect only computer programs simulating digital-like operations, the only ones supposed to be performed by human cognitive systems, according to the Physical Symbol System Hypothesis. If we allowed the introduction of continuous variables (like the ones managed by analog computers), noise and fluctuations, then the problems associated with complexity could be solved in a simpler way. In this case, however, usual computer programs could no more constitute a complete simulation of human cognitive operations, as they could only mimic in an incomplete way operations which cannot, in principle, be fully represented within a digital computer. Therefore the adoption of a non-computationalist view would imply that simulations implemented on a digital computer would be useless without the existence of a mathematical theory (the only context in which the previous concepts would have a well defined meaning), and that their role would mostly be the one of suggesting new insights for theory builders, as well as partially testing theoretical conclusions.

2.2 The implementation problem

The implementation problem is a many-faceted one. Roughly, we can say that it consists in the fact that the computationalist approach is unable to explain:

- the motivations underlying the choice of a given computational process and of the ways to implement it;
- the strong context-dependence of most cognitive processes.

In this regard, we remark that a complete Computation Theory should deal with three main aspects of computation processes: the *syntactic* one, related to the effects produced by computation rules on the symbol sequences to which they apply; the *semantic* one, related to goals underlying

computations and to the meanings attributed to symbol sequences; and the *pragmatic* one, related to the effects produced by a computation process on the surrounding environment (see, in this regard, Sloman, 2002). Unfortunately, actual Computation Theory, based on the fundamental works by Alan Turing, deals only with syntax, and so far little work has been done on the other two aspects. It is then obvious that such a circumstance makes Computation Theory useless for Cognitive Psychologists or, more in general, for Cognitive Scientists. As a matter of fact, no theoretical or experimental study on cognitive processing ever made resort to such a theory.

The lack of knowledge about semantic and pragmatic aspects of computation implies, in turn, a complete ignorance about the role played by the implementational setting of computation itself. Such a setting, of course, is related to the context in which computation takes place. The fact that most cognitive processes are strongly context-dependent is well known from long time and it doesn't need further exemplification. We will only limit ourselves to remark that even most symbol-manipulation devices currently used by human beings operate in a context-dependent way. Again, the typical example is given by language. As every schoolchild knows, most grammatical rules used in producing correct linguistic expressions (in whatever human language) are context-dependent. Thus, an exhaustive symbolic description of language production should necessarily include also a symbolic description of all possible contexts. Unfortunately, as we showed in the previous section, the latter is very difficult, if not impossible, to obtain. This is also shown by the difficulties encountered by children in learning their native language, or by adults in learning a new language. In all these cases, the difficulties stem not so much from rule understanding but, rather, from context-dependence understanding. Thus, in the light of these considerations, it is not surprising that this domain of Cognitive Science be characterized by a wide resort to models of a non-computational kind, such as connectionist models of language learning and use (see, e.g., MacWhinney, 1999). The unavoidable conclusion is that a strict computationalist approach cannot solve, in principle, the implementation problem.

2.3 The decomposition problem

The decomposition problem is a direct consequence of the identification of cognitive processes with symbolic manipulations performed by suitable programs running on digital computers. Any such program (except very simple cases) can be thought to be composed by a number of different subprograms, or *modules*, each one highly specialized in performing a given function. Moreover, these modules should be embedded within a hierarchical

architecture, which allows for interconnections between the modules themselves, both of a horizontal and of a vertical kind (Fodor, 1983). Therefore, if the Physical Symbol System Hypothesis holds, we should be able to detect the presence of these modules. The problem stems from the fact that experimental and theoretical studies on cognitive systems, and on the biological brain as well, were unable to provide convincing evidence for their existence.

We recall, in this regard, that the only tool available so far for detecting these modules is the so-called *method of dissociations*, also named, in other contexts, as *method of decompositions* (see Bechtel and Richardson, 1993). It can be used in two different ways:

- *top-down*, when applied to cognitive processes, as observed at a macroscopic level in psychological experiments;
- *bottom-up*, when applied to experimental and clinical data coming from Neuroscience.

The top-down version of decomposition method includes, in turn, two different sub-cases:

d.1) the model-based decomposition;

d.2) the task decomposition.

Method d.1) has been widely used within Artificial Intelligence. It consists in a decomposition of the procedure used to perform a given cognitive task, on the basis of purely *logical* arguments, into smaller sub-procedures, up to a level in which each sub-procedure can be implemented in an easy and intuitive way by a single module. This method permits the construction of efficient software programs, which are able to simulate the performance of the cognitive task under study. Unfortunately, even when the method works, it by no means grants that the human cognitive system be working in the same way as the computer program simulating its performance. In this connection, we recall that a number of arguments (the most celebrated one is Searle's Chinese room) shed light on the fact that, even if a computer program meets the requirements of the Turing Test with respect to a given cognitive task, the cognitive processes of a human being performing the same task might be utterly different (on Searle's argument there is a wide bibliography; see for reviews Hauser, 1997; Searle, 1999; Preston and Bishop, 2002).

As regards method d.2), it was proposed by Tulving (see Tulving, 1983), and it consists in considering pairs of different cognitive tasks, which share a given independent variable. If the manipulation of this variable produces an effect on the performance in one of the two tasks, but not in the other, then this is considered to be evidence for the fact that the two tasks are performed by two different modules. The problem with this method is that, even if we

assume its validity (which is far from being proved), it makes us detect modules whose nature is very different from the one we would expect on the basis of the Physical Symbol System Hypothesis. Namely, these modules are associated to functions of a very general kind, and they are linked by horizontal interconnections, whereas the modules expected within the computationalist approach should be highly specific, and mostly linked by vertical interconnections. The most celebrated example of modules detected by the application of method d.2) is given by Tulving distinction between episodic and semantic memory; but this very example illustrates how such general-purpose memory systems cannot be considered to provide compelling evidence for the existence of a modular architecture of the mind.

Let us now shortly discuss the bottom-up version of the decomposition method, usually named *method of dissociations* (see, for instance, Gazzaniga, 2004). This method can only be applied to subjects with a severe impairment in one specific cognitive ability. If impairment is associated to a lesion in a specific brain area, then we can identify this area as the seat of the module responsible for the ability under consideration. The method of dissociation, in association with different kinds of brain imaging techniques, recently led to singling out a number of brain modules devoted to specific cognitive tasks such as, for instance, face recognition, number manipulation, language understanding, reading. Even this method, however, is plagued by a number of shortcomings, such as:

s.1) the number of subjects to which the method can be applied is very small; namely it is very rare that a subject be characterized by the impairment of only one specific cognitive ability, or by a single localized brain lesion; most often the impairment concerns several different abilities, and the lesion is diffused instead of localized; in addition, the small number of available subjects implies the lack of statistical validity for the results obtained by this method;

s.2) the number of the modules which can be detected, if the method is applied without further restrictions, is very high; this appears to be somewhat unrealistic because, if the brain were characterized by such a complex modular architecture, it would be highly sensitive to noise, errors, and disturbances, as it is the case for a digital computer;

s.3) the application of the method itself could be biased by the belief in the existence of a modular cognitive architecture.

The foregoing discussion thus shows that none of the methods described is able to provide a reliable solution of the decomposition problem. Moreover, if we also consider the arguments of the previous sections (sec. 2.1 and 2.2), we reach a further general conclusion: within the computationalist view, it is in principle impossible to solve the problems of (i) complexity, (ii) implementation, and (iii) decomposition; therefore, we

are forced to abandon the computationalist approach (at least in its original form), and to resort to new approaches; if any, in order to deal with human cognitive processing in a more realistic way. The fundamental question is then: On what basis, different from computationalism, a new approach to the study of cognitive processing could be grounded?

3. BEYOND COMPUTATIONALISM

To start with, we recall that several alternatives to computationalism have already been proposed. Among them:

- the *connectionist* approach (McClelland and Rumelhart, 1986);
- the *dynamicist* approach (Port and Van Gelder, 1995; Van Gelder, 1998);
- the *embodied cognition* approach (Varela et al., 1991);
- the approach based on *Artificial Life* (Langton, 1989).

There are considerable overlaps between these different approaches, as well as a number of differences. Some of the common features are:

- the attention to biological implementation of cognitive processing;
- the attempt to describe cognitive processing within the context of organism-environment interaction;
- the use of analog computation (based on continuous quantities) and, sometimes, of noise or fluctuations;
- the view according to which cognitive abilities emerge from the interactions between microscopic units.

The presence of the last feature shows that, to displace computationalism, we need to first introduce different *levels of description* of cognitive processing, for instance the microscopic and the macroscopic one, so as to ground the observed features of cognitive systems on the *relationship* between these levels, rather than on the "flat" world of symbolic computation, where all symbols are in principle on the same foot. Such a view agrees with the one longly upheld by Physics, according to which the macroscopic features of physical phenomena are described by Thermodynamics, but they are then explained in terms of microscopic ones, which are related to behaviors of atoms and elementary particles. However, in Physics the relationship between microscopic and macroscopic features can be described in a rigorous way by Statistical Mechanics, while in Cognitive Science such a relationship is still largely unexplored, due to the fact that a "Statistical Mechanics of Cognitive Behaviors does not exist yet, despite claims to the contrary from the connectionist camp.

The main problem for every approach that aims to a radical displacement of computationalism consists in constructing a satisfactory theory of emergence of cognitive processes from interactions of suitable "microcognitive units". Such a theory is needed in order to account for a number of experimental facts, such as:

f.1) the existence of *long-range correlations* (both of temporal and spatial nature) and *large-scale coherence* of electroencephalographic signals (see, e.g., Freeman, 2000; Nunez, 2000), which is evidence for integration between different cognitive (Classen et al., 1998; Sarnthein et al., 1998), as well as affective (Hinrichs and Machleidt, 1992; Nielsen and Chénier, 1999) processes, in mental activity;

f.2) the existence of *long-range correlations* between the activities of different neuronal groups, sometimes interpreted as evidence for a synchronization of neuronal activities (see Rodriguez et al., 1999; see also the critique by Van der Velde and De Kamps, 2002);

f.3) the existence of (typically *middle-range*) correlations between different stimulation elements, shown by the celebrated Gestalt effects in visual perception, as well as by a number of other effects, which characterize visual attention and seem to favor global views on local ones;

f.4) the existence of a number of experimental effects, in psychology of language, learning and memory, which show that holistic features can influence local ones; sometimes these effects are interpreted as showing the importance of context in cognitive processes.

Even an outline of a theory of emergence is far beyond the limits of this paper. Thus, we just mention that a number of researchers tried to find an alternative to computationalism by resorting to a logical analysis of the models that allow for the existence of continuous quantities (such as models in Physics, neural networks introduced by connectionist people, and, more in general, all kinds of analog computation). The outcome of this analysis has been that most of these models describe processes which are not computable by a Turing machine. This produced a number of studies dealing with *hypercomputation* (Siegelmann, 1999; MacLennan, 2001; Stannett, 2001), claiming that a rejection of Turing-machine-based computationalism was the main recipe for building a new kind of Cognitive Science (see, e.g., Fetzer, 2001; Bringsjord and Zenzen, 2003). We stress here, however, that all computationalist models so far adopted within traditional Cognitive Psychology or Artificial Intelligence have never relied on considerations related to Turing-machine-computability. Thus, any consideration concerning this kind of computability seems useless, at least with respect to the problem at issue, i.e., the construction of a more realistic form of Cognitive Science. From the standpoint of a complete Computation Theory, the very notion of computation is related to the needs of the human subjects

who perform computations and, as such, it depends on the goals, knowledge, beliefs, and mental schemata, rather than being based on an absolute and objective standard. And it is well known that Turing's goal was mainly to show the intrinsic limitations of a specific notion of computation (a goal successfully reached) rather than to describe the notion of emergence in physical or biological systems (a theme to which Turing himself gave some contribution by employing, however, a totally different tool, i.e., differential equations). We can thus conclude that, rather wasting our time in the discussion of the pros and the cons of Turing-machine-computability, we would better engage in the concrete effort to build a theory of emergence of cognitive processes, which fulfill all the constraints set by the experimental findings mentioned above (see f.1 - f.4).

4. CONCLUSION

The previous discussion makes clear that any attempt to supersede computationalism should start with a theory of emergence of cognitive processes from interactions between microcognitive units and, ultimately, from features observed within biological stuff contained in biological brains. Constructing such a theory, however, will require a definite answer to a number of difficult questions, such as:

q.1) Is emergence of cognitive processes different from physical emergence commonly associated to *phase transitions*? Can the tools successfully employed by theoretical physics to deal with such phenomena also be used to describe cognitive or, even more generally, mental emergence?

q.2) How can we account for the fact that, most often, correlations displayed by psychological experiments are middle-range, not long-range correlations, which are typically displayed by physical phenomena?

q.3) How can we verify whether an alternative approach to computationalism is able to describe cognitive processes as effectively emergent, without being beset by the troubles of the complexity, implementation, and decomposition problems?

We feel that, to answer these questions, we need new conceptual and technical modeling tools. Only their introduction will likely ease the tremendous task of doing Cognitive Science while going beyond Good Old Fashioned Computationalist Approach (GOFCA).

REFERENCES

Bechtel, W., and Richardson, R., 1993, *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton Univ. Press, Princeton, NJ.

Bringsjord, S., and Zenzen, M., 2003, *Superminds: People Harness Hypercomputation and More*, Kluwer, Dordrecht.

Classen, J., Gerloff, C., Honda, M., and Hallet, M., 1998, Integrative visuomotor behavior is associated with interregionally coherent oscillations in the human brain, *Journal of Neurophysiology* **79**:1567-1573.

Fetzer, J. H., 2001, *Computers and Cognition: Why Minds are not Machines*, Kluwer, Dordrecht.

Fodor, J. A., 1975, *The Language of Thought*, Harvard University Press, Cambridge, MA.

Fodor, J. A., 1983, *Modularity of Mind*, MIT Press, Cambridge, MA.

Freeman, W. J., 2000, *Neurodynamics: an Exploration of Mesoscopic Brain Dynamics*, Springer, Berlin.

Gazzaniga, M.S., Ed., 2004, *The Cognitive Neurosciences III*, Third Edition, MIT Press, Cambridge, MA.

Hauser, L., 1997, Searle's Chinese box: debunking the Chinese room argument, *Minds and Machines* **7**:199-226.

Hinrichs, H., and Machleidt, W., 1992, Basic emotions reflected in EEG-coherence, *International Journal of Psychophysiology* **13**:225-232.

Hochbaum, D. S., Ed., 1997, *Approximation Algorithms for NP-Hard Problems*, PWS Publishing Company, Boston, MA.

Langton, C., 1989, *Artificial Life*, Addison-Wesley, Redwood City, CA.

Lenat, D., and Guha, R.,1990, *Building Large Knowledge Based Systems*, Addison-Wesley, Reading, MA.

MacLennan, B. J., 2001, *Transcending Turing Computability*, Technical Report UT-CS-01-473, Department of Computer Science, University of Tennessee, Knoxville, (Can be found in the website www.cs.utk.edu/~mclennan).

MacWhinney, B., Ed., 1999, *The Emergence of Language*, Lawrence Erlbaum, Mahwah, NJ.

McClelland, J. L., and Rumelhart, D.E., Eds., 1986, *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, MIT Press, Cambridge, MA.

Newell, A., and Simon, H. A., 1976, Computer science as empirical inquiry: symbols and search, *Communications of the ACM* **19**:113-126.

Nielsen, T. A., and Chénier, V., 1999, Variations in EEG coherence as an index of the affective content of dreams from REM sleep: Relationship with face imagery, *Brain and Cognition* **41**:200-212.

Nunez, P. L., 2000, Toward a quantitative description of large scale neocortical dynamic function and EEG, *Behavioral and Brain Sciences* **23**:371-437.

Port, R., and Van Gelder, T.J., Eds., 1995, *Mind as Motion: Explorations in the Dynamics of Cognition*, MIT Press, Cambridge, MA.

Preston, J., and Bishop, M., Eds., 2002, *Views into the Chinese Room: New essays on Searle and Artificial Intelligence*, Oxford University Press, Oxford, UK.

Pylyshyn, Z. W., 1984, *Computation and Cognition: Towards a Foundation for Cognitive Science*, MIT Press, Cambridge, MA.

Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., and Varela, F. J., 1999, Perception's shadow: Long-distance synchronization of human brain activity, *Nature* **397**:430-433.

Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., and Von Stein, A., 1998, Synchronization between prefrontal and posterior association cortex during human working memory, *Proceedings of the National Academy of Sciences USA* **95**:7092-7096.

Searle, J. R., 1999, *The Mystery of Consciousness*, A New York Review Book, New York.

Siegelmann, H. T., 1999, *Neural Networks and Analog Computation: Beyond the Turing limit*, Birkhäuser, Boston, MA.

Sloman, A., 2002, The irrelevance of Turing machines to AI, in: *Computationalism: New Directions*, M. Scheutz, Ed., MIT Press, Cambridge, MA, pp. 87-127.

Stannett, M., 2001, *An Introduction to Post-Newtonian and Non-Turing Computation*, Technical Report CS-91-02, Department of Computer Science, Sheffield University, Sheffield, UK.

Tulving, E., 1983, *Elements of Episodic Memory*, Oxford University Press, New York.

Van der Velde, F., and De Kamps, M., 2002, Synchrony in the eye of the beholder: An analysis of the role of neural synchronization in cognitive processes, *Brain and Mind* **3**:291-312.

Van Gelder, T. J., 1998, The dynamical hypothesis in cognitive science, *Behavioral and Brain Sciences* **21**:615-665.

Varela, F., Thompson, E., and Rosch, E., 1991, *The Embodied Mind: Cognitive Science and Human Experience*, MIT Press, Cambridge, MA.

PHYSICAL AND BIOLOGICAL EMERGENCE: ARE THEY DIFFERENT ?

Eliano Pessa
Dipartimento di Psicologia e Centro Interdipartimentale di Scienze Cognitive, Università di Pavia, Piazza Botta, 6 - 27100 Pavia, Italy

Abstract: In this paper we compare the features of models of emergence introduced within theoretical physics, mainly to account for phenomenology of second-order phase transitions, with the requirements coming from observations of biological self-organization. We argue that, notwithstanding the deep differences between biological and non-biological systems, the methods of theoretical physics could, in principle, account even for the main features of biological emergence.

Key words: emergence; reaction-diffusion systems; neural networks; quantum field theory; stochastic quantization.

1. INTRODUCTION

The last years were marked by a consistent growth of interest in emergence and self-organization, both from theoretical and experimental side. Such an interest, initially born within the context of Artificial Life models (which forced to introduce the concept of *emergent computation*), was further increased by the needs of nanotechnology, autonomous robotics, econophysics and other research domains. Despite the existence of different definitions of emergence (see Crutchfield, 1994; Bedau, 1997; Ronald et al., 1999; Rueger, 2000), most researchers agrees on characterizing the 'highest' and most interesting form of emergence (the one called *intrinsic emergence* by Crutchfield) in the following way:

- it occurs at a *macroscopic* level, that is at an observational level higher than the one commonly used to describe the single components of a given system;

- it consists un collective behaviors of these components, giving rise to the occurrence of macroscopic coherent entities;
- the specific features of these entities cannot be foreseen in advance, even if they are fully compatible with the models of the systems themselves; the latter, however, can state only what are the *potential* coherence phenomena, without a detailed prediction of *actual* ones;
- the occurrence of emergent coherence phenomena can modify the operation itself of the system under study.

These features, however, are too generic to be useful in building concrete models of emergence in specific systems. In this regard we remark that the only models of emergence endowed with these features and, at the same time, allowing for specific, experimentally testable, predictions have been introduced within theoretical physics to account for phase transition phenomena (mainly of second-order). These models have been highly successful in explaining relevant phenomena not only in condensed matter physics, but even in elementary particle physics, astrophysics, cosmology. This suggested to a number of physicists that they could be used to account for *whatever* kind of emergence, even in non-physical domains, such as biological, cognitive, social and economic ones. However the observational features associated to emergence in these domains (we will use the shortened expression *biological emergence* to label in a collective fashion these features) would seem, at first sight, to be deeply different from those characterizing *physical emergence* (as described by theories of second-order phase transitions). The latter circumstance raises the question of the difference between physical and biological emergence: can we resort to suitable generalizations of models describing physical emergence to account for observed features of biological emergence, or, on the contrary, to deal with the latter we need an entirely new approach, incompatible with the one adopted to model physical emergence?

This is a very difficult question, whose foremost importance for the development of science cannot be undervalued: answering it we will decide how to cope with the great challenge of next years, that is building a consistent theory of biological phenomena, ranging from viruses to brain cognitive operation. In this regard this paper will contain a number of arguments supporting the former of the two alternatives mentioned above. In other terms, we will suggest that the theoretical apparatus of phase transition theory could be able to account even for biological emergence, provided we generalize it in a suitable way.

2. THE INGREDIENTS FOR PHYSICAL EMERGENCE

Summarizing a large body of theories and mathematical models, it is possible to say that, in order to give rise to physical emergence (that is to models allowing for physical emergence) we need three different ingredients (see, in this regard, Haken, 1978, 1983, 1988; Mikhailov, 1990; Mikhailov and Loskutov, 1996):

- *bifurcation*;
- *spatial extent*;
- *fluctuations*.

We underline that physical emergence occurs if and only if *all three* ingredients are simultaneously present. Before going to our arguments we will spent some words on the meaning of each one of terms introduced.

a) Bifurcation

This name denotes a precise mathematical construct (see, for instance, Sattinger, 1978, 1980; Iooss and Joseph, 1981; Vanderbauwhede, 1982; Guckenheimer and Holmes, 1983; Glendinning, 1994). The latter is used within a context in which we model time evolution of suitable systems, in turn obeying suitable *evolution laws*, formalized through *evolution equations* (which can be of different kinds: differential equations, difference equations, recursion maps, and so on). Generally these equations are characterized by a number of *dependent* (or *state*) *variables*, by a number of *independent variables* (necessarily including time, or some substitute of it), and by a number of *parameters*. A bifurcation consists in a change of structural features of the solutions of these equations (describing system's *behaviors*) when a parameter crosses a *critical value*. There is a wide phenomenology of possible bifurcations and we will not insist further on the difficult underlying theory. However, we will underline that not all bifurcations are equally interesting for intrinsic emergence, but only the ones giving rise to a change (better, to an increase) in the number of possible equilibrium states of the system and, eventually, in their nature. Often these bifurcations are called *symmetry-breaking bifurcations*. When we speak of bifurcations, in practice we mean bifurcations of the latter kind.

b) Spatial extent

Effective models of intrinsic emergence describe always systems endowed with a spatial extension, that is having as independent variables, besides the temporal one, even spatial coordinates. Typically these systems are characterized by an *infinite* number of degrees of freedom (the possible values of dependent variables in each space location) and such a circumstance makes them endowed with a richness of possible

behaviors which cannot be emulated by systems with a finite number of degrees of freedom (such as the ones described by ordinary differential equations).

c) Fluctuations

Without fluctuations (whatever be their nature) we have only order, determinism, predictability. Therefore intrinsic emergence is not allowed. When introducing fluctuations we have a number of different possibilities. They can due, for instance, to *intrinsic noise*, or to the fact that the processes are of *stochastic* nature, as well as to the existence of *quantum uncertainty* or of *quantum zero-point fluctuations*. It can be shown that all these kinds of fluctuations give rise, in a way or in another, to some phenomenon of intrinsic emergence, provided the other two ingredients mentioned above be present.

There is a number of arguments supporting the claim that, if even only one of these ingredients is lacking, intrinsic emergence cannot occur. We will not review them here, by referring the reader to the existing literature (see, for instance, Nitzan and Ortoleva, 1980; Stein, 1980; Fernández, 1985; Sewell, 1986, 2002; Scott, 2003). Instead, we will remark that this doesn't imply that models containing only one or two of the ingredients above be devoid of interest: namely they have been very useful, showing that mechanisms such as bifurcation can explain *pattern formation* without the need for specific design rules (this is usually called *self-organization*). The most celebrated case is given by Prigogine's theory of Dissipative Structures (Nicolis and Prigogine, 1977), which are excellent models of self-organization, relying both on bifurcation and spatial extent (see, for instance, Belintsev, 1983; Beloussov, 1998; Mori and Kuramoto, 2001). Of course, introducing in them fluctuations, they could be used even to model intrinsic emergence.

3. CONTROLLING PHYSICAL EMERGENCE

In principle the word "control", when applied to intrinsic emergence, refers to a set of action potentialities relying on three different circumstances:

- the existence of a *theoretical model* of phenomena under consideration allowing to state what are the conditions granting for the occurrence of intrinsic emergence and the possible features of emergent behaviors;
- the existence of a set of *prescriptions* (possibly ensuing from the model above), stating what interventions (that is actions to be performed) to make on the system under study in order to give rise to intrinsic emergence;

- the existence of a set of *measurable quantities*, whose values let us detect the occurrence of intrinsic emergence.

In principle there are many different ways for trying to reach the goal of an efficient control of physical emergence. They differ mainly as regards the nature of theoretical models introduced. More precisely, within this context we can focus our attention on two relevant features, allowing for a possible classification of the models so far proposed: the first one describing the role played by *general principles*, and the second one related to the existence of *individual differences* between the single components of the system under study. The former feature leads to introduce a distinction between *ideal* and *non-ideal* models (Pessa, 2000). The latter, instead, suggests a distinction between *homogeneity-based* and *heterogeneity-based* models. In the following we will add some words of explanation to clarify the meaning of these attributes.

Ideal models. The attribute "ideal" will be here used to denote the models in which evolution laws, transition rules, constraints, boundary conditions, as well as every other circumstance underlying the production of system's behaviors, are nothing but a consequence of general principles, such as energy conservation principle, or least action principle. Perhaps the most typical examples of ideal models are the ones given by Quantum Field Theory (Itzykson and Zuber, 1986; Umezawa, 1993; Peskin and Schroeder, 1995; Huang, 1998; Lahiri and Pal, 2001). Anyway, we could rightly claim that most models and theories currently used in theoretical physics belong to this category.

Non-ideal models. Within this category we will collect all models in which behaviors (local or global) are nothing but a consequence of the introduction of suitable local evolution rules, supplemented by a fortunate choice of initial and boundary conditions, as well as of right parameter values. All models introduced in Artificial Life, for instance, fulfill this requirement, and the same can be said of models based on Cellular Automata or on Artificial Neural Networks. Even most models of self-organization based on differential equations, such as the Dissipative Structures quoted above, belong to this category, as the form chosen for their evolution equations is such that it cannot be derived from a general Variational Principle. Moreover, the nature of produced behaviors is strongly dependent on the kind of boundary conditions adopted.

Homogeneity-based models. Within them, all individual difference between the single elementary components at the lowest description level are neglected. Every component is identical to every other component and fulfills the same laws. This is the preferred choice for most ideal and non-ideal models, having the advantage of making simpler the mathematical analysis of model themselves.

Heterogeneity-based models. In the latter each elementary component is endowed with a sort of individuality, that is with individual features which, in principle, could differ from the ones of other elementary components. Even if this choice allows for a greater design flexibility, it makes, however, more difficult the mathematical investigation of these models. This strategy has been often adopted in (more or less) biologically inspired models. Among them we will quote, besides Artificial Life models, Artificial Neural Networks and Immune Networks.

Not all models belonging to these categories allow for a control of physical emergence, in the sense defined at the beginning of this section. This occurs almost exclusively for ideal and homogeneity-based models. This is proved, among the others, by the existence of consistent theories of special kinds of intrinsic emergence, such as the one associated to second-order phase transitions giving rise to superconductivity, superfluidity, ferromagnetism, laser effect (see, for instance, Goldenfeld, 1992). As a matter of fact, we are able to concretely induce such phenomena and to use them in an efficient way for technological purposes. It is, however, undeniable that non-ideal and heterogeneity-based models are more suited to describe biological emergence. On the other hand, the latter appears to be characterized by features which, at first sight, *cannot be present* in ideal and homogeneity-based models, such as:

- *medium-range correlations* (vs *long-range correlations* occurring in ideal models);
- *metastable states* (vs *stable ground states*);
- *heterogeneity* (vs *homogeneity*);
- *hierarchical organization* (vs *collective phenomena*);
- *interacting with the environment* (vs *working in the infinite volume limit*).

The question of the intrinsic difference between biological and physical emergence can be therefore reduced to the one of the irreducibility of non-ideal and heterogeneity-based models of biological emergence to the ones (ideal and homogeneity-based) of physical emergence. In this regard, we can claim that such an irreducibility doesn't hold if at least one of the following circumstances is verified:

a) it can be shown that models of biological emergence can be directly translated in the language of models of physical emergence; in other words both kinds of models are reciprocally equivalent (at least from a formal point o view);
b) it can be shown that the features characterizing models of biological emergence are nothing but macroscopic effects produced by a physical emergence occurring at a lower, microscopic level;
c) it can be shown that, introducing directly into models of biological emergence some features typical of models of physical emergence, the

former can be investigated through the same formal tools used for the latter.

The procedures for checking the occurrence of each one of the above circumstances gave rise to specific lines of research. The findings so far obtained will be briefly sketched, for each one of the cases a), b), c), in the following sections.

4. THE EQUIVALENCE BETWEEN DIFFERENT MODELS

Within this context two different questions must be answered:

1. can we find procedures to translate a model of biological emergence into a model of physical emergence?
2. can we find procedures to translate a model of physical emergence into a model of biological emergence?

If the answers to both questions would be positive, then we could claim that circumstance a) occurs. Here we will deal separately with each question, by starting with question 1). In this regard, among the many possible classes of biological emergence models, we will take into consideration only two of them: *reaction-diffusion systems* and *neural networks*. The former have been used to model, for instance, species evolution, swarm intelligence, morphogenesis. The above quoted Dissipative Structures also belong to this category. In general models of this kind are constituted by a set of basic components (which, to conform to a common usage, we will conventionally denote as *particles*) undergoing two types of processes: a random *diffusion* (which, for instance, can be modeled as a random walk on a suitable continuous or discrete space) and *reactions* between particles (the spontaneous decay of a particle, giving rise to new particles, being considered as a special case of a reaction). At any given time instant these models allow a suitable definition of system's microstate, characterized by the values of microscopic variables associated to the single particles. As the system evolution is stochastic, to each microstate α we can associate its probability of occurrence at time t, denoted by $p(\alpha,t)$. The latter must fulfill an evolution equation of the form:

$$\frac{dp(\alpha,t)}{dt} = \sum_{\beta} R_{\beta\alpha}\, p(\beta,t) - \sum_{\beta} R_{\alpha\beta}\, p(\alpha,t)$$

Here the symbol $R_{\beta\alpha}$ denotes the transition rate from the state β into α. The equation above is usually referred to as *master equation*. It is easy to

understand that, in general, the particle number is not conserved, owing to the fact that reactions can, in principle, destroy or create particles. This circumstance lets introduce the so-called formalism of *second quantization*, widely used in Quantum Field Theory (the prototype model of physical emergence). This can be done (the pioneering work on this subject was done by Doi, 1976; Peliti, 1985) very easily, for instance, in the case of particles moving on a discrete spatial lattice characterized by suitable *nodes*, labeled by a coordinate i (we will use a single label to save on symbols). In this context we can introduce the numbers of particles lying in each node (the so-called *occupation numbers*) and two operators, that is a *creation operator* $\alpha^{\dagger}_i$ and a *destruction operator* α_i, acting on system microstate and, respectively, raising the number of particles lying in the i-th node of one unity and lowering the same number of one unity. It is possible to show that these operators fulfill the same commutation relationships holding in Quantum Field Theory. They can be used to define a state with given occupation numbers, starting from the ground state, in which no particle exists, and applying to it, in a suitable way, creation and destruction operators. If we denote by $|\alpha,t\rangle$ a state characterized, at time t, by given occupation numbers (summarized through the single label α) system's state vector can then be defined as:

$$|\psi(t)\rangle = \sum_{\alpha} p(\alpha,t)|\alpha,t\rangle$$

By substituting this definition into the master equation it is possible to see that system's state vector fulfills a Schrödinger equation of the form:

$$\frac{d|\psi(t)\rangle}{dt} = -H|\psi(t)\rangle$$

where H denotes a suitable Hamiltonian, whose explicit form depends on the specific laws adopted for transition rates. In the case of a lattice of nodes, a suitable continuum limit then gives rise to a field Hamiltonian which can be dealt with exactly with the same methods used in standard Quantum Field Theory (on this topic see Cardy, 1996). Therefore the above procedure lets us perform a complete translation of models of biological emergence of reaction-diffusion type into models of physical emergence. It has been applied by a number of researchers to investigate the statistical features of models of interacting biological agents (see, among the others, Cardy and Täuber, 1998; Pastor-Satorras and Solé, 2001).

Let us now focus our attention on the case of neural networks, well known models allowing for emergent behaviors, in turn inspired by biological observations. In this regard it is convenient to restrict our consideration to the so-called Cellular Neural Networks (CNN), first introduced by Chua and Yang (see Chua and Yang, 1988). These latter can be described (for a short review see. Chua and Roska, 1993) as systems of nonlinear units arranged in one or more layers on a regular grid. The CNN differ from other neural networks since the interconnections between units are local and translationally invariant. The latter property means that both the type and the strength of the connection from the i-th to the j-th unit depend only on the relative position of j with respect to i. At every time instant to each unit of a CNN are associated two values: its (internal) state , denoted by $v^m_i(t)$, and its *output*, denoted by $u^m_i(t)$. Here the index m denotes the layer to which the unit belongs and the index i denotes the spatial location of the same unit within the layer. The general form of the dynamical laws ruling the time evolution of these functions is:

$$\frac{dv_i^m(t)}{dt} = -g\, v_i^m(t) + \sum_q \sum_k a_k^{qm}[u_{i+k}^m(t), u_i^m(t); P_{ak}^{qm}]$$

$$u_i^m(t) = f[v_i^m(t)].$$

In these formulae the function g describes the *inner dynamics* of a single unit, whereas f denotes the *output function* (often of sigmoidal type). Besides, the sum on the index q runs on all layer indices, and the sum on the index k runs on all values such that $i + k$ lies in the neighbourhood of i. Finally, the symbol $P^{qm}{}_{ak}$ denotes a set of suitable constant parameter values entering into the explicit expression of the *connection function* $a^{qm}{}_k$. Such parameters, whose values are independent from i, are referred to as *connection weights* or *templates* .

In the case of a single-layered CNN, in absence of inner dynamics, and when the output function coincides with the identity function, the previous laws can be reduced to the simpler form:

$$\frac{dv_i(t)}{dt} = \sum_k a_k[v_{i+k}(t), v_i(t); P_{ak}].$$

Let us now show, through a simple example, how a continuum described by partial differential equations can be approximated by a suitable CNN (see, in this regard, Kozek et al., 1995; Roska et al., 1995). To this end, let us choose a 1-dimensional medium ruled by the celebrated Korteweg-De Vries

equation, whose solution is a *solitary wave*, and, for reasons of convenience, cast under the form:

$$\frac{\partial \varphi}{\partial t} + \varphi \frac{\partial \varphi}{\partial x} + \delta^2 \frac{\partial^3 \varphi}{\partial x^3} = 0 .$$

Here δ^2 is a suitable parameter. If we introduce a discretization of spatial coordinates, based on a fixed space step Δx , the field function $\varphi(x,t)$ is replaced by a set of time-dependent functions $v_i(t)$ ($i = 1, 2, \ldots , N$), and the previous equation is replaced, in turn, by the set of ordinary differential equations:

$$\frac{dv_i}{dt} = -v_i \frac{v_{i+1} - v_{i-1}}{\Delta x} - \delta^2 \frac{v_{i+2} - 2v_{i+1} + 2v_{i-1} - v_{i-2}}{\Delta x^3} .$$

It is then easy to see that this set of equations describes the dynamics of a CNN, whose connection function is given by:

$$a_k = w_k v_{i+k} + r_k v_{i+k} .$$

In such a CNN the neighbourhood of the i-th unit goes from the (i+2)-th unit up to the (i-2)-th unit, so that the index k can assume only the values +2, +1, 0, -1, -2. A direct inspection shows that the values of connection weights w_k and r_k are given by:

$$v_2 = -\frac{\delta^2}{\Delta x^3}, \quad v_1 = \frac{2\delta^2}{\Delta x^3}, \quad v_0 = 0, \quad v_{-1} = -\frac{2\delta^2}{\Delta x^3}, \quad v_{-2} = \frac{\delta^2}{\Delta x^3}$$

$$r_2 = 0, \quad r_1 = -\frac{1}{\Delta x}, \quad r_0 = 0, \quad r_{-1} = \frac{1}{\Delta x}, \quad r_{-2} = 0.$$

So far we worked only within a classical framework. But is it entirely classical? The answer seems to be negative, as it has been shown (Vahala et al., 2003) that the solitary wave described by Korteweg-De Vries equation, just simulated by our CNN, can be reproduced by resorting to a suitable quantum lattice gas (a paper containing even a review on this topic is Yepez, 2002). We thus started from a biologically inspired model (the CNN), producing a behavior (the solitary wave) which can just be reproduced by a quantum system (the quantum lattice gas). We have here another case of a rule of translation from a model of biological emergence to a model of physical emergence.

Let us now turn to the question 2). In this regard, we must remember that it was known from long time (the pioneering work was done in Nelson, 1967) that both Quantum Mechanics and Quantum Field Theory can be rewritten under the form of classical systems endowed with a stochastic noise. Such a circumstance is just at the basis of the well known method of *Stochastic Quantization* (Itzykson and Drouffe, 1989a, 1989b; Parisi, 1998; Chaichian and Demichev, 2001). In it the quantum dynamics of a system described by a field $\varphi(x,t)$ is taken as equivalent to the one of a classical stochastic system obeying the Langevin equation:

$$\frac{\partial\varphi(x,t,\tau)}{\partial\tau} = -F[\varphi(x,t,\tau)] + \eta(t,\tau)$$

when the further fictitious variable τ tends to infinity. Here $F[\varphi(x,t,\tau)]$ denotes the form assumed by the left-hand side of motion equation (right-hand side being assumed as equal to zero) when time derivatives are replaced by derivatives with respect to $\tau = -i\,t$. Moreover, $\eta(t,\tau)$ denotes a suitable white noise source.

In this regard two remarks are to be made:

- a Langevin equation is nothing but a particular approximation resulting from a master equation;
- discretized versions of equations obtained from Stochastic Quantization method can be formulated under the form of laws of dynamical evolution of suitable neural networks.

Both remarks suggests that, in principle, a physical model of emergence (based on Quantum Field Theory) could be cast (via stochastic interpretation) under the form of a model of biological emergence. As it is well known, such an approach has been very fruitful in the domain of neural network models (see Amit, 1989; Dotsenko, 1994; Domany et al., 1996; Saad, 1998; Györgyi, 2001). Besides, it is to be recalled that the mathematical equivalence between stochastic partial differential equations and Quantum Field Theory (Fogedby, 1998; Fogedby and Brandenburg, 2002) throws a bridge between the world of stochastic models of biological phenomena and the one of models of physical emergence.

As a conclusion, all arguments so far presented in this section point to the existence of a number of precise translation rules of models of biological emergence into models of physical emergence and vice versa. The formal equivalence between these two kinds of models is therefore more than a hypothesis: it is a real fact!

5. QUANTUM APPROACHES TO BIOLOGICAL PHENOMENA

Since the Eighties the Quantum Field Theory-based approach to living matter is an active research domain (the best review is contained in Vitiello, 2001, Chapters 3-5). While it is impossible here to review all models and relevant experimental evidence, we will stress that it solves three main problems arising in the physical study of biological systems:

- biological processes usually require a very small amount of energy, often only slightly larger than average energy of a thermal fluctuation; this makes biological systems highly sensitive even to very small inputs; how can a so small energy produce so complex organizational effects?
- most biological processes require a non-dissipative and very efficient transmission of energy and information; how can this occur, despite the dissipation sources ubiquitously present in biological matter?
- most biological processes display a strong coherence; how can it be kept despite the ridiculously small decoherence time (of the order of 10 – 40 s) associated to physical features of biological molecules and to temperature characterizing biochemical reactions (see, in this regard, Tegmark, 2000; Alfinito et al., 2001)?

It can be shown that, if we take into account that most molecular components of living matter carry an electric dipole moment, we can introduce quantum fields describing the dipole excitations. In the case of water (the principal component of biological systems) both the equations describing these fields and their solutions are invariant with respect to rotations in 3-dimensional space. In water, however, are embedded quasi-monodimensional macromolecular chains, such as proteins. In the Seventies Davydov showed how a small energy release at one end of such chains, produced typically by ATP hydrolysis, gives rise to a collective mode (that is to a collective motion of atoms belonging to macromolecular chain) propagating along the chain itself under the form of a solitary wave. This non-dissipative form of energy transport (possible only within a quantum framework) triggers in turn a spontaneous symmetry breaking in the water molecules surrounding the chain. This means that the solutions of the equations describing dipole excitations are no longer invariant with respect to 3-dimensional rotations. However in Quantum Field Theory a spontaneous symmetry breaking is always associated to the occurrence of zero mass particles (the so-called Nambu-Goldstone bosons) carrying long-range interactions (for a deeper analysis of their role see Umezawa, 1993; Burgess, 2000; Brading and Castellani, 2003) . The effect of the latter is to induce a long-range ordering within the surrounding water, which in turn supports coherent phenomena in distant parts of the same biological system.

This description is further complicated by the existence of finite volume effects and by the presence of electromagnetic long range forces. These latter imply the existence of different and, more complex, organizational levels beyond the simple one so far described. Their exploration is still in progress, mainly with reference to brain dynamics (for a recent review see Jibu and Yasue, 2004; recent advances are contained in Pessa and Vitiello, 2004a, 2004b).

Summarizing the above discussion, we can say that models of physical emergence based on Quantum Field Theory appear as the only ones able to solve the problems quoted above. Namely:

- they don't require large energy amounts, as a collective process is based on the intrinsic features of dynamical laws fulfilled by quantum fields, and not on energy;
- collective processes, owing to their intrinsic coherence, grant for a non-dissipative energy transport (just as occurs for laser light);
- coherence is kept owing to the existence of Nambu-Goldstone modes (possible only within Quantum Field Theory) which counteract every action tending to destroy it.

We can thus conclude that Quantum Field Theory-based approach to living matter evidenced how suitable models of physical emergence be able to explain the main features of biological emergence. Even if most work must still be done, the road is open!

6. QUANTUM NEURAL NETWORKS

When focussing our attention towards a direct introduction of features typical of physical emergence into models of biological emergence, we are forced to limit our considerations to the field of quantum neural networks, so far the only one in which this topic was object of an intensive investigation. In this regard, we must distinguish between two approaches: the most popular one, in which quantum neurons are introduced from the start as physical realizations of quantum systems, such as multiple slits or quantum dots (for reviews see Behrman et al., 2000; Narayanan and Menneer, 2000), and the less studied one, in which quantum neurons are quantum dynamical systems whose laws are obtained by applying a quantization procedure to dynamical laws of classical neurons. Here we will take into consideration only the latter, by introducing a simple example to illustrate the nature of this approach. In this regard, let us start from the usual formulation of the dynamics of a McCulloch-Pitts neuron in terms of a differential equation:

$$\frac{ds_i}{dt} = -s_i + F(\sum_j w_{ij} s_j),$$

where s_i denotes the instantaneous output activity of the i-th neuron, w_{ij} are the connection weights, and the function F has the form:

$$F(x) = \frac{1}{1+\exp(-x)}.$$

By supposing that each single neuron be characterized by a spontaneous base activity, represented in a fictitious way through the introduction of a suitable *self-connection weight*, denoted by w, we will obtain that a single isolated neuron will obey the dynamical law:

$$\frac{ds}{dt} = -s + F(ws).$$

We can now consider the latter as a *first integral* of the 'true' dynamical equation, which can be easily obtained through a further derivation:

$$\frac{d^2 s}{dt^2} = s - F(ws) - wsF'(ws) + wF'(ws)F(ws).$$

Here the symbol F' denotes the derivative of the function F with respect to its argument. The latter equation can be written as:

$$\frac{d^2 s}{dt^2} = -\frac{\partial V}{\partial s},$$

where:

$$V = -\frac{s^2}{2} + sF(ws) - \frac{F^2(ws)}{2},$$

so that the quantity:

$$E = \frac{1}{2}\left(\frac{ds}{dt}\right)^2 + V$$

be conserved. It is tempting to call this quantity *energy*, but it is deeply different from the physical energy. Namely s doesn't denote a spatial coordinate, but an activation value.

In order to continue the implementation of our strategy, we need to take into account that the spatial coordinates, used in physics to specify the position of a given particle, are nothing but particular examples of *configurational variables*, that is independent variables, different from the time variable, which allow for a specification of the instantaneous configuration of the system under study. The fact that the configuration of a point particle is specified only through its spatial location must not prevent from the understanding that the configurations of systems differing from point particles could be, in principle, specified through other kinds of independent variables, even differing from spatial coordinates. In other words, we claim that classical mechanics, as well as quantization procedures, should be generalized in such a way as to include whatever kind of configurational variables, and not only spatial positions.

If we adopt such a point of view, then we must recognize that the configurational variable most suited to describe a neuronal unit such as our McCulloch-Pitts neuron is its output activity s and not its spatial location. Namely within this primitive model the indices associated to different neurons don't have a spatial meaning and can be considered only as labels useful to logically distinguish one neuron from another. Within this framework we could say that E is a sort of *configurational total energy*. This lets us introduce a *configurational Hamiltonian* and a *configurational wavefunction* $\psi(s,t)$. If we identify the *configurational momentum* p with ds/dt, we can now generalize the definition of two quantum operators: the one associated to the configurational coordinate s (to be denoted by S) and the one associated to the configurational momentum p (to be denoted by P) in this simple way:

$$S\psi(s,t) = s\,\psi(s,t), \quad P\psi(s,t) = -i\hbar\frac{\partial\psi(s,t)}{\partial s}.$$

It is then very easy to write the Schrödinger equation for a single quantum neuron:

$$-\frac{h}{8\pi^2}\frac{\partial^2\psi}{\partial s^2} + \left[-\frac{s^2}{2} + sF(ws) - \frac{F^2(ws)}{2}\right]\psi = i\frac{h}{2\pi}\frac{\partial\psi}{\partial t}.$$

If, in a first approximation, owing to the smallness of w, we write:

$$F(ws) \approx \frac{1}{2} + \frac{1}{4} ws ,$$

then this equation becomes formally identical to the one of a *quantum harmonic oscillator*, whose solutions are well known from longtime. It will be very easy, then, to derive an explicit form for the spectrum of *configurational energy levels* which characterize the behavior of our quantum neuron. We will not write here the explicit formula, which can be easily found in every textbook on QM.

In order to generalize our previous theory, we will allow for an interaction between different quantum neurons, of the kind described in the previous paragraph. If we suppose that the self-connection weight w be the same for all neurons, then the classical dynamical equation ruling the behavior of a generic neuron will assume the form:

$$\frac{d^2 s}{dt^2} = s - F(ws + I) - wsF'(ws + I) + wF'(ws + I)F(ws + I),$$

where I denotes the contribution coming from other neurons, which we will assume to be of the form:

$$I = \sum_j w_{ij} s_j .$$

Some mathematical considerations show that, in this case, the total potential energy of the neural network V is given by the sum of two terms:

$$V = \sum_i V_i + V_{\text{int}} ,$$

where:

$$V = -\frac{s_i^2}{2} + s_i F(ws_i + \sum_j w_{ij} s_j) - \frac{N}{2} F^2(ws_i + \sum_j w_{ij} s_j)$$
$$V_{\text{int}} = \sum_{ij} s_i s_j [F(ws_i + \sum_k w_{ik} s_k) + F(ws_j + \sum_k w_{jk} s_k)]$$

and N denotes the total number of neurons.

Computer simulations of the behavior of this quantum neural network, performed through Quantum Monte Carlo method, evidenced (Pessa, 2004) how the introduction of quantum features introduce long-range correlations

not present in its classical counterpart. Even if this is not a model based on Quantum Field Theory, but only on Quantum Mechanics, these first results suggest that the introduction of quantum features in a biologically inspired model lead to new models which, nevertheless, can be investigated with the same tools used in models of physical emergence.

7. CONCLUSION

The arguments presented in previous sections evidenced how all circumstances a), b), c), introduced in section 3, are undoubtedly verified. We therefore cannot assert, on the basis of these findings, that the modelling of biological emergence require new tools, irreducible to the ones used in modelling physical emergence. This conclusion entails that all theoretical work so far done within models of physical emergence be useful also to investigate biological emergence.

Despite this optimistic conclusion, however, we must not forget that the theoretical apparatus so far introduced to study physical emergence is largely inadequate and most problems still wait for a solution. Amongst the latter we will quote:

- how to build a comprehensive theory of defect formation in phase transitions;
- how to build a theory of multi-level emergence;
- how to build a theory of systems with variable kinds (and not only numbers) of components.

Any further progress along these directions will produce a significant improvement of our understanding of emergence, both in physical and in biological context.

REFERENCES

Alfinito, E., Viglione, R. G., and Vitiello, G., 2001, The decoherence criterion, *Modern Physics Letters B* **15**:127-136.

Amit, D. J., 1989, *Modeling Brain Function. The World of Attractor Neural Networks,* Cambridge University Press, Cambridge, UK.

Bedau, M. A., 1997, Weak emergence, *Philosophical Perspectives* **11**:375-399.

Behrman, E. C., Nash, L. R., Steck, J. E., Chandrashekar, V. G., and Skinner, S. R., 2000, Simulations of quantum neural networks, *Information Sciences* **128**:257-269.

Belintsev, B. N., 1983, Dissipative structures and the problem of biological pattern formation, *Soviet Physics Uspekhi* **26**:775-800.

Beloussov, L. V., 1998, *The Dynamic Architecture of Developing Organism*, Kluwer, Dordrecht.

Brading, K., and Castellani, E., eds., 2003, *Symmetries in Physics: Philosophical Reflections*, Cambridge University Press, Cambridge, UK.

Burgess, C. P., 2000, Goldstone and pseudo-Goldstone bosons in nuclear, particle and condensed-matter physics, *Physics Reports* **330**:193-261.

Cardy, J. L., 1996, *Scaling and Renormalization in Statistical Physics*, Cambridge University Press, Cambridge, UK.

Cardy, J. L., and Täuber, U. C., 1998, Field theory of branching and annihilating random walks, *Journal of Statistical Physics* **90**:1-56.

Chaichian, M., and Demichev, A., 2001, *Path Integrals in Physics. Volume 2: Quantum Field Theory, Statistical Physics and other modern applications*, IOP Press, Bristol, UK.

Chua, L. O., and Roska, T., 1993, The CNN Paradigm, *IEEE Transactions on Circuits and Systems* **40**:147-156.

Chua, L. O., and Yang, L., 1988, Cellular Neural Networks: Theory and applications, *IEEE Transactions on Circuits and Systems* **35**:1257-1290.

Cruchtfield, J. P., 1994, The Calculi of Emergence: Computation, Dynamics and Induction, *Physica D* **75**:11-54.

Doi, M., 1976, Second quantization representation for classical many-particle system, *Journal of Physics A* **9**:1465-1477.

Domany, E., Van Hemmen, J. L., and Schulten, K., eds., 1996, *Models of Neural Networks III: Association, Generalization, and Representation (Physics of Neural Networks)*, Springer, Berlin-Heidelberg-New York.

Dotsenko, V., 1994, *An Introduction to the Theory of Spin Glasses and Neural Networks*, World Scientific, Singapore.

Fernández, A., 1985, Global instability of a monoparametric family of vector fields representing the unfolding of a dissipative structure, *Journal of Mathematical Physics*, **26**:2632-2633.

Fogedby, H. C., 1998, Soliton approach to the noisy Burgers equation. Steepest descent method, *Physical Review E* **57**:4943-4968.

Fogedby, H. C., and Brandenburg, A., 2002, Solitons in the noisy Burgers equation, *Physical Review E* **66**: 016604,1-9.

Glendinning, P. ,1994, *Stability, Instability and Chaos: An Introduction to the Theory of Nonlinear Differential Equations*, Cambridge University Press, Cambridge, UK.

Goldenfeld, N., 1992, *Lectures on Phase Transitions and the Renormalization Group*, Addison-Wesley, Reading, MA.

Guckenheimer, J., and Holmes, P., 1983, *Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields*, Springer, Berlin.

Györgyi, G., 2001, Techniques of replica symmetry breaking and the storage problem of the McCulloch-Pitts neuron, *Physics Reports* **342**:263-392.

Haken, H., 1978, *Synergetics. An Introduction*, Springer, Berlin.

Haken, H., 1983, *Advanced Synergetics*, Springer, Berlin.

Haken, H., 1988, *Information and Self-Organization. A Macroscopic Approach to Complex Systems*, Springer, Berlin.

Huang, K., 1998, *Quantum Field Theory: From Operators to Path Integrals*, Wiley, New York.

Iooss, G., and Joseph, D. D., 1981, *Elementary Stability and Bifurcation Theory*, Springer, New York.

Itzykson, C., and Drouffe, J.-M., 1989a, *Statistical Field Theory: Volume 1, from Brownian Motion to Renormalization and Lattice Gauge Theory*, Cambridge University Press, Cambridge, UK.

Itzykson, C., and Drouffe, J.-M., 1989b, *Statistical Field Theory: Volume 2, Strong Coupling, Monte Carlo methods, Conformal Field Theory and Random Systems*, Cambridge University Press, Cambridge, UK.

Itzykson, C., and Zuber, J. B., 1986, *Quantum Field Theory*, McGraw-Hill, Singapore.

Kozek, T., Chua, L. O., Roska, T., Wolf, D., Tezlaff, R., Puffer, F., and Lotz, K., 1995, Simulating nonlinear waves and partial differential equations via CNN – Part II: Typical examples, *IEEE Transactions on Circuits and Systems* **42**:816-820.

Jibu, M., and Yasue, K., 2004, Quantum brain dynamics and Quantum Field Theory, in: G. G. Globus, K. H. Pribram and G. Vitello, eds., *Brain and Being. At the Boundary Between Science, Philosophy, Language and Arts*, Benjamins, Amsterdam, pp. 267-290.

Lahiri, A., and Pal, P. B., 2001, *A First Book of Quantum Field Theory*, CRC Press, Boca Raton, FL.

Mikhailov, A. S., 1990, *Foundations of Synergetics I. Distributed Active Systems*, Springer, Berlin.

Mikhailov, A. S., and Loskutov, A, Yu., 1996, *Foundations of Synergetics II. Chaos and Noise*, 2nd revised edition, Springer, Berlin.

Mori, H., and Kuramoto, Y., 2001, *Dissipative Structures and Chaos*, Springer, Berlin.

Narayanan, A., and Menneer, T., 2000, Quantum artificial neural network architectures and components, *Information Sciences* **128**:231-255.

Nelson E., 1967, *Dynamical Theories of Brownian Motion*, Princeton University Press, Princeton, NJ.

Nicolis, G., and Prigogine, I., 1977, *Self-organization in Nonequilibrium Systems*, Wiley, New York.

Nitzan, A., and Ortoleva, P., 1980, Scaling and Ginzburg criteria for critical bifurcations in nonequilibrium reacting systems, *Physical Review A* **21**:1735-1755.

Pastor-Satorras, R., and Solé, R. V., 2001, Field theory of a reaction-diffusion model of quasispecies dynamics, *Physical Review E* **64**:051909, 1-7.

Parisi, G., 1998, *Statistical Field Theory*, (New edition), Perseus Books, New York.

Peliti, L., 1985, Path integral approach to birth-death processes on a lattice, *Journal de Physique* **46**:1469-1483.

Peskin, M. E., and Schroeder, D. V., 1995, *An Introduction to Quantum Field Theory*, Addison-Wesley, Reading, MA.

Pessa, E., 2000, Cognitive Modelling and Dynamical Systems Theory, *La Nuova Critica* **35**:53-93.

Pessa, E., 2004, Quantum connectionism and the emergence of cognition, in: G. G. Globus, K. H. Pribram and G. Vitello, eds., *Brain and Being. At the Boundary Between Science, Philosophy, Language and Arts*, Benjamins, Amsterdam, pp. 127-145.

Pessa, E., and Vitiello, G., 2004a, Quantum noise, entanglement and chaos in the Quantum Field Theory of Mind/Brain states, *Mind and Matter* **1**:59-79.

Pessa, E., and Vitiello, G., 2004b, Quantum noise induced entanglement and chaos in the dissipative quantum model of brain, *International Journal of Modern Physics B* **18**:841-858.

Ronald, E. M. A., Sipper, M., and Capcarrère, M. S., 1999, Design, observation, surprise! A test of emergence, *Artificial Life* **5**:225-239.

Roska, T., Chua, L.O., Wolf, D., Kozek, T., Tezlaff, R., and Puffer, F., 1995, Simulating nonlinear waves and partial differential equations via CNN – Part I: Basic techniques, *IEEE Transactions on Circuits and Systems* **42**:807-815.

Rueger, A., 2000, Physical emergence, diachronic and synchronic, *Synthese* **124**:297-322.

Saad, D., ed., 1998, *On-line Learning in Neural Networks*, Cambridge University Press, Cambridge, UK.

Sattinger, D. H., 1978, *Topics in Stability and Bifurcation Theory*, Springer, Berlin.

Sattinger, D. H., 1980, Bifurcation and symmetry breaking in applied mathematics, *Bulletin of the American Mathematical Society* **3**:779-819.

Scott, A., 2003, *Nonlinear Science: Emergence and Dynamics of Coherent Structures*, Oxford University Press, Oxford, UK.

Sewell, G. L., 1986, *Quantum Theory of Collective Phenomena*, Oxford University Press, Oxford, UK.

Sewell, G. L., 2002, *Quantum Mechanics and its Emergent Macrophysics*, Princeton University Press, Princeton, NJ.

Stein, D. L., 1980, Dissipative structures, broken symmetry, and the theory of equilibrium phase transitions, *Journal of Chemical Physics* **72**:2869-2874.

Tegmark, M., 2000, Why the brain is probably not a quantum computer, *Information Sciences* **128**:155-179.

Umezawa, H., 1993, *Advanced Field Theory. Micro, Macro, and Thermal Physics*, American Institute of Physics, New York.

Vahala, G., Yepez, J., and Vahala, L., 2003, Quantum lattice gas representation of some classical solitons, *Physics Letters A* **310**:187-196.

Vanderbauwhede, A., 1982, *Local Bifurcation and Symmetry*, Pitman, Boston.

Vitiello, G., 2001, *My Double Unveiled*, Benjamins, Amsterdam.

Yepez, J., 2002, Quantum lattice-gas model for the Burgers equation, *Journal of Statistical Physics* **107**:203-224.

5

GENERAL SYSTEMS

INTERACTIONS BETWEEN SYSTEMS

Mario R. Abram
AIRS - Associazione Italiana per la Ricerca sui Sistemi, Milano, Italy

Abstract: On the basis of previous contributions about a methodology for the decomposition of systems, we attempt to investigate the interaction between subsystems. When considering the problem of decomposition of systems, different interaction levels can be evidenced and then a classification scheme can be defined. Some relation patterns in a decomposition can be interpreted as hierarchical levels showing the emergence of some hierarchical structure. The attitude (subjective or objective) and the methodological approach give the basic conditions for the application of these concepts. The preliminary analysis regarding the control of an industrial plant may be an useful example to test these ideas and to evaluate the interaction of critical subsystems, like human operator or critical control functions.

Key words: subsystems; decomposition; interaction; hierarchical level; control.

1. INTRODUCTION

The effort to find information about systems characterization and the definition of their properties, contributed to start various research paths developing different methodologies. The leading ideas that inspired this research are confronted with the practical problems to dominate complex aspects of reality, to gain information for controlling the evolution of difficult situations, to establish stability conditions, in few words the goal is to control our interaction with the environment and in general with the world around us. These problems are present in our daily life, as, for instance, the problems connected with the control of large structures such as railways, air traffic, power distribution and communications networks. The specific problems become more complex when these different systems interact each other.

Overcoming the reductive approach based on the study of the structure of the systems components, the system approach stresses the impact of the relations between the elements or subsystems, that constitute the system (Klir, 1991).

Starting from this point it is interesting to investigate the possibility to find information using the relations between the elements. In particular, the decomposition of a system into subsystems and the identification of the relations between them should give the possibility to go deeper into the practical system approach.

The system concept may then be a promising approach for answering some questions and it should be useful to develop some reflections about the concepts of subsystem, decomposition and interaction between subsystems.

Following these ideas, in a previous paper (Abram, 2002), some introductory elements for a methodology for systems decomposition were presented.

Relations between subsystems can be interpreted as action and reaction relations and, under particular conditions, they can be defined as interactions. The interactions can involve two or more subsystems. They can manifest different conditions that we call levels of interaction. Taking these ideas into account, when the set of relations is depicted as a whole, it is possible to see regular patterns. Specific properties as regularity, modular form, regular irregularities and similarity may emerge from a relation pattern developed in a decomposition process manifesting the presence of hierarchical structures. By applying this methodology, when the number of subsystems involved in a decomposition process increases, it is possible to see that in the patterns of the relations between subsystems some regular hierarchical structures may emerge.

This approach seems to be useful and reachable, because we propose to maintain ourselves at the basic level of relations. We will not investigate the analytical form that we can attribute to relations. We will limit ourselves to find evidence of special patterns in the relation matrices.

In the paper we recall the decomposition methodology and we give some representations of the decomposition process (section 2). Then we consider the interaction between systems (section 3) defining direct and indirect interactions and introducing the levels of interactions. Focusing on these concepts we see that in a decomposition process, hierarchical structures can emerge (section 4); these structures are related to subsystems ordering and relations patterns. In addition, the attitude to approach the decomposition process, "subjective" or "objective", is evidenced (section 5). The preliminary analysis for designing the control apparatus of an industrial plant (section 6) may give a realistic example for understanding the possible application of the previous ideas and for identifying the context in view of

their correct application. The role of operator-observer is considered (section 7). Some remarks (section 8) and conclusions (section 9) close the paper.

2. DECOMPOSITION OF SYSTEMS

With reference to a very general definition of system, $\boldsymbol{S} = (T,R)$, where T is a set of elements and R is the set of relations between the elements of T (Klir, 1991), a methodology for the decomposition of a system into subsystems was developed (Abram, 2002). Assuming the relations between the subsystems as oriented, we can specify the action and reaction relations. This distinction enabled to define the "topological" matrices describing the relation patterns of such a decomposition, as action and reaction matrices.

In this way it is possible to investigate the decompositions of a system into subsystems and to identify some basic properties of the relations between them. A system $\boldsymbol{S}$ may have many decompositions $\boldsymbol{D}_n$, that we represent graphically as in figure 1. For example, the system $\boldsymbol{S}_i$ can be decomposed into the subsystems $\boldsymbol{P}_i$ and $\boldsymbol{N}_i$; the system $\boldsymbol{P}_i$ owns the property Q_i ; the system $\boldsymbol{N}_i$ does not own the property Q_i. We call this elementary decomposition.

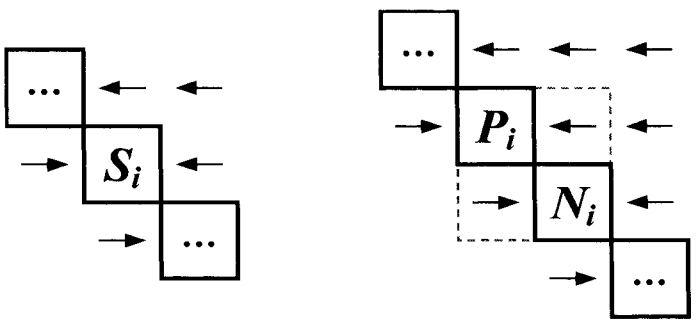


Figure 1. System decomposition.

It is evident that the number of subsystems involved in a decomposition process increases linearly, while the number of relations between the subsystems increases in a quadratic way. So the number of subsystems involved calls for managing a large amount of relations.

The decomposition methodology is based on the repetitive application of the elementary decomposition to the subsystems, focusing on the properties defining the subsystems and the identification of the relations involved with the subsystems.

We can operate in accordance with the following steps:

1. Apply the elementary decomposition to the system and put into evidence the relations that can be derived from the properties characterizing each subsystem.
2. Apply the elementary decomposition process more and more until the detail of the decomposition is adequate to visualize the relations involved in the problem.
3. Each elementary decomposition doubles the previous existing relations with the other subsystems and such relations are considered true for the new subsystems of the last decomposition step.
4. Reduce the doubled relations of the last decomposition by eliminating the relations that are not consistent with the properties of the subsystems.
5. If further subsystems are not useful to go deeper in the analysis, they can be composed back in one subsystem; the existing relations may then be grouped again.

The relations between the subsystems can be interpreted as action and reaction and they are represented by oriented lines or arrows (action: "→" and reaction "←").

The decomposition process can be represented in different ways. We choose the graphical method as in figure 1 because we want to visualize the subsystems involved, and the action and reaction relations can be evidenced by arrows showing a possible causality or functionality relationship.

More abstract representations are possible and can be described in a mathematical abstract form. The isomorphism between the possible and alternative representations involves, for example, the graphical and matrix representations, the block schemas representation, used in engineering, and the directed graphs representation.

In directed graphs the subsystems are identified by nodes and the relations are represented as arrows, or oriented lines, that connect the nodes.

It is evident how the organization and the ordering of the nodes may contribute to gain an easier or more difficult understanding of the emergence for hierarchical structures. The emergence of these structures comes into evidence by considering the specific patterns that describe them.

3. INTERACTIONS BETWEEN SYSTEMS

The idea of interaction comes into evidence when the action and reaction relations are present, involving the concept of mutual, reciprocal influence of one or more subsystems.

With reference to figure 2, we use the operator notation, so we represent a relation as an operator acting on a subsystem.

We may specify two types of interactions: direct and indirect interactions. Given a decomposition $\boldsymbol{D}_n$, we call $\boldsymbol{S}_\alpha$ and $\boldsymbol{S}_\beta$ two subsystems and $R_{\alpha\beta}\, R_{\beta\alpha}$ the two possible relations between the subsystems. Then we can point out the following definitions.

Definition 1: two subsystems $\boldsymbol{S}_\alpha$ and $\boldsymbol{S}_\beta$ interact directly if the two relations $R_{\alpha\beta}$ and $R_{\beta\alpha}$ exist and $R_{\alpha\beta}\, R_{\beta\alpha} = I$.

For the diagram in figure 2 (*a*) we can write:

$$
\begin{aligned}
S_\alpha &= I\, S_\alpha \\
S_\beta &= R_{\beta\alpha} S_\alpha \\
S_\alpha &= R_{\alpha\beta} S_\beta = R_{\alpha\beta}(R_{\beta\alpha} S_\alpha) = R_{\alpha\beta} R_{\beta\alpha} S_\alpha = I\, S_\alpha \\
I &= R_{\alpha\beta} R_{\beta\alpha}
\end{aligned}
$$

Definition 2: Two subsystems $\boldsymbol{S}_\alpha$ and $\boldsymbol{S}_\beta$ interact indirectly if it there exists a subsystem $\boldsymbol{S}_\gamma$ so that $\boldsymbol{S}_\gamma = R_{\gamma\alpha}\, \boldsymbol{S}_\alpha$ and $\boldsymbol{S}_\beta = R_{\beta\gamma}$ and $R_{\alpha\beta}\, R_{\beta\gamma}\, R_{\gamma\alpha} = I$.

With reference to figure 2 (*b*) we can write:

$$
\begin{aligned}
S_\beta &= R_{\beta\gamma} S_\gamma \\
S_\gamma &= R_{\gamma\alpha} S_\alpha \\
S_\beta &= R_{\beta\gamma} S_\gamma = R_{\beta\gamma}(R_{\gamma\alpha} S_\alpha) = R_{\beta\gamma} R_{\gamma\alpha} S_\alpha \\
R_{\beta\alpha} &= R_{\beta\gamma} R_{\gamma\alpha} \\
S_\alpha &= R_{\alpha\beta} S_\beta \\
&= R_{\alpha\beta}(R_{\beta\gamma} S_\gamma) \\
&= R_{\alpha\beta}(R_{\beta\gamma}(R_{\gamma\alpha} S_\alpha)) = R_{\alpha\beta} R_{\beta\gamma} R_{\gamma\alpha} S_\alpha \\
I &= R_{\alpha\beta} R_{\beta\gamma} R_{\gamma\alpha}
\end{aligned}
$$

Note that in figure 2, with reference to interactions, we ask that the relations involved a closed loop with the subsystems. For example, with reference to the decomposition diagram in figure 3(*a*), the subsystem $\boldsymbol{S}_\alpha$ interacts directly with the subsystems $\boldsymbol{S}_\beta$ and $\boldsymbol{S}_\gamma$. Instead, in figure 3(*b*), the subsystems $\boldsymbol{S}_\alpha$ and $\boldsymbol{S}_\beta$ interact indirectly by mean of subsystem $\boldsymbol{S}_\gamma$.

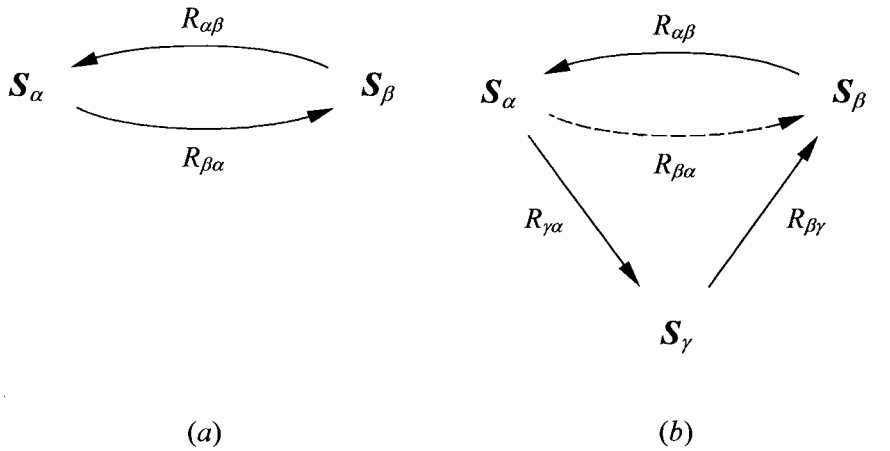


Figure 2. Direct interaction (*a*) and indirect interaction (*b*).

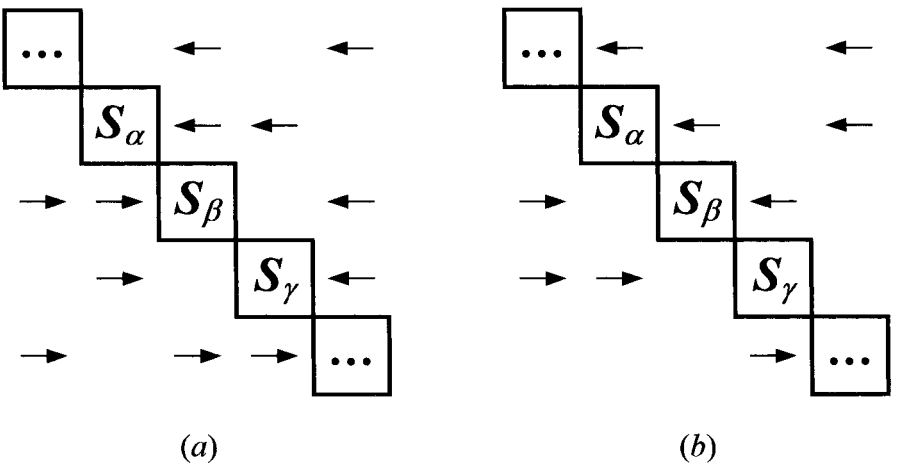


Figure 3. Direct interactions (*a*) and indirect interactions (*b*).

In a given decomposition, the number of subsystems that are involved in order to describe an interaction between the subsystems defines the level of interaction between the two subsystems. We introduce the following definitions.

Definition 3: The level of action L_A is the number of subsystems involved between two given subsystems that are in action relation with them.

Definition 4: The level of reaction L_R is the number of subsystems involved between the two subsystems that are in reaction relation with them.

Definition 5: The level of interaction L_I is the maximum of the level of action and the level of reaction $L_I = \max(L_A, L_R)$.

In this way the level of interaction between two subsystems is given by the number of subsystems that we need to interpose between the two subsystems in order to realize a closed loop of the relations involved. Because of that the identity condition is true.

In practical cases the direct relation may be decomposed in a finite number of elements by a finite number of decompositions steps. In real systems this case is frequently present.

So we may have interactions of level $0, 1, 2, \ldots, n$

The level of interaction is also the number of intermediate relations involved in the loop of interaction.

For example, in figure 2*a*, for the subsystems $\boldsymbol{S}_\alpha$ and $\boldsymbol{S}_\beta$, L_I is equal to zero. In figure 2*b* L_I is equal to one.

4. HIERARCHICAL STRUCTURES

The relation patterns of a decomposition give a global picture of the complete set of relations between the subsystems. The relations may appear to have regular configurations connected with rows, columns and diagonals. These patterns are useful to identify the status of interactions and, by evaluating a description of many interactions, it is possible to see the emergence of hierarchical structures in the relations patterns.

From the three regular patterns, three types of hierarchies may be present into a decomposition that we call Row, Column and Diagonal hierarchies that we identify as follows:

$H_R(i)$ Row hierarchy
$H_C(i)$ Column hierarchy
$H_{DA}(i)$ Action Diagonal hierarchy
$H_{DR}(i)$ Reaction Diagonal hierarchy.

Considering H_D a generic diagonal hierarchy, the diagonal patterns will be indicated conventionally as

$H_{DA}(i) = H_D(i,1)$ for $i > 1$
$H_{DR}(i) = H_D(1,i)$ for $i > 1$
$H_D(1) = H_D(1,1)$.

The indexes (i,j) express the starting point of the hierarchical level with reference to the corresponding element in the relations matrix.

Each type of hierarchy is characterized by the specific modality by which a subsystem is in relation with the other subsystems in the decomposition as shown in figure 4.

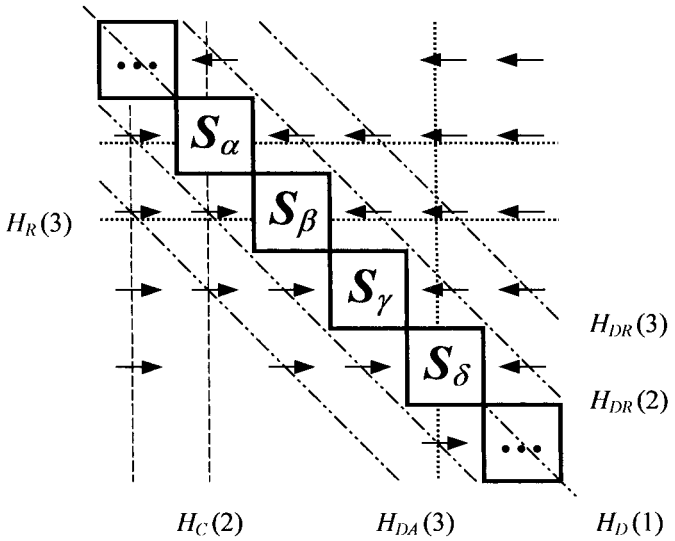


Figure 4. Hierarchical levels.

The presence of regular patterns, that are symmetric with reference to the diagonal, gives evidence of different levels of interaction between the subsystems.

$H_D(1)$ is the basic level; it is implicit and is due to the ordering coming out from the decomposition process.

The emergence of a hierarchical level is connected directly with the building of the relations paths. In particular it is interesting to note that the hierarchical levels are connected directly with the ordering properties of the subsystems in a decomposition. In general, a hierarchical level emerging from an ordered decomposition cannot be guaranteed from a different ordering of the same decomposition.

With these limits the relation pattern is a useful representation, a sort of "picture", of global properties emerging from the systems approach.

5. ATTITUDE

The decomposition methodology, the interaction levels and the hierarchical structures can be used from two different points of view.

If we pose ourselves in the position to watch the relations emerging from the analysis process developed decomposing a system into subsystems, we operate in order to see if a relationship between two or more subsystems exists. The key point is to find the more adequate decomposition of a system in order to find and to identify the relations between the subsystems. We maintain an "objective" point of view very similar to the scientist attitude to discover the laws undertaking phenomena.

An alternative point of view consists in specifying the relation patterns in a decomposition of a system. Given a decomposition we want to see or we want to impose a specific pattern. This is a position very similar to designing approach in which we specify and impose a set of attributes or requisites. This is a "subjective" point of view, very similar to an engineering approach. In this case we want to design and realize a specific system and, given a decomposition, we want to implement a specific relation patterns.

These two positions are present in our activities but we can speak of emergence of hierarchical levels only in the analytic process or "objective" position. Analysis makes hierarchical levels emergent.

6. APPLICATION TO PLANT CONTROL

The two points of view, active and passive, are present in the problem of controlling a system.

The reliable design of control strategies for the management of an industrial plant is often a difficult work. The complexity of phenomena imposes to use adequate and detailed modeling approaches. Nevertheless it may be interesting to investigate the possibility of developing a methodology for the study of the problem at a very general level, by concentrating on finding the key points of the problems and on identifying the interaction between the subsystems.

Recalling the application to the control of an industrial power plant, as it was described in a previous paper (Abram, 2002), the problem of controlling a system can be sketched starting from a decomposition process developing with the six steps reported in figure 5. Following the decomposition methodology, the relations between the subsystems are identified and settled as shown in graphical form in figure 6.

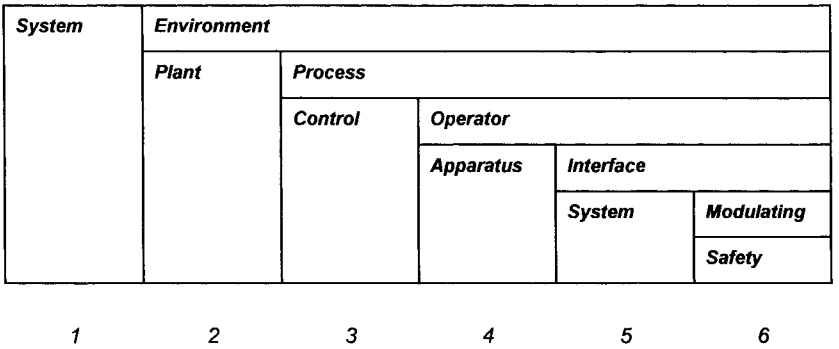


Figure 5. Example: decomposition steps for preliminary control analysis.

In designing the control strategy of the plant, the control subsystem must be designed in order to realize specific patterns of relations. In the case of a plant the phenomena are investigated and specific relations patterns can be discovered. In the case of designing control subsystems, specific relation patterns must be imposed, in order to have the desired evolution for all the subsystems.

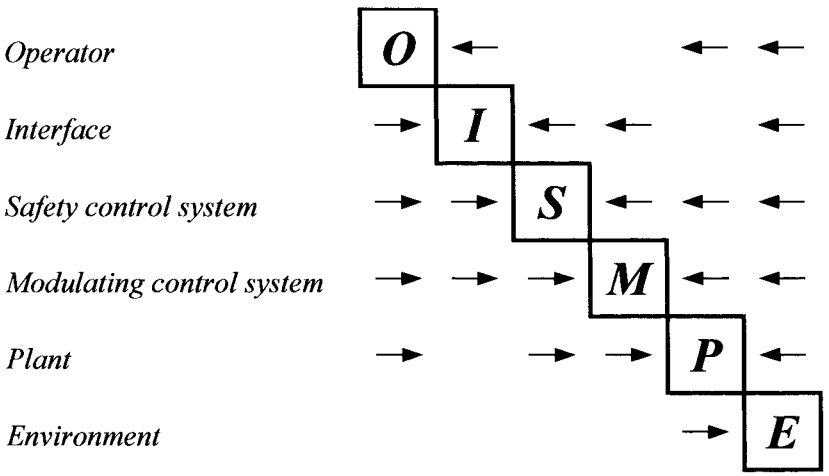


Figure 6. Example: decomposition for preliminary control analysis.

This is evident, for instance, when we design the control strategy of the plant, in which it is necessary to impose the different levels of a hierarchical control strategy. These hierarchies are described and are particularly evident in the specific relation patterns that are present in the topologic matrix of relations.

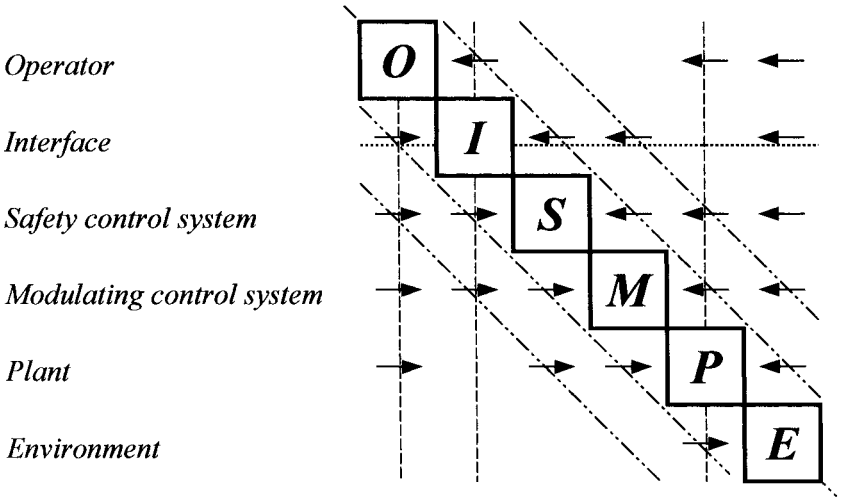


Figure 7. Interaction levels.

7. OPERATOR-OBSERVER FUNCTIONS

Going in a practical context when the decomposition is used to develop a preliminary analysis of the control subsystem, it is useful to give the right evidence to the role of the various subsystems. Even at this very general level we can describe the role of human operator managing the plant.

One problem that we face when designing the control subsystem, lies in defining the hierarchical levels that describe the action priorities of each subsystem on the other ones.

If an operator interacts with the process indirectly, by means of the various subsystems (Direct Action and Direct Reaction) this may be seen as the normal operating procedure: the hierarchy of subsystems is evidenced by the two diagonal path of action $H_{DA}(2)$ and reaction $H_{DR}(2)$ patterns (figure 7). In this case the interactions between the subsystems develop as a chain of direct interactions.

If the action and reaction patterns develop on the external diagonals, they develop as indirect interactions and evidence some difference emerging from the regular operating evolution. The chain of direct interactions is interrupted: the external diagonal puts into evidence the emergence of indirect interactions and the bypass of the functions of some subsystems.

If we give the operator the possibility to interact with the different subsystems, we give him the possibility to choose the level of interaction with them. The path of interactions between the operator and the subsystem is depicted by a column action pattern $H_C(1)$ or a row reaction pattern $H_R(1)$.

This extreme difference from the normal diagonal operating pattern can be seen as the critical or emergency operating condition (operators "planned" to operate in critical conditions).

It is evident that the topological matrix emerging from the analysis of the decomposition process implicates a passive attitude to find the possible relation patterns. On the contrary, if we impose the relation patterns as design requisites in defining the behavior and feature of the control systems, we assume an active attitude. Then we design the control subsystems and the interaction with the plant in order to guarantee the correct interactions between the subsystems, with the goal of controlling the plant not only in normal operating conditions (a continue direct interactions chain), but in every critical operating condition as well. This means that if we want that the operator interact with a subsystem, we must design the apparatuses to make this type of interactions really possible and applicable. In other words we must design and build the control functions in order to front all the levels of degraded operating conditions.

8. REMARKS

Some remarks connected with the decomposition process, interactions, emergence of hierarchical structures and their applications, are necessary to correctly define the status of our exposition. In this way, the following points may also be issues for future improvements.

- When the level of interaction is different from the hierarchical level. The level of interaction is connected with two specific subsystems and is expressed by the number of intermediate subsystems involved in building the relations that constitute the closed loop of interactions.
  In the other case the hierarchical level is connected with the complete relations pattern of the specific decomposition. In this pattern the relations may configure reciprocally in different and, more or less, regular pictures. The regularity of the pictures can be interpreted as the emergence of a specific hierarchy.
- The possibility of finding the "natural" decomposition is connected with the possible implications due to the different relation patterns generated by the different ordering configuration of the subsystems. It's possible to consider how the emergence of hierarchical structures is evidenced by the "right" ordering structure.
- An additional step in specifying the relation between heterogeneous systems is necessary to find a way to manage the different relation types. For example mass, energy and information can be used and it is

necessary to identify the correct laws to understand and use the specific transformations.

- Can the decomposition add information to the knowledge of the system? Yes, since the hierarchical structures emerging from the decomposition process give visibility to new information.
- Developing and completing a mathematical description of the decomposition process and of all their implications, should give the theoretical framework for the methodology and contribute to answering the open problems.

9. CONCLUSION

We presented some considerations about key points related to the interaction between subsystems. The choice to consider the relations level, without specifying the structure of each relation, enables to apply the decomposition process to systems and subsystems that are heterogeneous. In this way considerations and approaches are very general and the concepts of relation, action, reaction and interaction are used. They are meaningful in a very general context.

Maintaining the analysis on the relation level, we can evidence the possibility to consider the interactions between heterogeneous systems. In the examples introduced we considered the role of the operator as a synonym of "active observer" or "human factor". When the relations are specified and explicated, they assume the appropriate mathematical or physical form.

The previous remarks evidence how many open problems are present and how they can be possible paths for new research lines, especially if we have the goal to formalize mathematically the description of this methodology.

Making explicit and formalizing the relations and the interactions between heterogeneous systems may require the development of specific new modeling approaches.

Again, the possibility to choose the decomposition more adequate for studying a particular system means "to find" the decomposition more "interesting" and "economic" in connection with our problem. The ability to observe more general structures connected with interactions levels may help to see the emergence of hierarchical structures not visible in more traditional representations.

REFERENCES

Abram, M. R., 2002, Decomposition of Systems, in: *Emergence in Complex, Cognitive, Social and Biological Systems*, G. Minati and E. Pessa, eds., Kluwer Academic, New York, pp. 103-116.

Godsil, C., and Royle, G., 2001, *Algebraic Graph Theory*, Springer, New York.

Klir, G. J., 1991, *Facets of Systems Science*, Plenum Press, New York.

Mesarovic, M. D., and Takahara, Y., 1989, *Abstract System Theory*, Springer, Berlin.

Minati, G., and Abram, M. R., 2003, Sistemi e sistemica, *Rivista AEI* **90**(1/2):41-50.

TOWARDS A SYSTEMIC APPROACH TO ARCHITECTURE

Valerio Di Battista
Politecnico di Milano - Dipartimento Building Environment Science and Technology – BEST

Abstract: The historical difficulty in defining architecture corresponds to the complexity and variety of the actions implied, to their multiple interests and meanings, to the different knowledge and theories involved and to its great functional and cultural implications for human life and society. Vitruvius gave a notion of architecture as the emerging system of the main connections of *firmitas*, *utilitas* and *venustas*. A more recent and flexible definition is William Morris's, who conceived architecture as something regarding "all the signs that mankind leaves on the Earth, except pure desert". Today we could agreed on a definition of architecture as *a whole of artifacts and signs that establish and define the human settlement*. To explore its dimensions, performances, multiple values, we need a systemic approach allowing us to recognize and act more consciously in the whole of its variables.

Key words: architecture; project; cognitive system.

1. A DEFINITION OF ARCHITECTURE: PROBLEMS, REFERENCES, HYPOTHESES

The etymology of the word *architecture*, from the Latin term *architectura* which, in turn, comes from the Greek *arkitecton*, identifies, at the beginning, an activity that "nascitur ex fabrica et ratiocinatione"[10]: that is, putting together building practice and theory. This notion has been developed in different languages, often interlacing various meanings, such as *dwelling* and *building*, as well as *structuring*, *ordering* and *measuring*.

This historical difficulty in definition and etymology corresponds to the complexity and variety of the actions implied, to their multiple interests and

[10] "it comes from practice and ratiocination", Vitruvius, De Architectura, I, 1.

meanings, to the different knowledge and theories involved and to the great moment of the functional and cultural implications of architecture, in every place and dimension of human life and society.

Francesco Milizia (1781) says that "architecture may be called twin-sister to agriculture, since to hunger, for which man gave himself to agriculture, must be also connected the need for a shelter, whence architecture came"; but we also find: "construire, pour l'architecte, c'est employer les matériaux en raison de leur qualities et de leur nature propre, avec l'idée préconcue de satisfaire à un besoin par les moyens les plus simplex et les plus solides" (Viollet-le-Duc), or – poetically – "l'architecture est le jeu savant, correcte et magnifique des volumes assembles sous le soleil", but also "the Parthenon is a product of selection applied to a standard. Architecture acts on standards. Standards are a matter of logic, of analysis, of painstaking study; they are established upon well set problem. Research definitively settles the standard" (Le Corbusier). These are just a few examples of the available definitions; Bruno Zevi (1958), even though dwelling on architecture as "the art of space", suggests to distinguish among "cultural, psychological and symbolic definitions"; "functional and technical definitions" and "linguistic definitions".

I believe that architecture is so difficult to define because of the many elements it implies; many and complex are the needs that prompt it and its field is extraordinarily wide, regarding actions, reasons, implementation, knowledge, emotions, symbols, values. We should remember the ancient, masterly Vitruvian notion of architecture – that has been a common notion for many centuries, up to the Industrial age – where it could be perceived as the emerging system of the main connections of *firmitas*, *utilitas* and *venustas*. *Firmitas* means steadiness and long standing; *venustas* means beauty, order and representativeness; *utilitas* means serviceability, good performance: all these qualities are vital to good architecture.

I tried to find, in the recent tradition of architecture, a more flexible definition, that could go beyond the conceptual limit of physical space as a separate entity; I think I have found it in what William Morris said of architecture as something regarding "all the signs that mankind leaves on the Earth, except pure desert". This definition anticipates the notion of material culture and recognizes in every building a *sign* of human action; it encompasses every construction, from a single artifact to the whole landscape, as a product of the human activity, in opposition to a place without any human action: that is, pure desert. The word *sign* underlines the load of communication, of human information embodied in the building production, it indicates the material culture references that can be found in all artifacts or group of products; it also points out information about

materials, workings, uses, processes, rules that have been employed in the production.

The anthropic environment (architecture = all signs put on Earth by mankind) can always be described by the complex connections between artifice (artificial environment) and nature – natural environment. It is neither the relationship between "*imitatio naturae*" (an artifact that imitates nature) and "uncontaminated" nature, nor between an artificial thing and its natural background. It is, rather, a completely different "world", where human actions have deeply modified, used, controlled some of the natural conditions to build up, often in a very long time, a different system, a different world that we can define "built environment".

If we accept this meaning of the word *Architecture*, i.e. "the whole built environment", then we can recognize information and meaning in every product of mankind. All these products refer to connections between people and things: survival (protection, comfort), business (actions done, hosted, exchanged), symbolic value (belonging, identity) or other multiple values: emotional, religious, cultural, familiar, social, material (use value, exchange value, economic value). This system of human signs is produced by and for the construction of our settlements and it is intrinsic to the places where we live. To these places, Morris confronted a non-artificial place, i.e., the desert, as the absence of human traces: nature without any sign of human presence.

This idea, suited to nineteen-century culture, cannot satisfy us any more. Today no place is totally void of human presence; even where it is only temporary (sky, sea, desert, polar regions) it is more and more frequent and organized. We do not just deposit signs, because our artifacts bring modifications, alterations, emissions; these things do not indicate "excluded" places, but rather "different" places that reach any faraway spot on our planet. We reach anywhere as *direct observers*, either continuous and systematic– with satellite monitoring – or temporary – as part of the growing tourist flows. We often are *indirect observers* as well, by recorded images; we are deep in a multiple network of observation – both scientific and functional - for multiple purposes (geographical or military surveys, mineral resources, communication, transport ...); we enjoy wider and growing opportunities of knowing (thus generating interests, memory, emotions, amazement, fear ...) and of committing to them multiple meanings. We can view, visit, enjoy, read, know these many places, differently shaped by our actions. We can distinguish, among them, either those with prevailing "natural" characteristics (shapes, materials, light, colors, movements of waves, haze, clouds ...) or those with prevailing "artificial" features, deeply marked by human culture and history.

Even places where "natural" features prevail, often show signs – feeble as they may be – of human action; we should decide whether they all belong

to architecture: is a temporary connection (a ship against the skyline, a parked car) or an occasional presence (a wreck in a desert or in a sea depth) enough to define a built environment? How long and strong must a relation between artifact and place be, to define such a vision? There is no doubt that an archaeological site in a uninhabited place characterizes the place as "built environment": a trace of human activity, although feeble and remote, marks a place with a permanent result. We accept as architecture every artifact that establishes relations with a place, becoming a part of it; these relations may be quite settled, but also undergo many changes; there are long term relations and temporary events that leave marks and prints, that may be diversely irreversible (absolute irreversibility is impossible).

We have evoked a system of connections between artifacts and place, that could refer to *venustas* (signs and their meaning), *firmitas* (steadiness, continuity, duration) and *utilitas* (satisfaction of a need).

This vision also implies that every architecture becomes historical and perceivable when different settlements structure (and in turn are structured by) social forms, be they simple or complex; these forms memorize, preserve and stock their past marks, so that they become built environment. The built environment is a whole of works (buildings, walls, fences, roads and bridges, shelters, canals and terraces...), physical elements linked together in different forms, depending from a lot of local characteristics and conditions: land, geography and climate; requirements and characteristics of the inhabitants, local resources (materials, energy, knowledge). All these connections, that determine architecture as an artifact that cannot be separated from its context, underline that the difficulties and contradictions in its definitions come mainly from the overwhelming importance that has traditionally been given to its exceptional products (the *monuments*, as we often call them), rather than to the rich fabric of signs and actions generated from all the connections between artifacts and natural environment. Today, we find it is very difficult to think of architecture as a whole; we see a split image, through a number of different disciplines and sciences: the science of *firmitas*, the culture of form and function (*utilitas*), the more recent proud calls for autonomy by the culture of *venustas*. All of them are partial points of view, and some of them seem to rely upon the idea of the architectural project as something so complex to be committed to intuition, to the pursuit of a mythical sublime vision, rather than recognize in it a territory to be explored with new instruments.

For this reason, I think it could be useful both to retrieve a possible link with a millennial tradition and to look at the new problems with instruments fit for the multi-systemic processes and the multiple observation that we so badly need when architecture is intended as a *system of systems* (*places*, *signs*, *performances*) *of the built environment*.

If we agreed on this tentative definition of architecture: "A whole of artifacts and signs that establish and define the human settlement", we could try to explore its dimensions, performances, multiple values; we could conceive new means of description and allow us to recognize and act more consciously in the whole of variables that build up the multi-layer realm of architecture.

2. BUILT ENVIRONMENT, SETTLEMENT ENVIRONMENT, SETTLEMENT SYSTEM[11]

Every anthropic territory can be described by a physical system, a social system, an economical system (built environment, settlement environment); it is a place where one can find generating inputs and generated outputs, inside connections and outside connections. This *settlement system* takes different dimensions, borders, characters according to *observation*. It is up to the observation system (disciplines, for instance) to select the description levels; through an approach that is consistent with its cognitive system, it defines the boundaries necessary and sufficient to the needed observation. At the same time, the characteristics of the observed system itself – such as identities and recognized values – do suggest different ways of observation; these, in turn, can highlight unsuspected relations that are fed back as generators. This settlement system can be defined as an open and dynamic relationship between observing system and observed system, and it individuates the context and meaning of architecture.

Settlements seem to be described by geography (De Matteis, 1995) at a descriptive level (where things happen), a syntactic level (how things happen), a symbolic level. But geography cannot trace the anthropic processes in their diachronic development, where artifacts stock up over time according to land use, generated by local cognitive and social conditions and, in turn, become generators of further means of transformation of the natural and built environment and of new social and cognitive conditions.

Every settlement system – be it simple or complex – can be usefully described by geography, but its complex connections are so multiple – connecting issues that are anthropological, social, cultural, religious, symbolic, political, to resources, business, commerce etc. – that any description needs the contribution of many different disciplines and observations. What's more, in a settlement system it is very difficult to trace any cause-effect connection that is steady, linear and long-standing. A settlement seems to allow description as a system of signs; that is, a

[11] Di Battista Valerio (1992).

reductive but useful representation of a settlement system could pick out the connections that are most significant in a given situation; for instance, among physical system, social system, economic system. Anyway, this system of multiple connections seems to admit only some partial form of government and self-government.

Architecture is representation, system of signs, physical and organizational connotation of the place: it can be read as the visible emergence of the very complex weave that is a human settlement.

Anthropic physical systems: the built environment. An anthropic physical system could be described first by the way land is used and built upon. The built environment shows networks and artifacts, made over time by the settled population; all these human works satisfy needs and give performances which, in turn and over time, further modify internal and external connections in the system. Some of the artifacts seem to proceed rather by accumulation (the networks, for instance), than by life-cycles, as many utilitarian buildings do. Some of them become long-duration structures that survive their original function an meaning, some others live and die, in a progressive substitution process that is very often casual.

The network systems. The networks system encompasses all the structures that connect functions and places; they are built by stocking over time. It also comprises facilities needed to distribute such things as water, power, sewage, communication etc. Type, dimension and efficiency of the network system generate constrains and freedoms; the self-regulating capability of network systems depends upon the capability of self-regulation of the whole settlement (relationship among network management and local authorities).

The system of buildings. Buildings are our artifacts more familiar and important as symbols, being a system of spaces made to give us shelter and protection, and to host our activities. These activities, and all people performing them, need to be protected and safe, to enjoy a comfortable environment, to act in spaces that are suitable for dimensions, shape, connections etc. Every building is characterized by an outer envelope (walls, roofs, windows etc.), and defines spaces and social meanings according to its use; uses may either vary over time, causing the building to become obsolete, or cease at all; thus a building may be either reconverted or demolished; a refurbishment often changes not only the physical character of a building, but its complex meanings as well.

Every building, be it a part of a thick urban fabric or a secluded one, has always a tight connection with the environment it belongs to, and contributes to its settlement; every variation in its shape, frame, use, also modifies its context, and every context variation modifies its social, economic, cultural and symbolic values. The relationship between architecture and context

corresponds to the one between a word and a sentence inside a linguistic environment; architecture may lose its meaning, or gain multiple purely potential ones, when taken outside its cognitive and temporal settlement. The anthropic system is a historically determined product of human presence; this presence is always characterized by social structures that, together with geographic characteristics, gives shape to settlements; every social system interacts with its settlement at every scale and through many and complex connections. Every social system represents itself by its most important buildings, which are dedicated to special functions. The built environment shows different levels of "thickness", relations with open spaces, peculiarities in shapes, volume and building characteristics; these differences come from uses, opportunities, means of self-representation, restraints or freedom of action that belong to the settled social groups, their activities, customs and traditions... These cultural schemes also vary over time, nevertheless they may have a strong or weak inertia: both change and permanence happen all the time in the built environment, according to varied and heterogeneous causes. Signs and meanings vary over time, in different cycles and at different speeds, in a sort of continuous symphony connected to memory, intentions, expectations that, in our case, builds up an environment of solid configurations. This environment represents the life-cycle of the system itself: it confirms its identity and acknowledged values by means of its own long duration; it recognizes its new symbols and updates its language by means of continuous change.

Every built artifact is conceived, made and managed in its life-time according to the processes that happen in its settlement system; in turn, every building interferes – by permanence or change – with the local characteristics of the system itself.

The simplified description of a settlement system through the interactions of the three sub-systems – physical, social, economical – indicates a great number of connections inside every human settlement. In the physical one, flows and exchange of energy/matter; in the social one, flows and exchange of information, values, rules; in the economic one, flows and exchange of resources, work, assets etc. Input and output flows are many and various, according to local characteristic and generate a great number of combinations. Regarding architecture, we are especially concerned with energy, rules, representation, symbolic values, economic values, use values.

The interaction of these many variables and combinations that define a settlement system, may be conceived as a continuous, non-linear deposition of complexity over time, corresponding to a growing complexity of cognitive space. Thus, we can imagine an evolution of our species: from acting in a physical space, to acting in a cognitive space strongly conditioned by physical space, to acting in different interconnected cognitive spaces (or

in different portions of time and space), which, in turn, can modify physical space and may also coincide with it. The space of settlement, however, interprets, materializes, interacts with, confirms – by redundancy and convergence – cognitive systems, giving them conditions of temporary stability. Yet, as the processes that generate settlements are open and dynamic, they generate as well – over time – conditions of inconsistency with the cognitive systems themselves; this possible occurrence can be resolved at a higher level in emergence processes.

Architecture organizes and represents the settlement system; it interprets, materializes, interacts with and confirms the references of cognitive systems, and projects (foresees) and builds "coherent occurrences" (steadiness, confirmation) and "incoherent occurrences" (emergence) in the settlement itself. Architecture operates in the interactions between mankind and natural environment with "coherent actions" (communication; consistent changes; confirmation of symbols and meaning) and "incoherent actions" (casual changes, inconsistent changes, new symbols and meanings). Coherent actions are usually controlled by rules and laws that guarantee stability to the system (conditions of identity and acknowledged values); incoherent actions generally derive from a break in the cognitive references (breaking the paradigm) or from the action of "implicit projects" (Di Battista, 1988). These are the result of multiple actions by different subjects who operate all together without any or with very weak connections and have different – sometimes conflicting – interests, knowledge, codes, objectives. Implicit projects always act in the crack and gaps of a rule system; they often succeed, according to the freedom allowed by the settlement system.

Perhaps, the possible virtuous connections of this project, in its probable ways of organization and representation, could identify, today, the boundaries of architecture that, with or without architects, encompass "the whole of artifacts and signs that establish and define the human settlement".

REFERENCES

De Matteis, G., 1995, *Progetto Implicito*, F. Angeli, Milano.

Di Battista, Valerio, 1988, La concezione sistemica e prestazionale nel progetto di recupero, *Recuperare* **35**:404-405.

Di Battista Valerio, 1992, Le discipline del costruito e il problema della continuità, in: *Tecnologia della Costruzione*, G. Ciribini, ed., NIS, Roma.

Le Corbusier, 1929, *Oeuvre Complete*, Zurich.

Milizia, Francesco, 1781, *Principij di Architettura Civile*.

Viollet-le-Duc, Eugène-Emmanuel, 1863-1872, *Entretiens*.

Vitruvius, *De Architectura*, I, 1.

Zevi, Bruno, 1958, entry "Architettura", in: *E.U.A.*, Sansoni, Firenze.

MUSIC, EMERGENCE AND PEDAGOGICAL PROCESS

Emanuela Pietrocini
Accademia Angelica Costantiniana, Centro "Anna Comneno"
Piazza A. Tosti 4, Roma RM, Italy
http://www.accademiacostantiniana.org, email: emanuela.pietrocini@libero.it
AIRS – Italian Systems Society, Milan, Italy

Abstract: Music presents features typical of complex systems, whether for the multiple aspects it contains or the type of connections it establishes between systems that are seemingly far apart in terms of context, problematic and local characteristics. Actually, in music it is detected the necessary persistence and coexistence of contrasting or apparently irreconcilable elements whose interaction gives rise to what we call "beauty"; this can be more accurately defined, by way of complexity, as an emergent property of artistic production. In this sense, music can help us to redefine "cognitive paths" in virtue of its profound ability to represent and make emergent cognitive processes. Perception, representation, abstraction, creativity and non-linearity are, among the emergent properties of the music-system, those which are most consistent with the process of learning. A didactics of music based on complexity as methodological point of reference shapes the pedagogical process as an interaction in which teacher and student are involved in a reciprocal relationship. From the superposition of their roles in real experience and from the relational aspects, a form of self-defined learning process arises taking on the character of emergent property of the system.

Key words: music; emergence; pedagogical process; complexity.

PRÆLUDIUM

"Music, like the soul, is a harmony of contrasts, the unification of many and the consonance of dissonance."
Philolaus of Croton, ca. Fifth century BCE

1. INTRODUCTION

Complexity and emergence are often revealed in the perception of a fleeting dissonance, in the sudden awareness of a contradiction.

There is no error or incoherence, only a particular detail that, for an instant, stands out from the marked path.

This is a particular form of implicit learning (Bechara, Damasio and Damasio, 2000) that occurs when one is before a work of art: there's an emotive exchange, a sense of wonder, of astonishment. In essence, it predisposes one to a conscious perception.

This probably happens because art manifests itself through the representation of a unified experience, of a space-time bond between the corporeal self, consciousness and the mind (Solms and Turnbull, 2002).

The creation of art necessarily contains, in principle, the persistence and coexistence of contrasting or seemingly irreconcilable elements whose interaction gives rise to what we call "beauty"; this can be more accurately defined, by way of complexity, as an emergent property of artistic production.

The occurrence of processes of emergence is the key property of complex systems. Processes of emergence are based on the fundamental role of the observer able to realize them by using its cognitive models (Baas, 1997; Pessa, 2002).

Music contains and represents aspects that are typical of complex systems: openness, fuzzy boundaries, non-linear relationships among elements, and behavioral characteristics such as chaotic, adaptive, anticipatory behavior (Flood and Carson, 1988).

From this point of view, it's possible to understand the nature of the musical phenomenon by way of the link between the various aspects that characterize it, thus overcoming the limits created by an epistemological reading.

These connections can involve vastly different domains, systems that are apparently far apart from one another in terms of context, problematic and local characteristics. The functional link between them can be defined in terms of coherence: the emergent properties of individual systems, or parts

of them, are shared. From this standpoint, the observer's role is enhanced by the property of identifying *coherence* in the network of connections between systems; one could say that this very characteristic develops as emergence and contributes towards defining a higher level of abstraction in which there is awareness and perception of the unity of the whole. Emergence, in fact, consists in recognizing coherence in processes that were not initially recognized as such by the observer-listener.

The crucial point of this discussion is developed around two basic considerations:

1. Music contains and involves multiple aspects that, for complexity, can be likened to elements of the development of cognitive processing. "Thinking musically," means, in fact, having an integrated experience of the perceived world in every moment of the temporal continuum; this is brought about by an alternative source of information that involves both cognitive and affective aspects (Solms and Turnbull, 2002) and represents a unitary image (i.e. trans-disciplinary, in the systemic sense. Recall that in the trans-disciplinary approach, distinguished from the inter-disciplinary one, when considering systemic properties are considered per se rather than in reference to disciplinary contexts) of the contents, meanings and forms.
2. Music can help us redefine the pedagogical approach since it induces processes activating a form of self-defined learning, that, due to the cognitive style, modality of approach and attractors of interest, are consistent with the learner's personality (Ferreiro and Teberosky, 1979).

Learning assumes the aspect of an emergent property of a system that takes complexity as its methodological reference.

In this article I'll seek to highlight certain features of a pedagogical process through which the *music hour* and *piano lesson* may be considered as integrated in the global process of learning. The first part takes a close look at the systemic aspects linked to musical action and thinking and introduces, in an historical-critical context, some of the aesthetic and structural aspects of music that best highlight the themes of complexity; the second part examines its links with cognitive processes of learning.

The point of view presented here is that of a musician —performer and interpreter of early music— and teacher who places her own reflections and experience in the systemic context and sees them enhanced by the added value of awareness.

The references to cognitive processes are not of a scientific nature, but are used to recall certain aspects of the developmental path within the didactic experience.

2. PART ONE: DE MUSICA

2.1 Multiplicity and complexity in the history of music theory

In the past, reflection upon the phenomenon of music focused on the multiplicity of the aspects that characterize it. From time to time, these have been examined using methods of analysis and research that have decisively contributed to the acquisition of valuable tools for understanding creative forms and models.

The area in which we can, perhaps, historically locate a meaningful convergence of the experiences of complexity and at the same time outline its features, is that of esthetic reflection.

It's worth noting that, in speaking of the esthetics of music, we refer to the development of a philosophy of music and set aside the definition that idealist historiographers identified in the theories of Baumgartner, Kant and in pre-Romantic philosophy (Fubini, 1976).

The sources of the esthetics of music, especially those that refer to Ancient philosophy, reveal an attempt at reconciling various aspects, especially as regards the semantic problem.

In Greek thought, for example, music was seen as a factor of civilization. They gave it an educational function as a harmonizing element for all the human faculties. Yet, they also saw it as a dark force, equally capable of raising man to the heights of the gods, or casting him down among the forces of evil (Fubini, 1976).

The myths of Orpheus and Dionysus represent music as a supernatural power, capable of combining the opposing principles apparently underlying all of nature: life and death, good and evil, reason and instinct.

Orpheus is the human hero who changes the course of natural events, crossing the gates to hell with the help of the synergic power of his notes and words (citarodia, 'string instrument'); Dionysus is the god who puts the primordial forces into play, routing reason's control with just the power of his flute (auletica, 'wind instrument').

Pseudo-Plutarch reconsiders the "rational" and ethical aspect of music, attributing the paternity of the auletica and the citarodia to Apollo, thus giving them once again "a perfect dignity in every aspect" (Pseudo-Plutarch, De Musica).

In truth, it's with Pythagoras that music takes a central position due to the metaphysical and cosmogonic concept of *harmony.*

For the Pythagoreans, harmony is above all the unification of opposites, and extends to the universe considered as a whole, thus determining a dynamic principle of order between the opposed forces that govern it.

The concept of harmony is completed with that of number: "Nothing would be comprehensible, neither things nor their relations, if there was not number and its substance... in the soul, this harmonizes all things with perception..." (cf. Stobeo Eclogae I).

Music reveals the deep nature of number and harmony since the relations between sounds, which can be expressed in numbers, can also be included in the model of the same universal harmony.

Hence, the music of spheres, which is produced by stellar bodies rotating in the cosmos according harmonic laws and ratios, is "harmony that the inadequacy of our nature prevents us from hearing" (Porfirio, Diels Kranz 31B129).

The metaphysical concept of music frequently arises in Plato, who notes its ethical and cathartic value. The sense of hearing has a rather secondary role; indeed, it can distract from the true understanding and exercise of music which is and must be a purely intellectual activity.

As it is, he makes a clear distinction between music one hears and music one doesn't hear: the former, understood as *technè*, does not have the standing of science and can be likened to a technique whose usefulness resides in producing pleasure (cf. Gorgias); the latter "represents divine harmony in mortal actions" (cf. Timeus) and is defined as the highest form of human education. It is in this educational function of music that one can discern a convergence between the two opposed concepts: in Laws as in the Republic, Plato considers the possibility that music may pass from the purely intelligible to sensible reality on the basis of its normative and ethical nature.

It's no coincidence that Aristotle chooses book VIII of Politics to introduce his own discussion of music, counting it among the most important subjects in education since it is suitable for an elitist practice of *otium*, activity "worthy of a free human being" (cf. Politics VIII).

In Aristotle, we find a definition of pleasure as a factor as mush organically connected to the musical function, as the affinity between sound and the world of emotions. The relation with ethos is by nature formal and indirect and based on the concept of motion, which implies an idea of order and, ultimately, harmony.

The soul experiences pleasure in ordered motion since order conforms to nature and music reproduces it in the most varied way (Aristotle, Problems, problem 38).

In this sense, the Pythagorean concept joins the Platonic discipline of ethos in an eclectic synthesis that accentuates the psychological and empirical aspects of the musical fact.

Ancient music theory was able to grasp and describe all the principal aspects of music, highlighting the relations it establishes between the various systems of representation of reality.

We will see below how the structural characteristics of the composition and evolution of musical language draw on the same references and trace out procedural models and effects typical of complexity.

2.2 Composition, execution, improvisation

All forms of human expression (language, music, painting...) can be defined as symbolic, and hence as falling within the domain of semiology, to the extent that one can identify in it the following three dimensions (Molino, 1982):

1. The *poietic process*, that is, the set of strategies by which an original product comes to exist through a creative act.
2. The material object, or rather, the tangible product in symbolic form that exists only when executed or perceived (the object's immanent organization).
3. The *esthesic process*, that is, the set of strategies set into action by the perception of the product.

The references to theories of communication are nearly too obvious and may not be shared if taken as foundational principles of a deterministic poetics of music.

It remains that the musical product, understood as an artistic fact, is the result of a particularly rigorous process of construction which guarantees it recognizability and effective expressiveness.

Likewise, there is no musical phenomenon without a referent that accomplishes an esthesic analysis of the work.

In systemic terms, we could say that the overlap of functions of the composer/performer and listener shapes the role of the observer with respect to the musical phenomenon; in fact, the discovery of emergent properties occurs through processes that can be compared in terms of complexity and which converge towards defining a unitary representation of the work irrespective of the analytic model and the modalities of interaction.

The history and evolution of language and musical codes underscore an on-going search for formal models that define the creative process: "[T]he movements of music, regardless of the period and geographical location, are guided by a kind of order of things...that appears beyond philosophical speculation" (Brailoiu, 1953).

Over and above the cultural differences, the substance of the *cum ponere* always took shape through the modelization of processes: there is in fact a basic operative level that organizes the sonorous material in ordered sequences according to the variables of pitch, intensity, duration, and timbre in the discrete space-time that's represented, for example, by a musical

phrase. The subsequent level is identified by systems theories that involve syntactic-like structural parameters such as melody, harmony and rhythm, defining them by modalities of relation based on models that can be mathematically described: modality, tonality and seriality are only some examples that belong to Western culture.

Even "expressive" elements, inherent to the music's phraseology, were codified and placed in sequence, just as in *dispositio* oratory (cf. Cicero, De Oratore). In fact, in any musical passage one's able to find a proposal and response, a thesis and antithesis, a beginning and conclusion.

The musical equivalents of the rhetorical forms are particularly clear in the early Barocco's compositions "in sections", such as the Toccata (Raschl, 1977): the repetition (*anaphora*) marks the need to highlight a topical moment and "gives greater force to expression" (cf. Beethoven, 1885, Treatise on Harmony and Composition); the abrupt, unexpected interruption caused by a pause (*abruptio*) irresistibly attracts attention and underscores the expressive elements just discussed; the dynamic emphasis upon a stress (*exclamatio*), very effectively calls up the "surge of affections" (Caccini, 1614). The use of such expressive forms can contextually assume a didactic and esthetic value from the moment in which a feeling arises by way of analogy.

Despite the clear linearity of the procedural framework, at the moment of the creative elaboration there inevitably appear completely unforeseen variables and solutions, modifications and chance events that, although of little importance, determine for that context and time irreversible, evolving consequences.

Paradoxically, this trait is more clearly seen in the so-called theoretic works of famous authors such as Johann Sebastian Bach: the Art of the Fugue and the Musical Offering, composed according to a more rigorous contrapuntal style (Fig. 1), bear the traits of a definitive, unalterable fading of the tonal system in favor of a "serial" approach that anticipates the compositional models and poetics of the nineteenth century.

In general, the evolution of formal and structural aspects in compositional techniques shows how the tendency towards a system's stabilization proceeds alongside the gradual loss of its characteristic parameters (Bruno-Pietrocini, 2003).

This is due in good part to the phenomenological characteristics of the variation. This process consists in transforming, through various devices, a thematic element consisting of a cell or musical phrase (Fig. 2).

Thematis Regii
Elaborationes Canonicae

1 - 5. Canones diversi super Thema Regium

Canon 1. a 2 *cancrizans*

Canon 2. a 2 Violini in unisono

Canon 3. a 2 per Motum contrarium

Thema

Figure 1. J. S. Bach: Musikalisches Opfer BWV 1079.

The operative modules can be classified according to the structural aspect upon which they intervene; rhythmic variations, for example, modify the characteristics of note duration and can alter the meter, stress and tempo.

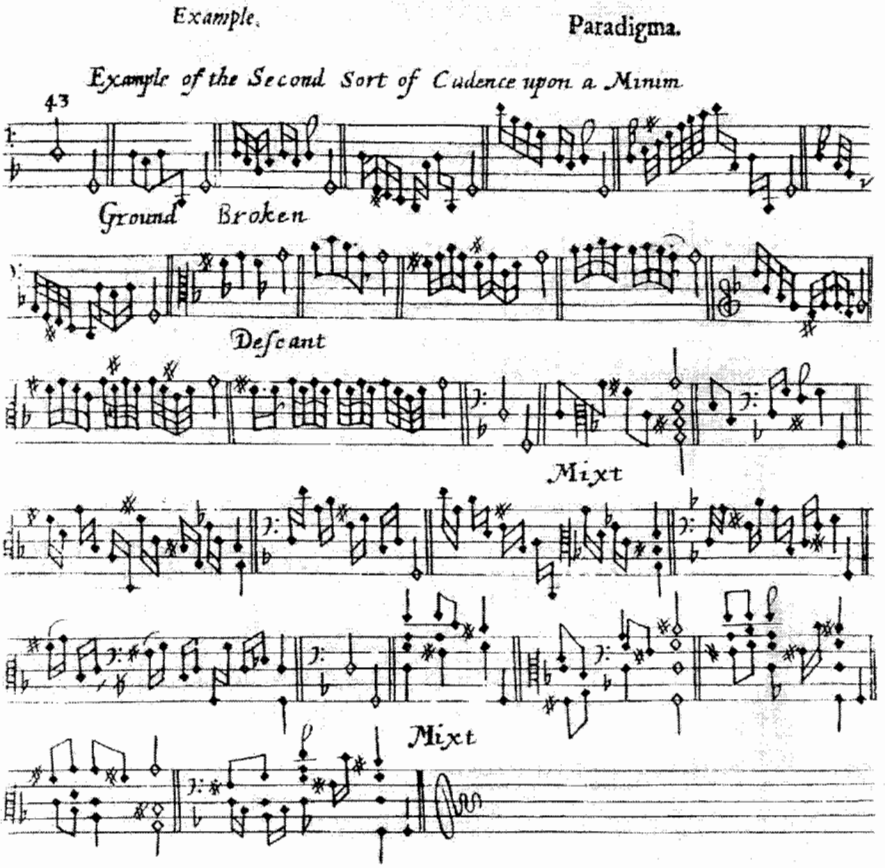


Figure 2. C. Simpson: The Division-Viol or the Art of playing extempore upon a ground (1667).

Similarly, the melodic transformation makes use of techniques such as *inversion* and *retrogradation* (Fig. 1) that, in changing the relationships of succession between notes through inversion or retrogradation, define a different serial identification of the phrase.

In improvisation, which provides for extemporaneous forms of variation immanent to the execution, one can clearly note the lack of what could be called stable equilibrium, without this condition having the least impact upon the coherence, intelligibility or artistic quality of the product. Improvisation, which is reminiscent of the genesis of the creative musical act, exemplifies the model of a system subject to continuous modifications.

It may be worthwhile to consider in detail some of the specific aspects with reference to the performance practice of Basso Continuo (Thorough-

Bass), a form of extemporaneous accompaniment that was particularly popular in European musical culture from the end of the sixteenth century to the early 1800s.

The need to construct a structure supporting a melodic line, a legacy of polyphony's structural contraction, takes shape through the definition of a bass, obbligato or free, and a harmonic-contrapuntal realization performed by a polyvocal instrument (Del Sordo, 1996).

The improvisational character of the Continuo, strictly fashioned by the laws of the theoretic and stylistic system of reference, can give rise to a succession of variations that may substantially modify important aspects of the work being performed: the conducting of the parts, agogics, dynamics and rhythm contribute, as unpredictable factors, to the elaboration of an original, coherent and meaningful artistic product.

Similar characteristics are present in more recent musical production, in particular jazz, whose linguistic matrix can be traced to forms of collective improvisation. The system's codification, beginning already in the early 1920s, reveals highly rigorous forms and structures that lose, however, any static connotation from the moment that the extemporaneous solo begins to gradually modify the path's parameters to the point of making it nearly unrecognizable.

The random and unpredictable tempo that characterizes this genre of musical production describes a *local disorder*. However, if we consider the musical event along the only inalienable dimension, space-time, it regains a *global stability*.

This is an emergent property of the system that defines a non-linear behavior in the relations between elements and gives rise to a "double loop" that modifies not only the parameters, but also the rules, thus generating new structures (Minati, 2001).

2.3 Perception, emotion, cognitivity

"Music is a strange thing. I'd dare say it's a miracle, since it's halfway between thought and phenomenon ... a kind of nebulous mediator that is both like and unlike each of the things it mediates ... we don't know what music is" (H. Heine).

These words present once again the theme of complexity by way of two of the most problematic aspects of musical perception: the cognitive and expressive-emotive elements.

Music can determine an authentic perceptual spectrum that goes from the reception of sensory-auditory data to impressions that, in sensitive indviduals, can be so evocative as to escape any kind of description (Critchley, 1987).

The most direct and widespread approach to listening to music does not require action. And, upon closer examination, not even attention; only acceptance.

It's easy and natural to let oneself be led for endless, enchanting discursions while listening to a delicate mazurka, a rhapsody or simply a beautiful song.

A strophe, ritornello...a melody that returns, a small variation, a new tonality; silence. And then once again the familiar phrase, the main theme that rewards the memory of a recent treasure (Deutsch, 1987), but so stable to be not soon forgotten (Solms and Turnbull, 2002).

Changeable, mysterious and at the same time reassuring, music can pass through us with a sense of levity, and remind us of a state of "fusion" and identification (Critchley, 1987).

It's then that emotions resonate, like an enormous, powerful tuning fork.

What is meant by "emotion"?

One might say that emotions are a sensory modality directed inwards; a kind of "sixth sense" of our conscious existence. They provide information on the current state of the corporeal Self which is compared to the state of the external world, even though they constitute the aspect of consciousness that would remain if all the content deriving from the external world were eliminated.

Knowledge of our inner state, of our nuclear Self, is guaranteed by emotions that can be evoked in an extraordinary way by music. Similarly, the birth by gemmation of new creative processes in composition can be associated with a phenomenon of "horizontal" amplification (redundancy) of emotions.

One can speculate that, when listening to a musical passage, a particular model of relations between inside and outside the self is established.

Actually, a first level of knowledge of music is related to the substantial role of emotions in cognitive processes (Damasio, 1994, 1996) and is made explicit in the intuition that outlines the emergent features of a mental representation of the phenomenon.

Recurrence, contrast and variation, the basic constitutive principles of composition (Bent, 1980) are preceived, regardless of the listener's musical expertise, as elements of reference implying a structure. This condition can in turn raise a need of analysis that goes beyond the esthetic process in itself, but leads to the cognitive investigation and to the meta-cognitive development of the processes (which will be discussed below).

In this sense, we could say that music suggests a form of implicit learning (Bechara, Damasio and Damasio, 2000) that can be extended to various aspects of complexity since it identifies a non-linear method for knowing.

3. PART TWO: DE ITINERE

3.1 The pedagogical process and meta-cognitive methodology

In order to define pedagogy, "one adopts a kind of moving, circular explanation; an explanation in which, in order to understand the phenomenon, one goes from the parts to the whole, and from the whole to the parts" (Morin, 1993).

Pedagogical research is also enriched by complexity and outlines the content and forms of a unifying and transversal project of education.

The reference to the dynamic and circular properties of the pedagogical process seems to be particularly apt for describing the essential elements of a model of interaction involving teachers and students in a reciprocal relationship

If we consider learning as a shared condition of development which is manifested through the stable and generalized modification of behavior on the basis of experience, we're able to notice the characteristics of the active process in the relational system itself and in the coherence of the functions.

One who learns acquires information (input), transforms it, elaborates it, applies it (output) and in the end verifies and checks the suitability of the elaboration and application (Derry, 1990). This process is part of both the student's and teacher's experience: the latter identifies his pedagogical function by in turn "learning" the student's cognitive path.

One may say that the teacher's essential task consists in detecting, activating and managing emergences in the learning system in order to develop a coherent and effective educational strategy.

It is precisely these aspects one refers to in speaking of the self-defined path of learning. The outline below seeks to provide a concise description of the stages and characteristics:

1. The student identifies the centers of interest (attractors) on the basis of his own cognitive needs and on the affective-emotive assonance he establishes (Solms and Turnbull, 2002), following a direction that's consistent with his own personality. The teacher collects the data resulting from observation and uses them to configure a field of reference (integrative background) in which to situate the experiences.
2. The acquisition and re-elaboration of information define an outline of the knowledge path from which emerge skills and strategies. The teacher organizes and indirectly provides opportunities, means and operative tools based on the developmental characteristics he gathers *in itinere*, observing the student's behavior and cognitive style. At the same time, he identifies and reinforces those relations, established between different

aspects and contents, which can possibly lead to new attractors (Benvenuto, 2002).

3. Along the knowledge path, the student consciously learns to manage and control the unfolding of his own cognitive process as well as to generalize the use of strategies. This aspect calls for significant involvement of meta-cognitive skills (which will be outlined below) whose use and practice contribute to the teacher's action.
4. The verification and evaluation of the knowledge path occur contextually in its evolution by way of the superposition of the student's and teacher's analytic function; in terms of complexity, one can say that they essentially consist in the observation and detection of emergences in the system. In this sense, even learning develops as something emerging from the pedagogical process.

The references to meta-cognitive methodology, which may be defined as one of the most interesting and useful developments in the cognitive psychology of education, are necessary in order to better understand certain characteristic aspects of the pedagogical process at issue.

In the first place, going beyond cognition means acquiring the awareness of what one is doing, why one is doing it, how, when it's appropriate to do it, and under what conditions. Moreover, the meta-cognitive approach tends to shape the ability to directly "manage" one's cognitive processes by actively directing them through one's own evaluations and operative indications (Cornoldi and Caponi, 1991)

The teacher working in a meta-cognitive manner contributes to the pedagogical process according to a method that's clearly expressed by the etymological meaning of the term: by "accompanying" the student along the learning process.

Let's take a look at a brief example of how music can identify the integrative background of a learning process.

The example will be given in narrative form, as notes in a teacher's diary, with the intention of leaving the reader the pleasure of interpreting and recognizing the traces of complexity.

The protagonist of this account is a young piano student, whom we'll call C.

In order to better understand the text, a small glossary of technical musical terms has been added at the end of the article.

3.2 The piano lesson

"With which piece do you want to start?"

As almost always happens, C doesn't answer. She never speaks much. She is now nine years old, and has played since she was five.

She simply takes the school-bag and slowly pulls out books and music scores; one by one she arranges them on top of the piano in a small, well-ordered stack. Only when her bag is empty and all the music is in place does she decide and choose: she runs her finger over the scores until she finds the right one, finally pulls it out, opens it, and places it on the stand.

"Ah, Corelli's Saraband. Fine, let's start."

No problems, you can tell she likes this piece; the tempo[1] is correct, the phrasing[2] is fluid and the sound clean: she feels confident... What happened? She made a mistake with the cadence[3] rhythm; perhaps she was distracted... these notes aren't in the score, now she's adding them to the melodic phrase[4]: she neglected to read the full score[5] carefully... she's using a different fingering from the one indicated on the consecutives sixteenth quarter notes[7]... 3-4-3-4 instead of 1-2-3-4, why?... But she doesn't stop, she continues to play with the same confidence, just as at the beginning. I don't want to interrupt her...

"O.K., that was very nice. Can you play the cadence of the first part for me again?"

The same mistake in rhythm... so, it wasn't an accident.

"Why do you play it that way? It seems that the score indicates another rhythm... let's check."

C. looks at me, she doesn't seem at all confused, and then effortlessly and correctly solmizates[8] the rhythmic phrase in question, keeping perfectly to the score.

"There, in the score it's written exactly the way you just read it out. Try to play it again now."

No use: the mistake is still there. And yet she understood the rhythm...

"Do you know you're playing the cadence differently from what's written?"

"I don't know, when I play, it comes out that way..."

While C. answers, I start "to see" the image of the countless Sarabands[9] that I've played... by adopting the antique way of performing[10], all cadences can be performed with a shift of stress, with a change in beat[11] from triple to double, regardless of the annotated rhythm, which remains unchanged. The addition of passing notes[12] in the melodic phrase can now be explained: they're used to "fill in" the empty spaces, just as with the technique of diminution[13]... and the fingering? It's that of the keyboard of the early Barocco, the very period of Corelli... The literature for piano includes

certain "ancient" works but, usually, it's presented in a modern form, and hence with revised and corrected modalities according to a different performance practice. C.'s scores are for piano...

"I really like the way you play this Saraband: do you want to try to play it on the harpsichord, too?..."

4. CONCLUSIONS

Anyone who studies music is immersed in complexity, often without being able to give a name, a reason to the intuition that appears at the margins of consciousness, insistently, like an idée fixe in the Sinfonie Fantastique by Berlioz.

Perhaps it's in the moment in which that particular relation with the other than self—called "teaching"—is established that we're able to grasp the true nature of music: the uncertain, random and unpredictable content of the semantic origin of music finds explanation and acceptance.

The systems approach offers a great opportunity for reflection since it represents music as a unified and trans-disciplinary experience, and makes it accessible as a tool for conceiving and managing knowledge. In this way, it provides music with an additional value, enriched by the discovery that one can situate it next to the sciences in the evolution of the research process.

A SMALL GLOSSARY OF TECHNICAL-MUSICAL TERMS IN ORDER OF APPEARANCE

1. *tempo*: refers to the agogical indications of a musical piece or, more simply, the speed at which it's played (adagio, moderato, allegro etc.)
2. *phrasing*: a way of expressively articulating the execution of a passage, respecting its syntactic structure and the musical discourse.
3. *cadence*: the form of the conclusion of a musical passage or a single phrase.
4. *melodic phrase*: a syntactic part of the musical discourse distinguished by a melodic line, from 4 to 8 beats (corresponding, in spoken language, to a phrase composed of a subject, verbal predicate and complement).
5. *score*: annotated text of a musical composition
6. *fingering*: a means of ordering the articulation and sequence of fingers upon the keys; in modern fingering for keyboards, the fingers of the two hands are symmetrically indicated with the numbers 1 through 5.

7. *consecutives sixteenth quarter notes*: a rhythmic-melodic figuration consisting of four contiguous sounds (e.g. do re mi fa) whose rhythm is distinguished by the subdivision of a given duration into four equal parts.
8. *solmization*: oral exercise of reading music.
9. *saraband*: a dance having a slow and solemn tempo in triple time, that became part of the instrumental literature in the seventeenth century.
10. *performance practice*: method of performing a musical piece on the basis of written indications, stylistic directions and interpretive conventions characteristic of the period in which it was written.
11. *beat*: scansion of the musical time; it's also referred to for identifying the sequence of strong and weak stress in the measures (e.g. double time—March: S-w/S-w; triple time—Waltz: S-w-w/S-w-w).
12. *passing notes*: melodic connecting notes between unlinked chords; they're used to "fill in the gap" between distant notes.
13. *diminution*: in the seventeenth and eighteenth centuries they indicated an improvisational procedure of melodic variation that consisted in the addition of short passing notes and ornamentations along the original line.

REFERENCES

Apel, W., 1967, Geschichte der Orgel und Klaviermusik bis 1700, Kassel: Bärenreiter-Verlag.

Apel, W., 1962, *Die Notation der Polyphonen Musik, 900-1600*, Breitkopf & Härtel Musikverlag, Leipzig.

Aristotele, *Problemi musicali*, Ed. G. Marenghi, Sansoni, Florence, (1957).

Baas, N. A., and Emmeche, C., 1997, On Emergence and Explanation, *Intellectica* **25**(2):67-83, (also published as: the SFI Working Paper 97-02-008, Santa Fe Institute, New Mexico).

Bach, C., Ph, E., 1753, *Versuch über die wahre Art das Clavier zu spielen*, (Italian translation: *L'interpretazione della musica barocca-un saggio di metodo sulla tastiera*, Gabriella Gentili Verona, ed., Edizioni Curci, Milan, 1995).

Bechara, A., Damasio, H., and Damasio, A. R., 2000, Emotion, decision making and the orbitofrontal cortex, *Cerebral Cortex* **10**:295-307.

Bellasich, A., Fadini, E., Leschiutta, S., and Lindley, M., 1984, *Il Clavicembalo*, EDT, Turin.

Bent, I., and Drabkin, W., 1980, *Analysis*, Macmillan Publishers, London, (Italian translation: Annibaldi, C., ed., 1990, *Analisi musicale*, EDT, Turin).

Benvenuto, S., 2002, Caos e mode culturali, *Lettera Internazionale* **73-74**:59-61.

Brailoiu, C., 1960, La vie antérieure, Histoire de la musique, Des origines à J, S, Bach, in: *Enciclopédie de la Pléiade*, vol IX, R. Manuel, ed., Gallimard, Paris, pp. 118-127.

Bruno, G., 2002, Caos: il linguaggio della natura, *Lettera Internazionale* **73-74**:56-58.

Bruno, G., and Pietrocini, E., 2003, Complessità, pensiero non lineare e musica, in: *Arte e Matematica: un Sorprendente Binomio*, E. Rossi, ed., Mathesis Conference Proceedings, Abruzzo Regional Council Presidency, Vasto, April 10-12, 2003.

Caccini, G., 1614, *Le Nuove Musiche et Nuova Maniera di Scriverle*, facsimile of the original at the Florence National Library, Archivium Musicum S.P.E.S, Florence, 1983.

Cornoldi, C., 1990, Metacognitive control processes and memory deficits in poor comprehenders, *Learning Disability Quarterly* **13**.

Cornoldi, C., 1995a, *Matematica e Metacognizione*, Erickson, Trento.

Cornoldi, C., 1995b, *Metacognizione e Apprendimento*, Il Mulino, Bologna.

Cornoldi, C., and Caponi, B., 1991, *Memoria e Metacognizione*, Erickson, Trento.

Critchley, M., 1987, "Esperienze estatiche e sinestetiche durante la percezione musicale." La musica e il cervello, Piccin Nuova Libraria, Padua.

De Beni, R., and Pazzaglia, F., 1995, *Lettura e Metacognizione*, Erickson, Trento.

Del Sordo, F., 1996, *Il Basso Continuo*, Armelin Musica - Edizioni Musicali Euganea, Padua.

Derry, S. J., 1990, Remediating academic difficulties through strategy training: the acquisition of useful knowledge, *Remedial and Special Education* **11**(6).

Deutsch, D., 1987, *Memoria e Attenzione nella Musica. La Musica e il Cervello*, Piccin Nuova Libraria, Padua.

Diels, H., and Kranz, W., 1956-59, *Die Fragmente der Vorsokratiker*, vols, 1-3, Berlin.

Ferreiro, E., and Teberosky, A.,1979, *Los Sistemas de Escritura en el Desarrollo del Niño*, ed., Siglo Veintuno, Cerro del Agua, Mexico.

Flood, R., and Carson, E., 1988, *Dealing with Complexity, An Introduction to the Theory and Application of Systems Science*, Kluwer Academic/Plenum Publishers, New York.

Fubini, E., 1976, *L'estetica Musicale dall'Antichità al Settecento*,Einaudi, Turin.

Minati, G., 2001, *Esseri Collettivi*, Apogeo, Milan.

Molino, J., 1982, Un discours n'est pas vrai ou faux, c'est une construction symbolique, *L'Opinion* (Morocco) **8**(January),15.

Morin, E., 1993, *Introduzione al Pensiero Complesso, Gli Strumenti per Affrontare la Sfida della Complessità*, Sperling e Kupfer, Milan.

Nattiez, J. J., 1977, *Il Discorso Musicale*, Einaudi, Turin.

O'Connor, J., and Mc Dermott, I., 1997, *The Art of Systems Thinking*, (Italian translation: *Il Pensiero Sistemico*, Sperling e Kupfer, Turin, 2003).

Pessa, E., 2000, Cognitive modelling and dynamical systems theory, *La Nuova Critica* **35**:53-93.

Pessa, E., 2002, What is emergence?, in: *Emergence in Complex Cognitive, Social and Biological Systems*, G. Minati and E. Pessa, eds., Kluwer Academic/Plenum Publishers, New York.

Plato, *Opere Complete*, Laterza, Bari, (1971).

Pseudo Plutarco, *Della Musica*, Ed. L. Gamberini, Olschki, Florence, 1979.

Raschl, E., 1977, Die musikalisch-rhetorischen figuren in den veltlichen vokalwerken des Giovanni Felice Sances, Studien zur Musikwissenschaft, Beihefte der Denkmaler Tonkunst den Osterreich, XXVIII, pp.29-103.

Simpson, C., 1667, *The Division-Viol or The Art of playng extempore upon a Ground*, lithographic facsimile of the second edition, J. Curwen & Sons, London.

Solms, M., and Turnbull, O., 2002, The Brain and the inner world; (Italian translation: *Il cervello e il mondo interno*, Raffaello Cortina, Milan, 2004).

Von Bertalanffy, L., 1968, *General Systems Theory*, George Braziller, New York.

INTRINSIC UNCERTAINTY IN THE STUDY OF COMPLEX SYSTEMS: THE CASE OF CHOICE OF ACADEMIC CAREER

Maria Santa Ferretti and Eliano Pessa
Psychology Department, University of Pavia
Piazza Botta 6, 27100 Pavia, Italy

Abstract: Usually the uncertainties associated to modeling complex systems arise from the impossibility of adopting a single model to describe the whole set of possible behaviours of a given system. It is, on the contrary, taken as granted that, once chosen a particular model (and whence renouncing to a complete knowledge about the system itself), every uncertainty should disappear. In this paper we will show, by resorting to an example related to the choice of academic career and to a structural equations modeling, that, even in this case, there is a further intrinsic uncertainty associated to the fact that algorithms used to make previsions give different answers as a function of adopted software, of the algorithm details, and of the degree of precision required. Such a further uncertainty prevents, in principle, from any attempt to reach a complete elimination of uncertainty within the study of complex systems.

Key words: structural equation models; intrinsic uncertainty; choice of academic career; decision making.

1. INTRODUCTION

According to a widespread opinion, social and cognitive systems are qualified as complex. Besides, some believe that, as a consequence of their complexity, systems of this kind cannot be fully described through a unique model, owing to the unavoidable uncertainties encountered in describing them. Notwithstanding the fact that the assessment of correctness of such beliefs be of capital importance for systemics as well for many disciplines (such as economics, psychology, sociology and so on), this appears as a very difficult enterprise, owing to the lack of a rigorous definitions of complexity

suitable for these systems. To this regard, this paper tries to contribute to a better understanding of the problem by discussing the uncertainties which occur in an unavoidable way within a particular case study: the one connected to the choice of the academic career by high school students. These uncertainties could be representative of the ones occurring even in the study of a larger class of systems of this kind and the results presented in this paper could, thus, be useful for setting a general classification scheme relative to the uncertainties to be expected when dealing with other systems.

Before entering into details, we remind that most researchers try to understand the phenomena involved in the study of choice of academic career by resorting to the so-called Socio-cognitive theory (Bandura, 2000), according to which the structuration of cognitive system is a byproduct of the social interaction. An important construct proposed within this theory is the one of self-efficacy, denoting the beliefs of an individual about his own ability in performing a given task.

A number of studies evidenced how Socio-cognitive theory can be applied also to career decision making (Krumboltz, Mitchel and Jones, 1976; Krumboltz, 1979); moreover the construct of self-efficacy has been used to explain the formation of vocational interests and values and the professional growth itself (Lent et al., 1994). According to Socio-cognitive theory these constructs are influenced by some cognitive mediators such as outcome expectancy. The latter, related to a given professional activity, refers to expected and imagined outcomes of given action.

According to this theory, self-efficacy and self-confidence would exert a considerable influence on vocational interests' formation. Individual abilities, as claimed also by Barak (1981), surely influence self-efficacy development, but the latter should constitute the main determinant for vocational interests consolidation. As regards the students, school Self-efficacy is to be identified with the ability of knowing themselves, of being aware of his/her own personal abilities to successfully face up the learning tasks. It includes the control and the cognitive abilities, the determination, the emotional stability and studying habits.

The above assertions could be considered as a synthetic description of a rough and qualitative model of academic career choice. In principle we could agree on the fact that such a model captures only particular features of this process, the latter presumably being highly complex and whence impossible to describe by resorting to a single model. On the other hand, we could expect that this be the only possible cause of the uncertainty encountered when trying to model the process itself, and that, once chosen a particular model between the many possible ones, all uncertainties be removed and the model itself give well defined and unique (even if incomplete) experimental predictions. In this paper we will show that this is

not the case and that, even when a particular model choice has been made, another kind of intrinsic uncertainty enters into play: the one associated to practical computation of experimental predictions. This means that, in the study of a complex system, we must take into account not one but two kinds of different uncertainty sources, which contribute both to prevent from any possibility of an exhaustive description of any system of this sort.

2. THE GENERAL MODEL

The above assumptions can be used to generate a more detailed model of academic career choice, whose general structure is graphically depicted in the Figure 1.

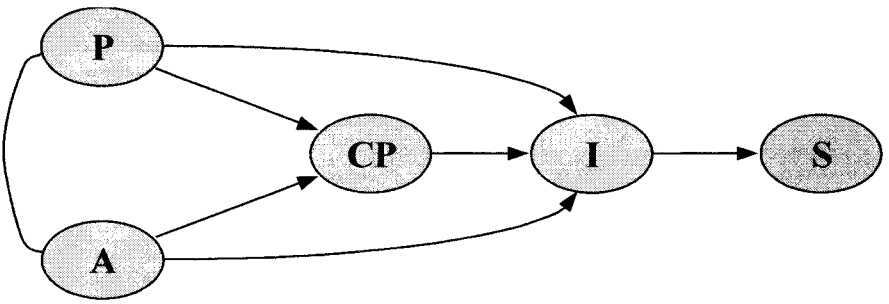


Figure 1. The general model.

The meaning of the symbols is the following: A = school self-efficacy, P = school performance, CP = perceived abilities, I = interest for the academic domain, S = choice (identified with the intention to register at a particular university course).

As we can see from the Figure 1, we hypothesize that school self-efficacy be correlated with school performance. Besides, the performance in the disciplines related to the chosen academic domain should influence both the perceived abilities as well as the interest. However, we expect that the latter be mostly influenced by school self-efficacy.

In order to transform the abstract scheme depicted in the Figure 1 in a complete quantitative model, able to do predictions, we resorted to a technique widely used within social sciences and known as *Structural Equations Modeling* (Jöreskog, 1986). Within the latter the relationships between the different variables (both the observed and the *latent* ones) are interpreted as true cause-effect relationships and expressed through linear multivariate regression equations. The numerical values of the coefficients

of these equations are found through a process of *minimization* of the distance between the variance-covariance matrix derived from a suitable set of experimental data and the variance-covariance matrix predicted on the basis of model equations.

By referring to the quoted literature for further technical details we will limit ourselves to mention that the use of Structural Equations Modeling is based on the following steps:

a) formulation of a general model about the relationships between given variables, such as the one depicted in Figure 1;
b) carrying out of experiments so as to obtain a set of data about the values of some variables contained in the general model;
c) computation of the variance-covariance matrix for the obtained experimental data;
d) starting of an iterative procedure for computing the values of model coefficients in such a way as to minimize the distance between the variance-covariance matrix generated by the model and the one relative to experimental data; such a procedure starts from arbitrarily chosen initial values of model coefficients and changes them at every step according one of the traditional optimization methods, such as, e.g., gradient descent (see standard textbooks such as Gill et al., 1981; Scales, 1985; an excellent and synthetic review is contained in Mizutani and Jang, 1997);
e) once found the model coefficients corresponding to the minimum distance, computation of a number of indices (later specified) measuring the goodness of fit given by the quantitative form of the model so found to the experimental data.

As this technique has been designed to produce the best available (linear) model compatible with existing experimental data, we expect that, once fixed the general model structure and the data set, it should give rise to a unique result, that is the model deriving from these choices. The existence of a plurality of possible models of the same complex system, whence, should depend only on the different possible choices both of model structure and of data set. In the next sections we will show that this expectancy is completely wrong.

3. THE EXPERIMENTAL DATA

The experimental data were consisting in answers to questionnaires given to students of province of Pavia (Italy), enrolled in the last year of high school.

3.1 Participants and Procedure

The total number of students answering to questionnaires was 123. It is to be mentioned that different kinds of high school were represented. Participation in this study was voluntary, and each student had at disposal a time of 2 hours to complete the questionnaires.

3.2 Questionnaires

We used five different questionnaires to gain information about the model variables. They can be shortly described as follows:

1. *School Self-efficacy*
   Measured through the questionnaire "Come sono a scuola?"(Zanetti and Cavallini, 2001), containing 30 items, subdivided into 6 different areas: control abilities, determination, cognitive abilities, studying habits, relational abilities and emotional stability. The self-ratings were expressed on a 4-point Likert scale.
2. *Perceived abilities* (*regarding the personal performance in the future university studies*) Self-made questionnaire related to academic interests, containing 43 items, and built in accordance with the criteria introduced by the Italian Ministry for University, Education and Research (MIUR) in order to characterize the different kinds (from the disciplinary side) of graduate studies. The self-ratings were expressed on a 6-point Likert scale.
3. *Interests towards the five main disciplinary areas individuated by MIUR*
   Self-made questionnaire on academic and professional interests, containing 128 items. The self-ratings were expressed on a 6-point Likert scale.
4. *School performance*
   Measured through the school marks in the first half of school year, obtained in italian language, latin language, history, philosophy, english language, arts, mathematics, physics, sciences.
5. *Choice of a University course*
   Self-made questionnaire in which students were asked whether they decided on future university career. If they indicated that they had made a career choice, they were asked to list the course of study they had selected. If they indicated that they had not yet decided on a career, they were asked to list the courses of study they were considering. Career choice was operationalized by giving to each considered area a score of $1/N$, if the student had taken into consideration N different possible choices.

4. STRUCTURAL EQUATIONS MODELING THROUGH SEPATH SOFTWARE

Neglecting here all aspects related to descriptive statistics of the data set and to internal consistency of the used questionnaires, we will focus our attention only on a first attempt to perform a Structural Equations Modeling by resorting to the computer software SEPATH 6.0. In order to estimate model coefficients this latter used the *Maximum Likelihood* method while the null hypothesis, asserting that the difference between theoretical and observed variance-covariance matrix was due only to chance, was tested through the usual chi-square statistics. The Goodness of fit was evaluated through the following indices (we refer to quoted literature for further details):

- RMSEA (*Root Mean Squared Error of Approximation*)
- RMS (*Root Mean Squared Residual*)
- GFI (*Goodness of Fit Index*)
- AGFI (*Adjusted Goodness of Fit Index*).

The obtained values were: $\chi^2 = 1479.2$ (df = 248; p =.001), RMSEA = .196, RMS = .15, GFI = .50, AGFI = .40. They evidence how the model produced by SEPATH doesn't fit in a satisfactory way the experimental data, as the value of RMSEA is too high (it should be lesser than .05) and the values of GFI and AGFI are too small (they should be both very close to 1).

Such a circumstance forced us to introduce five different choice models, one for each main disciplinary area, all having the same general structure as before, but with model coefficients estimated, for each model, only on the basis of the data related to the students considering, as a choice possibility, the area related to the model under study. Such a strategy was very effective, as, for each one of the five models so individuated, the goodness-of-fit indices RMSEA and RMS assumed very small values, whereas GFI and AGFI were very close to 1.

Without illustrating the single models thus obtained, we will focus here on a particular one, related to scientific area, as its indices evidence a very good fitting of experimental data. The model and its coefficients are represented in Figure 2.

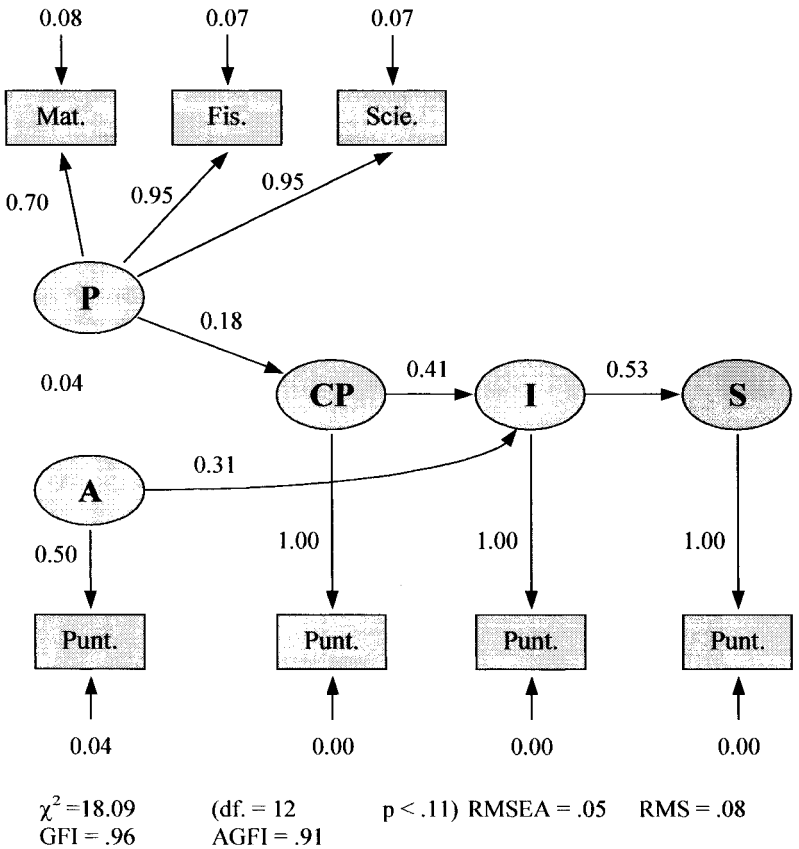
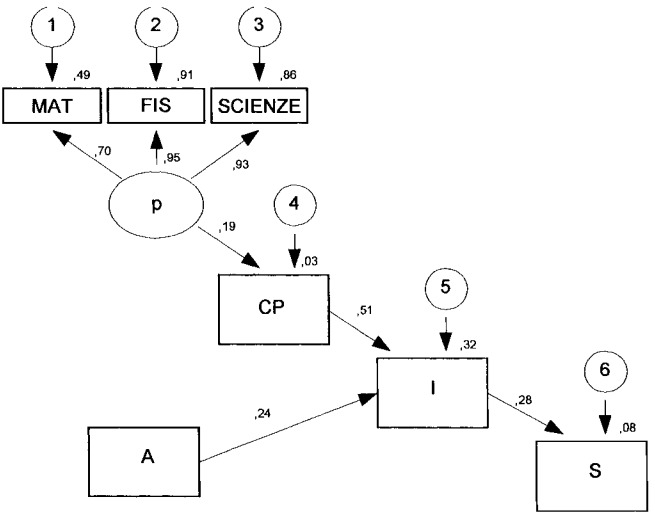


Figure 2. Model of academic career choice for the scientific area

A comparison between the Figure 2 and the Figure 1 shows how the model for scientific area is somewhat different from the general model initially postulated. Namely, whereas the influence of school self-efficacy on interests (a cornerstone of socio-cognitive theory) is still present, other influences (as the reciprocal ones between school performance and school self-efficacy, and the one of school self-efficacy on perceived abilities, as well as the one of school performance on interests) are associated to so small values of the associated coefficients as to be neglected.

5. THE NON-UNIQUENESS OF OBTAINED MODEL: THE CASE OF AMOS SOFTWARE

The results so far obtained appear to partly contradict our initial expectancies. Namely, instead of obtaining a unique model fitting the experimental data, we obtained five different models, each one relative to a different choice context. Such a circumstance already seems to be paradoxical, as, in principle, one should be inclined to think that a choice process is based on a very general mechanism, independent on the disciplinary area taken into consideration. Indeed, this was the general framework underlying the formulation of the general model depicted in Figure 1.



χ^2 =26.034 (df. = 14 p < .03) GFI = .943 AGFI = .887

Figure 3. Model of choice related to scientific area obtained through AMOS software

Anyway, we could continue to assert that model uniqueness is still holding, even if restricted to a particular context, the one of disciplinary area under consideration. It is possible, however, to show that even this assertion is wrong. When using a different software, the well known AMOS (Arbucle, 2003), for doing Structural Equations Modeling on the same data, starting from the same general model, we obtained a very different model for the choice related to scientific area, shown in the Figure 3.

By comparing the Figure 3 with Figure 2, we can observe how the structure of both models is essentially identical, whereas the model

coefficients differ in a significant way. We remark that the situation is even worse in the models related to the other four disciplinary areas, where the application of AMOS gave rise to deep changes even in the structure of models obtained through SEPATH.

6. DISCUSSION AND CONCLUSIONS

What is the cause for the difference between the two models of the same phenomenon (based on the same data set)? The answer is very simple if we take into account two facts:

1. the algorithms for minimizing the distance between the variance-covariance matrix of experimental data and the one predicted by the model are slightly different in the two software programs we used (SEPATH and AMOS);
2. the minimization problem is characterized by the existence of multiple minima.

The latter circumstance is connected to the number of free parameters existing in the models under consideration. It is easy to see that such a situation occurs in most descriptions of complex systems (neural networks are a typical case). It destroys any hope of obtaining a unique solution of whatsoever minimization problem related to these systems. Even if this fact is very well known from decades to all people dealing with machine learning problems, we will add here a further elementary example related to the minimization of the quadratic form $(\Sigma_i w_i - 1)^2$. In this example the index i runs from 1 to 10 and the goal of minimization procedure is to find the set of values of the 'weights' w_i granting that the value of the quadratic form be zero or lesser than a given value. In our example we choose such a value as given by 0.005. The minimization was performed through a standard gradient descent method, implemented through a Widrow-Hoff rule in which the weight variation at every step was given by the elementary expression:

$$\Delta w_i = -\eta(\Sigma_k w_k - 1),$$

where η denotes a suitable learning parameter. Besides, we started always this procedure from the same initial weight values, given by –0.1 for the weights corresponding to the indices 4,5,6,7, and given by 0.1 for all other indices. Despite the simplicity of this problem, by changing only the value of learning parameter we obtained different solutions. More precisely, when $\eta = 0.1$, our procedure converged in only two steps to values of weights given by –0.02 for the indices 4,5,6,7, and by 0.18 in all other cases. Instead, when $\eta = 0.01$, our procedure converged in 21 steps to values of weights

given by –0.02875352 for the indices 4,5,6,7, and by 0.1712465 in all other cases. A not so great difference which, however, should not exist in a so simple case if the hypothesis of the existence of a unique model would be true!

The conclusion stemming from the arguments presented within our paper seems now unavoidable: the derivation of a unique model from a given set of experimental data is, in the case of complex systems, in principle impossible. Thus a new form of intrinsic uncertainty adds to the ones already known: not only we have the uncertainties associated to the features and the variables we choose to introduce in our model, as well as the uncertainties associated to the choice of experimental data, but, *ceteris paribus*, there is a further uncertainty associated to the fact that algorithms will never give a unique answer for model coefficient values, just owing to the complexity of the system itself (and whence of any model trying to describe it). Even if this form of uncertainty was known from longtime by all mathematicians working in Numerical Analysis, now we need to take it into account and to evaluate more carefully its role in the future modelling of complex systems.

REFERENCES

Bandura, A., 1986, *Social Foundations of Thought and Action: A Social Cognitive Theory*, Prentice Hall, Englewood Cliffs, NJ.

Bandura, A., 1997, *Self-efficacy: The Exercise of Control*, W.H. Freeman, New York.

Bandura, A., 2000, *Autoefficacia. Teoria e Applicazioni*, Erickson, Trento.

Barak, A., 1981, Vocational interests: a cognitive view, *Jour. of Vocational Behavior* **19**:1-14.

Gill, P. E., Murray, W., and Wright, M. H., 1981, *Practical Optimization*, Academic Press, New York.

Krumboltz, J. D., Mitchell, A. M., and Jones, J. D., 1976, A social learning theory of selection, *The Counselling Psychologist* **6**:71-81.

Krumboltz, J. D., 1979, A social learning theory of career decision-making, in: *Social Learning and Career Decision Making*, A. M. Mitchell, G. B. Jones and J. D. Krumboltz, eds., Carrol Press, Cranston, RI.

Jöreskog, K. G., 1986, *Lisrel VI: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables and Least Squared Methods*, Uppsala University, Dept. of Statistics, Uppsala.

Lent, R. W., Lopez, F. G., and Bieschke, K. J., 1993, Predicting mathematics-related choice and success behaviors: test of an expanded social cognitive model, *Journal of Vocational Behavior* **42**:223-236.

Mizutani, E., and Jang, J.-S. R., 1997, Derivative-based optimization, in: *Neuro-Fuzzy and Soft Computing*, J.-S. R. Jang, C.-T. Sun and E. Mizutani, eds., Prentice Hall, Upper Saddle River, NJ, pp. 129-172.

Scales, L. E., 1985, *Introduction to Nonlinear Optimization*, Macmillan, London.

Zanetti, M. A., and Cavallini, E., 2001, "Come sono a scuola?" Credenze di efficacia e successo scolastico: uno strumento di autovalutazione, in: *Pre Atti del Congresso Nazionale Orientamento alla Scelta*, R. Nardelli, ed., Padova, 25-27 October.

A MODEL OF HYPERTEXTUAL STRUCTURE AND ORGANIZATION

Maria Pietronilla Penna[1], Vera Stara[2], Daniele Costenaro[1] and Paolo Puliti[2]
[1]*Università degli Studi di Cagliari, Facoltà di Scienze della Formazione, Dipartimento di Psicologia - Email: maria.pietronilla@unica.it, costenaro@email.it;*
[2]*Università Politecnica delle Marche, Facoltà di Ingegneria, DEIT Email: v.stara@univpm.it, p.puliti@univpm.it*

Abstract: Nowadays, many educational services are offered through computers, and the fast spread of hypertexts impose a careful reflection to all experts, both on its structure and its organization design to reach an effective adaptation to the customer target and to the context of use. The problematic gains its greatest importance in the specific case of introduction and use of hypertextual support in the primary school, where structure, navigability, friendliness and coherence represent open key issues (Panto, 1999). Unfortunately, a standard of reference does not exist currently; it would predispose modalities of navigation, define rules and conventions on the quality of the information provided, or outline the relationship previewed in every node (Norman, 1994). Therefore, it is difficult to translate the educational plan into an effective and functional unit without a tutorial guidance. The objective of this contribution concerns the location of a "reasoned" hypertextual structure and organization in a systemic view of all the features, in order to identify a careful instructional model design.

Key words: hypertext; instructional strategies.

1. THE MODEL

The model would help to understand the role of complexity while reflecting on didactic tool quality improving the multimedia educational value.

Model construction follows the steps:

- the target analysis phase, oriented to learn about the audience;

- the design phase, centred on selection, writing and designing the hypertextual environment and contents;
- the prototyping phase, oriented to develop the hypertext; according to the ADDIE (Analyze, Design, Develop, Implement and Evaluate) Model.

Each step previews a systemic analysis of emergent aspect, used to allow an overall view of the tool building process:

- Cognitive aspect;
- Instructional strategies design aspect;
- Text understanding aspect;
- Usability aspect.

1.1 Step 1st. The Target Analysis Phase

"Know your user" is a categorical imperative while planning. Every human-machine artefact should be adapted to its real target, in order to ensure a friendly interaction; furthermore, a preventive analysis of background knowledge, user experience and content definition could determine a clear asses of the learner's needs and preferences in order to create an adequate balance between human factors, system design and instructional design.

Hypertext should provide concrete advantages for improving contents accessibility given similarities between the artefact and human information process (Mantovani, 1995), its interactive characteristics could change users into "active and aware agent" by maintaining attention active in global meaning.

If those considerations are right, it is concretely possible to develop a technological product to improve learning process: a good educational training is not achieved by means of the most modern technological support system but through the closer approximation of human factor with machine factor.

Education quality can only increase if the user's characteristics correspond to the instrument's characteristics.

The target analysis helps, moreover, to identify those structural aspects that characterize the educational product: analogy (level of affinity between the truth and that one represented from the product), interactivity (level of interaction between customer and product), freedom (level of active action during navigation); in this way, every choice and action is always taken considering the final customer.

1.2 Step 2nd. The Design Phase

Although different types of hypertext exist, the most used structure and organization applied in primary school is the closed mode, due both to user's insufficient skills and to class management reason.

This particular design is suitable to represent the contents and the structure helps to not "get lost in hyperspace".

The computer product has its most important aspect in the elaboration of the interface, being the interaction interface between the contents and people who must learn, regardless of its type (closed, open or mixed) and its destination (adult target or not).

Design an interface "adapted to child" could seem trivial, however the adaptation of computer world to children's world does not seem to be easy, not only for the introduction to informatics but for a particular user's model problem.

Interface and contents, in fact, can not leave aside a perfect accordance with the cognitive processes that are active during the vision and the manipulation of the product: every graphical or textual aspect must be adapted to perception, attention, memorization and learning dynamics.

This involves a preventive communication and visualization choice, that guides the designer during the definition of the interface's characteristics in order to create an easy connection between the learning set through the instrument and its practice.

A key question would be at the basis of all the interface design: which mind processes are stimulated? Which cognitive structures are activated? Is it possible to find an equivalence between the structural aspects of the didactic instrument and these cognitive structures?

Once again, considering the user and the specific context of use, the product would have to follow some "directions of common sense" like these main points:

- the font must be readable and understandable preventing, as possible, the cognitive overload;
- the chromatic choices must respect the requirements of visual perception;
- there must be an opportune balance between several communicative channels and contents offered;
- the navigation style must be intuitive;
- every used metaphor must be shared with the user.

The second crucial node is the elaboration of the contents.

"Instructional strategies" are used to introduce a content or a unit and to show the goals student must manage.

How can I design the instructional message? It is possible to create an instructional strategy to focus the learner's attention on the instruction, signaling different aspects of the text through words and typography, and, using picture to enhance learner understanding. It is necessary, in fact, to create a set of distinctive elements that signal the structure of the contents to learners manipulating text's schema with explicit signals, like pointer words (Meyer 1985), and with typographical signals to identify text's changes in topic from the surrounding words.

The use of pictures or graphics instruction, as well, illustrates verbal information and make it more meaningful to the learner.

1.3 Step 3rd. The Prototyping Phase

In this phase we translate the previous steps into an effective product: an hypertext.

The hypertextual architecture is mixed: one first section is closed type, and carries the user to the second section that is open type.

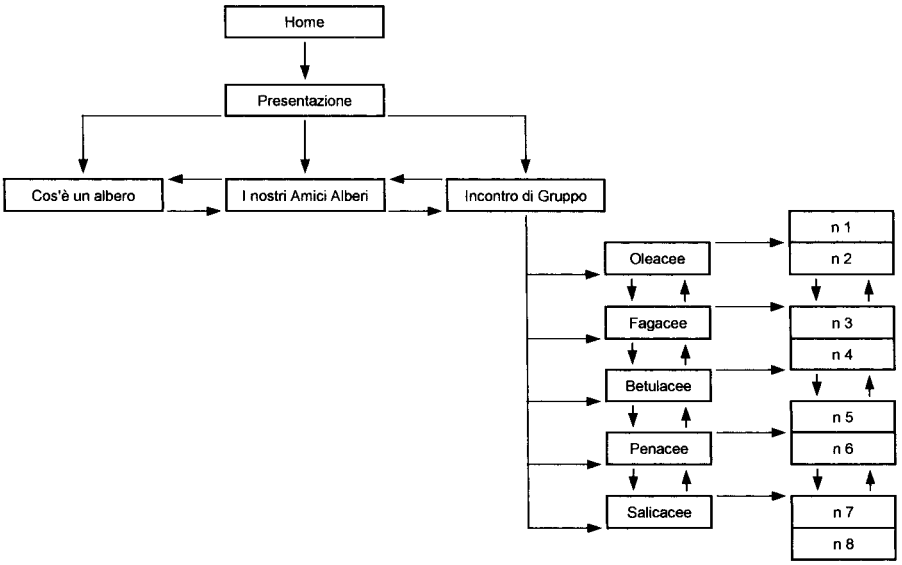


Figure 1. Hypertext architecture.

Purely didactic reasons have determined this choice: the content's exposition presupposes an introduction to the thematic in which lesson's objectives are defined and, through them, learners approach the principles

that represent the background knowledge for the information in the several pages of the hypertext.

Contents and learning objectives presentation is a key point in a step-by-step user approach to the problematic; in such a way the learner calibrates the cognitive effort demanded and he builds a mental path to achieve the educational goal.

Using a visual identification, a logo, inside the interface's space and in the main sections of the hypertext can stress more this representation and follow the reader in every navigation step.

The visual metaphor, obviously, has been chosen referring to the content in order to create a clear connection between content-interface-user, which are the three poles of every instructional computer product.

Figure 2. The Logo.

Figure 3. The main menu.

An "instructional strategies design" constantly accompanies the readers through the use of introduction questions (preinstructional strategies). At the beginning of every page, the reader enters the argument through short questions that induce reflection or stimulate memory if information is already known.

The use of some structure signaling of the text through explicit marks and typographical marks in order to communicate the subdivision of important contents. To improve the perception of the text's structure, we introduced

different strategies of text representation concerning the background and the different elements constantly present in the interface; in particular every page introduces its key words with different fonts and dimensions underlining the target concept of the visualized area.

The use of images in GIF and JPEG format was introduced for decorations and content representations.

Moreover, the contents have been developed considering some aspects of language understanding: the text organization in the whole and sentences in their context.

The text has a hierarchical structure that proceeds from the simple to the complex and pushing gradually the reader's attention in the details of the argument.

The single sentences maintain a coherent configuration with the hierarchical structure through various types of recurrent relations that concur to the organization of the same phrases in the overall structure of the text: the relations present in the prototype are "answer", "specific", "explanation", "succession", "cause" types (Meyer, 1974).

The importance of such relations resides in the clues provided to the reader concerning the particular way in which the phrase, containing the same relation, must be used.

Table 1. Type of relations between phrases.

Type of relation	Description
Answer	presentation of a question to which answer follows one
Specific	presentation of specific information as a result of a general point
Explanation	explanation of a point
Succession	presentation of the points in their temporal succession like a system
Cause	presentation of an event like cause of an other one

The attention given to the information architecture of the interface, to the instructional strategies, to the designs and to the written language, would contribute to end product usability, that has been monitored during planning, making mainly reference to the parameters of a good informative stability and an easy reading through an opportune trade-off between figures and background.

2. CONCLUSIONS

The construction of hypertexts for children requires a careful analysis of its components, that are characterized in structure (type of connections between the nodes) and in content organization.

A preventive analysis of such levels could lead to an identification of rules and conventions on the quality of the didactic products and permit to verify which are the modalities that, supplying a mental representation of the educational curriculum, allow learners to effectively exploit the potentialities of the technological supports in the daily management of the class and in the instruction processes.

Follow some fundamental steps could be useful in order to guarantee, in primis, the effectiveness of the product in terms of educational objects achievements and in terms of the functionality of the same application. These steps could represent the guidelines of the plan hic et nunc, in absence of directives.

The congruence of aspects like cognitive aspect, instructional strategies design and usability, could improve the quality of the educational technology.

Obviously the problem of the effective use of technology, of whatever kind, cannot be resolved exclusively by planning effective, efficient and satisfactory applications, but it presupposes a careful analysis on the dynamics of value and importance attribution to the instruments used by subjects.

A hypertext, in this study case, can be designed in consideration of all the variables that are considered fundamental for the design of good formative material, but for sure its use in school context will depend on teacher's and students' ability in taking advantage from it.

The technology application field in the school system appears still unconscious of the risks of a dogmatic integration: the introduction of a new instrument, should always be followed by an essential vademecum, including a description of the risks and the advantages of the technology, that is intended to understand the real impact that technology could produce in daily actions.

The education field can not leave aside technology, but it would be desirable to find an opportune setting in order to activate the reflection about a coherent definition of "computer science integration", "evaluation of quality of the formative participation through the technology"; probably meaning sharing could carry to an inner reference standard that clears the field from false believes and activates a net of strategies oriented to the improvement of education, thanks to the contribution that technological integration can offer.

REFERENCES

Anderson, J. R., 1990, *Psicologia Cognitiva e sue Implicazioni*, Zanichelli, Bologna.

Costa, R., 2002, La valutazione dei media didattici, *Pedagogia Sperimentale e Ricerca Didattica* **1**(1).

De Beni, R., and Pazzaglia, F., 1995, *La Comprensione del Testo*, UTET, Torino.

Meyer, B. J. F., 1985, Signaling the structure of text, in: *The Technology of Text*, D. H. Yonassen, ed., Educational Technology, Englewood Cliffs, NJ, pp. 64-89.

Norman, K. L., 1994, Navigating the educational space with hypercourseware, *Hypermedia* **6**(1).

Pantò, E., 1999, Il valzer dei siti, in: www.graffinrete.it/tracciati/storico/anno99/valzer.htm

Van Nimwengen, C., Pouw, M., and Van Oostendorp, H., 1999, The influence of structure and reading-manipulation on usability of hypertexts, *Interacting whit Computers* **12**:7-21, in: http://www.istruzione.it/innovazione/didattica/valutazione.shtml.

6

LEARNING

TEACHERS IN THE TECHNOLOGICAL AGE: A COMPARISON BETWEEN TRADITIONAL AND HYPERTEXTUAL INSTRUCTIONAL STRATEGIES

Maria Pietronilla Penna[1], Vera Stara[2] and Daniele Costenaro[1]
[1]*Università degli Studi di Cagliari, Facoltà di Scienze della Formazione, Dipartimento di Psicologia, Email: maria.pietronilla@unica.it, costenaro@email.it*
[2]*Università Politecnica delle Marche, Facoltà di Ingegneria, DEIT Email: v.stara@univpm.it*

Abstract: The educational process occurs within a very complex system, in turn characterized by the cognitive systems of both teachers and learners, as well as by their interactions and by the technological tools which mediate these latter. In order to gain some knowledge about the interrelationships on which such a system is grounded a possible strategy is to start by investigating, through the methods of experimental psychology, the role played by specific instructional strategies in improving the educational process itself. In particular, we compared two very popular strategies, the one based on usual frontal lesson done by a single teacher, and the other based on a suitable computer-supported hypertextual structure (even in the latter case, however, under teacher supervision and assistance).

Key words: instructional strategies; computer based learning; technological education.

1. INTRODUCTION

In order to do a correct comparison between the two strategies we must, however, take into account that the hypertext-based one requires a particular technological tool – the computer – and, as such, its effectiveness depends, in turn, on teacher's technological know-how about the tool itself. Such a circumstance highlights a further problem: how much the teacher's cognitive framework is suited to take advantage of the new technological tools? A

tentative answer would be: very little. Namely, despite the fact that the introduction of personal computers within school contexts dates back to, more or less, forty years ago, and notwithstanding the spread of multimedia and new educational technologies, we still need a new configuration of teacher know-how. Namely, within a technological context, the teacher should act as a sort of individual or group tutor who is able to make easier the use of the new technologies, to redesign the relational networks inside the classroom-group and to behave as a "guide for student psychosocial evolution"(Parisi, 2000). The most important task for such a teacher would be the one of guiding students in information selection, according to previously stated educational goals, while avoiding, at the same time, both informational losses and cognitive overloads.

However, within the actual school organization, the knowledge transmission is still viewed as an information exchange through frontal lessons, a circumstance which entails a specific way of characterizing the teacher role, his/her tasks and activities according to a strict know-how, which interferes with other different educational strategies. So the effective use of new technologies involves, in fact, a change in school habits that can confuse the teachers, who are forced to redefine their role according to innovative patterns (Calvani, 1998). For this reason the research described in this paper studied the effects of the two strategies – the traditional and the hypertextual one – relying not only on traditional behavioural indicators, such as the communication style adopted, but also on an investigation about teacher's technological competence and experience, as well as on his/her opinions and expectations about the use of computer-based educational technologies.

2. THE RESEARCH

According to what stated in the introductory section, our research was performed through three successive steps:

1. Assessment of teacher's technological expertise, as well as of his/her representation of latent potentialities of the technological tools in an educational context;
2. Design of a lesson about a specific topic, to be done in one of two different ways: a hypertext-based one, and a traditional one;
3. Ecological observation about the communication styles adopted during the lesson.

The research involved 12 teachers of an elementary school, randomly extracted from a sample used for a previous research (Penna et al., 2002). Of these teachers, 6 were operating in a school without computer laboratory.

The three steps previously quoted will be described in detail in the following.

2.1 First step

The assessment specified before was obtained by submitting each teacher to a questionnaire including 26 items. These latter were designed both to measure teacher's *user experience* and to survey his/her opinions and expectations about topics such as:

a) The preference of traditional vs. computer-aided teaching contexts;
b) The influence that computer-aided teaching contexts have, compared to the one exerted by traditional contexts;
c) The changes in teacher-students relationships that could occur within computer-aided teaching contexts.

2.2 Second step

It was devoted to the design of a lesson about a scientific subject: the trees and their botanical features. The design regarded both teaching methods: the hypertext-based and the traditional one.

The context of both methods was organized so as to obtain a fair balance between textual and graphic contents: both the hypertext and the paper support (the latter designed for the traditional lesson) included a reading part and a specific image-looking part. The language adopted in both cases was tailored for the target users: 199 children of the III, IV and V class of elementary school, 92 attending the hypertext-based lesson, and 166 attending the traditional lesson (Penna et al., 2003).

The organization as well as the presentation format of both the hypertext and the paper text prepared for the traditional lesson were based on three rhetorical elements (Meyer, 1985a):

1. the covariance, which induces a causal relationship (*paratactical* or *hypotactical*) between events described within the text;
2. the description, that is the presentation of a topic and of its features;
3. the question, which creates a relationship between two elements within the text, the one representing the question and the other its answer.

Moreover, we choose to use suitable graphical features in order to direct the reader's attention upon a number of elements which we considered as primary within both paper text and hypertext.

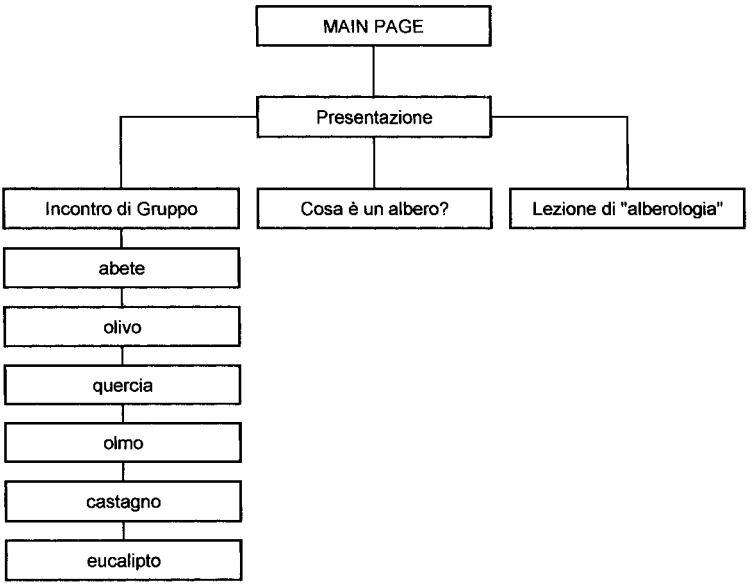


Figure 1. The architecture of the hypertext we adopted in our research.

The hypertext structure adopted was the one of a "*horizontal tree*" (see Figure 1), so as to allow a small number of paths. In this way most information can be accessed already from the first pages. We choose this type of architecture for two reasons: first, this kind of hypertext includes not much pages; in the second place, an extended multilinear hypertext represents an average synthesis between simplicity and hypertextuality.

We asked every teacher to feel free in interpreting the lesson and, as regards the hypertext-based lesson, the navigation was oriented according to teacher's user experience.

2.3 Third step

In this step we asked each teacher to do a lesson about the botanical features of trees; of the 12 chosen subjects 6 did the lesson in a hypertextual way and 6 in a traditional way. An observational grid, including four areas and 15 categories, was built; its design was based on the idea that, as regards the relationships between students, and between students and teacher, the synergy, supported by computer availability, between playing, studying, collaborating and meeting can foster class cohesion, and can be useful in achieving learning targets (Schimmenti et al., 1996). This occurs because the presence of a computer should allow a better communication in school

environments, orienting it towards a multidirectional dynamics (Olguin et al., 2000).

As a consequence, the observational grid we introduced was designed to detect differences as regards the construct "direction of educational communication", differences which could be attributed to the influence of the following factors:

- the way of teaching, hypertext-based or traditional
- school class attended by the students (III, IV or V).

To be more precise, the proposed observational grid was subdivided in four areas, according to a Buxton (1996) classification, regarding:

- *unidirectional communication (low dialogue),* in turn described through four categories – or moments – all dealing with a situation in which the communication is only from teacher to students;
  A1.0 Recalls the attention of students on text
  A1.1 Recalls the attention of only one student on text;
  A1.2 Recover a previously specified content
  A1.3 Integrate previously specified concepts among them
- *bidirectional communication* (medium dialogue), described through four categories – or moments – dealing with a situation in which a reciprocal communication between teacher and students is active;
  A2.0 Gives a feedback to strains of students
  A2.1 Involves students in conversation
  A2.2 Answers to questions and curiosities
  A2.3 Organizes the discourse taking care of student's demands
- *multidirectional communication* (high dialogue), described through four categories – or moments – dealing with a situation in which the conversation is active both between teacher and students and among students;
  A3.0 start up discussions
  A3.1 Confront about some concepts
  A3.2 Importance is given to experience of students
  A3.3 A problem (ex. of comprehension) is solved
- *various*, the last area which includes three categories describing pause, silence, and chaos moments.

3. DATA ANALYSIS AND DISCUSSION

The user experience of teachers was homogeneous throughout the chosen sample. An analysis of questionnaire answers, done through descriptive statistics, a χ^2 and a factor analysis, evidenced how all teachers, independently from the existence of a computer laboratory within their

school, attended a computer literacy course and evaluated the level of their knowledge of this topic as oscillating between adequate and insufficient. Moreover, they felt difficulty in assessing their subjective basic expertise.

In spite of this, for what concern their preferences for technological or traditional teaching contexts area, all subjects appeared as convinced of potential advantages stemming from the use of technological tools in the school contexts, as well as of the need for introducing them. In particular, teachers tell that attention, recall and plainness of arguments would have been improved by a technology-aided teaching (See Figure 2).

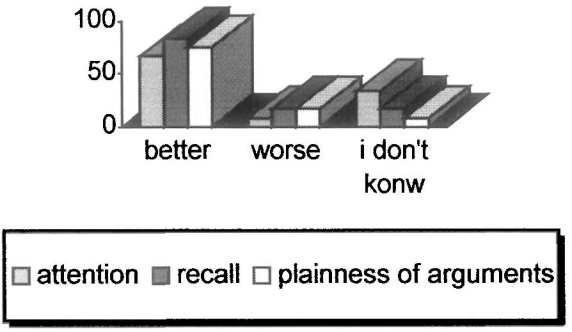


Figure 2 Percentages of opinions regarding the effects of a technology-aided teaching.

As regards the opinions about the relationships among students throughout the group activities, the technology-aided context was evaluated as better than the traditional one, owing to the fact that it should improve the positive features of collaborative relationship (See Figure 3).

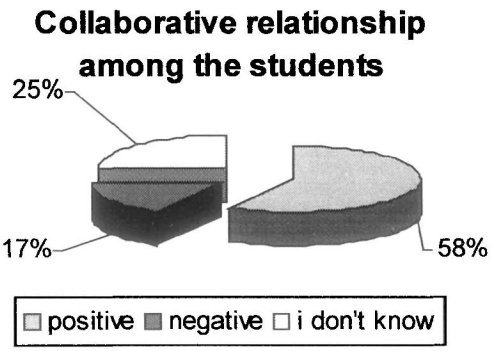


Figure 3. Percentage of opinions regarding the effect of technology-aided context on the collaborative relationship among the students.

As regards the teacher's role in the class context in presence of a technology-aided teaching, the majority of subjects felt that it should be more oriented to promotion of meeting and of the use of new technologies, with a plain improvement of performance and with a lowest tiring in daily working. A role which can be overall qualified as positive (see Figure 4).

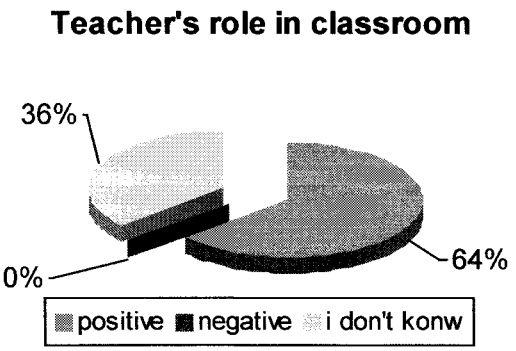


Figure 4. Percentage of opinions regarding the new teacher role as a consequence of a technology-aided context.

As regards the data collected, with the aid of our observational grid, from the ecological observation of the lessons concretely done, an ANOVA didn't evidence significant differences in direction of educational communications between the two teaching ways (hypertext-based and traditional).

An overall descriptive statistics regarding both kinds of lesson evidenced how the unidirectional communication was active in 44% of cases, the multidirectional communication was active in 30% of cases, the bidirectional communication was active only in 18% of cases, whereas the remaining 8% of cases were belonging to the fourth area (see Figure 5). As previously said, we didn't find significant differences for the "kind of lesson" factor and the "class" factor; it is, however, to be remarked that the behavior labeled as "recalls attention of class on the text", within the area of "unidirectional communication", showed a significant growth from the third to the fifth year in primary school ($F = 13.452$; $df = 2$; $p < 0.02$).

These outcomes evidence that teacher's attitudes towards different kinds of instructional strategies are substantially invariant, a circumstance which probably could be related to a low confidence in using technological tools.

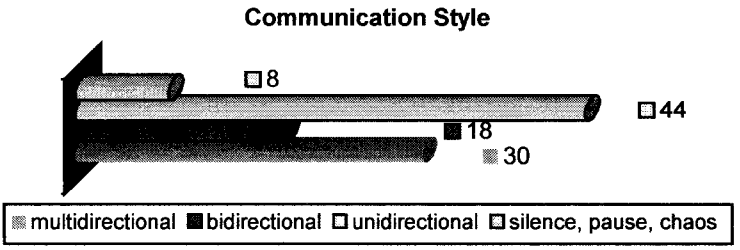


Figure 5. Percentages of the observed communicative behaviors belonging to the four areas included within our observational grid.

All was going as if teachers, when using a new technology, were adopting a traditional paradigm in which contents were simply transferred on video and the actions of the students were reduced to a passive reception of the text. This 'traditional' use of a new technology erased the possibility of promoting a modality of collaborative and active learning that should characterize the educational process in presence of hypertextual supports. If used in a proper way, the hypertext would provide an environment for developing a new way of teaching, in which there would be a synergy between theory and practice, and multiple interconnections. Probably, in this case the application to the hypertextual context of instructional methodologies typical of a traditional context would seem to be finalized to the maintenance of the class control: in fact, no teacher allowed the students to freely navigate during the lesson, preferring to centralize the communication on unidirectional styles rather than to take advantage of the possibility to teach according to multidirectional paradigms, that should characterize the computer-supported instruction.

Besides, this attitude evidences the fear of finding themselves in a new situation, in which the escape solutions are not known, probably for a technical ability lived as inadequate or for a rigid reference to old knowledge and to fixed schemata. When relying on flexible schemata, we lose certainty and perhaps this it is the main difficulty experienced by most teachers when approaching the new technologies. The ability to cope with such a difficulty is, however, one of the main requirements for the future teachers, which will be forced to take into consideration the changes that the educational use of the technology can determine in the teaching style, rethinking the entire organization of the educational processes.

4. CONCLUSIONS

A new shaping of teacher's know-how isn't enough, insofar as it's needful. The concept itself of teacher training should be extended far beyond the "informatics literacy" one currently used. The purpose of this new, and more effective, kind of training should be the one of providing an amount of knowledge such as to allow for a productive use of computer in the classroom. It's probably a utopia to believe that teachers can actually acquire new *know-how* in absence of a more articulated learning path (Calvani, 1998).

It would be more useful to design the informatics literacy course so as to include a critical examination of concrete teaching situations, so that every teacher can have a support from specialists and therefore can have the opportunity to criticize methods and strategies (Schlager et al., 1998). Of course, the better way to perform such a design would be the one of resorting to a systemic framework. Namely, as evidenced by a study of Antonietti and Cantoia (2001), the simple introduction of a computer does not necessarily carry *per se* a dramatic improvement in school activities. A technological tool can contribute to improve the learning process, but only when a correct representation of its potentialities is embedded within a concrete educational design. On the other hand, teacher common opinions are full of erroneous beliefs, often strengthened by a superficial literature on computer based education, seeming to give credit to the hypothesis of the existence of different learning levels, which should depend only on the choice between traditional lessons and hypermedial ones (Gineprini, 2000).

To take the maximal advantage of the human-machine interaction, the introduction of technology in school environments should be supported by a t*echnology education program*, both for students and for teachers. Only in this way the computer use could support a new form of learning, through the discovery of new schemata. Once again, the outcomes of our research evidence how the lacking of a systemic approach in the introduction of a new technology - this time in the educational domain – is useless or can even result in a failure. In our opinion, only by adopting a systemic framework we can ground on computer-supported instruction a deep transformation of educational processes.

REFERENCES

Antonietti, A., and Cantoia, M., 2001, *Imparare con il Computer, Come Costruire Contesti di Apprendimento per il Software*, Erickson, Trento.

Buxton, W., 1996, The emergence of communications study psychological welfare or scientific thoroughfare?; http://www.cjc-online.ca/~cjc/BackIssues/21.4/buxton.htm .

Calvani, A., 1998, Nell'illusione tecnologica c'è un pericolo per chi insegna, *Telema* **12**.

Correard, I., 2001, Twelve years of technology education in France, England and the Netherlands: how do pupils' perceive the subject?; http://www.iteawww.org .

Gineprini, M., 2000, Multimedialità e centralità della didattica; http://www.pavonerisorse.to.it http://www.ed.gov/technology/TechConf/1999 .

Meyer, B. J. F., 1985a, Prose analysis: purposes, procedures, and problems, in: *Analyzing and understanding expository text*, B. K. Britton and J. Black, eds., Erlbaum, Hillsdale, NJ, pp. 11-64, 269-304.

Meyer, B. J. F., 1985b, Signaling the structure of text, in: *The Technology of Text*, D. H. Jonassen, ed., Educational Technology, Englewood Cliffs, NJ, pp. 64-89.

Olguin, C. J., Delgado A. L. N, and Ricarte, I. L. M., 2000, An agent infrastructure to set collaborative environment, *Educational Technology & Society* **3**(3).

Parisi, D., 2000, *Scuola.it*, Mondadori, Milano.

Penna, M. P., and Stara, V., 2003, Un esempio di rappresentazione cognitiva dello strumento tecnologico nel sistema scolastico italiano, in: *Congresso Nazionale "Contesto, Cultura, Intervento. Quale psicologia per la scuola del futuro"*, Lecce 20-22 Giugno 2003, pp. 238-240.

Penna, M. P., Stara, V., and Bonfiglio, N., 2002, A systemic proposal on the use of a new technology as a learning tool in school context, in: *Emergence in Complex, Cognitive, Social, and Biological Systems*, G. Minati and E. Pessa, eds., Kluwer, New York , pp. 153-157.

Schimmenti, V., D'Alessio, M., and Schieda, A. M., 1996, *Il Computer: Rappresentazione e Apprendimento nell'Età Scolare*, Franco Angeli, Milano.

Schlager, M., Fusco, J. and Schank, P., 1998, Cornerstones for an on-line community of education professionals, *IEEE Technology and Society* **17**(4).

THE EMERGENCE OF E.LEARNING

Maria Pietronilla Penna[1], Vera Stara[2] and Paolo Puliti[2]
[1]*Università degli Studi di Cagliari, Facoltà di Scienze della Formazione, Dipartimento di Psicologia, Email: maria.pietronilla@unica.it;*
[2]*Università Politecnica delle Marche, Facoltà di Ingegneria, DEIT Email: v.stara@univpm.it, p.puliti@univpm.it*

Abstract: The innovation process, which took place in recent years, in Information and Communication Technology (ICT) had a strong impact on training activities. Namely the push towards knowledge commercialisation seems to be oriented to a setting which is free from both spatial and temporal constraints. This, in turn, implies a redefinition of the concept of Open Learning that, in its ideal representation, should be the solution to all distance learning problems. The new form of computer based knowledge involves a different approach to didactics, through the learning object practice, and, in particular, a careful reflection on technological equipment intended to: 1) assure accessibility to every user, 2) grant for the fruition of a useful product to all recipients, 3) assure a perfect integration between the knowledge content and the different devices employed. The aim of this contribution is to propose a systemic reflection on some emerging aspects in e-learning contexts, related to the role of usability, and to accessibility and cognitive processing in designing phases.

Key words: e.learning; usability; accessibility.

1. A SYSTEMIC REFLECTION ON THE GUIDELINES FOR E.LEARNING PROJECTS

As any other cyber-idea, e.learning must put the user at the center of every design framework, in the necessary effort to approximate the human model, not only the model of the user and of the person who plans the specific course, but also the model of the person who implements the platform. The course's effectiveness depends on the contents, but also on user's understanding on how to "handle" them in his/her own knowledge management. The platform planning phase is therefore the fundamental link

between a good outcome of the overall information architecture and the choices made while defining links, icons, search engines and the various multimedia devices; these choices will determine the effective dialogue level between instruments and the customer.

Beyond any thought on e.learning effectiveness, its advantages or drawbacks, it is important to consider how the problem of cognitive factors in the formation process are "complicated" by the problems arising from the nature itself of human-machine interface. So, several forces act in the complex scernario of the new formative settings, where the users try to control at the same time the attention both to the lesson and to the interface manipulation. The friendlier the interface is, the fewer users will spend energies in looking for correct actions in order to select information on the screen.

An ancient problem of every formation basic technology is the quality assessment of efficient service and effective attainment of instructional's goal. Maybe the usability could be, to this regard, the basic parameter for the evaluation of e.learning technologies and systems (Zaharas, 2002) but it is necessary to know first what is its effective role within the context of e.learning design even if, unfortunately, we lack usability studies related to distance education.

Probably, the most important reflection is the one leading to choose a specific real meaning of "usability for e.learning", where "usability", according to ISO 9241, is a complex outcome of *effectiveness* (the user's ability to achieve specific goals in the environment), *efficiency* (the resources used, such as time, money and mental effort), and *satisfaction* (the user's comfort level and acceptance of the system). These features are present also in the Formative Evaluation where *effectiveness* is the achievement of instructional objectives, *efficiency* represents how quickly learning objectives are achieved and *satisfaction* describes the user's interest and desire in learning. Lohr (2000) tries to integrate the basic of ISO 9241 with the formative evaluation (Figure 1), defining the "Instructional Interface Design Process". In this special context *effectiveness* describes how much the learner interprets correctly instructional interface functions, *efficiency* defines the learner's experience of a minimal frustration in using instructional interface elements, and, at the end, *satisfaction* concerns how much learners feel comfortable in the overall environment.

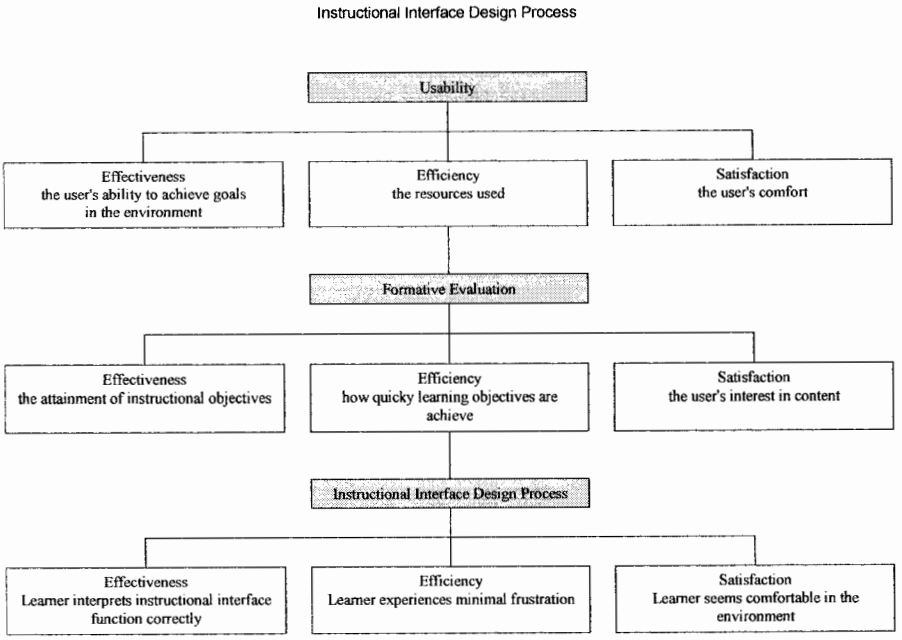


Figure 1. Instructional Interface Design Process.

2. PLANNING AN E.LEARNING INTERFACE: A SYSTEMIC DESIGN MODEL

In order to realize a learner-centered design, it is necessary to analyze learner's attitudes towards technology that could be key determinants in predicting student motivation and success. Understanding learners' profiles is the best way to create useful designs, styles and tones, but, when delivering training via online learning, there are some special design concerns that represent other potential benefits in planning; they start from a common step: select a delivery technique or combination of techniques in order to define, a priori, a user interface design. In synchronous training as well as in asynchronous training, user interface design refers to the overall look of the program that allows learner to access information (Hall, 1997). Identifying what navigation tools are most user-friendly and where to place information are the main problems associated to this kind of design.

In designing the interface we must remember the following basic prescriptions, according to Van Rennes and Collis (1998), stemming from usability heuristics for interface design of software systems:

- Use a obvious design that helps learners to find their own way through the environment giving them the control of navigation.
- Use clear and consistent navigation systems that indicate always the site status.
- A visible indication (such as title page, menu bars, clear name links, etc) helps users to identify where they are.
- Use screen-friendly fonts and web-safe colors in order to create a standard consistent look.
- Learners are sensitive to the readability of on-screen text; therefore formatting and spacing of the text as well as colour are important; moreover a common look helps users to distinguish course pages from external linked hyper-pages.
- Pages must fit onto all monitors and browsers; it is fundamentally necessary to build flexible and efficient web pages that ensure accessibility for all audience.
- Provide quick download times and help users by providing printer-friendly pages.
- Learners do not like studing texts from the screen and they study away from the computer; when they search for information they do not want to go more than three clicks, so they need a navigation frame always available.
- Learners are always in search of something new inside the web; it is important to update frequently contents and news and also give a direct indication of what is new as soon as possible.

These guidelines must to be read while taking into consideration the existence of several, mutually dependent, layers of usability (Kukulska-Hulme and Shield, 2004), that is:

- the context-specific layer, in which case courses have their own needs and outcomes;
- the academic usability layer, which deals with educational issues;
- the general usability layer, reflecting general HCI concepts;
- the technical usability layer, centered on common functional usability problems.

3. CONCLUSIONS

Human factors, system design and instructional design are the principal components in a e.learning device, but "workers in HCI and educational computing areas rarely speak to each other or take note of others'work" (Squires, 1999). In order to achieve effectiveness, efficiency and satisfaction

and to make e.learning methodology more effective, we should need a good web design while taking into account usability issues, but this circumstance is still far away from taking place. A good instructional design can advantage learners in having a successful and enjoyable experience, but this is possible only if the learning environment does not become a frustration barrier (Smulders, 2001). It is just in this delicate field that HCI, educational experts and designers must collaborate to understand how to plan a user centered e.learning website.

REFERENCES

Clark, R., 2002, Applying cognitive strategies to instructional design, *Performance Improvement Journal* **41**(7):8-14.

Guidelines for e-learning projects in the Public Sector, 2003, CNIPA, Roma, http://www.cnipa.gov.it/site/_files/ENG%20linee%20guida%20e-learning.pdf.

Hall, B., 1997, *Web-Based Training Cookbook*, John Wiley & Sons, New York.

Kruse, K., 2004, Designing e-Learning User Interfaces. http://www.e-learningguru.com/articles.htm .

Kukulsa-Hulme, A., and Shield, L., 2004, The Open University, http://www.shef.ac.uk/nlc2004/Proceedings/Individual_Papers/Kukulska_shield.htm .

Lohr, L. L., 2000, Designing the instructional interface, *Computers in Human Behavior* **16**(2):161-182.

Penna, M. P., Stara, V., Farci, E., and Moreno, A., 2003, La progettazione di un corso online per insegnanti. In: *Atti del 7° Congresso Nazionale AIDA*, Centro Stampa Nuova Cultura, Roma.

Smulders, D., 2001, Web Course Usability; http://www.learningcircuits.org/2001/aug2001/elearn.html .

Squires, D., 1999, Usability and educational software design, *Interacting with Computers* **11**(5):463-466.

Van Rennes, L., and Collis, B., 1998, *User Interface Design for WWW-based Courses: Building upon Student Evaluations*, AE Enschede, The Netherlands. Educational Science and Technology, University of Twente. (ERIC Document Reproduction Service No. ED 428 731).

Zaharias Panagiotis, 2002, Usability and E.learning; http://www.japit.org/zaharias_etal02.pdf.

SPATIAL LEARNING IN CHILDREN

Barbara Lai[1], Maria Pietronilla Penna[1] and Vera Stara[2]
[1]*Università degli Studi di Cagliari, Facoltà di Scienze della Formazione*
[2]*Università Politecnica delle Marche, Facoltà di Ingegneria, DEIT*

Abstract: The scope of research described in this contribution is to verify if the performance in a spatial reconstruction task is determined from mechanisms of visuo-spatial memory and from the characteristics of the subject, like the age and the sex, rather than from those of the environment, like landmarks or elements contained within it (in turn characterized by shape, size and color).

Key words: memory; visuo-spatial memory in children; landmark; spatial reconstruction task.

1. INTRODUCTION

An important aspect of our daily activities is to remember where we left recently used objects or a route. Without some sort of spatial memory we would be continuously engaged in active search for our keys or the road when we go in office. Fortunately, we do have this kind of mental ability. The models of working memory originally included visuo-spatial memory as an undifferentiated component, but recent work indicated that spatial location is separable from other visual characteristics (Baddeley, 1991, 1995). The spatial memory contains only the information related to position, distance and heading of an object with respect to another; for example it lets us describe and travel familiar ways, move from a place to the other. On the contrary the visual memory records only the visually determined characteristics, like the form, the colour and the size of an object or of a person; for example it allows the recognition of a face and of its features. The modalities of visuo-spatial memory operation are evidenced from the ability to navigate and reconstruct the structure of a route. One of the fundamental concepts introduced to this purpose is the one of cognitive map

(Tolman, 1948), that is an abstract representation of position of important places you meet following a path, stored into memory.

An important source of information for the specification of orientation in space and for spatial memory is provided by external sensory signals obtained instantaneously at each position in space. This local position information comes from all sensory readings as a function of observer position and orientation, and underlies the most general concept of landmark information. In vision, the landmark at a particular point is a view or a snapshot, that is a raw image of an object acting as spatial cue along a path. There are different kinds of spatial cues. For instance, a proximal landmark is an object that is so close to the goal of navigation that the subject simply must orient itself and approach it. On the contrary a distal landmark is much further away from the goal but is still close enough to it to provide some information about the distance to the goal itself (Leonard and McNaughton, 1990). The studies on the role played by landmarks (Mallot and Gillner 2000; Shelton and McNamara, 2001) indicated that these latter facilitate a memory storage based on characteristics of environment and their successive recall. Landmark information thus seems to be an important factor in orientation during childhood. In the actual literature we found evidence of the fact that children can make use of both types of landmarks, but this finding requires further specifications. Allen and Ondracek (1995) studied improvement in children's performance in relation to age-sensitive cognitive abilities and found that the acquisition of landmark knowledge improved with age. They showed that spatial memory was significantly involved in the acquisition of route knowledge and that both improved with age. It must remembered that, as regards the modalities of development of visuo-spatial memory, several researches evidenced how memory abilities improve with cognitive development, above all between the six and the ten years of age.

There are many ways to classify spatial reference systems. In this regard Levinson (1996) draws a distinction between egocentric and environmental reference systems. Egocentric reference systems specify location and orientation with respect to observer body-centred coordinates. Environmental reference systems define spatial relations with respect to elements of the environment, such as the perceived direction and landmarks. Environmental research theory distinguishes between two types of environmental knowledge: route knowledge and configurational knowledge. Route knowledge includes important landmarks in the environment, the routes connecting them and the order of route turns (relational directions such as right, left, straight ahead) in way finding. Configurational knowledge refers to a more "global" representation of the environment according to an Euclidean reference system. Cardinal directions and metric distances serve

as coordinates to map spatial relationships among distinctive locations within a network of routes (Schmitz, 1999).

The hypothesis is there are mechanisms of visuo-spatial memory that depend both on subject characteristics and on its environment. We designed our experiment to test whether: I) the performances of subjects differ in a visuo-spatial memory task as a function of age and sex; II) in a reconstruction task the elements playing the role of landmarks are recalled first and put in a correct location, and if they help to recall the surrounding elements; III) some characteristics of elements, like form, size, color, help subjects to put them in the right location.

2. RESEARCH

2.1 Subjects

Subject sample was constituted by 100 children, 48 male and 52 female, of age between the six and ten years old, attending at an elementary school in Cagliari. Every age class included twenty children, sorted with random sampling.

2.2 Material

Since this research was done on school children, we used easy and amusing material. For this reason the stimulus consisted of a image coming from a comic strip (Lupo Alberto), depicting a small farm with a way. The image was composed by 59 elements (houses, trees, bridges, etc), eight of which were characters. Those characters were identified with the landmarks, because they had different physical characteristics as regards colour and form, and in the image they were far one from another. The stimulus represented a survey knowledge and this allowed to see the main components and their interrelationships in space.

The stimulus image was manipulated through a computer program and we made other two stimuli corresponding to the two difficulty levels used in the experimental task. In the first difficult level there were the characters and the way, and we eliminated the other elements like houses, trees, etc. In the second difficult level there was only the way, and we eliminated all other elements.

2.3 Method

The experiment was consisting of two phases. In the first one the children saw the stimulus image for 10 minutes. Subsequently we introduced an interference task that consisted in singing a song for ten minutes. After we gave them the image corresponding to the first difficulty level and a small bag containing the elements previously cut. The task was consisting in the correct re-location of the elements previously observed and all elements that were re-located must be labeled according to the recall sequence.

After a week, we gave directly to the children the stimulus corresponding to the second difficulty level and the small bag with the missing elements. The task was like the first one. In this case the children didn't see the stimulus image because we investigated if the landmark helped the re-located in a correct way.

The administration was collective but the children performed the task individually and were placed at a suitable spatial distance one from the other.

3. DATA ANALYSIS AND DISCUSSION

The data on performances of subjects have been analyzed through a multivariate ANOVA for 5 (class) × 2 (level) × 2 (sex) × 6 (form), a multivariate ANOVA for 5 (class) × 2 (level) × 2 (sex) × 3 (colour), a multivariate ANOVA for 5 (class) × 2 (level) × 2 (sex) × 2 (size); whereas we performed a one way ANOVA for evidencing the presence of significant differences for sex and age in performance on elements located immediately near the landmark in first and second phase separately.

For the single factor class we have obtained significant differences in all three multivariate ANOVA. In the first it was significant the factor form ($F=5.5$; gdl=5,450; $p<0.001$), in the second the factor colour ($F=178.4$; gdl=2,180; $p<0.001$), in the third the factor size ($F=1108.6$; gdl=1,90; $p<0.001$), (see Figure 1).

As regards the factor form we have divided the elements in six categories: the tree, the house, the field, the bush, the bridge and the stockades. For the factor colour we used three categories: brown, red and green. For the factor size, the elements have been subdivided in two categories: large and small; as criterion of separation between the two categories we used the diagonal of the rectangle containing the element, compared to the one of elements of average size.

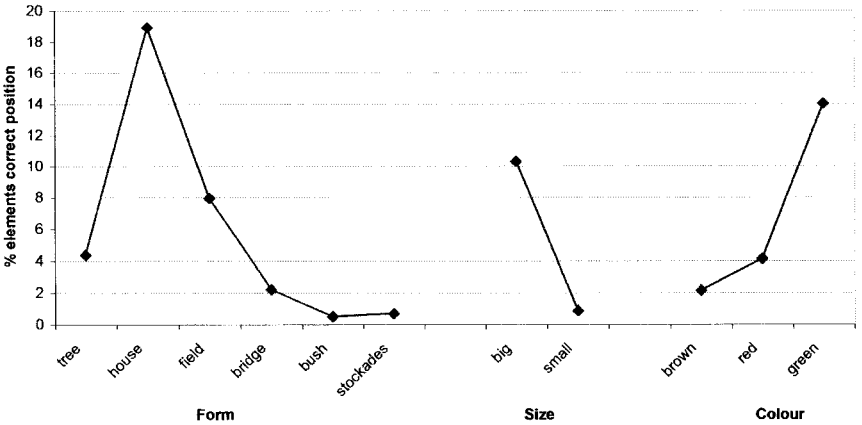


Figure 1. Factors: form, size, colour.

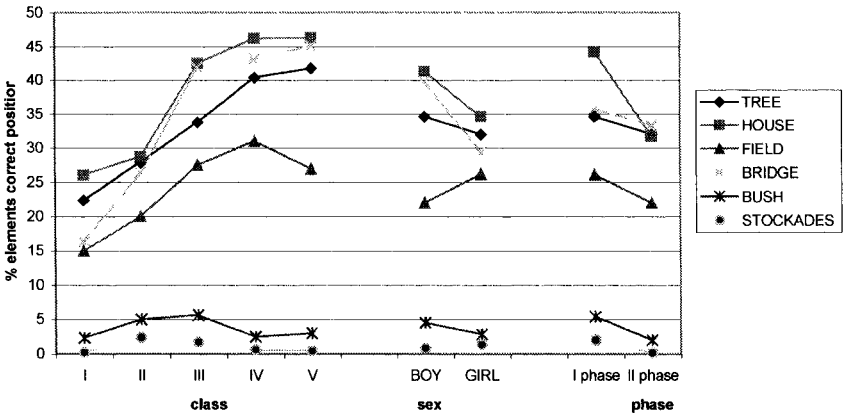


Figure 2. Form × Class; Form × sex; Form × phase.

Significant differences for the following combinations of factors have been obtained: form x class ($F=3.94$; gdl=20,450; $p<0.001$), form x sex ($F=4.77$; gdl=5,450; $p<0.001$), form x level ($F=5.70$; gdl=5,450; $p<0.001$), colour x class ($F=4.15$; gdl=8,180; $p<0.001$), colour x sex ($F=8.55$; gdl=2,180; $p<0.001$), size x class ($F=16.14$; gdl=4,90; $p<0.001$), as evidenced from Figures 2,3,4

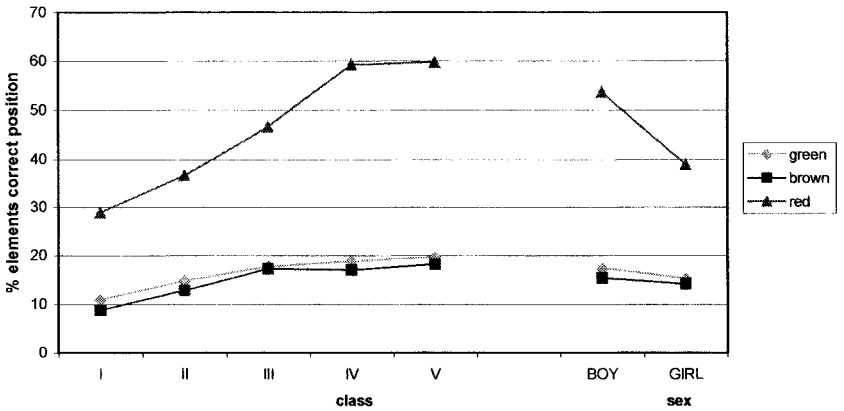


Figure 3. Colour × Class; Colour × Sex.

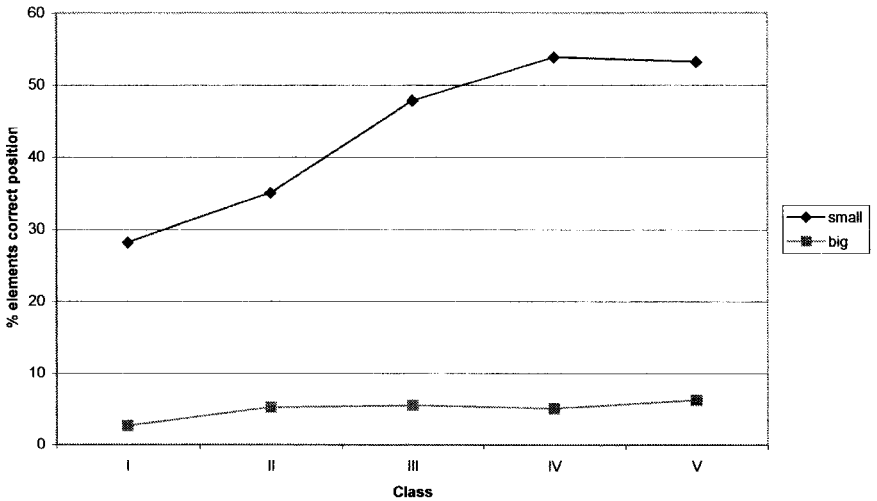


Figure 4. Size × Class.

As regards the relation between the percentage of correct positioning of single landmarks and the factors constituted from the type of landmark and the class, the ANOVA evidenced significant differences both in the first and in the second phase. In particular in the second phase it was significant the effect of the type of landmark ($F=109.96$; $gdl=7,792$; $p<0.001$), while for all landmarks it was significant the factor class.

Finally we obtained that the percentage of times in which the children re-located between the first ones a landmark depends on the type of the landmark itself, as evidenced from Figure 5.

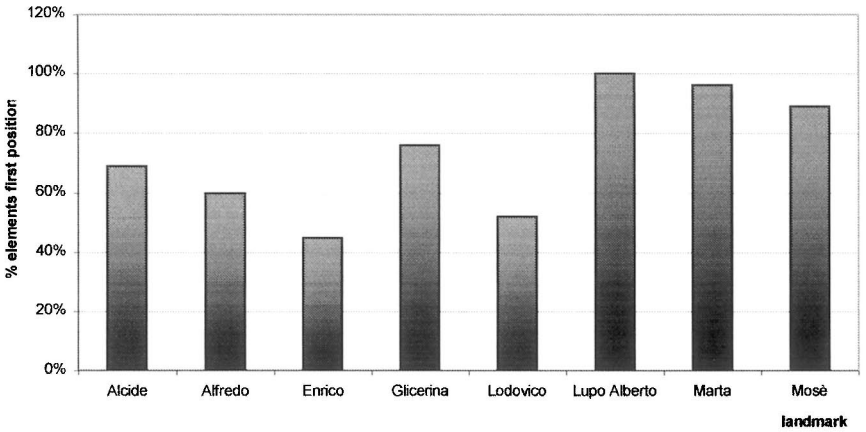


Figure 5. Sequence Landmark Position.

4. CONCLUSION

The data analysis confirmed our hypothesis that the performances of the subjects in tasks that imply the visuo-spatial memory differ as a function of the age; however the spatial competence and the knowledge about environment organization are already present in the children of six years. Then they develop during all the evolutionary age, also improving their use and the efficiency of metacognitive strategies.

As already evidenced from DeLoache and Brown (1983), another influence exerted by age was on recall criterion. The older subjects performed so well in the landmark condition because they actively integrated the landmark cue with the stimulus image. The younger subjects may have failed to encode the spatial relationships starting only from the position of the elements, and may have neglected to draw on this information for retrieval.

Moreover the results confirm that in a task of reconstruction of spatial structure that used a survey knowledge the elements playing the role of landmark are essential not only for the acknowledgment of a distance but also for the correct location of elements placed in their neighborhood. Besides these findings show that landmark elements were chosen not only

for an emotional-affective valence but also because they had particular physical characteristics, like size and colour. The data analyses showed that the main character of the comic strip (Lupo Alberto) has been positioned as first in 100% of cases, while the character with the smaller size (Enrico, the mole) has been positioned as first in a smaller percentage of cases. These data confirm the hypothesis that some characteristics of the elements, like form, size and colour, effectively help the subjects to put in the right position the elements. Namely the elements of smaller size were associated to worse performances, while those of greater size were better remembered.

REFERENCES

Baddeley, A. D., 1991, *La Memoria di Lavoro*, Raffaello Cortina, Milano.

Baddeley, A. D., 1995, *La Memoria Umana*, Il Mulino, Bologna.

Allen, G. L., and Ondracek, P. J., 1995, Age sensitive cognitive abilities related to children's acquisition of spatial knowledge, *Developmental Psychology* **31**(6):934-945.

DeLoache, J. S., and Brown, A. L., 1983, Very young children's memory for the location of object in a large-scale environment, *Child Development* **54**:888-897.

Flavell, J. H., 1985, *Cognitive Development*, Prentice-Hall, Englewood Cliffs, NJ.

Lehnung, M., Leplow, B., Friege, L., Herzog, A., Ferstl, R., and Mehdorn, M., 1998, Development of spatial memory and spatial orientation in preschoolers and primary school children, *British Journal of Psychology* **89**:463-480.

Logie, R. H., 1994, *Visuo-Spatial Working Memory*, Psychology Press, Hove, UK.

Mallot, H. A., and Gillner, S., 2000, Route navigation without place recognition, what is recognised in recognition-triggered responses?, *Perception* **29**:43-55

Pessa, E., and Penna, M. P., 2000, *Manuale di Scienza Cognitiva*, Laterza, Roma.

Piaget, J., 1976, *Memoria e Intelligenza*, La Nuova Italia, Milano.

Shelton, A. L, and McNamara, T. P., 2001, Systems of spatial reference in human memory, *Cognitive Psychology* **43**:274-310.

Shumann-Hengsteler, R., 1992, The development of visuo-spatial memory: how to remember location, *International Journal of Behavioral Development* **15**:455-475.

Schmitz, S., 1999, Gender Differences in Acquisition of Environmental Knowledge Related to Wayfinding Behavior, Spatial Anxiety and Self-Estimated Environmental Competencies, *A Jounal of Research Issues*.

7

MANAGEMENT

DYNAMICS OF STRATEGY: A FEEDBACK APPROACH TO CORPORATE STRATEGY-MAKING

Vittorio Coda[1] and Edoardo Mollona[2]
[1]*ISEA, Università Commerciale Luigi Bocconi, Milano, Italy*
[2]*Department of Computer Science, Università degli Studi di Bologna*

Abstract: The object of the article is a company's strategic management processes. The aim is to propose a dynamic model to explain how a company's realised strategy does emerge from interactions of purposes, tensions, and pressures dynamically interplaying. The paper contributes to strategy literature in two directions. First, we expect the model will be useful to management as a reference frame for understanding and efficiently governing a company strategy-making behaviour, both in cases in which the aim is to transform it radically, and when it is to be innovated by means of gradual evolutive change. Second, the model constitutes a set of hypotheses to orient further empirical and theoretical analysis. The analysis which we conduct, examining theoretic contributions and empirical settings, is strongly influenced by the assumption that the subject of the strategic government of companies may benefit from a systemic approach which considers the dynamic interaction among the many processes which impact a company's situation.

Key words: strategic and organisational change; strategic management; system dynamics; organisational complexity; chaos.

1. INTRODUCTION

The object of the article is a company's strategic management processes. The aim is to propose a dynamic model to explain how a company's realised strategy does emerge from interactions of purposes, tensions, and pressures dynamically interplaying. The paper contributes to strategy literature in two directions. First, we expect the model will be useful to management as a reference frame for understanding and efficiently governing a company

strategy-making behaviour, both in cases in which the aim is to transform it radically, and when it is to be innovated by means of gradual evolutive change. Second, the model constitutes a set of hypotheses to orient further empirical and theoretical analysis.

The analysis which we conduct, examining theoretic contributions and empirical settings, is strongly influenced by the assumption that the subject of the strategic government of companies may benefit from a systemic approach which considers the dynamic interaction among the many processes which impact a company's situation. Markedly, the strategic processes we focuse on are the learning processes which lie at the origin of top management's strategic intents and mental models; the managerial processes in which top management's actions are made clear; the entrepreneurial behaviour both induced by top management or which develops autonomously. All these processes unfold in environmental contexts which are usually intensely changeable.

2. THE STRATEGIC MANAGEMENT PROCESS IN LITERATURE

Literature on firms' strategic management process has investigated how a company's both realised and intentional strategy is defined and what the relevant activities are in the strategy-making behaviour.

In particular, the theoretical contributions on the subject of strategy process have assumed different positions with respect to the following problems:

1. interpretation of the strategic management process as a purely analytical-rational process or as a complex learning by doing process;
2. interpretation of the strategic management process as a top-down or bottom-up process;
3. interpretation of the role of top-management in governing the process.

For example, in the Harvard tradition, which gave rise to the schools which Mintzberg (1990a, b) re-christened Design School (Andrews, 1971) and Planning School (Ansoff, 1965, 1979, 1984, 1991), the strategic management process is of an analytic-rational type in its formulation phase and also in its realisation phase (this latter essentially intended as the design and implementation of an organisational structure, in the broad sense, logically descending from the content of strategic choices). This set-up is decidedly top-down and is based on the hypothesis that the decisions are totally rational and the realisations are logically consequential.

On the other hand, Normann (1977) highlights the fact that the formation within firms of a business idea (in other words of a successful strategy) is always a learning by doing process. On similar lines, Mintzberg (1978, 1979, 1984) eliminates the separation between thought and action. Strategy becomes the resultant of a learning process which proceeds along a parallel path: that of a decided strategy which embodies the top-down, analytical-rational aspect and that of an emergent strategy, the fruit of a trial by error process in which there is a strong bottom-up component (see Figure 1).

According to Bower (1970), Burgelman (1983a, b, c, 1991, 1994) and Noda & Bower (1996), the strategic process is essentially bottom-up and CEOs, although, on the one hand, playing a fundamental role in establishing a company's strategic and organisational context, in which the strategy takes shape, on the other, limit themselves to "adjust" company strategy a posteriori, relying on and rendering official the results of the strategies which have survived the selective pressure of the company's strategic and organisational context. Within this group of contributions, which intend strategy as the result of a continuous learning process, rather than as the result of an a priori analytical process, we may also position the contribution of Quinn (1980, 1981), which sees strategy as a logical incrementalism process in which company leaders channel flows of activity and events into conscious strategies[12].

Lastly, along these lines, a valuable contribution to the analysis of firms' strategy-making derives from evolutionary economists Nelson and Winter (1982).

According to their theory, environmental changes impose learning processes in which inefficient routines are replaced by efficient ones and a firm's top management has the task of accompanying the learning process by facilitating the elimination of inefficient routines, removing the difficulties in transmitting the change within the organisation and stimulating the change of the routines by means of innovation and imitation (Mintzberg, Ahlstrand and Lampel, 1998). However, the real agents of the evolution of company

[12] As Mintzberg, Ahlstrand and Lampel (1998: 180-182) point out, Quinn can be considered as an exponent of the Learning School because of the emphasis placed on the incremental component of strategy. However, the authors highlight a certain ambiguity which could position half way between the Learning School and the Design School. In fact, in certain contributions, Quinn describes the shaping of strategy as a process of which the CEO has a very clear idea, a priori, of company strategy and incrementalism is the fruit of the realisation effort, which has to pass through the gradual creation of the necessary political conditions for the strategy to be accepted.
Therefore, incrementalism could be said to be the fruit not so much of a learning process within the strategy definition process, as the outcome of the difficulty in controlling political coalitions within the company.

strategy are said to be the subsystems in which the organisational routines are localised.

Because of the emphasis assigned to learning processes in forming company strategy, the contributions we have listed – from Norman to Mintzberg, Bower, Burgelman and Noda and possibly Quinn, and finishing with Nelson and Winter – can be traced back to a line of thought which Mintzberg, Ahlstrand and Lampel (1998) have called the Learning School.

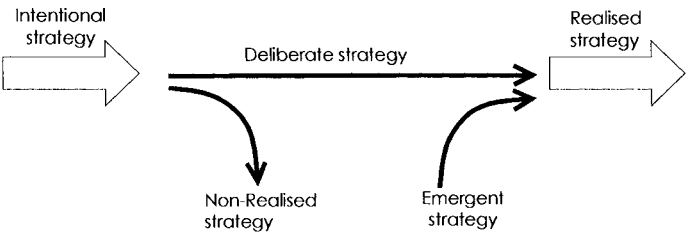


Figure 1. Mintzberg's Model of Strategy-Making (1978).

3. A NUMBER OF OPEN PROBLEMS

As we have seen, the contributions of Norman, Mintzberg, Bower and Burgelman highlighted a number of fundamental aspects, like the decisive role of learning and the spontaneous, emergent component of the company's strategic activity. In addition, these contributions have helped us to re-interpret the role of the company leader, re-dimensioning his "heroic" content (Burgelman, 1983a) as the enlightened guide, in perfect control of the situation, and highlighting his no-less important qualities as the designer, or the architect, of complex systems.

Analysis of strategy-making processes is closely related to the investigation of firms' adaptation capability to environmental changes. Mintzberg (1991) and Ansoff (1991) debated on whether a top-down strategy-making process or a learning-by-doing, emergent process was better suited to face dynamic environments and Burgelman (1991) suggested that radical strategic renewal is possible only by means of bottom up, strategic processes. In the mentioned literature, the interpretation of the link between the morphology of a firm's strategy-making process and the firm's adaptation performances takes two avenues.

First, the feature of strategy-making processes affect adaptation behaviours by defining a specific trade-off between top-down/global versus bottom-up/local rationality in shaping the direction of change. Top-down

driven adaptation behaviours have the advantage of relying on global rationality and information sets and of taking advantage of full control on resources. On the other hand, bottom-up driven adaptation relies on local rationality and information, and may dissipate resources on projects dissaligned from a firm's core competencies. However, bottom-up adaptation may be more effective when radical change is required. In these cases, cognitive inertia is an impending menace and self-referential attitudes are dangerous.

Second, the characteristics of strategy strategy-making processes influence adaptation behaviours by differently dealing with an organisation's structural and cognitive inertia thereby deciding the pace of change.

Yet, the literature lacks a way to rigorously compare the structures of different strategy-making processes in the light of diverse weight assigned to global versus local rationality or to clearly elicit the loci where structural and cognitive inertia can be observed within a strategy-making process.

3.1 Local vs. Global rationality and learning: A feedback approach:

Although Normann and Mintzberg and Bower-Burgelman clearly highlight the spontaneous, emergent component of strategy, based on learning-by-doing processes, the nature and protagonists of these processes need to be further clarified.

It is one matter to say that a CEO learns because he, or she, observes the result of the enacted strategic action. It is another matter to state that the CEO monitors, approves and includes *a posteriori* into a company strategy the results of emergent strategic initiatives generated by front-line managers or other collaborators like researchers or people close to customers and the market, who are not necessarily members of the company top management.

In the first case, we shall not distance ourselves greatly from the existence of a unique and global rationality as the motor of strategic change; the learning process is an individual one: the leading player in the CEO who, by observing the facts and emergent information, re-examines his strategic intents, or does not re-examine them, but learns to be more effective in the actions taken to realise them. In this case, we shall have a process in which, starting from Mintzberg's model in Figure 1, information deriving from the realised strategy reaches the CEO who processes it and promotes the generation of emergent strategies (loop 1 in Figure 2). Secondly, the new realised strategy which now also incorporates the results of the emergent strategies can help make a contribution towards changing the intentional strategy (loop 2, Figure 2). This situation, in which the top manager is personally involved in a entrepreneurial-like, strategic innovation activity,

remind us of typical situations of small-medium sized companies in which the articulation of the hierarchical levels and degree of complexity to be managed are of a low level.

The process is substantially different if we consider large, complex companies with an articulated organisational structure. In this case, the CEO has a different role in governing retroaction loop 1 in Figure 2. In the first place, the CEO increasingly manipulates the strategic-organisational context in order to induce[13], rather than to develop personally, emergent strategic initiatives in which the contents come under an 'umbrella strategy'[14].

Then, the CEO decides to what extent to approve or discourage strategic behaviour or initiatives which, on their conception, do not come under the company's umbrella strategy. If he decides to allow strategic initiatives of this type to germinate, via loop 2 of Figure 2 the CEO will have to adjust the umbrella strategy a posteriori in order to incorporate the content of these emergent strategic innovations.

It is clear that in the second interpretation of the strategic learning process, the agents who contribute to learning are distributed throughout the entire company system and the functioning of loop 1 appears to be the result of a choral effort. There are actors who, bottom upwards, through loop 1, gradually enrich the operating strategy with new contents, both within the confines established by the company management and in totally new directions; and, on the other hand, top management, through loop 2, makes its contribution both by deliberating and realising intentional strategy, thereby creating the premises for learning and emergent strategies, and by adjusting, along the way, the strategic aims in the presence of emergent initiatives. These latter, although they were not originally included under the umbrella strategy, may appear to be valid and promising.

A sort of 'specialisation' is established in which the role of top management is to conceive of the strategic-organisational context, designing company strategy, for example by outlining the company business portfolio strategy, with the middle and front-line managers engaged in developing specific strategic initiatives like, for example, the development of new products. In this case, learning is articulated within the whole corporate and strategy evolves as a result of the aggregation of contributions received from various areas and different hierarchical levels within the organisation.

A third interpretation of the strategic learning process, which could be configured as an extreme case of 'specialisation', is the one described by

[13] For this type of emergent strategic initiatives, Burgelman uses the concept of induced strategic initiatives (Burgelman, 1991).

[14] For this type of emergent strategic initiatives, Burgelman uses the term 'autonomous strategic initiatives' (Burgelman, 1991).

Burgelman for the Intel case (1991): having gradually shifted its production from semiconductors to microprocessors, after ten years, the company officially states that it has left the semiconductor business[15].

The nature of this last type of learning process could by described, in Figure 2, by saying that the company management governs loop 2 while the front-line management or other members of the organisation manage feedback loop 1. In other words, we may say that the Intel case appears as an extreme version of Burgelman's description of the strategy formation process in which the sharing of tasks in the context of the strategy formation process is taken to it extreme consequences and a company management which governs loop 2 and other figures performing a leading role are identified in loop 1.

In the company reality, it is probable that learning processes, in which the contribution and role of top management has different weight, co-exist and become confused. However, with a view to understanding and governing the mechanisms at the base of the evolution of company strategy, a distinction should be drawn between the characteristics of processes of a different nature.

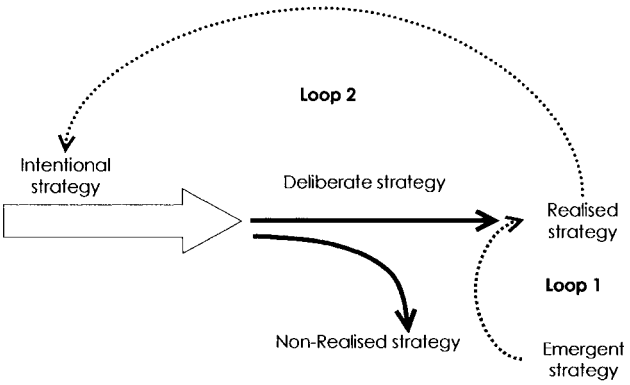


Figure 2. Feedback Thinking in Strategy-Making.

Mintzberg does not use the concept of the feedback loop and focuses his contribution on identifying a generic learning process in which it is difficult

[15] Looking closely, this second feedback loop, which describes the mechanism which generates and adapts strategic intents, was not considered by Mintzberg who, nevertheless, albeit at an implicit level, considers loop 1 when he explains that emergent strategy takes shape from the learning triggered by the attempt to realise the strategy. Therefore, we imagine that, while the strategy is realised progressively, it produces observable results which become the starting point for learning.

to single out the truly spontaneous and evolutionistic component of strategy formation.

On the other hand, Burgelman is interested precisely in highlighting the feedback mechanisms which lie at the base of the formation processes of strategic aims and of emergent strategies in a big company in which different levels of management and an articulated network of players must be identified. In the contribution of the Author, these are presumed to contribute in different ways and to a different extent to the formation of company strategy.

3.2 Exploring sources of inertia with stock and flow diagramming

In management literature, symbolic languages used to describe cause-effect relationships seldom includes the notion of stock and flow variables. Yet, to provide managers with appropriate operational and conceptual tools to govern strategy dynamics it becomes essential to distinguish between flow variables, which may be employed to represent strategic processes, and stock variables, which are results of strategic processes and generate inertia in a firm's strategic behaviour.

Returning to Mintzberg's diagram shown in Figure 1, we note that some of the concepts illustrated refer to processes and that others appear to be more specifically the observable results of the processes themselves.

For example, considering intentional strategy, this would appear to be an observable result of such processes as planning or environmetal analysis.

As far as deliberate strategy, we might consider it as a construct encapsulating a number of implementation processes leading to the realisation of intentional stategy. On the other hand, we might well consider deliberate strategy as a further stock variable in which intentional strategy is translated into an officially stated list of objectives and projects explicitly communicated to, for example, shareholders. In this case, we ought to elicit the processes that transform intentions into deliberate strategies and those latter into implemented strategic actions.

Concerning realised strategy, it is the product of implementation processes, these latter observable for example in a company's physical-technical, organisational and cltural endowment. But what should we understand by emergent strategy? Should we intend the processes which modify realised strategy from the bottom up as a result of strategic management, or as both?

The distinction between processes and results is a necessary starting point for an accurate description of the strategic management mechanisms in a company. If we consider emergent strategy as a product, it is interesting to

understand where and how we observe this product. For example, we could identify emergent strategies with the various initiatives enacted without official support from top management, with the ongoing research and development projects or the experiments and trials that top management itself wishes to encourage in order to adjust strategy 'along the way'. If, on the contrary, we consider emergent strategies as processes, we must understand the morphology of these processes and identify the relevant sub-processes.

Describing appropriately a feedback dynamics means to rigorously distinguish, and tease out the interplay between, processes and products of processes; the former observed over periods of time and the latter in different points in time. For example, intentional strategy, as well as being the product of certain processes (strategic planning, visioning, etc.), defines a desirable situation which orientates and directs managerial actions aimed at achieving it. Realised strategy, as well as being the product of top-down and bottom-up executive processes, defines a context from which learning processes unfold in-field, resulting in operational innovations and strategic initiatives.

3.3 Complexity and chaos theory

To investigate the variety of possible emerging strategic behaviours and to address the paradox of reconciling order and experimentation in strategy making, a number of management scholars have borrowed conceptual tools from complexity and chaos theory[16]. Complexity and chaos theory have handed down useful metaphors to support theory creation by analogical reasoning. A fundamental line of enquiry is the investigation of how firms ought to manage the trade-off between order and control, on the one hand, and experimentation and emerging, self-organised, behaviours, on the other. Chaos is a state in which a dynamic system does not tend to a stable equilibrium, does not evolve along an unstable exploding trajectory and is not attracted towards a defined oscillatory path. Rather, in the system, stability and instability coexist and are so tightly intertwined that emerging behaviour is unpredictable and seemingly random. As complex dynamic systems show chaotic behaviours when they bounce among different types of attractors, similarly organisations are potentially chaotic systems since they are characterised by counteracting forces. Thietart and Forgures (1995), for example, suggest that organisations are contemporaneously pushed by opposite forces: planning, structuring and controlling, which push towards

[16] Stacey (2003) provides a review of contributions in which complexity and chaos theory inspire enquiry into management issues.

stability and order, and, on the other hand, innovation, initiative and experimentation, which push towards instability and disorder. Along similar lines, Brown and Eisenhardt (1998) advise that organisations ought to be managed 'at the edge of chaos', that is, in a state where structure and freedom coexist. Key issue for top managers is to find the appropriate balance between self-emerging and top-down-induced behaviours. Aiming towards a chaotic state allows firms to discover new and unexploited strategies, yet, stopping just before prevent firms from dissipating resources into undisciplined strategic behaviours.

Yet, we think that the concept of 'edge of chaos' might be developed in, at least, three directions.

First, a question emerges concerning leadership and the role played by top managers. To what extent managers can govern companies at the edge of chaos? How do managers sense when are they structuring too much their organisations or when they are dangerously relaxing control? In the light of chaos theory, leadership seems to be associated to system understanding and design capabilities.

Second, Brown and Eisenhardt suggest that organisational structure is the lever to be manipulated in order to push an organisation to the edge of chaos calibrating the level of stability and instability. Yet, organisations may have heavy structures and, yet, be unstable because they repeatedly change their structures. Indeed, as Hannan and Freeman suggest (1984), as organisations get structured and inertial, change becomes difficult and its outcomes are unpredictable and seemingly random. Thus, instability accrues from rate of organisational restructuring as much as it derives from lack of structuring.

Third, restructuring may be the product of top down processes, when top managers conceive of and implement change, or the result of bottom up pressures, when change results from internal venturing and entrepreneurship (Burgelman, 1983 a,b,c), this latter more or less decoupled from official organisational strategy (Burgelman, 1991). Thus, we expect organisational change to have a different impact on the degree of stability and instability within organisations, depending on whether it is top-down or bottom-up. We suppose that the more restructuring is conceived bottom up, the more emerging organisational change will be disordered and its outcomes unexpected.

Concluding, we suggest that to explore further the concept of managing at the edge of chaos, fruitful directions are both the analysis of the rate of organisational change and the logic by which the change is conceived.

4. GOVERNING THE DYNAMICS OF STRATEGY: A SYSTEMIC MODEL

In this section, we propose a model of the strategy's dynamics. The description of the model is articulated into three phases.

The first phase highlights the level variables, which represent the state of a system at a certain point in time as the result of one or more processes.

The question we asked ourselves in this first phase is the following: if we imagine a company's strategy as a dynamic system, in which various types of processes are entwined, at a certain point in time, what are the observable products of these processes? In other words, if we imagine that we can freeze the company strategy system in a certain moment, what are the stock variables which will crystallise its state?

The second phase in the construction of our model is the description of the processes, the flow variables, which both are influenced by information concerning the state of the system, the stock variables, and affect the state of the same.

Lastly, the third phase consists of highlighting the causal chain which links stock variables and flow variables[17].

The model proposed is rooted both in relevant literature and in a set of longitudinal case studies. In particular, the construction of the model and the selection of the variables is based on a grounded-type approach (Glaser and Strauss, 1967), based on analyses of the cases of companies involved in important strategic-organisational change processes[18].

[17] In this paper use is made of a symbolic language based on the distinction between flow variables and stock variables in order to represent economic processes. In proposing such a logic, we make explicit a number of assumptions that, we think, permeate, implicitly, thinking in strategic management. Not only this is true as far as anglosaxon management literature is concerned but also in European tradition. For example, the representation of the dynamism of economic processes based on the flow-stock dichotomy was implicitly inserted into the heart of Italian business economics tradition as early as Zappa. On the one hand, Zappa claimed that the movement is usually represented as a "sequence of states" in which the accumulation of previous variations is observed (Zappa, 1957: 930-931), and on the other, he suggested that in order to fully understand production phenomena, the mere association of a sequence of states is not sufficient, but that the definition of times and durations is also necessary in order to describe the processes "within the unit of time". The portrayal of dynamic phenomena centring on the distinction between sequence of states, which vary in their accumulation, and processes, defined in the unit of time, is therefore similar to the representation based on stock levels, which at a certain moment in time represent the state of a system following successive accumulations and on flow variables which describe the rate of variation of the stock variables within a certain period of time.

[18] The central body of the empirical research consists of a clinical analysis conducted with a 'grounded-type approach (Glaser and Strauss, 1967) to the case of IBM between 1993 and 2000. The analysis of IBM's case was conducted by means of open-ended interviews

4.1 Stock variables

Guided by the analysis of the literature and the information drawn from the empirical analysis, we selected four significant stock variables to describe the state of the company strategy system at a certain moment in time[19]: the Mental models of CEO, Intentional strategy of the CEO, Realised strategy and the Portfolio of strategic initiatives.

Mental models of CEO encompasses its key strategic orientation (Coda, 1989), it is a stock variable which incorporates the key beliefs and cognitive framework of top management, his or her basic values, convictions and attitudes formed over time as a result of accumulated experience. Although not tangible, CEO's mental models are pillars upon which hinges a company's strategy system; in fact, all the formulation and realisation processes of strategic intents, analyses, and the interpretation and control of results, are permeated by interpretative patterns which have been consolidated over time (Argyris, 1982; Argyris and Schon, 1978).

The second stock variable which we consider to be relevant is the CEO's intentional strategy. This variable includes both the strategic goals and intents and the possible plans for achieving them. The concept may also include the strategic intent proposed by Hamel and Prahalad (1989), which evokes a desired market leader position and the criteria for monitoring the approach to this position. Strategic intents can be drawn, for example, from official documents like the report to shareholders or the statements issued by top management during interviews, press conferences, meetings with collaborators and other events in communication.

The third stock variable we shall deal with is realised strategy. This variable defines the structure of the company system operating at a certain

---

(about 20 interviews with managers who in the period considered held important positions), the study of balance sheets, the examination of internal procedures and documents (memos, e-mails, manuals) and the collection of information published in newspapers, specialised magazines and previous studies. The analysis of IBM's case also formed the starting point for the development of subsequent clinical analyses of company case histories aimed at corroborating the constructs and the relations between constructs. In addition, in order to make the description of the grounded model more vivid, where this appeared to be plausible, reference was made to empirical cases, like that of General Electrics, for example, analysis of which was conducted above all on secondary sources.

[19] By analysing the literature, we identified a series of wide categories of concepts (for example the intentional profile of company strategy is linked - albeit with different facets - to both Mintzberg's concept of intentional strategy and to Hamel and Prahalad's strategic intent (1989)). The categories formed in the analysis of literature were then compared with the constructs and relations between constructs which emerged from the empirical analysis.

point in time in a given environmental context – therefore, for example, strategic positioning, the organisational set-up, actual company culture and values – and the variables which indicate its economic-financial, competitive and social performance.

In the description of the state of the system, the latter may come into play as rates, or rhythms, but are observable results when accumulate into stock variables (for example, monthly profit or loss production rates, or monthly billing production rates) or of relative performance levels (for example, level of customer satisfaction or level of staff motivation). Realised strategy is therefore an extremely complex aggregate variable which indicates the concrete situation in which the company's top management finds itself at a given moment in time.

Lastly, the final stock variable we highlight is the portfolio of strategic initiatives and innovations; it embraces: projects and business ideas, in the experimentation and development phase.

4.2 Flow variables

Having described the stock variables, we now pass to a description of the processes, represented by flow variables which, over a certain interval of time, modify the state of level variables.

We consider five groups of processes which modify the state of the four stock variables described previously: mental model learning processes; intentional strategy realisation processes, innovation generation processes and innovation selection and realisation processes.

The first group includes CEO's learning processes, in other words, the processes which affect top management's mental models, enriching, modifying and/or strengthening its content. By observing the results of its decisions incorporated into the realised strategy, the members of top management learn and evaluate the suitability of their mental patterns. For example, they adjust their ambitions according to whether they have been seen to be unachievable or not very challenging, or they adapt their basic beliefs to the attitudes of members of the organisation or other stakeholders, having observed their behaviour.

The second group considers the intentional strategy formation processes, i.e. all the processes which are responsible for the formation of the contents of a desirable strategic outlook and therefore worthy of being achieved. We may assume that the strategic intention formation process emerges as the combination of various subprocesses.

In our empirical analysis, we reported, for example, the following sub-processes: company and environmental analysis subprocess; subprocess in which the top management's ambition for challenging goals takes shape,

which we define as visioning; sub-processes of benchmarking competitors to elicit strategic directions; analytic-rational and organisational strategic planning subprocesses and a-posteriori learning subprocesses through which top managers include retroactively within the company's official strategy results from bottom-up strategic initiatives (Burgelman, 1991).

Our empirical analyses revealed the presence of these subprocesses with differing levels of importance and in different proportions. Dissimilarities may be explained, in our study, by referring to the degree of organisational formality/informality, to the level of top management's ambition, to the more or less participatory style of leadership, to the ability to conduct an in-depth analysis of the problems, to the CEO's conception of his role and his way of interpreting it.

In a third group, we include intentional strategy implementation processes, i.e. the managerial processes which result from the strategic intentions and are aimed at ensuring that the latter are realised.
These processes can be ascribed to the following classes:

a) processes of communication and sharing the intentional strategy;
b) processes for structuring a company's business portfolio;
c) processes that set up, or adjust, organisational structures and operating mechanisms;
d) processes that encapsulate the launching of company challenges and the projects into which these latter translate;
e) processes that crystallise managerial decision-making, such as planning, budgeting, controlling, and staff management.

The fourth group is that of the generation processes of innovations which include processes aimed at creating operational innovations and internal entrepreneurship processes which generate strategic innovations. The innovation generation processes are, in various ways, stimulated by the environmental opportunities and cultural and morphological characteristics of the company context. By morphology of the organisational context, we intend, for example, the characteristics of the mechanisms, formal or informal, via which internal entrepreneurship is stimulated or discouraged. As far as the informal mechanisms are concerned, the culture, history and folklore which permeate the life of a organisation and form a layer of accumulated information which reveals the widespread attitudes towards innovation.

On the other hand, these attitudes are frequently formalised into programmes or routines, systems of remuneration, promotion and stimulation. Suffice it to think, for example, of the 'melting-pot of ideas' created by General Electric at the end of 1988, which involved the constitution of a periodic forum among employees in which the latter could

present ideas and proposals on how to make their business more effective and have an immediate reaction to the initiatives presented. Inversely, at 3M, the stimulus to produce innovation is also created by the "15 per cent rule" which enables employees to allocate 15% of their time to work on ideas which they believe to have some development potential, and by the goal, imposed by the divisions, of having at least 30% of turnover originate from products introduced in the last four years.

The last group of processes to which we refer are the processes for realising and selecting innovations. These functions act as a filter on the various emergent initiatives. These 'filters' may be of an official type and be incorporated into formalised routines In this case, they assume the form, for example, of processes of periodic assessment of the economic-financial, commercial and strategic potential of the single emergent strategic initiatives or of feasibility studies of the latter. Or, the selection processes may concern the evaluation of costs, opportunities and the possible recovery of efficacy and efficiency made possible by innovations of an operational type.

These processes are usually linked to resource allocation mechanisms which enable the initiatives to survive, grow stronger and finally emerge. On the other hand, there are also informal mechanisms which stimulate or discourage emergent strategic and operational initiatives. For example, as Burgelman (1991) explains, it sometimes happens that strategic initiatives can survive and be finalised outside the official evaluation and selection mechanisms.

4.3 Map of cause-effect relationships

Following the description of the flow and stock variables, we can proceed to link the variables by drawing the retroaction loops.

We show four fundamental loops: the strategic control loop, the strategic intent formation loop, the entrepreneurial loop and the learning loop of mental models.

4.3.1 Strategic control loop

The first feedback is loop 1 via which the realisation of intentional strategy is controlled. Once intentional strategies have been modelled and possibly articulated into strategic plans, the resulting realisation processes are oriented towards reducing the gap between strategic intents and realised strategy (Figure 3). Each time, in order to govern the dynamics of the system, the distance between goals and results is gauged and reduced by the means of realisation. Loop 1 is a mechanism which performs a strategic

control activity[20] and therefore behaves as a thermostat: it aims to preserve the homeostasis of the company strategy system when it is felt that that entrepreneurial formula in operation does not require adjustment, or to place the system on a new trajectory of evolution if the company management intends to change the operating entrepreneurial formula. We may liken this loop to the first-order learning loop of Argyris (1982).

The strategic control loop describes a company's ability to execute a certain strategy promptly and efficaciously. For example, on his arrival at IBM in 1993, Gerstner indicated dangerous shortcomings in the effective realisation of the strategy. According to Gerstner, the result was that the strategic plans were left on the shelf and never realised. It was not by chance that one of Gerstner's interventions was to introduce the concept of *execution*, i.e. the ability to realise or execute strategic intents rapidly and efficaciously, into management performance assessment[21].

In loop 1, the gap created between the desired situation and the actual situation, between goals and results, is a fundamental variable. Not only the size, but also the quality of this gap should be considered. In fact, on the one hand we expect that loop 1 is aimed at keeping the gap under control, at minimising it, so we are led to hope for a situation in which the gap is limited.

On the other hand, given the function of a stimulus performed by needing to close the difference between goals and results, it is a physiological fact that the gap is never totally eliminated and we must ask ourselves questions about the quality of the existing gap. This quality depends above all on the depth of the analysis of the situation to be strategically managed and on the values and ambitions at the basis of strategic intents: a great deal of ambition and superficial analysis or in-depth analyses not supported by an ethical conception of the company (warped by the interests of the controlling group) give a negative character to this gap, from which destructive tension is released.

Inversely, as an example of a gap creating creative tension, we should consider the one perceived by Hayek when, in 1984, he assumed the leadership of SMH. The quality of this gap is marked, on the one hand, by an in-depth analysis of the competitive problems and the situation of Swiss

[20] To increase the efficacy of the control loop management can use traditional strategic control tools or diagnostic control systems (Simons, 1995).

[21] Lucio Stanca, currently the Italian Minister for Information Technology, was until 2000 Chairman and General Manager of IBM EMEA (EMEA stands for Europe, Middle East e Africa); he reported to us the atmosphere of those years: "Gerstner told us, "You should not create strategy. I, myself, and the BRAND managers will create strategy. You carry out." Stanca adds "Previously, we all created strategy. We had bands of planners! In IBM Italy alone we had 300-400 planners. Gerstner forced us to emphasise execution".

watchmaking companies and, on the other, by basic values and beliefs which, in the light of the facts, proved to be extremely valid.

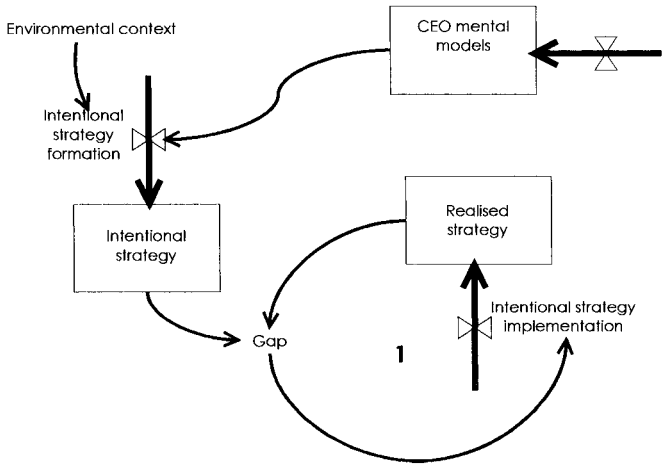


Figure 3. Strategic Control Loop.

The protraction of a gap which creates constructive tension is the starting point, or the spark, which starts off a process generating company efficiency and development. For example, in the case of General Electric, the strategic intent of being the number 1 or 2 in its businesses and the analysis of the actual competitive situation led to decisions regarding the company's portfolio (disinvestment of 200 businesses and 370 acquisitions); the goal of being a 'lean and flexible' business, compared with the high degree of bureaucracy previously existing in the company, led to reflections on organisational structure, the thinning out of jobs with the cutting of 50% of the strategic planning group's employees and the reduction of hierarchical levels from 9 to 4. The intent of becoming 'lean and flexible' then resulted in the definition of the 'improved practices' programme and the launch of the company challenge called 'surpassing ourselves'.

At IBM, in 1994, well ahead of other companies in the sector, Gerstner's strategic intent of winning the leadership in the business of services linked to the Internet or, more generally, to connectivity technologies, resulted in the decision to shift 25% of the research and development budget to projects related to the Internet and the creation of a study group which was to prefigure the characteristics of the new emergent sector and of the new products which had to be developed. After about a year's work, in September 1995, the study group presented its conclusions and in October

the fundamental decision was taken to allocate three hundred million dollars for the creation of the Internet Division.

4.3.2 Strategic intent formation loop

This second loop represents the process whose protagonists are the top-managers who draw useful indications from the realised strategy in order to re-examine and adjust strategic intents (Figure 4). We note that, in our model, we keep the two processes of the formation of strategic intents and the learning of values and mental patterns distinct.

Our hypothesis is that the fact that strategic intents can change according to the observation of the results of past actions does not necessarily implies that the basic beliefs and values of top management must also change. For example, a company can re-dimension its goals in terms of market share when it has seen that it was unable to achieve these goals.

Re-examining the goals in the elucidation of the strategic intents may have the aim of not 'stressing' the organisation at a given point in time and granting it the time to reorganise its forces, to then re-attempt to achieve the most challenging goal. Nevertheless, all this takes place without losing the profound belief that the ambitious market share goal can sooner or later be achieved.

For these reasons, we use the term 'learning' where there is a real adjustment of mental patterns and inversely, we use the expression 'formation of strategic intents' when the change to the strategic intents is not a result of updating mental models but, rather, of a gradual elucidation and awareness of one's own mental patterns or of tactical needs for managing the gap.

A concrete example of how this motor works is provided by an interview which Jack Welch, General Electric's CEO, released in the late 80's when GE's restructuring process had already been under way for a number of years: "In the mid-80's, the hardware part, or the organisational structure, was more or less at a satisfactory stage. We were pleased with our businesses. The time had come to tackle the software". In this interview, it emerges how the observation of what had been achieved led to the gradual enrichment of the content of the strategic intents without, however, changing the basic goals.

In conclusion, whereas in the strategic control loop observation of realised strategy is used to monitor the degree of realisation of the strategic goals, the latter, contained in the intentional strategy, in the strategic intent formation loop, observation of the realised strategy is preparatory to the adjustment of the goals themselves. In the first case, the goals remain

stationary and act as a reference point for control, in the second case, the goals evolve as the realised strategy changes.

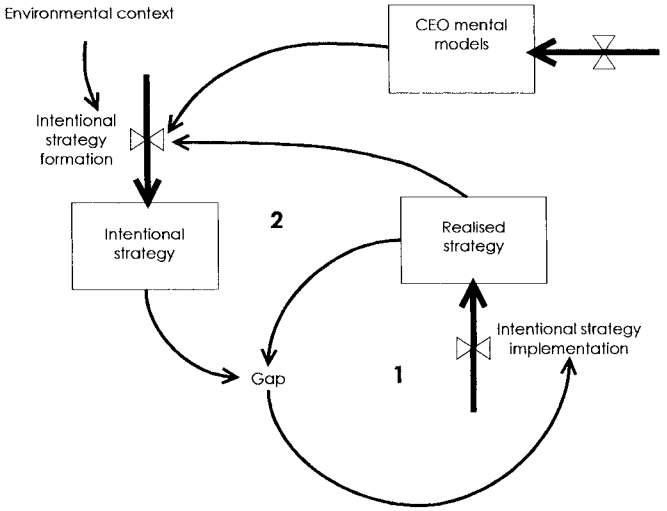


Figure 4. Strategic Intention Formation Loop.

4.3.3 Entrepreneurial loop

The third loop describes the bottom-up innovation processes which are the expression of internal entrepreneurship. In the diagram shown in Figure 5, loop 3 consists of a series of elements. The process pivots on a stock variable: the strategic and operational initiatives, the latter describe the results of the subprocesses which, positioned upstream and downstream of the stock, modify its level. The choice of a stock variable of strategic initiatives is the response to a precise demand for research: at a certain moment in time, what describes the energies, tensions and resources which are operating in order to generate innovations in the strategy of a company?

For example, the patents owned by a company represent the results of innovative initiatives after the latter have been selected, funded and realised and have become part of the realised strategy. On the other hand, the ideas and projects in support of which resources and energies have not yet been added indicate the richness and cultural fertility of a certain organisational context and are therefore elements of the realised strategy, although they do not yet constitute 'initiatives'. By representing the strategic and operational initiative variable, an attempt is made to 'photograph' the intermediate moment in time at which the stimuli and incentives present in the

organisational context have taken shape and combined into initiatives which have nevertheless not yet changed the strategic–organisational context and are still in the development phase. This photograph makes it possible to observe the processes upstream and downstream of the stock of strategic and operational initiatives. Upstream, the initiative generation processes which take place in the strategic-organisational context, feeding the stock of strategic and operational initiatives; downstream, the selection and realisation processes, by means of which the single initiatives are assessed and funded; these latter processes empty the stock of initiatives since, once selected and realised, single initiatives help to modify realised strategy and becomes and integral part of it.

In this way, realised initiatives define the cultural environment in which the subsequent initiatives will be conceived. This description is coherent, for example, with the contribution of Burgelman (1983a, b, c; 1991). The latter highlights how the strategic initiatives generated inside a company, which are fundamental elements of both incremental and revolutionary strategic innovation, are at the same time products of certain strategic-organisational contexts and triggers for changing these contexts.

Loop 3 describes the potential of large organisations for renewal. In fact, the behaviour of companies, and in particular, the ability to generate strategic and operational initiatives from innovative contents may remain confined to trajectories defined by the company's past history, with obvious problems of self-reference, or they may emerge as self-organised phenomena, totally new and unpredicted, in the sense that they originated not in a top-down fashion, or as the product of top management's rationality alone, but as a result of the repeated interaction of a strategic-organisational context with individual and local behaviour.

An example of how the mechanism of loop 3 can function is provided by the well-known case which describes the conquest of the USA motorcycle market by Honda in the late 50's – early 60's (1984).

The intentional strategy in 1958, when Honda had become the domestic market leader, was to embark on a process of internationalisation, starting from the California coast of the United States. This simple strategic intent resulted in decisions and actions which led to the setting up in Los Angeles of a tiny bridgehead made up of just a few men, with very few financial resources and a modest stock of motorcycles of all capacities, headed by a director in whose ability to get by with the few resources placed at his disposal, Mr Honda and his partner, Takeo Fujisawa, placed their complete trust.

On the field, this small nucleus of men was able to develop a radically innovative learning process which soon led to the discovery of the existence of a market for low-capacity motorcycles in the USA, about which no-one

had previously even thought, to the widening of the outlet for medium and large capacity motorcycles in a new segment, to the opening of new channels and to tip relations between contractual strength and the trade over in its favour. In conclusion, the bottom-up working of loop 3 provided the stimulus to explore a segment that was marked by a usage function so far neglected (the use of motorcycles as a means of healthy amusement) and rather more extensive that the one in which European manufacturers and Harley Davidson had been positioned.

This reported case of Honda is useful for exemplifying and putting into focus loop 3, precisely: the criticality of "realised strategy" as a body of variables which configure the strategic-organisational context in which "learning by doing" takes place. Within an firm's organisational context, strategic innovations emerge which constitute the heart of a successful entrepreneurial formula. Loop 3 helps to highlight the criticality of the relationship between top management and front-line management in defining the quality of the behavioural context in which the latter operates and concretely shaping the process of selecting and retaining innovations.

There are situations in which, within the frame of a realised strategy, both because the strategic intents are only generally outlined – for example in situations of environmental uncertainty - and because in the realisation of the strategy new opportunities emerge which enrich or change the contents of the strategic intents, the role of the patrimony of internal entrepreneurship diffused inside a company becomes fundamental.

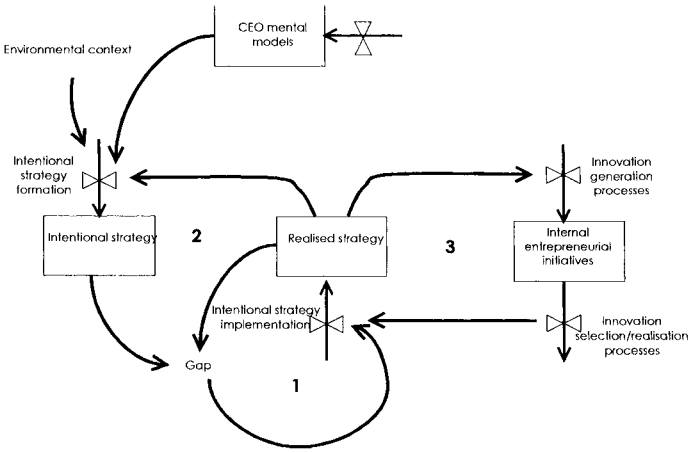


Figure 5. Entrepreneurial Loop.

In the case of Honda, the managers who were sent to the USA were able to exploit the resources placed at their disposal within the sphere of the realised strategy, by means of a learning process, by trial and error, behaving as entrepreneurs in the true sense of the word.

4.3.4 The loop of mental model learning

The process represented in loop 4 of Figure 6 highlights the impact of analysis of past strategic action on top management learning. Compared with the mechanism described in loop 2, the learning process shown in loop 4 goes to greater depth because it changes management's mental patterns, i.e. it goes to the roots of strategic intents. Therefore, loop 4 describes a mechanism which is very similar to the second-order learning of Argyris (1982).

For example, in IBM's restructuring process started by Gerstner at the end of 1993, the motor which was loop 4 made it possible to constantly refine the basic beliefs concerning IBM's field of activities, its role in the competitive arena(s) in which it operates, its way of being and operating and the meaning assumed by profit as a measure of success.

Concerning the field of activity in particular, Gerstner, following an in-depth, evolving analysis of IBM's distinctive competencies and of competitive arenas in which the company was operating, realised that the information technology sector was evolving profoundly and that value for the customer was created not so much in production as in assembling services.

Gerstner's strategic goal was therefore to make IBM a company providing integrated services and able to create value by producing new solutions to old problems and creating new competencies.

Subsequently, Gerstner captured the trend of the ICT industry towards the network-centric computing (NCC), i.e. the possibility of communicating and exchanging various types of digitalised information like video, high resolution images, voice and music by means of interconnected networks of computers. Following his analysis of environmental evolution, Gerstner further refined IBM's field of activity and, around 1995, the idea that IBM was essentially a service company became even clearer and developed into the strategy of making IBM not only an important company in the Information Technology Sector but also, and above all, the biggest service company in the network-centric computing. Gerstner realised that network-centric computing and the tool which is its prime motor, Internet, would lead to a profound revolution in world's culture and way of life and therefore, in the business strategies of client companies. So IBM's mission would become to accompany companies in this technological and cultural transition.

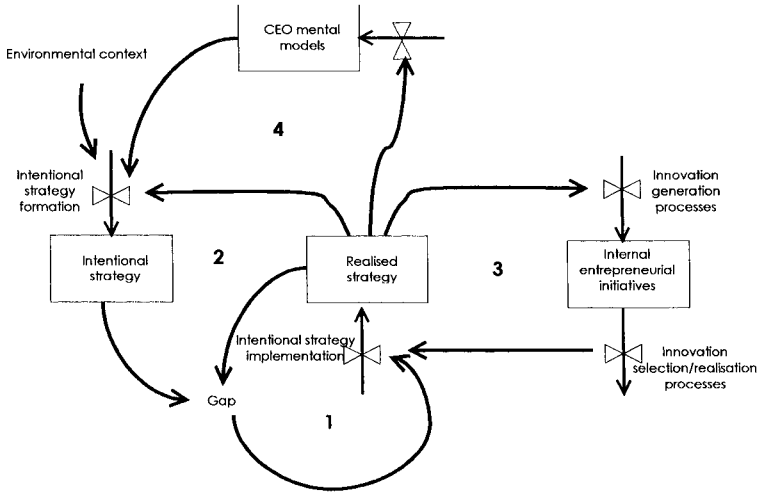


Figure 6. Mental Models Learning.

In conclusion, the model suggests that, in order to manage the strategy, both as a continuous process and as a 'one-off' discrete transformation process, we must acknowledge the existence of the four motors described.

The first motor highlights top management's ability to create, more or less efficaciously, managerial actions aimed at achieving the contents the intentional strategy.

The second motor refers to top management's ability to update, if required, the strategic intents, taking account of the structural changes within the environmental context and company situation. Also by this means, the gap is controlled, aiming to keep the level of motivation of collaborators high without causing stress.

The third motor makes it possible to achieve the potential for innovation built into the company's articulated human and organisational chain, to the extent that energy, know-how and creativity are released in the direction marked by a productivity and development growth strategy into new spaces for entrepreneurial initiative and responsibility. The possibility that this strategy can be shaped "bottom-up" increases the company system's adaptability, making it quicker in perceiving the changes under way in the environment and in framing suitable responses.

Lastly, the fourth motor describes top management's ability to open itself to questions and to learn, challenging its own mental patterns.

5. CO-ORDINATE THE MOTORS OF STRATEGY DYNAMICS

In the foregoing, we presented a qualitative model to investigate relationships between the morphology of a strategy-making process and a firm's adaptation behaviour. To generate the model, we used an inductive grounded approach. The model is the first step in the 'theory articulation' phase of an ongoing research project. In the following, we present a number of hypotheses which which we expect might be the base for further empirical research.

In order to successfully manage a company's strategy, it is necessary to learn how to orchestrate the simultaneous operation of the four motors, controlling two delicate areas.

The first area concerns the co-ordination of the strategic control motor (loop 1) on the one hand, and on the other, of the motors for forming strategic intents and learning of mental and BSO patterns (loops 2 and 4).

The working of the first motor work leads to reduction of the gap between realised and intentional strategy. Jack Welch of General Electric and Lou Gerstner of IBM are examples of leaders able to achieve strategic intentions effectively, shifting resources within the organisation, redesigning operational mechanisms and creating the necessary motivation for pursuing the new goals with determination.

On the other hand, high profile strategic management also requires the company's top management to be able to govern loops 2 and 4 in order to re-open the gap between realised strategy and intentional strategy, creating constructive tension which pushes towards challenging goals. Insisting in our metaphorical usage of complexity theory, we refer to McKelvey's idea of 'adaptive tension' as resulting from energy differentials within an organisation. Along similar lines, we expect that the existence of a difference between desired and reported state of key variables, creates a useful tension within organisations.

Therefore, the first area for attention in governing the dynamics of strategy can be defined as 'management of the gap'. The company management must be able to govern the gap: a persistent and significant gap between strategic intentions and their achievement can be the result of shallow analysis and excessively high ambitions which lead to the definition of unachievable goals or to inconsistent plans and executive actions and in all cases, results in the generation of non-productive stress and negative tension within the organisation.

On the other hand, the existence of a comfortable situation of well-being with no gap can be the indication of a dangerous state of equilibrium,

featuring no positive tension, and the company is drawn towards a state of inertia which is detrimental to its very survival.

The second area to which attention must be paid concerns loops 1, 2 and 4 on the one hand and loop 3 on the other. In fact the first group of motors is indirectly driven by the CEO who assumes a fundamental role in directing the movement. By means of loop 1, the CEO governs the achievement of strategic intents; via loop 2 he adjusts strategic intents and finally, by means of loop 4, he reviews and updates his mental patterns and, as a consequence the intentional strategy. Inversely, loop 3 is only indirectly governed by the company's top management.

The potential protagonists of motor 3 are distributed more or less throughout the organisation and are all those who are able to develop new ideas and initiatives, stimulated by the knowledge that develops in-field and by a strategic and organisational context which rewards co-operation and widespread initiative. The company's top management influences this loop only indirectly as an architect or design engineer who designs and shapes the strategic and organisational context, making it a 'behavioural environment' (Bartlett and Goshal, 1995) which is more or less favourable for the generation of operational and strategic innovations. Let's think, for example, of the impact of a leadership style consisting of 'wandering around' in production departments and offices, asking questions and looking into the innovations achieved and problems of improvement. In general, the organisational context can be shaped in order to leave more or less liberty in exploring new business areas outside the dominant strategy and the 'core' competencies of the company and in using resources for experiments and research even if the latter do not promise tangible results in the short term.

The greater the freedom of action assigned to loop 3, the greater stimulus will be given to liberalising creative energies and entrepreneurial behaviour but, on the other hand, the greater the disorder and dissipation of resources and energies might be in non-correlated directions and with no exploitation of synergies.

The systemic model drawn up, with identification of the motors at the base of the company's strategy dynamics, highlights a fundamental problem which characterises the study of the strategic behaviour of companies: the role and space for manoeuvre of top management in shaping company strategy in the face of the emergence of bottom-up self-organisation processes, of the inertia which characterises the variation of stocks of resources, be they tangible or intangible, or of consolidated mental patterns and the difficulty in anticipating the consequences of the decisions taken within the context of dynamic, complex company systems.

On the basis of the model developed, the two areas of attention considered now cross in a matrix (Figure 7).

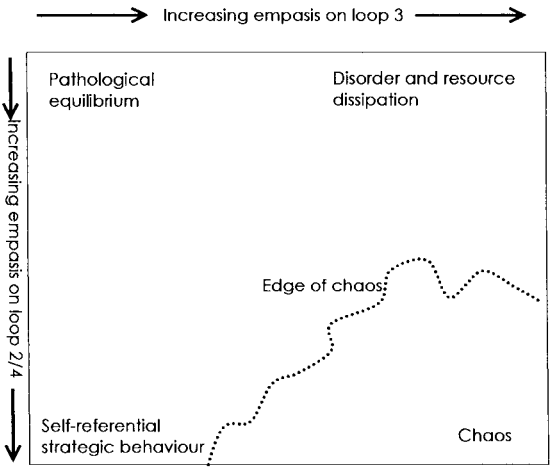


Figure 7. Map of Emerging Dynamics of Corporate Strategy.

The first dimension, on the vertical axis, concerns the use of motors 1, 2 and 4. The more intense the function of the learning loops of mental patterns and the formation of strategic intents is, the greater the tendency will be to abandon situations of equilibrium and to move towards the exploration of new territories, new business areas, new products, new technologies and new management systems, taking up new opportunities but also running the risk of neglecting or underestimating the existing situation and of conceiving strategic intents which cannot be achieved or are too distant from the company's basic competencies or of creating too much stress within the organisation.

On the other hand, a management which focuses exclusively on loop 1 risks paralysing the company in a situation which is comfortably balanced but potentially dangerous because it is prone to change into inertia and the inability to face the challenges laid down by the discontinuities in the environmental context (competitive, social, juridical-institutional, etc.) in which the company operates.

The second dimension, on the horizontal axis, concerns the strength, within an organisation, of the entrepreneurial loop. The greater is the use made of the motor in loop 3, by moving – in the diagram – from left to right towards areas of creativity, freedom of experimentation and disorder, the more the company' strategic behaviour will be left to be the result not only of the unique rationality of the company's top management but also of the local rationality of the other members of the organisation.

So it will be necessary to decide to what extent to liberate or bridle loop 3. The company abandons innovation and creativity potential if loop 3 is excessively restrained, or generates too much disorder and resource dissipation if loop 3 is allowed to unfold without instilling it with the necessary discipline.

So the model contributes to explain how, by steering the four strategy motors, to approach a zone which we call the threshold of chaos (Pascale, 1999; Pascale, Milleman and Gioja, 2000), or the edge of chaos (Brown and Eisenhardt, 1998), which represents a condition, an intermediate, permeable stage between order and disorder and which is the place in which innovation is produced.

As figure 7 illustrates, top managers face a very delicate endeavour: to maintain their firms within an area which we named edge of chaos by manouvering the described four startegy-making motors. Loosing control of one of the four motors entails placing the company in one of the four corners of matrix in figure 7, each having its pitfall and threats.

For example, considering the pathological state of equilibrium in the top left-hand corner, we can hypothesise that the more companies emphasise the processes for achieving strategy to the detriment of learning and entrepreneurship strategies and the more they adopt top-down control logics, the more they run the risk of entering states of entropy intended as states of equilibrium void of usable energy. Through the lenses of the theory of complexity, the notion of equilibrium takes on shades of negative connotations because it is a thermodynamic equilibrium, i.e. a state of stagnation reached in a closed system when entropy is at its highest and the ability to produce energy has fallen to its lowest level (Prigogine and Stenger, 1979); in this state, the system is inert and we are approaching an inescapable state of degradation (Monod, 1970:187). In this light, the challenge of steering the company towards the edge of chaos becomes the compulsory path for importing energy into the company, stimulating organisational learning[22] and entrepreneurial processes thereby avoiding the entropic decay which is typical of closed systems.

[22] The contributions produced in the chain of literature on organisational learning are a useful support for understanding the problems relating to the control and design of organisational learning processes. More precisely, certain contributions have created particularly important areas of research regarding the strategic-organisational change of a company. For example, as early as 1988, Nonaka dealt with the subject of the management of orders and chaos in organisational learning processes (Nonaka, 1988). Inversely, March tackled the problem of how to balance the exploitation of existing knowledge with the exploration of new terrain (1991). Nonaka and Takeushi (1995) analysed the link between the production of knowledge inside an organisation and the generation of innovation by the organisations itself, whereas Spender (1996) laid the foundations for a dynamic company theory based on knowledge.

The bottom left corner is an area where are organisations which may be very good in adapting to changes within the firm's core strategy and competencies. Yet, given their self-referential top-down approach, they may fail when facing competence-breaking environmental shifts.

The top right corner is an area where organisations are located that both are driven by a strong bottom-up, entrepreneurial motor of change and lack top management learning. These companies might face two orders of problems. On the one hand, emerging strategic initiatives face the opposition of top management, which is not capable to update its mental models. On the other hand, in presence of a weak top management, emerging strategic initiatives may be able to survive and growth by enacting loop 1, for example by exploiting unabsorbed slack resources or by creating political pressures at the level of middle management.The results are expected to be, in the first case, disruptive tension within organisations and, in the second case, unclear corporate mission and resource dissipation.

Finally, firms should avoid the lower right-hand corner of the matrix, where the level of disorder is unsustainable. Top down and bottom up strategic initiatives overlaps, corporate strategic behaviour emerges as unintended and unfolding results become ambiguous.

In short, innovation can be produced neither in a well-balanced and comfortable situation void of stimuli nor in chaotic environments in which there is strong tension pushing towards change and innovation but failing to channel itself in a constructive manner. If this is true, the task of the top management governing a company is to introduce into successful and consolidated entrepreneurial formula elements of disequilibrium.

If, on the contrary, the management finds itself managing a company in a situation of strategic disorientation, the top management's task is to introduce elements of equilibrium by enacting loop1. In order to avoid falling prey to the opposing dangers of deadly equilibrium and pathological chaos, management should be clearly aware (i) of where the obstacles lie which prevent it from governing the company along the paths of innovation; (ii) of what the inescapable elements of discipline of the behavioural context are, which are not only compatible with the need to innovate, but are also functional in terms of innovative processes.

6. CONCLUSIONS

From a theoretical point of view, the contribution of the article is twofold.

First, the article provides a platform for organising and interpreting literature on the strategic process of companies, by using the symbolic language of feedback loops and the logic of analysis crystallised in them. In this sense, the work presented proposes an example of how feedback loops can be used to represent and communicate theories of the strategic behaviour of companies. In fact, although literature on strategic management dedicates increasingly more space to approaches to research which are influenced by studies on dynamic systems and complexity and are therefore characterised by non-mechanistic, interpretative logics in which increasing attention is paid to relations of circular-type, rather than one-way, among variables, the circular nature of the causal model used frequently remains implicit, concealed in the web of narrative theories and not represented and communicated in an explicit and rigorous way (Farjoun, 2002). In this respect, we suggest that futher studies may go in the direction of formalising and further articulate the theory through computer simulation, thereby testing its internal coherency and honing hypotheses for empirical testing. Along these lines, we hope the article will provide a theoretical reference to guide longitudinal clinical studies finalised to interpret emerging strategic behaviour of large organisations.

Second, although the work presented does not contain specific operational indications on, for example, what the optimal equilibrium between the various loops is and how this equilibrium can be achieved, from a more applicative point of view it provides a tool for diagnosing the strategic behaviour of companies. Given that the representation logic used pivots on the concept of a complex dynamic system, typically characterised by its capacity for self-organisation in constantly new and unexpected states, we felt that it was interesting to investigate the dimensions for creating a conceptual space in which to analyse the dynamics of strategy, showing the tension, pressures, forces and processes in play.

The positioning of the matrix in Figure 7 forms the basis for dynamic analysis of pathological aspects or of the traps which characterise the trajectory of a company's strategic behaviour, receiving indications as to which loop has to be stimulated or slowed down. In addition, the article facilitates the metaphorical use of a number of concepts which management studies have borrowed from chaos theory and the theory of complexity.

REFERENCES

Andrews, K. R., 1971, *The Concept of Corporate Strategy*, Irwin, Homewood, IL.

Ansoff, H. I., 1965, *Corporate Strategy*, McGraw-Hill, New York.

Ansoff, H. I., 1979, *Strategic Management*, MacMillan Press, Ltd., London and Basingstoke.

Ansoff, H. I., 1984, *Implanting Strategic Management*, Prentice-Hall, Englewood Cliffs, NJ.

Ansoff, H. I., 1991, Critique of Henry Mintzberg's 'The Design School: reconsidering the basic premises of strategic management', *Strategic Management Journal* **12**(6):449-461.

Argyris, C., 1982, *Reasoning, Learning, and Action: Individual and Organizational*, Jossey-Bass Inc, San Francisco.

Argyris, C., and Schon, D. A., 1978, *Organizational Learning: A Theory of Action Perspective,* Addison-Wesley Publishing Co., Reading, MA.

Bartlett, C. A., and Goshal, S., 1995, Rebuilding behavioral context: turn process reengineering into people rejuvenation, *Sloan Management Review* (Fall).

Bower, J. L., 1970, *Managing the Resource Allocation Process: A Study of Corporate Planning and Investment*, Harvard University Press, Boston, Mass..

Brown, S. L., and Eisenhardt, K., 1998, *Competing on the Edge: Strategy as Structured Chaos*, Harvard Business School Press, Boston.

Burgelman, R. A., 1983a, A process model of internal corporate venturing in the diversified major firms, *Administrative Science Quarterly* **28**:223-244.

Burgelman, R. A., 1983b, Corporate entrepreneurship and strategic management: insights from a process study, *Management Science* **29**(12):1349-1364.

Burgelman, R. A., 1983c, A model of interaction of strategic behavior, corporate context, and the concept of strategy, *Academy of Management Review* **8**(1):61-70.

Burgelman, R. A., 1991, Intraorganizational ecology of strategy making and organizational adaptation: theory and field research, *Organization Science* **2**(3):239-262.

Burgelman, R. A., 1994, Fading memories: a process theory of strategic business exit in dynamic environments, *Administrative Science Quarterly* **39**:24-56.

Coda, V., 1989, *L'Orientamento Strategico di Fondo delle Imprese*, UTET.

Farjoun, M., 2002, Towards an organic perspective on strategy,Strategic, *Management Journal* **23**(7).

Forrester, J. W., 1961, *Industrial Dynamics*, Productivity Press, Cambridge, MA.

Forrester, J. W., 1968a, *Principles of Systems*, Productivity Press, Cambridge, MA.

Forrester, J. W., 1968b, Market growth as influenced by capital investment, *The Industrial Management Review* **9**(2):83-105.

Forrester, J. W., 1969, *Urban Dynamics*, Productivity Press, Cambridge, MA.

Forrester, J. W., 1971, Counterintuitive behaviour of social systems, *Technology Review* **73**(3):52-68.

Forrester, J. W., 1973, *World Dynamics*, Productivity Press, Cambridge, MA.

Forrester, J. W., 1992, Policies, decisions and information sources for modelling, *European Journal of European Research* **59**(1):42-63.

Glaser, B. G. and Strauss, A. L., 1967, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Aldine de Gruyter, New York.

Hamel, G., and Prahalad, C. K.,1989, Strategic intent, *Harvard Business Review* **67**(3):63-76.

Hannan, M. T., and Freeman, J., 1984, Structural inertia and organizational change, *American Sociological Review* **49**:149-64.

March, J. C., 1991, Exploration and exploitation in organizational learning, *Organization Science* **2**(1).

McKelvey, B., Dynamics of new science leadership: strategy, microevolution, distributed intelligence, complexity, in: *Mobilizing the Self-renewing Organization*, A. Y. Lewin and H. Volberda, eds., Sage, Thousand Oaks, CA.

Mintzberg, H., 1978, Patterns in strategy formation, *Management Science* **24**:935-948.

Mintzberg, H., 1979, *The Structuring of Organizations*, Prentice-Hall, Englewoods Cliffs, NJ.

Mintzberg, H.,1985, Of Strategies, Deliberate and Emergent, *Strategic Management Journal* **6**:934-948.

Mintzberg, H., 1967, Crafting Strategy, *Harvard Business Review* **65**(4):66-75.

Mintzberg, H., 1990a, The design school: reconsidering the basic premises of strategic management, *Strategic Management Journal* **11**(3):121-195.

Mintzberg, H., 1990b, Strategy formation: ten schools of thought, in: *Perspectives on Strategic Management*, J. Fredrickson, ed., Ballinger, New York.

Mintzberg, H., 1991, Learning 1, planning 0 reply to Igor Ansoff, *Strategic Management Journal* **12**(6):463-466.

Mintzberg, H., Ahlstrand, B., and Lampel, J., 1998, *Strategy Safari*, Prentice Hall Europe.

Monod, J., 1970, *Il Caso e la Necessità*, Arnoldo Mondatori, Milan, Italy.

Morecroft, J. D. W., 1984, Strategy support models, *Strategic Management Journal* **5**(3):215-229.

Nonaka, I., 1988, Creating organizational order out of chaos: Self-renewal in Japanese firms, *California Management Review* **30**.

Nonaka, I., and Takeuchi, H., 1995, *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press, New York.

Normann, R., 1977, *Management for Growth*, Wiley, New York.

Pascale, T. R., 1984, Perspective on strategy: the real story behind Honda's success, *California Management Review* **26**(3):47-72.

Pascale, T. R., 1999, Surfing the edge of chaos, *Sloan Management Review*, (Spring).

Pascale, T. R., Millemann, M., and Gioja, L., 2000, *Surfing the Edge of Chaos*, Crown Business, New York.

Prigogine, I., and Stengers, I., 1979, *La Nouvelle Alliance. Metamorphose de la Science*, Gallimard, Paris.

Quinn, J. B., 1980, *Strategic Change: Logical Incrementalism*, Prentice-Hall, Englewood Cliffs, NJ.

Quinn J. B., 1981, Formulating strategy one step at time, *Journal of Business Strategy* **1**(3):42-63.

Spender, J. C., 1996, Making knowledge the basis of a dynamic theory of the firm, *Strategic Management Journal* **17**:45-62, (Special Issue: Knowledge and the Firm).

Stacey, R. D., 2003, *Strategic Management and Organisational Dynamics*, 4th edition, Pearson Education Limited.

Thietart, R. A., and Forgues, B., 1995, Chaos theory and organisation, *Organisation Science* **6**(1):19-31.

Zappa. G., 1957, *Le Produzioni*, Volume 2°, UTET.

A COGNITIVE APPROACH TO ORGANIZATIONAL COMPLEXITY

Guido Fioretti[1] and Bauke Visser[2]
[1]*Università di Bologna, Dipartimento di Scienze dell'Informazione, Mura Anteo Zamboni 7, 40127 Bologna, Italy; E-mail: fioretti@cs.unibo.it.*
[2]*Department of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, E-mail: bvisser@few.eur.nl.*

Abstract: Organizational theory has construed complexity as an objective characteristic of either the structure or the behaviour of an organization. We argue that in order to further our understanding complexity should be understood in terms of human cognition of a structure or behaviour. This cognitive twist is illustrated by means of two theoretical approaches, whose relationship is discussed.

Key words: organizational complexity; human cognition; objective property; cognitive map; detail of information.

1. INTRODUCTION

Organization theory presents complexity as an objective property of the organization, much in the same way as, e.g., its degree of centralisation and formalisation. It is viewed as an objective characteristic of the structure of an organization, defined and measured in terms of the number of its constituent parts, their diversity and relationships (e.g. Lawrence and Lorsch, 1967, Thompson, 1967; Galbraith, 1973, 1977; Jablin, 1987; Daft, 1989). In the 1990s, complexity also becomes identified with intricate organizational behaviour, with small changes at the unit or employee level possibly giving rise to 'complex' aggregate patterns (e.g., Anderson, 1999; Lissack, 1999; Marion, 1999).

The aim of this paper is to argue that complexity should neither be defined nor measured in term of its source, be it an objectively given feature

of the structure or of the behaviour of an organization, but instead in terms of its effects on human cognition. Organization theorists have been careful in pointing to the decision context within which the concept of complexity plays a role. Briefly, complexity as numerosity, diversity, and inpredictability matters because of the increasing demands it imposes on decision makers concerned with attaining overall organizational effectiveness. But such demands are cognitive in nature. It therefore only seems natural to take the analysis one step further by detaching the notion of complexity from its objective source and instead attaching it to its consequence on the cognitive effort exerted by the decision maker to come to grips with her decision problem. That is, an organization is complex to the extent that a human being -e.g., an organizational designer or an outside observer-·has to exert a certain degree of cognitive effort in coming to grips with a decision problem.

Various reasons support this cognitive turn. First, this approach fits well with the general outlook of the theories in which the objective notions of complexity appear. As noted, complexity as a structural feature figures within well-defined decision theoretic approaches to organizations and it therefore seems natural to construe complexity cognitively. The same applies to behavioural complexity. Second, it helps unify existing organization theory. In particular, we illustrate how this approach may shed light on the discussion whether structural complexity stems from the number of horizontally organized units and vertically organized layers or from the connections between these parts. Moreover, it helps show commonalities between theories focusing on complex behaviour and theories focusing on complex structures. Third, it is in line with major accounts of complexity in general. Finally, a cognitive approach is in line with what one commonly understands to be 'complex'. `This is complex' is an utterance typical in situations when we do not understand something, as is clear from the fact that without any change in the observed phenomenon or the problem at hand, all at once we may consider it trivial or at least manageable.

After an overview of the main ingredients of contemporary thought on complexity in the realm of organization theory, the cognitive turn is presented in general terms. In section 4, the first operationalization is presented that starts with the cognitive map held by a human being who faces an organizational problem. Complexity is defined in terms of the level of dissatisfaction with the explanatory power of a cognitive map. The second operationalization, discussed in section 5, defines complexity in terms of information required to solve an organizational problem. An application to the problem of selecting projects by a firm illustrates its main features. The relationships between these two approaches are discussed in the concluding section.

2. THE TRADITIONAL APPROACH TO ORGANIZATIONAL COMPLEXITY

There are two strands in the organizational literature that use the words 'complex' and 'complexity' to characterize organizations. The first one refers to structural features of an organization. In particular, an organization is called complex if it is large and consists of several subsystems (e.g., R&D, manufacturing, sales, finance) that differ from each other "in terms of subsystem formal structures, the member's goal orientation, member's time orientations and member's interpersonal orientations" (Lawrence and Lorsch, 1967, p. 1). Thompson (1967, pp. 55-59) and Galbraith (1973, pp. 46-66) characterise a complex organization very much in the same vein. These authors stress the importance of the relationships among organizational units as being at the root of complexity.

With a slightly different emphasis, Jablin and Daft focus on the number of organizational parts in their definitions of complexity. Jablin (1987, pp. 400-404) uses complexity to depict "the structural components/units into which organizations and their employees may be categorized". He distinguishes vertical and horizontal complexity. Vertical complexity is the outcome of vertical differentiation and "is an indication of the number of different hierarchical levels in an organization relative to its size". Similarly, horizontal complexity measures "the number of department divisions in an organization". One such measure is the "number of different occupational specialties or specialized sub-units at a given hierarchical level". That is, complexity refers to the number of parts in an organization, with each part specialising in some activity. Daft (1989, p.18) uses very similar definitions.

In this line of researchy, complexity matters because the implied differentiation allegedly requires integration for the organization to perform well. Lawrence and Lorsch (1967, p.1) consider differentiation and integration to be antagonistic states, and study ways in which organizations assure integration. Thompson and Galbraith observe that differentiation and heterogeneity make coordination necessary, with the intensity of coordination being dependent on the type of interdependence. They come up with a classification of complexity on the basis of these differences in the intensity of coordination and the ensuing differences in information processing demands. Thompson (1967, pp. 55-59) distinguishes pooled, sequential, and reciprocal interdependencies between subsystems as the basis for his classification of degrees of complexity. Galbraith (1973) distinguishes lateral relationships of varying intensity. Types of coordination differ in communication and decision load, and, Thompson (1967, p. 56) adds, "[t]here are very real costs involved in coordination.".

Since the end of the 1980s, a second strand in the literature on organizations uses the notion of complexity. Here, it is related to the behaviour originating from the interactions of the many parts of a `complex' system. Building upon early studies on self-organization (Nicolis and Prigogine, 1977; Prigogine and Prigogine, 1989; Haken, 1983, 1987), several models investigated the formation of structures between a large number of interacting particles and the ensuing properties of the behaviour of the system as a whole (Fontana, 1991; Kauffman, 1993). By analogy, one started to suspect that similar phenomena were widespread in natural and social systems alike. Complexity came to be identified with a special kind of behaviour, i.e., with intricate aggregate patterns emerging from the interaction of the constituent parts of an organization that themselves followed relatively simple behavioural rules. It is evident that these insights and their related methodologies are relevant to organizational problems (Anderson, 1999; Frank and Fahrbach, 1999; Lissack, 1999a, 1999b; Marion, 1999; Borel and Ramanujam, 1999).

According to this strand of literature, complexity matters to organization theory because it makes organizational behaviour subject to surprises and hard to predict (Anderson, 1999, pp. 216-217), rendering the attainment of organizational effectiveness nonobvious. Consequently, decision-makers should become aware of the limits of their knowledge and engange in a learning process with the complex system they are facing (Allen, 2000, 2001; Cilliers, 2002; Allen and Strathern, 2003). Indeed, complexity is not seen as a set of rules to solve a particular set of problems, but rather a perspective that may provide a new understanding to problems (Lissack and Letiche, 2002).

3. A COGNITIVE TURN

The shift from a view of complexity based on the number of component parts to the intricacy of microbehaviors to the current emphasis on the methodology of complexity suggests us the possibility of a cognitive turn in the interpretation of this concept. From understanding organizational complexity in terms of the structure or the behaviour of an organization to its effects on human cognition.

We want to argue that the very reason that makes complexity important to organizational theory also points to a cognitive conception of complexity. There is wide agreement that complexity matters because of the resulting difficulties when it comes to questions of organizational design and decision making. The multiplicity of subsystems, their diversity, the linkages among them, and the erratic aggregate behaviour that results make designing

'effective' organizations and taking decisions involving organizations hard. That is, complexity matters only because of cognitive problems it give rise to. It is therefore only natural to define and measure complexity in cognitive terms.

In other words, this cognitive turn implies that complexity should not be seen as an objective feature of some organizational characteristic, but rather as relative to a decision problem or to the representation a decision maker has of this problem. Apart from being an arguably natural part of the organizational theories in which it plays a role, a cognitive conception of complexity has got three other advantages.

First, it sheds light on various questions that come up when studying the extant literature on organizational complexity, two of which will be presented here. Recall that in the literature on complexity as a feature of organizational structure, there was little agreement on the measurement of complexity. Some argued that complexity is captured by the number of horizontally arranged units and vertical layers, while others insisted on the type of connections between these units and layers. We illustrate in section 5 how explicitly taking into account the cognitive requirements an organizational designer faces may resolve the dilemma of whether numbers or connections matter. A further ambiguity concerns the relationship between complexity as an aspect of organizational structure on the one hand, and complexity as organizational behaviour on the other hand. Little has been said about their relationship, though in both cases reference is made to numerous parts that are somehow connected. What unites both approaches is the implicit assumption that numerosity and connectedness make understanding more complicated. Thus, moving to the level of cognition allows to unify these two strands of literature.

Second, support for a cognitive view of complexity in the realm of organizational studies is provided by several conceptions of complexity in general, to begin with Dupuy (1982) and Rosen (1985) but also including Crutchfield and Young (1989) and Gell-Mann (1994), who moved from the idea of computational complexity (Solomonoff, 1964; Kolmogorov, 1965; Chaitin, 1966). These insights impact on a discipline where it is widely recognized that strategy-making is tightly linked to both cognitive and organizational problems (Anderson and Paine, 1975). So among organization theorists, Simon (1999; p. 215) directs our attention to complexity as a characteristic of a description instead of complexity as an objectively given characteristic: "[h]ow complex or simple a structure is depends critically upon the way in which we describe it". Rescher (1998, p. 1) is even more explicit about the role of cognition in the definition of complexity. Writing within the context of complexity in general, he states "[o]ur best practical index of an item's complexity is the effort that has to be

expended in coming to cognitive terms with it in matters of description and explanation".

Third, explicit recognition of the cognitive nature of complexity fits very well with the use of the word "complex" in common parlance. Suppose one were to find a more cogent, less complicated account of some (organizational) phenomenon. Then, in common parlance, one would say that the phenomenon itself has become less complex. Even if a phenomenon is not changing one may consider it "complex" at a certain point in time and "simple" at a later time, if in the meantime sufficient information has been received and a proper reformulation of the problem allowed to come to grips with it.

In the next two sections, we provide two ways of exploring organizational complexity in a cognitive way. We illustrate how one could operationalise a cognitive approach to complexity in the realm of organization theory. The two models that we expound tackle the same problem from different sides and thus provide complementary points of view.

4. WHAT A DECISION-MAKER DOES NOT KNOW...

The first approach to operationalize a cognitive view of complexity starts from observing and modelling the way decision-makers represent problems in their minds. If this representation has been able to suggest the correct behavior, a decision-maker will not say that he is facing a complex reality. If, however, this representation suggested a behavior that induced an outcome very different from the intended one, then a decision maker may not be confident to have framed the decision-problem in the best possible way. To the extent that he doubts his own representation of a decision problem, he will say that this is a complex one.

Let us look more closely at the process by which mental representations of decision problems arise. Options, objectives and strategies are not self-evident. Rather, they result out of cognitive processes of information categorization and construction of causal relations between these categories, processes whose ultimate outcome is a cognitive map (Hebb, 1949; Haye, 1952).

A cognitive map is a network of causal relationships between options and objectives that one can safely trust, if not always, at least most of the times. The cognitive map of a company entails the options that it envisions, the objectives that it wants to pursue, and a network of causal links from options to objectives along paths that represent available strategies.

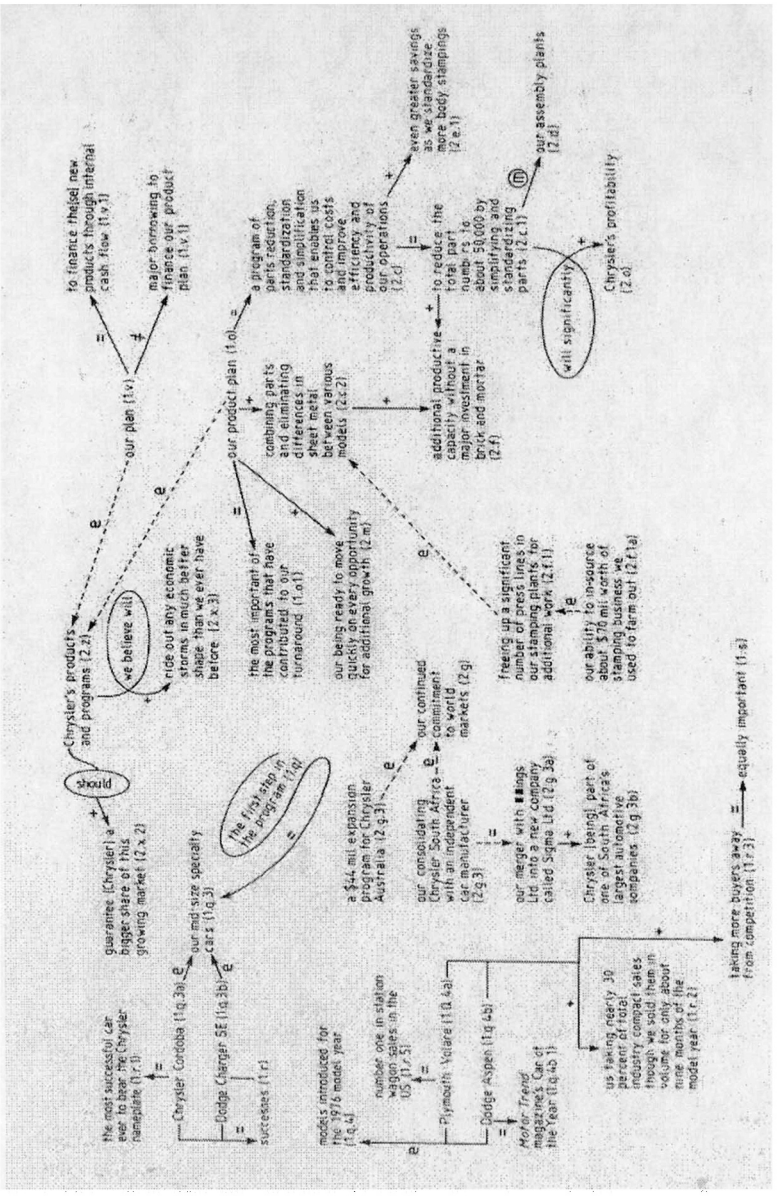


Figure 1. Chrystler cognitive map, 1976. Positive causal links are denoted by a '+', negative causal links are denoted by a '-', equivalences are denote by a '=' and examples are denoted by an 'e'. Dashed lines denote inferred relationships (reproduced with permission).

Corporate cognitive maps can be reconstructed by means of a linguistic analysis of letters to shareholders and other corporate documents (Sigismund-Huff, 1990). For instance, figure 1 illustrates a portion of

Chrysler's cognitive map evinced out of speeches to securities analysts in 1976 (Sigismund-Huff and Schwenk, 1990).

Observing figure 1 it is evident that, in 1976, standardization and reduction of parts were high on the agenda. For instance, "to reduce the total number of part numbers to about 50,000" is deemed to generate "additional productive capacity without a major investment in brick and mortar". Here we can see a causal link from the option of reducing parts to the objective of increasing productive capacity, which, taken together with all other causal links, illustrates Chrysler's strategy. Although figure 1 also includes a number of equivalence relations and supportive examples, these are rhetorical devices that have been included in the speech with the purpose of stressing causal links. Ultimately, causal links constitute the backbone of a cognitive map.

Thus, the structure of a cognitive map can be seen as a set of links between options and objectives as depicted in figure 2. Note that in order to draw figure 2 from a part of figure 1, options and objectives had to be compounded using the equivalence and example relations.

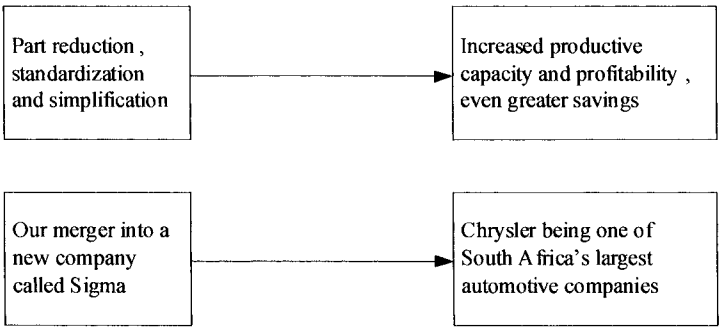


Figure 2. A portion of the causal links depicted in figure1, after compounding some of the items.

The most important fact about cognitive maps is that causal relations between options and objectives are not conceived independently of options and objectives themselves. For instance, the causal link between "reducing the number of parts" and "increasing profitability" was not conceived independently of the idea of reducing the number of parts and increasing profitability to the levels attained by Japanese competitors. In fact, managers define options by lumping a number of detailed actions at the shop-floor level into broad categories (e.g. "reducing the number of parts"). Similarly, they define objectives by lumping a number of accounting variables together (e.g. into "increasing profitability"). Options and objectives are mental

categories for a number of actions to undertake and a number of indicators to observe, which are defined having in mind a possible causal link between them. Thus, the categories employed by a decision-maker cannot be investigated in isolation from one another, but rather within the network in which they are embedded.

Since causal links are tailored to specific options and objectives, emergence of novelties is detected when the chosen option did not reach the expected objective. For instance, Chrysler's experience after restructuring was that, although its reorganization program brought it close to the standards of its Japanese competitors, lack of standardized communication procedures kept it short from reaching their levels of performance (Sobek, Liker and Ward, 1998). In other words, novelties call attention upon themselves because causal links appear, that are different from the expected ones. This observation is crucial for the concept of complexity presented in this section.

In fact, let us define complexity as the extent to which empirical experience runs contrary to the expectations embedded in a cognitive map. It is when novelties emerge that the causal relations that one expects may not hold. On these occasions, a decision-maker is likely to say that he is facing a "complex" environment.

Given this definition, complexity can be measured by looking at the structure of causal links in a cognitive map. At a first glance, it may seem straightforward to measure complexity by means of an index of the extent to which empirical experiences make a cognitive map intertwined: the more distant from a simple network of one-to-one correspondences, the higher is complexity. However, a second factor should be considered, namely, that a cognitive map where highly intertwined blocks are separated by sparse links should yield a lower complexity than one where not even blocks are distinguishable. For instance, figure 3 illustrates a situation where complexity should be zero (left), maximum (center) and intermediate (right). In this last case, complexity is lower because interwined causal links are arranged in blocks. Although figure 1 does not depict a very intertwined cognitive map, we can distinguish blocks of linkages: the largest one refers to parts reduction and standardization, a second one focuses on the merger with a South African car manufacturer, and two other blocks are concerned with compact and mid-size cars, respectively. A detailed mathematical account of the proposed measure can be found in Fioretti (1999).

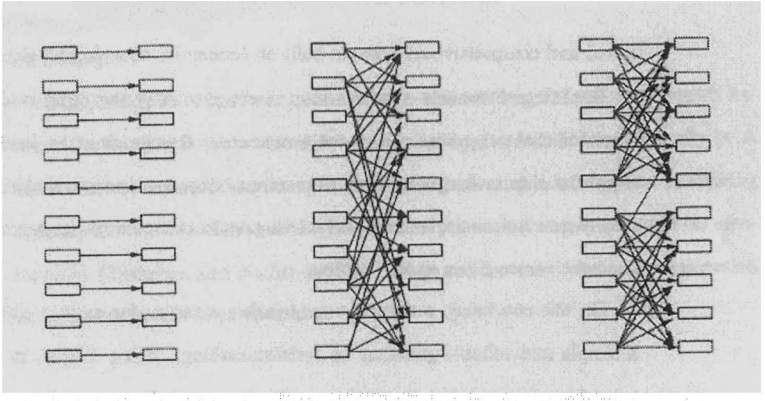


Figure 3. Three cognitive maps illustrating causal links between options and objectives. From left to right, a cognitive map that works perfectly envisions a simple world, a cognitive map that is not able to provide any orientation envisions a very complex world, and a cognitive map where causal links have a structure envisions a world of intermediate complexity.

This approach to complexity is useful in order to explain sudden jumps in decision-making, from an established course of action to realizing not to have a direction to follow when novel events destroy our certitudes, and back to a novel vision of what are the right things to do. Changing one's vision corresponds to changing one's cognitive map, either as a consequence of loosing faith in an established one or because of restored confidence in a new one (Sigismund-Huff and Huff, 2000).

Organizations may take a long time in order to realize that their cognitive frame should be changed, and they may take an even longer time in order to change it. Evidence suggests that changing a cognitive map is not triggered automatically by a single instance of an unexpected reaction of competitors or by other changes in the environment. Organizations go through phases of shock and defensive retreat; continue doing the same, but more; want to "weather the storm" before acknowledging changes and adapting to the new reality by "unlearning yesterday" and "inventing tomorrow" (Fink et al., 1971; Hedberg et al., 1976; see also Ford, 1985; Ford and Baucus, 1987; Mone et al., 1998). For instance, the first reaction of the US automotive industry to the entrance of Japanese rivals in the US market was to blame its difficulties "on the government, on unfair trade, on brickheaded workers, on snooty American consumers, and (...) on the "congenital sickos" in the media" (Ingrassia and White, 1994, p. 456; see also Womack et al., 1990). Only after years of irregular and dwindling profits it was realised that Detroit's view of a reasonably predictable world and competitive advantages built on economies of scale had given way.

Realizing that one's cognitive map is not providing the right guidelines is the stimulus that triggers a search for a new one. Recognizing the inadequacy of a cognitive map to deal with novel phenomena, stopping decision-making in order to formulate a new cognitive map and starting again as soon as it has become available, is a continuous, never ending process.

On the one hand, a cognitive map makes sense as far as it is able to provide a simple and reliable guidance to decision-making. Being simple, in the sense of providing clear-cut directions of causality, is part of the very nature of a cognitive map (Axelrod, 1976). On the other hand, simplicity and coarseness make a cognitive map obsolete when novelties appear. When the causal links that a cognitive map proposes in order to interpret reality are at odds with real experiences, then a decision-maker has a complicated, intertwined map, one that says that for as a consequence of any option anything may happen. Such a cognitive map is useless. When managers do not know what to do, when they prefer waiting to acting, then they would say that the situation is complex. It is wiser to wait and see, postponing decision-making until a new, reliable cognitive map has become available. It is the source of liquidity preference in the face of too uncertain investment prospects (Keynes, 1936).

A cognitive map that is able to provide a sensible orientation in decision-making is a simple one, one that connects options to objectives with a few, clear-cut causal links. Novelties may emerge, that eventually generate additional causal relationships. This means that it becomes very difficult for managers to make a decision, since they foresee many different, even opposite outcomes for each single option. Decisions are likely not to be made until a new cognitive map has been developed. Clearly, we cannot predict the new map that will be conceived, but we can measure the extent to which the present one is far from being simple, and call this magnitude complexity.

Complexity, as it has been expounded in this section, denotes a mismatch between the world as it is envisioned by decision-makers and reality as it actually is. A basic tenet of organization theory is that an organization's information processing capacity should be tailored to the information processing requirements of its environment (Tushman and Nadler, 1978). This is not the case when an organization views its environment as complex.

5. ... AND WHAT A DECISION-MAKER SHOULD KNOW

In the previous section complexity is defined in terms of key managers lacking confidence in the cognitive map that they had been using hitherto. Sooner or later action should be taken in order to solve this problem. As we shall see, this observation suggests a second and complementary view of complexity.

Through collection of new data, brain-storming and discussion, managers sooner or later will be able to construct a cognitive map that provides an explanation for confusing facts. It is not our purpose to describe the creative aspects of the process by which a novel cognitive map is formulated. Rather, we pick the point in time when a novel cognitive map has just been conceived and we focus on its implementation. Once a new cognitive map is there, managers face well-defined organizational problems in search for relevant information.

Supposing that a just-forged cognitive map must be applied to concrete situations, one can think of organizational complexity as of the amount of information that is necessary in order to solve a given class of decision problems. One can think of a collection of tasks, employees, structural features that can be arranged in different ways in order to perform adequately in the situations that are envisaged by the accepted cognitive map. For the purpose of this paper, 'organizational structure' refers to the assignment of tasks and authority to employees; to the grouping of these employees in organizational units and other work relationships; and to the connections between these units and their overall arrangement. The goal is to attain a good fit between the elements that make up an organizational structure and its environment.

An organization will need information about certain characteristics of its tasks, its employees and its own structural features in order to fit its employees with the tasks they face in particular organizational positions. However, the required information on the specific abilities of single employees may well depend on the many possible ways of arranging these employees. Moreover, information produced by individual employees who are dispersed throughout the organization will have to be aggregated in one of many possible ways in order to be useful to top managers. Thus, the amount of information that is required, and the number of employees that should be consulted in order to solve a decision problem, is likely to be affected by organizational structure. Two examples illustrate these points.

Example 1: In a production process where various tasks must be performed sequentially, inventories reduce the amount of information on the

timing of individual tasks that is required to run the overall process smoothly. On the contrary, Just-in-Time production systems require transmission of detailed information between production units (kanban), but also that managers have detailed knowledge of the features of each production unit. In fact, to the extent that production units cannot be made as flexible as to perform any task and process lots of any size, the set of possible sequencing paths has to be planned by management in much greater detail than in the case inventories are there to buffer mistakes and suboptimal arrangements. Thus, the particular decision problem of managing production can be solved by organizational arrangements that choose different combinations of the amount of inventories and the amount of information needed to eliminate inventories. This is akin to Galbraith's (1973, pp. 14-19) account of the effects of, on the one hand, the introduction of slack resources on the need for information processing and, on the other hand, the creation of lateral relations on the capacity to process information.

Example 2: According to Alfred Chandler (1962), the main reason the functional form gave way to the multi-divisional form was that the latter structure solved two problems the former created: information overload at the top management level, and lack of information on product line profitability. Relevant information on product profitability could actually be produced by functional structures, but only at a very high cost since it had to be pulled out of many functions. On the contrary, within a multi-divisional structure such information is readily generated as part of the financial information on which divisional managers base their decisions. Ultimately, the multi-divisional structure allowed to solve decision problems related to product lines by means of less information, because it only produced the needed one.

In general, since employees differ from one another with respect to the ability by which they perform a specific task, their assignment to positions in an organizational structure is likely to determine its overall performance. However, correct assignment of employees requires information on their specific abilities, generally in varying degrees of detail depending on organizational structure. Therefore, the amount of information on the abilities of employees that is required by an organizational designer in order to solve his assignment problem induces an ordering of organizational structures. This amount of information can be used as a measure of organizational complexity.

To show how one could operationalise such an approach, consider the problem of a firm that contemplates the introduction of a new product. As it is unclear whether this product will be good (g) and give rise to a profit, X, or bad (b) and lead to a loss, - Y, different departments of the firm run a number of tests, t. There is an a priori probability of α of a product being

good. All tests are imperfect in that bad products may pass a test favourably (A or Accept), while good products may receive a negative verdict (R or Reject). That is, a test t can be described by the pair (p_t^b, p_t^g), the probabilities with which it accepts bad and good projects. Let us call a test t better than t' if t correctly rejects more bad projects than t', and good products also pass more often favourably t than t'. To reduce errors, test results obtained by the departments are combined. The goal of the firm is to maximize profits on products offered. Let us assume that the firm can choose from the sequential structures depicted in Figure 4.

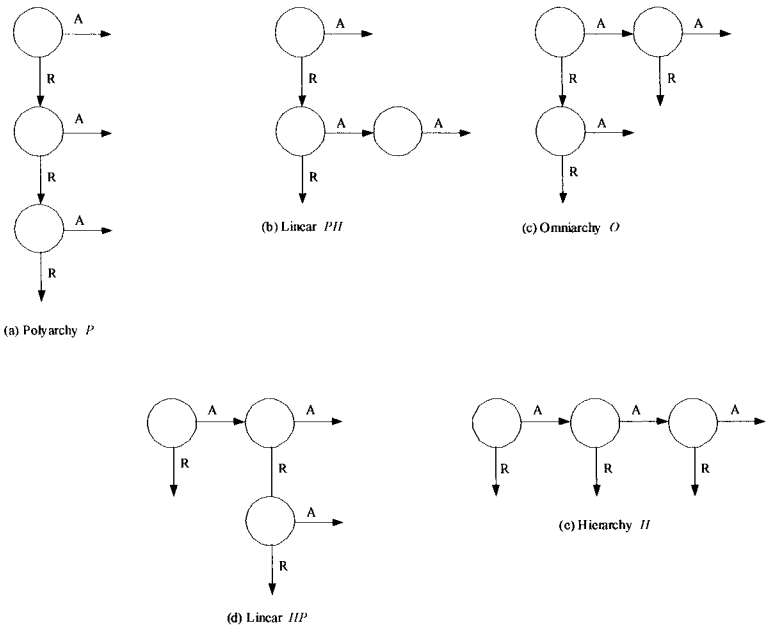


Figure 4. Five structures.

The firm then faces the twin problems of (a) determining for every structure which department should first run a test, which next, and (b) which of the structures to use. As the purpose of this example is to illustrate the use of a cognitive notion of complexity we focus on the first question - the second question, and the interplay of performance with complexity and robustness, is addressed in Visser (2002).

The complexity of this decision problem is measured by the level of detail of information about the individual tests that is necessary and sufficient to determine the optimal ordering. It can be shown that it is not so much the number of tests, but rather the way in which tests are run

consecutively by the departments that determines the kind of information that the firms needs to possess about the qualities of the test in order to position them correctly.

If departments are organized in a simple sequence like the ones depicted in Figures 4 (a) and (e), where a department runs a test only if preceding departments have either all accepted or all rejected the product, no information about the qualities of individual tests is required to attain the best performance. This is easy to see for the sequence of Figure 4 (a), as the probability of final implementation is simply the product of acceptance of individual tests, $p_{t1}\ p_{t2}\ p_{t3}$. Changing the order of departments leaves the overall probability of acceptance unaffected as multiplication is a commutative operator. The same applies to Figure 4 (e), where the probability of acceptance equals $1-(1-p_{t1})(1-p_{t2})(1-p_{t3})$.

If, however, departments are arranged like in Figures 4 (b) and (d), with alternating connections between departments, but where any test can still be final, profit maximization requires the firm to be able to order tests in terms of their characteristics as the best test should be used first. That is, the firm needs ordinal information about the quality of the tests run by its departments. If tests cannot be ordered using the "*better*" criterion because, say, a test t run by a department has a higher probability of accepting both good and bad projects than some other test t', the firm needs cardinal information, i.e., information about the numerical values of the characteristics of the tests.

Finally, in structures like the one depicted in Figure 4 - where at least one test t is always followed by some other test t', irrespective of the outcome of test t - cardinal information is required. That is, the firm needs not only to be able to order the tests in terms of their characteristics, but also to know the precise probabilities of acceptance. Clearly, by moving from the structures depicted in Figures 4 (a) and (e) through those in Figures 4 (b) and (d) to the one in Figure 4 (c), the firm's problem becomes more complex since cardinal information is more detailed than ordinal information.

Once an ordinal/cardinal distinction has been made, the amount of information required to solve an organizational problem can be measured in terms of number of items to be measured. Within structures that require, say, ordinal information, one may usefully distinguish between numbers of tests to be ordered. Recall that in Figures 4 (b) and (d) one only needed to be able to identify the best test. That is, one had to be able to distinguish two specific groups of tests: the best test, and the others. In larger structures that have different connections between successive desks, like the ones depicted in Figures 4 (b) and (d), the number of groups of tests that one should be able to identify can easily grow. Although ordinal information is still sufficient, the increase in the number of groups that one needs to identify does imply a

correspondingly harder task for the firm. In this sense, the size of the organization affects the difficulty of the organizational design problem. In other words, one could think of the *type* of connections to induce a qualitative classification of required information, and the *size* of the organization as inducing a quantitative refinement of that classification.

Among those who approached organizational complexity as a characteristic of the structure of an organization there is wide agreement that this characteristic is in fact objectively given and that it captures differences among organizational units, but there is little agreement regarding the details. This comes clearly to the fore when discussing degrees of complexity and operational measures of it. Both Thompson and Galbraith emphasize relationships between units, not their mere number, as a measure of complexity. For instance, Thompson (1967; p. 74) claims that "size alone does not result in complexity". Also for Scott (1998) organizational complexity has to be identified with the type of relationships among organizational parts. The contrast with, e.g., Jablin (1987) and Daft (1989) could hardly be starker: they express horizontal and vertical complexity in terms of numbers of units and layers, respectively.

We claim that this issue can be addressed within the framework developed above. As in Thompson and Galbraith, it is the type of relationships between successive units that determines complexity when this is measured in terms of the organizational designer needing ordinal or cardinal information in order to arrange the parts. As in Jablin, it is the size of an organization that determines complexity when this is measured in terms of its designer needing information on a number of tests in order to arrange them. If we stipulate that any amount of cardinal complexity is larger than any amount of ordinal complexity, these two measures do not contradict one another.

Clearly, we presented a very stylized model. However, one could conceivably expand the above scheme to tests or agents that classify projects on the basis of higher dimension categories, providing richer judgements than a simple accept/reject dichotomy. In the case of decision makers, the combination of individual mental categories would yield a cognitive map of the kind illustrated in figure1, one which would moreover explicitly depend on organizational structure.

6. CONCLUDING REMARKS

The two approaches presented above are distinct yet not opposite to one another. In fact, in section 4 we defined complexity in terms of inadequacy of what is currently known in order to solve an ill-defined problem.

Subsequently, in section 5 we defined complexity in terms of what should be known in order to solve a well-defined problem. In between, the task of transforming ill-defined problems into well-defined problems by means of the new interpretation provided by a novel cognitive map, was left unspecified. However, in spite of a missing link the above approaches are complementary in the sense that the first one aims at assessing improper framing of decision problems, whereas the second one attempts to provide operative solutions once problems have been reframed.

These perspectives are not independent of one another, because problem framing depends on organizational structure. Consider that the cognitive map of an organization is the result of organizational interpretation and information processing with a view to building causal relationships needed to guide decision-making. Both interpretation and information processing are intimately tight up with organizational structure. It determines how information is aggregated, coded and classified; it influences the options considered and the criteria used in such considerations by decision makers throughout the organization; and it regulates how and which conflicts over interpretation, decision, and implementation are referred to higher levels for resolution, thereby affecting in turn what is being learned by whom (Hammond, 1994). This is likely to affect to a considerable degree the organizational view on causation, on the relationships between the options open to the organization on the one hand and the envisioned outcomes on the other hand, in short on the cognitive map.

This also suggests that the relationship between the two concepts of complexity is perhaps not as simple as depicted above. Hammond's analysis directs attention to the convoluted nature of this relationship, with the organizational structure influencing the cognitive map, and the cognitive map influencing the search for an adequate structure in its turn.

REFERENCES

Allen, P. M., 2000, Knowledge, ignorance, and learning, *Emergence* **2**(4):78-103.

Allen, P. M., 2001, What Is complexity science? Knowledge of the limits to knowledge, *Emergence* **3**(1):24-42.

Allen, P. M., and Strathern, M., 2003, Evolution, emergence, and learning in complex systems, *Emergence* **5**(4):8-33.

Anderson, P., 1999, Complexity theory and oganization science, *Organization Science* **10**:216-232.

Anderson, P., and Paine, F. T., 1975, Managerial perceptions and strategic behavior, *Academy of Management Journal* **18**:811-823.

Atkin, R., 1974, *Mathematical Structures in Human Affairs*, Crane, Russak, New York.

Atkin, R., 1981, *Multidimensional Man*, Penguin Books, Harmondsworth.

Axelrod, R., 1976, *Structure of Decision. The Cognitive Maps of Political Elites*, Princeton University Press, Princeton.

Chaitin, G. J., 1966, On the length of programs for computing finite binary sequences, *Journal of the Association for Computing Machinery* **13**:547-69.

Chandler, A., 1962, *Strategy and Structure. Chapters in the History of the Industrial Enterprise*, MIT Press, Cambridge, MA.

Cilliers, P., 2002, Why we cannot know complex things completely, *Emergence* **4**(1/2):77-84.

Clark, A., 1993, *Associative Engines: Connectionism, Concepts, and Representational Change*, The MIT Press, Cambridge, MA.

Crutchfield, J. P., and Young, K., 1989, Inferring statistical complexity, *Physical Review Letters* **63**:105-108.

Daft, R. L., 1989, *Organization Theory and Design*, West Publishing Company, St. Paul.

Dupuy, J. P., 1982, *Ordres et Désordres*, Editions du Seuil, Paris.

Fink, S., Beak, J., and Taddeo, K., 1971, Organizational crisis and change, *Journal of Applied Behavioral Science* **7**:15-27.

Fioretti, G., 1998, A concept of complexity for the social sciences, *Révue Internationale de Systémique* **12**:285-312.

Fioretti, G., 1999, A subjective measure of complexity, *Advances in Complex Systems*, **4**:349-370.

Fontana, W., 1991, Algorithmic chemistry, in: *Artificial Life II*, C. G. Langton, ed., Addison-Wesley, Redwood City, pp. 159-209.

Ford, J. D., 1985, The effects of causal attributions on decision makers' responses to performance downturns, *Academy of Management Review* **10**:770-786.

Ford, J. D., and Baucus, D. A., 1987, Organizational adaptation to performance downturns: an interpretation-based perspective, *Academy of Management Review* **12**:366-380.

Frank, K. A., and Fahrbach, K., 1999, Organization culture as a complex system: balance and information in models of influence and selection, *Organization Science* **10**(3):253-277.

Galbraith, J., 1973, *Designing Complex Organizations*, Addison-Wesley, Reading, MA.

Gell-Mann, M., 1994, *The Quark and the Jaguar: Adventures in the Simple and the Complex*, Freeman, New York.

Haken, H., 1983, *Synergetics: An Introduction*, Springer Verlag, Berlin.

Haken, H., 1987, *Advanced Synergetics*, Springer Verlag, Berlin.

Hammond, T.H., 1994, Structure, strategy, and the agenda of the firm, in *Fundamental Issues in Strategy*, R. P. Rumelt, D. E. Schendel and D. J. Teece, eds., Harvard Business School Press, Boston, MA, pp. 97-154.

von Hayek, F., 1952, *The Sensory Order*, Routledge & Kegan Paul, London.

Hebb, D.O., 1949, *The Organization of Behavior*, John Wiley & Sons, New York.

Hedberg, B., Nystrom, P., and Starbuck, W., 1976, Camping on seesaws: prescriptions for a self-designing organization, *Administrative Science Quarterly* **21**:41-65.

Ingrassia, P., and White, J. B., 1994, *Comeback. The Fall and Rise of the American Automobile Industry*, Simon and Schuster, New York.

Jablin, F. M., 1987, Formal organization structure, in: *Handbook of Organizational Communication. An Interdisciplinary Perspective*, F. M. Jablin et al. ,eds., Sage Publications, Newbury Park, pp. 389-419.

Kauffman, S., 1993, *The Origins of Order*, Oxford University Press, Oxford.

Keynes, J. M., 1936, *The General Theory of Employment, Interest and Money*, MacMillan, London.

Kolmogorov, A. N., 1965, Three approaches to the quantitative definition of information, *Problems of Information Transmission* **1**:1-7, (Reprinted in 1968, *International Journal of Computer Mathematics* **2**:157-68).

Lawrence, P. R., and Lorsch, J. W., 1967, Differentiation and integration in complex organizations, *Administrative Science Quarterly* **12**:1-30.

Lissack, M. R., 1999a, Complexity and management: it is more than jargon, in: *Managing Complexity in Organizations. A View in Many Directions*, M. R. Lissack and H. P Gunz, eds., Quorum Books, Westport, Conn., pp. 11-28.

Lissack, M. R., 1999b, Complexity: the science, its vocabulary, and its relation to organizations, *Emergence* **1**(1):110-126.

Lissack, M. R., and Letiche, H., 2002, Complexity, emergence, resilience, and coherence: gaining perspective on organizations and their study, *Emergence* **4**(3):72-94.

Marion, R., 1999, *The Edge of Organization: Chaos and Complexity Theories of Formal Social Systems*, SAGE Publications, Thousand Oaks.

Mone, M. A., McKinley, W., and Barker III, V. L., 1998, Organizational decline and innovation: a contingency framework, *Academy of Management Review* **23**:115-132.

Morel, B., and Ramanujam, R., 1999, Through the looking glass of complexity: the dynamics of organizations as adaptive and evolving systems, *Organization Science* **10**(3):278-293.

Nicolis, G., and Prigogine, I., 1977, *Self-Organizatiion in Non-Equilibrium Systems: From Dissipative Structures to Order Through Fluctuations*, John Wiley & Sons, New York.

Prigogine, I., and Prigogine, G., 1989, *Exploring Complexity: An Introduction*, Freeman, N.Y.

Rescher, N., 1998, *Complexity. A Philosophical Overview*, Transaction, New Brunswick, NJ.

Rosen, R., 1985, *Anticipatory Systems*, Pergamon Press, Oxford.

Sah, R. K., and Stiglitz, J. E., 1986, The architecture of economic systems: hierarchies and polyarchies, *American Economic Review* **76**:716-727.

Scott, W. R., 1985, *Organizations. Rational, Natural and Open Systems*, Prentice Hall, Upper Saddle River, NJ.

Sigismund-Huff, A., 1990, Mapping strategic thought, in: *Mapping Strategic Thought*, A. Sigismund-Huff, ed., John Wiley & Sons, Chirchester, pp. 11-49.

Sigismund-Huff, A., and Huff, J. O., (with P. Barr), 2000, *When Firms Change Direction*, Oxford University Press, Oxford.

Sigismund-Huff, A., and Schwenk, Ch. R., 1990, Bias and sensemaking in good times and bad, in: *Mapping Strategic Thought*, A. Sigismund-Huff, ed., John Wiley & Sons, Chirchester, pp. 89-108.

Simon, H. A., 1999, *The Sciences of the Artificial*, 3rd ed., The MIT Press, Cambridge, MA.

Sobek II, D. K., Liker, J. K., and Ward, A. C., 1998, Another look at how Toyota integrates product development, *Harvard Business Review* **76**:36-49.

Solomonoff, R. J., 1964, A formal theory of inductive inference, *Information and Control*, **7**:1-22 and 224-54.

Thompson, J. D., 1967, *Organizations in Action. Social Science Bases of Administrative Theory*, Mc Graw-Hill, New York.

Tushman, M. L., and Nadler, D. A., 1978, Information processing as an integrating concept in organizational design, *Academy of Management Review* **3**:613-624.

Visser, B., 2001, *Classifying Organizations of Boundedly Rational Agents*, Mimeo, Erasmus University, Rotterdam.

Visser, B., 2002, *Complexity, Robustness, Performance: Trade-Offs in Organizational Design*, Tinbergen Institute discussion paper.

Womack, J. P., Jones, D. T., and Roos, D., 1990, *The Machine That Changed the World*, Rawson Associates, New York.

NORMATIVE COMMITMENT TO THE ORGANIZATION, SUPPORT AND SELF COMPETENCE

Adalgisa Battistelli[1], Marco Mariani[2] and Benedetta Bellò[3]
[1]*Department of Psychology and Cultural Anthropology, University of Verona, Via S. Francesco 22, 37129 Verona VR; E-mail: adalgisa.battistelli@univr.it*
[2]*Department of Sciences of Education, University of Bologna, Via Zamboni 34, Bologna;*
[3]*Department of Psychology and Cultural Anthropology, University of Verona, Via S. Francesco 22, 37129 Verona VR*

Abstract: The main objective of this study was to examine the relationship between perceived organizational support (POS) and organizational commitment; more precisely, using a sample of 687 employees, we conducted a study to examine the relationship between POS and a) affective commitment and b) normative commitment; the study also aimed to examine the relationship between POS and self competence as its antecedent. Data confirm the hypotheses that: affective and normative commitment were found to be influenced by POS in an equal manner; POS was found to be influenced by self competence.

Key words: perceived organizational support; self competence; affective commitment; normative commitment.

1. INTRODUCTION

Recent changes of the work market, due to socioeconomic changes, globalization, technological progress, have enacted the fundamental value of human resource and the importance of the acquisition and development processes of competences, in an optics of continual growth and development of people, but they also change the nature of the employment relationship having consequences in the way individuals involve themselves in the organizations they work for.

Meyer and Allen (1997) argue that there are several reasons why employee's attachment to the organization is still important and will probably become even more important in the future (Pratt, 2000a). The research points out organizational commitment and organizational support as two of the most relevant dimensions in the relationship between the organization and its members. Organizational commitment above all is essential for the organizations, whereas it seems that the employees modify and adjust their level of organizational commitment as a function of the way they interpret and make sense of their work context (Vandenberger and Self, 1993) and, consequently, they alter their actions within or toward the organization. Organizational commitment is arguably one of the most important factors involved in employees' support especially in time of changes; for example it has been incorporated recently into various theoretical models of change as the Armenakis, Harris e Field's one (1999) of system readiness for change.

Among the antecedent factors, that influence levels of organizational commitment, organizational support has a significant weight, whereas, by virtue of reciprocity norm (Gouldner, 1960), employees who feel supported by their organization will attempt to repay their debt through commitment (Settoon, Bennett and Liden, 1996).

The organizational commitment is a psychological state that characterizes the employee's relationship with the organization where they work and has implications for the decision in continuing to remain members within.

Meyer and Allen (1991) in trying to identify a univocal definition of the term "organizational commitment", noted that the various definitions reflect three broad themes as indicated by the category labels: a) an affective orientation toward the organization, b) a recognition of costs associated with leaving the organization, c) a moral obligation to remain with the organization; so they proposed a three-component model of organizational commitment: affective commitment, continuance commitment and normative commitment. They argue that it was more appropriate to consider the three mentioned as components because an employee's relationship with an organization might reflect varying degrees of all three.

Affective commitment (AC) refers to the employee's emotional attachment to, identification with, and involvement in the organization. Employees with a strong affective commitment continue employment with the organization because they want to do so.

Continuance commitment (CC) refers to an awareness of the costs associated with leaving the organization. The employees remain because they need to do so.

Normative commitment (NC) denotes to what extend a person feels committed to the organization because of normal values and normative beliefs. The employees feel that they ought to remain in it.

Affective organizational commitment was found to be related to a wide variety of correlates. The literature suggested that affective commitment is related to both demographic characteristics and work experience; it was also found to be positively related to performance (Meyer, Paunonen, Gellatly, Goffin and Jackson, 1989).

To understand the different aspects of the relation between individuals and their work organization that can have consequences on the employees' organizational behaviour, reference is often made to the social exchange theory and its related norm of reciprocity, which make people respond positively to favourable treatment received from others (Blau, 1964; Gouldner, 1960). Using a social exchange framework, Eisenberger, Huntington, Hutchinson and Sowa (1986) proposed that members' beliefs concerning perception organizational support are the foundation for perceptions of employees' commitment.

The perceived organizational support (POS) or "employer commitment", indicates an employees' perception concerning the extend to which the organization values their contribution and cares about their well-being (Eisenberger et al., 1986); it pertains to any action taken by the organization or its representatives to benefit the employees whereas the development of perceived organizational support is afforded by the natural tendency of employees to personify their organization by ascribing humanlike characteristics to it.

Shore and Shore (1995) suggested that high organizational support (POS) is a characteristic of a secure, positive environment and that supportive organization could be considered synonymous with a caring workplace. Supportive organizations are those perceived to care about the welfare and needs of their members (Eisenberger et al., 1986). Employees trust the organization that will keep providing rewards and the organization trust employees that will continue performing well (Shore and Wayne, 1993).

The research strongly suggested that perception of POS generates members' feelings of reciprocity thus promoting their commitment towards the organization (Eisenberger et al., 1986; Settoon, Bennett and Liden, 1996).

Organizational support theory suggests that perceived organizational support would reinforce affective commitment to the organization (Eisenberger et al., 1986, 2001; Rhoades, Eisenberger and Armeli, 2001; Shore and Shore, 1995). Several researches emphasized the strong relation between perceived organizational support and continuance commitment (Shore and Tetrick, 1991) and between perceived organizational support and

affective commitment (Eisenberger et al, 1990; Settoon, Bennett and Liden, 1996; Wayne, Shore and Liden, 1997); Rhoades, Eisenberger and Armeli (2001) demonstrate that perceived organizational support led to a temporal change in organizational affective commitment, and not the reverse (Stinglhamber and Vandenberghe, 2003), indicating that perceived organizational support is an antecedent of organizational affective commitment (Rhoades et al., 2001).

H1. Perceived organizational support influences affective commitment.

In reference to Meyer and Allen (1991) and their three dimensions of the organizational commitment (affective, normative and continuance), there is lack of empirical evidences in POS research on the relation with normative commitment.

H2. Perceived organizational support influences normative commitment.

H3. The influences between perceived organizational support and normative commitment and between perceived organizational support and affective commitment are equal.

Self competence is defined as an individual's subjective evaluation of his/her task related ability, self esteem and what he/she is able to do in a specific situation (Ford, 1985; McCombs, 1986). It is considered as the psychological component of the competence and it is related to the capacity to do an activity aimed at an object and the efficacy of the individual's behaviour (Ford, 1985). Self competence is also considered as performance expectancies; persons who perceive themselves as being competent expecting to reach a high level of performance. In this perspective self competence contains the concept of self efficacy (Bandura, 1977). The perception of self efficacy and self competence are related to the perception of "I can do it" and "I am effective" (Markus, Cross and Wurf, 1990; Williams and Lillibridge, 1990). Self competence is specific to a given situation (it is related to a specific task or setting) and it is changeable (it is influenced by the effect of the situation and by dispositional factors). It is also influenced by past successes, but this influence could also be modified by present experiences and by future expectancies.

Some researches on the perceived self competence highlighted the relationship with attitudes toward the organization (Mathieu and Zajac, 1990). As perception of personal capacity, self competence has a real effect on the employee's feelings and behaviours in the organization; it will be interesting to study the relationship between self competence and organizational support.

The human natural tendency to ascribe humanlike characteristic to the organization, based on organization's personification that bring employees to consider the treatment received from the organization as an indication that it favours or disfavours them, could also bring employees to believe that their

competence (perceived self competence) can arouse the organization's interest in them and, as a consequence, that it increases the care and the support that the organization will show for them; briefly, we hypothesize that perceived self competence influences perceived organizational support.

H4. Perceived self competence influences perceived organizational support.

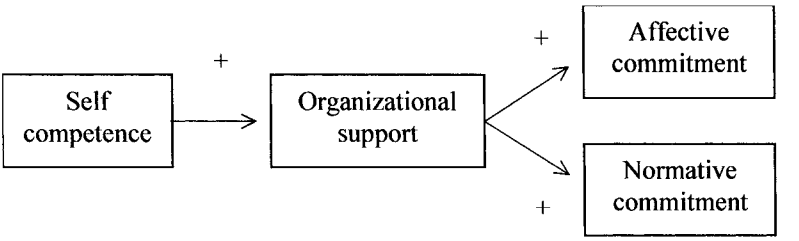


Figure 1. The hypothesized model.

2. METHOD

2.1 Sample and procedure

The hypothesized model was tested on a group of firm workers and then it was cross-validated with two groups of workers employed by training agencies and by high school works.

All subjects of the research are Italian employees (N = 687; M = 41 year old; SD = 9.35); 54% are males. They work in three types of organizations. 47% work in the field of production and distribution (46% of them in commercial distribution and 54% in industrial production); 24% are employees of training agencies; 29% are high school teachers.

Every sample subject answered items on a questionnaire (five-point Likert-type scale). Validity of scale was analyzed by principal components factor analysis and the hypothesis was tested by structural equation models.

2.2 Measures

As the study was conducted in an Italian-speaking-context, all measures were translated from English to Italian by two translators.

We used a questionnaire made up of four scales:

- Organizational support; it was assessed using 4 items of the Eisenberger's 17-item scale (Eisenberger et al., 1986). The scale used was the Survey of Perceived Organizational Support (SPOS). For additional validation evidence see Shore and Tetrick (1991) and Eisenberger et al. (1990). The 4 items used are: "the organization takes pride in my accomplishment at work", "my supervisor is proud that I am part of this organization", "the organization cares about my opinion", and "the organization values my contribution to its well-being". (Alfa of Cronbach: .73);
- Self perceived competence; the scale used was made up of 6 items (es: "I can give my colleagues good advice on the job", "I have all the necessary competence to do my work well"). (Alfa of Cronbach: .84);
- Organizational commitment; the scale used is the "Organizational Commitment Scale" (Meyer and Allen, 1991).Three factors (as reported in the original version) were extracted by principal component analysis (49% of variance) which were rotated by oblimin method. In this research we use only five items of the affective commitment factor (20% of total variance); they are: "I do not feel "emotionally attached" towards this organization", "I do not feel a strong sense of belonging to my organization", "I do not feel "part of the family" at my organization", "This organization has a great deal of personal meaning for me", "I really feel as if this organization's problems are my own" (Meyer and Allen, 1991) (Alfa of Cronbach: .75). Four items of the normative commitment factor were used (18% of total variance): "I would not leave my organization right now because I have a sense of obligation to the people in it", "Even if it were to my advantage, I do not feel it would be right to leave my organization now", "This organization deserves my loyalty", "I owe a great deal to my organization" (Meyer and Allen, 1991) (Alfa of Cronbach: .76).

2.3 Results

2.3.1 The background variables

This section begins with a descriptive analysis of self perceived competence, organizational support and organizational commitment in regard to organizations. Next we present results the hypothesized model. Four one-way anova were performed to determine the effect of organization type on four constructs (Table 1).

Table 1. Average responses on scales by organization (1 = very low; 5 = very high). [* = p < .05; ** = p < .01]

	Firms		Training agencies		High schools	
Self perceived competence scale**	M	SD	M	SD	M	SD
Organizational support scale**	3.69	0.68	3.40	0.59	3.21	0.62
Affective commitment scale	3.11	0.75	2.77	0.75	2.65	0.69
Normative commitment scale**	3.56	0.81	3.67	0.76	3.52	0.70
	3.05	0.91	2.97	0.84	2.72	0.82

The results showed that the effect was significant at the .01 level for three scales: Self perceived competences ($F_{2.664}$ = 35.77), Organizational support ($F_{2.664}$ = 26.53) and Normative commitment ($F_{2.664}$ = 8.58). The firm participants showed higher value in the three constructs.

2.3.2 The hypothesized model test

Initially the cases of the firms (the largest sample) served as calibration sample on which the hypothesized model was tested and on which post hoc analyses were conducted in the process of attaining well-fitting model (Byrne, 2001). Afterwards model cross-validation was conducted with the cases of training agencies and high schools.

Firstly we tested the hypothesized model about the model on firm sample, the Chi^2 statistic test and the other indices show that the model must not be rejected. The goodness-of-fit statistics have good values (Byrne, 2001) (table 2).

Moreover, we did an analysis of standardized regression weights between perceived organizational support and normative commitment and between perceived organizational support and affective commitment; they were equal on the firm sample.

Table 2. Model fit indices on firm sample.

	Model Indices	
	No Equality Imposed constraints	Equality Imposed constraints
Chi^2	0.58	0.69
Df	2	3
P	0.75	0.88
Chi^2 /Df	0.29	0.23
GFI	1.00	1.00
AGFI	0.99	0.99
CFI	1.00	1.00
RMSEA	0.00	0.00

To test this hypothesis we compared the Chi^2 of the model with not equality imposed constraint (0.58) with the Chi^2 of the model with equality imposed constraint (0.69) (table 3).

The difference is not significant, so we chose model 2 (with equality imposed constraints) which is the simplest model (Arbuckle and Wothke, 1999).

Table 3. Model comparisons: goodness-of-fit statistics [$\Delta\chi^2$ = Difference in χ^2 values between models; Δ *df* = Difference in degree of freedom values between models].

	Equality Imposed constraints	χ^2	*df*	$\Delta\chi^2$	Δ *df*	p
Model 1	No	0.58	2			
Model 2	Yes	0.69	3	0.11	1	NS

Results on firm sample are in line with the research hypothesis which claims that self competence influences perceived organizational support (H4) and that perceived organizational support influences normative commitment (H2) and affective commitment (H1). Moreover, results on firm sample show that model with equality imposed constraints (an equal relationship between perceived organizational support and normative commitment and between perceived organizational support and affective commitment were equal on the firm sample) is the best model (H3).

2.3.3 The cross-validation of the model

Previously the originally hypothesized model was tested on the data from calibration sample of firm employees. Here we tested to find the validity of casual structure in the different groups.

Table 4. Simultaneous test of the model on the three samples.

	Model Indices
Chi^2	21.45
Df	9
P	0.01
Chi^2 /Df	2.38
GFI	0.98
AGFI	0.95
CFI	0.97
RMSEA	0.05

The goodness of fit of the model for the three groups in combination is good (table 4): results show an invariance of the model across groups. In this case too, the comparison between this model, with imposed equal influence of organizational support on the two constructs of organizational

commitment, (Chi^2 = 21.45; df = 9) and the model with not constraints (Chi^2 = 19.60; df = 6) brings forward to choose the simplest model (H3).

Table 5. Standardized path coefficients and Squared Multiple Correlations. [** = p < .01; * = p < .05]

Standardized Regression Weights		Firms	Training agencies	High schools
Self perceived professional competence	→Perceived organizational support	0.16**	0.03	0.49**
Perceived organizational support	→Normative commitment	0.41**	0.52**	0.30**
Perceived organizational support	→Affective commitment	0.46**	0.58**	0.34**
Squared Multiple Correlations				
Perceived organizational support		0.03	0.00	0.24
Normative commitment		0.17	0.34	0.12
Affective commitment		0.21	0.27	0.09

Subsequently, we tested for invariance of parameters across the three sample. In this phase, firstly we tested for the invariance of relationship between self competence and perceived organizational support across the three sample (H4). To test this hypothesis we compared the Chi^2 of the model with not equality imposed constraint across groups (21.45) with the Chi^2 of the model with equality imposed constraint across groups (42.74) (table 5). The difference is significant, so we chose model 1 which has a better fit (Arbuckle and Wothke, 1999). So we can claim that the level of influence of self competence on perceived organizational support changes in the three samples. Besides standardized regression weights of this relationship change from .03 to .49 in the three samples; moreover, the standardized regression weights of the training agencies sample was not significant (Table 6).

Table 6. Model comparisons: goodness-of-fit statistics [$\Delta\chi^2$ = Difference in χ^2 values between models; Δ *df* = Difference in degree of freedom values between models].

	Equality Imposed constraints	χ^2	*df*	$\Delta\chi^2$	Δ *df*	p
Model 1	No	21.45	9			
Model 2	Yes	42.74	11	21.29	2	Sig. 0.00

The following test was for the invariance of relationship between perceived organizational support and the two aspects of organizational commitment across the three samples. To test this hypothesis we compared

the Chi^2 of the model with not equality imposed constraint across groups (21.45) with the Chi^2 of the model with equality imposed constraint across groups (28.98) (table 6). The difference is significant, so we chose model 1 which has a better fit (Arbuckle and Wothke, 1999). Thus we can claim that the level of influence of perceived organizational support and the two aspects of organizational commitment changes in the three samples. In this case even if these parameters change in three samples they are always significant (table 7) (H1, H2, H3).

Table 7. Model comparisons: goodness-of-fit statistics (abc) [$\Delta\chi^2$ = Difference in χ^2 values between models; Δ d*f* = Difference in degree of freedom values between models].

	Equality Imposed constraints	χ^2	d*f*	$\Delta\chi^2$	Δ d*f*	p
Model 1	No	21.44	9			
Model 2	Yes	28.98	11	7.54	2	0.02

In summary, hypothesis H1 (perceived organizational support influences affective commitment) and H2 (perceived organizational support influences normative commitment) have been confirmed in the three samples; even if these relationship were not stable in the samples.

Hypothesis H3 which claims that the influences between perceived organizational support and normative commitment and between perceived organizational support and affective commitment are equal, has been confirmed in the three samples.

Finally, hypothesis H3 (perceived self competence influences perceived organizational support) has been confirmed in only two samples.

3. DISCUSSION AND CONCLUSION

In time of change, psychological attachment is important for the core members of the organization; these changes make employees perceiving flexibility and precariousness as giving them a feeling of uncertainty. Meyer and Allen (1997) suggested that employee's attachment to the organization is still important and will probably become even more important in the future for the individual's well being and the organization's success.

Our study aimed to understand which variables can facilitate the employee's attachment to the organization. The research hypotheses are confirmed. Perceived organizational support influences affective commitment according to the results of previous researches (e.g. Eisenberger et al., 1990; Settoon, Bennett and Liden, 1996; Shore and Tetrick, 1991; Wayne, Shore and Liden, 1997).

In addition, even normative commitment seems to be influenced by perceived organizational support.

Finally, the relation between perceived self competence and perceived organizational support seems to be significant in only some organization types.

If employees perceive a high self competence they also perceive to be supported by the organization; it increases the attachment between members and organizations and, according to a social exchange framework, the employees feel part of the organization as a family, feel a sense of belonging to it, feel the organization's problems as their own. On the other hand, they can feel a sense of obligation towards the organization; so, in any case, the relationship between organization and members seems to be stronger; the employees can perceive themselves as a resource for the organization and their expectation to remain within could increase. Moreover, workers believe that their competence (perceived self competence) arouses the organization's interest in them increasing the care and support that the organization will show for them, but similarly, we can say that workers who receive high organizational support, probably think they are important for the organization and so they perceive an high level of self competence. Thus we think that there is a double influence between the two variables; in the future, further researches should study and investigate in greater depth the relationship between perceived self competence and perceived organizational support to understand the direction of influence and possible consequences and implications.

REFERENCES

Arbuckle J. L., and Wothke W., 1999, *AMOS 4.0 User's guide*, Small Waters Corporation, Chicago.

Armenakis, A. A., Harris, S. G., and Feild, H. S., 1999, Paradigms in organizational change: change agents and change target perspective, in: *Hanbook of Organizational Behaviour*, R. Golembiewski, ed., Marcel Dekker, New York, pp. 631-658.

Bandura, A., 1977b, Self-efficacy: toward a unifying theory of behavioural change, *Psychological Review* **84**:191-215.

Blau, P. M., 1964, *Exchange and Power in Social Life*, Wiley, New York.

Byrne, B. M., 2001, *Structural Equation Model with AMOS: Basic Concepts, Applications and Programming*, Erlbaum, New York.

Eisenberger, R., Fasolo, P., and Davis-LaMastro, V., 1990, Perceived organizational support and employee diligence, commitment, and innovation, *Journal of Applied Psychology* **75**:51-59.

Eisenberger, R., Huntington, R., Hutchinson, S., and Sowa, D., 1986, Perceive organizational support, *Journal of Applied Psychology* **71**:500-507.

Ford, M. E., 1985, The concept of competence: Themes and variations, in: *Competence Development*, H. A. Jr. Marlowe, and R. B. Weinberg, eds., C.C. Thomas, Springfield, IL, pp. 3-49.

Gouldner, A. W., 1960, The norm of reciprocity, *American Social Review* **25**:165-167.

Markus, H., Cross, S., and Wurf, E., 1990, The role of the self-system in competence, in: *Competence considered: perception of competence and incompetence across the lifespan*, R. Sternberg, and J. Jr Kolligian, eds., Yale University Press, New Haven, CT.

Mathieu, J. E., and Zajac, D., 1990, A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment, *Psychological Bulletin* **108**:171-194.

McCombs, B. L., 1986, The role of the self-system in self regulated learning, *Contemporary Educational Psychology* **11**:314-332.

Meyer, J. P., and Allen, N. J., 1991, A three-component conceptualization of organizational commitment, *Human Resource Management Review* **1**:61-89.

Meyer, J. P., and Allen, N.J., 1997, *Commitment in the workplace: Theory, research, and application*, Sage, Newbury Park, CA.

Meyer, J. P., Paunonen, S. V., Gellatly, I. H., Goffin, R. D., and Jackson, D. N., 1989, Organizational commitment and job performance: It's the nature of the commitment that counts, *Journal of Applied Psychology* **74**:152-156.

Pratt, M. G., 2000a, Building an ideological fortress: the role of spirituality, encapsulation, and sensemaking, *Studies in Cultures, Organizations, and Societies* **6**:35-69.

Rhoades, L., Eisenberger, R., and Armeli, S., 2001, Affective commitment to the organization: the contribution of perceived organizational support, *Journal of Applied Psychology* **86**:825-836.

Settoon, R. P., Bennett, N., and Liden, R. C., 1996, Social exchange in organizations: perceived organizational support, leader-member exchange, and employee reciprocity, *Journal of Applied Psychology* **81**:219-227.

Shore, L. M., and Shore, T. H., 1995, Perceived organizational support and organizational justice, in: *Organizational Politics, Justice, and Support: Managing the Social Climate of the Workplace*, R.Cropanzano and K. M. Kacmar, eds., Quorum Books, Westport, CT, pp. 149-154.

Shore, L. M., and Tetrick, L. E., 1991, A construct validity study of the survey of perceived organizational support, *Journal of Applied Psychology* **76**:637-643.

Shore, L. M., and Wayne, S. J., 1993, Commitment and employee behaviour: comparison of affective and continuance commitment whit perceived organizational support, *Journal of Applied Psychology* **78**:774-780.

Stinglhamber, F., and Vandenberghe, C., 2003, Organizations and supervisors as sources of support and targets of commitment: a longitudinal study, *Journal of Organizational Behavior* **24**:251-270.

Vandenberger, C., and Self, R. M., 1993, Assessing newcomer's changing commitments to the organization during the first 6 months of work, *Journal of Applied Psychology* **78**:557-568.

Wayne, S. J., Shore, L. M., and Liden, R. C., 1997, Perceived organizational support and leader-member exchange: a social exchange perspective, *Academy of Management Journal* **40**:82-111.

Williams, K. J., and Lillibridge, J. R., 1990, The identification of managerial talent: a proactive view, in: *Psychology in Organizations: Integrating Science and Practice*, K. Murphy and F. Saal, eds., Erlbaum, Hillsdale, NY, pp. 69-94.

A MULTIVARIATE CONTRIBUTION TO THE STUDY OF MOBBING, USING THE QAM 1.5 QUESTIONNAIRE

Piergiorgio Argentero and Natale S. Bonfiglio
Università degli Studi di Pavia

Abstract: The aim of this work was to study the phenomenon of mobbing by adopting a systemic approach. This, in turn, implies a multivariate framework, relying on a structural equation model. In this regard several variables were taken into consideration and their consistence was evaluated through the QMP 1.5 questionnaire. We hypothesized the existence of a difference between victims and non-victims, related to variables like: isolation, exclusion, communication, overload, disheartenment, enthusiasm, and inadequacy. The model obtained in this study evidenced the systemic nature of the interrelations between the different variables: some of them, taken alone or summed together, could not predict the occurrence of mobbing, while this prediction could be based only on their reciprocal relations.

Key words: mobbing; structural equation model.

1. INTRODUCTION: THE MOBBING PROBLEM

During the last years many researchers have considered mobbing as one of the most important problems in workplaces. This is because it involves several factors like physic and mental stress, social relations between people, psychological differences, organizational strategies, and so on.

The mobbing condition was firstly studied in northern Europe by Heinz Leymann (1984, 1986) and, after him, several researchers were interested in studying this problem related to the different country contexts.

The troubles for studying mobbing were and, still are, connected to 1) the *definition* of mobbing and 2) to the *measurement* of it.

The most common definition of mobbing was first proposed by Leyman (1990) description, which considers mobbing as a sequence of vexation conditions that persists in time from, at least, six mounts and with a frequency of, at least, once a week and against the same person, aiming to injure the victim and to keep him out of the workplace.

Some authors have lightly changed this definition, adapting it to their own particular country job context of study. Ege (1998), in particular, had considered the mobbing as a six phases model (despite the four phases model of Leymann), in relation to the Italian situation, and that is: 1) "*zero condition*": that represents the non-victim condition in Italian work context, where is non-victim to be in conflict situations. This doesn't mean that people are vexed or are in a mobbing situation. What is a mobbing indicator in European context workplaces (that is, a conflict condition) is a norm indicator in Italian workplaces; 2) "*mobbing beginning*": where a victim is identified; 3) "*somatic symptoms*": which means that the victim begins to develope some psychological and psycho-physical symptoms that are indicators of psychological troubles; 4) "*personnel manager errors and abuses*": job conditions have decreased and Human Resources are involved. Sometimes the Personnel Manager takes measures against the victim, causing him further troubles (changing of job duty, leaving workplace for a period because of health problems, and so on); 5) "*the psycho-physical condition grow worse*": sometimes the above situations lead to psychological or physical pathologies; 6) "*leaving workplace*": the victim leaves his/her job as their own decision or as a management decision.

As difficult it is the definition of mobbing as complex are the measurements and the methodology used for measuring the problem, because, until today, the methodology used are not quietly shared by all authors (Cowie at al., 2000).

In this work we used the questionnaire methodology for measuring the variables involved in mobbing with the aim to determine the amount of importance of some indicators as predictors of mobbing. We suppose that there are several indicators of trouble conditions and personal and social difficulties, each of them with different importance. Therefore, mobbing is not determined only by the number of indicators as a result of quantity, but also by the relations between these variables, in the sense of dependence and independence that link each other.

For doing this, we have used an equational structural model with the multi-group methodology, in order to compare different models between subjects with mobbing and subjects without mobbing. In that way, we hope to find difference between the two different groups related to some parameters considered in the model.

2. THE QAM 1.5 QUESTIONNAIRE

We used, for the purpose of this work, the QAM 1.5 (Questionnaire for Mobbing auto-Perception) developed by Argentero and Bonfiglio (2004). This instrument is composed of 81 items which evaluate the perception of mobbing that subjects have about ourselves, plus 12 items for a lie scale.

The QAM 1.5 is composed of eight sections, that are: 1) relationships area: composed by items related to communication and relationship sentences and containing two dimensions, that are *isolation* and *communication*; 2) professional area: containing sentences about job exclusion and working difficulties, composed by two dimensions called *exclusion* and *overload*; 3) attacks on person area: composed by items related to oppression situations and injuries and including two dimensions, one for the *hostility* and the one for the *attacks*; 4) double mobbing area: referring to situations that could verify out of the workplace like relations with partner, parents and friends or interests for out of work life; 5) cognitive area: investigates the memory, attention and concentration conditions of subjects and is formed by a dimension for *mental efficiency* and a dimension for *concentration*; 6) psycho-emotive area: containing sentences about the sentimental and emotional conditions of the subjects, related to the workplace and composed of three dimensions, that are *enthusiasm*, *tension* and *self-esteem*; 7) psycho-physic area: presents a set of symptoms related to cardio-respiratory and muscle-skeleton systems and to the sleep rhythms and composed by a dimension for *cardio-respiratory symptoms*, one for *sleep troubles* and another for *muscle-skeleton symptoms*; 8) humour area: structured for revealing the humour of the subjects related to job situations with two last dimensions called *disheartened* and *inadequacy*.

The modality of responses were, for the first four areas, of two types: the first use a frequency Likert scale of five points, and the second use a duration Likert scale of five points too. From the fifth to the seventh areas, the modality of responses were composed by a Likert scale of five points. For the last eighth scale the modality of responses were arranged on a semantic differential.

3. SUBJECTS

We gave the QAM 1.5 questionnaire to 109 subjects, some of them were diagnosed as mobbing victims (N=51) by a Psychological Assistance Centre of Pavia (and defined here as victim subjects).

The total number of males was 55 (34 for non-victims group and 21 for victims group) and the total number of females was 54 (24 for non-victims

group and 30 victims group). The mean age was 39,3 (37,3 for non-victims group and 41,3 for victims group).

Furthermore, 33 subjects had an inferior secondary school education, 56 had a superior secondary school education and 19 had a degree.

4. RESULTS

To evaluate the importance of the variables considered in this study, we proved several structural equation models. In this section we propose the model which better satisfies the condition of parsimony, that better fits and is adequate to the aim of this work. The Figure 1 shows the model.

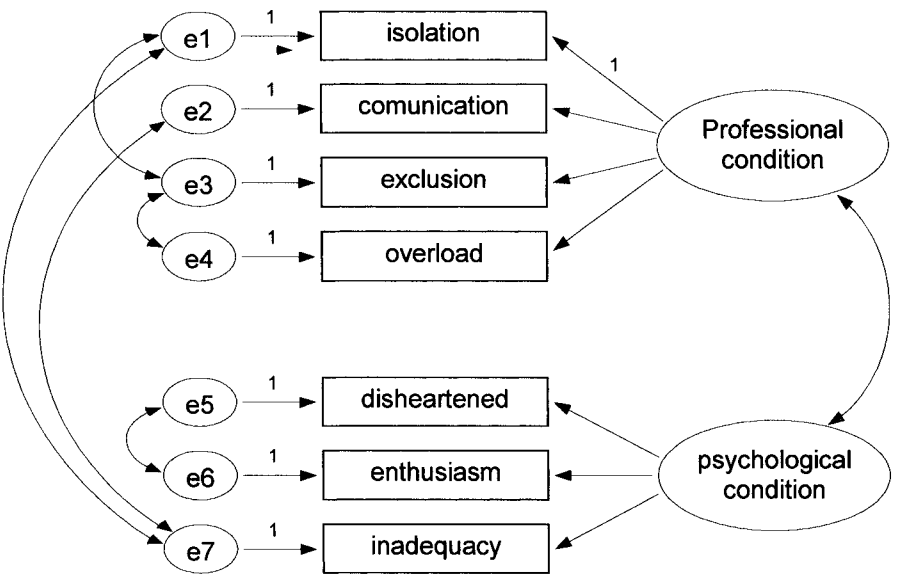


Figure 1. The structural equation model.

The table below (Table 1) shows the model comparison; as you can see, we have compared an unconstrained model, a measurement weights model, a structural covariances model and a measurement residuals. None was significant (except the measurement residuals).

Table 1. The used models

Model	*CMIN*	*DF*	*p*
Unconstrained	22,007	24	0,579
Measurement weights	34,019	29	0,239
Structural covariances	42,313	32	0,105
Measurement residuals	104,043	42	0,000
Saturated model	0,000	0	
Independence model	226,055	42	0,000

The table below (Table 2), shows the fit indexes obtained. As you can see, every models (except the measurement residuals) seem to fit well.

Table 2. Indexes of fit for each model

Model	*RMR*	*GFI*	*AGFI*
Unconstrained	,063	,947	,876
Measurement weights	,104	,917	,839
Structural covariances	,102	,899	,823
Measurement residuals	,135	,806	,741
Saturated model	,000	1,000	
Independence model	,215	,579	,439

The next table (Table 3) shows the differences between the models assuming the unconstrained model to be correct.

Table 3. Confronted models assuming the unconstrained model to be correct.

Model	*CMIN*	*DF*	*p*
Measurement weights	5	12,012	,035
Structural covariances	8	20,305	,009
Measurement residuals	18	82,036	,000

The next table (Table 4), instead, shows the difference between models assuming measurement weights model to be correct.

Table 4. Confronted models assuming measurement weights model to be correct.

Model	*CMIN*	*DF*	*p*
Structural covariances	3	8,294	0,040
Measurement residuals	13	70,024	0,000

After, we evaluated the regression weights (Table 5) between non-victim and victim subjects. You can see two remarkable differences: the one for the influence of professional condition versus the overload (that is not present for victim subjects), and the second for psychological condition versus disheartened (that is lower for victim subjects).

Table 5. Regression weights confronting non-victim with victim subjects

			Non-victims	Victims
overload	←	professional condition	0,469	0,157
exclusion	←	professional condition	0,710	0,730
isolation	←	professional condition	0,734	0,751
communication	←	professional condition	0,421	0,557
inadequacy	←	psychological condition	0,616	0,469
enthusiasm	←	psychological condition	0,832	0,984
disheartened	←	psychological condition	0,905	0,624

Furthermore, the correlation indexes between psychological and professional condition were 0.49 for non-victim subjects and 0.24 victim subject.

5. DISCUSSION

To prove the existence of a difference between non-victim and victim subjects we have trained a multigroup equational structural model.

The analysis used have shown that we can suppose the existence of a difference between subjects with a suspected mobbing situation and subjects without, on the basis of some variables that we have named as psychological condition and professional condition. Instead, the relation between these two variables is not as strong for victim subjects than for non-victim.

Furthermore, the latent variables, for each group, seems to show a strong correlation with the described indicators.

We can conclude, on the basis of a remarkable difference between the models obtained for each group, that the mobbing problem seems to be related, in part, to the professional situation (work stress, vexations, isolations, etc.) and in part to the psychological condition. Instead, for victim subjects this problem is much more related to an individual condition (psychology, personality), in fact we haven't obtained a strong correlation between psychological condition and professional situation.

The model presented in this study has shown very important relations that could be explained in a systemic view: some variables, take alone or summed together, could not predict mobbing, but only the relation between some of them could predict.

To understand better what we mean, try to suppose some situations.

1. Where a subject is lightly psychological weak. This situation could potentially become mobbing because of the stress conditions that involve the subject.
2. Where the workplace context is at risk, because of conflict relations. In this case a subject could become a victim of the situation.

3. Where the organization intentionally makes use of mobbing to keep out a worker. In this case a subject becomes a victim because a situation and a conflict contexts are created.

In the first situation mobbing is potentially possible, but, if the situation is not at risk because of the absence of conflict conditions, mobbing remain in a potential state.

In the second situation mobbing is potentially possible too, but, if a subject is not particularly weak, the context remain the same without consequences.

Let's see, now, the third condition. In this case, if an organization would like to create a mobbing situation it has to create, firstly, a conflict context situation (it couldn't create a victim first!) as a result of this, a victim can be vexed, but only if the victim is particularly weak and stressed.

So, the model showed in this study is able to: a) evaluate different possibilities that could exploit in mobbing situation or are potential mobbing situations; b) single out situation of trouble for both a subject or an organization; c) be adherent with reality.

Finally, we would like to spend some words for the potentiality of the QAM 1.5 questionnaire. His versatility and elasticity linked to his simplicity of compilation and to the variety of variables and situation measured, gives us the possibility to have many indicators and to test many models in order to find the ones that fit better with data, in a multivariate view.

We intend to, for the future, evaluate other relations that are not conceptually described here, like the possibility of a falsification of own health conditions (for a worker) or work conditions (for an organization). We intend to use the methodology of structural equation model to validate the QAM 1.5 deleting the variables that doesn't agree with the theoretical model presented here, in order to present a questionnaire easier to use and to evaluate.

REFERENCES

Argentero, P., Bonfiglio, N. S., and Zanaletti, W., 2004, La percezione del mobbing: alcune determinanti individuali ed organizzative, *Risorsa Uomo - Rivista di Psicologia del lavoro e dell'organizzazione* (in press).

Ceresia, F., and Lupo, I., 2003, Un contributo empirico e metodologico per l'indagine della fenomenologia del mobbing, *Risorsa Uomo - Rivista di Psicologia del Lavoro e dell'Organizzazione* **9**:1,39-53.

Corbetta, P., 1992, *Metodi di Analisi Multivariata per le Scienze Sociali*, Il Mulino, Bologna.

Cowie, H., Naylor, P., Rivers, I., Smith, P. K., and Pareira, B., 2000, Measuring workplace bullying, *Aggression and violent behavior* **7**:33-51.

Ege, H., 1998, *I Numeri del Mobbing. La Prima Ricerca Italiana*, Pitagora Editrice, Bologna.

Leymann, H., 1990, Mobbing and psychological terror at workplaces, *Violence and Victims* **5**:119-126.

Leymann, H., 1986 *Vuxenmobbning, om Psykiskt Valdiarbetslivet* (Mobbing, psychological violence at work places), Studenlitteratur, Lund.

Leymann, H., and Gustafsson, B., 1984, *Psykiskt valdiarbetslivet. Tva esplorativa undersokningar* (Psychological violence at work places. Two esplorative studies), (Undersokningsrapport 42), Arbetarskyddsstyrelsen, Stokholm.

Zapf, D., Knortz, C., and Kulla, M., 1996, On the relationship between mobbing factors and job content, social work, environment, and health outcomes, *European Journal of Work and Organizational Psychology* **5**:215-237.

REPRESENTATION IN PSYCHOMETRICS: CONFIRMATORY FACTOR MODELS OF JOB SATISFACTION IN A GROUP OF PROFESSIONAL STAFF

Maria Santa Ferretti and Piergiorgio Argentero
Department of Psychology, University of Pavia

Abstract: This article falls within the scope of one of the most recent interpretations of psychometrics, characterized by the concept of "representation". Formal models and theories to define latent structures represent a particularly important research tool in modern psychometrics, in which psychological disciplines interact with disciplines that envisage formalization, such as mathematics and statistics. The study in question was used as a means of identifying a model that represents the structure of the typical constructs of a set of items in a job satisfaction scale. The results, consistent with the current literature, demonstrate the multidimensional nature of the construct even among the sample population of managerial staff, yet also identify some characteristic dimensions that, while distinct, are strictly correlated.

Key words: psychometry; structural models; job satisfaction; professional staff.

1. INTRODUCTION

According to a common approach in modern psychometrics, researchers in this sector are primarily concerned with "representing" the relations between the variables that are being examined, by defining formal models to assess and represent psychological processes and phenomena. This is achieved by applying the methods and language of disciplines such as mathematics, statistics and computer science to various aspects of psychometric research. In this perspective it is apparent that all areas of psychology may include aspects of interest that can be represented using formal models. It thus follows that the use of formal representation models is

an interdisciplinary approach, not only because it makes use of the methods and language of disciplines that envisage formalization, but also because it is suitable for use in the various areas of psychological research such as, in our case, to assess job satisfaction.

The concept of job satisfaction has been the subject of organizational psychology research programmes for a long time; initially assessed alongside the concept of "morale" (Viteles, 1932), it was later analyzed from various research perspectives in order to examine the possible causes as well as its effects on the individual's behaviour at work. Some of the main theories in this area include the *situational theories* (e.g. Herzberg's two-factor theory, 1967, or Hackman and Oldham's job characteristics model, 1976), *dispositional theories* (Judge, Locke and Durham, 1997; Staw et al., 1986), *interactive theories* (Lawler, 1973; Locke, 1976), *social information processing theories* (Salancik and Pfeffer, 1978; O'Reilly and Caldwell, 1985) and *person-environment fit* theories (Lofquist and Dawis, 1969; Wanous, 1978, 1992).

Without dwelling upon the pertinence of the various theoretical approaches to job satisfaction (for a recent review of these see Judge et al., 2001), it is important to note that job satisfaction is a central aspect of many present-day models or theories regarding work behaviour and attitudes, and carries important practical implications for improvements to the quality of work life and organizational efficiency.

However it is defined[23], satisfaction is a variable that depends upon and embraces different aspects and facets of the relationship between individual and work. The factors that may influence job satisfaction can be classified, for example, as objective and subjective factors. The former refer to working conditions, career opportunities, salaries and benefits, relationships (defined on the basis of various classifications and hierarchies). The latter refer to variables such as age, gender, seniority on the job, education and training, position or grade in the hierarchy. Whereas in the past satisfaction was essentially studied and observed according to a one-dimensional approach, nowadays researchers recognize its complex and multi-dimensional nature. We have moved on from some general and simplistic observations to a series of increasingly detailed conceptual and empirical definitions. As a whole, the studies and theories in this field have enabled us to understand and identify the many factors that affect satisfaction and its possible

[23] Among the many definitions in the literature, the following are widely recognized as being generally accepted: according to Locke (1976) satisfaction is "a positive emotional state resulting from the appraisal of one's job or job experiences", and Fourgous and Itturalde (1991) define job satisfaction on the basis of how one perceives one's own professional context: "satisfaction is measured by comparing what one has with what one wants, because it seems fair and because it seems desirable or represents a value".

consequences. According to Locke (1976) these are: turnover, absenteeism, physical health, mental health, complaints and grievances, attitude towards work, self-esteem. The sources of satisfaction – or factors that may affect it – have been classified in many different ways. Herzberg (1967) and Locke (1976) are the authors of some of the most important of these classifications. According to Herzberg dissatisfaction is caused by factors that are related to the work environment. These are known as "hygiene factors" or *extrinsic* factors and include, for example, company policies and administration, technical supervision, salary, interpersonal relations, working conditions. The sources of satisfaction, called "motivators" or *intrinsic* factors, are related to the content of the job: the work itself, responsibilities, advancement, recognition, achievements. Locke suggests a different method for classifying the factors that affect satisfaction. He defines three groups of factors. The first group consists of variables relating to *working conditions* such as: working conditions, the work itself, salary, advancement, recognition, benefits. The second group consists of variables regarding *co-workers*. Finally, Locke recognizes a third set of variables regarding *individual characteristics*: age, gender, seniority on the job, education and training, position in the hierarchy.

2. OBJECTIVES

The literature on job satisfaction has examined many occupational categories (e.g. labourers, office workers, health service employees), but there are very few works that specifically address the issue of job satisfaction among managerial staff.

For this reason we decided to assess job satisfaction at managerial level, with regard to both the dimensional and structural aspects.

The main objective of the study was to define and use the appropriate formal models and analysis procedures in order to represent sets of variables that are important in job satisfaction assessment. The approach based on the concept of representation facilitates the reliable formal modelling of the empirical results.

The term LISREL (Linear Structural Relationship), first used in the early Seventies as the name of a computer programme developed by the Swedish statistician Karl Joreskog and his team to perform maximum likelihood estimation of factor loadings (Joreskog and van Trillo, 1973), quickly evolved into a more general procedure for structural equation models, while continuing to distinguish between observed and latent variables (Joreskog, 1973). Nowadays LISREL is the name most commonly used to describe the general theoretical approach that may comprise different methods: factor

analysis, measurement models, path analysis, structural equation models, structural covariance analysis, etc.

This approach responds to the need of social science researchers for a set of valid and reliable measurement tools capable of expressing the relationships between indicators that are used (observed variables) and latent variables, as well as suitable instruments and methods for empirically measuring the existence of any causal relationships between theoretically-formulated variables. To satisfy these requirements LISREL consists of two parts: a measurement model and a structural model. The first describes how the latent or underlying variables are measured using the observed variables, the second specifies the causal relationships between the latent variables and describes the causal effects and assigns the total unexplained variance (Joreskog and Sorbom, 1982).

The term factor analysis refers to a series of statistical techniques that are used to express a set of observed variables in terms of a smaller number of latent (hypothetical) variables or factors. Factor analysis, especially using the confirmatory model, enables researchers to use statistical tools to define the most suitable structure of the dimensions that represent the constructs being analyzed.

The purpose of this study is to use the confirmatory approach to assess the factorial structure of the satisfaction scale by defining and comparing different models. The parameterization of the confirmatory factor models is a means of operationalizing hypotheses concerning the structure of the tool. However, in order to avoid repetitions, the theoretical and methodological reasons for choosing these models are illustrated below.

3. METHOD

3.1 Subjects

The study involved a group of 800 *professional staff* from a multinational engineering firm based in Italy (*professional staff* means employees covering the role of "manager" or "executive"). The study involved all of the firm's professional staff (and thus addressed the entire company population) employed in the technical, production, sales and personnel sectors.

The majority of subjects were males (93%), aged more than 40 (72%), employed as managers (92%) in the technical (26%) and production (35%) sectors with several years in the company (79% more than 10 years). The distribution of the individual characteristics partly reflect the industrial sector in which the company operates (engineering) and partly the high level

in the hierarchy of the company population that was studied (managerial staff).

3.2 Research tool

The tool used to conduct this study consisted of an anonymous questionnaire containing 16 items with multiple-choice answers in which the subjects were asked to indicate their level of agreement with each statement (on the basis of their own experience with the company) using an agreement scale from 1=strongly disagree to 5=strongly agree. All of the statements were worded positively so that a value of 1 always indicated the lowest level of satisfaction and a value of 5 indicated maximum satisfaction.

The statements in the questionnaire were taken from a pilot study, consisting of individual and group interviews with a sample of subjects, from which we were able to define the dimensions deemed important in terms of job satisfaction. These were classified (according to the recent literature in the specific field, see Xuang and Van de Vliert, 2003) as intrinsic and extrinsic factors; the first group included characteristics related to the job itself and to interpersonal relations with superiors; the second group included work aspects related to salary, status, career.

The questionnaire measured the following constructs:

- *Satisfaction related to opportunities for growth and professional development*, meaning the use of one's skills and abilities also by means of adequate training opportunities and mobility. Example of item: "I think this company offers me good opportunities for training and professional development";
- *Satisfaction concerning information received*, meaning with regard to corporate strategies and the company's actual achievements. Example of item: "I feel that I am well-informed as regards the company's achievements";
- *Satisfaction related to salaries and benefits*. This area refers to recognition at work of personal contributions in terms of competence and suggestions for alternative solutions. Example of item: "I am satisfied that my personal contributions are adequately recognized by my company";
- *Satisfaction concerning relations with superiors*. This refers to the existence of a professional relationship with a superior who dialogues, supplies concrete support and promotes professional development of subordinates. Example of item: "Whenever I am in difficulty or doubt I can count on the concrete support of my direct superior".

In the specific sample population that was studied, the work aspects that are sources of professional satisfaction, according to the subjects, relate to

four main areas. These are a significant reflection of the professional level of the subjects themselves in that, for example, importance is given to factors concerning the amount of information that is received, while no reference is made to interpersonal relations - except for those with superiors.

3.3 Procedure

The questionnaire was presented to the *professional staff* during the course of specific meetings organized by the management. At these meetings the contents and objectives of the questionnaire were explained and the subjects were instructed on how to answer the questions. The questionnaires were completely anonymous and those taking part in the study were assured that it would be impossible for the management to trace their identity. The meeting, which lasted approximately one hour, was held during working hours. The questionnaires were collected within a one-and-a-half-month period.

3.4 Data analysis

Data were analyzed in two stages:

1. identification of the latent structure using different confirmatory factor analysis models;
2. multisample factor analysis to assess the factorial invariance of the scale.

We used the "cross-validation" method according to Cudeck and Browne (1983; Bagozzi and Baumgartner, 1994), which consists of randomly dividing the overall sample into two subsets, developing a model on the first and then assessing its generalizability. The group of 800 subjects was thus randomly divided into two subsets each of 400 subjects.

Three different models were tested on the first group. The first to be evaluated was the single-factor model (Model 1). This model is based on the hypothesis that the variance of the answers to the questionnaire can be divided into a general factor plus the error variance associated with each single item. It is standard practice to assess the fit of a single-factor model in that this is the most parsimonious of all the possible models.

Model 2 is based on the hypothesis that the satisfaction scale measures 4 independent (orthogonal) factors (Salaries, Relations with superiors, Information, Personal development) that correspond to the four areas of the questionnaire. Moreover, in this model the items of the 4 factors load only on the respective factors.

Model 3 is based on the hypothesis that the 4 factors are correlated.

The confirmatory factor analysis model is illustrated in Figure 1.

The theoretical models that were taken into consideration were assessed by analyzing the respective structural equation models using the AMOS 5.0 software package (Arbuckle, 2003). The fit of the confirmatory factor analysis models was assessed using the $\chi 2$ test. Fit is considered satisfactory when $\chi 2$ is not significant; however, given its dependence on the size of the sample, other independent indices were also considered, namely the CFI - Comparative Fix Index (Bentler, 1990), the TLI - *Tucker-Lewis Index* (Tucker and Lewis, 1973) and the RMSEA - *Root Mean Square Error Approximation* (Steiger, 1990). For the first two indices, using values of between 0 and 1, values of more than .90 were considered to be satisfactory, as recommended by Bentler (1990). As regards the RMSEA index we followed the advice of Browne (1990), according to whom values of less than .08 should be considered satisfactory and those of less than .05 good (Marsh, Balla and Hau, 1996). The tables containing the *fit* indices for the models produced also include the $\chi 2$ values, which were not deemed essential in that this test proved to be highly sensitive to the size of the sample. It may also be significant for small discrepancies between hypothetical and observed variables (Bollen and Long, 1993; Primi, 2002).

Finally, we assessed hypotheses regarding the actual factorial invariance across the two subsets.

This procedure enabled us to analyze the factorial invariance of the structure across the groups (Joreskog and Sorbom, 1989; Bagozzi and Foxall, 1995).

Firstly, we postulated that factor loadings were invariant across the groups: acceptance of this hypothesis signifies that the measurements reflect the same constructs across the samples. Secondly, we formulated the hypothesis that the measurement errors were invariant: acceptance of this hypothesis confirms the same reliability of the measurements across the samples. Finally, the last hypothesis related to the equivalence of the variances and covariances of the latent factors: the acceptance of this hypothesis means that the constructs are correlated in the same way across both groups. Usually only verification of the factor loading invariance hypothesis is requested to assess the generalizability and stability of the significance of the constructs across the samples. However, the stability of the error components and relations between constructs strengthen confidence in the validity of a set of measurements (Bagozzi and Foxall, 1995).

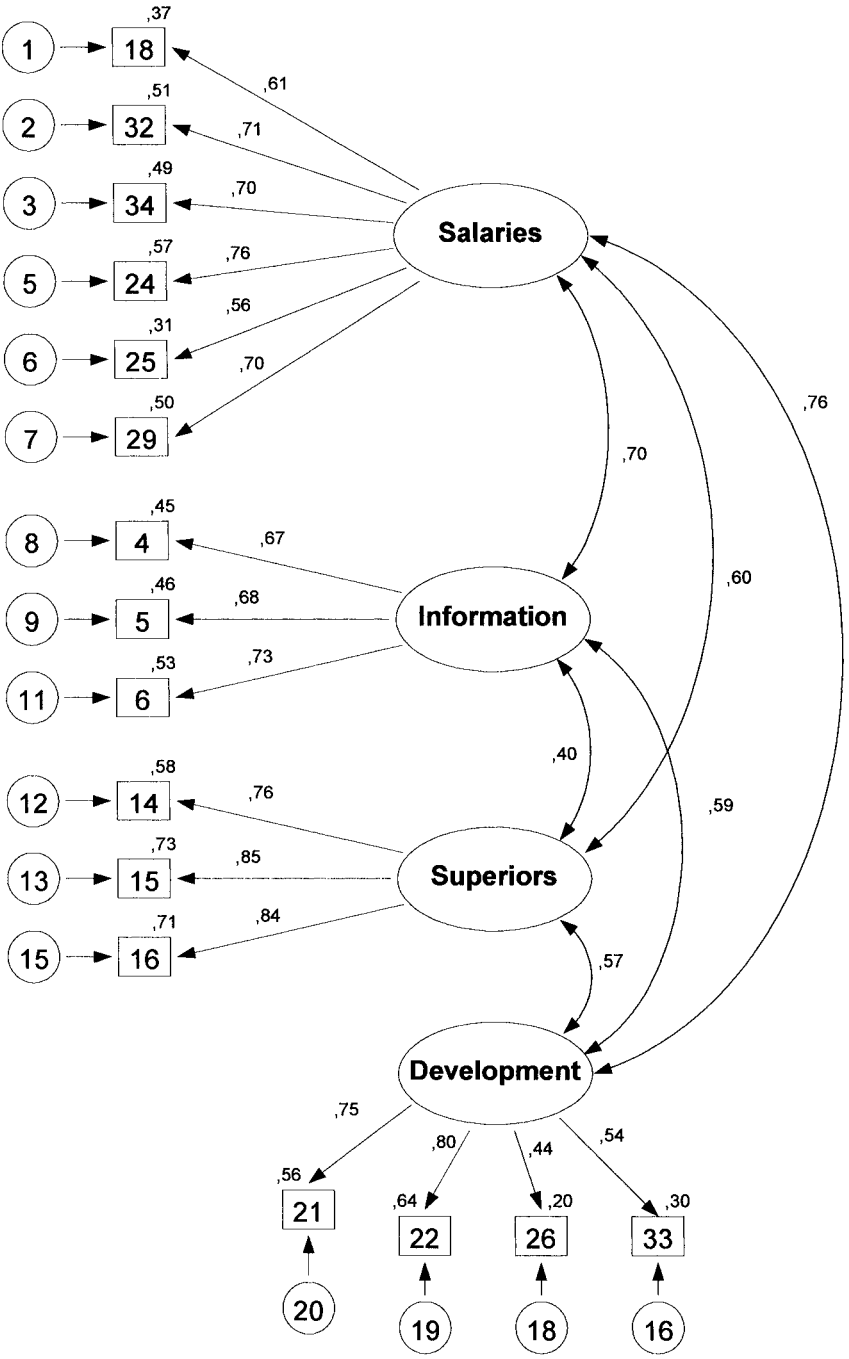


Figure 1. The confirmatory factor analysis model.

4. RESULTS

Three different models were used to assess the factorial structure.

The first single-factor model consisted of one latent variable and 16 observed variables.

Table 1 summarizes the fit statistics for the confirmatory models. The general factorial model (Model 1) has a low fit index; however, loadings on this factor are high: average loading is .59. Thus, even if the hypothesis that satisfaction is a one-dimensional measurement is untenable, there is evidence of a substantial common variance between the items.

Table 1. Fit indices for confirmatory factor analysis models (best fitting model in bold).

	χ^2	gl	χ^2/gl	RMSEA	TLI	CFI
Model 1 (Single factor)						
Sample 1	672.280	104	6.464	.117	.733	.769
Model 2 (4 orthogonal factors)						
Sample 1	657.583	104	6.323	.116	.740	.775
Model 3 (4 correlated factors)						
Sample 1	**177.899**	**98**	**1.815**	**.045**	**.960**	**.967**
Sample 2	212.323	98	2.167	.054	.935	.947

We then tested Model 2, based on the hypothesis that the questionnaire measures 4 independent factors. The model thus comprised 16 observed variables and 4 latent variables (Figure 1). The results show no improvement to the fit indices.

The correlated factors model (Model 3) produced much higher indices than the independent factors model, demonstrating that the hypothesis of independence between scales is untenable.

The fit indices that were chosen all indicate that the model reports data well. The comparative fit index (Bentler, 1990) is more than .90 (CFI= .97), the RMSEA index (Root Mean Square Error of Approximation, Steiger, 1990) is less than .08 (RMSEA = .04), and the Tucker and Lewis index (1973) is more than .90 (TLI= .96).

Table 2 summarizes the factor loadings of each item. As regards the validity of the measurements, the convergent validity is demonstrated by the fact that each item only loads on its respective factor (Table 2). All loadings are high and significant. Discriminating validity is demonstrated by the fact that the correlations between constructs are either not significant (not different to zero) or less than 1.00.

Table 2. Confirmatory Factor Analysis: 4 correlated factors. Structural coefficients: standardized estimates.

Parameters			
Salaries		Superiors	
V29	.704	V16	.842
V25	.559	V15	.853
V24	.756	V14	.758
V34	.698		
V32	.712	Development	
V18	.608	V26	.445
		V33	.545
Information		V22	.801
V6	.726	V21	.749
V5	.779		
V4	.663		

If all correlations between factors are significant as in this case (Table 3), the confidence interval that is obtained by considering two standard errors above and two standard errors below the hypothetical correlation ($p = .05$) does not contain the perfect correlation. The highest Φ regards the variables for "Salaries" and "Relations with superiors" ($\Phi i,d = .76$); thus also in this case the correlation differs significantly from 1.00 (the confidence interval is: $.61 \leq \Phi i,d \leq .91$).

Table 3. Factor correlations (standard errors in brackets).

Group 1 (N = 400)

	Salaries	Information	Superiors	Development
Salaries		.696 (.068)	.762 (.077)	.404 (.066)
Information			.596 (.072)	.592 (.071)
Superiors				.572 (.080)
Development				

Group 2 (N = 400)

	Salaries	Information	Superiors	Development
Salaries		.539 (.053)	.760 (.071)	.228 (.058)
Information			.533 (.066)	.606 (.064)
Superiors				.503 (.075)
Development				

If the structure is valid, it must prove to be stable across the subsets. As shown in Table 1, both models are satisfactory as regards the values for CFI, TLI and RMSEA, and are therefore acceptable.

Having assessed the theoretical model, we tested the hypothesis of structural invariance between the different samples. This was performed according to the procedure defined by Reise, Widaman and Pugh (1993). We

then tested a second model (M1) based on the assumption of invariance of the hypothetical factor *loadings* on the 2 models. The invariance hypothesis is acceptable if the difference between the values for x2 in model M1 and the values in the Bs (Baseline) model is not significant for a number of degrees of freedom equal to the difference between the degrees of freedom of the two models.

Factorial invariance was tested by simultaneously analyzing the model illustrated in Figure 1 in the two groups. First of all the baseline model was tested, in which in both groups the factors corresponded to the same indicators but there were no restrictions concerning the equivalence of parameters across the samples. Apart from being a means of comparison when investigating different invariance hypotheses, the baseline model can also be seen as a means of verifying the presence of the same number of factors across the samples. The chi-square value in the baseline model is significant ($p < .001$), but the other indices that were measured demonstrated the practical goodness-of-fit of the model in question (see table). Having established that the observed variables are at the base of the same factors in the two samples, it was possible to examine the factorial variance of the scale in greater detail. Table 4 summarizes the hypothetical invariances that were tested. First of all the baseline model (M1) was compared against the model (M2) which included an additional restriction concerning the equivalence of factor loadings across the samples. M2 verifies whether the indicators have the same amount of true variance in the two samples, and thus whether the significance of the latent factors is invariant across the two groups.

Table 4. Results of testing for differences between multisample confirmatory factor analysis models.

Model	χ^2	RMSEA	CFI	TLI	Hypothesis test
M1: Unconstrained model	390.222	.035	.958	.949	
M2: measurement weights (factor loadings)	400.159	.034	.958	.952	M2-M1 X2 = 9,937 P < .62
M3: structural covariances	412.877	.033	.958	.954	M3-M2 X2 = 12.718 P < .24
M4: measurement residuals	430.373	.032	.957	.956	M4-M3 X2 = 17.496 P < .35

The difference between the chi-square values in M1 and M2 indicates that the loading invariance hypothesis must be accepted. Model 3 establishes the invariance of the correlations between the constructs, obtaining a satisfactory degree of adequacy, as shown by the RMSEA index that is less than .08 (Browne, 1990). The comparison between M4 and M3 adds to the restrictions on M3 the invariance of the error components across the samples and the comparison indicates its adequacy.

5. DISCUSSION AND CONCLUSIONS

We were able to draw a number of important conclusions from data that had previously been gathered and analyzed.

First of all we were able to define the important satisfaction factors for people high up in the work organizational hierarchy. The areas examined in the questionnaire using items that had been selected previously during the course of a preliminary study represent the four latent constructs of job satisfaction (Salaries, Information, Relations with Superiors and Personal Development). The analyses that were conducted, by processing different models, excluded the possibility that job satisfaction in the specific sample group can be considered as a one-dimensional construct. These results, while confirming research described in the specific literature with regard to the multi-dimensional nature of the construct, also reveal a number of special aspects regarding the characteristics of the important dimensions. Satisfaction related to Information is identified as a dimension that contributes to the definition of the construct.

The outcome of the comparison between the orthogonal model and the correlated factors model lies in favour of the latter, confirming the existence of the four constructs that, while distinct, are closely correlated.

Finally, multisample analysis models achieve a good level of generalizability of the emerging structure. The results demonstrated far more than is required to assess the generalizability of the construct measurements. Not only was each factor loading invariant across the groups, but even the error variances and variances-covariances between latent factors were identical across the two subsets. Further research is required in order to confirm these findings, more specifically by taking into consideration a population of the same level but from a different work organization.

In conclusion, the results of this study demonstrate that more knowledge and understanding can be obtained when, in addition to developing adequate psychological theories, we also construct and apply formal theories that can be used in the various areas of psychological research.

BIBLIOGRAPHY

Adigun, I. O., and Stephenson, G. M., 1992, Sources of job motivation among British and Nigerian employees, *Journal of Social Psychology* **132**:369-376.

Bentler, P. M., 1995, *EQS Structural Equations Program Manual*, Multivariate Software Encino, CA.

Bentler, P. M., and Mooijart, A., 1989, Choice of structural model via parsimony: A rationale based on precision, *Psychological Bulletin* **106**:315-317.

Biggs, D. B., et al., 1991, A method of choosing multiway partitions for classification and decision trees, *Journal of Applied Statistics* **18**:49-62.

Carmines, E., and McIver, J., 1981, Analyzing models with unobserved variables: analysis of covariance structure, in: *Social Measurement*, G. Bohnrstedt and E. Borgatta, eds., Sage, Newbury Park, CA, pp. 65-115.

Fourgous, J. M., and Itturalde, B., 1991, *Mesurer et Améliorer le Climat Social dans l'Entreprise*, Les Editions d'Organisations, Paris.

Francès, R., 1981, *La Satisfaction dans le Travail et l'Emploi*, Presses Universitaires de France, Paris.

Hackman, J. R., and Oldham, G. R., 1976, Motivation through the design of work: test of a theory, *Organizational Behavior and Human Decision Processes* **16**:250-279.

Harding, S., Phillips, D., and Fogarty, M., 1986, *Contrasting Values in Western Europe*. The Macmillan Press, London.

Herzberg, F., 1967, *Work and the Nature of Man*, World Book, Cleveland.

Hofstede, G. H., 1991, *Cultures and Organizations: Software of the Mind*, McGraw-Hill, London.

Hofstede, G. H., 2001, *Culture's Consequences: Comparing Values, Behaviours, Institutions, and Organizations across Nations*, Sage, Thousand Oaks.

Hu, L. T., and Bentler, P. M., 1998, Fit indices in covariance structure modeling: Sensitivity to underparametrized model misspecification, *Psychological Methods* **3**:424-453.

Hu, L. T., and Bentler, P. M., 1999, Cutoff criteria for fit indexes in covariances structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* **6**:1-55.

Huang, X., and Van de Vliert, E., 2003, Where intrinsic job satisfaction fails to work: national moderators of intrinsic motivation, *Journal of Organizational Behavior* **24**:159-179.

Inglehart, R., 1997, *Modernization and Post Modernization: Cultural, Economic and Political Change in 43 Societies*, Princeton University Press, Princeton.

Joreskog, K. G., 1973, A general method for estimating a linear structural equation system, in: *Structural Equation Models in the social sciences*, A. S. Goldberg and O. D. Duncan, eds., Seminar Press, New York, pp. 85-112.

Joreskog, K. G., and van Thillo, M., 1973, *LISREL: a General Computer Program for Estimating a Linear Structural Equation System Involving Multiple Indicators of unmeasured variables*, Research report, 73-5, Uppsala University, Dept. of Statistics, Uppsala.

Judge, T. A., Locke, E. A., and Durham, C. C., 1997, The dispositional causes of job satisfaction: a core evaluations approach, *Research in Organizational Behavior* **19**:151-188.

Judge, T. A., Parker, S. K., Colbert, A. E., Heller, D., and Ilies, R., 2001, Job satisfaction: a cross-cultural Review, in: *Industrial, Work and Organizational Psychology*, N. Anderson, D. S. Ones, H. K. Sinangil and C. Viswesvaran, eds., Sage, London.

Lawler, E. E., 1973, *Motivation in Work Organization*, Brooks/Cole, Monterey.

Locke, E. A., 1976, The nature and causes of job satisfaction, in: *Handbook of Industrial and Organizational Psychology*, M. D. Dunnette, ed., Rand McNally, Chicago.

Lofquist, L. H., and Dawis, R. V., 1969, *Adjustment to Work*, Appleton, New York.

Maslow, A. H., 1954, *Motivation and Personality*, Harper and Row, New York.

Salancik, G. R., and Pfeffer, J., 1978, A social information processing approach to job attitudes and task design, *Administrative Science Quarterly* **23**:224-253.

Schein, E. H., 2000, *Culture d'impresa*. Raffaello Cortina Editore, Milano.

Staw, B. M., Bell, N. E., and Clausen, J. A., 1986, The dispositional approach to job attitudes: a lifetime longitudinal test, *Administrative Science Quarterly* **31**:437-453.

Stone-Romero, E. F., Stone, D. L., and Salas, E., 2003, The influence of culture on role conceptions and role behavior in organizations, *Applied Psychology: an International Review* **52**:328-362.

Veenhoven, R., 1991, Is happiness relative? *Social Indicators Research* **24**:1-34.

Viteles, M. S., 1932, *Industrial Psychology*, Norton and Company, New York.

Wanous, J., 1978, Realistic job preview: can a procedure to reduce turnover also influence the relationship between abilities and performance? *Personnel Psychology* **31**:249-258.

Wanous, J., 1992, *Organizational Entry: Recruitment, Selection, Orientation and Socialization of Newcomers*, Addison-Wesley, Reading.

Wasti, S. A., 2003, The influence of cultural values on antecedents of organizational commitment: an individual-level analysis, *Applied Psychology: an International Review* **52**:533-554.

8

SOCIAL SYSTEMS

THE IMPACT OF EMAIL ON SYSTEM IDENTITY AND AUTONOMY: A CASE STUDY IN SELF-OBSERVATION

Lucio Biggiero
Università dell'Aquila; Dipartimento di Sistemi e Istituzioni per l'Economia
Piazza del Santuario 19, 67090 Roio Poggio (AQ); biggiero@ec.univaq.it

Abstract: Although at the core of systems theorizing, the concepts of system's identity and autonomy are still lacking remarkable empirical tests. Both the characteristics come from recursive self-organizing and self-referential processes. In the case of human systems, they are mainly based on self-observation. This property takes place through cognitive and communication patterns. The present paper analyzes the COMMORG case, an international research team which, besides the formal identity given by the European Union administration, built up its own identity and autonomy during its working life. The COMMORG system set up its self-observation by means of three different methodological tools: emailing list analysis, genre repertoire analysis, and self-survey analysis. They show the structure and evolution of the communication patterns forging its identity and autonomy. The use of an emailing list revealed as the central means of communication, which enabled the development of the system's identity in the form of a (virtual) international research team. The system identity, the identification and trust of its members, the communication patterns, and the semiotic patterns are recursively related to structural and social aspects, and change over time.

Key words: autonomy; email; emailing list analysis; genre repertoire; group identification processes; international research team; knowledge transfer; organizational identity and trust; research methodologies; second-order cybernetics; self-observation; self-organization; self-reference; system's identity; virtual team.

1. INTRODUCTION

The underlining of the self-referential characteristics of many cybernetic systems is one of the main innovation of the second-order cybernetics,

respect to the early stage of cybernetics from mid 40' to mid 70'. When concerning social systems, self-reference assumes specific peculiarities (Biggiero, 1997, 2001a; Geyer, 2002; Geyer and van der Zouwen, 1990), among which the most important is the ability of self-observation (Geyer, 2002), that is, the ability of the whole system -or of some of its parts- to observe itself. The self-observation can be aware or unaware, intentional or unintentional, formal or informal. When there is a certain degree of awareness, in the sociological literature that ability is usually addressed to as self-reflexivity.

These properties are evidently showed by any social system at any level of description/aggregation (Biggiero, 1997, 2001a): plans or reflections or perceptions of a self-employer or consultant or just single worker or person on/of himself; plans or reflections or perceptions of a (member of a) group (department, office, etc.) on the whole group or parts of it; and so on, until the highest level, which is represented by supra-national organizations, like UN, OECD, etc., which can observe themselves or some of its members. At the level of the whole society and in the evolution of sociology the self-reference and the self-organization properties became permanent and distinctive characteristics (Leydesdorff, 2001). However, excepted for the human society taken as a whole system, the single social systems cannot be considered as autopoietic systems (Biggiero, 2001c).

These properties have a lot of methodological, epistemological and operational consequences on the studying, the evolving, the learning and the managing of the social systems (Biggiero, 1997, 2001a; Foerster, 1982; Geyer, 2002). Maybe the most important one is that systems'a identity and autonomy are both related to, and dependent on, self-reference (Biggiero, 1999; Luhmann, 1990). While it is possible to conceive autonomy without self-reference, it is impossible to get an identity without self-reference. Another strict implication is that self-reference is based on communication. It could also be said that self-referring is communicating with ourselves. When this mechanism takes place in human systems, then communication is mostly non-physical and non-biological: it is mostly verbal (oral and written), para-verbal (i.e. the tone of the voice), and non-verbal (i.e. the gesture). Thus, in order to study the creation, the structure and the evolution of system's identity and autonomy, it is crucial to understand also the means and the patterns of communication by which the self-referential processes do occur.

Just recently it is becoming apparent the self-referential characteristics of the knowledge growth at the level of single organizations (Nonaka and Takeuchi, 1995). Biggiero (1999, 2001b) and Sammarra and Biggiero (2001) evidenced the effects of these properties at different systemic levels: organizational, inter-organizational, and district levels. A lot of interesting

research questions stem from these consequences: in which ways does the self-reference or the self-observation take place in different social systems and in different contexts? Is there any disadvantage in getting very high or very low degree of self-reference? To what extent is it possible to steer these processes? What type of consequences do have computer-mediated communication technologies in the specific way of the realization of the self-reference? Are virtual teams special cases? How do measure and operationalize systems closure, identity and self-reference? And many other more. To some of these research questions we can answer discussing this case study, whose results are definitely far from being generalized, but they can be a first brick in the perspective of giving empirical content to some of the central issues of cybernetics and systems sciences.

Notwithstanding the relevance of self-observation, there is a substantial lack of empirical research. On one side organization, management, and social science studies in general, deal implicitly with these properties in almost any study, but usually they overlook or neglect them. This way those studies are impoverished and limited in its explanatory power. On the other side, cybernetics studies remain almost always only on the theoretical aspect, lacking the ability to apply and to test the key concepts of socio-cybernetics, which often become a sterile repetition of well known concepts. Or, as it happens too often in many conferences, such a lacking gives rise to Byzantine reasoning without any empirical ground. Indeed a clear sign of pathological self-reference.

This paper contributes to fill in the gap. It refers to the COMMORG case, an international research team, that was doing an European research project in the field of computer-mediated communication technologies: the organizational consequences of email introduction and diffusion. The group decided to include itself among the case studies to be investigated (Biggiero et al., 2004).

Three types of self-observation have been carried on, using the following three different methodologies of research:

1. *emailing list analysis*, which means that the number, origin and direction of email exchanged through the emailing list have been analyzed. The statistical analysis evidenced the structure, intensity and evolution of the communication patterns;
2. *genre repertoire analysis*, which means that the content of the emails exchanged through the emailing list were categorized and analyzed, so to find emergent patterns of the meanings of communications;
3. *self-survey analysis*, which means that, through a questionnaire of 34 questions submitted to all members (anonymously filled in) intra-group differences and dynamics of participation, trust, identity, and knowledge transfers have been examined.

The COMMORG group expected to repeat the analysis and to feed back its results three times into the group, so to activate the mechanism of the double loop learning, but substantial delay in other tasks of the project allowed it just for the emailing list results. Thus, the latter two methodologies have been applied just once, during last 12 months of the project.

Even taken singularly, these methodologies are very interesting cases of self-observation, especially because they have been developed in a non-hierarchical organization, like the COMMORG team. Indeed, the usual (and unaware) application of self-observation (or self-reflexivity) occurs within hierarchical organizations, like almost all profit or non-profit organizations. It is made through hierarchical practices: the top management asks some manager to analyze some department or process or unit of the organization (or the whole organization) and to prepare a report. In these cases someone (or one group) observes others. On the contrary, in the COMMORG case, COMMORG observes itself without any hierarchical command or task or position. The only form of hierarchy was concerning the role of the representative played by one – usually, the senior researcher, who was appointed by his/her institution – in each site (see next section). However, this form of hierarchical relationship and its effects on the whole system's identity and autonomy and on the self-referential processes have been explicitly taken under study by means of the third methodology (the self-survey analysis).

Besides this interesting peculiarity, the COMMORG case promises other very interesting findings because three methodologies have been applied, and their complementary and simultaneous use could highly enrich the research findings. Moreover, although it has not yet been done, the data would allow to explore these issues also through the emailing list and the genre repertoire analysis. Indeed, a deep triangulation analysis with all the three methodologies is still lacking, but present in the research agenda.

This paper summarizes the most important results that can be grasped at the first sight, because most data have become available only very recently. A better integration of these three types of method and, above all, more sophisticated statistical techniques of data analysis will give richer results in the next future. The interpretation of the findings will benefit significantly from the realization of other similar case studies. Actually, missing the comparisons with other cases, the interpretation should be rather cautious and temporary.

2. THE COMMORG CASE STUDY

Even if in some social science dates already long, the application of textual analysis to organization and management sciences is relative recent (Fiol, 1989; Yates and Orlikowski, 1992). Discourse analysis and narrative approaches are based on the idea that a social organization is a cognitive system whose members communicate through verbal, non-verbal and para-verbal messages. Communication patterns become fundamental aspects of organizational life and working, and its analysis appears crucial to understand its behavior, structure and evolutionary patterns. Although, at the beginning of this research field, communication was mainly took in its technological sense as means of communication (information systems, database mining, computational power, etc.), more recently it has been intended as media mix (especially between traditional and computer-mediated communication technologies), and finally as a semiotic text of verbal, non-verbal and para-verbal messages[24].

Understanding communication patterns helps to explain decision making processes and organizational behavior. What people decide and do is channeled through their forms and contents of communication. However, a further acknowledgement is that communication patterns are not merely the mirror of structural and behavioral aspects of organizations. Social constructivist and adaptive structuration approaches (Contractor and Seibold, 1993; Contractor, Seibold and Heller, 1996; DeSanctis and Poole 1994; Giddens, 1979, 1984; Gopal et al., 1993; Poole et al., 1991; Poole and DeSanctis, 1990; Zack and McKenney, 1995) argue that also the reverse does hold: communication patterns frame, constraint and address structural and behavioral aspects of organizations. There is a recursive causation, a feedback process, between structural, behavioral and communication aspects. Thus, the understanding of communication patterns is crucial either as the result and the mirror of organizational structure and behavior, or as one of its causes.

The COMMORG case study deals just with the attempt to investigate these recursive processes, and to underline its effects on two central issues in cybernetics and systems sciences: the formation and evolution of system identity and autonomy. For the case refers to an international research team, which lived temporarily for the time of finalizing an European research project, it is even more interesting because it has implications also for the

[24] Actually something very similar occurred also in the evolution of cybernetics and systems sciences. Socio-cybernetics (Geyer and Zouwen, 2002) and cyber-semiotics (Brier, 1995, 1997) are the most recent developments of a field of study imprinted by natural and formal sciences.

debate on the virtual organizations. By the way, it is also noteworthy that the project itself concerned the organizational consequences of email communication, including organization and group identity.

According to its own research design, and in line with the most recent and interesting methodologies, COMMORG group carried on a self-analysis, that is it decided to study itself, and to feedback the corresponding empirical results into the group itself. Indeed, the peculiarity of the COMMORG group as a case study made possible not only to do that, but also to apply some methodological tools that COMMORG has not been able to apply to the other case studies, because of the protection of the privacy concerning the content and the flows of emails. Actually, the COMMORG group is an international research team which used heavily email for communication and coordination, and as such it is a virtual team. A sub-group of COMMORG prepared the three specific methodological tools, among which a survey (called self-survey), that was submitted to all the members (that sub-group included).

The COMMORG group was made by 17 people rather stable (20% turnover) over the 30 months of length, who used 7 co-ordination (usually both scientific and management) meeting, 6 of which lasting more than two days. The COMMORG group was made by young people: 71% are less then 40 years old. They are all highly educated and remote workers: 27% of them make at home between 40 and 80% of the work. It is also a well balanced group in terms of gender.

The project was organized with a supervisor of each site –who was also the corresponding representative- and a work package leader. A work package identifies a set of strictly interrelated activities. In a work package do work usually all the partners, but with a major involvement of some of them. The work packages were the following: project management (WP1), research design (WP2), preparation of the methodologies for the fieldwork (WP3), activities concerning fieldwork (WP4), analysis of data (WP5), policy issues and implications (WP6), self-evaluation and assessment (WP7), and finally the dissemination and exploitation (WP8).

These work packages constitute very different activities, in terms of task complexity, strategic role for the success of the whole project, and duration. Such differences should be taken into a very relevant account to explain the following research findings. The sites were corresponding to the partners, but their size varied a little bit: two of them (Surrey and Nijmegen) were a one-woman site, Manchester group was made by two people, Amsterdam by three people, Patras by two people, Milan by two (sometimes four) people, and Rome by three (sometimes four) people.

3. EMAILING LIST ANALYSIS

The statistical analysis of the COMMORG emailing list at commorg@hyperion.math.upatras.gr has been performed three times throughout the project[25], and the results have been reported to the members of the group in the meetings. Two types of analysis have been performed: the first one is focused on the threads, that is, on the total single themes-dialogues occurred in the emailing list communication, and the second on the participants. The former "explores the extent to which dialogues were conducted through the list (as opposed for example to separate announcements, unrelated to each other). In all periods reported, both the percentage of e-mail messages belonging in a thread (a theme-dialogue), and the average number of messages in a thread are quite high, indicating the dialogical role of the list" (Vasileiadu, 2003).

As confirmed also by the self-survey analysis[26] the emailing list presents a high level of interactivity, and it has been the almost exclusive communication medium. It was used intensively for all the types of communication: operative information (when and where to meet, which format use for editing documents, etc.), scientific and management discussions and opinions; documents exchanges (Boudourides and Mavrikakis, 2001; Mavrikakis, 2002; Mavrikakis, 2003). The main findings can be summarized as follows:

1. all along the project there has been a very high average of mailings per day, increasing from 2.23 to 2.36 to 2.47 during the three checked periods;
2. the number of threads containing more than one mailing was also rather high: 59%, 57% and 73% during the three periods;
3. the average number of mailings per thread changed from 3.118 to 2.756 to 5.25 in the third period;
4. the communication intensifies right before and after scientific meeting.

This latter result is particularly interesting for the debate on the substitution effect, that is, on questioning the extent to which email can replace face-to-face communication. Commorg case shows that the meetings played a crucial and non-replaceable role in (re)orienting, evaluating and

[25] The first period is from 8 months before up to 8 months after the project's start-up; the second is from month 9 to month 21; and the third is from month 22 to month 28.

[26] It should be noted that the group size is rather different in the emailing list and in the self-survey analysis, because the mailing list included some people who "collaborated externally" to the project, but who were not true project members. However, especially during the last period, the two sizes almost identified (19 in the mailing list and 17 in the self-survey).

revitalizing the work of the virtual team, even when it is an heavily email-based group.

The first three results show that the communication interaction, already high in absolute values, intensified greatly during the project. Despite this very good sign of participation to project decision making and interaction, the analysis focused on the participants shows that a process of marginalisation and polarisation of the communication between project members occurred, because:

- the emailings per communicant decreased from 38 to 36 to 24;
- the first three and first five communicants accounted respectively for 40% and 57% in the first period, 43% and 58% in the second period, and 50% and 67% in the third period.

4. GENRE REPERTOIRE ANALYSIS OF THE COMMORG EMAILING LIST

In the perspective of cyber-semiotics (Brier, 1995, 1997) and of semiotic approaches of organizations (Fiol, 1989), a genre repertoire analysis of the COMMORG emailing list has been carried on. It investigates both forms and contents of emails exchanged during the duration of the project. Merely the emails channeled through the emailing list have been analyzed, but that limitation doesn't appear too hard, because the vast majority of emails have been exchanged through the emailing list. Although there are no quantitative data to demonstrate it, this impression was shared by all the group members. Actually, this is just by itself a relevant result, demonstrating the transparency and democratic way in which the team worked. Indeed, the sending of emails to single members (or few informal sub-groups) and the frequent use of BCC would indicate an anti-democratic use of email communication, also favoring the formation of clans and coalitions. The so-called "petty tyranny" (Romm and Pliskin, 1999) represents the opposite of the common place according to which the introduction of the email communication into organizations determines a democratizing evolution.

Genre repertoire analysis of communication patterns in organizations (Orlikowsky and Yates, 1994; Yates and Orlikowski, 1992) studies the forms and contents of verbal communication, and focusing on recursive processes it tries to evidence emergent patterns. This interactive process of structuration leads to the creation of a common set of shared meanings in which the uniqueness of each organization or community resides (Brier, 2000). "The contextual embeddedness of social shared meanings implies that each sign, each metaphor, each socially constructed practice fully deploys its significance only within the context in which it is generated. In cybernetic

terms, this is a reinforcement mechanism of the self-reference of social systems and may have relevant impacts on repertoire's structuration" (Muzzi and Dandi, 2004).

Two different genre repertoire analysis have been done by COMMORG members: Vasileiadu, Mavrikakis and Kalamaras (2003) made an extensive analysis of 2412 emails exchanged during three years, focused on the form and purpose of messages, while Muzzi and Dandi (2004) considered only 583 emails exchanged during early eight months of the project. However they studied also the relationships between the form and purposes of messages and the complexity of the tasks they were referring to. Both analysis used advanced statistical techniques. They also coded messages in a slightly different way, and this has implications for the consequent analysis.

Vasileiadu et al (2003) list the following characteristics: attachments, CC to others, embedded message, emotions, forwarded message, general opening, humour, informal/colloquial, personal addressing, reply subject, decision related, distributing responsibility, FYI (For your Information), proposal/suggestion, question, reminder, reporting work done, response, scientific argumentation, sending comments, socialising.

Using factor analysis six genres have been identified: 1. dialogue (form: reply subject and embedded message; purpose: response) 2. sending work (form: attachment; purpose: sending work) 3. decisions (no form; purpose: scientific argumentation, proposal, decision-related and reporting work done) 4. reminder (no form; purpose: reminder and distributing responsibility) 5. socialising (form: informal language and humor; purpose: socialising) 6. FYI (form: no question; purpose: FYI).

Their main results are the following:

a) the team is highly dialogical. That means that most issues are discussed within the COMMORG list, and most issues followed by the team. This relates to the high degree of interactivity of the list, presented in the other sections of this self-observation analysis.
b) Most of the completed work is sent in attachments, instead of using e.g. the COMMORG website, fax or any other medium.
c) There are decisions made within the list, which actually respond to one of our original research questions, whether email is mainly a coordination medium, or whether decisions are also made through it. The other characteristics related to decisions are reporting work done, scientific argumentation and proposals/suggestions. This means that suggestions in the team, as well as scientific argumentation and reports on work performed so far are followed up by a discussion and then a decision making process.
d) The usage of the list is not only for working-related messages, but for emails of social character as well. The team is using email as a means to

build up a common identity, and trust, to socialise and build up a team-spirit. This is a distinctive characteristic of this emailing list, not to be found in other project-related emailing lists, or in relevant studies.

It becomes clear that using genre analysis we can identify not only communication patterns in a team, but also, and mainly working patterns, and ways in which a working group functions. We have seen with the above analysis that the group is very dialogical, and also democratic, using the e-mailing list to make decisions (which means that they are not limited to face-to-face decisions only). If issues emerge, decisions about suggestions and work are made via the e-mailing list. Finally, the group is not limited to a working-group only, but also socialises, building up a team spirit through the emailing list" (Vasileiadu, et al., 2003).

Crossing the 14 tasks[27] with messages coded by purpose and form, Muzzi and Dandi (2004) found five major genre, which they place into four classes of genres they pre-defined. These pre-defined genres are:

1. *mechanic genres*, which are the genres of communication associated to bureaucratic and hierarchical coordination, where standardization of tasks and command-and-control are prevalent;
2. *task-oriented genres*, which have the same purpose of the previous ones but present less formalization. Actors here tend to use command-and-control forms of coordination (information requests for example) but prefer to use an informal style of communication;
3. *organic genres*, which are associated to participatory mechanisms of coordination. Interaction is no more based on asymmetry but occur at the same level of authority. Formalization is low as tasks require easier and faster communication. These genres are associated to complex tasks;
4. *formal participation genres*, which are as participatory as the previous ones but more formalized. This means that tasks are highly interdependent but some degree of standardization of the style is present" (Muzzi and Dandi, 2004).

In terms of purpose items, the messages are coded as following: question, answer, ballots, management, disagreement, internal report, meta-communication, technical support, proposal; while in terms of form items the messages are coded as follows: closing formula, emoticons, reply with embedded message, lack of text structure, openings and greetings, signature.

Through cluster analysis Muzzi and Dandi have been able to identify the following five genres of email communication in COMMORG emailing list:

[27] The activities of the COMMORG group were structured into 8 work packages, articulated into tasks. The sum of all tasks is 14.

Genre 1, [which] represents the expression of disagreement on a given issue. This is usually a reply message that contains the original message. No indications on the degree of formalization. It can be ascribed to the class of organic genres as disagreement is a form of participation. The task where disagreement was expressed most was 13 (brochure preparation): more than one third of messages belonging to this genre concern the brochure preparation. This task was considered highly critical, as it concerned the presentation of the project to external organisations (especially those the project wanted to analyse). In order to be effective, the research team concentrated a lot of efforts and interactions in a task that otherwise and elsewhere would be considered trivial. Some disagreement therefore was likely to occur. Another task with some presence of disagreement is task 2, that is project management (almost one third of the messages belonging to Genre 1), where the expression of different points of view is, however, self-explaining: budget issues and decisions concerning the division of labour are often potential sources of conflicts.

Genre 2, [which] is related to technical and meta-communication issues. These messages are very informal (no opening or closing greetings) and embed the original message, so that we could define them as answer to technical questions. They are composed of explications on how to use media of communication (web-site, blackboard) and suggestions/thoughts about how to behave through these media (meta-communication). It is worthy to remember that COMMORG is a research project aimed to studying computer-mediated communication: the team is encouraged to practice self-observation and this study is part of this strategy. The half of messages in this genre occurred for communicating information concerning task 2 (project management), while 13% of them concerns task 13 (brochure preparation). These tasks have been coded as complex because they required high levels of interaction. Meta-communication can be viewed as a means for reducing ambiguity and uncertainty as it frames the situations (Weick, 1979).

Genre 3, [which] is composed of the messages aimed to sponsoring interaction and participation: ballots and proposals. We don't have information about the form of these messages, so we can say that they are close to the mean of the other messages under this issue. One fourth of the messages belonging to Genre 3 occurred for communicating information regarding task 2 (project management), while about 16% of them concerns messages regarding task 3 (literature review and key issues), 15% regarding task 12 (dissemination/exploitation activities) and 16% regarding task 13 (brochure preparation).

Genre 4, [which] represents quick operational questions, without formality and aimed to obtaining fast exchange of information. We can

ascribe these genres to the class of task-oriented genres. Genre 4 occurred mostly in messages concerning again task 2 (project management, 20% of messages in Genre 4), task 5 (sampling level one, 15%) and task 6 (sampling level two, 15%).

Genre 5, [which] concerns messages aimed at directing people (management) or reporting to others. Formality is obtained through including opening and closing greetings. However this genre is associated with lack of structure, therefore this is a moderately strong form of mechanic genre. Almost 30% of messages belonging to this genre concerns task 3 (literature review and key issues), while task 6 (sampling level one) and task 2 (project management) are also highly represented in this genre (about 16% of messages in Genre 5, each). This can be explained by the fact that task 3 is the perfect example of pooled interdependence (members wrote separately their literature reviews by topic), while task 6 has a high presence of report messages (people reported to the emailing list the improvements in approaching multinational companies, in order to gain access to the same organisation throughout all the four countries involved in the project).

To sum up these results we matched each genre to each category of our classification (Tab. 1):

Table 1. Classification of the COMMORG genres.

		Form	
		Low formalization	High formalization
Purpose	Command and control	2. Task-oriented genres: Genre 4	1. Mechanic genres: Genre 5
	Participatory interaction	3. Organic genres: Genre1, 2 and Genre 3	4. Formal participation genres:

It is indeed interesting to look at the evolutionary side of the communication patterns, because the ix of these five genres changes over time. Genre 1 almost disappears, and the organic genre prevails during early months of the project. These changes seem to reflect an evolutionary model of the intra-group communication: at the beginning, people don't know each other, and so they explore and try to adapt to each other, searching for common views and codes. That adaptation process implies also a high share of disagreement. However, over time, to the extent that the tasks should be finalized and accomplished, some agreements and common views should be reached, and so the communication becomes more standardized, and disagreements – at least, the expression of disagreements – sharply decrease. Accordingly, genre repertoire mutates.

5. THE SELF-SURVEY ANALYSIS

The survey investigates members' thinking and perceptions concerning all the sphere of the group working life and personal relationships. Actually, the self-evaluation of project performance and the analysis of knowledge sharing and transfer and creation between different disciplines were some of the specific innovation brought into this research project. Related to this analysis was also the attempt to use its results to foster and speed learning processes of COMMORG group, according to the logic of double loop learning, but that goal has not been accomplished, because of delays in other project's tasks. The survey was divided into the following seven sections: A) project members background, B) task analysis, C) degree of autonomy, D) types of communication media, E) knowledge issues, F) identification and trust, G) project performance and value.

We started the analysis trying to find, through the cluster analysis, some specific typology in COMMORG group. The idea was to get from eventual typology some indications for further investigations. However, considering all the high number of questions, a first attempt of cluster analysis did not evidenced any typology among COMMORG group members. It means that, in general, the group is rather homogeneous. Therefore we carried on a simple basic analysis based on frequency tables.

1. Most members (87%) felt mostly involved in the activities concerning data gathering in fieldwork, data analysis, and finally dissemination and exploitation. The lowest degree of involvement was recorded into research design and the setting up of different methodologies for the fieldwork. This result could be explained by the fact that, being at the upstream of the project, the latter two activities (and especially the research design) were considered very crucial and so it was coped mostly by project leaders.
2. The degree of autonomy from the site supervisor was much higher than from the work package leader, especially for dissemination and exploitation. The lowest values are for the project management, where actually had to be made decisions just by the site supervisors. The lowest degree of autonomy from work package leader was recorded in research design, because of its crucial role.
3. COMMORG group was made by heavy computer-mediated communication users, especially through the emailing list. It was concerning both simple communication or scientific discussions or even retrieval and allocation of information and knowledge. On the contrary, within each site the communication was mostly face-to-face and secondary by direct email. It is interesting noting that almost no any other means (like fax, telephone, etc.) was used.

4. There was a significant difference in media use between the work packages. The communication by mailing was never low, but it was: extremely high in more complex work packages: half members used it between 40 and 60% of all communications in the project management activities, between 65 and 95% in the research design, between 65 and 100% in fieldwork, and between 42 and 52% in the research design. It was lower in less complex work packages (evaluation and assessment, and dissemination and exploitation) or in less interactive (among sites) work packages (the setting up of different methodologies for the fieldwork, and the policy issues and implications). The communication face-to-face was almost always low, medium-low or well distributed. An anomaly was in the research design activities, where face-to-face was extremely low (75% of members used it less than 33%). This result is due to the fact that a large part of that work package was done for the literature review, which was a rather individualistic task. In general, these results confirmed that: 1) computer-mediated communication (and specifically the emailing list) was the most important medium used by COMMORG; 2) task complexity didn't increase significantly face-to-face.
5. The group was not very open to exchange the scientific issues of COMMORG with non-members experts, while a certain degree of openness was occurring when concerning the sharing of personal experience of the project with non-expert people, supposedly friends.
6. Most members argue that they increased their practical knowledge, especially in the project management and data analysis, while less in the research design and in the setting up of different methodologies for the fieldwork. Likely this corresponds to the minor degree of involvement in these work packages (see point 1). Because of different reasons, the increase of practical knowledge was less significant and less diffused among members in reference to the fieldwork, and the dissemination and exploitation. Likely it is due to the fact that fieldwork and dissemination activities are more commonly experienced activities of researchers, even when they are young.
7. In terms of theoretical learning, COMMORG members experienced a less remarkable and less diffused increase than for practical knowledge. Some higher values are recorded in the field of trust/substitution, and organizational theory and behavior. Likely this result is due to two reasons: 1) the most critical phase –the research design- which was also at higher scientific-content degree, was involving less people (see point 1), and 2) due to time shortage almost no room has been devoted to scientific discussion both after completing literature review and after research findings.

8. Most members perceived a medium-high impact of project work, especially in terms of the ability to team working. This is true also for technical learning, which otherwise did not improved.
9. Most people (80%) perceived the group as open to knowledge sharing, and with seldom conflicts about that. The majority (60%) perceived the group very open to knowledge sharing and medium-high degree of frankness (available to eye-to-eye solutions when conflicts emerged). While conflicts about scientific issues occurred seldom, the ones concerning management issues happened very often, according to the perception of most members (73%). Half group perceived that these last problems implied difficult solutions.
10. Almost no one felt any information overload, and most people consider that the precision, the timeliness and the usefulness of information were appropriate and satisfying (medium-high).
11. The identification with the whole COMMORG group-project was rather high for 60% of members, but the percentage raises slightly when considering the identification at the site level. As concerning trust at site level, it was perceived very high by at least 60% of members, while at consortium (whole group) level 53% of members considers the others only moderately trustworthy and/or getting the right job without monitoring and/or for job support if missing their own contribution. These percentage raises considerably (67% with medium-high score) when considering to be confident in the other group members' willingness to make decisions taking into account also their own interests and points of view and when considering (80% with medium-high score) others' integrity. However there have been two members clearly non-identified with the group and who definitely didn't trust all the others. These two "outlayers" gave negative scores to all these sets of questions.
12. The three quarters of people argue that the project fully achieved all the goals committed with EU, but that share reduces slightly to 60% when considering the project highly success in general terms –that is, not solely from the point of view of EU- and when considering their own satisfaction. This result confirms the widely opinion that the administrative rules and constraints of EU Fifth Framework Programme slightly frustrate the needs and the degree of freedom necessary for a good scientific research. Approximately the same number of people consider that they have personally gained from the project, that the project promoted them professionally, and that they would like to repeat the same type of experience. Indeed, this share is not higher because of the above mentioned sense of frustration due to the then current administrative rules. In the same light should be interpreted their moderate and medium-low perception of respectively cost- and time-

efficiency of the project. Most people believe that the administrative rules not solely damaged the scientific effectiveness, but also resulted in lower efficiency.

6. CONCLUSIONS

COMMORG self-observation shows that an emailing list can guarantee an effective communication of a virtual complex organization, as an international research group is. However that outcome is possible only if it is punctuated and grounded in periodic meetings, dedicated to solve the main intervening problems and to agree about the planning of future activities. On the other hand it seems that these meetings are crucial events, during which problems can be solved but also created. The phases between two adjacent meetings are characterized by almost solely email communication, which can both increase or reduce conflicts eventually raised during the meetings, but never solve them when they are considered sensitive. In this sense it is rather interesting that the members of COMMORG group, which was a scientific research team, were considering sensitive above all management and budget problems.

COMMORG self-case shows also that:

1. the most participating – in terms of the amount of emails- members are not necessarily also the most collaborative, and neither the most face-to-face participating;
2. the group was very interactive, dialogical and democratic, dealing with all scientific, management and inter-personal issues;
3. although a virtual team, and although belonging to different scientific communities, it developed a rather high group identity;
4. it enhanced significantly members' practical knowledge and ability to team working.

As concerning the specific debate in sociocybernetics, from which we started in this paper, we can argue that:

a) self-organization, self-observation and self-reference can be measured and operationalized for the empirical research;
b) temporary and virtual organizations can develop a significantly high degree of identification and trust;
c) both these properties can be operationalized and measured;
d) the system identity, the communication patterns, and the semiotic patterns are recursively related to structural aspects, i.e. the relevance and complexity of the activities to be performed;

e) the system identity, the communication patterns, and the semiotic patterns change over time;
f) rather interestingly, the autonomy (closure) of the system was higher under the scientific aspects of the working life then the management aspects, even if they were less sensitive. Indeed they were more salient and distinctive;
g) within the system the autonomy of its sub-systems depends heavily on the social and structural characteristics of the corresponding activities.

REFERENCES

Biggiero, L., 1997, Managerial action and observation: A view of relational complexity, *Sistemica* **12**:23-37.

Biggiero, L., 1999, Markets, hierarchies, networks, districts: a cybernetic approach, *Human Systems Management* **18**:71-86.

Biggiero, L., 2001a, Sources of complexity, in human systems, *Nonlinear Dynamics, Psychology, and Life Sciences* **5**(1):3-19.

Biggiero, L., 2001b, Self-organizing processes in building entrepreneurial networks: a theoretical and empirical investigation, *Human Systems Management* **20**:209-222.

Biggiero, L., 2001c, Are firms autopoietic systems?, in: *Sociocybernetics: Complexity, Autopoiesis, and Observation of Social Systems*, F. Geyer and J. Van der Zouwen, eds., Greenwood Press, Westport, pp. 125-140.

Biggiero, L., Sammarra, A., Muzzi C., and Dandi, R., 2004, Organizational consequences of E-mail adoption and diffusion: theoretical issues and empirical results, in: *eAdoption and the Knowledge Economy: Issues, Applications, Case Studies*, P. Cunningham and M. Cunningham, eds., IOS Press, Amsterdam, pp. 134-140.

Boudourides, M., and Mavrikakis, M., 2001, *First Internal Report on Mailing list Analysis*, COMMORG internal report.

Brier, S., 1995, Cyber-semiotics: on autopoiesis, code-duality and sign games in bio-semiotics, *Cybernetics & Human Knowing* **3**:3-15.

Brier, S., 1997, Cyber-semiotics, *Systems Research.*

Brier, S., 2000, Trans-scientific frameworks of knowing: complementarity views of the different types of human knowledge, *Yearbook Edition of Systems Research & Behavioral Science* **17**:433-458.

Contractor, N. S., and Seibold, D. R., 1993, Theoretical frameworks for the study of structuring processes in group decision support systems: Adaptive structuration and self-organizing systems theory, *Human Communication Research* **19**:528-563.

Contractor, N. S., Seibold, D. R., and Heller, M. A., 1996, Interactional influence in the structuring of media use in groups: Influence of members' perceptions of group decision support system use, *Human Communication Research* **23**:451-481.

DeSanctis, G., and Scott Poole, M., 1994, Capturing the complexity in advanced technology use: Adaptive structuration theory, *Organization Science* **5**:121-147.

Fiol C. M. 1989. A Semiotic Analysis of Corporate Language: Organizational Boundaries and Joint Venturing. *Administrative Science Quarterly* **34**:277-303

Foerster H. 1982. *Observing Systems*, Intersystems, Seaside.

Geyer, F., 2002, The march of self-reference, *Kybernetes* **31**(7/8):1021-1042.

Geyer, F., and van der Zouwen, J., eds., 1990, *Self-Referencing in Social Systems*, Intersystems Publications, Salinas (CA).

Giddens, A., 1979, *Central Problems in Social Theory*, University of California Press, Berkeley.

Giddens, A., 1984, *The Constitution of Society*, University of California Press, Berkeley.

Gopal, A, Bostrom, R. P., and Chin, W. W., 1992-1993, Applying adaptive structuration theory to investigate the process of group support systems use, *Journal of Management Information Systems* **9**:45-69.

Leydesdorff, L., 2001, *A Sociological Theory of Communication: The Self-organization of the Knowledge-based Society*, Universal Publishers.

Luhmann, N., 1990, *Essays on Self-reference*, Columbia UP, NY.

Mavrikakis, M., 2002, *Second Internal Report on Mailing list Analysis*, COMMORG internal report.

Mavrikakis, M., 2003, *Third Internal Report on Mailing list Analysis*, COMMORG internal report.

Muzzi, C., and R. Dandi, R., 2004, *Self-organizing of Genre Repertoire in a Virtual Research Team*, Paper presented at ISA World Congress.

Nonaka, I., and Takeuchi, H., 1995, *The Knowledge-creating Company*, Oxford UP, Oxford.

Orlikowski, W., and Yates, J., 1994, Genre repertoire: Structuring of communicative practices in organizations, *Administrative Science Quarterly* **39**:541-574.

Poole, M. S., and DeSanctis, G., 1990, Understanding the Use of Group Decision Support Systems: The Theory of Adaptive Structuration, in: *Organizations and Communication Technology*, J. Fulk and C. Steinfield, eds., Sage, Newbury Park, pp. 173-193.

Poole, M. S., Holmes, M., and De Sanctis, G., 1991, Conflict management in a computer-supported meeting environment, *Management Science* **37**:926-953.

Romm, C. T., and Pliskin, N., 1999, The office tyrant: social control through E-mail, *Information Technology & People* **12**:27-43.

Sammarra, A., and Biggiero, L. 2001, Identity and identification in industrial districts, *Journal of Management and Governance* **5**(1):61-82

Vasileiadu, E., 2003, *Emailing List Analysis*, Internal report of the COMMORG Project.

Vasileiadu, E., Mavrikakis, M., and Kalamaras, D., 2003, *Genre Analysis of the COMMORG mailing list*, Internal report of the COMMORG Project.

Weick, K. E., 1979, Cognitive processes in organizations, in: *Research in Organizational Behavior*, B. M. Staw, ed., JAY Press, Greenwich (Connecticut).

Yates, J., and Orlikowski, W., 1992, Genres of organisational communication: A structurational approach to studying communication and media, *Academy of Management Review* **17**(2):299-326.

Zack, M. H., and McKenney, J. L., 1995, Social context and interaction in ongoing computer-supported management groups, *Organization Science* **6**:394-422.

SOME COMMENTS ON DEMOCRACY AND MANIPULATING CONSENT IN WESTERN POST-DEMOCRATIC SOCIETIES

Gianfranco Minati
Italian Systems Society, president, Via P. Rossi, 42 - 20161 - Milano, Italy, www.airs.it, Tel./Fax: +39-2-66202417, E-mail: gianfranco.minati@AIRS.it http://www.geocities.com/lminati/gminati/index.html

Abstract: In the history of western democracies, now degenerating into post-democracies, it is possible to identify a first phase during which aspirant leadership has been attempting to influence and involve people, by forcing masses to do something, to believe something, rather than by getting consent. The second phase relates to democracies where aspirant leadership must get consent thorough formal elections. The mass dimension is not related anymore to involvement, but to getting formal consent. Manipulating social techniques, based on sophisticated research in cognitive science and applied by using the mass-media, have been and are used for marketing exploiting knowledge of complex human behavior in order to turn individuals into customers. Similar technologies are used for to influence people to buy a political offer and leadership. The mass dimension is not anymore a warranty of democracy, but rather the basis for applying marketing techniques that make consent buyable. Democratic societies became degenerated post-democratic societies. The most significant aspect of such manipulating techniques is the manipulation and control of language used by applying approaches based on cognitive science. Some of those approaches are introduced. The purpose of this contribution is to focus on how the systems community may make people aware of the manipulating processes and able to recognize them, with special reference to language. The possibility to make consent buyable may be the end of the classic idea of democracy and this must be taken in count when dealing with so-called emergent social systems, often assumed to be non-democratic by western societies.

Key words: consent; democracy; involvement; language; manipulation.

1. INTRODUCTION

The concept of democracy is a very complex one. We introduce some reflections on the situation of western societies where, thanks to sophisticated marketing techniques and usage of mass media, the mass dimension, *warranty* of classic democracies, is now the entry point for manipulation. This is also very important for dealing in the next future with emergent and economically very important non-democratic (in western meaning) societies, like China. In the past the idea was to *export* democracy by considering that equivalent to free people, to design freedom for everyone. Now, which is the expected reaction of *post-democratic* societies, based on consent manipulation, interacting with *non-democratic* societies?

2. FROM INFLUENCING TO GETTING CONSENT

There is a difference between

- *influencing* the behavior of masses and
- *manipulating* to get consent.

The two strategies are not always clearly distinct and may be applied in different ways depending on the context and on the specific objective.

Examples of the first one are given by an unfortunate usage of religions, diffusing manipulated information, rallies, and manipulating economical situations. The purpose is not to get consent (formal consent is not necessary), but to involve, forcing masses to do something, to believe something, like it took and take place in dictatorships.

Mass behavior may be influenced by usage of *symbols* like flags, hymns, uniforms, habits, ways of dressing in general making evident social roles, dietetic habitudes like usage of alcohol and drugs, pornography, and so on. Mass behavior is induced as amplifications or replications of specific manifestations like parades and social events.

Events have the purpose to produce effects at the mass level. Individuals participate, perform roles that have been not singularly discussed, but assumed as standard, usual behavior.

Influencing has the purpose to make difficult to do not adapt, do not assume the behavior assumed as standard. Who does not accept the standard is considered deviated, psychologically or mentally ill (like dissidents in Stalin times) because it was not acceptable to discuss or to refuse the standard behavior. Better, acceptance is not matter of convincement, a cultural or political decision, so it is not appropriated to discuss it.

Leaders or dictators use a kind of engineering, an architecture of influencing by combining several available social factors like Hitler for instance did with the racial and nationalistic arguments.

It is matter of involvement, rather than of convincement.

Human attitude is then such that it is possible to *make rational*, to self-convince that an involvement is also a convincement: in this process there is production of symbols, values, standards able to (self-) confirm and reproduce the involvement.

Subject of this contribution is the second case, when formal consent is necessary, like in democratic societies.

The very basic idea at the core of the concept of democracy is that the need to get consent with free voting, having each participants the same weight, having each member of the society the right to vote, requires participation of the total population.

The mass dimension of democracy was the warranty of the impossibility of establishment of leaderships based on nobility, money, and power. The leadership is elected by populations.

The current situation in western societies is such that the stronger aspect of democracies, that's their need of popular voting base, became the entry point for processes of *consent manipulation*. This took place because of the availability of many sophisticated techniques carried out specially for marketing reasons. They are also based on the availability of technologies allowing diffusion of information tailored for a specific purpose and able to multi-dimensionally (reference is to different social roles, as discussed later) reach *each* component of the social system. They were originally used for advertising products and services. It's usually accepted to have advertises inserted in information diffused by using any technologies, misunderstanding it as an informational service.

Information are tailored in such a way that they are homogeneous with advertising: *information look as advertising* and *advertising look like information*.

We are just making reference to the formal, syntactical aspects and not yet to contents.

One important step is based on not diffusing or making easily available data, original texts (laws and rules, for instance), but to make available only comments, extrapolations, judgments and estimations. In this way people are asked to take non-influential positions, or to select between pre-established choices depending on third part interpretations. This is a good way to make people non autonomous in judging.

This strategy is combined with focusing on details, on events such as crime news, and details that people may directly comment, reinforcing the sense of freedom.

Political, economic, and social strategies are never subjects of public discussion: they have not to be shared. Public discussions regard *comments*.

Getting consent through manipulation is the new way to involve.

3. MANIPULATING CONSENT

As we said, the new strategy to manage and control social systems is not based anymore on involvement, rather on manipulating by using sophisticated technologies used to diffuse information.

Anyway the core of the new manipulating processes is not constituted by technologies: they are the tools able to apply, amplify, and integrate manipulating techniques in social life. The crucial point may be realized by considering the approaches of cognitive science. It refers to the ability to influence the cognitive models, the cognitive processing carried out by agents. In this way *the emphasis is not anymore on the information, but on the way by which the information is proposed, on the cognitive processing induced*: in short, on software rather than on data. Information are not anymore used for convincing, but for induce behaviors like to buy, select, and give consent.

In the traditional view, the same cognitive system may adopt different models because of cultural, political or religious reasons. In the new approach, *the strategy is to force the cognitive system to adopt such a way of processing information that the results do not depend on the models adopted.*

The purpose is to generate the *illusion of freedom, reduced to the possibility to select* and to judge only among a closed set of pre-arranged possibilities. In this way social systems are in the trap of the *one-dimensional* freedom (Marcuse, 1964).

This reminds the difference between *First-Order* and *Second-Order Cybernetics*. The *First-Order* Cybernetics, in short Cybernetics, has been introduced by Norbert Wiener (1894-1964) with the goal of studying the processes of control and communication in animals and machines (Wiener 1948; 1961). It was identified with the information theory (Ashby 1956; Heims 1991). Generally in this approach, named *First-Order Cybernetics*, feedback *single* loops are based on using output variables to control, regulate the process.

Heinz von Foerster (1911-2002) with scientists like Warren McCulloch, Norbert Wiener and John von Neumann developed the so-called *Second-Order Cybernetics*, or Cybernetics of Cybernetics (Von Foerster, 1974, 1979), focusing on self-referential systems for the explanation of complex phenomena.

This approach may be represented by using feedback double loops based on using output variables not only to control and regulate the process, but also for redesigning the process itself in front of the objectives of the observer-designer.

The fundamental difference between the two approaches is that in the first case it is assumed that the process is based on applying rules (even having learning, adapting capabilities), while in the second new rules may be invented, being the focus not on maintaining stable a process, but the purposes of the observer-designer.

The situation taking place in western societies is the reverse: from systems designed to try to carry out the purposes of the observer-designer (democratic societies) to systems designed to maintain, regulate and control pre-established processes (post-democratic societies).

By dynamically using different cognitive models human agents are able to make emergent multiple systems. Multidimensional social systems, corresponding to multiple systems, have been introduced as Collective Beings (complex, multidimensional societies). Interacting agents exhibiting emergent collective behavior, simultaneously belong to different systems and share, in time, the same cognitive models (Minati 2001). Collective Beings are multiple systems constituted by agents playing same role in different times and different roles at the same time, like parents, buyers, workers, players, and soldiers constituting families, markets, corporations, and armies, behaving always *simultaneously* belonging to all those systems.

Those multidimensional systems are now forced to become a single, one-dimensional system, that's passing from dynamically sharing the same models in time, to sharing the same model.

In *post-democratic* societies, that's *formal* democratic societies, human agents are expected to make emergent different multiple systems by using the same model in time, that's by using the same way of processing information, the same standardized approach and language.

In one-dimensional societies, cognitive processing of information (Lindsay and Norman, 1977) is regressed and limited:

- Agents process certain information in a very limited way, by assuming, for instance, that only comments are possible and that decisions, actions are not possible. That's combining the lack of suitable institutions and organizations with the convincement that this role should be for somebody else, expert and supposed to officially take care of the problem. Examples are given by ecological problem, or by accepting to don't know official documents (formally available), programs and rules, but only knowing comments or journalistic reductions.

- Agents consider only information suitable to be processed in standardized way, producing some possible outputs, such as to take consumer decisions.
- Agents process information and problems in standardized ways only, like looking for simplification, repetition, optimising, and solving rather than to be open to consider new approaches, complexity and non-linearity. We may say that the standardized way of processing is algorithmic, by combining standardized approaches.

In literature there have been introduced terms like cognition, cognitive systems, and cognitive model now disciplinary scientific fields.
"Cognition is a biological phenomenon and can only be understood as such; any epistemological insight in the domain of knowledge requires this understanding." (Maturana, 1970, p. 3).
"Living systems are cognitive systems, and living as a process is a process of cognition." (Maturana and Varela, 1980, p. 13).
In this view "A cognitive system is a system whose organization defines a domain of interactions in which it can act with relevance to the maintenance of itself, and the process of cognition is the actual (inductive) acting or behaving in this domain." (Maturana and Varela, 1980, p. 13).

A Cognitive System is intended as a system constituted of interactions among activities such as the ones related to attention, perception, language, affective-emotional sphere, memory, and the inferential system, as well expressed in (Pessa, 2000).

Cognitive models refer to models of some specific processes of the general cognitive activity, intended not separated but interacting. By considering, in a reductionistic way, cognition equivalent to computation, a cognitive model may be intended as a computer program even if based on non symbolic approach like using neural networks. In this approach a cognitive model is an abstract device that simulates user behavior (Howes, 1995).

By assuming, as in literature, but first of all for exemplificative purposes, equivalence between Cognition and Computation a cognitive system may be assumed as a system of models interacting in a cognitive architecture (Anderson J. R., 1983).

Agents provided with artificial cognitive system are based on computational modelling of cognitive processes. This approach is peculiar for studies in Artificial Life (Rodney, 1992; Hemerlijk, 2001) and in robotics (Cao et al., 1977; Cliff et al., 1993). Another example is the study of emergence of computational ecologies (Huberman and Hogg, 1993).

In my view, and on the basis of research mentioned later, the key point to reach such a standardized cognitive modelling is primarily based on the control of language.

4. USAGE OF LANGUAGE TO MANIPULATE

Language is a very difficult topic. It has and is studied from many different disciplinary approaches by considering, for instance just to give an idea of the complexity, formal, context-dependent, context free, and natural languages as well syntax and semantics.

The study of natural languages is approached by interdisciplinary research like in Psycholinguistic, Neurolinguistic (Bouton, 1991), Sociolinguistic (Cicourel, 1973), and Linguistic.

By assuming equivalence between Cognition and Computation the influence of the programming language on software designing is very strong. The crucial point is not on the *formal* ability of a language to represent something (*translation* is the non-algorithmic process able to reproduce same meaning in different languages), but on designing, that's on the inducing and influencing what may be expressed. This is much more than *suitability*. When it is said that "the tools we use have a profound influence on our thinking habits and therefore on our thinking abilities" (Dijkstra, 1975), the reference is to the meaning of the *Sapir-Whorf hypothesis*, introduced later. In this context the interest for considering language is related to the designing of cognitive models.

Formal computing programming languages are designed to deal with different specific problems. We do not refer to formal differences and different power in representing, but to different power to induce usages and understanding.

Cognitive modelling and language processing are activities of the cognitive system

The interest in this contribution is to focus on usage of natural language for the carrying out (or *emergence*) of cognitive models.

A very effective approach to influence the processing of information is to control the language available to represent, to design, to think.

We make reference to old, controversial approaches such as the one expressed in the Sapir-Whorf hypothesis. Specifically, in 1929, the American linguist and anthropologist Edward Sapir (1884-1936) wrote:

"*Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression for their society. It is quite an illusion to imagine that one adjusts*

to reality essentially without the use of language and that language is merely an incidental means of solving specific problems of communication and reflection. The fact of the matter is that the 'real world' is to a large extent unconsciously built up on the language habits of the group."

Sapir, in collaboration with his pupil, Benjamin Lee Whorf (1897-1941), developed the so-called *Sapir-Whorf hypothesis* introducing linguistic determinism (Caroll, 1956). We just mention that Bertalanffy dedicated an entire session of chapter 10 of his fundamental book *General System Theory: Foundations, Development, Applications* (Bertalanffy, 1968) to this hypothesis. Modern science, specially cognitive science, accepts the weaker version of the theory by retaining that process of thinking is influenced by the linguistic systems available and used. The strong version, named linguistic determinism, leads to consider that people having a sophisticated language may think more and better than others having a less sophisticated one. This could be a way to support kinds of linguistic and cultural racism.

Language and linguistics, relations between language and human behavior, linguistics and cognitive science, are subjects of intense interdisciplinary research (Bickerton, 1992; 1996; Chomsky, 1986; Deacon, 1997; Givon, 1998; Muller, 1996; Pinker and Bloom, 1990; Wilkins and Wakefield, 1995) focusing on many aspects related, for instance, to human interactions and communication, general semantics theory, knowledge representation, linkage of human processes of thinking to human behavior, linkage of process of thinking to language, man-machine interfacing, representing meaning through language, linguistic construction of reality (Grace, 1987), translation, and so on.

Representing is a crucial issue in modern science. Focusing on representing something with something else allows for fast computing processing, for modelling and simulating. Knowledge representation is a subject interconnected to the concerns of Artificial Intelligence and Cognitive Science.

The introduction of the Sapir-Whorf hypothesis generated many fruitful disciplinary scientific researches. *Interest in this contribution is not at all in applying it to human behavior control, rather in understanding and realizing how human behavior is controlled in such a way.* Our real freedom may be intended as related to the ability to *design the future* (see, e.g., Berger and Luckmann, 1966) and not reduced to the possibility to select among pre-established choices.

How in social system natural language may be controlled, that's reducing its power of representing making easier to representing something rather than something else?

A first step is to reduce the vocabulary used in the daily spoken language, to induce usage of an inaccurate syntax. *Influencing the syntax by which*

elements are composed influences the emerging system. An appropriated and accurate usage of syntax allows to properly describe, represent, and communicate. How to get this regress? For instance by emulation, by diffusing poor usage of language in advertising, songs, movies and popular publications. Advertising introduced standardized usages of language. Popular songs also introduced standardization. They both introduce expressions usually repeated, used by people in communicating and, inevitably, in thinking. By using those reduced, stereotyped, standardized subsets of a natural language is possible to only represent, to think something specific and not something else, that's making agents to think in similar way.

It has been noted that young people approximate, use grammatically incorrect expressions, syntactic reductions, make poor usage of verbs, emulating language used for advertising, songs, and popular TV programs.

In Italy we are experiencing in the daily spoken language that the number of words used is reduced and simplified.

How this process has been established? In the past the problem was both to reduce ignorance and to have unified, national languages instead of dialects. The purpose of schools and literary production was to educate people to read, write, and speak more sophisticated, and appropriated real, complex languages by avoiding dialectical reductions and approximations. In this new age based on controlling and reducing usage of cognitive systems, the situation is reversed. The controlling, simplifying, reducing and limiting usage of cognitive systems take place with a contextually simplified, reduced and limited usage of natural language. Official language is not anymore the written or spoken language. It's the language of mass media, especially television. The mass media purpose is to have the larger possible audience, reaching the maximum number of people with advertising messages and then taking advantage for getting advertisers. The purpose of reaching the larger audience as possible is not due to following the original missions (still formally valid) having democratic goals, like to diffuse information, knowledge, educational activities, artistic productions, and politic issues. *The attempt of reaching the larger audience as possible is related to business purposes. The target is not users, but potential buyers, the market.*

What is diffused is constituted by details, *marketing suggestions* about products and services, events and not information to be processed by a *complex software* (that's complex cognitive models).

The daily spoken language is more and more the language made official by mass media. *At this regard it suggests that social systems are regressing to involvement as discussed in session 1. New behavioral standards used for getting involvement are not anymore flags, hymns, uniforms, and habits, but*

linguistic standards. The very first way and more important technology to diffuse such standards is TV, combining sounds and pictures and simulating real presences having such a perceived authority because non-interacting. Combining sounds and picture and simulating real presences gets full attention of the cognitive system of audience.

The same effect was not possible with the radio involving only listening. The problem is not the technology itself, as usual, but its usage. There are several reasons for people to watch TV, like getting information, watching sport games, watching documentaries, having entertainment, and so on. People may watch TV alone or not. In the first case a single person may watch TV for having, that's simulating, company, in order to avoid loneliness. In the second case a little community is dedicated to watching TV, simultaneously synchronized to the TV program.

We would also like to mention the common misunderstanding about watching TV for having relaxing time. Usually time is assumed relaxing when spent doing something pleasant. It may require concentration, like for videogames or crosswords, but not complex thinking, that's just reacting, finding words and not dealing with strategies taking place instead when plying, for instance, chess or cards. The same takes place when time is spent in repeating very simple actions, such as *peeling a piece of wood*, inducing sleeping. By watching TV something else is taking place, that's *delegating thinking to external processes*. Watchers are assumed to just wait for what is going on.

Movies tells stories, TV simulates the involvement of the watchers when organizers and participants directly watch the camera giving the impression of involvement. *Time spent by watching TV programs is intended relaxing because it takes control of the mind activity substituting, controlling and deciding the thinking activity*. That's the meaning of the expression watching TV in general and not watching a specific, selected, program.

Relaxing may be intended spending time being not involved in complex thinking, just enjoying, reacting, selecting or doing simple, repeated activities. *Having cognitive processes managed, decided, and substituted by external events is very different than relaxing*.

In such a way the familiar reduced television language assumed in daily life, because of its reduced representation and description power due to constant combination with pictures, *is suitable to process events and facts only, not abstractions*. Interactions among people represented and conceived by using this language are then reduced to be events driven. Problems discussed by using this language are related to details, materialistic factors (like events but not information, prices but not economics, discoveries but not research, decisions but not strategies, and so on) because this language can represent them only. Usage of this language makes societies unable to

discuss at mass level of ideals, designs, leaving this dimension perhaps to religions only.

We just mention here the availability in the next future of new technologies in controlling human will. They are based on a very sophisticated physics. It is in this area of research the question related to the ethical acceptance of invasive approaches for human care, such as implant of the bio-electronic devices in the body. For now, the technology has been used to act and repair the body. Modifications have been made by surgery and transplantations. Artificial devices have been implanted to support physically impaired functions (for the ailing hearth, for example).

The new possibility, however, is to integrate the body with external and active devices able to perform functions that the natural body is not anymore able to do. *Active* makes reference to the influence on the neurological system, on the functioning of the brain. The problem is that this external device may be a way to introduce manipulating input into the body and the mind and its functioning may be influenced and controlled from outside. Thanks to the knowledge of the physics of mental process this approach is going to be possibly implemented even without invasive approaches. The ethicality of these studies must be carefully considered.

5. BUYING CONSENT

In the past, in democratic societies, the efforts of political competitors were oriented toward communicating to diffuse their programs and proposals and for convincing voters to support them. Competitors used all available communication techniques, psychological effects, advertising, promising, lying, saying only part of the truth, and so on. Always the financial effort to support such political campaigns has been very high, and competitors found different ways to get founds, even if usually from large corporations having interest in influencing their political program.

In the new, post-democratic approach, the process seems to be the same, but the game is played at another level. Competition requires now to play by influencing and interacting with manipulated cognitive models by using sophisticated and very expensive techniques, like influencing language. Because of that very few may play. It's a formal competition between private interests disguised as political programs, and only designers of the system may participate.

The system of rules established in this way has the purpose to make consent buyable (Minati, 2004). The participants in the game may set new rules without making them public.

In this situation only few, loser, competitors may think that they play the usual game while the game is really changed. On one side unaware competitors may only lose because they play the wrong game and on the other side they contribute to make people thinking that the game is always the same, genuine democratic competition. The new game is based on establishing and playing new implicit and very expensive rules: the winner is expected to be the designers of the game, the owners of the very private, expensive, sophisticated knowledge and technology involved. In this case, they are winner no matter who the competitors are.

A couple of examples of such a strategy are:

- Reduce the possibility of selection, by modifying, forcing the context in such a way that only one (the desired one) action is suitable. Often the action induced in this way is to be used as cause or excuse for other actions that designers have in mind for their interests. Selections of actions are allowed only after an artificial setting of the context. It is possible, for instance, to make people desperate because of no houses, no jobs, no rights by controlling inflation rate, the national debt or organizing embargos, and then creating extraordinary laws against them, to control and put down revolts to be used for specific purposes, like make lands uninhabitable or justify racial discriminations.
- Make available a lot of detailed information without any ranking, structure or ordering, that's making impossible any searching or extraction of meaning, if not presuppositions, judgements and general comments. *The availability of too many, fragmented information is worst than lack of information.* In this case people have the illusion of having a detailed full set of information. It's the same that giving pieces of a puzzle mixed with other extraneous components. Observer receive neither the picture neither the effective possibility to build it. Any tentative to reconstruct the global picture produces only deformed caricatures used for discussing, making judgements and taking *non-dangerous* (for the owners of the game) decisions.

Social control in post-democratic societies is based on inducing agents to process available information in the same, desired way. Therefore, the output doesn't depend on the model adopted, being them all equivalent with reference to some kind of information.

Focus of political actions is not on convincing, but on forcing people (considered agents, like buyers, consumers, investors, anonym users of services) to use a standardized and adequate (for private and not social interests) cognitive model. This may be the end of classic western democracies because real consent is not anymore necessary. Western democratic societies are becoming post-democratic societies. *Manufacturing*

consent (Herman and Chomsky, 1988) is one of the big issues of modern sociology, psychology and cognitive science. *Freedom to design degenerates in freedom to select between standardized, pre-established, alternatives*. That's equivalent to freedom in shopping, when availability of concurrent products or services make possible optimisation and not designing. Freedom is in this case an illusion because reduced to selection among logically equivalent choices, not allowing the possibility to design, change rules and game.

George Orwell, in his novel "1984", forecasted this situation when introducing the metaphor of the "big brother". He conceived the existence of an owner of all the media (In Italy, and not only, now a reality!), a big brother controlling all mass communications. In this framework it is really possible to apply the Sapir-Whorf hypothesis. Conceivably, they (i.e. the owner of the game, that's of the technologies able to influence the way of thinking, the cognitive models adopted) control language and induce people to think in a standardized way.

We need to also note how western democracies are managed by leaders elected by decreasing percentages of voting population. The leaderships are elected by a small percentage of population, anyway majority of the voters.

Technologies for controlling and manufacturing consent dramatically contrast with the original principles and values that generated democratic societies.

6. CONCLUSIONS

We think that the classic concept of democracy emerged from centuries of history of western societies, their philosophical and cultural background, materialized in different political experiences and different national constitutions, should be reconsidered in the new technological context allowing manipulation and usages of the core principles (such as equality of rights) against themselves as introduced above.

Are still classic democracies, when manipulations through technology are possible, the political form of government more suitable for dealing with global problems like ecological ones, overpopulation, sustainability of economical development; scientific projects crucial for the human beings like life in space, move human life on other planets and life manipulation?

The distance between cultural, philosophical principles and real applications is increasing. The way to apply is going to be more important than what is expected to be applied. This relates to the problem of transforming ethical principles in rules, that's normative ethics.

It has been well stated how it is unethical the designing of a system for somebody else (Banathy, 1996).

The ethical approach has the purpose to help and facilitate the emerging of more suitable systems for people having different culture.

Western countries assumed the idealistic approach that what is suitable for them must be suitable for anybody else.

We assumed that our view of democracy is like the truth: we reached it and now we have the mission to diffuse it to anybody else.

This approach is now weaker than in the past because of the internal problems in western democracies like the ones mentioned above.

Moreover the economic, military and social importance of some so-called developing countries are in the focus. China and India have a very important process of economic growth and western countries interact both taking advantage and trying to influence cultural and social aspects introducing, for instance, consumerism by avoiding, because of interests, any interference into political aspects.

Who's designing the strategy to interact with those social systems? Corporations interested in amplify their markets, in having cheaper labor, in natural resources, or military forces interested in having strategic positions or religions suitable to specific form of power?

The real lack of democracy is given by the absence of calling for ideas about these problems: there are not public places (like social institutions, universities, schools, and publications) where to be informed of this kind of problems nor where to discuss them. They are kept private, diffusing the idea that those kinds of problems are for experts only, for specialists having not social responsibility, but private political, economical, military responsibility only. There is no, or very limited, or wrong information available. The same for models and descriptions. They are available to marketing specialists, economists, and military. Public information are the non strategic, ineffective and marketable ones typically diffused trough images in TV, movies, advertising, and magazines, etc.: fragments without any meaning useful to build a global picture and understanding.

Is it compatible with original democratic principles that a part of the society (decision-makers, scientists, military, social designers, managers, etc.) schizophrenically (because of manipulating processes) lives taking advantage of another one manipulated (consuming, producing, making wars and doing the *dirty jobs*)?

Can the systems community play a role in trying to break this situation by explaining wretched, unethical usage of some knowledge and helping people to refuse to be manipulated?

REFERENCES

Anderson, J. R., 1983, *The Architecture of Cognition*, Harvard University Press, Cambridge, MA.

Ashby, R., 1956, *An Introduction to Cybernetics*, John Wiley, New York.

Banathy, B. H., 1996, *Designing Social Systems in a Changing World*, Plenum, New York.

Berger, P. L., and Luckmann, T., 1966, *The Social Construction of Reality*, Doubleday and Co., Garden City, New York.

Bickerton, D., 1992, *Language & Species*, University of Chicago Press, Chicago.

Bickerton, D., 1996, *Language and Human Behavior*, (The Jessie and John Danz Lectures), University of Washington Press, Washington.

Bouton, C. P., 1991, *Neurolinguistics: Historical and Theoretical Perspectives*, Plenum, New York.

Cao, Y. U., Fukunaga, A. S., and Kahng, A. B., 1977, Cooperative mobile robotics: antecedent and directions, *Autonomous Robots* **4**:7-27.

Caroll, J., ed., 1956, *Language, Thought and Reality: Selected writings of B. L. Whorf*, John Wiley & Sons, New York.

Chomsky, N., 1986, *Knowledge of Language: Its Nature, Origin and Use*, Praeger, New York.

Cicourel, A., 1973, *Cognitive Sociology. Language and Meaning in Social Interaction*, Penguin, Harmondsworth.

Cliff, D. T., Harvey, I., and Husbands, P., 1993, Explorations in Evolutionary Robotics, *Adaptive Behavior* **2**:73-110.

Deacon, T. W., 1997, *The symbolic species*, W.W. Norton & Company, New York.

Dijkstra, E. W., 1975, How do we tell truths that might hurt?, in: *Selected Writings on Computing: A Personal Perspective*, Springer Verlag, pp.129-131.

Givon, T., 1998, On the co-evolution of language, mind and brain, *Evolution of Communication* **2**:45-116.

Grace, G., 1987, *The Linguistic Construction of Reality*, Croom Helm, London.

Heims, S. J., 1991, *The Cybernetics Group*, MIT Press, Cambridge.

Hemerlijjk, C. K., 2001, Computer Models of Social Emergent Behavior, in: *International Encyclopaedia of the Social & Behavioral Sciences*, Elsevier Science Ltd.

Herman, E. S., and Chomsky, N., 1988, *Manufacturing Consent*, Pantheon Books, New York.

Howes, A., 1995, An Introduction to Cognitive Modelling in Human-Computer Interaction, in: *Perspectives on HCI – Diverse Approaches*, A. Monk and N. Gilbert, eds., Academic Press, London, pp. 97-120.

Huberman, B. A., and Hogg, T., 1993, The Emergence of computational ecologies, in: *1992 Lectures in Complex Systems*, L. Nadel and D. L. Stein, eds., SFI Studies in the Sciences of Complexity, Lectures Vol. V, Addison-Wesley, Reading, MA, pp. 185-205.

Lindsay, P. H., and Norman, D. A., 1977, *Human Information Processing*, Academic Press, New York.

Marcuse, H., 1964, *One Dimensional Man*, Beacon, Boston.

Maturana, H., 1970, Neurophysiology of cognition, in: *Cognition: a Multiple View*, P. Garvin, ed., Spartan Books, New York, pp. 3-23.

Maturana, H., and Varela, F. J., 1980, *Autopoiesis and Cognition*, Reidel, Dordrecht, Holland.

Minati, G., 2001, *Esseri Collettivi*, Apogeo, Milano.

Minati, G., 2004, Buying consent in the "free markets", *World Futures* **60**(1-2):29-37.

Müller, R. A., 1996, Innateness, autonomy, universality - neurobiological approaches to language, *Behavioral and Brain Sciences* **19**:611-631.

Pessa, E., 2000, Cognitive modelling and dynamical systems theory, *La Nuova Critica* **35**:53-93.

Pinker, S., and Bloom, P., 1990, Natural language and natural selection, *Behavioral and Brain Sciences* **13**:707-727.

Rodney, A. B., 1992, Artificial life and real robots, in: *Towards a practice of autonomous systems: Proceedings of the first European Conference on Artificial Life*, J. F. Varela and P. Bourgine, eds., MIT Press/Bradford Books, pp. 3-10.

Von Bertalanffy, L., 1968, *General System Theory: Foundations, Development, Applications*, George Braziller, New York (revised edition March 1976).

Von Foerster, H., ed., 1974, *Cybernetics of Cybernetics. Biological Computer Laboratory.* Dept. of Electrical Engineering , University of Illinois, Urbana, Report 73.38.

Von Foerster, H, 1979, Cybernetics of Cybernetics, in: *Communication and Control in Society*, K. Krippendorff, ed., Gordon and Breach, New York, pp. 5-8.

Wiener, N., 1948, *Cybernetics or Control and Communication in the Animal and the Machine*, Massachusetts Institute of Technology Press, Cambridge.

Wiener, N., 1961, *Cybernetics: or Control and Communication in the Animal and the Machine*, 2nd edition, Massachusetts Institute of Technology Press, Cambridge.

Wilkins, W., and Wakefield, J., 1995, Brain evolution and neurolinguistic preconditions, *Behavioral and Brain Sciences* **18**:161-182.

METASYSTEM TRANSITIONS AND SUSTAINABILITY IN HUMAN ORGANIZATIONS. PART 1 - TOWARDS ORGANIZATIONAL SYNERGETICS

Graziano Terenzi
ATESS - Territorial Agency for Energy and Sustainable Development, Frosinone, Italy
AIRS - Italian Systems Society, Milan, Italy

Abstract: This two-paper series deals with the problem of understanding the relations between viability and sustainability in the context of social metasystem transitions occurring in a global environment. PART I of the paper argues that metasystem transitions in human organizations can be better understood by resorting to concepts from Synergetics, such as those of slaving principle, order parameter and control within order parameter equations. The subject of Organizational Synergetics is thus introduced. After recognizing the centrality of observation processes, an integration to Stafford Beer's Viable System Model is proposed, which includes a dedicated observing subsystem for the detection of both internal and environmental order parameters. The signal flow graph of the subsystem is constructed and a scheme of its overall transfer function is also derived.

Key words: emergence; metasystem transitions; viability; sustainability; Viable System Model; organizational synergetics; control of order parameters.

1. INTRODUCTION

Even at a first glance all the systems in reality appear to be embedded in other systems. Systems that are formed as the result of a stable pattern of interaction among other systems, the latter regarded as its subsystems, are called metasystems. They are the result of the definition of a new level of control (Turchin, 1977) and on turn define a new level of semantic closure

(Pattee, 1991, 1997). One of the basic features of natural metasystems is their ability to self-assembly and to evolve in ever more complex forms of organization. Self-assembly and complexification of metasystems is thus one of the key issues in Systems Science as well as in all of the disciplines which treat the evolution of complexity and organization as their fundamental subject. It has been largely recognized in the systemic community however that the problem of the evolution of complexity is coextensive with the problem of its **emergence**; and the two terms are almost generally used interchangeably. The construction and emergence of metasystems is today usually referred to as a Metasystem Transition. Metasystem transitions are therefore a crucial issue for understanding many of the astonishing phenomena occurring in the natural as well as in the human world (Heylighen, 1995). But in what sense exactly are metasystem transitions relevant to human systems? In the context of human systems, indeed, metasystem transitions appear in all their dramatic force (Heylighen and Campbell, 1995); they also raise a number of questions that call for an urgent answer. What are the forces that drive social evolution and how are they bound to subsystems interacting in an environment? Where does the ability of human systems to innovate themselves rely? What is the relationship between viability and sustainability? How does it happen that specific synergistic processes occur within human systems? How can they be influenced in order to make the metasystem transition sustainable for all the subjects involved in a way that may be regarded as scale-invariant?

In this paper I try not to answer all of the aforementioned questions directly and systematically, but I rather put forward a set of suggestions about how to approach a possible answer. I begin by introducing what will be referred to as Organizational Synergetics, which tries to synthesize concepts from Synergetics with concepts from the Theory of Organization. After considering the issues of metasystem transition and emergence, I argue that self-organization of human organizations is modulated and shaped (even though not completely determined) by the relationships that they bear to their environment: this latter is regarded, in a systemic sense, as anything external to the model of the system under study but on which, nevertheless, its functioning depends. Dynamical evolution of a system, indeed, is strongly influenced by control parameters that exemplify the relations the system bears to its environment. Anyone who could in principle manipulate these parameters could also control its dynamical evolution. Moreover, internal relationships can induce catalytic processes within the system itself, which, on turn, can also induce the instauration of collective modes of behavior at all of its levels of organization. We'll see also that the control of these collective modes plays a crucial role in the induction of important restructuring processes within the systems. I'll remark, however, that the

ability of human systems to spontaneously innovate themselves resides in their intrinsic "logical openness" which leads to the identification and implementation of novel paths of evolution. Finally, it will be proposed an integration to Stafford Beer's Viable System Model (Beer, 1989) that takes into account a suitable process of internal and external observation deemed exactly to disclose organizational collective modes and their relations.

2. EMERGENCE AND METASYSTEM TRANSITIONS

The view according to which Human Systems can be better understood as complex self-organizing entities, emerging out of the interaction among a large number of individual human agents and their environment, is one of the basic tenets of the Systems movement (von Bertalanffy, 1968; Boulding, 1978; Beer, 1975, 1981; Checkland, 1981). One of the essential premises of this view is that the concepts of *self-organization* and **emergence** play a central role in the definition of what a human system is and how it could effectively be operated on. As stated in the previous section, self-organization and emergence are also strictly tied to the concept of Metasystem Transition, in that both of them concern the process of construction of novel systems. It is then clear why any approach which tries to apply systems concepts and techniques to human systems should as a first step also devote much effort to the clarification of the meaning of the terms "*self-organization*" and "*emergence*". In this section I try to synthesize the main ideas concerning Emergence and Self-Organization in order to delineate the general conceptual framework that justifies the arguments to come.

2.1 Self-Organization and Emergence

Both the terms self-organization and emergence refer to the problem of explaining the spontaneous appearance of novel and ever more complex forms of organization in the context of the interactions among a system's constituents. By taking this issue seriously we unavoidably ask ourselves a number of fundamental questions: *how do highly ordered collective behaviors come into light in such a way as to exhibit long-range space-time correlations and sophisticated hierarchical arrangements? How can they be detected, reproduced or explained? Are there any generating mechanisms? And if so, what kind of mechanism we refer to? What role does the observer play in the explanation of such mechanisms? What general properties should*

they possess in order to explain the appearance of ordered collective behaviors?

All of this amounts to formulating the **problem of explanation** of emergence, that on turn requires an appropriate definition apt to establish whether "something" is emergent or not, and how it happens to be so. Many efforts have actually been devoted to the solution of this problem during the past century. By referring to the most recent lines of research which date back to the pioneers of Second Order Cybernetics, emergence can be better understood as a property of the system which comprises both the *observer* and the *observed system*. According to von Foerster (1972) the properties that are constitutively ascribed to observed systems essentially depend on the existence of an observing system with which they interact. In the last decades many attempts have been made in order to better understand the role of the observer with respect to the modelling of emergent phenomena. Worth mentioning contributions are those by Haken (1977; 1988), Heylighen (1991), Cruchfield (1994), Baas and Emmeche (1997) and Pessa (1998). In accordance with the aforementioned contributions, it has been argued elsewhere (Terenzi, in preparation) that the concept of *emergence* is strictly tied to the detection, performed by an **observing system**, of structural novelties within the causal organization of **logically open systems** which also support **phase transitions** and the integration of stable patterns of behavior on different scales (i.e. different mesoscopic levels). It is a common view that spontaneous emergence of ordered behaviors and macroscopic structures is strictly dependent on local interactions among a system's constituents. However, it is also clear from Theoretical Physics (Synergetics and Quantum Field Theory) that local interaction alone is not enough in order to produce macroscopic structure. For this reason, local interaction should be supplemented in some way. Indeed, the appearance of macroscopic order occurs only in *far from thermal equilibrium* systems, such as *open systems* dynamically coupled to an environment with which they exchange energy, matter and information (dissipative systems). These exchanges are usually formalized by a set of *control parameters*. By studying the change of the system's response on a variation of the control parameters, one can focus on situations in which the external influences result in a macroscopic ordered reorganization of the state of the system. This amounts to saying that the system undergoes a *non-equilibrium phase transition*, by switching from one state to another one. During phase transitions, new quantities, the so-called *order parameters*, emerge. They do not need to be physical quantities, nevertheless they describe and govern the new evolving state as a whole. Emergence can thus be seen as the coordination (coherence) of a system's constituents by means of **ordering information** that is rooted in the signals that constituents send each other

while interacting and exerting *mutual constraint* on each other. This essentially means that ordering information is displayed by the constituents' behavior that can be detected by other constituents' perceptual apparatuses. It must be remarked that ordering information is a consequence of the reorganization of the system in the context of a phase transition. However, emergence arises only in the context of the interaction among observing systems, and emergent quantities are observer-dependent. Summing up, self-organization and emergence are jointly tied to the existence of the following things:

1. **far-from-equilibrium open systems** (e.g. by energy, matter, information supply)
2. **local interaction** among subsystems which exert **mutual constraints**
3. **ordering information** which coordinates individual behaviors (macroscopic order)
4. critical values of control parameters in **phase transitions**
5. **observing systems** which detect emergent quantities (order parameters)
6. **logically open systems**, which only can account for true novelty.

The concepts just introduced are also crucial for the formation of completely novel systems departing from a repertoire of existing variants, like in what has been referred to as **metasystem transitions**. In this context I stress that the formation of novel systems requires the existence of a mechanism for the generation of variety within a repertoire of low-level systems together with a process of self-organization and emergence that guides the construction and retention of higher-level entities. The possibility of resorting to ever changing repertoires of variants is guaranteed by the intrinsic logical openness of the context that hosts the formation of new systems. To the logical openness is also due the ability of systems to evolve in completely novel patterns of organization.

2.2 Order Parameters Equations and Control: Towards Organizational Synergetics

In this section I suggest to use the term "*Organizational Synergetics*" to refer to a possible line of research, which is meant to extend concepts and methods from Synergetics to the study of the structure and functioning of organizations in general. Here, in fact, the concept of "order parameter" is a subject of great relevance. By order parameter in the theory of phase transitions is meant a variable the average value of which provides a *measure* (signature) of the *order* of the system. Order is typically conceived of as a consequence of a **spontaneous symmetry breaking** process (where *symmetry* means *disorder*, and *broken symmetry* or *less symmetry* means

order). Then, it is clear why collective behaviors that become unstable in the vicinity of a critical value of a control parameter within far-from-equilibrium systems are referred to as order parameters. Once established by mutual constraint of interacting constituents, order parameters determine the behavior of the constituents themselves. This means that in the vicinity of a critical value of a control parameter, order parameters "enslave" the individual behaviors, i.e. by modifying the value of order parameters in the vicinity of an instability all system's constituents undergo a modification that is a function of the change of the order parameter. This is referred to as the **slaving principle** (Haken, 1977). The slaving principle realizes what is also known as **circular causality**, according to which the behavior of the constituents determines the occurrence and the properties of the order parameters, and vice versa. The major benefit of the concept of an order parameter is its **information compression** capacity with respect both to quantitative analysis and control of complex dynamics. Indeed the dimensionality of order parameter equations is less than the dimensionality of the original system's equations. This is profitable if we seek to control complex dynamic systems. As a matter of fact, the control of few order parameters is much easier than the control of whole multi-dimensional complex systems, and the learning capability that evolved in the critical regimen of biological evolution also seems to show that abstraction of regularities within systems' dynamics proved to be adaptive, as a consequence of the control that organisms could exert over their environment. Indeed, also the notion of control is crucial both with respect to the evolution of complex systems and to their survival. According to Joslyn (1991), control requires *constraint*, i.e. *selection* or *reduction in variety* of the possible states of an entity, driving the dynamical evolution of the entity towards a final state that remains stable in the face of factors that should produce variability (Marken, R. 1988). The controlled variable remains stable as a consequence of the interaction between a *controller* and the *controlled system*. Controlled systems are typically far-from-equilibrium systems. Moreover, the controlling system is usually more powerful than the controlled one, in that it can suppress any attempt by the other to impose its preferences. This fact translates into an *asymmetry* in the control loop and an *amplification* of ordering information through the control system.

As noted above, order parameter equations describe the behavior of a system in the vicinity of an instability. Since the dimensionality of order parameter equations is less than the dimensionality of the original system's equations, then it is convenient to introduce suitable *control mechanism* within order parameter equations, and to use the obtained results for the control of the original system. This is the subject of a number of studies that demonstrate the feasibility of such methods (Levi, Schanz, Kornienko and

Kornienko, 1999). The relevance of this approach can be recognized with respect to the analysis, synthesis and control of complex dynamic systems, human systems included.

2.2.1 Hierarchical Control of Human Systems and Social Order Parameters

Emergent phenomena are ubiquitous both in nature and society. But in what sense the concepts of emergence and self-organization can be applied to human systems? A simple example would help here. At a first glance *societies* and *companies* intended as distinct *juridical subjects* are entities endowed with qualities not inherited by their members. If a company produces a particular commodity, e.g. a Boeing 747, no one would say that each of the company's members and employees physically produces that commodity (Boeing 747). Indeed, societal action is a kind of collective action (performed by means of a particular organization) that is ruled by a system of rules and that is the result of local interaction among organization's constituents. Moreover, these new qualities are detected by observing systems (i.e. other human beings) which operate upon them and "control" their behavior in order to guarantee reciprocal interests (that is, mutual constraint), e.g. by negotiating a suitable legislation. New levels of organization are established this way. Production of commodities is thus linked to (even though not the same as) the concept of order parameter for the societal level, in that production of commodities is determined by local interaction, and by modifying production also the company must be reorganized as a whole (circular causality). Another example is given by social institutions, which emerge in the context of human interaction as *control mechanisms* deemed to guarantee the survival in an environment of a community as an autopoietic system (Maturana and Varela, 1980).

3. ORGANIZATIONAL SYNERGETICS AND THE VIABLE SYSTEM MODEL

So far, I have examined the meaning of what I have called Organizational Synergetics and I have emphasized its relevance for understanding emergence in human systems. In this section, I turn to examine the subject in relation to the issue of viability. To this aim, however, we need to consider in more detail the relationship between concepts from the field of Synergetics and the Viable System Model (also VSM).

Connections between Synergetics and the Theory of Organization are evident and rather direct. Indeed, they appear in all their clarity if we

examine the VSM. Stafford Beer's marvellous insight is that every organization could increase its success and viability by revising its structure and functioning according to a suitable model of the functioning of organic systems. Effective organizations, in fact, seem to share a set of organizational principles with organic systems (like the human body). By resorting to homeomorphism one is thus allowed to reconstruct the features of the abstract model that describes it (see figure 1).

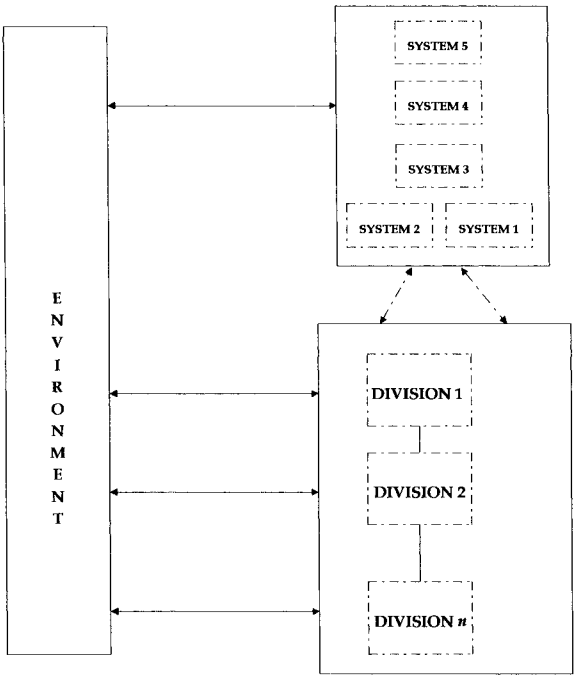


Figure 1. Schematic outline of the Viable System Model; in DIVISION n, typically $5 \leq n \leq 7$ (adapted from Beer, S 1981, p. 157).

According to Beer's view (Beer, 1981, 1989), all the viable systems are characterized by the existence and interaction of five subsystems, which together are jointly involved in defining and maintaining the system's identity independently of other organisms within a shared environment. The process of their formation is thus a metasystem transition. The first subsystem of any viable system (S1) consists of those elements that produce it (i.e. its autopoietic generators); they are the operative units that carry out the primary activities of an organization, i.e. the main organizational tasks, in an environment. According to Beer's theory all viable systems contain and are contained in viable systems; thus each operative unit must be a viable system on its own; if not so the condition of viability would be

violated. The second subsystem of any viable system (S2) consists of those elements and processes that coordinate the activity of the operative units, by resolving conflicts eventually arising in their interaction. The third subsystem involved in any viable system (S3) consists of those elements that monitor and control the activity of the operative units and system S2, by interpreting their alerting signals and by issuing worst-case scenario instructions. System S3 also passes information to the fourth system involved in any viable system (S4). S4 on turn is devoted to link the process of internal regulation to the process of observation of the external environment, and also to a process of future planning which is meant to allow for organizational adaptation; it also passes information to the fifth and last subsystem involved in any viable system (S5) and receives information from it. Finally, the fifth subsystem, S5, consists of those elements that oversee the entire system and are devoted to set out system's purposes and its general adaptive strategies, which, on turn, are passed back to the other subsystems and revised according to incoming information. Clearly, these last two subsystems must possess a significant ability to learn as well as huge adaptive skills.

3.1 Indicators as Order Parameters

A crucial role is played in the VSM by the communication channels and by the information systems. Whereas communication channels are the *condicio sine qua non* of focused activity, information systems allow the overall system to work effectively and minimize the risk of instability. However, in order for each subsystem to work properly and effectively in the context of the whole system, information must exist that constitutes a real-time model of the goings of the organization and that describes the behavior of the interacting subsystems as a whole. Then it makes sense to use indicators of some sort (such as performance indicators) and to assess their change in order to alert the system in case of necessity (e.g. when something wrong is happening).

It turns out, however, that this kind of information (the so called "algedonic" signals) seems to be strictly tied to what we have been calling, up to here, order parameters. Indeed, indicators can be considered as quantities that describe the overall behavior of a subsystem in the context of the higher-level system. In fact, they are determined by lower level interactions among elements in the subsystem and typically are considered as describing an ordered and normal condition if they fall within a precise range. They describe the degree to which a set of functions is properly executed for maintaining the whole system ordered and stable. Finally, if one

were allowed to actually modify the value of an indicator, then the overall behavior of the elements of the subsystem would also result to be modified.

This amounts to saying that indicators are the same as the order parameters. However, if we want to operate modifications on low-level behavior then we have to know how does low-level behavior map to the corresponding higher-level order parameters. Indicators, as order parameters, are thus a crucial element for the effective functioning of any organization.

3.2 Internal Observation and Learning of Order Parameters Relations

Suppose that we could identify, within our organization, some of the order parameters that are informative of the functioning of its operative units. Suppose, moreover, that we could analyse the patterns of interaction between the order parameters at stake, and that interesting regularities emerged from our analyses (that is: if I alter parameter P1 this way then I get parameter P2 and P3 altered that way). Then, it is clear that simply by knowing the relationships between these global variables and by suitably intervening on the values of some, we could in principle affect the value of the others. In fact, according to the *slaving principle* the behavior of the components of the system undergoes a modification that is a function of the modification of the order parameter itself. For example, if we want Productivity to get improved and we know that Productivity is affected positively, up to a certain point, by increasing Motivation and Morale then we could try to find strategies for improving motivation and morale and finally implement them, without the need of reorganizing production processes. However, for this argument to be valid we must be able 1) to identify proper order parameters in the context of the organization on the various mesoscopic levels, 2) to measure their values, and 3) to identify the precise relationships that hold between them on their specific level.

I have argued that emergent quantities, i.e. order parameters, are measures of the ordered behavior of a system as a whole and that they describe the role played by the system as a whole in an higher level of interaction. As such they are detected by observing systems. Consequently, in order to detect order parameters in our organization we need a process of *internal observation and measurement* occurring on a continuous basis and on different meso-levels (since emergence may be recursive and structured). Moreover, observation must be paralleled by learning, because only through learning novel regularities can be recognized and actual relationships between parameters can be identified properly. This also means that both observation and measurement play a double role in metasystem transitions

occurring in organizational settings, as it appears to be if we consider the VSM.

3.3 External Observation and Organization's Sustainability

Viability is strictly tied to sustainability. Whereas viability is the systemic capability of independent and persistent existence, sustainability is the ability to survive extended to the higher-level system that comprises it. Another important function of the observation process is thus that of assessing the sustainability of the organization's activity in relation to its environment. Environment indeed is not to be considered as something totally external to the organization. A great quality of the VSM is that of having stressed the essential role of the environment in the definition of the viability of any organization. But without the appropriate relation with its environment the organization simply could not exist. So what exactly we refer to when we talk of "appropriate" relation? As it often happens, only a little fragment of this relation is taken into account (e.g. profits and costs); the rest is considered as irrelevant. The idea of including internal performance indicators, different from money, in the assessment of achievement is of great importance. However, they only remark indirectly the ways in which internal activity depends on external processes, but do not say anything on how it bears on the external environment. Clearly, a more complete picture has to be depicted. And the reason is primarily that an organization cannot survive in an environment rendered unsupportive by its activity. Therefore, the ability of the environment to support the activity of the organization must be assessed, continuously monitored and anticipated.

In this sense, the observation process must also be directed to what happens externally to the organization, both spatially and temporally. External observation, thus, amounts to 1) build a model of the environment of the organization inclusive of its ecological and social aspects 2) monitor the status and evolution of the environmental and social variables in relation to organizational activity 3) measure the rate of production, rate of consumption, availability and costs of the resources (both natural and human) used by the organization in its primary activities 4) assess sustainability.

3.4 The Observing Subsystem within the Organization

According to what we have argued in previous sections, observation processes are necessary not only for assessing and auditing the standard

functioning of a viable organization but also for enhancing its potential by disclosing organizational and environmental order parameters, and their relations. By forwarding this kind of information to specific subsystems within the organization when needed, very important modifications can be operated on the complex system without going down blindly to all the low-level details. Figure 2 shows an integration of the VSM with a suitable observing subsystem. Figure 3 discloses in more detail the operation of the latter.

Outlining the role played by a dedicated observing subsystem within a viable organization is very important for one basic reason: standard observation processes take into account only signals and indicators already coded in the cybernetic structure of the organization. So, an ability to code for novel information must not only be considered for anticipating future environmental events but also in order to enhance both organizational effectiveness and organizational change. Indeed, by modifying order parameters on a given mesoscopic level, the lower-level structure results also automatically modified (given that we already know the involved mappings).

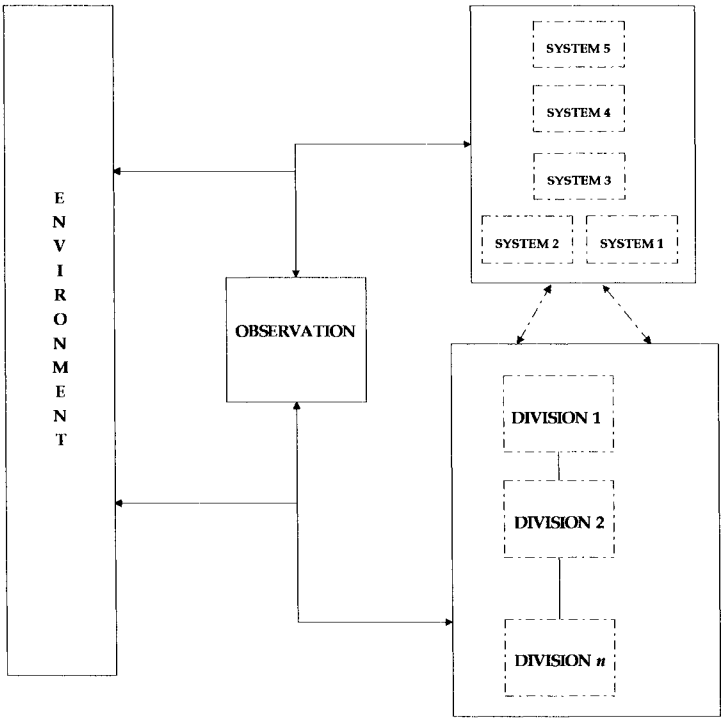


Figure 2. VSM with observing subsystem.

Operations Control of the Observation Process (System 1 and 2 omitted):
1. Building a model of the environment of the organization, inclusive of its ecological and social aspects.
2. Monitoring the status and evolution of the environmental variables and quantities.
3. Measuring the rate of production, rate of consumption, availability and costs of the resources used by the organization in its primary activities.
3A. Filtering information from operative units.
4. Assessing sustainability and identifying environmental order parameters. Forwarding information to S4 and S5 .
5. Identifying order parameters in the context of the organization on the actual level of recursion.
5A. Identifying the relationships which hold between order parameters identified at the divisional level of recursion.

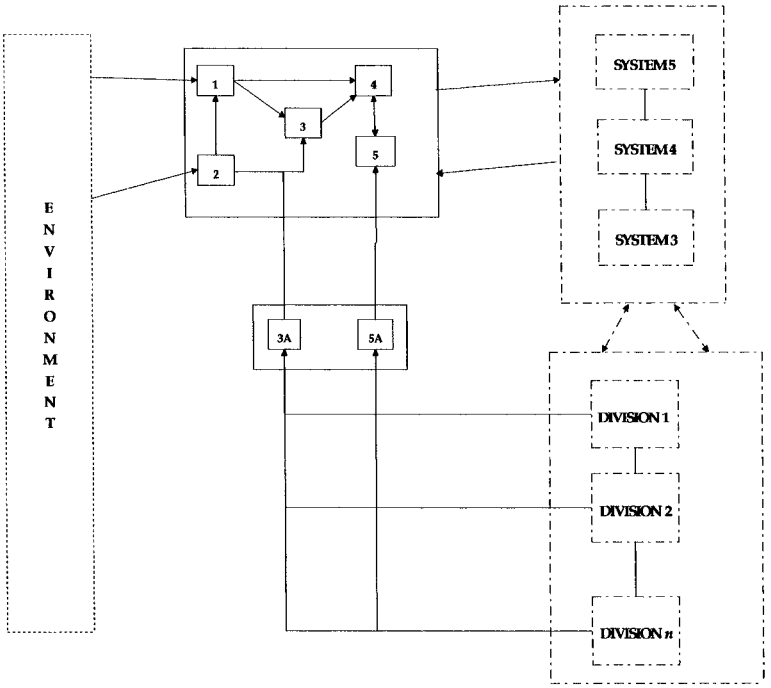


Figure 3. Operation of the Observing Subsystem.

3.4.1 Operations Control of the Observing Subsystem

The goal of this discussion is not that of designing in the detail all of the subsystems involved in the observing system of the organization but simply that of outlining its importance and role within the cybernetics of the organization. To this aim the Signal Flow Graph of the observing subsystem considered as a whole is presented in figure 4. Here F1(s), F2(s), F3A(s), F3(s), F4(s), F5A(s), F5(s) are the transfer functions of the corresponding subsystems. By resorting to the Mason's Rule (Mason, 1953, 1956) we are thus allowed to construct the overall Transfer Function

$$F(s) = \frac{1}{\Delta}\sum_{k=1}^{p} M_k \Delta_k = \frac{1}{\Delta}(M_1\Delta_1 + \cdots + M_p\Delta_p) \qquad (1)$$

where, as usual, p is the number of forward paths in the graph, and Δ = 1 – (sum of all individual loop gains) + (sum of products of all non-touching loops, taken two at a time) – (sum of products of all non-touching loops, taken three at a time) + (sum of products of all non-touching loops, taken four at a time), and so on.

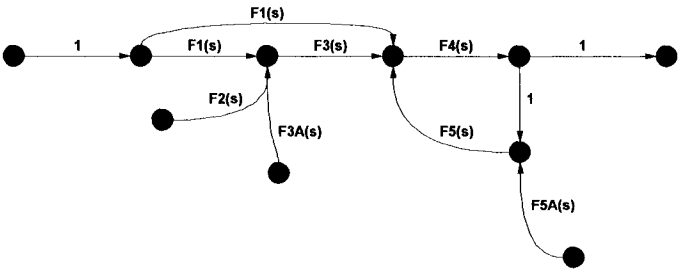


Figure 4. The Information Flow Graph of the Observing Subsystem.

The Transfer Function F of the whole subsystem is thus

$$F=\frac{F1F4+F1F3F4+F2F3F4+F3AF3F4+F5AF5F4}{1-F4F5}. \tag{2}$$

So far I have described the operations control of the dedicated observing subsystem, but I have not said anything about how the latter is operated on by other subsystems, such as Systems 1, 2, 3, 4 and 5 within the whole organization. Here I limit myself to remark that the overall structure and the detailed functioning of the observing subsystem can be changed dynamically by the aforementioned subsystems, and that its operation is particularly useful to System 4 and System 5. If we refer to part of the operation of these latter systems (let's call them M) as to $F_m(s)$, and if we refer to $f_o(s)$ as the transfer function of the observing subsystem (O), we can write the following expression

$$f'_o(s)=F_m(f_o(s),r,c,g), \tag{3}$$

where $f_o'(s)$ is the a revised version of $f_o(s)$, r is a quantity corresponding to available resources of the organization, c is the cost of the operations of O and g expresses the goals of the organization. This essentially means that the intervention of subsystem M in O might result in a selective or a general reorganization. Thus, is up to M to constrain and also to drive, even though

indirectly, the operations of O. After all it couldn't be reasonably different, in that it is a basic function of the "brain" of the organization to continuously assessing the needs of the organization as a whole and to driving its global behavior in order to gain viability.

4. CONCLUSIONS

In this first paper of this two-part series we have dealt with the problem of enhancing viability in the context of metasystem transitions occurring in individual organizations. I have argued that an important role in this process is played by what I have called Organizational Synergetics, which discloses the meaning of order parameters and emergence in human organizations. The recognition of the importance of order parameters and their control in human systems, together with the recognition of the centrality of observation processes in their detection, led us to introduce a suitable integration to the VSM, which, indeed, includes a dedicated observing subsystem for the detection of internal and external emergents. In this context, I have argued that the identification of emergents and the study of their relationships could enhance both organizational effectiveness and transformation. I have also argued that dedicated observation processes are not only necessary to improve organization's viability but also to assess organization's sustainability, this latter considered as a kind of viability extended to the environment of the organization.

REFERENCES

Baas, N. A., and Emmeche, C, 1997, On emergence and explanation, *SFI Working Paper*, 97-02-008, Santa Fe Institute.

Beer, S, 1975, *Platform for Change*, John Wiley, Chichester.

Beer, S, 1981, *Brain of the Firm*, 2nd edition, John Wiley, Chichester.

Beer, S, 1989, The Viable System Model: its Provenance, Development, Methodology and Pathology, in: *The Viable System Model: Interpretations and Applications of Stafford Beer's VSM*, R. Esperjo and R. Harnden, eds., John Wiley, Chichester.

Boulding, K. E, 1978, *Ecodynamics*, SAGE, London.

Checkland, P, 1981, *Systems Thinking, Systems Practice*, John Wiley, Chichester.

Cruchfield, J. P, 1994, The calculi of emergence: computation, dynamics and induction, *Physica D* **75**:11-54.

Forrester, J. W, 1971, *World Dynamics*, Wright-Allen Press, Cambridge, MA.

Haken, H, 1977, *Synergetics*, Springer Verlag, Berlin.

Haken, H, 1988, *Information and Self-Organization: A Macroscopic Approach to Complex Systems*, Springer Verlag, Berlin.

Heylighen, F, 1991, Modeling Emergence, *World Futures* **31**:89-104.

Heylighen, F, 1995, Metasystems as Constraints on Variation: a classification and natural history of MST's, (special issue), *World Futures* **45**(1-4):59.

Heylighen, F., and Campbell, D. T, 1995, Selection of organization at the social level: obstacles and facilitators of MST's, (special issue), *World Futures* **45**(1-4):181.

Heylighen, F., and Joslyn, C, 2001, Cybernetics and second order cybernetics, in: *Encyclopaedia of Physical Science and Technology*, Meyers, R. A., ed., (3rd ed.), Vol. 4, Academic Press, New York, pp. 155-170.

Joslyn, C, 1991, Control theory and meta-system theory, in: *Workbook of the 1st Principia Cybernetica Workshop*, F. Heylighen, ed., Principia Cybenetica, Brussels.

Joslyn, C,, 2001, The semiotics of control and modeling relations in complex systems, *BioSystems* **60**:131-148.

Levi, P., Schanz, M., Kornienko, S., and Kornienko, O., 1999, Application of order parameter equations for the analysis and the control of non-linear time discrete dynamical systems, *International Journal of Bifurcation and Chaos* **9**:1619-1634.

Marken, R. S, 1988, The nature of behavior: control as fact and theory, *Behavioral Science* **33**:196-206.

Mason, S. J, 1953, Feedback theory-some properties of signal flow graphs, *Proceedings of the IRE* **41**(9):1144-1156.

Mason, S. J, 1956, Feedback theory-further properties of signal flow graphs, *Proceedings of the IRE* **44**(7):920-926.

Maturana, H. R., and Varela, F. J., 1980, *Autopoiesis and Cognition: the Realization of the Living*, D. Reidel, Dordrecht, Holland.

Meadows, H. D., Meadows, L. D., Randers, J., and Beherens, W. W. III, 1972, *The Limits to Growth*, Universe Books, New York.

Minati, G., Penna, M. P., and Pessa, E, 1998, Thermodinamical and logical openness in general systems, *Systems Research and Behavioral Science* **15**:131-145.

Paritsis, N, 2002, Holonic property: expected consequences on society and globalization, *Res Systemica* **2**, (Special Issue; Proceedings of the 5th European System Science Congress).

Pattee, H., 1991, Laws, controls, measurements, and symbols, *Applied Mathematics and Computation*, (Special issue on biophysicalism, M. Conrad, ed.)

Pattee, H., 1997, The physics of symbols and the evolution of semiotic controls, in: *Proceedings of the Workshop on Control Mechanisms for Complex Systems*, Addison-Wesley.

Pessa, E, 1998, Emergence, self-organization and quantum theory, in: *Proceedings of the First Italian Conference on Systemics*, G. Minati, ed., Apogeo, Milan.

Terenzi, G., 2003, Global evolution of human systems: a prototype model, in: *Proc. of the 4th Annual Conference of the International Society for the Systems Science*, Hiraklion, Crete.

Terenzi, G, Emergence and logical openness in general systems, (in preparation).

Terenzi, G, Metasystem transitions and sustainability in human organizations - a heuristics for global sustainability, (in this volume).

Turchin, V, 1977, *The Phenomenon of Science. A Cybernetic Approach to Human Evolution*, Columbia University Press, New York.

von Bertalanffy, 1968, *General System Theory*, Braziller, New York.

von Foerster, 1972, Notes on an epistemology for living things, *BCL Report, No 9.3* Biological Computer Laboratory, Dept. of Electrical Engineering, University of Illinois, Urbana.

METASYSTEM TRANSITIONS AND SUSTAINABILITY IN HUMAN ORGANIZATIONS. PART 2 - A HEURISTICS FOR GLOBAL SUSTAINABILITY

Graziano Terenzi
ATESS - Territorial Agency for Energy and Sustainable Development, Frosinone, Italy
AIRS - Italian Systems Society, Milan, Italy

Abstract: This two-paper series deals with the problem of understanding the relations between viability and sustainability in the context of social metasystem transitions occurring in a global environment. In PART 1 of the series, the subject of Organizational Synergetics has been introduced, and an integration to Stafford Beer's Viable System Model has been proposed. PART 2 argues that the identification of emergents and their relationships is not only necessary to improve organization's viability but also to assess organization's sustainability. The distinction is made, then, between local sustainability and global sustainability. Whereas the former is amenable to a standard computational treatment, the latter, it is argued, is an undecidable property of the global system comprising both the system in focus and its global environment. In force of its undecidable and holonic character, global sustainability can only be attained by resorting to suitable heuristics designed to guide global evolution. Finally, a strategy and a general heuristics, which is based on the concept of a Viable Holonic Network, are proposed.

Key words: emergence; metasystem transitions; viability; local and global sustainability; Viable System Model; Organizational Synergetics; control of order parameters; Viable Holonic Networks; holonic property.

1. INTRODUCTION

In the first paper of this two-paper series, we have discussed the issue of metasystem transition and emergence in the context of human organizations.

Even though the emergence of complexity would be hardly considered as a threat by a supposed "totally external" observer who would play no role in the process of construction, it is instead of great relevance for "internal" observers who take part to the same process. From the point of view of internal observers, in fact, the issue of the emergence of metasystems is inseparable from the issue of their viability (i.e. the capability of independent and persistent existence) and from the issue of their sustainability (i.e. the ability to survive extended to the higher-level system that comprises it). In order to be "viable and sustainable" a metasystem must not only possess the capability both to maintain and to regenerate its organization in time, but must also guarantee viability for the environment that puts it up. If a metasystem transition were not viable for an internal observer who is also supposed to take part to the process of construction, then the latter simply would soon cease to be part of the metasystem under construction. And if it were unsustainable for the external environment upon which its existence strictly depends, then the system as whole would soon collapse and lose its long-term viability. Insofar as the separation between external and internal observers is a consequence of an arbitrary decision on behalf of an observing system, how should we consider the issues of metasystem transitions when we turn our attention to Global Human Systems?

In this second paper, after clarifying some aspects of sustainability, I'll turn to discuss a strategy and a general **heuristics** for redirecting the evolution of human systems in a sustainable way. Simply stated, I'll put forward the idea that by intervening "suitably" both on control and on order parameters one would be allowed to influence and redirect the evolutionary trajectories of the human system *in focus* in such a way as to induce the restructuring of an higher-level metasystem shaped according to the principles of freedom, viability and sustainability. The idea, introduced in another paper (Terenzi, 2003), is expanded and discussed here in more detail; it is based essentially on the concept of a **viable holonic network**, an organizational heuristics which is meant to enhance global human evolution.

2. LOCAL AND GLOBAL SUSTAINABILITY

I have argued so far (Terenzi, this volume) that, whatever the level of their organization, human systems are *far-from-equilibrium open systems* that are coupled to an environment that feeds them with matter, energy and information. Control parameters, which characterize in a simplified manner the relations that the human system bears to its environment, essentially concern exchanges of matter, energy and information. Whoever could in principle manipulate those control parameters could also induce specific

patterns of dynamical evolution within the system. In the course of phase transitions collective modes of behavior emerge that describe the evolving state of the system as a whole, and we have referred to the quantities that describe those collective modes as to order parameters. As a consequence of affecting order parameters, also low-level behavior would result to be affected. Order parameters, on their side, are detected by means of suitable observation and measuring processes. These latter can be internal as well as external; in the first case they allow for a better internal regulation and adaptation; in the second case they foster external harmonization and adaptation. According to the latter consideration and following Stafford Beer's Theory, I have argued that human organizations in order to be effective and survive in their environments must be viable. Long-term viability, however, is not attainable if the system is not sustainable.

But how are viability and sustainability related? Whereas individual human agents, restricted social groups and also business organizations and corporations can in many cases definitely be considered as distinct viable systems, it is an open question whether the whole system that comprises them is to be considered as such. Indeed, even though individual business organizations and corporations are amenable to such a kind of analysis on their own right, the same cannot be said of the evolving global scenario, where all of the organizations interact with all of the other actors involved in an integration process which is giving rise to unavoidable conflicts. It is well known from the theory of viable systems that this kind of systems are autopoietic units that regenerate themselves through self-production of their own elements and of the network of processes that produces them. Cooperation among elements and operational closure are essential features of this kind of processes. Also conflicting processes do play a role in the competition for resources among subsystems and threatening exogenous agents. But, in both cases the result of mutual adjustments is the *benefit* of the whole system. Cells, Organs and Organisms fit closely this characterization, as also viable organizations do. However, the global human system does not fit this image at all. Wars, social and economic conflicts, environmental disasters caused by those who also will suffer from their consequences are just a number of the things that prevent the global human system from being considered as a super-organism or an autopoietic system on its own right. Indeed, viability and sustainability, which on their side are two basic features of any healthy relation between an organism and its environment, are seriously challenged. Thus, or are we finding ourselves in the context of a process of evolution which only through a lot of difficulties will bring forth a new integrated humanity, or are we finding ourselves in the same condition of an organism challenged by something like a serious autoimmune disease. Once again this fact highlights that individual

organizations are at best allowed to assessing effectively organizational viability as well as sustainability from the necessarily restricted point of view of the organization itself.

It is now clear, then, that at least two meanings of the term sustainability seem to coexist, one linked to individual long-term viability and another linked to collective long-term viability. I'll refer to the former as to "local sustainability", and to the latter as to "global sustainability". On the one hand, the assessment of local sustainability can gain benefit from the identification, analysis and control of organizational and environmental order parameters, which on turn constrain organizational behavior by anticipating useful information for the design of possible strategies of intervention. On the other hand, however, global sustainability seems to call for the respect of precise as well as unpredictable environmental constraints, which can only be assessed and measured through higher-level observation processes, which go well beyond the boundaries imposed by closed organizations. To cope with these global constraints is thus one of the primary goals of any strategy for inducing metasystem transitions in human systems. This problem appears in all its seriousness when we turn to consider human evolution in the context of the complex global economy, where issues of ethics come into play as a consequence of the strict interrelation between human activities and ecological processes. In this context, we cannot avoid to taking into account the fact that at some extent we all are affected by many local decisions and that we all are **global shareholders** of the global effects of any local activity (Paritsis, 2002).

3. VIABILITY, SUSTAINABILITY AND DECIDABILITY: THE NEED FOR ORGANIZATIONAL HEURISTICS

I have defined the sustainability of a system as some kind of viability extended to the environment of the system itself. In this sense, sustainability is a property of the global system that comprises both the system in focus and its environment and thus applies to the relations that hold between them. Specifically, the system is sustainable if it enhances the ability of an environment to support its activity, and conversely it does not compromise the conditions upon which its viability depends. According to the view put forward by Stafford Beer (Beer, 1981, 1989), viability seems to be a decidable property of organizations in that all viable systems disclose a recursive, and thus computable organizational structure. But is sustainability also recursive or, equivalently, decidable? Whereas idealized system-

environment relations may be regarded as perfectly decidable, and may be fixed in advance by means of a suitable computational model, real world system-environment relations are much less amenable to recursive computational modelling. Unless the environment is considered as a logically closed system, whose behaviors are completely predictable by resorting to known rules and by applying them to a known set of initial conditions, the system-environment relations are largely unpredictable and non-computational. This means that the global system comprising both the system in focus and its environment must be considered as a logically open system at some degree. Even though viability still remains a recursive property of an organization, the organization itself is to be considered as a logically open system. This fact has been recognized for the first time by Stafford Beer himself who correctly outlined that not everything that makes sense for an organization can be expressed in the language of the organization, and that a metalanguage is needed to express it; this is true even if we move up to higher levels of logical description. This amounts to saying that human organizations are logically open systems that are collocated at the top of the hierarchy of logical openness, i.e. logical open systems of degree n, with $n = \infty$. But logically open systems at an infinite degree of logical openness cannot be modelled but only approximated by logically closed models of finite degree (Minati et al., 1998). If sustainability concerns system-environment relations in the context of a logically open system, and since they are undecidable, then we are forced to conclude that sustainability is an undecidable property. This however does not mean that sustainability cannot be attained and that therefore it makes no sense to pursue it. It only means that it does not exist a general algorithm for establishing here and now what is and what is not to count as sustainable for each possible system-environment relation. Clearly in order to maximize their sustainability and thus their viability, organizations are allowed to continuously monitor significant environmental parameters and variables, and to operate accordingly in order to seek their moving target. As we have seen, to this aim the organization needs general adaptive and learning capabilities. We know what sustainability is in specific cases but as far as we can see a generally applicable algorithm does not exist, and it is then an undecidable property of system-environment interaction. In section 4 I'll turn to discuss a strategy and a general heuristics for a sustainable metasystem transition in global settings.

4. A STRATEGY FOR INDUCING SUSTAINABLE METASYSTEM TRANSITIONS IN GLOBAL ENVIRONMENTS

It is now clear why we urgently feel the need for reuniting sustainability and viability in the context of the construction of human systems. In this section, the issue of inducing sustainable metasystem transitions in global settings is discussed in the context of a general strategy that might be assessed by means of computer simulations of human systems dynamics.

The primary goal of the strategy is to *induce* an *overall reorganization* of the human system *in focus* by promoting *an Ethics of Sustainable Development* at each of its levels of organization. This is done initially by concentrating on lowest levels, where bonds with higher-level structures are much weaker and problems are more directly perceived. Doing so amounts to operating *bottom-up actions* that are meant to create the most general consensus and activate novel forms of relationship among all of the actors involved in the development of the human system *in focus*. If, on the one hand, bottom-up actions are performed in order to induce *self-organization* of sustainable human systems, on the other hand negotiation with standard processes of power and control is required to foster both integration of interests and higher-order processes of self-organization.

The can thus be summarized in the following Table 1:

Table 1. Phases of the strategy.

Phase I	Phase II
1) Identify the *human system* in focus	5) Expand control parameters (external influences) by further analysis
2) Identify its *control parameters*	6) Design external systems devoted to redirect in a *sustainable* way the dynamical evolution of the original system by modifying *control parameters* and thus giving rise to novel patterns of organization
3) Identify its actual *order parameters* (*actual emergents*)	
4) Identify the *control mechanisms* rooted in higher order processes	

Whereas phase 1 is referred to as the *analytic phase* of the strategy, phase 2 is referred to as the *synthetic phase*. Both phase 1 and phase 2 are extremely important and the effort that should be devoted to clarify their realization even in restricted territorial or human systems goes beyond the scope of this paper, which, on its side, aims only at providing a conceptual and rather sketchy treatment of the subject. Moreover, the strategy is *general*

in that it could be used in order to carry out actions on each level of organization of any human system. With respect to the synthesis phase it should firstly be noted that it calls for a further specification of the preliminary analysis such that eventually the relations the human system under study bears to its environment could also be modelled explicitly. Soft systems methodology (Checkland, 1981) could be used in order to provide a conceptual (first-sketch) characterization of system-environment relationship as a *human activity system*. More sophisticated modeling techniques, like Multi Agent Systems, Artificial Life and Neural Networks could be used as a further specification. As a consequence of this, the whole system could be considered as a system in which sub-systems that populate the environment interact in such a way as to exchange energy, matter and information in a far-from-equilibrium regimen and drive the human system under study (i.e. the original system) towards a critical point that enables a phase transition. In this context, particular subsystems and evolutionary heuristics should be studied and included into the environment in order to give rise to a sustainable evolution of the system as a whole. Summing up, our strategy consists of introducing suitable "*agencies*" within the environment that feeds human systems in focus in such a way that, by intervening directly over the control parameters of these latter systems, they could induce novel forms of organization (i.e. reorganization) and new order parameters that together should redirect the global dynamics into a sustainable global evolution.

4.1 An Heuristics For Global Sustainability: Viable Holonic Networks

Finding out possible ways of intervention for redirecting evolution of human systems is a rather difficult issue. How could that be achieved? Promoting a true sustainable development amounts to seeking primarily organizational innovation in a stable metasystem. But, in order to be sustainable, innovation must be socially and ecologically viable both on the local and on the global scale, as we all are global shareholders of the effects of any global or even local decision. For this reason, some kind of "agency" is needed in order to catalyze and induce this process. This kind of "agency" must be a powerful one, in that it should strongly influence the human system on which it operates; it must be strongly coupled to the system in focus. Moreover, it must account for the global consequences of its local activity and it must induce a similar pattern of behavior in all of the processes it promotes. In a word, it must possess a Holonic character (Paritsis, N. 2002). According to the strategy developed above, the only way one could hope to do so is by intervening directly on control and order parameters. The idea is, then, about introducing within the global

environment a *network of systems* that would mirror the structure of the human systems *in focus* and could induce on their own level of operation - municipal, provincial, regional etc.- *self-organization* and *participatory design* of *sustainable human systems*. With reference to its environment, each operative unit is characterized by a *horizontal level* of operation that is determined by the wholistic properties of the particular human system on which it *directly* operates (e.g. a regional system, including local institutions, local companies and private citizens). However, each unit is also characterized by a *hierarchy* of *vertical levels* of operation which are determined 1) by the consequences of the unit's actions over the local human system which reflect on the global system (e.g. nation and world system), and 2) by a number of relationships with higher as well as lower-level structures, including higher- (lower)-level institutions (e.g. national and international governments), organizations (e.g. companies, NGO etc.) and *operative units* (Figure 1 (a)).

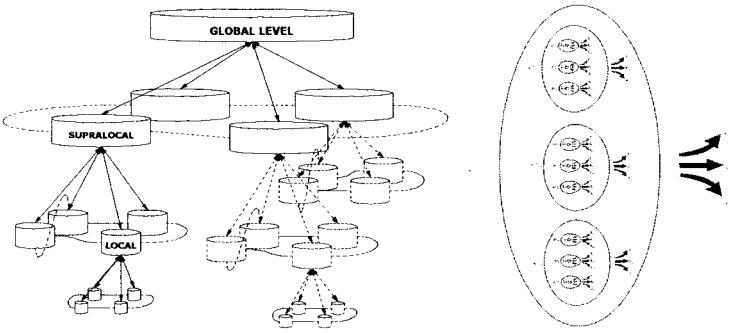


Figure 1. (a) Multi-level structure of the network of organizations. (b) Holonic character of the network with respect to the relationship between sub-networks (including operative units and multiple stakeholders); Ax stands for X-Level-Organization, whereas Sx stands for the set of Stakeholders that should be considered with respect to level X: each X-level- Organization should be committed not only towards X-Level stakeholders but even towards higher as well as lower level ones.

The network is then **multi-level** in that one's allowed to identify, within it, distinct structures (i.e. units) that, by operating on a specific scale, are in dynamical and hierarchical controlling relation with structures operating on different scales. Indeed, it exists a mechanism of communication between all of the distinct levels. In this sense lower-level information bears on higher-level actions, and, conversely, higher-level information bears on lower-level actions; an algedonic signalling system together with a mechanism of bottom-up strategic planning would be much useful. It is **holonic** because the same organizational structure operating on a given level operates on

lower and upper levels as well; moreover, systems that operate on a given level account for the consequences of their activity on all of the other levels, and to some extent they also consider as shareholders all of the stakeholders involved not only on its level but on each other level even indirectly affected by its activity (see Figure 1 (b)). Finally, lower-level structures and higher-level structures are strictly interdependent both in terms of material and non-material resources. In order to be **viable** the network must be constructed according to the viable systems model, and all of the operative units must on turn be on their own viable. The viability of the network is thus extended in the direction of sustainability by means of its holonic character: stakeholders as local environmental agents are considered as shareholders at the global level.

4.2 An Example

The process described above is meant to promote the linking of organizations to a system coherently seeking global evolution, thus minimizing risks for the actors involved and providing a basis for a sustainable global development. An example of the operation of such a network is given by the construction of a Territorial Energetic System based on renewable energy sources. It would be desirable, both socially and ecologically, to reconvert a territorial energetic system in such a way as to diminish the dependence on external agents and polluting energy sources. By involving all of the actors of a particular territorial system in a "common" energy reconversion programme carried out by the constitution of a suitable higher-order organization, one could hope to build up a renewable energetic system which would hopefully nullify external dependence and thus enrich, on the long run, all of the actors involved. The earnings of the organization would thus be distributed among the shareholders, and a minimal part of them could also nourish a Local "Holonic" Fund, for the benefit of the local community. Moreover, one could think of a replication of a similar solution in other contexts and with respect to different activities. Thus, many diverse Local Funds could be instituted; and the many local funds could devolve a certain amount of their capital to a suitable Higher-Order "Holonic" Fund, and so on. This would constitute a feasible realization of the concept of Holonic Property as it has been recently introduced by N. Paritsis (2002).

5. CONCLUSIONS

In this two-part series, we have dealt with the problem of understanding the relations between viability and sustainability in the context of

metasystem transitions occurring in the global human system. In PART I, I have argued that a crucial role in clarifying this relation is played by what I have called Organizational Synergetics, which discloses the meaning of order parameters and emergence in human organizations. The recognition of the importance of order parameters and their control in human systems, together with the recognition of the centrality of observation processes in their detection, led us to introduce a suitable integration to the VSM, which, indeed, includes a dedicated observing subsystem for the detection of internal and external emergents. In this context, I have argued that the identification of emergents and the study of their relationships could enhance both organizational effectiveness and transformation. I have also argued that dedicated observation processes are not only necessary to improve organization's viability but also to assess organization's sustainability, this latter considered as a kind of viability extended to the environment of the organization. PART 1 of the series, showed, however, that sustainability is an undecidable property of the global system comprising both the system in focus and its environment. Therefore in order to foster sustainability an organization requires strong adaptive skills that rely on suitable heuristics and observing capabilities. But since, as I have argued, environmental issues possess an holonic character, and we all are global shareholders of the effects of any local decision, then a complex system must be included in the global environment to manage dynamically all of the organization-environment trade-offs, in such a way as to decrease global entropy and redirect global evolution towards the goal of global sustainability; this, indeed, is the meaning of the heuristics I have called "Viable Holonic Network". It is our hope that the application of the aforementioned heuristics will appear to be feasible in real settings, and that a truly systemic approach to global sustainability will emerge as a consequence of developing future applications.

ACKNOWLEDGEMENTS

Among all of the friends who contributed to the maturation of the ideas presented in these papers, I would like to express my gratitude particularly to Prof. G. Minati for his inspiration, and to Prof. N. Paritsis for a long and stimulating conversation we had in Crete.

REFERENCES

Baas, N. A., and Emmeche, C, 1997, On emergence and explanation, *SFI Working Paper*, 97-02-008, Santa Fe Institute.

Beer, S, 1975, *Platform for Change*, John Wiley, Chichester.

Beer, S, 1981, *Brain of the Firm*, 2nd edition, John Wiley, Chichester.

Beer, S, 1989, The Viable System Model: its Provenance, Development, Methodology and Pathology, in: *The Viable System Model: Interpretations and Applications of Stafford Beer's VSM*, R. Esperjo and R. Harnden, eds., John Wiley, Chichester.

Boulding, K. E, 1978, *Ecodynamics*, SAGE, London.

Checkland, P, 1981, *Systems Thinking, Systems Practice*, John Wiley, Chichester.

Cruchfield, J. P, 1994, The calculi of emergence: computation, dynamics and induction, *Physica D* **75**:11-54.

Forrester, J. W, 1971, *World Dynamics*, Wright-Allen Press, Cambridge, MA.

Haken, H, 1977, *Synergetics*, Springer Verlag, Berlin.

Haken, H, 1988, *Information and Self-Organization: A Macroscopic Approach to Complex Systems*, Springer Verlag, Berlin.

Heylighen, F, 1991, Modeling Emergence, *World Futures* **31**:89-104.

Heylighen, F, 1995, Metasystems as Constraints on Variation: a classification and natural history of MST's, (special issue), *World Futures* **45**(1-4):59.

Heylighen, F., and Campbell, D. T, 1995, Selection of organization at the social level: obstacles and facilitators of MST's, (special issue), *World Futures* **45**(1-4):181.

Heylighen, F., and Joslyn, C, 2001, Cybernetics and second order cybernetics, in: *Encyclopaedia of Physical Science and Technology*, Meyers, R. A., ed., (3rd ed.), Vol. 4, Academic Press, New York, pp. 155-170.

Joslyn, C, 1991, Control theory and meta-system theory, in: *Workbook of the 1st Principia Cybernetica Workshop*, F. Heylighen, ed., Principia Cybenetica, Brussels.

Joslyn, C,, 2001, The semiotics of control and modeling relations in complex systems, *BioSystems* **60**:131-148.

Levi, P., Schanz, M., Kornienko, S., and Kornienko, O., 1999, Application of order parameter equations for the analysis and the control of non-linear time discrete dynamical systems, *International Journal of Bifurcation and Chaos* **9**:1619-1634.

Marken, R. S, 1988, The nature of behavior: control as fact and theory, *Behavioral Science* **33**:196-206.

Mason, S. J, 1953, Feedback theory-some properties of signal flow graphs, *Proceedings of the IRE* **41**(9):1144-1156.

Mason, S. J, 1956, Feedback theory-further properties of signal flow graphs, *Proceedings of the IRE* **44**(7):920-926.

Maturana, H. R., and Varela, F. J., 1980, *Autopoiesis and Cognition: the Realization of the Living*, D. Reidel, Dordrecht, Holland.

Meadows, H. D., Meadows, L. D., Randers, J., and Beherens, W. W. III, 1972, *The Limits to Growth*, Universe Books, New York.

Minati, G., Penna, M. P., and Pessa, E, 1998, Thermodinamical and logical openness in general systems, *Systems Research and Behavioral Science* **15**:131-145.

Paritsis, N, 2002, Holonic property: expected consequences on society and globalization, *Res Systemica* **2**, (Special Issue; Proceedings of the 5th European System Science Congress).

Pattee, H., 1991, Laws, controls, measurements, and symbols, *Applied Mathematics and Computation*, (Special issue on biophysicalism, M. Conrad, ed.)

Pattee, H., 1997, The physics of symbols and the evolution of semiotic controls, in: *Proceedings of the Workshop on Control Mechanisms for Complex Systems*, Addison-Wesley.

Pessa, E, 1998, Emergence, self-organization and quantum theory, in: *Proceedings of the First Italian Conference on Systemics*, G. Minati, ed., Apogeo, Milan.

Terenzi, G., 2003, Global evolution of human systems: a prototype model, in: *Proc. of the 4th Annual Conference of the International Society for the Systems Science*, Hiraklion, Crete.

Terenzi, G, Emergence and logical openness in general systems, (in preparation).

Terenzi, G, Metasystem transitions and sustainability in human organizations - a heuristics for global sustainability, (in this volume).

Turchin, V, 1977, *The Phenomenon of Science. A Cybernetic Approach to Human Evolution*, Columbia University Press, New York.

von Bertalanffy, 1968, *General System Theory*, Braziller, New York.

von Foerster, 1972, Notes on an epistemology for living things, *BCL Report, No 9.3* Biological Computer Laboratory, Dept. of Electrical Engineering, University of Illinois, Urbana.

9

SYSTEMIC APPROACH AND INFORMATION SCIENCE

SCALE FREE GRAPHS IN DYNAMIC KNOWLEDGE ACQUISITION

I. Licata[1], G.Tascini[2], L. Lella[2], A.Montesanto[2] and W. Giordano[2]
[1]*Istituto di Cibernetica Non-Lineare per lo Studio dei Sistemi Complessi, Via Favorita 9, Marsala (TP);*
[2]*Università Politecnica delle Marche, D.E.I.T., Via Brecce Bianche, 60131 Ancona, Italy*

Abstract: Classical representation forms are not suited to represent knowledge as human mind does. In tasks as discourse comprehension knowledge stuctures have to adapt themselves on the basis of the objectives, the past experiences and the particular context. So we have developed a modular knowledge acquisition system based on cognitive criteria, that dynamically updates a representation by the use of a scale free graph model.

Key words: knowledge acquisition; dynamic representations; discourse analysis; scale free graphs.

1. INTRODUCTION

Our work is based on an observation made by Kintsch (1998) who first noticed that classical forms of representation presented in literature and especially used in IA, such as "associative networks" (Meyer and Schvaneveldt, 1971), "semantic networks" (Collins and Quillian, 1969), "frames" (Minsky, 1975) and "scripts" (Schank and Abelson, 1977), are not suited to represent human knowledge. In particular they all lack of dynamic properties i.e. their structure doesn't adapt to the context of use.

On the contrary human mind, in analysing new information, generates dynamic structures that are adapted to the particular context of use. In order to reproduce as exactly as possible the mechanisms involved in discourse comprehension, they developed the so called "networks of propositions".

This particular model extends and combines the advantages of the classical forms of representation. They are based on the predicate-argument

schema because their atomic parts are propositions linked each other by weighted and not labelled arcs.

The meaning of a node and of its correspondent concept is given by its position in the network, because it is constrained by the close nodes. During the comparison process are not used all nodes to specify the sense of a node, but only those which are activated by the particular context of use given by objectives, accumulated experiences, emotional and situational state etc. So the meaning of every concept in the network is not fixed and permanent but it is always built in a particular mnestic structure called "working memory" by the activation of a particular subset of the neighborhood.

The adoption of a network of propositions for the knowledge representation presents certainly great advantages in comparison to classic formalisms. While semantic networks, frames and scripts organize knowledge in a more ordered and logical way, the networks of propositions are definitely more disorganized and chaotic, but show the not negligible advantage of being capable to vary dynamically not only in time, on the basis of past experiences, but also on the basis of the perceived context.

To specify the activation modalities Kintsch and Ericsson (Kintsch, Patel, and Ericsson, 1999) introduced the concept of "Long Term Working Memory" (LTWM). This is a part of the "Long Term Memory" (LTM) that is the entire network of propositions. LTWM is generated by the short term part of the working memory (STWM). This process is allowed by fixed and stable memory structures called "retrieval cues" that link the objects present in the STWM to other in the LTM.

After defining their model Kintsch and Ericsson tried to implement it. Two particular problems had to be solved, the creation of the LTWM and the formation of retrieval cues.

To define the LTWM Kintsch developed two methods. The first, worked out with Van Dijk (van Dijk and Kintsch, 1983), is a manual technique that starts from the propositions in the text (micropropositions) and by using some organizing rules arrives to the definition of "macropropositions" and "macrostructures" and even to the definition of LTWM.

The second is based on the "Latent Semantic Analysis" (LSA) (Landauer, Foltz, and Laham, 1998). This technique can infer, from the matrix of co-occurrence rates of the words, a semantic space that reflects the semantic relations between words and phrases. This space has typically 300-400 dimensions and allows to represent words, phrases and entire texts in a vectorial form. In this way the semantic relation between two vectors can be estimated by their cosine (a measure that according to Kintsch can be interpreted as a correlation coefficient).

In the second solution the modalities of the information retrieval from the semantic space are not clearly specified. When the "textbase", i.e. the

representation obtained directly from the text, is sufficiently expressed, the retrieval of knowledge from the LTM is not necessary. In other cases a correct comprehension of the text (or the relative "situation model") requires the retrieval of knowledge from the LTM.

After the creation of the LTWM the integration process begins, i.e. the activation of the nodes correspondent to the meaning of the phrase. Kintsch uses an activation diffusion procedure that is the simplified version of the McClelland and Rumelhart classical one (McClelland and Rumelhart, 1986). Firstly, is defined an activation vector whose elements are indexed over the nodes of LTWM. Any element's value is "1" or "0" depending on the presence or the absence of the corresponding node in the analyzed phrase (i.e. in the STWM). This vector is multiplied by the matrix of the correlation rates (the weights of the links of the LTWM) and the resulting vector is normalized. This becomes the new activation vector that must be multiplied again by the matrix of the correlation rates. Such procedure goes on until the activation vector becomes stable. After the integration process, the irrelevant nodes are deactivated and only those which represent the situation model remain activated.

There is also another problem that must be considered. Theorically the position occupied by a word in the LTWM is determined by a lifetime experience, i.e. by the continuous use that is made of it. Obviously this kind of knowledge cannot be practically reached and Kintsch build his semantic space using information taken from a dictionary. It is worth noticing that such operation is done only once and the semantic space is not further updated. So this kind of implementation loses some of its dynamic properties.

1.1 An alternative implementation of the LTWM model

We think that all the previous problems can be fully solved only by dropping the intermediate representation of the semantic space and trying to find a method that directly define the retrieval cues. Unfortunately the lack of adequate textual parsers able to convert the paragraphs of a text on the correspondent atomic propositions has driven us to develop simple dynamic models of associative networks that are based on the LTWM model of Kintsch and Ericsson.

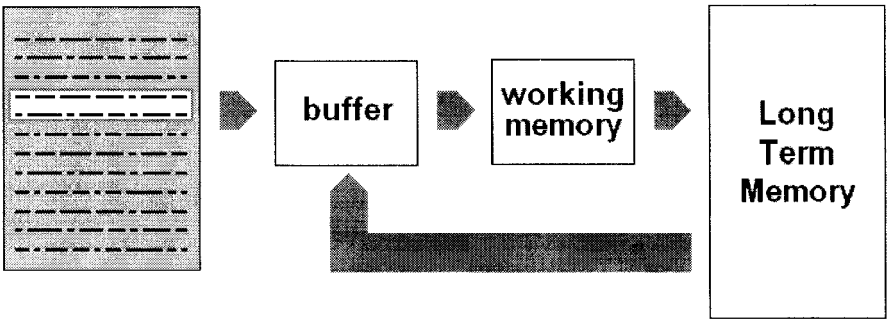


Figure 1. A possibile architecture of a system for the dynamical acquisition of knowledge from a repository of documents.

Fig. 1 describes the system that we used for the extraction of knowledge from textual documents. The part of the document that is analysed – a section, a paragraph or a simple group of words enclosed in a window – is stored temporarily in a buffer. Its content must be codified on the basis of the context before being elaborated by the working memory block. This has been implemented by a simple scale-free graph model (Albert and Barabasi, 2001), due to the fact that recently it has been found that human knowledge seems to be structured in this way (Steyvers and Tenenbaum, 2001).

The analysis is performed over all the paragraphs transferring their content in the buffer. This structure contains not only the words of the analyzed paragraph, but also words retrieved from the LTM by the diffusion of an activation signal starting from the LTM nodes that represent words in the buffer. Theorically the buffer should contain also words activated during the analysis of the previous paragraph, but this aspect of Kintsch model has not been considered for its computational complexity.

The buffer, the working memory and the activated part of the LTM block can be compared (but they are not the same structure) to the LTWM defined by Kintsch and Ericsson.

During the acquisition of the paragraph is used a stoplist of words that must not be considered (as articles, pronouns etc.). For any word in the text, the paragraphs where it has appeared (or where it has been inserted after the retrieval procedure) are stored. When the entire text has been parsed and the data of all the N not filtered words have been memorized, the formation of the network of concepts in the working memory begins.

The model adopted is similar to the one defined by Bianconi and Barabasi (2001). The process starts with a net consisting of N disconnected

nodes. At every step $t = 1..N$ each node (associated with one of the N words) establishes a link with other M units ($M = 5$). If j is the selected unit, the probability that this node establishes a link with the unit i is:

$$P_i = \frac{U_i k_i}{U_1 k_1 + \ldots + U_N k_N}$$

where k_i is the degree of the unit i[1], i.e. the number of links established by it, while U_i is the fitness value associated to the node that can be computed as the ratio between the number of paragraphs that contain both words i and j and the number of paragraphs that contain either i or j.

LTM is an associative network that is updated with the content of the WM. Whenever a link of the WM corresponds to a link present in the LTM, the weight of this one is increased by "1". For example if the WM links "market" to "economy" and in the LTM "market" is linked to "economy" with weight "7" and to "stock" with weight "4", in the updated LTM "market" is linked to "economy" with weight "8" and to "stock" with weight "4" (unchanged). To perform the diffusion of the activation signal all the weights must be normalized. In this case "market" must be linked to "economy" with weight 8/(8+4) and to "stock" with weight 4/(8+4).

Since the scale free network that represents the content of the WM is used to update the content of LTM, this associative networks should take the form of a scale free graph. Unfortunately the modalities of evolution of the LTM does not allow the definition of a simple equivalent mathematic model, that is necessary to make useful previsions about its evolution.

While in the scale free graph models proposed by literature at each temporal step M are added new nodes to the graph, with M defined beforehand, in the system that we have developed, after the analysis of a new document the links related to an unknown number of nodes of the LTM network are updated on the basis of the content of the WM. This is the number of the words that have not been filtered by the stoplist.

Another important difference with other scale free models presented in literature (Dorogovtsev and Mendes, 2001) is the particular fitness function that is used. This function does not depend on a single node but on the considered pair of nodes, i.e the fitness value of a word is not constant but depends on the other word that is linked to it. For example the noun "car" should present for the link with "engine" a fitness value greater than the one presented for the link with "economy".

[1] Each node is connected to itself by a loop.

2. EVALUATION OF SCALE FREE PROPERTIES

To test the validity of the scale free graph model adopted for the WM, we gave 100 files of the Reuters Corpus[28] as input to the system disabling the retrieval of information from the LTM. Two versions of the model have been tested, one with bidirectional links and the other with directed links (in this case we considered $k_i = k_{i\,(\mathrm{IN})} + k_{i\,(\mathrm{OUT})}$).

In Fig. 2 it is represented an example of a network with bidirectional links. The economic bias of the articles justifies the presence of hubs as "interest rate", "economy", etc., while other frequent words as "child", "restaurant", etc. establish less link with the others.

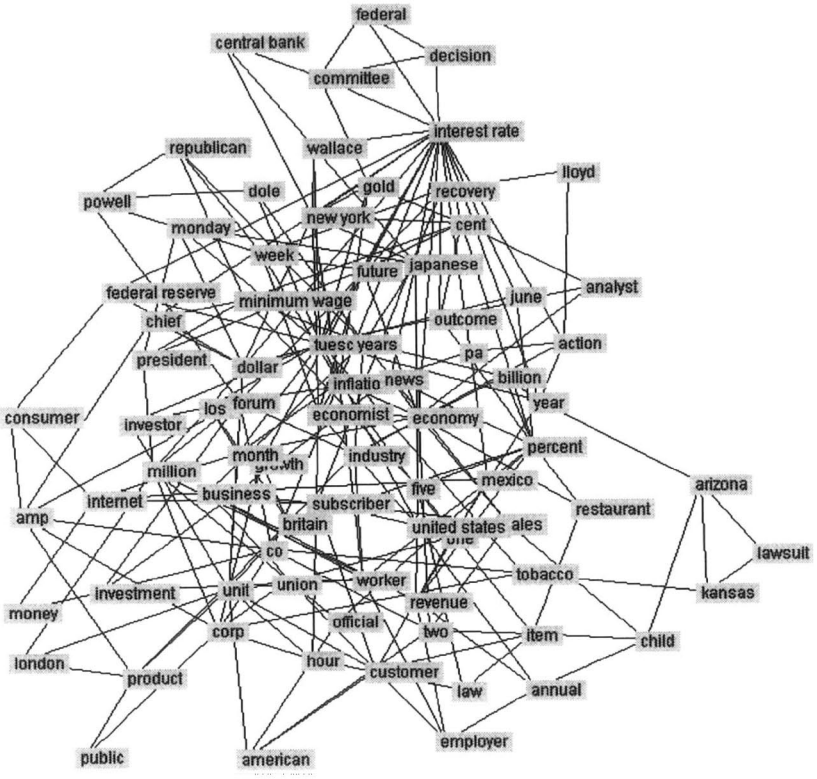


Figure 2. A network with bidirectional links obtained by the analysis of 100 files of the Reuters Corpus.

[28] Reuters Corpus, Volume 1, English language, 1996-08-20 to 1997-08-19, http://about.reuters.com/ researchandstandards/corpus.

The following graphs report the average path length between each pairs of nodes (Fig. 3) and the clustering coefficients (Fig. 4). Some networks with different sizes have been considered filtering less or more words in the analysed texts.

All the results seem to confirm those reached by Bianconi and Barabasi. The trend of the average path related to random graphs having the same dimensions of the considered scale free graphs has an higher slope. The clustering coefficient of the scale free graph model has an higher order of magnitude in comparison with the one computed for random networks.

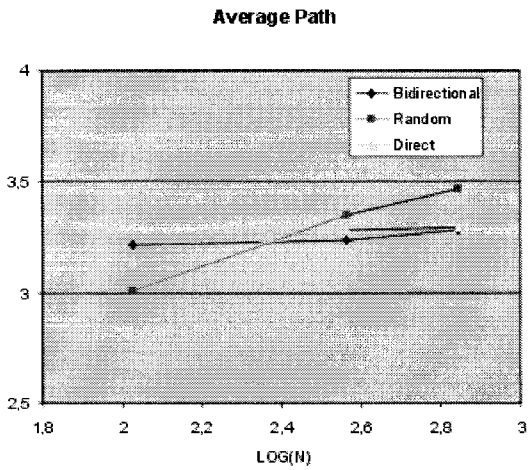


Figure 3. Comparison of average path lengths of different types of networks.

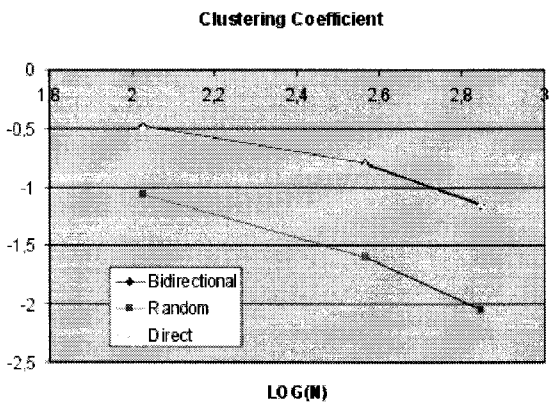


Figure 4. Comparison of clustering coefficients of different types of networks.

Fig. 5 and Fig. 6 show the degree distribution of a graph with bidirectional links and directed links respectively. In the first case the degrees distribution decays as $P(k) \approx k^{-G}$ with G = 3.2657. In the second case the power law trend has a coefficient G = 2.3897.

We also analyzed the structure of LTM associative network that, as expected, kept the features of a scale-free graph (Tab. 1). The system was tested enabling the retrieval of information from LTM and the analysis was repeated 30 times computing the coherence rate of the final knowledge represenations.

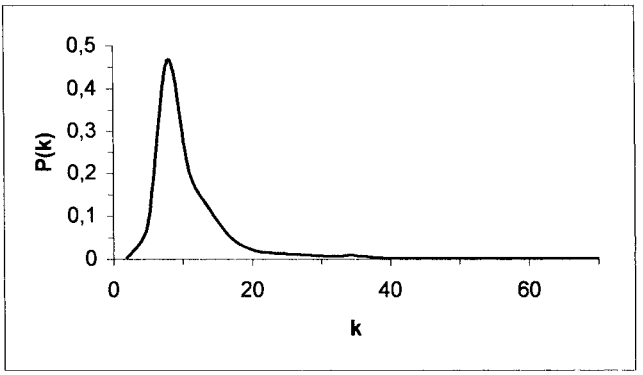


Figure 5. Degree distribution of a graph with $M = 5$ and bidirectional links.

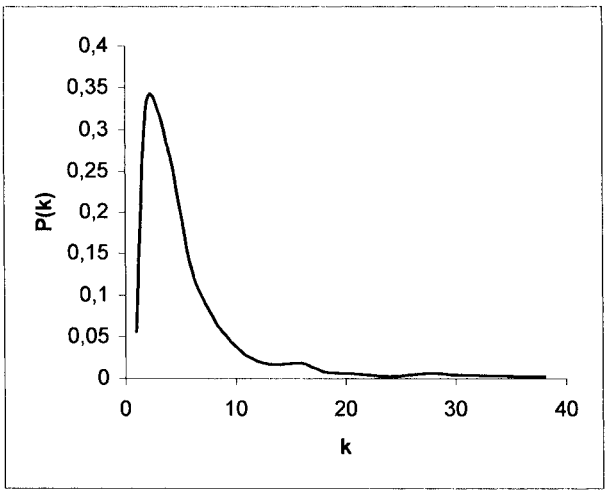


Figure 6. Degree distribution of a graph with $M = 5$ and bidirectional links.

Table 1. LTM with 40 nodes.

M	Average path length	Average degree	Clustering coefficient
1	2.56	5.95	0.32
2	2.49	6.50	0.34
3	2.27	8.30	0.45
4	2.25	9.50	0.43
5	2.23	9.85	0.43

The coherence rate is obtained by correlating the LTM ratings given for each item in a pair with all of the other concepts[29].

The average coherence rate (0.45) has confirmed that the conceptualization, i.e. the evolution of the associative network, was made by the system on the basis of a precise inner schema. Now we are going to evaluate the correctness of this schema by comparing the final LTM representation to other associative networks obtained from a group of human subjects who will read the same texts.

3. CONCLUSIONS

We have presented an innovative knowledge acquisition system based on the long term working memory model developed by Kintsch and Ericsson. The knowledge of the system is structured as an associative networks that is dynamically updated by the integration of scale-free graphs that represent the content of the new analyzed documents. By the diffusion of an activation signal in the LTM associative network, all the information necessary to identify the context of the analyzed concepts (terms) is retrieved.

The analysis of the WM and LTM networks have confirmed that they both are examples of scale-free graphs. The computation of the coherence rate of LTM networks revealed that the system acquires knowledge on the base of precise inner schema whose correctness will be evaluated by the comparison with the other associative networks obtained from human subjects.

Certainly our system is susceptible of improvements. Maybe the presence of the external feedback of the human user could help the system to model correctly his knowledge. For example the links in the LTM could be strenghten only when the knowledge representation is used to filter or retrieve documents correctly. Furthermore the association of an age to the links of the LTM could guarantee more plasticity to its structure. This further information could be used in the computation of the fitness values as in the Dorogovtzev models (Dorogovtsev and Mendes, 2000).

[29] This rate was computed by the software PCKNOT 4.3, a product of Interlink Inc.

We think that our knowledge acquisition system can be effectively used for the semantic disambiguation, that is the first phase of the analysis in the most recent systems for the extraction of ontologies from texts (Navigli, Velardi, and Gangemi, 2003). We are also considering the possibility to extract taxonomical representations from the knowledge stored in the LTM.

REFERENCES

Albert, R., and Barabasi, A. L., 2001, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**:47-97.

Bianconi, G., and Barabasi, A. L., 2001, Bose-Einstein condensation in complex networks, *Physical Review Letters* **86**(24).

Collins, A. M., and Quillian, M. R., 1969, Retrieval from semantic memory, *Journal of Verbal Learning and Verbal Behaviour* **8**:240-247.

Dorogovtsev, S. N., and Mendes, J. F. F., 2000, *Evolution of reference networks with aging*, arXiv: cond-mat/0001419.

Dorogovtsev, S. N., and Mendes, J. F. F., 2001, *Evolution of networks,* arXiv: cond-mat/0106144, (submitted to *Adv. Phys.*).

Kintsch, W., 1998, *Comprehension. A Paradigm for Cognition*, Cambridge University Press.

Kintsch, W., Patel, V. L., and Ericsson, K. ., 1999, The role of long-term working memory in text comprehension, *Psychologia* **42**:186-198.

Landauer, T. K., Foltz, P. W., and Laham, D., 1998, An introduction to latent semantic analysis, *Discourse Processes* **25**:259-284.

McClelland, J. L., and Rumelhart, D. E., 1986, *Parallel Distributed Processing*, MIT Press, Cambridge, MA.

Meyer, D. E. and Schvaneveldt, R. W., 1971, Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations, *Journal of Experimental Psychology* **90**:227-234.

Minsky, M., 1975, A framework for representing knowledge, in: *The Psychology of Computer Vision*, P. H. Winston, ed., McGraw-Hill, New York.

Navigli, R., Velardi, P., and Gangemi, A., 2003, Ontology learning and its application to automated terminology translation, *IEEE Intelligent Systems*, (January/February 2003):22-31.

Schank, R. C. and Abelson, R. P., 1977, *Scripts, Plans, Goals, and Understanding*, Erlbaum, Hillsdale, NJ.

Steyvers, M. and Tenenbaum, J., 2001, The large-scale structure of semantic networks, (Working draft submitted to *Cognitive Science*).

van Dijk, T. A. and Kintsch, W., 1983, *Strategies of Discourse Comprehension*, Academic Press, New York.

RECENT RESULTS ON RANDOM BOOLEAN NETWORKS

Roberto Serra and Marco Villani
Centro Ricerche e Servizi Ambientali Fenice
Via Ciro Menotti 48, I-48023 Marina di Ravenna, Italy

Abstract: Random boolean networks (RBN) are well known dynamical systems, whose properties have been extensively studied in the case where each node has the same number of incoming connections, coming from other nodes chosen at random with uniform probability, and the updating is synchronous. In the past, the comparison with experimental results has been limited to some well-known tests; we review here some recent results that demonstrate that the availability of gene expression data now allows further testing of these models. Moreover, in this paper we summarize some recent results and present some novel data concerning the dynamics of these networks in the case where either the network has a scale-free topology or the updating takes place asynchronously.

Key words: genetic networks; scale-free; attractor; DNA microarray.

1. INTRODUCTION

There are at least two kinds of reasons why genetic networks should be considered particularly interesting: they represent a paradigmatic example of a complex system, where positive and negative feedback loops interact in a high dimensional nonlinear system, and they can model biological systems of great complexity and importance.

The technology of molecular biology is flooding databases with an unprecedented wealth of data, and new methods are required to make sense out of all these data. Genetic networks represent indeed one of the most interesting candidate concepts for this purpose. While sequencing genomes has become common practice, a different set of data is nowadays produced by microarray technology, which provides information about which genes

are expressed under specific conditions or in specific kinds of cells. In multicellular organisms the existence of different cell types is indeed due to the different patterns of activations of the same genome.

It is also well known that the expression level of a gene is influenced by the products of other genes (proteins) and by the presence of certain chemicals in the cell; this chemical environment is in turn affected by the presence of enzymes produced by other genes, so genes influence each other. This kind of relationship is captured in genetic networks, simplified descriptions where only the gene activations are considered. Each gene is a node in the network, and there is a directed link from node A to node B if (the product of) gene A influences the expression of gene B. The best known example of a model of this kind is that of random boolean networks, where each node can be either active (in state 1) or inactive (in state 0). The model is described in several excellent reviews (Kauffman, 1993; Aldana et al., 2003), so we will limit to a brief outline here (see section 2).

A peculiar feature of this model is that it has been introduced not to capture the details of a particular genetic circuit, but rather to explore the generic properties of networks with similar structures; this is the first and still best known example of the "ensemble approach" to the study of these kinds of complex systems. This approach looks for widespread system properties, by examining the behaviour of statistical ensembles of networks, which share some property (e.g. number of nodes, average connectivity per node) but which are otherwise generated at random. The inventor of the RBN model, S. Kauffman, examined how the number of attractors (which are cycles in finite networks) and their typical length scale with the number of nodes (Kauffman, 1993). By comparing these data to those concerning the way how the number of different cell types, and the typical cell cycle length scale with the total DNA content in organisms belonging to different phyla, he found an interesting similarity (see below).

Recently, we have analyzed data concerning the response of S. Cerevisiae cells to gene knock-out (Serra, Villani and Semeria, 2004). In each of these experiments a single gene is silenced, and the expression levels of the other genes are measured (by comparison with the expression level of the same gene in the unperturbed cell). We introduced different statistical measures of the cell response to perturbations, and computed them on the available data of S. Cerevisiae. The results (briefly summarized in section 3) show surprisingly good agreement on the first order statistics. Second order statistics show some differences that can be explained and which point to the opportunity of studying also networks with higher connectivity and with different topology.

As far as this latter aspect is concerned it is interesting to observe that the topology may affect the system dynamical properties. In particular, it is

possible to modify the network construction algorithm in such a way as to obtain a scale-free (i.e. power law) distribution of outgoing links, instead of the usual poissonian distribution. This does not seem a purely mathematical exercise, since in many natural and man-made networks such a scale-free topology has been actually found. It is interesting to observe that such a change in topology has a great effect on the dynamics, i.e. on the number of attractors, cycle length and transient duration (Serra, Villani and Agostini, 2004). The scale-free network presents a much more ordered behaviour than the corresponding poissonian network: the main results concerning this aspect are briefly reviewed in section 4. Of course the issue whether real genetic networks are more closely described by scale-free or by random networks is still open (other possibilities cannot also be excluded).

Due to the importance of the properties of attractors in RBN, it is important to ascertain how robust their scaling properties are with respect to some changes in the model. An important point is the updating strategy, since the synchronous updating used in the original model does not seem particularily realistic. RBN with asynchronous updating have been investigated in (Harvey and Bossomaier, 1997), where is introduced the notion of "loose attractors". By concentrating on the easier task of establishing how the number of fixed points scale with the number of nodes, these authors came to the rather surprising result that the number of fixed points seems to stay almost constant. By exploring larger networks we show in section 5 that this result is true only up to a certain network size, and that above that size the number of fixed points declines sharply.

The observations of sections 3-5 show that Random boolean networks, although very simple, represent a useful tool to explore the behaviours of large genetic networks, both from the viewpoint of the application to real gene networks and for system theoretical reasons.

2. RANDOM BOOLEAN NETWORKS

Let us consider a network composed of N genes, or nodes, which can take either the value 0 or 1. Let $x_i(t) \in \{0,1\}$ be the activation value of node i at time t, let $X(t)=[x_1(t), x_2(t) \ldots x_N(t)]$ be the vector of activation values of all the genes, and let there be a directed link, from node A to node B, if the product of gene A influences the activation of gene B. To each node a (constant) boolean function is associated, which determines the value of its activation at time $t+1$ from those of its input nodes at time t. The network dynamics is discrete and synchronous, so all the nodes update their values at the same time. In a classical RBN each node has the same number of

incoming connections kin, while the other terminals of the connections are chosen with uniform probability among the other nodes.

In order to analyze the properties of an ensemble of random boolean networks, different networks are synthesized and their dynamical properties are examined. The ensembles differ mainly in the choice of the number of nodes N, the input connectivity per node k, and the choice of the set of allowed boolean functions. While individual realizations may differ markedly from the average properties of a given class of networks, one of the major results is the discovery of the existence of two different dynamical regimes, an ordered and a disordered one, divided by a "critical zone".

In the ordered region, attractors tend to be stable, i.e. by flipping one of its nodes it often relaxes back to the original attractor. Moreover, both the number of attractors and their typical period scale slowly with the network size (i.e. as a power of the number of nodes N). Also their basins of attractions are regular, so it often happens that two initial states which are very close to each other tend to the same attractor.

In the disordered state the attractors are often unstable, close states usually tend to different attractors, and the typical duration of a cycle attractor scales exponentially with the network size. The border between the ordered and the disordered regime depends upon the value of the connectivity per node k and upon how the boolean functions are chosen. Kauffman has proposed that real biological networks are driven by evolution in an ordered region close to the border between order and disordered regimes (the "edge of chaos") (Kauffman, 2000).

The scaling properties of the average number of attractors and average cycle length with the number of nodes N in this region have been compared (Kauffman, 1993) to actual data concerning the dependence, upon the total DNA content, of the number of different cell types (which should correspond to the number of attractors) and of the duration of the cell cycle (which should correspond to the typical length of the attractor cycle). The agreement appears satisfactory for data that span several orders of magnitude, over different organisms belonging to different phyla.

3. THE SIMULATION OF GENE KNOCK-OUT EXPERIMENTS

In a series of interesting experiments, Hughes et al (Hughes et al., 2000) have measured with cDNA microarray techniques, the expression profiles of 6312 genes in *Saccharomyces Cerevisiae* subject to 227 different gene knock-out experiments (ie silencing of selected genes, one at a time). In a typical knock-out experiment, one compares the expression levels of all the

genes in cells with a knocked-out gene, with those in normal ("wild type") cells. Since microarray data are noisy, it is required that a threshold be defined, such that the difference in expression level (between the knocked-out and the wild type cell) is regarded as "meaningful" if the ratio is greater than the threshold θ (or smaller than $1/\theta$) and neglected otherwise.

In order to describe the global features of these experiments, two important aggregate variables are the so-called avalanches (which measure the number of genes whose expression level has been modified in a knock-out experiment) and susceptibilities (which measure in how many different experiments a single gene's expression level has changed). For a precise definition see (Serra, Villani and Semeria, 2003, 2004).

In order to simulate the dynamical features of gene regulatory networks, model boolean networks with a high number of genes were computer-generated and several simulations were performed, aimed at reproducing the experimental conditions. For reasons discussed in detail in (Harris et al., 2002), the study concentrated upon networks with input connectivity $k_{in} = 2$, which lie in the ordered phase, and particular attention was paid to networks where the non-canalyzing functions (which are XOR and NOT XOR in the two-input case) as well as the NULL function are not allowed. Knocking-out was simulated by clamping the value of one gene to 0 in the attractor cycle with the largest basin of attraction.

In order to compare perturbed and unperturbed samples, for each gene the ratio between the expression level in the perturbed sample and the expression level in the unperturbed sample is computed. When dealing with oscillating genes, the expression level is equal to its average expression level taken over the period T of the attractor of interest. Since boolean functions are not well suited to deal with threshold effects, in the simulation every gene whose expression level is different in the two cases (perturbed and unperturbed) is considered as "affected" (i.e. the threshold for synthetic networks is equal to zero).

The main results can be summarized as follows (for further details and comment see (Serra, Villani and Semeria, 2004):

- the distributions of avalanches and susceptibilities are very similar in different network realizations. This robust behaviour is different from most properties of RBN (eg number and length of attractors) which vary largely in different networks. The importance of this observation is that it points to these distributions as candidates to be robust properties, largely unaffected by the network details
- the average distribution of avalanches in synthetic networks is definitely close to the one observed in actual experiments, except for the smallest avalanches, i.e. those of size 1. Synthetic networks overestimate the fraction of such small networks with respect to biological data

- the average distribution of susceptibilities in synthetic networks is even closer (than that of avalanches) to the one observed in the experiments.

The agreement is indeed surprising, even more so if one realizes that there are no adjustable parameters here (there is indeed one parameter in the data preparation, i.e. the threshold chosen on the change in expression level, above which a gene is considered as modified; for a discussion see (Serra, Villani and Semeria, 2003)). It is therefore important to perform further and more stringent statistical analysis of the response of gene expression levels to the knock-out of single genes, in order to compare the behaviour of model and real networks.

Meaningful measures of "distances" between the expression patterns of pairs of genes, and between the expression profiles in different experiments were introduced in (Serra, Villani, Semeria and Kauffman, 2004), and the results for both real and synthetic networks were compared. The similarities are remarkable also in this case, but there are also some differences which can be (tentatively) explained by the facts that

- simulated networks present a much higher fraction of small avalanches of size 1 (roughly 30% of the total) than those which are found in real data (15%).
- the maximum value of the susceptibility is smaller in synthetic than in real networks.

It can be conjectured that the difference between the number of small avalanches in the two cases is related to the fact that the number of incoming connections, k, is held fixed to the value 2 in the simulations. This is certainly not a biologically plausible assumption, since it is well known that there are some genes whose expression can be influenced by a higher number of other genes. It would be therefore interesting to explore the behaviours of networks with higher connectivities, although in this case the computational load associated to the simulation of a network with 6000 nodes would be much heavier.

Moreover, the presence of unusually highly susceptible genes in biological networks suggests the opportunity to test models where the constraint that all the genes have the same number of input is removed. It would be interesting to analyze the behaviour of random boolean models with exponential or scale-free connectivities.

4. THE DYNAMICS OF SCALE-FREE RBN

In the classical RBN model (shortly, CRBN), the number of incoming links is the same for every node, while outgoing connections turn out to follow a Poissonian distribution; however, there is growing evidence (Amaral et al., 2000) that many natural and artificial networks actually have a scale-free topology, where the nodes may have different connectivities, with a power law probability distribution $p(k) \sim k^{-\gamma}$. It has been observed that the dynamical properties of nonlinear systems may be affected by the topology of the corresponding networks (Serra and Villani, 2002; Strogatz, 2001). Scale-free RBN have also been studied with analytical methods and simulation, and it has been shown that their dynamical properties differ from those of classical RBN (Aldana, 2002; Fox and Hill, 2001). An interesting finding is that the region of parameter space where the dynamics is ordered is larger in these models than in classical RBN.

These studies have concerned a scale-free distribution of incoming links. However, since most previous studies of RBN concern the case where kin is the same for all the nodes (Kauffman, 1993), a proper comparison of scale-free vs. poissonian random networks could be better performed by fixing the value of kin equal for all the nodes and introducing a scale-free distribution of the outgoing links. In the model, which has been called SFRBN (scale-free RBN), there are kin incoming links for each node, exactly like in the original RBN, but the distribution of the other terminals is scale-free. The boolean function for each node is chosen at random, as in RBN. The algorithm for generating such a scale-free distribution of outgoing links has been presented elsewhere (Serra, Villani and Agostini, 2004).

Using this algorithm, extensive simulations have been performed, which have shown an impressive difference of the dynamics of SBRN, compared to that of the corresponding CRBN. In particular, it has been shown here that the change in the topology of the outgoing connections causes profound modifications in the phase portrait of the system:

- the number of attractors is much smaller
- this number is almost independent of the network size for networks up to 20.000 nodes
- the period of asymptotic cycles is much shorter, and grows very slowly with the network size
- the duration of the transients are also shorter than in classical RBN.

Further investigations should be performed to confirm these findings for different values of the number of ongoing connections k_{in}. It should be interesting to analyze also the effects of the introduction of a cutoff in the maximum number of allowed links per node. Despite these limitations, this

work already provides a clear indication concerning the existence of a mechanism potentially able to control the behavior of a growing and sparsely interconnected system.

5. ASYNCHRONOUS UPDATING

The importance of the number of attractors, and of their scaling properties has already been remarked. The CRBN model is based upon synchronous updating, which is certainly not a realistic description of how genes work. So one should consider the way how the phase portrait changes if this assumption is relaxed. Unfortunately, in the case of asynchronous updating the very notions of attractor and basin of attraction need to be reconsidered, in the most interesting case where the node to be updated is chosen at random at every time step. Suppose that the system has settled to an attractor state, and consider a state $X = X(t)$ and its successor $X' = X(t+1)$. Now suppose that at further time $t+q$ the system is found again in state X: $X(t+q) = X$. Since the node to be updated is chosen at random and independently, in general $X(t+q+1)$ will be different from X'. Moreover, it is well known that, with random asynchronous updating, some initial states may evolve to one attractor or another depending upon which node is updated first (Serra and Zanarini, 1990).

In order to deal with these difficulties, Harvey and Bossomaier (Harvey and Bossomaier, 1997) introduced the notion of a "loose" attractor, which is a set of points that may entrap the system after transients have died out. It is much harder to identify and count loose attractors than usual attractor cycles, but this difficulty disappears if one considers fixed points only. It has often been found that systems with asynchronous updating tend to have a larger proportion of attractors that are indeed fixed points (Serra and Zanarini, 1990). Harvey and Bossomaier (Harvey and Bossomaier, 1997) therefore studied how the number of fixed points scales with the network size, and came to the rather surprising conclusion that the number is almost constant. Since the average period of cyclic attractors in CRBN grows with the network size, this may point to a real difference between the phase portraits in the two cases.

However, we have recently shown (Serra, Villani and Benelli, in preparation) that the number of fixed point attractors in asyncronous networks actually decreases sharply with the network size, after a certain size has been exceeded, so the approximate constancy of the number of fixed point attractors that was found by Harvey and Bossomaier holds only for sufficiently small networks.

6. CONCLUSIONS

We have summarized here some recent results, and presented a new result, concerning different properties of random boolean networks. Although they are oversimplified models of real genetic networks, it has been shown that they can describe with a good approximation some key statistical features of the response of biological cells to perturbations. The use of such simplified models may provide good hints for finding properties which are robust with respect to the model details, and knowing these "generic" properties would be of primary importance in order to understand the robustness of real biological systems. Even if in the end it would be proven that the role of generic properties (if any) is very limited, their search would provide a heuristic guide for meaningful experimentation.

A further remark is related to the methodology: since we are dealing with models which are oversimplified, it is necessary to look for properties which are not too critically dependent upon the peculiar features of a particular kind of model (or otherwise to provide good reasons why one has to choose exactly that member of the set of similar models). This is why the analysis of the dynamical properties in the case of different topologies, different updating (and other modifications of the basic model of CRBN) is important in studying genetic networks.

REFERENCES

Aldana, M., 2002, Dynamics of Boolean Networks with Scale-Free Topology, (available at: http://arXiv:cond-mat/0209571 v1).

Aldana, M., Coppersmith, S., and Kadanoff, L. P., 2003, Boolean dynamics with random couplings, in: *Perspectives and Problems in Nonlinear Science*, E. Kaplan, J. E. Marsden and K. R. Sreenivasan, eds., Springer, (Also available at http://www.arXiv:cond-mat/0209571.

Amaral, L. A. N., Scala, A., Barthelemy, M, and Stanley, H. E., 2000, *Proceedings of the National Academy of Sciences USA* **97**:11149-11152

Fox, J. J., and Hill, C. C., 2001, From topology to dynamics in biochemical networks, *Chaos* **11**:809-815.

Harris, S. E., Sawhill, B. K., Wuensche, A., and Kauffman, S. A., 2002, A model of transcriptional regulatory networks based on biases in the observed regulation rules, *Complexity* 7:23-40.

Harvey, I., and Bossomaier, T., 1997, Time out of joint, attractors in asynchronous random boolean networks, in: *Proceedings of The Fourth European Conference on Artificial Life (ECAL97)*, P. Husbands and I. Harvey, eds, MIT Press, Massachusetts, pp. 67-75.

Hughes, T. R., et al., 2000, Functional discovery via a compendium of expression profiles, *Cell* **102**:109-126.

Kauffman, S. A., 1993, *The origins of order*, Oxford University Press.

Kauffman, S. A., 2000, *Investigations*, Oxford University Press.

Serra, R., and Villani, M., 2002, Perturbing the regular topology of cellular automata: implications for the dynamics, Springer Lecture Notes in Computer Science 2493 168-177.

Serra, R., Villani, M., and Agostini, L., 2004, On the dynamics of random Boolean networks with scale-free outgoing connections, *Physica A* (in press).

Serra, R., Villani, M., and Semeria, A., 2003, Robustness to damage of biological and synthetic networks, in: *Advances in Artificial Life*, W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim and J. Ziegler, eds., Springer, Heidelberg, pp. 706-715.

Serra, R., Villani, M., and Semeria, A., 2004, Genetic network models and statistical properties of gene expression data in knock-out experiments, *Journal of Theoretical Biology* **227**:149-157.

Serra, R., Villani, M., Semeria A., and Kauffman, S. A., 2004, Perturbations in genetic regulatory networks: simulations and experiments, (Submitted).

Serra, R., and Zanarini, G. 1990, *Complex Systems and Cognitive Processes*, Springer, Heidelberg.

Strogatz, S. H., 2001, Exploring complex networks, *Nature* **410**:268-276.

Wagner, A., and Fell, D., 2000, *The small world inside large metabolic networks*, Tech. Rep. 00-07-041, Santa Fe Institute.

COLOR-ORIENTED CONTENT BASED IMAGE RETRIEVAL

Guido Tascini, Anna Montesanto and Paolo Puliti
Dipartimento di Elettronica, Intelligenza Artificiale e Telecomunicazioni
Università Politecnica delle Marche, 60131 Ancona, Italia

Abstract: The aim of this work is to study a metrics that represents the perceptive space of the colors. Besides we want to furnish innovative methods and tools for annotate and seek images. The experimental results have shown that in tasks of evaluation of the similarity, the subjects don't refer to the most general category of "color", but they create subordinate categories in base to some particular color. Those categories contain all the variations of this color and also they form intersections between categories in which any variations are shared. The perception of the variations is not isometric; on the contrary that perception is weighed in different manner if the variations belong to a particular color. So the variations that belong to the intersection area will have different values of similarity in relation to the own category. We developed a system of color-oriented content-based image retrieval using this metrics. This system analyzes the image through features of color correspondents to the own perception of the human being. Beyond to guarantee a good degree of satisfaction for the user, this approach furnishes a novelty in the development of the CBIR systems. In fact there is the introduction of a criterion to index the figures; it is very synthetic and fast.

Key words: color perception; non-isometric similarity metrics; human subjects; content based image retrieval.

1. INTRODUCTION

The recognition of an object as similar to another, therefore as belonging to the same category, depends on cognitive strategies that are extremely effective to gather constancy, invariance and regularity. The perceptive representation is not necessarily holistic; it could be a schematic appearance of the perceptive state extracted through the selective attention and stored in

the long-term memory. Establishing the criterions in base to which judge the degree of similarity between two or more objects is not simple. Our mind uses something more complex of only one distinctive feature. During this processing the so-called relational and structural features are examined. As regards the similarity between the elements it could be derived (Smith, Shoben and Rips, 1974) from a precise feature (e.g. an ellipse) or from a "configuration" (a particular rule that is followed in arranging the distinctive features).

The similarity between the objects could not be attributed considering the objects separately: for instance an ambiguous object loses his ambiguity in the moment in which is compared (Medin, 1993). Naturally also the factors stimulus and task could address the formation of the concept of similarity forcing the subject to consider only some features of the objects. The information around the similarity between two objects is well represented through the conceptual spaces (Gardenførs, 1999), representations based on notions that are both topologic a geometric. At the conceptual level, the information is relative to a domain that could be represented from the "quality dimensions" that form the conceptual spaces.

The similarity could be seen like a distance in the conceptual spaces. For instance in the representation of the colors based on "Hue, Saturation and Brightness," there exist two vertices, that is "White" and "Black", but inside the concept of "Skin" the "White" and the "Black" are not those absolute. In this situation is like if to the inside of the general fuse of representation of the colors there is an other small fuse that has to his inside all the representation of the colors (red, yellow, white, black, ...) but only reported to the concept of "skin." A valid measure for that kind of representation of the information is the scaling (Torgerson, 1965; Nosofsky, 1991): it allows representing from a metric point of view the psychological measures. If an object A has judged similar to B for the 75% of the cases, and the object C has judged similar to B for l'85% of the cases, with the scaling we could establish than the distance AB is bigger than the CB distance: through the law of the categorical judgments of Torgerson (1965).

Various fields, like art, medicine, entertainment, education, and multimedia in general, require fast and effective recovery methods of images. Among these is Content Based Image Retrieval, in which images are described not by keywords but by content. A main approach is using low-level characteristics, like color, for segmenting, indexing and recovering. This work presents a method for annotating and recovering images that uses a new evaluation method for the similarity between color hues that corresponds to human color perception. In addition a fast and effective method for image indexing, is presented. In literature many methods are presented to this aim. A simple and fast method is based on a set of key

words that describes the pictorial content (Y. Rui, and al. 1999). The drawbacks of this approach are various: the method is hard for the big databases; the quality of key words is subjective; the search by similarity is impossible. A more general approach to multimedia recovering is different from those based on visual or acoustic data. The main difference depends on extraction of features. A popular approach is the Query By Example (QBE), where the query is a key object, in particular an image, in the database or depicted at query time. The content-based methods allow recovering images by the visual language characteristics, like similarity, approximation and metric relations, research key as figures, structures, shapes, lines and colors. As consequence they are many modes of indexing, storing, searching and recovering visual data. More refined are the methods in which the images may be analyzed in the query phase; the corresponding software is called: Content Based Image Retrieval (CBIR) Systems. As the Query by Color, two types of approaches are important: 1) retrieval of images with global color distribution similar to the query image one, interesting for the pictorial data bases; 2) recovering of an object in a scene, using its chromatic features. (Smith, 1997) We will briefly describe some of most popular CBIR. QBIC, that means Query By Image Content (Flickner and al., 1995), uses various perceptual characteristics and a partition-based approach to the color. Introduces the Munsell transformation and defines a color similarity metric (Bach and al., 1996). The system is limited in the search of spatial characteristics. Virage (Bach and al., 1996) that supports the query about color, structure and spatial relations operated on the following four primitives: Global Color, Local Color, Structure e Texture. Photobook (Pentland, 1996) is an interactive set of tools developed at M.I.T. Media Laboratory on the Perceptual Computing. The system interacts with user by Motif interface. The matching is performed on the feature vectors extracted by considering invariance, scaling and rotation. VisualSEEk (Smith and al., 1996a, Smith and al., 1996b) and WebSEEk (Smith et al. 1997) are academic information systems developed at the Columbia University. VisualSEEk is a hybrid image recovery system that integrates the feature extraction using the color representation, the structure and the spatial distribution. The recovering process is enhanced with algorithms based on binary trees. WebSEEk instead is a catalogue-based engine for the World Wide Web; it accepts queries on visual properties, like color, layout correspondence and structure. ImageRover (Sclaroff and al., 1997) is an image recovery tool developed at the Boston University. This system combines visual and textual queries for the computation of the image decompositions, associations and textual index. The visual features are stored in a vector, using color and histograms texture-orientation; the textual one are captured by using the *Latent Semantic Indexing* on the association of

the words contained in the HTML document (La Cascia and al. 1998). The user refines the initial query using the *relevance feedback*. The Munsell color space is a three-dimensional polar space, the dimensions being Hue, Value and Chroma. Value represents perceived luminance represented by a numerical coordinate with a lower boundary of zero and an upper boundary of ten. Chroma represents the strength of the color, the lower boundary of zero indicating an entirely achromatic color such as black, white or grey. The upper value of Chroma varies depending upon the Value and Hue coordinates. The Hue dimension is polar and consists of ten sections that are represented textually each with ten subsections represented numerically. Our work considers the only dimension Hue, while maintains constant the other 2 variable Saturation and Intensity. Differently from the Munsell color space, it considers the space of this single dimension 'not-uniform' and 'not-linear' that is 'not-isometric'. A main difference with the Munsell space is the following: our work do not evaluates the belonging of a color to a 'nominal' category, that may be invalidated also by conceptual structures related to social history of examined population. We evaluate how much it is similar to a target-color a variation of it, performed in the only hue dimension. Then the not-linearity is related to the color similarity and not their categorization. The results are related to the subject judgments on the similarity evaluation between two colors, and not on the hue. If the Munsell space is perceptually uniform and linear, then a variation Δh of hue would be proportional to the related similarity variation Δs: the results have shown that this direct proportionality, between hue-variation and similarity-variation, of two colors do not exists.

2. THE SIMILARITY FUNCTION

We perform an experiment with multiple conditions within the subjects, to determine the form of the function that ties the independent variable (hue of the color) with the dependent variable (similarity). To appraise the similarity with the image target we present to the subjects 21 similar images, but they are perturbed in the color. Such images maintain spatial fixed disposition. For color and spatial fixed disposition is understood the same of the target image and for perturbation is understood the graduation of the color (clustering). Everything is repeated for the three colors used in the search: the yellow, the red and the blue. They are the three fundamental colors from the point of view of the physical pigment.

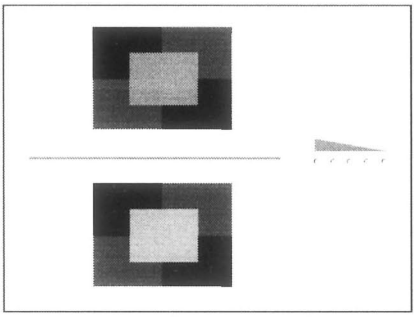


Figure 1. The test image for the perception of the colors.

2.1 Human evaluation of the similarity

We use the images shown in the Fig. 2 for an experiment on color similarity evaluation. They vary only the color of the central rectangle. To the subject are show simultaneously the image target and his variation. These 63 couples of images are randomized for each proof (each set of 21 corresponds to a relative target image). The subject sees all the mixed images that are not subdivided for typology of basic color. The subject have to attribute a coefficient of "similarity to the target" of the images using scale from 1 to 5: 1 = nothing, 2 = very little, 3 = little, 4 = enough and 5 = totally. Once submitted the test of evaluation of the similarity to 12 subjects, we compute the frequencies of attribution of the 63 stimuli at the 5 categories. The distributions of the average answers of the subjects are not homogeneous and very asymmetrical.

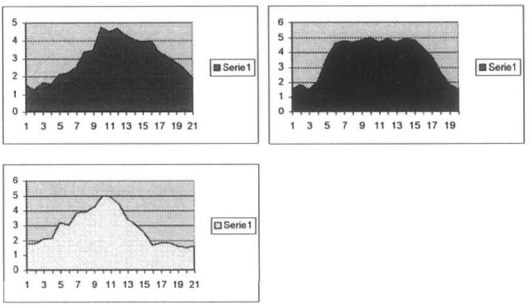


Figure 2. The three graphic represent the frequencies of attribution to the categories of similarity.

Accordingly to choose the corresponding values to the colors representing respectively the 80% and the 60% of the answers the procedure is the following:

- we take the average of the reference value. For instance for the yellow with hue 42, the corresponding value was of 4.92; for the Blue with hue 165 was 5 and for the Red with hue 0 was 4.5.
- we compute the 80% and the 60% of such middle values to have a homogeneous reference to the inside of the asymmetrical curves.

The resultant values from this analysis represent the colors that could be definite perceptively similar, as regards the target color, at 80% and at 60% both increasing that decreasing the hue. We have to notice as each color already has, to the level of the simple frequencies, a different representation of the similarity. Those also if in theory the variation step of the color was originally equal for the three colors.

3. ONE-DIMENSIONAL SCALING

The judgment on similarity for each color has a frequency trend, which may be viewed as a cognitive representation of the different colors. The application of the one-dimensional scaling gives us a measure of the distance between variation steps for each color.

The single dimension choice depends on the assumption that the only variation is that one of the hue, while the others two dimensions are fixed. The scaling is applied to each color, so giving different measurement scales for the Red, the yellow and the Blue.

In this work we assume that the difference among colors depends on: 1- the physical property of the bright energy, 2- psychological perception and 3- different entity of the same stimuli. The colors could not be measured on the same scales, but they are subject to some transformations by considering the sense of category belonging.

3.1 Results of the scaling

From the one-dimensional scaling we obtain the quantification of the qualitative data of the first level: by comparing the graphic of the two levels underline their homogeneity. We notice that the graphic of the first level is different from those of the second one: this last shows more precision in the metrics.

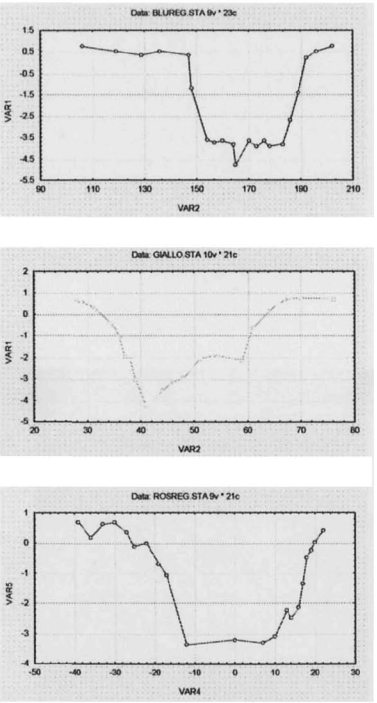


Figure 3. The representation of the results of the one-dimensional scaling: the curves show the relationship between the value of similarity deriving from the scaling and the HSI hue.

4. THE SIMILARITY FUNCTION

It is necessary now to find a model for representing the relationship between the HSI hues and the related values of similarity. A not linear regression is used to find the function able to interpolate the similarities points (y, dependent variable) in relationship with the hues (x, variable independent). The similarity function distribution, based on hue, is a polynomial. The value of the single parameters depends on the derivation; each color will weigh the similarity value of hue variation in different manner for the different values of the parameters.

The resultant functions are the followings:

Yellow: $y = -.00026\,x^3 + .046504\,x^2 - 2.6155\,x + 43.9572$

Blue: $y = .000061\,x^3 - .02622\,x^2 + 3.64571\,x - 163.38$

Red: $y = .000001\,x^3 - .00096\,x^2 + .198553\,x - 4.2372$

We underline the discrepancies between the form of the functions and the real disposition of the points in the space; to overcome this problem we have defined some broken functions to have a better representation of the data. So we have different range of hue (h= hue) and they are partitioned in three fundamental color ranges: a) Blue: $100 < h < 204$; b) Yellow: $26 < h < 78$; c) Red:- $42 < h < 22$, that coincides to $(0 < h < 22)$ OR $(198 < h < 240)$.

These functions allow, given a hue value with saturation 240 and brightness 120, to determine to which color this hue is similar and how much is similar in comparison with the reference color. For example:

- $H = 0 \Rightarrow$ Red with sim.= -3.199093;
- $0 < h < 26 \Rightarrow$ Red with sim.= $0.016065\ h^2 - 0.16173\ h - 3.1260$;
- $25 < h < 42 \Rightarrow$ Yellow with sim.= $-0.03198\ h^2 + 1.83212\ h - 25.677$.

5. DEVELOPMENT OF THE COLOR-ORIENTED CBIR

A second aim of our work is the development of a system for image recovering based on visual content (Content Based Image Retrieval, CBIR). By considering that a retrieval based on the RGB metric space usually is not satisfactory for the user, we will go toward a metrics representative of the perceptual color space.

The two main steps of a CBIR system are

- the decomposition of the image in a suitable data structure;
- the comparison between these data structures for detecting similar images.

The first step is related to the image segmentation, and for this we use a QuadTree approach. The second step includes the idea of similarity matching'.

5.1 Construction of the similarity_QuadTree

We define a Similarity_QuadTree that is a quaternary tree structure that decomposes the whole image in 4 quadrants, each quadrant again in 4 sub quadrants and so on. When we reach an homogeneous quadrant we stop our subdivision: the leaves so obtained are aimed to contain the similarity values of the three colors respect the related three reference node different from leaves contain the pointers to the children by adopting for each quadrant the convention of the four cardinal points: NO, NE, SE and SO. We use three colors of the RYB space (Red, Yellow, and Blue), while the three reference

values are extrapolated during the development of the metrics based on the perceptual color space:

- Red: ⇒ ref. = -3,199093;
- Yellow: ⇒ ref. = -3.544842;
- Blue: ⇒ ref. = -4.768372 .

The procedure of compute the similarity computation receives a leaf node of the RGB_QuadTree in input extracts the three RGB values of it and extrapolates the value of the hue (h). We check the belonging range of the hue, and then we calculate the similarity to the color by using the suitable features that are obtained from the perceptual color space. The numerical values 0, 1, 2 has used for memorizing the color by pointing out respectively red, yellow and blue (figure 4).

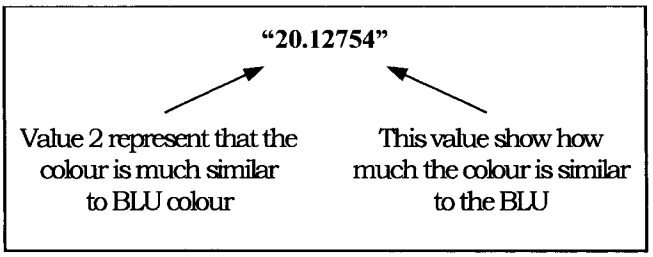


Figure 4. Example of what is contained on the knot leaf.

If the drawn out RGB color from the RGB_QuadTree tree, is not perceptually valid (for instance little saturated or too bright), it is not inserted inside of the leaf of the similarity value. We bring again the method that calculates the hue of the given RGB color in input, now by maintaining a constant saturation of 240 and a brightness of 120, suitable with 1 and 0.5 in relative coordinates. The main procedure that concerns the similarity tree creation receives in input the roots of the trees: RGB_QuadTree and Similarity_QuadTree. When the RGB_QuadTree has reached a leaf node, then the procedure insert the value of similarity in the Similarity_QuadTree. If it is not in a leaf node then the four children of the quadtree are created.

Now in the figure 5 we show an example of the decomposition of an image with the quadtree. We analyze the first leaf node that is indexed always from No, that in the image corresponds to the part aloft to left, his value is " 0-3.199093," that is:

- the first term " 0" points out that the red color is dealt with;
- the negative number "-3.199093" corresponds to the value of similarity, which in this case is quite the reference one.

If we take in consideration an other node, for instance that indexed from SE- SE- SE- SO, that is the blue quadrant of small dimension, it has a value of " 2-4.05779," where:

- the first term " 2" points out that the blue color is dealt with;
- the negative number "-4.05779" corresponds to the value of similarity, that is very near to the reference one.

The nodes of the figure "Low Saturation" they correspond to situations of white, black or grey; therefore we will find again them in all the cases in which intensity and brightness are in border conditions.

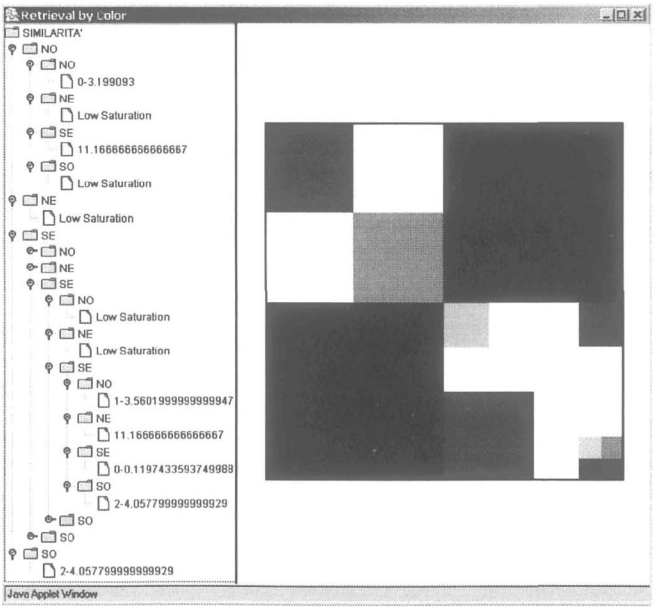


Figure 5. Example of the creation of a Similarity_QuadTree.

5.2 Label Extraction of the perceptually valid color

The visit of the similarities tree of the colors of the image respects at three references is useful to extrapolate the features:

- the average value of the similarities as regards the red color;
- the average value of the similarities as regards the yellow color;
- the average value of the similarities as regards the blue color;
- the increasing order of the colors respects the quantity.

This method allows extracting and understanding the content of the leaf node. The knot leaf is transformed in a vector containing in the zero position the label of the color (0: red, 1: yellow, 2: blue), while in the one position the value of similarity to this color. Now we compute the number of pixels that

belong that quadrant, because such number becomes smaller and smaller by increasing the decomposition level of the quadtree. The number of these pixels is in a vector:

- position " 0" for the red pixels;
- position " 1" for the yellow pixels;
- position " 2" for the blue pixels.

The average similarity is computed and then memorized it in a suitable vector. In case of color we assign an unlikely value of similarity, for example 10000. For the calculus of the order of the colors, we count the percentage of pixels that belongs each color by discarding all the pixels considered in "Low Saturation". Then we use the three methods maximum, medium and minimum to origin to the vector order, which contains the three values 0, 1, 2, organized in an increasing way.

6. INDEXING THE WHOLE DATABASE

We associate with the database of imagines a file containing the data deriving from the segmentation. This file, which is created before the retrieval allows a rapid access when these features are required. The structures of file data will be composed by a sequence of information.

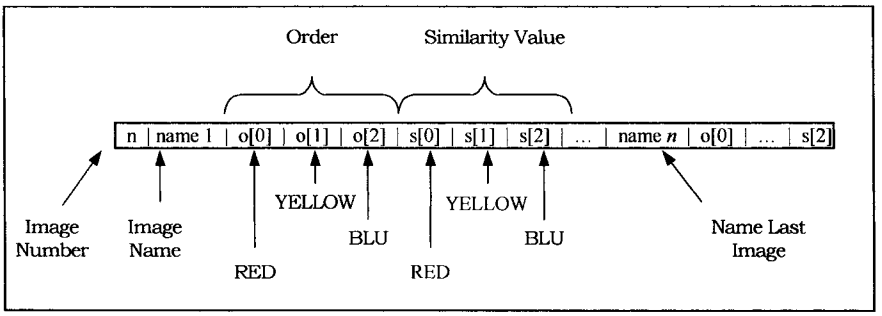


Figure 6. The structures of file data.

In conclusion our software segment the whole database, and points out that all the components images of the same database may be analyzed through matching and sorting procedures that draw out fundamental properties of the color. Here is described the methodology of selection of the image to use as query for the retrieval, with which compare the property of the figures of the database to get a matching on the similarity. Our software

gives the possibility of designing an image through a Java application. The user also can insert an image or a photo to use like query.

6.1 Matching and Sorting

We compute a score that identifies the similarity of a generic image in comparison with the select one as search key by assigning a value that is understood from 0 (image completely different), to 10000 (image perfectly similar). This last value is divided in the following way:

- 4000 points derive from an appropriate comparison respect the order of quantity of the principal red, yellow and blue colors;
- 6000 points result from the difference between the similarities of the two images for every main color.

As the first point the two variables belong colorOrd and vectOrd that represent the vectors that describe which is the order of the colors in relationship to the quantity. If the first two colors have identical quantity, then we assign, to the variable score, 2000 credits, while, if also the seconds are equal then we assign 1250 credits, and so on. This part has conceived to overcome the situation like that one that's considers very similar two images with pixel of a certain color similar to the average of all the colors of the other. The more consistent part of the score is accredited in the second part, where firstly the relative differences of the similarity of the three colors are computed. Then a weighed quantity formula is applied, that increases the value of the variable score. The real matching is the method that is concerned with the composition of the vectors, aimed to contain the property of the images organized by similarity. By analyzing the code we find four important vectors: a vector containing the score of sorted similarity in an increasing way, a vector containing the runs that identify the sorted images, a matrix containing the three average values of similarity (k = 0 red, k = 1 yellow, k = 2 blue) sorted respect the image like from point 1 and a matrix containing the three values of the order of quantity (k = 0 red, k = 1 yellow, k = 2 blue) sorted respect the image. The dimension of these vectors depends on a constant [max], which decides how much images show to the user in the phase of visualization. After the computation of score similarity, of the contained images in the file that represents our database, we go on for searching the optimal position for inserting such score in the vector of the query image. If this position is valid, so inferior to max, then we insert the relative data score, name, similarity and order in the respective four vectors.

7. PRACTICAL RESULTS

We use two kind of image like keys of search in a small database. A subject gives a perceptive judgment on the similarity of the images to the keys, by concluding retrieval: this may be considered the optimal result that our program should reach. Then we perform the same two searches using our software. Fig. 7.1 and in Fig. 7.2 show the images choices as keys of search. The subjects assign a score from 1 to 10 for each element of the database, where 10 stand for the best similarity and 1 for the smallest one. The total score associates a meaning degree of similarity to all database respect to the key image. Figure 8 shows the database images used in the example.

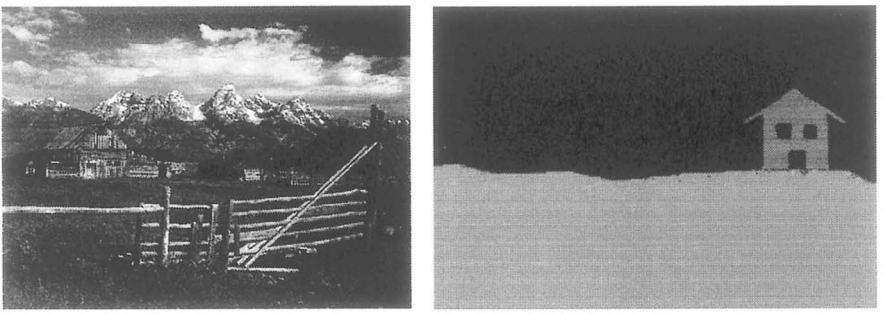


Figure 7. Image key 1 and Image key 2.

7.1 Comparison between human decision and software

We want to compare the results obtained with human subjects and with the software we examine the two tests separately.

In the Table 1 it is listed the numbers of the sorted figures, departing from the more similar to the key image 1 up to the less similar. The score that we find in the last column it is the associated value to the position of the orderly images from our CBIR system as regards the order of the same result from the Test. The used rule is that of assign 10 points to the images that they are in the same row, while two points for each line of distance are scaled. The last row corresponds simply to the percentage calculus of the score, therefore represents the empirical evaluation of the percentage of correctness in the measure of similarity of the software.

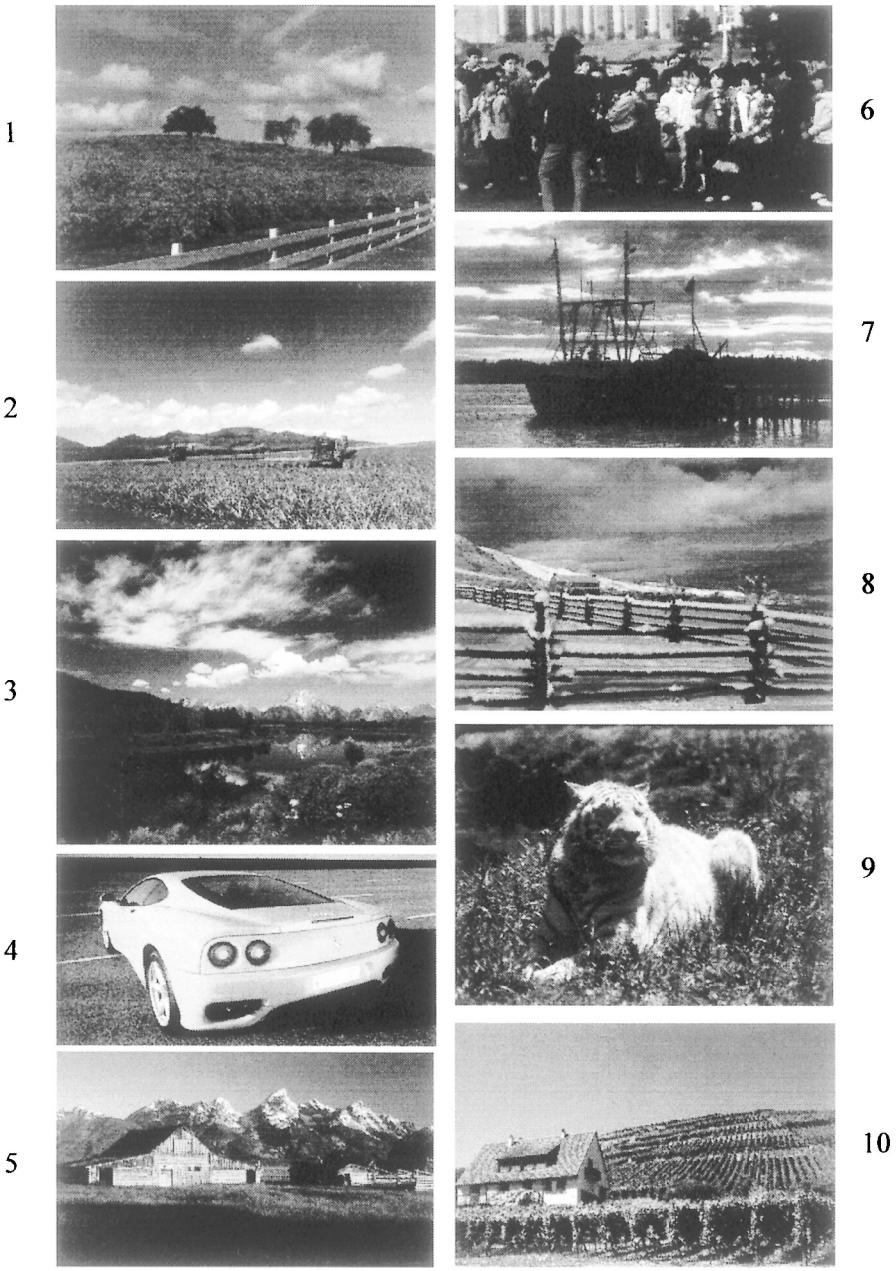


Figure 8. The images that constitute the database.

Table 1. Results of testing the two key images.

Image key 1				**Image key 2**			
Similarity Order	Image n°		Score	Similarity Order	Image n°		Score
	Human	Software			Human	Software	
1	5	5	10	1	10	10	10
2	3	3	10	2	2	2	10
3	2	2	10	3	5	3	8
4	1	10	8	4	3	5	8
5	10	1	8	5	1	9	8
6	8	8	10	6	8	1	6
7	9	6	8	7	9	6	6
8	6	9	8	8	7	8	8
9	7	4	8	9	6	7	6
10	4	7	8	10	4	4	10
Percentage of fidelity:			88%	Percentage of fidelity:			80%

This technique allows to have a range of percentages from 0 to 100 points because allows of assign also score negative, therefore the value gotten like a valid respect of the correctness of the created program could be considered. In tab.1 we compared the two methodologies using first as key the photo of the test 1. In this case we obtain a score of 88%, which points out that the program possesses a good analogy with the real perception human. Even in tab.1 we can see the test numbers 2, where the image key is a sketch. We used the same methods for the calculus of the score. The percentage of correctness in the measure of similarity of the software is 80%, so we have a good level of fidelity.

8. CONCLUSIONS

The paper has presented a method of Content Based Image Retrieval, whose originality is related to two main aspects: 1) the definition of a *perceptual approach* that allows to build a new method for the similarity between color hues evaluation, and that represents the abstraction of the content in a simple and efficient way; 2) the introduction of a new *methodology for indexing* the images, based on the Similarity Quad-Tree. So we can extract the properties only related to the color, excluding the features like form and spatial relationships between objects of the image. The efficiency of this idea derives from the similarity function application. This function is derived from experimental evaluation of the perceptual metrics used by the human while judge the similarity between colors. The indexing methodology quality is related to a fast access to the representative features of an image that are stored in a vector: this necessarily involves high computation speed and cost minimization. Preliminary results give on 8000

image data-base, about 15 seconds of image-seek, for an ancient Pentium III at 750 MHz, where about 9 seconds are used for loading graphic interface, 2 second for feature extraction and less the 4 seconds for searching in the data-file and for matching. In a commercial PC with over 3 GHz of clock we go down under 2 second for all computation.

REFERENCES

Bach, J. R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., and Shu, C. F., 1996, The Virage image search engine: An open framework for image management, in: *Proceedings of the Storage and Retrieval for Still Image and Video Databases IV*, San Jose, CA, pp. 76-87.

Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., and Yanker, P., 1995, Query by image content and video content: The QBIC System, *IEEE Computer* **38**:23-31.

Gärdenfors, P., 1999, *Conceptual Spaces*, MIT Press, Cambridge, MA.

La Cascia, M., Sethi, S., and Sclaroff, S., 1998, Combining textual and visual cues for content-based image retrieval on the world wide web, in: *Proceedings of the IEEE Workshop on Content-based Access of Image and Video Libraries*, Santa Barbara, CA, pp. 24-28.

Medin, D. L., Goldstone, R. L., and Gentner, D., 1993, Respects for Similarity, *Psychological Review* **100**:254-278.

Nosofsky, R. M., 1991, Stimulus bias, asymmetric similarity, and classification, *Cognitive Psychology* **23**:94-140.

Pentland, A., Picard, R. W., and Sclaroff, S., 1996, Photobook: Content-based manipulation of image databases, *International Journal of Computer Vision* **18**:233-254

Rui, Y., Huang, T. S., and Chang, S. F., 1999, Image retrieval: Past, present, and future, *Journal of Visual Communication and Image Representation* **10**:1-23.

Sclaroff, S., Taycher, L., and La Cascia, M., 1997, Imagerover: A content-based image browser for the world wide web, in: *Proceedings of the IEEE Workshop on Content-based Access of Image and Video Libraries*, San Juan, PR, pp. 2-9.

Smith, E. E., Shoben, E. J., and Rips, L. J. ,1974, Structure and process in semantic memory: A featural model for semantic decisions, *Psychological Review* **81**:214-241.

Smith, J.R., 1997, *Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression*, PhD thesis, Graduate School of Arts and Sciences, Columbia University.

Smith, J. R. and Chang, S. F., 1996a, Querying by Color Regions using the VisualSEEk Content-Based Visual Query System, in: *Intelligent Multimedia Information Retrieval*, M. T. Maybury, ed., AAAI/MIT Press, Cambridge, MA, pp. 23-42.

Smith, J. R., and Chang, S. F., 1996b, VisualSEEK: a fully automated content-based image query system, in: *Proceedings of the 4th ACM International Conference on Multimedia*, Boston, MA, pp. 87-98.

Smith, J. R., and Chang, S. F., 1997, Visually searching the web for content, *IEEE Multimedia Magazine* **4**:12-20.

Torgerson, W. S., 1965, Multidimensional scaling of similarity, *Psychometrika* **30**:379-393.

10

THEORETICAL ISSUES IN SYSTEMICS

UNCERTAINTY AND THE ROLE OF THE OBSERVER

Giordano Bruno[1], Gianfranco Minati[2] and Alberto Trotta[3]
[1]*Dept. Memomat - School of Engineering, University "La Sapienza",*
Via A. Scarpa, 16 - 00161 Roma, Italy
Tel. +39-6-49766876, Fax +39-6-49766870, e-mail: bigi@dmmm.uniroma1.it
http://www.dmmm.uniroma1.it/~bruno
[2]*Italian Systems Society, Via P. Rossi, 42 - 20161 Milano, Italy*
Tel./Fax:+39-2-66202417, e-mail: gianfranco.minati@AIRS.it, www.airs.it
http://www.geocities.com/lminati/gminati/index.html
[3]*ITC "Emanuela Loi" - Nettuno (RM), Italy*
Tel. +39-06-9864039, e-mail: albertotrotta@virgilio.it

Abstract: In this paper we consider the correspondence between the centrality of the role of the observer for the concepts of probability and emergence. We base our considerations on the fundamental insight of the Italian mathematician Bruno de Finetti who introduced the concept of probability of an event as the observer's degree of belief. This correspondence is very important for dealing with modern problems of uncertainty related to chaos and complexity and to the modelling emergence.

Key words: coherence; emergence; information; probability; systemics; uncertainty.

Probability doesn't exist!
(Bruno de Finetti)

1. INTRODUCTION

Since the mechanistic approach - based on principles according to which the microscopic world is simpler than the macroscopic one, and that the macroscopic world may be explained through an infinitely precise knowledge of details - has been put in crisis due to various problems such as

the so-called Three Body Problem, new approaches have been introduced based on the theoretical role of the observer, non-linearity, uncertainty principles, constructivism, systemics as well as emergence - to mention just a few. We want to stress in this paper the need for developing a language and logic of uncertainty based on subjective probability and to highlight the essential role of the one who evaluates probabilities, whom we'll call the 'observer'. The logic of uncertainty explores the context of the possible: while acknowledging its inability to make predictions, it makes tools available for calculating the configurations of events with their probabilities. Traditionally, the concept of probability has been considered to be a consequence of our ignorance, our limitations: in precisely the same way that uncertainty principles were considered by mechanistic thinking as a limit of our knowledge of infinite details. We may now consider probability as our way of describing nature. The approach used is based on *coherence* in assigning probabilities to logically connected events. Moreover, it must be possible to coherently update assigned probabilities when *new information* are (or better are supposed to be) available about the event in consideration.

In this paper we will show how the problem regarding the observer's coherent assignment of probabilities to a set of events is related to *emergence* as understood in modern science.

Note that for the purposes of this paper *emergence* (Corning, 2002; Minati and Pessa, 2002; Pessa, 1998, 2002) may be considered as a process for the formation of new (i.e. requiring a level of description different from the one used for elements) collective entities - such as swarms, flocks, traffic, industrial districts, markets, as well as collective effects such as superconductivity, ferro-magnetism, and the laser effect - established by the coherent (as detected by an observer) behaviour of interacting components. With reference to the cognitive model (Anderson, 1983) used by the observer, emergence may be considered as an observer-dependent process; that is, by considering that collective properties emerge at a level of description higher (i.e. more abstract) than the one used by the observer for dealing with components, and that collective properties are detected as being new by the observer depending upon the cognitive model assumed, one is able to detect the establishment of coherence.

2. EVENTS AND UNCERTAINTY

2.1 Dealing with uncertainty

The paradigms that increasingly serve as guides for scientific inquiry and epistemological reflection are chaos and complexity, which if not the only paradigms, are certainly the main ones.

In the twentieth century a view of nature and the world began to be adopted that resembled less and less the view that had held sway for previous centuries. An explanation of phenomena was assumed to provide (deterministic) laws encompassing them; it was in this way that things progressed in all so-called scientific fields. These laws, in order to be considered as such, necessarily had to be objective, that is to say, independent of the observer, of the researcher or, better yet, had to be valid for all observers.

Chaos and complexity are showing us that it is not all so easy, that within systems (even those that aren't especially complicated) certain regularities continue to hold and that predicting their behaviour is not a simple matter of solving differential equations, but that at most one can use stochastic-like techniques. The disorder underlying the reality that we generally tend to consider as being ordered has become the source of new knowledge and a new way of thinking about and interpreting the world. The passage from ordered systems to disordered systems is based on the recognition that uncertainty prevails in the latter, in the sense that, while for the former we can say that once the state at time t is known it's possible to establish a priori the state at time $t+1$, the same does not hold for disordered systems. For the latter, we can only make predictions by calculating, for example, the probability that the system enters state $i+1$ at time $t+1$, assuming that at time t it was in state i. Dealing with uncertainty thus takes on a primary role. Even in this area, however, science, which is still dominated by the need to provide objectively valid answers-if not truths-has developed a number of methods and techniques of a statistical-probabilistic nature, and has identified a number of statistical-probabilistic models that claim in some way to represent and describe phenomena exhibiting uncertainty. In doing so, it gives the impression that these phenomena, by their nature, may be part of those models, and thus presents once again a kind of objective view of reality. In contrast, the mathematician Bruno de Finetti's fundamental insight, expressed in his claim that the probability of an event is nothing more than the *degree of belief* a *coherent individual* (as considered in probability theory introduced in section 3) has in the event's occurrence (based on his information about it), has brought to the forefront in this domain as well the observer's role (in this case, the individual who expresses

his degree of belief), just as happens in modern theories of complex systems. We will show by way of examples how the subjective view of probability has a systemic validity, in that the individual who evaluates a probability, the observer, plays an essential role in the emergence of a system of coherent probabilities (they are thus in no way independent of the observer). We will also address the problem of teaching probability which, in our opinion, needs to be re-examined in order to help students acquire a language and logic of uncertainty, with the aim of making them accustomed to dealing with situations exhibiting uncertainty and thus providing them a basis for making consistent, non-contradictory decisions.

We want to just mention here another different conceptual approach related to dealing with the future, one based on the so-called *Anticipatory Systems* introduced by Robert Rosen. In this approach, systems are considered as having predictive models of future evolution and their behaviour is based of such predictions.

"Strictly speaking, an anticipatory system is one in which present change of state depends upon future circumstances, rather than merely on the present or past. As such, anticipation has routinely been excluded from any kind of systematic study, on the grounds that it violates the causal foundation on which all of theoretical science must rest [...]. Nevertheless, biology is replete with situations in which organisms can generate and maintain internal predictive models of themselves and their environments, and utilize the predictions of these models about the future for purpose of control in the present. Many of the unique properties of organisms can really be understood only if these internal models are taken into account. Thus, the concept of a system with an internal predictive model seemed to offer a way to study anticipatory systems in a scientifically rigorous way." (Rosen, 1985).

2.2 Introductory theoretical considerations

Recall that in classical logic (other approaches are possible such as the "Fuzzy Logic" (Zadeh and Klir, 1996) related to the belonging of an element to a set, introduced in 1965 by L. A. Zadeh (Klir and Bo, 1995) by event we understand a logical proposition that can only be true or false, that is, one for which we have an unambiguous criterion that makes it possible to establish its truth or falsehood either now or in the future. It's well known that one can introduce into a family of events the logical operations of negation, implication, union, and intersection, and that this makes a Boolean algebra possible.

What is generally not adequately acknowledged, and in our view this is a grave mistake in the teaching of probability, is that in assigning a measure

for uncertainty to a family of events, whose truth value we do not know, we may introduce among them a natural *order* that corresponds to what we often do everyday.

If in fact we assume, as is valid, that one event in comparison to another may be either *more possible* or *less possible* or *equally possible*, then we can easily construct an *ordering relation* that we will call *no less possible* as regards the family of events being considered.

Such a relation has the following properties:

An uncertain event E_h of the family is always more possible than the impossible event and less possible than the certain event, and is no less possible than itself.

If an event E_h is no less possible than E_k, then E_k cannot be no less possible than E_h, unless E_h and E_k are equally possible.

If an event E_h is no less possible than E_k and E_k is no less possible than E_i, then E_h is no less possible than E_i.

If the events E and E_k are mutually exclusive, as are the events E and E_i, and E_k is no less possible than E_i, then the union of E and E_k is no less possible than the union of E and E_i.

Taking these properties as axioms, one can construct the theory of probability, ultimately obtaining the **theorem of total probability**.

In this way a qualitative measure of the uncertainty of an event is introduced.

Moreover, in instances involving the partition of a certain event into cases that are judged as being equally possible, it follows that the qualitative order that's introduced can be immediately translated into a quantitative measure of their uncertainty in terms of the ratio between the number of favourable cases and the number of possible cases.

If one then considers a further property relating to conditional events:[30]

if the events E_h and E_k imply E, then $E_h|E$ is no less possible than $E_k|E$ provided that E_h is no less possible than E_k,

taking it as an axiom along with the ones above, one can qualitatively develop the entire theory of probability (de Finetti, 1937).

Clearly, going from a qualitative measure of uncertainty to a quantitative one is not straightforward, except in the case cited earlier, while the opposite direction quite obviously is.

It is easy to see that there can be an infinite number of different qualitative evaluations that accord with the above axioms.

[30] Recall that the conditional event $E|H$ is a (three-valued) logical entity which is true when H is true and E is true, false when H is true and E is false, and indeterminate when H is false.

3. UNCERTAINTY AND PROBABILITY

How can one assign a quantitative measure of uncertainty? Once an event E is identified, quantitatively measuring its uncertainty means attributing a numeric value to the degree of belief an individual has in its being true (or in its occurrence)!

In order to make this measurement one needs a suitable *instrument*. Such an instrument must guarantee that the individual's evaluation reflects, in terms of a numeric scale, which he assesses qualitatively. To this end, Bruno de Finetti, the founder of the subjective theory of probability, proposed two equivalent criteria of measurement: bets and penalties (de Finetti, 1974).

We will only consider here the case of bets, which seems more natural to us since it also reflects the actual historical development of probability theory.[31]

Suppose you had to wager a certain amount of money in order win more money should a certain event occur. This is more or less what typically happens in all those instances in which betting takes place. Placing a *bet* on an event E means that one is willing to pay *part* of a certain amount S, which we will indicate by pS, in order to receive S if E occurs, and 0 otherwise.

If, as regards the bet, we introduce the *gain* function G_E, we obtain the following:

$$GE = \begin{cases} S - pS & \text{if } E \text{ occurs} \\ -pS & \text{otherwise} \end{cases}$$

We must be sure, however, that this criterion accurately measure what is precisely the most delicate step to be taken: the move from a qualitative to a quantitative evaluation! It's quite clear that in betting our goal is to maximize our gains; this fact could therefore lead us to distort our evaluation. How can we avoid this possibility so that the betting instrument does not become completely arbitrary and hence ineffective in fulfilling its role?

First of all, one needs to guarantee that for one who bets pS in order to receive S (from the other bettor) in case E occurs, must similarly be willing to pay S in order to receive pS if E occurs, that is, to exchange the terms of the bet with the other bettor. This will guarantee that the individual's evaluation reflects his degree of belief without being influenced by the desire to make greater gains, which otherwise might be made by the other. For

[31] As is well known, the theory of probability was first developed in connection to games and their respective bets that some players presented to, for example, Galileo and Pascal.

example, assume that in betting on E I estimate that I can pay 70 and receive 100 should E occur; I might try to increase my hypothetical gain (here equal to 30) by saying that I'm willing to pay 40. But if I'm willing to change betting positions with the other competitor, then my hypothetical gain could notably diminish (-60)!

But this is not enough. One must also guarantee that the possible values of the gain G_E do not both have the same sign, since only then would there be no certain loss or win, independently of whether or not E occurs. It's only with this guarantee that an individual would be willing to bet. This situation was aptly labelled *coherence* by de Finetti, which is nothing more than common sense applied to the calculation!

As is well known, coherence in betting on an event E let's one establish that, letting S equal 1 (but it also holds for $S \neq 1$), the amount p that an individual is willing to pay in order to receive 1 if E occurs, is always between 0 and 1. Moreover, if E is certain then p must necessarily be equal to 1, and if E is impossible then p must necessarily be equal to 0.

One immediately sees, however, that $p = 1$ does not imply that E is certain, nor does $p = 0$ imply that E is impossible (for a further discussion of this see de Finetti, 1974).

One also sees that this observation makes us reflect on the freedom and responsibility that the *observer* has in evaluating uncertainty.

Following once again de Finetti, we will say that the probability of an event E (i.e. the numeric measurement of the uncertainty over E) is the amount p that an individual is willing to pay in a coherent bet in order to receive 1 if E occurs and 0 otherwise.

As regards E there will therefore exist an infinite number of coherent evaluations, provided they are between 0 and 1!

How does the observer choose one? *S*/He will have to do so on the basis of the *information* s/he has regarding E and express that information in terms of a number. Clearly, the *richer* the information, the less doubt s/he will have in choosing one number from an infinite range of them!

One also notes that, as always happens in measuring information, an individual does not have an objective criterion of evaluation. Indeed, all one's personal convictions, moods, and various aspects that contribute to forming a judgment come into play. Thus, s/he needs only express objectively that which he subjectively evaluates!

In some cases it will obviously be much simpler: for example, if one judges as being equally possible the occurrence of E and its negation E^C, then one will assign the probability of 1/2.

On the other hand, if one judges that E is five times more probable than E^C then one will assign a probability of 5/6 to E and 1/6 to E^C.

In other cases one can appeal to evaluations that are based on the observed frequency, but only when the event E is part of a family of *exchangeable* events, as de Finetti accurately pointed out; that is, only when the evaluation of the probability of any n-tuple of events in the family considered depends only on the number of events fixed and not on the particular events themselves. In short, it only depends on *how many* and not on *which* events are considered (de Finetti, 1974)!

For example, if in extracting with replacement balls from an urn whose composition is unknown - that is, in which the total number of balls is known but not the percentage that are red - we wanted to evaluate the probability of extracting a red ball on the nth attempt, and we had extracted $(n\text{-}1)/3$ red balls out of a total of $n\text{-}1$, we could evaluate the probability of extracting a red one on the nth attempt as being equal to $(n\text{-}1)/3$ since the events of the family in consideration are interchangeable (we're only interested in how many red balls are extracted). But if we wanted to know the probability that a given boxer wins the 101st match of his career, would it suffice to know that s/he had won 85 of the preceding 100 matches, and thus evaluate the probability as being equal to 85/100? Obviously not, since he might have, in the worst case, lost all of the last 15 bouts!

We have so far looked at those considerations having to do with single events or families of similar events. However, we often come up against random events, which in some cases can be described by random numbers, for example: the number of automobile mortalities in a year involving people who were not wearing a safety belt, and which can be considered as events of the type $(X = n)$. Or we may be interested in more complex phenomena involving events of another type, for example: given a determinate city in Italy, whether in the next year the amount of rainfall will be greater; whether the average level of humidity will be lower; whether the level of smog will increase or the consumption of fuel remain the same.

In the first case, regarding random numbers, various probabilistic models have been created that allow us to evaluate the probability of any event composed of these, but we must not forget that these evaluations are not objective, as they may apparently seem, since once we have the model we need only apply the formulae in order to derive them: and it is always the observer who chooses, on the basis of his information, the model s/he believes best describes the phenomenon under consideration.

What happens in the second case?

The formulation adopted by de Finetti is, in our view, the one most apt for guaranteeing a *systemic* attitude and procedure since it is *open*, based on a *non-linear* approach and, more importantly, highlights the role played by the *observer* (the individual who evaluates).

Recall how one proceeds in the classic approach (often used in applications) towards random phenomena: first, one defines a space Ω of outcomes or possible basic cases, that is, one constructs a partition of the certain event; next, one attributes a probability to each of the cases (or *atoms*); and since any event referring to that phenomenon can be obtained as the union of elements, a probability is attributed, in a linear manner, to each of these (of course, the problem of how to assign probabilities to the elements remains open!).

In contrast, de Finetti bases his theory on the fact that every event is unique and for each of these we can express our degree of belief by way of a qualitative or quantitative evaluation. If, in addition to finding ourselves faced with a unique event and having to evaluate its probability, we need to assign probability to future events, how ought we to proceed?

Various cases may arise. If we have a family of events E_i that form a *partition* of the certain event, one can easily prove that, because of coherence, the sum of the probabilities of the single E_i must be 1, and that the probability of the union of *n mutually exclusive* events must be equal to the sum of the individual probabilities.

Given once again n events and a coherent assignment of probabilities, then the probability of an event that linearly depends on the first n is immediately assigned.

In cases in which there is a *logical* connection between the events E_i and a new event E, then, once again due to coherence, one needs to proceed in the following manner: the elements of the given family are first constructed (that is, all the possible intersections between the events such that in each there is either one of them or its negation, for example, $E_1 \wedge E_2 \wedge \ldots \wedge E^C_h \wedge E_{h+1} \wedge \ldots \wedge E_{n-1} \wedge E_n$); one then identifies the two events E' and E'' that are, respectively, the maximum event which is the union of all the elements implying E and the minimum event which is the union of all the elements implied by E; finally, the probability of E will necessarily be included in the closed interval $[P(E'), P(E'')]$.

Even in this situation the probability of E is not uniquely determined; it may simply be restricted to the interval [0,1] within which the observer may evaluate it in order to be coherent.

But the more interesting situation arises when dealing with single events that are examined one by one, starting from the first event. On the basis of the information s/he has, the observer attributes a coherent probability to each event, i.e., a value between 0 and 1. In this case, however, there may be logical connections between the events of which s/he is not aware or for which the information was available only when all the events had been introduced. How does one then check whether the evaluation is coherent as a whole?

One needs to construct the atoms on the basis of the events being considered, and since each of the latter will be the union of some of atoms, its probability must be equal to the sum of the probabilities (not yet determined, and hence unknown) of these elements. In this way one will obtain a system of n equations with s ($\leq 2^n$) unknown x_i , and with the constraints that $x_1 + x_2 + \ldots + x_s = 1$ and $x_i \geq 0$. If there is an s-tuple solution to the system, then the *evaluation* can be said to be *coherent* (Scozzafava, 2002).

It's interesting to note that the system may not have any solutions; in this case the evaluation would be incoherent. The system may also have a unique solution, or there may be several solutions, that is, a set of different evaluations that are all coherent.

We'll clarify this latter, interesting aspect by way of examples.

Let there be three events A, B, C and an observer who has evaluated their probability in the following way: $P(A)= 1/2$, $P(B) = 2/5$, $P(C) = 1/5$ (clearly, each of these represents a coherent evaluation!).

Let $ABC = \phi$ (where $ABC \equiv A \wedge B \wedge C$)

Then the possible elements are

$$Q_1 = A^C BC, \quad Q_2 = AB^C C, \quad Q_3 = ABC^C, \quad Q_4 = AB^C C^C,$$
$$Q_5 = A^C BC_C, \quad Q_6 = A^C B^C C, \quad Q_7 = A^C B^C C^C$$

In order to establish then whether, under the given conditions, the overall evaluation expressed by $P(A) = 1/2$, $P(B) = 2/5$, $P(C) = 1/5$ is coherent, one needs to determine whether the following system has at least one solution:

$$\begin{cases} x_2 + x_3 + x_4 = 1/2 \\ x_1 + x_3 + x_5 = 2/5 \\ x_1 + x_2 + x_6 = 1/5 \\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 1 \\ x_i \geq 0 \quad i = 1, \cdots, 7 \end{cases}$$

Letting $x_1 = 0$, $x_3 = 0$, $x_6 = 0$, as we may, after a few simple calculations one obtains the following solution to the system: $x_1 = 0$, $x_2 = 1/5$, $x_3 = 0$, $x_4 = 3/10$, $x_5 = 2/5$, $x_6 = 0$, $x_7 = 1/10$, thus the probabilities assigned to the events A, B, C determine a coherent overall evaluation!

Note that, had we let $x_2 = 0$, $x_3 = 0$, $x_6 = 0$, we would have obtained a different solution to the system $x_1 = 1/5$, $x_2 = 0$, $x_3 = 0$, $x_4 = 1/2$, $x_5 = 1/5$, $x_6 = 0$, $x_7 = 1/10$, and even in this case the overall evaluation would have turned out to be coherent!

Moreover, had we initially let $P(A) = \alpha$, $P(B) = \beta$ and $P(C) = \gamma$ with the condition that $ABC = \phi$, we would have obtained the following system:

$$\begin{cases} x_2 + x_3 + x_4 = \alpha \\ x_1 + x_3 + x_5 = \beta \\ x_1 + x_2 + x_6 = \gamma \\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 1 \\ x_i \geq 0 \quad i = 1, \cdots, 7 \\ 0 \leq \alpha \leq 1, \quad 0 \leq \beta \leq 1, \quad 0 \leq \gamma \leq 1, \end{cases}$$

and setting the selected $x_1 = 0$, as was done in the two cases studied above, we would have, respectively, the following solutions:

$$x_1 = 0, x_2 = \gamma, x_3 = 0, x_4 = \alpha - \gamma, x_5 = \beta, x_6 = 0, x_7 = 1 - (\alpha - \beta)$$

which would yield a coherent overall evaluation, provided $\alpha \geq \gamma$ and $\alpha + \beta \leq 1$.

$$x_1 = \gamma, x_2 = 0, x_3 = 0, x_4 = \alpha, x_5 = \beta - \gamma, x_6 = 0, x_7 = 1 - (\alpha + \beta),$$

which would yield a coherent overall evaluation, provided $\beta \geq \gamma$ and $\alpha + \beta \leq 1$

In any case, as seen in the examples described, one obtains a set of events that can be transformed into a system of events when, due to the observer, there *emerges* a *coherent evaluation* of their probability. Generally, as we have seen, there can be more than one coherent evaluation; it is therefore the observer's *responsibility* to choose the one s/he believes to best represent the *status of his information* regarding the set of events being considered![32]

[32] The problem of updating evaluations under conditions of uncertainty, which is performed by way of conditional probabilities, and its systemic validity, is addressed in a subsequent paper.

4. CONCLUSIONS

From a systemic viewpoint we may say that in this logic, interactions between agents (events, in this case) are required to be *coherent*, rather than linear or mutually deducible. In this case, *coherence* is not something related to formal, logical rules as in deduction, i.e. deterministic computation, but is instead related to *emergence*. Coherence is not something deterministically computed and detected: it is designed, learned, experienced and *then* formalized in modelling and simulation, as in the case of fuzzy logic (Zadeh et al., 1996).

According to this view, elements of a system are events. The observer is modelling the emergent system - i.e., a configuration of interacting events - through probabilities provided by the observer himself rather than through physical interactions. This approach seems to be necessary for the crucial theoretical role of the observer in processes of emergence and their modelling. We note that, in order to model a process of emergence, an observer, who is an integral part of this process, must first model himself by selecting, for example, his own logic which must be compatible with his role in the process. In dealing with systems considered as emerging from interacting components (e.g. physical, biological, and social systems), assuming in an objectivist way that one is not part of the system, or that one is part of it while adopting for the observer a logic that is incompatible with systems thinking (i.e. linear, in assuming it acts in a deterministic space), are two ineffective, though not incorrect, strategies.

The approach introduced by de Finetti considers systems of probabilities, and moves the focus from single events to systems of events.

We think that adopting the *language and logic of uncertainty* is crucial in everyday life for coping with complexity and for helping various disciplines to systemically interact by coherently using concepts, analogies, correspondences, and invariants. We believe that such a language and logic must be introduced and developed when one first begins speaking about probability, thus providing students and future citizens a power logical-critical tool without out which they may easily end up with errors, contradictions and misconceptions.

REFERENCES

Anderson, J. R., 1983, *The Architecture of Cognition*, Harvard University Press, Cambridge, MA.

Antomarini, B., 2003, Pensare con l'errore, *Lettera Internazionale*, 75, Roma.

Barrow-Green, J., 1996, *Poincaré and the Three Body Problem*, Amer. Math. Soc.

Benvenuto, S., 2002, Caos e mode culturali, *Lettera Internazionale*, 73/74, Roma

Bruno, G., 2002, Caos: il linguaggio della natura, *Lettera Internazionale*, 73/74, Roma

Bruno, G., 2002 , Logica dell'incerto e didattica della matematica, *Atti Convegno Nazionale UMI*, Loano.

Bruno, G., 2004, Logica dell'incerto: un'esperienza didattica, in: Atti Convegno: *Insegnare la matematica nella scuola di tutti e di ciascuno*, Bari.

Carnap, R., 1936, Testability and Meaning, *Philosophy of Science* **3**.

Coletti, G., and Sozzafava, R., eds., 1995, Mathematical Models for Handling Partial Knowledge, in: *Artificial Intelligence*, Plenum Press.

Coletti, G., Scozzafava, R., 2002, *Probabilistic Logic in a Coherent Setting*, Kluwer Academic Publishers.

Corning, P., 2002, The re-emergence of "emergence": A venerable concept in search of a theory, *Complexity* 7(6):18-30.

Davydov, A. S., 1979, Bioenergetics and the mechanism of muscle contraction, *International Journal of Quantum Chemistry* **16**:5-17.

Davies, P., 1989, *The New Physics: a Synthesis*, Cambridge University Press, Cambridge.

de Finetti, B., 1937, La prevision: ses lois logiques, ses sours subjectives, *Annales de l'Institut Poincaré* 7(1):1-68.

de Finetti, B., 1972, Subjective or objective probability: is the dispute indecidable?, *Symposia Mathematica*, Academic Press, London.

de Finetti, B., 1972, *Probability, Induction and Statistics*, Wiley, New York.

de Finetti, B., 1974, *Theory of Probability: A Critical Introductory Treatment*, 2 volumes (translated by A. Machi and A. Smith), Wiley, London.

de Finetti, B., (a cura di M, Mondadori), 1989, *La logica dell'incerto*, Il Saggiatore, Milano.

Del Giudice, E., Doglia S., Milani M. and Vitiello G., 1988, Structures, correlations and electromagnetic interactions in living matter: theory and applications, in: *Biological Coherence and Response to External Stimuli*, H. Froelich, ed., Springer-Verlag, Berlin, pp. 49-64.

Feynmann, R. S., 1965, *The Character of Physical Law*, MIT Press, Cambridge.

Greco, P., 2003, Termodinamica e complessità, *Lettera Internazionale*, 75, Roma.

Heisenberg, W., 1971, *Physics and Beyond*, Harper & Row, New York.

Keynes, J.M., 1921, *A Treatise on Probability*, Macmillan, London.

Klir, G. J., and Bo, Y., 1995, *Fuzzy sets and Fuzzy Logic: Theory and applications*, Prentice Hall, Englewood Cliffs, NJ.

Minati, G., and Pessa, E., eds., 2002, *Emergence in Complex Cognitive, Social and Biological Systems*, Proceedings of the Second Conference of the Italian Systems Society, Kluwer Academic/Plenum Publishers, New York.

Nicolis, G., and Prigogine, I., 1989, *Exploring Complexity*, Freeman, New York.

Penrose, R., 1989, *The Emperor's New Mind*, Oxford University Press, New York.

Peskin, E. M., and Schroeder, D. V., 1995, *An Introduction to Quantum Field Theory*, Westview Press.

Pessa, E., 1998, Emergence, Self-Organization, and Quantum Theory, in: *Proceedings of the First Italian Conference on Systemics*, G. Minati, ed., Apogeo, Milano, Italy.

Pessa, E., 2002, What is emergence?, in: *Emergence in Complex Cognitive, Social and Biological Systems*, G. Minati and E. Pessa, eds., Kluwer Academic/Plenum Publishers, New York.

Pievani, T., 2003, Una nuova biologia della complessità, *Lettera Internazionale*, 75, Roma.

Poincaré, H. 1908, *Science et méthode*, Flammarion, Paris.

Prigogine, I., and Stengers, I., 1979, *La Nouvelle Alliance*, Gallimard, Paris.

Prigogine, I., 1980, *From Being to Becoming*, Freeman, New York.
Prigogine, I., 1993, *Le leggi del caos*, Editori Laterza, Bari.
Rosen, R., 1985, *Anticipatory Systems*, Pergamon, New York.
Scozzafava, R., 2002, *Incertezza e Probabilità*, Zanichelli.
Snow, C. P., 1959, *The Two Cultures and the Scientific Revolution*, Cambridge University Pres, New York.
von Bertalanffy, L., 1968, *General Systems Theory*, George Braziller, New York.
von Foerster, H., 2003, *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer-Verlag, New York.
Voltaggio, F., 2002, L'enigma della complessità nella storia del pensiero, *Lettera Internazionale*, 73/74, Roma.
Zadeh, L. A., Klir, G. J., (ed.), Yuan, B., (ed.), 1996, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh*, World Scientific, Singapore.

TOWARDS A SECOND SYSTEMICS

Gianfranco Minati
Italian Systems Society, president, Via P. Rossi, 42 - 20161 Milano, Italy,
www.airs.it, Tel./Fax: +39-2-66202417, E-mail: gianfranco.minati@AIRS.it
http://www.geocities.com/lminati/gminati/index.html

Abstract: General Systems Theory produced many cultural and scientific results and approaches based on some fundamental aspects like the interaction between components, distinguished from relation. Inter-disciplinarity is introduced as the disciplinary study of systemic properties. Trans-disciplinarity is introduced as the study of systemic properties in general and of relationships among them. Finally Systemics is introduced as cultural generalization of the principles contained in the General Systems Theory. In reference to many new scientific disciplinary results we introduce the need to update the concepts and models of Systemics. We introduce a short review of those results, like Collective Phenomena; Phase Transitions in physics; Dynamical Usage of Models (DYSAM); Multiple systems, emerging from the same elements, but having simultaneous different interactions among them; Uncertainty Principles; Laws of scaling; Modelling emergence; Systemic meaning of new theorizations like Quantum Field Theories (QFT) in physics with related applications in biology, in studying the brain, consciousness, and in dealing with long-range correlations. The study of emergence undertaken in many disciplinary fields, like Physics, Biology, Artificial Life, Information Technology and Economics, has been realized focusing on the web of fundamental problems of General Systems Theory like the transition between non-systemic and systemic phases. The problem of modelling emergence relates to modelling processes of interaction between components and the observer. Dealing with those new problems and results calls for new theoretical approaches for Systemics. The change is expected to be so innovative to name this process with particular reference to emergence: Systemics of emergence or Second Systemics. We stress the need that the systems community, honouring its tradition and mission, be active part and leads this process, while at the present the most important research activities on Systemics seem to take place disciplinarily, out from the system community.

Key words: emergence; interaction; inter-disciplinarity; systemics; theory; trans-disciplinarity.

1. INTRODUCTION

The General Systems Theory, as introduced in the well-known book published by Ludwig von Bertalanffy in 1968 has been a very important *cultural* and *scientific* event (Von Bertalanffy, 1968).

The *cultural aspect* officially broke the assumption based on the Cartesian approach that in general the *macro level* may be explained through the *micro level* and that the *macro level* may be effectively managed by acting on the *micro level*. The mechanistic view, based on this assumption, considered that the microscopic world is simpler than the macro and that the macroscopic world may be explained through an infinite knowledge of the microscopic one. The possible ineffectiveness of this approach should be related to an inadequate and not sufficient knowledge of details. More appropriately this view has been named *reductionism* in reference to the belief that problems at the macro level may me reduced and explained by details and components. The problem was not the conceptual wrongness of this approach in general, but to assume it as a philosophical one, having general validity in any contexts. Mistake that, however, we may make by assuming valid *in principle* any approach, the systemic one also. The reductionism anyway allowed very important progresses in suitable dominium of science. Between the more important exponents of this approach we mention Newton (1643-1727) and Laplace (1749-1827). The cultural impact of the introduction of the General Systems Theory was produced not as much as revelation of discoveries, but as organization, interpretation, formalization, theorization and understanding of problems and results produced by disciplines and technology. The General Systems Theory allowed the establishing of new and more effective approaches, and the concept of system became a fundamental conceptual cultural and disciplinary resource in general.

Because the usual, daily life (often events-driven) deals with such *levels of description* that non-systemic, reductionistic approaches are effective (e.g. dealing with machines; procedural problems; local economic processes; simplified human relationships; cause-effect political problems; and in medicine too by focusing on symptoms), then the systemic approach is considered for sophisticated thinkers, dealing with specific *complex*, often confused with *complicated*, problems.. Actually, *Systems Science* is neglected in educational programs and only disciplinary taught (e.g. in engineering and management). The non availability of systems thinking at common level is a way to make social systems more subject to processes of manipulation, having people difficulties to understand or to hinder decisions taken in a systemic view by political leaders, marketing designers, media communicators, economical and financial institutions (Minati, 2004, 2005).

The theoretical and scientific aspects of the systemic approach have been carried out in many disciplinary fields, specially in physics, biology, engineering, cognitive science and information technology. The concept of system has been formulated in many different disciplinary ways and it is used in different contexts to describe interacting elements making emergent a system having properties different and unpredictable from the ones of components. In this process it has been considered the role of the observer (Von Foerster, 1981; 2003), integrant, theoretical part of the process itself, in the context of Uncertainty Principles introduced and used in Physics (Heisenberg, 1971).

The process of establishment (technically emergence as we'll see later) of systems based on interactions among components detected by an observer equipped with a sufficiently sophisticated cognitive model, is a conceptual framework of crucial importance in scientific disciplinary applications and theories.

The general conceptual framework based on interaction between components used to describe the establishment (*emergence*) of systems having properties different and non deducible from the ones of the components became the basis of the systemic approach.

The establishment of systems is related to the concept of interaction. In physics an interaction is assumed taking place when the behavior of an elements affects the behavior of other elements, as introduced in the original definition of system by Von Bertalanffy (Von Bertalanffy 1968). Interacting elements may be elements considered by physics (e.g. particles), biological elements (e.g. cells), chemical elements (e.g. molecules), living beings (e.g. ants), human beings (e.g. workers), words (e.g. in a novel), and sounds (e.g. in music) detected by an observer, equipped with a suitable cognitive model, as a system having properties (like physical, biological, chemical, behavioral, semantic, and artistic) different and not linearly deducible from the ones of components.

It has to be mentioned the difference between the processes of *interacting* and *composing*. Composition takes place when elements compose, that's when they are having, by reacting, positions in a new structure, as in chemical processes. Similar processes having the same nature, but different names in relation to different effects are merging and diluting. Composition gives rise to the establishment of new entities having properties different from ones of the components (like a molecule of water H_2O rises from the composition of two atoms of hydrogen and one atom of oxygen bonded together). Between molecules, the hydrogen bonds create a tetrahedral arrangement, caused by the hydrogen's attraction to neighbouring molecules' oxygen. The ideal bond angle for a single molecule of water is 109.47. Composition is the *result* of processes of interaction. Anyway there are

chemical reactions that establish a system by showing *continuous interaction*, like chemical (e.g. the Belousov-Zhabotinsky reaction generating showy chromatic changes, periodic liquid phase reactions (Belousov, 1959; Zhabotinsky, 1964)) and biological oscillating phenomena (Zaikin and Zhabotinsky, 1970). The story of the research that converged to the Belousov-Zhabotinsky reaction may be found in (Winfree, 1984).

It's important to also notice that the systemic approach is a *balance* between different cognitive strategies as objectivistic and non-objectivistic view, constructivism and non-constructivism, different levels of descriptions, giving up looking for the best choice in general and dealing with multiple modeling, as with the *Dynamic Usage of Models* (DYSAM) introduced later.

2. INTER- AND TRANS-DISCIPLINARY APPROACHES

Mono-, Multi- and inter-disciplinarity make reference to the way of dealing with disciplinary problems and to the levels of description assumed by the observer.

We had the opportunity to introduce how multi-disciplinarity is a need in the modern, complex world where it's important to synchronically and simultaneously use different disciplinary resources to deal with specific problems, like in management dealing with technical, economic, legal and social problems (Minati, 2001; Minati and Collen, 1997).

When the conceptual schema of interaction is applied to disciplines instead of physical agents the process of interacting is named inter-disciplinarity. We have inter-disciplinarity when interaction is conceptually taking place between different disciplinary approaches. This is possible when applying same systemic concepts in different disciplines, like *openness* in physics, biology and music. Models and approaches used by one discipline may be applied by another one because of the same systemic meaning. Phenomenological analogies and correspondences are generalized and made epistemologically robust when considered due to the same systemic property. Disciplinary problems are in this case usually described by using concepts of General Systems Theory, systemic properties and issues as: adaptive, allopoietic, anticipatory, attractors, autonomous, autopoietic, balanced, bifurcation, chaotic, complex, composite, connessionistic, deterministic, dissipative, equifinal, ergodic, far from equilibrium, goal-seeking, growing vs. developing, heterogeneous, heuristic, hierarchic, homeostatic, in equilibrium, open and closed, oscillating, self-

organized, symmetry breaking, systemic structure and organization (Minati and Pessa, in progress).

Inter-disciplinaruty takes also place in education when a discipline is taught by another one, like teaching mathematics when dealing with physics; history when dealing with economics; anthropology when dealing with geography.

Some problems are inter-disciplinary in principle, requiring cognitive, cultural, technical, scientific usage of same resources in different disciplinary ways, like in environmental research where disciplinary problems related to physics, biology, chemistry, economics and law are described as interconnected. In this case problems are inter-disciplinary in principle and not because of specific approaches. In this case inter-disciplinarity relates to the nature of the problems themselves, because same systemic aspects are described by using different disciplinary languages and because of effectiveness reasons, like for environmental research and cognitive science.

A very good example of research for inter-disciplinary application of disciplinary knowledge, physics and economics in this case, and not just combining results and approaches in a local inter-disciplinary view, is presented in a publication of the *Santa Fe Institute* (Yegorov, 2002) available on line. Another example of inter-disciplinary research is given by the *Annual Business Network Meetings of the Santa Fe Institute* on subjects like "Ways in which scaling laws can arise. A comparison of different explanations that are offered for scaling laws in a variety of fields." dealing with the Zipf's law. The research related to biological systems may be found in (West 2000). Similar research meetings are organized by the *New England Complex Systems Institute* (NECSI).

Trans-disciplinarity studies systemic properties *per se*. Trans-disciplinarity deals with systemic properties and problems in general, that's as properties of models and representations used with no reference to a specific disciplinary case nor to specific disciplinary cases simultaneously considered as in inter-disciplinarity. Trans-disciplinarity deals with questions like how do systems growth?, what the difference between growth and development?, how manage systems?, how describe systems?, how modelling and simulating? Trans-disciplinarity also allows dealing with relations between systemic properties, like openness and complexity. The general usability of systemic concepts is related to the possibility to describe problems in systemic way, that's by using the model of system. Systemic concepts and approaches apply because the description is common and not thanks to analogies or metaphors. This approach has been introduced by Von Bertalanffy in (Von Bertalanffy, 1968) particularly in chapter 3, in the

sessions 7 dedicated to isomorphism in sciences and 8 dedicated to unity of sciences.

Systemic properties, unlike problems, are not disciplinary, inter-disciplinary or trans-disciplinary per se: this relates to the level of description assumed in the process of emergence by the observer.

Inter-disciplinarity allows questions such as: How models from Physics may describe biological processes? How models describing processes of biological aggregation may be used to model socio-economic processes? Is Game Theory sufficient to model decision-making processes or do we need to take in count problems introduced by Cognitive Science? What the difference between systemic openness in biology and in physics?

Trans-disciplinarity allows questions such as: What is openness? What is learning? How systemic properties are related? How systemic properties may be induced or regulated? How systemic properties may me measured? Which relationships between adaptivity and openness? All those questions in reality refers to a single, specific, crucial theoretical issue: *modelling emergence*.

Trans-disciplinarity also is a very powerful research approach allowing to assume, in name of the *unity of science* and of the world (very important framework for the *General Systems Theory* as introduced in Von Bertalanffy 1968), the validity of inquiring about the meaning of a systemic property in a field where this property has not been detected or considered yet. Because of that trans-disciplinarity allows questions such as: What openness is for music? What openness is for natural languages? What is equilibrium in architecture? For instance processes of growing and opening/closing are trans-disciplinary when considered in general, as properties of systemic models and not of models carried out in a specific disciplinary or even inter-disciplinary approach. They are also studied as interrelated, networked systemic properties.

Trans-disciplinarity is constituted of related, interacting, *possible worlds*, one recognizable in another, in such an architecture of continuity that is promise of unity.

Practice of trans-disciplinarity takes place when considering different levels of representation, of reality, using disciplinary and inter-disciplinary knowledge and then generalizing by producing theories at an higher level of abstraction.

3. RESEARCH ON TRANS-DISCIPLINARITY

Considering, for instance, openness in different disciplinary systems (e.g. physical, biological, chemical and social) allows to establish inter-

disciplinary approaches. Considering openness as systemic property in general and its relationships, for instance, with complexity and self-organization allows to establish trans-disciplinary approaches.

Research on trans-disciplinarity relates to architectures of abstractions, methodologies, logics, models and representations of systemic properties in general, like the ones listed above.

Research at trans-disciplinary level is at an higher level of abstraction than disciplinary and inter-disciplinary research, but not independent. If it is considered independent we don't *watch into the telescope* as Galileo's detractors did when the current culture was based on the assumption to be autonomous, independent from experience and any other level of knowledge.

Telescope of trans-disciplinarity is the disciplinary and inter-disciplinary research.

We need to clarify that trans-disciplinary approaches relate to the establishment of robust theoretical generalization and not to a *metaphorical* usage of systemic knowledge.

Generalizing asks for a crucial theoretical effort, while *making generic, metaphoric* allows to extend the usage of the concept trading with less rigorousness, less specificity, and lower theoretical level. The misunderstanding of considering making generic and metaphoric equivalent to generalizing, must be removed.

Systemic issues, considered in general and not at specific level of description or of theorization gave rise to the more general aspect of the approach named *Systemics* in English, *Systémique* in French, *Sistemica* in Italian and Spanish (Minati and Collen, 1997; Collen and Minati, in progress), intended as a cultural generalization of the principles contained in the General Systems Theory.

Systemics should be source of concepts, language, and representations to be used in the daily life inducing people to think systemically, avoiding, for instance, processes of *consent manipulation* through language (Minati, 2004), influencing the social software (Minati, 2001, 2002) adopted.

I proposed at the 2004 conference of the *International Society for the Systems Science* (ISSS), and I would like to introduce the same proposal to the entire systems community, to establish an *observatory* on disciplinary results, problems, approaches, methodologies, and perspectives. The mission of the observatory should be to detect, realize, and understand the systemic meaning of disciplinary issues. The mission of the observatory should be to produce resources for systems researchers, to support research in systems theory, and to constantly revise systems thinking. Research on systems theory is supported by fostering new abstractions and conceptual architectures, and by introducing new correspondences, models and approaches arising from (not coinciding with) disciplinary results having

systemic contents. We need to organize and theorize the systemic contents of disciplinary results at an higher level, the one of trans-disciplinarity.

Systems scientists need to know at an adequate level of detail disciplinary results to go beyond inter-disciplinarity, necessary to deal with complexity, towards trans-disciplinarity. Trans-disciplinarity must have profound roots into disciplines and not leave them aside.

Divulgation doesn't help: we have to search by ourselves, as systems researchers, for the appropriate level of detail.

The observatory is expected to organize disciplinary workshops, a specific journal or a session in a systemic journal already established, and to organize a network of researchers sharing the common interest to identify the systemic meaning of their disciplinary activities. The observatory may also yearly publish a summary of its activity of identifying the systemic meanings of disciplinary results and disciplinary publications selected because of their relevance for systems theory. The observatory should be lead by a *manager of knowledge* and driven by a team of systems researchers having different disciplinary backgrounds.

4. SYSTEMICS OF EMERGENCE.

Emergence is the name of a line of research introduced firstly by C. L. Morgan in 1923 (Morgan, 1923) based on the idea that emergent properties are ones present at a certain levels of complexity but not at lower ones. Emergent properties were considered as *unpredictable* properties, arising in biological evolution. The concept related to the so-called *emergent evolutionism*. Von Bertalanffy (which was himself a biologist) implicitly proposed the same view when introducing the General Systems Theory. In his approach a system is introduced as an entity having properties that components have not.

A more precise definition arose when researchers in chaos theory began to carry out models of Artificial Life. The term emergent computation (Forrest, 1990) denoted the appearance of new, unexpected computational abilities generated (emergent) by cooperative action of individual agents interacting by simple rules.

It has been realized how the concept is *indissolubly* and *theoretically* related to the observer and its cognitive model. The observer carries out a model of the system and detect emergent phenomena by necessarily using two levels of observation. The observer is considered to model the system by introducing rules and symmetries according to the principles assumed and the expected behavior. In such a way a behavior is intended to be emergent if different from the one expected by the observer using a specific model.

Crutchfield introduced three different kinds of definition of emergence (Crutchfield, 1994):

1. *intuitive definition*, corresponding to the rough identification of "emergence" with "unpredictable";
2. *pattern formation*. A pattern is said to be "emergent" when an observer identifies "organization" in a dynamical system; auto-organizing processes in dissipative structures, like whirlpool (Nicolis and Prigogine, 1977);
3. *intrinsic emergence*, when occurrence of a particular, even compatible with the model used, behavior cannot be predicted in advance, and its appearance gives rise to a deep modification of the system's structure, requiring the formulation of a new model of the system, like flocks.

According to the definitions introduced by Crutchfield systemic properties, intended as non reducible to the properties of components, but emerging from their interactions, are generated by processes of *intrinsic emergence*.

I just mention the importance of the concept of emergence for Collective Phenomena, like superconductivity, ferromagnetism and laser effect, all manifestations of collective effects; and Collective Behaviors, like traffic, flocks, herds, and swarms from the interactions between single agents such as autonomous vehicles, birds, horses, and bees.

For Systemics the interest is on the study of emergence specifically focusing on the process of trans-formation of interacting elements into a new entity, like a system is. Researchers focus on what happens between the two states, non-system and system, disorder and order; during the transition, emergence of coherence; how to model, control, keep, regulate, simulate and manage this process.

This theoretical interest relates to some crucial questions like the change between non-living and living, between phases of the matter as studied in physics by phase transitions.

Emergence is now the name of research on processes that are the fundamental, core processes of the General Systems Theory.

In the last years General Systems Theory, became a system of scientific disciplinary theories based on the general approach introduced by L. von Bertalanffy.

We are now facing the process by which General Systems Theory is more and more becoming a *Theory of Emergence* looking for suitable models and formalizations of its fundamental bases.

The subject has been studied by disciplines often unaware of the systemic potentialities. We may refer to Physics (the quantum theory of collective phenomena, the theory of symmetry-breaking phase transitions, the reformulation of Quantum Field Theory, the theory of non linear phenomena, the study of classical and quantum chaos, the theory of

dissipative structures, the birth of Synergetics), Biology (the birth of structuralism), Neuroscience (the discovery of long range correlations in the brain, the discovery of role of chaotic processes in olfactory bulb, the birth of psychoneuroimmunology), Cognitive Science (the introduction of connectionist models of cognitive processing), Artificial Intelligence (the introduction of neural and neuro-fuzzy networks, soft computing, evolutionary algorithms and Artificial Life), Engineering (the birth of nanotechnology, quantum computing and self-designing machines), Philosophy (the analysis of Binding Problem, Symbol Grounding Problem and concepts such as Coherence and Consciousness) (Pessa, 2002).

Emergence (Minati and Pessa, 2002) refers to the core theoretical problems of the processes of arising of systems as introduced by the General Systems Theory. Correspondingly we need to look for, to be ready for the establishment of a *Second Systemics*, a *Systemics of Emergence* relating to new crucial issues like the ones of:

1. Collective Phenomena,
2. Phase Transitions, like in physics (e. g. transition from solid to liquid) and in learning processes,
3. Dynamical Usage of Models (DYSAM),
4. Multiple systems, emerging from the same elements but simultaneously having different interactions among them,
5. Uncertainty Principles,
6. Laws of scaling,
7. Modelling emergence,
8. Systemic meaning of new theorizations like ones of the Quantum Field Theories (QFT) in physics with related applications (e.g. biology, brain, consciousness, dealing with long-range correlations).

5. THE PROBLEMS OF EMERGENCE AS THE PROBLEMS OF GENERAL SYSTEMS THEORY

We are now going to explain with more details the single points listed above illustrating how the theoretical problems of Emergence are the core, fundamental problems of General Systems Theory.

1. We have examples of establishment of coherence between agents in different contexts like in Physics when giving arise to Collective Phenomena considering superconductivity (emergence of superconductivity in certain metals when cooled below a critical temperature), ferromagnetism, laser effect, which are manifestations of collective effects and cannot be described by using the traditional models of Physics (Haken, 1981, 1987); in Social Systems we may consider the

example of the emergence of industrial districts (Pyke and Sengenberger, 1992). We may consider the case of the establishment of swarms, anthills, herds, traffic, markets, biological growing, and societies as very deeply systemically shown by theoretical models in (Mikhailov and Calenbuhr, 2002).

2. In Physics, a phase transition is the transition of a thermodynamic system from one phase to another. Reference is to transitions between solid, liquid, and gaseous phases (boiling, melting, sublimation, etc.); ferromagnetic and paramagnetic phases of magnetic materials. Phases are sometimes called *states of matter*, but this may be misleading by introducing a confusion with *thermodynamic states*. For instance, two gases are in different thermodynamic states when maintained at different pressures, but they have the same state of matter. Phase transitions take place with the sudden changing in one or more physical properties when small changes occur for instance in a thermodynamic variable, such the temperature (Goldenfeld, 1992). The theoretical schema of the phase transition process has been considered in other domains, like in education by considering learning as a phase transition process (Penna and Pessa, 1995).
3. On the basis of the research made in different fields like the *Evolutionary Game Theory* (Maynard-Smith, 1982; Weibull, 1995); the so-called *Evolutionary Stable Strategies* applied to model ecosystems, (Huberman and Hogg, 1988, 1993), biological systems (Hines, 1987; Schuster, 1998), and markets (Gints, 2000); the so-called *iterated prisoner dilemma* game of great interest for game theorists (Pessa et al., 1998); it has been well established how, in games with incomplete information and having an enough level of complexity (such as the iterated prisoner dilemma) it's impossible to have a single equilibrium point. Only a multiplicity of different equilibrium points (Nash, 1950a, 1950b, 1951) is possible. On this basis it has been introduced the so-called *Dynamic Usage of Models* (DYSAM), related to the fact that dealing with processes of emergence and *multiple systems* (introduced later) it is ineffective the strategy based on looking for the more effective one. The strategy introduced with DYSAM is based on the *simultaneous* usage of multiple models allowing usage of errors and redundancy (like in models of cognitive science used in evolutionary age) instead of having the only strategy to avoid them (Minati, 2001; Minati and Brahms, 2001, 2002).
4. The concept of multiple-systems has been introduced several years ago in different fields, like in psychology with multiple-memory-systems (Tulving, 1985). The concept also relates to multiple belonging of elements. Multiple systems are considered emerging from the same elements when simultaneously some of them are coping with different

kinds of interactions as introduced in (Minati, 2001), naming them *Collective Beings*. The concept has been introduced specially for dealing with agents equipped with cognitive models and able to simultaneously handle different kinds of interactions. Examples of Collective Beings are constituted by human beings giving arise to different systems (like families, professional communities, markets, and educational communities), and behaving by always considering simultaneous multiple belonging.

5. *Uncertainty Principles* arose in situations in which the processing of observing itself was detected interfering with the system under study. It is well known how this has been the case in physics in studying phenomena at atomic scale as introduced by Heisenberg in 1927 (Heisenberg, 1971). Similar approaches has been introduced in more general contexts with reference to problems of cognitive science, when science studies itself as in the fundamental contributions of H. von Foerster (Von Foerster, 1981, 2003).
6. The laws of scaling applied to the systems growing. Systems have been empirically detected to not grow continuously, but following laws of scaling. The Zipf's law (Zipf, 1949) was formulated in the 1940's by Harvard linguistics professor George Kingsley Zipf (1902-1950) as an empirical generalisation. This empirical law states that the m-th most frequent word in a language shows up with frequency 1/m. It is important to notice that at the moment the so-called Zipf's law is an experimental law, not a theoretical one. The causes of Zipfian distributions in real life are matter of some controversy and should be matter of study of systems researchers. Zipf introduced his observation by considering the English language. The probability of encountering the rth most common word is given roughly by $P(r) = 0.1/r$ for r up to 1000. The law breaks down for less frequent words. The law is applied in many fields like growing of cities (Gabaix, 1999) and in biology (West and Brown, 2000).
7. The problem of modelling emergence is a very crucial problem in modern science. The issue relates to modelling a process and its observer. One approach is introduced in this Congress (Vitiello and Minati, 2005). Modelling emergence is a still open issue. To cope with this problem we need more than dynamic modelling. Technologies for user modelling are available, but in this case we need not only to model a user, but the observer and the observed process both changing interactively.
8. Systemic meaning of new theorizations like ones of the *Quantum Field Theories* (QFT). We make reference for instance to the concept of quantum, quantic vacuum, the simultaneity of effects, the long range correlations (Pessa, 1998b). Which relations, for instance, with the concept of *logical openness* (Minati et al., 1998)? These principles are

already disciplinary applied not only in Physics, but in Biology, in the study of the brain and in theories about consciousness (Vitiello, 2001).

The urgency to adapt and improve the systemic culture comes both from the need to

- avoid to be self-referential by just re-using principles having exhausted innovating power (like the whole is greater than the sum of its parts);
- deal with some disciplinary problems as the ones listed above and with results reached in disciplinary research (like in Physics) having such a level of architectural abstraction calling for their re-formulation in a systemic view, suitable for inter- and trans-disciplinary usage more than to be just divulgated or generalized. This recalls when sophisticated musical architectural designs (Minati, 2002) have the purpose to express meaning without making symbolically (Pessa, 1994) explicit the message trough text.

6. CONCLUSION

The systems community having expression in national and international societies, books, journals and conferences, are assumed to have the mission to continue the approaches introduced by L. von Bertalanffy not only by applying and divulgating, but by innovating them in the context of new disciplinary results.

It is some time that the traditional systems community seems to be unbalanced towards some specific disciplinary attitude like the so-called humanistic, social and managerial interests. On the other side some societies have interests in technological issues only, like engineering and information sciences.

Scientific issues like in physics, mathematics, chemistry, biology, cognitive sciences, linguistics, economics and medicine, are less and less considered by the traditional systems community. They are dealt with by specific institutions having little or no relationships with the community expected to take care of trans- and inter-disciplinarity. It's not a problem of communication, but of different interests. Some time the systems community seems partitioned between humanistic and scientific communities. In the new stage of *Systemic of Emergence* there are more theoretical and conceptual contributions from disciplinary societies than from traditional systems societies. This a betrayal of the mission that Von Bertalanffy gave to the developing systems communities.

I think that we have not the mission to just diffuse a methodology, but to care about procedures of scientific and artistic production and application

dealing with technological, scientific, ethical, and humanistic aspects. Trans-disciplinarity is the core value and to honour it we should continuously deal with disciplinary results more and more obtained thanks to systemic approaches applied in disciplines. We have the duty and not the option of trans-disciplinarity, having deep roots into disciplines and not leaving them aside.

The very first step is to deal with the new crucial and having a lot of implicit systemic meaning, problems of modern science, as mentioned above.

The systems community has now the fantastic opportunity to give a start to the *Systemics of Emergence*.

The *Italian Systems Society* has this goal.

REFERENCES

Belousov, B. P., 1959, *A periodic chemical reaction and its mechanism*, Sb. Ref. Radiats. Med. Medgiz, Moscow, pp. 145-147.

Chance, B., Pye E. K., Ghosh A. K., and Hess, B., eds., 1973, *Biological and Biochemical Oscillators*. Academic Press, New York.

Cruchtfield, J. P., 1994, The Calculi of emergence: computation, dynamics and induction, *Physica D* **75**:11-54.

Forrest, S., ed., 1990, *Emergent Computation*, North Holland, Amsterdam.

Gabaix, X., 1999, Zipf's law for cities: an explanation, *Quarterly Journal of Economics* **114**:739-767.

Gintis, H., 2000, *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*, Princeton University Press, Princeton.

Goldenfeld, N., 1992, *Lectures on Phase Transitions and the Renormalization Group*, Perseus Publishing.

Haken, H., 1981, *Erfolgsgeheimnisse der Natur*, Deutsche Verlags-Anstalt, Stuttgart.

Haken, H., 1987, Synergetics: an approach to self-organization, in: *Self-Organizing Systems: The Emergence of Order*, F. E. Yates, ed., Kluwer Academic/Plenum Publishers, New York.

Heisenberg, W., 1971, *Physics and Beyond*. Harper & Row, New York.

Hines, W. G., 1987, Evolutionary stable strategies: a review of basic theory, *Theoretical Population Biology* **31**:195-272.

Huberman, B. A., and Hogg, T., 1993, The emergence of computational ecologies, in: *1992 Lectures in Complex Systems*, L. Nadel and D. L. Stein, eds., SFI Studies in the Sciences of Complexity Lectures, Vol. V, Addision-Wesley, Reading MA, pp. 185-205.

Huberman, B. A., and Hogg, T., 1988, The behavior of computational ecologies, in: *The Ecology of Computation*, B. A. Huberman, ed., Elsevier North Holland, Amsterdam, pp. 77-115.

Maynard-Smith, J., 1982, *Evolution and the Theory of Games*, Cambridge University Press, Cambridge.

Mikhailov, A. S., and Calenbuhr, V., 2002, *From Cells to Societies*, Springer Verlag, Berlin.

Minati, G., 2004, Buying consensus in the "free markets", *World Futures* **60**(1-2):29-37.

Minati, G., 2001, *Esseri Collettivi*, Apogeo, Milano.

Minati, G., 2002, Music and systems architecture, in: *Proceedings of the 5th European Systems Science Congress*; http://www.afscet.asso.fr/resSystemica/Crete02/Minati.pdf .

Minati, G., 2005, Manipulating consent and democracy, in *Proceedings of the Third Italian Systems Conference*, G. Minati and E. Pessa, eds., Kluwer Academic/Plenum Publishers, London.

Minati, G., and Brahms, S., 2001, Experimenting with the DYnamic uSAge of Models (DYSAM) approach: the cases of corporate communication and education, in: *Proceedings or the 45th Conference of the International Society for the Systems Sciences* (ISSS), USA.

Minati, G., and Brahms, S., 2002, The DYnamic uSAge of Models (DYSAM), in: *Proceedings of the Second Italian Systems Conference*, G. Minati and E. Pessa, eds., Kluwer Academic/Plenum Publishers, London.

Minati, G., and Collen, A., 1997, *Introduction to Systemics*, Eagleye Bks Int., Walnut Creek, CA (Collen, G., and Minati, G., revised edition, in progress).

Minati, G., and Pessa, E., eds., 2002, *Emergence in Complex Cognitive, Social and Biological Systems*, Kluwer Academic/Plenum Publishers, London.

Minati, G., and Pessa, E., (in progress), *Collective Beings*, Kluwer Academic/Plenum Publishers, London.

Minati, G., Penna, M. P., and Pessa, E., 1998, Thermodynamic and logical openness in general systems, *Systems Research and Behavioral Science* **15**:131-145.

Morgan, C. L., 1923, *Emergent Evolution*, Williams & Norgate, London.

Nash, J., 1950a, The bargaining problem, *Econometrica* **18**:155-162.

Nash, J., 1950b, Equilibrium points in *n*-person games, in: *Proceedings of the National Academy of Sciences of the United States* **36**:48-49.

Nash, J., 1951, Non-cooperative games, *Annals of Mathematics* **54**:286-295.

Nicolis, G., and Prigogine, I., 1977, *Self-organization in nonequilibrium systems*, Wiley, New York.

Penna, M. P., and Pessa, E., 1995, Can learning process in neural networks be considered as a phase transition?, in: *Neural Nets, Proceedings of the 7th Italian Workshop on Neural Nets*, M. Marinaro and R. Tagliaferri, eds., World Scientific, Singapore, pp. 123-129.

Pessa, E., 1994, Symbolic and sub-symbolic models, and their use in systems research, *Systems Research* **11**:23-41.

Pessa, E., 1998, Emergence, self-organization, and quantum theory, in: *Proceedings of the First Italian Conference on Systemics*, G. Minati, ed., Apogeo, Milano.

Pessa, E., 2002, Systems community should focus on emergence in complex cognitive and biological systems; http://www.airs.it/airs/indexit.htm .

Pessa, E., Penna, M. P., and Montesanto, A., 1998, A systemic description of the interactions between two players in an iterated prisoner dilemma game, in: *Proceedings of the First Italian Conference on Systemics*, G. Minati, ed., Apogeo, Milano.

Pyke, F., and W., Sengenberger, eds., 1992, *Industrial districts and local economic regeneration*, International Institute for Labour Studies, Geneva.

Schuster, P., 1998, Evolution at molecular resolution, in: *Nonlinear Cooperative Phenomena in Biological Systems*, Leif Matsson, ed., World Scientific, Singapore, pp. 86-112.

Tulving, E., 1985, How many memory systems are there?, *American Psychologist* **40**:385-398.

Vitiello, G., 2001, *My double unveiled*, John Benjamins Publishing, Amsterdam, The Netherlands

Vitiello, G., and Minati, G., 2005, Mistake making machines, in: *Proceedings of the Third Italian Systems Conference*, G. Minati and E. Pessa, eds., Kluwer Academic/Plenum Publishers, London.

Von Bertalanffy, L., 1968, *General System Theory: Foundations, Development, Applications*, George Braziller, New York, (revised edition, March 1976).

Von Foerster, H., 1981, *Observing Systems. Selected Papers of Heinz von Foerster*, Intersystems Publications, Seaside, CA.

Von Foerster, H., 2003, *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer-Verlag, New York.

Weibull, J. W., 1995, *Evolutionary Game Theory*, The MIT Press, Cambridge, MA.

West, G. B., and Brown, J. H., eds., 2000, *Scaling in Biology*, Santa Fe Institute Studies in the Sciences of Complexity, Oxford University Press.

Winfree, A. T., 1984, The prehistory of the Belousov-Zhabotinsky oscillator, *J. Chem. Educ.* **61**:661-663.

Yegorov, Y., 2002, *Social Dynamics and Emergence of Order: Building Theory of Field in Economics*; http://www.santafe.edu/sfi/education/international/intlfellows/intlfal02/fellows/files/yegorov_2003.pdf.

Zaikin, A. N., and Zhabotinsky, A. M., 1970, Concentration wave propagation in two-dimensional liquid phase self oscillating system, *Nature* **225**:535-537.

Zhabotinsky, A. M., 1964, Periodic liquid phase reactions, in: *Proceedings of the Academy Sciences*, USSR, **157**:392-395.

Zipf, G. K., 1949, *Human Behavior and the Principle of Least Effort*, Addison-Wesley.

IS BEING COMPUTATIONAL AN INTRINSIC PROPERTY OF A DYNAMICAL SYSTEM?

Marco Giunti
Università di Cagliari, Italy

Abstract: I consider whether or not a discrete dynamical system has two isomorphic representations, one recursive and the other non-recursive; if it does not, the system can be said to be an intrinsic computational system. I prove that intrinsic computational systems exist, as well as non-intrinsic ones, and I finally argue that some representation of a non-intrinsic computational system is not effective with respect to the state-space structure of the system.

Key words: dynamical systems theory; discrete system; computational system; computation; computability theory; recursive function; effective procedure; representability.

1. INTRODUCTION

By the term *computational system* I refer to any device of the kind studied by standard (or elementary) computation theory. Thus, for example, Turing machines, register machines, cellular automata, and finite state automata are four different types of computational systems. Discreteness and determinism are two properties shared by all such devices. Therefore, so called analog computers are not computational systems in this sense.

Computational systems can in fact be thought as dynamical systems of a special kind. From an intuitive point of view, the computational systems can be identified with those discrete, deterministic, dynamical systems that can be described or represented effectively. Elsewhere, I gave a formal explication of this intuitive concept, and I showed that Turing machines and all other systems that have been actually studied by standard computation theory (register machines, cellular automata, monogenic production systems, etc.) satisfy the formal definition (Giunti, 1992, 1997).

According to this definition, being a computational system is a property that reduces to the existence of an *effective representation* of a discrete dynamical system. The condition of the existence of an effective representation can be made precise by requiring that the discrete system admit at least one recursive isomorphic system, *i.e.*, an isomorphic system (i) whose state space is a recursively enumerable[1] subset of the natural numbers and (ii) whose state transitions are recursive.

However, it is quite natural to ask whether or not the property of being computational is *intrinsic* to the dynamics of a discrete system. In fact, a discrete system might admit two isomorphic numeric representations, such that one is recursive and the other is not. In this case, the property of being computational could not be said to be intrinsic to the dynamics of the system, for it would depend on the numeric representation of the dynamics we choose.

In sec. 2 of this paper, I will lay down the formal apparatus necessary to discuss this problem and, in sec. 3, I will prove that some computational systems are intrinsic, but some are not. I will also raise the question whether intrinsic non-computational systems exist, *i.e.*, whether there are discrete systems whose numeric representations are all non-recursive.

Finally, in sec. 4, I will make a few remarks about a different way of looking at this kind of problem. One of the results of sec. 3 show that some discrete systems are non-intrinsic computational systems, for they admit two numeric representations such that one is recursive and the other is not. One may wonder whether this result might depend on the fact that at least one of the two *representations* is not intrinsic to the dynamics of the system.

By an *intrinsic representation* of a discrete[2] dynamical system DS I mean a pair $(u, DS_{\#})$ such that: (i) $DS_{\#}$ is a dynamical system whose state space is the set of the natural numbers Z^{+}; (ii) u is an isomorphism of $DS_{\#}$ in DS (and thus, u is a bijective enumeration from Z^{+} to the state space M of DS); (iii) the enumeration $u: Z^{+} \rightarrow M$ can be constructed effectively by means of a mechanical procedure that takes as given the *whole structure* of the state space M, and *nothing more*.

A precise definition of the intuitive idea of an intrinsic representation of a discrete dynamical system, however, requires some further general concepts of dynamical systems theory and graph theory, as well as the new notion of an *enumerating machine*, *i.e.*, a machine that effectively produces an enumeration of the state space by moving from state to state in accordance with the state transitions determined by the dynamics of the system. These developments go beyond the limits of the present work, and will be the subject of a forthcoming paper.

2. DYNAMICAL AND COMPUTATIONAL SYSTEMS

A dynamical system is a mathematical model that expresses the idea of an arbitrary deterministic system, either reversible or irreversible, with discrete or continuous time or state space (Arnold, 1977; Szlensk, 1984). Let Z be the integers, Z^+ the non-negative integers, R the reals and R^+ the non-negative reals; below is the exact definition of a dynamical system.

[1] *DS is a dynamical system* iff there is M, T, $(g^t)_{t \in T}$ such that $DS = (M, (g^t)_{t \in T})$ and

1. M is a non-empty set; M represents all the possible states of the system, and it is called the *state space*;
2. T is either Z, Z^+, R, or R^+; T represents the time of the system, and it is called the *time set*[3];
3. $(g^t)_{t \in T}$ is a family of functions from M to M; each function g^t is called a *state transition* or a *t-advance* of the system;
4. for any $t, v \in T$, for any $x \in M$,
   $g^0(x) = x$;
   $g^{t+v}(x) = g^v(g^t(x))$.

[2] A *discrete*[4] *dynamical system* is a dynamical system whose state space is finite or denumerable, and whose time set is either Z or Z^+; [3] a *continuous dynamical system* is a dynamical system that is not discrete; [4] a *cascade* is a dynamical system with discrete time, *i.e.*, whose time set is either Z or Z^+. Thus, all discrete dynamical systems are cascades, but the reverse is not true.

[5] A dynamical system is *reversible* iff its time set is either Z or R; [6] it is *irreversible* iff its time set is either Z^+ or R^+. If a dynamical system DS is reversible, then any state transition is bijective, and the set of all state transitions $\{g^t\}_{t \in T}$ is a commutative group with respect to the composition of functions; the unit is g^0 and, for any $t \in T$, the algebraic inverse of g^t is g^{-t} = the inverse function $(g^t)^{-1}$. If DS is irreversible, $\{g^t\}_{t \in T}$ is a commutative monoid with respect to the composition operation, with unit g^0.

Any t-advance ($t > 0$) of an irreversible cascade can always be thought as being generated by iterating t times a given function $g: M \rightarrow M$ (thus, $g^1 = g$). As for a reversible cascade, the generating function $g: M \rightarrow M$ must be bijective; the positive t-advances are then obtained as before, while the negative ones ($t < 0$) are generated by iterating $|t|$ times the inverse function g^{-1}.

[7] $DS = (M, (g^t)_{t \in T})$ is a *possible dynamical system* iff DS satisfies the first three conditions of def. 1. We can now define the concept of an isomorphism between two possible dynamical systems as follows. [8] u *is an isomorphism of* DS_1 *in* DS_2 iff $DS_1 = (M, (g^t)_{t \in T})$ and $DS_2 = (N, (h^v)_{v \in V})$ are possible dynamical systems, $T = V$, $u: M \rightarrow N$ is a bijection and, for any $t \in T$, for any $x \in M$, $u(g^t(x)) = h^t(u(x))$.

[9] DS_1 *is isomorphic* to DS_2 iff there is u such that u is an isomorphism of DS_1 in DS_2. It is easy to verify that the isomorphism relation is an equivalence relation on any given set of possible dynamical systems. (The concept of *set of all possible dynamical systems* is inconsistent, and we must then take as the basis of the theory of dynamical systems a specific, sufficiently large, set of possible dynamical systems.)

It is also not difficult to prove that the relation of isomorphism is a congruence with respect to the property of being a dynamical system, that is to say: if DS_1 is isomorphic to DS_2 and DS_1 is a dynamical system, then DS_2 is a dynamical system. This allows us to speak of abstract dynamical systems in exactly the same sense we talk of abstract groups, fields, lattices, order structures, etc. We can thus define: [10] an *abstract dynamical system* is any equivalence class of isomorphic dynamical systems.

[11] A *representation* of a dynamical system DS is a pair $(u, DS_\#)$ such that u is an isomorphism of $DS_\#$ in DS. [12] A *numeric representation* of a dynamical system DS is a representation $(u, DS_\#)$ of DS such that the state space of $DS_\#$ is a subset of Z^+. By def. 12 and 1, it immediately follows that any discrete system has a numeric representation. [13] A *recursive representation* of a discrete dynamical system DS is a numeric representation $(u, DS_\#)$ of DS such that (i) the state space of $DS_\#$ is a recursively enumerable subset of Z^+; (ii) any state transition of $DS_\#$ is a recursive function.

Note that, since any discrete system is a cascade, condition (ii) of def. 13 reduces to the following condition. Let $g_\#$ be the generating function of the positive state transitions of $DS_\#$. Then, if DS is irreversible, (ii) is equivalent to requiring that $g_\#$ be recursive; if DS is reversible, (ii) is equivalent to requiring that both $g_\#$ and its inverse $g_\#^{-1}$ be recursive. However, by condition (i), the domain of $g_\#$ is a recursively enumerable subset of Z^+; this, together with the recursivity of $g_\#$, entails that $g_\#^{-1}$ is recursive too. Therefore, condition (ii) is equivalent to requiring that $g_\#$ be recursive.

[14] A *canonic numeric representation* of a dynamical system DS is a numeric representation $(u, DS_\#)$ of DS such that either the state space $Z_\#$ of $DS_\#$ is an initial segment of Z^+, or $Z_\# = Z^+$. By def. 14 and 1, it is obvious that any discrete dynamical system has a canonic numeric representation. [15] A *recursive canonic representation* of a discrete dynamical system DS

is a canonic numeric representation $(u, DS_{\#})$ of DS, which is also a recursive representation of DS.

Obviously, the recursive canonic representations are a proper subset of the recursive representations. Nevertheless, if a discrete dynamical system has a recursive representation, it has a recursive canonic representation as well. This is established by the following theorem.

THEOREM 1 [RECURSIVE CANONIC REPRESENTABILITY]

For any discrete dynamical system DS, if DS has a recursive representation, then DS has a recursive canonic representation.

PROOF

Let $DS = (M, (g^t)_{t \in T})$ be a discrete dynamical system, and let $(u, DS_{\#})$ be a recursive representation of DS, where $DS_{\#} = (N, (g_{\#}{}^t)_{t \in T})$. The proof is trivial if M is finite. Let us then assume that M is denumerable. By def. 13, N is recursively enumerable. By means of this fact, we construct a canonic numeric representation of DS, and we then show that it is recursive.

Since N is recursively enumerable, there is $e: Z^+ \to N$ such that e is surjective and recursive. Let us define $c: Z^+ \to N$ as follows: for any $m \in Z^+$, $c(m) = e$(the least $n \geq m$ such that, for any $k < m$, $c(k) \neq e(n)$). By its definition, c is a bijection from Z^+ to N and, since e is recursive, c is recursive as well. Since the domain of c is the whole Z^+, its inverse c^{-1} is recursive too. Let us then define $u': Z^+ \to M$ as follows: for any $m \in Z^+$, $u'(m) = u(c(m))$. By its definition, u' is a bijection from Z^+ to M. Let $g_{\#}': Z^+ \to N$ be defined as follows: for any $m \in Z^+$, $g_{\#}'(m) = u'^{-1}(g(u'(m)))$, where $g = g^1$ is the generating function of DS. Let $DS_{\#}' = (Z^+, (g_{\#}'^t)_{t \in T})$ be the discrete dynamical system generated by $g_{\#}'$. Then, by construction, $(u', DS_{\#}')$ is a canonic numeric representation of DS.

In addition, by the definitions of u' and $g_{\#}'$, and since u is an isomorphism of $DS_{\#}$ in DS, it follows that, for any $m \in Z^+$, $g_{\#}'(m) = c^{-1}(g_{\#}(c(m)))$. Thus, being a composition of recursive functions, $g_{\#}'$ is recursive. Therefore, $(u', DS_{\#}')$ is a recursive representation of DS. Since it is also a canonic numeric representation of DS, by def. 15, the thesis follows.

Q.E.D.

Finally, we can state the precise definition of the concept of a computational system. [16] DS is a *computational system* iff DS is a discrete dynamical system, and there is a recursive representation of DS.

Before concluding this section, let me remark that, according to theorem 1, computational systems admit a *uniform* recursive representation. In fact,

all computational systems whose state space has the same cardinality have a recursive canonic representation with *identical* state space (either the same initial segment of Z^+ for all finite systems with the same number of states, or Z^+ itself for all systems with a denumerable number of states). Thus, the recursive canonic representations of any two such systems only differ for the *state space structure*, that is, for the family of recursive state transitions that determine the specific dynamics of the different systems.

3. INTRINSIC VS. NON-INTRINSIC COMPUTATIONAL SYSTEMS

We are now in a position to distinguish two types of computational systems, according to whether the property of being computational is intrinsic or not. [17] *DS* is an *intrinsic computational system* iff *DS* is a discrete dynamical system, and any numeric representation of *DS* is a recursive representation of *DS*. [18] *DS* is a *non-intrinsic computational system* iff *DS* is a computational system, and there is a numeric representation of *DS* that is not a recursive representation of *DS*. [19] *DS* is an *intrinsic non-computational system* iff *DS* is a discrete dynamical system, and any numeric representation of *DS* is not a recursive representation of *DS*.

Note that computational systems and intrinsic non-computational systems constitute a partition of the set of all discrete dynamical systems. However, while the set of all computational systems is certainly not empty, it is an open question whether intrinsic non-computational systems exist.[5]

Analogously, intrinsic computational systems and non-intrinsic computational systems form a partition of the set of all computational systems. However, all *denumerable*[6] computational systems are non-intrinsic, and this entails that the set of all intrinsic computational systems is identical to the set of all finite discrete dynamical systems. This is the content of the theorem below.

THEOREM 2 [FINITENESS OF ANY INTRINSIC COMPUTATIONAL SYSTEM]

1. If *DS* is a denumerable computational system, then *DS* is a non-intrinsic computational system;
2. *DS* is an intrinsic computational system iff *DS* is a finite discrete dynamical system.

PROOF OF 1

If *DS* is a denumerable computational system, it is always possible to find a non-recursive numeric representation $(u, DS_{\#})$ of *DS*. Take the state space N of $DS_{\#}$ to be an arbitrary non-recursively enumerable subset

of Z^+. If M is the state space of DS and g is the generating function of DS, choose any bijection $u: N \rightarrow M$, and generate the state transitions of $DS_\#$ by $g_\#: N \rightarrow N$ such that, for any $n \in N$, $g_\#(n) = u^{-1}(g(u(n)))$. By construction, $(u, DS_\#)$ is a non-recursive numeric representation of DS.

PROOF OF 2

The left/right implication follows from thesis 1. The converse is obvious.

Q.E.D.

According to theorem 2, all denumerable computational systems are non-intrinsic (thesis 1). However, if we look at the proof of thesis 1, we realize that it depends crucially on condition (i) of def. 13, *i.e.*, on the requirement that the state space of any recursive representation be a recursively enumerable subset of Z^+. One may then think that the definition of an intrinsic computational system (def. 17) is too strong, for it fails for all denumerable systems *just because*, for any such system, there is a numeric representation whose state space in not recursively enumerable. But, as far as we know, this numeric representation might very well satisfy condition (ii) of def. 13. It is then natural to ask whether, by appropriately limiting the scope of the relevant numeric representations, we might get a refined concept of intrinsic computational system, for which the somewhat trivial proof of thesis 1 does not go through.

In effect, it is possible to get such a refined concept by just considering the *canonic* numeric representations, and not *all* the numeric representations like def. 17 does. The new definitions (where "*c*" is a reminder that these concepts are limited to canonic numeric representations) are as follows. [20] DS is a *c-intrinsic computational system* iff DS is a discrete dynamical system, and any canonic numeric representation of DS is a recursive representation of DS. [21] DS is a *c-non-intrinsic computational system* iff DS is a computational system, and there is a canonic numeric representation of DS that is not a recursive representation of DS. [22] DS is a *c-intrinsic non-computational system* iff DS is a discrete dynamical system, and any canonic numeric representation of DS is not a recursive representation of DS.

I will now show that the set of the *c*-intrinsic computational systems, as well as the one of the *c*-non-intrinsic computational systems, is not empty, and that both sets admit members whose state space is denumerable.

The proof concerning the *c*-intrinsic computational systems takes into account the discrete dynamical system $DS_1 = (Z^+, (i^n)_{n \in Z^+})$, generated by the identity function $i: Z^+ \rightarrow Z^+$. It is then almost immediate to show that DS_1 is a *c*-intrinsic computational system.

As for c-non-intrinsic computational systems, I will show that the discrete dynamical system $DS_2 = (Z^+, (s^n)_{n \in Z^+})$, generated by the successor function s: $Z^+ \rightarrow Z^+$, is a c-non-intrinsic computational system.

THEOREM 3 [EXISTENCE OF BOTH C-INTRINSIC AND C-NON-INTRINSIC COMPUTATIONAL SYSTEMS]

$DS_1 = (Z^+, (i^n)_{n \in Z^+})$ is a c-intrinsic computational system;
$DS_2 = (Z^+, (s^n)_{n \in Z^+})$ is a c-non-intrinsic computational system.

PROOF OF 1

Obviously, DS_1 is a computational system, for (i, DS_1), where i is the identity function on Z^+, is a recursive representation of DS_1.
Note that an arbitrary canonic numeric representation of DS_1 is of the form $(p, DS_\#)$, where p: $Z^+ \rightarrow Z^+$ is an arbitrary bijection and $DS_\# = DS_1$. Thus, by def. 13, $(p, DS_\#)$ is a recursive representation of DS_1. Therefore, by def. 20, DS_1 is a c-intrinsic computational system.

PROOF OF 2

Obviously, DS_2 is a computational system, for (i, DS_2), where i is the identity function on Z^+, is a recursive representation of DS_2.
For any bijection p: $Z^+ \rightarrow Z^+$, let s_p: $Z^+ \rightarrow Z^+$ such that, for any $m \in Z^+$, $s_p(m) = p(s(p^{-1}(m)))$. Let $DS_p = (Z^+, (s_p{}^n)_{n \in Z^+})$ be the discrete dynamical system generated by s_p. Then, by construction, (p^{-1}, DS_p) is a canonic numeric representation of DS_2.
Note that, for any p, s_p can be thought as a "new successor function" on Z^+, corresponding to the order induced by p on Z^+. The first element of this order, so to speak the "new zero element", is $p(0)$, the "new 1" is $p(1)$, and so forth, so that, for any $m \in Z^+$, $p(m) = s_p{}^m(p(0))$. It is then easy to verify that, for any two different bijections p and q, $s_p \neq s_q$.[7]
Consequently, there are as many functions s_p as there are bijections p: $Z^+ \rightarrow Z^+$. But the number of such bijections is non-denumerable. Hence, there is p^* such that s_{p^*} is not recursive. It thus follows that the canonic numeric representation (p^{*-1}, DS_{p^*}) is not recursive. Therefore, by def. 21, DS_2 is a c-non-intrinsic computational system.

Q.E.D.

4. TOWARD A THEORY OF INTRINSIC REPRESENTABILITY

That the system generated by the identity function be a *c*-intrinsic computational system was to be expected. On the contrary, the proof that the computational system DS_2 generated by the successor function is *c*-non-intrinsic is surely surprising. There is a feeling of oddity in realizing that a dynamical system like DS_{p^*}, which has exactly the same structure as the sequence of the natural numbers, is generated by a non-recursive *pseudo-successor* function s_{p^*}, and that (p^{*-1}, DS_{p^*}) thus constitutes a *bona fide* non-recursive canonic representation of DS_2, which, in contrast, is generated by the *authentic* successor function that is obviously recursive[8].

One may wonder that, after all, (p^{*-1}, DS_{p^*}) is not a *bona fide* representation of DS_2. That this way of looking at the problem might be promising is confirmed by the following observation. While it is obvious that, if we are given the whole structure of DS_2 (*i.e.*, the successor function s: $Z^+ \rightarrow Z^+$), we can mechanically produce the identity function i (by simply starting from state 0 and counting 0, then moving to state $s(0) = 1$ and counting 1, and so forth), it seems that, by just moving back and forth along the structure of DS_2 and counting whenever we reach a new state, in no way can we produce such a complex permutation of Z^+ like the bijection p^{*-1} (see fig. 1 below).

Also observe that the situation is exactly symmetrical if, instead, we imagine that we are given the whole structure of DS_{p^*} (*i.e.*, the pseudo-successor function s_{p^*}: $Z^+ \rightarrow Z^+$). In this case, we can easily produce p^* (by starting from state pseudo-0 = $p^*(0)$ and counting 0, then moving to state $s_{p^*}(0)$ = pseudo-1 and counting 1, and so forth), but it seems that, by just moving back and forth along the structure of DS_{p^*} and counting whenever we reach a new state, in no way can we produce such a simple enumeration of Z^+ like the identity function.

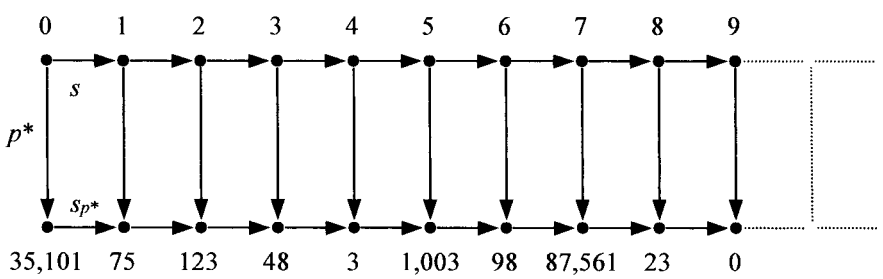


Figure 1. A hypothetical initial segment of p^*.

Thus, summing up the two previous observations, we can describe the situation as follows: (i, DS_2), but not (p^{*-1}, DS_{p^*}), is a *bona fide* representation of DS_2; conversely, (p^*, DS_2), but not (i, DS_{p^*}), is a *bona fide* representation of DS_{p^*}; where, by a *bona fide representation* of a discrete dynamical system $DS = (M, (g^t)_{t \in T})$, I mean a canonic numeric representation $(u, DS_{\#})$ of DS, such that the bijection $u: Z^+ \rightarrow M$ can be constructed effectively by means of a mechanical procedure that takes as given the whole structure of the state space M, and nothing more. In other words, a *bona fide* representation of DS is a canonic numeric representation $(u, DS_{\#})$ of DS, whose enumeration[9] $u: Z^+ \rightarrow M$ is effective with respect to the structure of the state space M. Let us stipulate that the term [23] *intrinsic representation* of DS is a synonym for *bona fide* representation of DS.

Note that, as it stands, def. 23 is not formally adequate, for I have not precisely defined the idea of an enumeration $u: Z^+ \rightarrow M$ that is effective with respect to the structure of the state space M. However, a precise definition of the intuitive idea of an intrinsic representation of a discrete dynamical system requires some further general concepts of dynamical systems theory and graph theory, as well as the new notion of an *enumerating machine*, *i.e.*, a machine that effectively produces an enumeration of the state space by moving from state to state in accordance with the state transitions determined by the dynamics of the system. These developments go beyond the scope of the present work, and will be the subject of a forthcoming paper.

NOTES

1 In previous works (Giunti 1992, 1995, 1996, 1997, 1998), I required that the state space of the representing system be recursive, and not just recursively enumerable. From an intuitive point of view, the recursivity of the state space may seem too strong, for the important issue is that there exist an effective procedure for generating all the numbers that represent states of the system, and not that can we decide whether an arbitrary number stands for some state. In effect, however, it does not matter which of the two requirements we choose, for the two corresponding definitions of computational system are equivalent. This is an immediate consequence of the theorem of canonic recursive representability (th. 1), and of the fact that the state space of any canonic recursive representation is a recursive subset of the natural numbers Z^+ (because, by def. 15 and 14, such a state space is either finite or identical to Z^+).

2 At the moment, I only consider *denumerable* discrete systems, *i.e.*, discrete dynamical systems with a denumerable number of states. However, the complete definition of an intrinsic representation must also apply to the somehow trivial case of *finite* discrete systems.

3 It is important to keep in mind that T is not a *bare* set, but rather, a set on which we implicitly assume the whole usual structure of, respectively, the (non-negative) integers or the (non-negative) reals. More precisely, we could say that T is the domain of a model $(T, (\sigma_i)_{i \in I})$ of, respectively, the theory of the (non-negative) integers or the theory of the (non-negative) reals.

4 The term "discrete dynamical system" is often used as a synonym for "cascade", *i.e.*, a dynamical system with discrete time; see for example Kulenovic and Merino (2002), Martelli (1999), and Sandefour (1990). My use of the term "discrete dynamical system" is in accordance with Turing (1950).

5 In Giunti (1998, note 7) I claimed that the existence of intrinsic non-computational systems can be proved. I am no longer so confident that such a proof can be given.

6 By a *denumerable* (or *finite*) *dynamical system* I mean a system whose state space is denumerable (or finite).

7 The actual proof that if $p \neq q$, then $sp \neq sq$ is by reductio and by cases. In the first place, under the assumption $p \neq q$, we assume for reductio $sp = sq$. We then consider the two cases $p(0) \neq q(0)$ and $p(0) = q(0)$. In either case, keeping in mind the observation in the text, a contradiction readily follows.

8 But we may get even more surprised when we realize that this proof also entails the following: $(Z^+, s_{p^*}, p^*(0))$ is a model of Peano's axioms such that its successor function s_{p^*} is not recursive! This means that the property of being recursive is not an arithmetical property of the successor function, where by an *arithmetical property* I mean any property of a numeric entity which is invariant for all (isomorphic) models of Peano's axioms. To put it in a different way: the recursivity/non-recursivity of the successor function seems to depend on the model of arithmetic we choose. (And, quite obviously, this consideration can then be extendend to any other numeric entity to which the property of being recursive/non-recursive applies.)

9 If DS is a *finite* discrete system, then the enumeration u is not a bijection from the whole Z^+ to M, but from a finite initial segment of Z^+ to M.

REFERENCES

Arnold, Vladimir, I., 1977, *Ordinary Differential Equations*, The MIT Press, Cambridge.

Kulenovic, Mustafa R. S., and Orlando, Merino, 2002, *Discrete Dynamical Systems and Difference Equations with Mathematica*, Chapman & Hall/CRC, Boca Raton.

Giunti, Marco, 1992, *Computers, Dynamical Systems, Phenomena and the Mind*, Ph.D. dissertation, Bloomington, in: Department of History and Philosophy of Science, Indiana University, (Microfilm published in 1993, University Microfilms Inc., Ann Arbor, MI).

Giunti, Marco, 1995, Dynamical models of cognition, in: *Mind as motion: Explorations in the dynamics of cognition*, in: R. F. Port and T. van Gelder, eds., The MIT Press, Cambridge, pp. 549-571.

Giunti, Marco, 1996, Beyond computationalism, in: *Proceedings of the 18th annual conference of the Cognitive Science Society*, Cottrel, Garrison W., ed., L. Erlbaum Associates, Mahwah, NJ, pp. 71-75.

Giunti, Marco, 1997, *Computation, Dynamics, and Cognition*, Oxford University Press, New York.

Giunti, Marco, 1998, Is computationalism the hard core of cognitive science?, in: *Prospettive della logica e della filosofia della scienza: Atti del convegno triennale della Società*

Italiana di Logica e Filosofia delle Scienze, Roma, 3-5 gennaio 1996, V. M. Abrusci, C. Cellucci, R. Cordeschi and V. Fano, eds., Edizioni ETS, Pisa, pp. 255-267.

Martelli, Mario, 1999, *Introduction to Discrete Dynamical Systems and Chaos*, Wiley, New York.

Sandefur, James T., 1990, *Discrete Dynamical Systems: Theory and Applications*, Oxford University Press, New York.

Szlensk, Wieslaw, 1984, *An Introduction to the Theory of Smooth Dynamical Systems*, John Wiley and Sons, Chichister, England.

Turing, Alan M., 1950, Computing machinery and intelligence, *Mind*, LIX, **236**:433-460.

THE ORIGIN OF ANALOGIES IN PHYSICS

Enzo Tonti
Università di Trieste; e-mail: tonti@units.it

Abstract: The paper gives the reasons for analogies in physics showing that they arise from the natural link between global physical variables and the space elements, i.e. points, lines, surfaces and volumes.

Key words: analogies; topological equations; constitutive equations.

1. INTRODUCTION

It is a common experience that when we are faced with a new problem, our mind runs quickly to our previous experiences to find similar problems already solved. A similar behaviour of the human mind arises also when we study a new field: in this case we search for something analogous, not necessarily equal, to the field we must study. The term "analogous" means that, to each element of one field there corresponds an element of the other (the corresponding elements are said to be *homologous*) in such a way that the *relations* between the homologous elements be the same. Such a correspondence is called *analogy*.

Analogies play a fundamental role in scientific research as well as in scientific education. An analogy between two phenomena permits one to make one phenomenon a model of the other; it permits one to transfer without difficulty many notions that are already known in one phenomenon to the other. Analogies are like roads in our mind that make the understanding of field of knowledge easy by using our previous familiarity with another field of knowledge.

Analogies have been used in engineering for *analogical models* of various phenomena. The most common among them is the electrical network model of discrete systems.

Analogies between different physical fields make various formalisms possible. Among them we mention

- the formalism of **dynamical systems theory**;
- the formalism of **generalized networks theory**;
- the formalism of the **mathematical field theory** that starts with variational principles (Lagrangians, Hamiltonians, conservation laws, Noether's theorems, etc.);
- the formalism of **irreversible thermodynamics** (generalized forces, generalized fluxes, phenomenological equations, the Onsager reciprocity principle, etc.);
- the formalism of the **first quantization** for dynamical systems
- and that of the **second quantization** for fields.

In physics analogies are usually grasped by comparing the equations of different fields of physics. The most widely known analogy is the one based on Poisson's and Laplace's equations which are common to electrostatics, magnetostatics, irrotational perfect fluid flow, stationary thermal conduction, gravitation, diffusion, torsion, capillarity, percolation, etc. Starting from the similarity of the equations of different physical fields one can infer the physical variables that are homologous in those fields. In this way one can build up a correspondence table among physical variables of the two fields.

In this approach one starts with the similarity of the equations and deduces the homologous variables.

In physics analogies are usually taken for granted: any attempt to explain their existence has been fruitless.

One of the most distinguished physicist of our time, Richard Feynman, raise the question: *Why are the equations from different phenomena so similar? We might say: "It is the underlying unity of nature." But what does that mean? What could such a statement mean? It could mean simply that the equations are similar for different phenomena; but then, of course, we have given no explanation. The "underlying unity" might mean that everything is made of the same stuff, and therefore obeys the same equations. That sounds like a good explanation, but let us think. The electrostatic potential, the diffusion of neutrons, heat flow — are we really dealing with the same stuff? Can we really imagine that the electrostatic potential is physically identical to the temperature, or to the density of particles? Certainly ϕ is not exactly the same as the thermal energy of particles. The displacement of a membrane is certainly not like a temperature. Why, then, is there an underlying unity? .... Is it possible that this is the clue? That the thing which is common to all the phenomena is the space the framework into which the physics is put?* (Feynman, vol.II, p.12-12)

We might ask ourselves the question: *Why is it that different physical fields display a common mathematical structure?* In other words: ***What is the origin of analogies in physics?*** Hence we can no longer accept the analogies as they arise from a comparison of different fields of physics, but, rather, we must ask why they exist.

This paper aims to give a possible explanation.

2. THE STARTING POINT

Since analogies in physics are revealed by the similarity of the equations in two different physical fields, it appears natural to start our investigation from an analysis of physical equations. Moreover, since physical equations link **physical variables**, it seems better to analyse physical variables directly. In fact an analogy is usually expressed by a correspondence table between the physical variables of two theories.

The new viewpoint we take in this analysis is to consider the process of creation of physical variables and, consequently, its measuring process.

Both a flux and a flow refer to **surfaces**. So we measure the magnetic flux through a surface, and the mass flow through a surface. The same is for the entropy flow, the momentum flow, the electric flux, the vortex flux, etc.

The entropy content, the mass content, the momentum content, the charge content, etc., all refer to **volumes**.

The line integral of a force, i.e. work, the line integral of velocity in fluid dynamics, the voltage in electrostatics, the electromotive and the magnetomotive forces in electromagnetism, all refer to **lines**.

Temperature, electric potential, displacement in continuum mechanics all refer to **points**.

In forming these examples we see the important fact that

the very definition of a physical variable and, consequently, its measuring process, leads us to the introduction of physical variables which refer to the four space elements to be defined as points, lines, surfaces and volumes.

We shall denote these four space elements with the symbols $\boldsymbol{P}$, $\boldsymbol{L}$, $\boldsymbol{S}$, $\boldsymbol{V}$ respectively. With the seeming exception for the variables associated with points, one sees that the other variables are not functions of points, but functions of lines, surfaces and volumes, i.e. they are **domain functions**.

From these domain functions we generally deduce the corresponding densities, by carrying out the ratio between a domain function and the extension of the associated space element (length, area and volume). In this way we obtain, for example, the mass density ρr, the magnetic flux density

Br, the electric field vector ***Er***, which are functions of point, i.e. the common *field functions*.

3. GLOBAL VARIABLES

We shall call *global variables* those variables that are associated with the four space elements *without being densities of another variable*. In this fashion we see that those variables which refer to points, such as temperature, electric potential, displacement of a point in continuum mechanics, are global variables since they are not densities. Global variables contain the *integral variables*, with the advantage that the adjective "global" makes no reference to the integration of field functions, i.e. to the reconstruction of a domain variable from its density. We consider densities to be obtained from global variables and not vice versa.

From what we have stated in the preceeding section, global variables are associated with space elements. So a flux is associated with a surface; an electromotive force with a line; etc. The density of a global variable *inherits* a reference to the space element with which the corresponding global variable is associated. So, since the magnetic induction vector $\boldsymbol{B}$ is the density of the magnetic flux, which refers to surfaces, we say that $\boldsymbol{B}$ *inherits* an association with surfaces. Since the charge density ρ is the density of the electric charge content, that refers to a volume, we say that ρ *inherits* an association with volumes. This inherited association will be denoted by putting the space element into square brackets. Thus we can write both $\rho\boldsymbol{r}$ and $\rho[V]$, $\boldsymbol{B}(\boldsymbol{r})$ and $\boldsymbol{B}[S]$, $\boldsymbol{Er}$ and $\boldsymbol{E}[\boldsymbol{L}]$.

We can now state our first rule: *the homologous variables of two fields of physics are those associated with the same space element.*

4. ORIENTATION OF SPACE ELEMENTS

Space elements, i.e. points, lines, surfaces and volumes can be endowed with orientation. There are two kinds of orientation: the *inner* and the *outer* one, as shown in Fig. 1. We shall avoid giving here a detailed presentation of these two notions as they would need a longer paper, and we send the reader back to some of the previous papers by the author (Tonti, 1976, 2001, 2001a). We hope that Fig. 1 gives an intuitive idea of the distinction between the two notions.

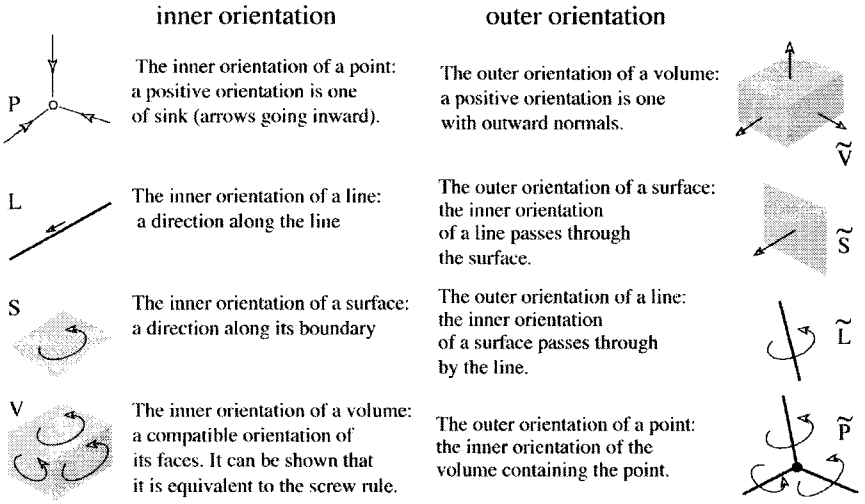


Figure 1. The two kinds of oriented space elements: (left) inner orientation; (right) outer orientation.

A deeper analysis of global physical variables shows that some of them are associated with space elements endowed with inner orientation, while others are endowed with outer orientation. Thus a matter flow requires a surface with an outer orientation because matter passes through the surface from one side to the other. On the contrary, a magnetic flux requires a surface endowed with inner orientation because the sign of the flux associated with a loop depends on the direction of the current in the loop. In a magnetic flux nothing passes through the surface and, therefore, an outer orientation is not involved. In an analogous way the electromotive force is associated with a line endowed with inner orientation, while a magnetomotive force requires a line endowed with outer orientation: this can be seen by analysing the Biot-Savart law (see Tonti, 2001).

5. CONFIGURATION, SOURCE AND ENERGY VARIABLES

Physical quantities are divided into two main classes: physical *parameters* and physical *variables*. Material constants, physical constants, universal constants, system parameters, etc., belong to the class of parameters.

All quantities that describe the state and the configuration of a system (displacement, velocity, strain, etc.); the sources of a field (forces, charges,

currents, pressures, etc.); all kinds of energy (kinetic, potential, internal, chemical, etc.); the various potentials (electric, gravitational, chemical, thermal, etc.), belong to the class of variables.

Physical variables, in turn, can be divided into three classes:

- **configuration** variables that describe the *configuration* of a system or of a field;
- **source** variables that describe the *sources* of a field;
- **energy** variables obtained as a product of a configuration for a source variable.

So in particle mechanics the radius vector and the velocity are configuration variables, force and momentum are source variables, while work, kinetic and potential energy are energy variables.

At this point one comes to a remarkable discovery:

global configuration variables are associated with space elements endowed with inner orientation while global source variables are associated with space elements endowed with outer orientation.

This property, which the author of this paper is unable to justify, allows one to give a detailed classification of physical variables for all physical fields.

To this end it is essential to make use of a subdivision of a space region into cells, as we shall describe in the following section.

6. CELL COMPLEXES

Space elements, i.e. points, lines, surfaces and volumes, are clearly exibited by means of a *cell complex*. If we divide a space region into cells of whatever shape and dimensions, e.g. into cubic cells, we can consider vertices, edges, faces and the cells themselves as representative of points, lines, surfaces and volumes respectively. In so doing we discretize the space region. In algebraic topology vertices are called 0-cells, edges are called 1-cells, faces are called 2-cells, the cells themselves are called 3-cells. In general one speaks of p-dimensional cells or, briefly, of p-cells.

If we now consider the barycenters of the cells, and connect each one to those of the adjacent cells, we build up a second cell complex, called the *dual complex*, as shown in Fig. 2, *left*.

The first cell complex, called the *primal* complex, can be endowed with an inner orientation. To this end one must assign an inner orientation to all its vertices, edges, faces and cells. As can easily be seen in Fig.2, the cells, faces, edges and vertices of the dual complex are automatically endowed with an outer orientation.

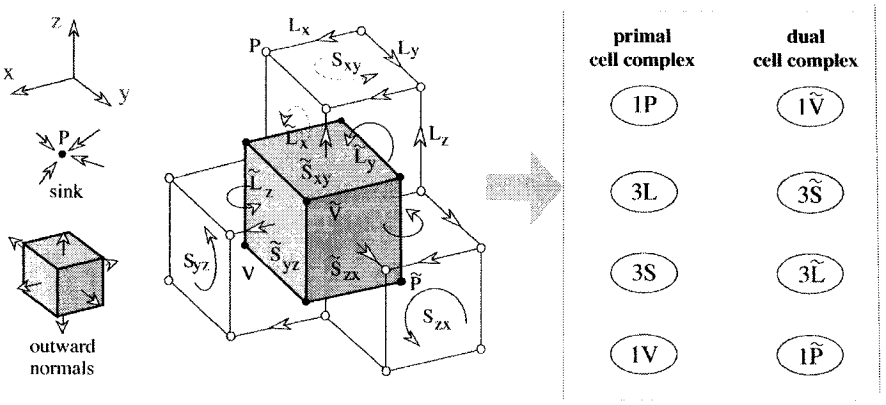


Figure 2. A cartesian cell complex showing a primal and a dual complex. The inner orientation of the primal complex induces an outer orientation on the dual.

This leads us to consider a diagram like the one of Fig. 2, *right* which organizes the 4 x 2 = 8 space elements into boxes ordered in two columns: the left column containing the space elements of the primal complex, endowed with inner orientation, while the right column contains the space elements of the dual complex, endowed with outer orientation.

At this point we have obtained a classification diagram of *space elements*.

7. A CLASSIFICATION DIAGRAM

Since global physical variables are directly associated with space elements, and since field functions have an inherited association with space elements, it becomes natural to use the classification diagram of space elements just obtained to classify physical variables of every physical theory. The procedure is that of inserting each variable of a given field into the box of the corresponding space element. In this way we obtain a classification diagram of *physical variables* like the one of electrostatics shown in Fig. 3.

A cell complex is particularly suited to describe the mathematical structure of a physical theory. Let us analyse some advantages. First of all, in a discrete formulation of physics, cell complexes play the same role that coordinate systems play in the differential formulation. The fact that global variables are directly associated with space elements implies that, to every element of a cell complex, say vertex, edge, face and cell, there is a corresponding physical variable. It follows that, for a variable associated with lines endowed with inner orientation, such as voltage in electrostatics,

one can consider the amount of this variable associated with each edge of the complex.

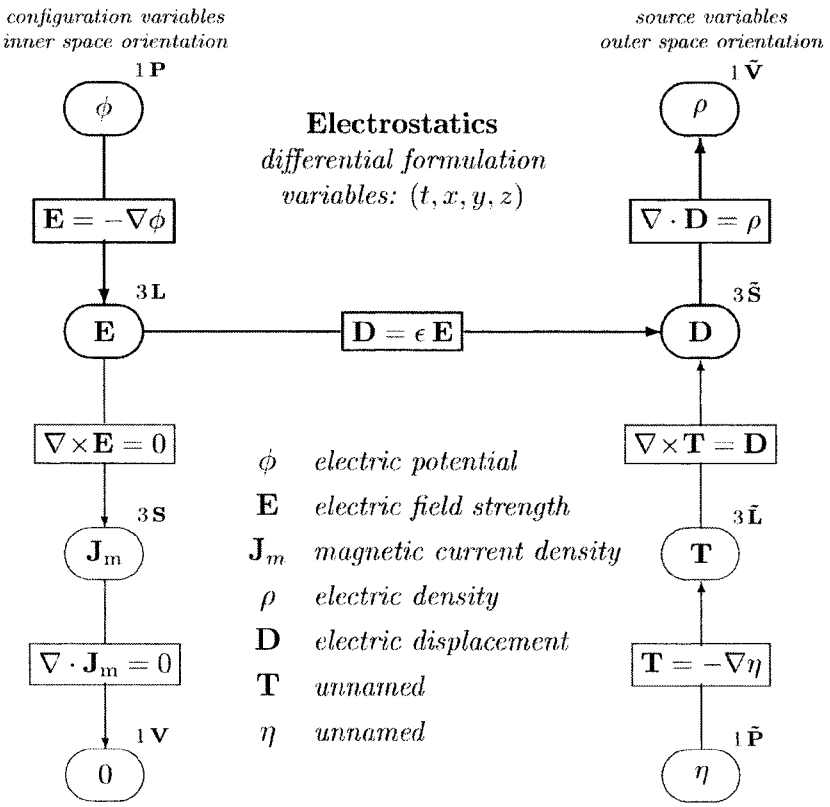


Figure 3. The diagram of electrostatics. In each box we put the field variables that inherited an association with the corresponding space element.

This distribution of the voltage on the various 1-cells (edges) of a cell complex will be called *one dimensional distribution*; the distribution of a flux on each 2-cell (face) will be called *two dimensional distribution*; etc.

In *algebraic* topology, one of the three branches of topology (the other being *analytic* topology and *differential* topology), such p-dimensional distributions have a strange name: p-dimensional *cochains*. This term means "complementary" to a chain which is a collection of cells.

Hence, instead of considering point functions (field functions), which are typical of the differential formulation, one is led to consider p-dimensional distributions.

Physics has, up to now, been described by using the differential formulation, i.e. by using total and partial derivatives, differential equations, differential operators such as gradient, curl and divergence. In order to do so we need field variables, i.e. point functions.

Since we measure mainly global variables, in order to obtain field variables we must introduce densities. In so doing we strip global variables of their geometrical content. Thus in thermodynamics temperature T, pressure p and density ρ are considered as intensive variables, which is quite right, but pressure is obtained from the normal force on a surface and therefore it is associated with surfaces, while mass density is the ratio between mass and volume and therefore is associated with volumes. Hence we can write $T\,[\boldsymbol{P}]$, $p[\boldsymbol{S}]$, $\rho\,[\boldsymbol{V}]$. This relation with space elements is essential for the classification of such physical variables. The use of a cell complex makes a proper collocation of global physical variables and their densities into the classification diagram possible.

One of the consequences of this classification is that it allows one to separate the basic equations of every physical theory into two large classes: the *topological* and the *constitutive* equations.

8. TOPOLOGICAL EQUATIONS

Let us start by considering a balance law, say the balance of mass, of energy, of electric charge, of entropy, of momentum, of angular momentum. A balance links different aspects of a common extensive physical variable. Thus, entropy balance states that the entropy *production* inside a volume during a time interval is split into two quantities: the entropy *stored* inside the volume in the time interval and the entropy which has flowed outside the boundary of the volume, the so called *outflow*, during the time interval. In short

production = **storage** + **outflow**.

What is remarkable in a balance is that the shape and the extension of the volume and the duration of the time interval are arbitrary. No metrical notions, such as length, area, volume (say cubic meters) are involved and no measure of duration is required (say, seconds). This means that a balance is a **topological** law.

Moreover, it is not an experimental law but, rather, an *a priori* law of our mind. Only experience can say whether there is a storage, an outflow or a production. So, if the production vanishes, there is *conservation* (e.g..

electric charge); if the storage vanishes, the system is stationary; if the outflow vanishes, the system is *closed* (adiabatic, isolated).

Topological equations, when written in a differential formulation, give rise to the divergence, the curl and the gradient, and this introduces metrical notions. Hence the differential formulation mixes topological and metrical notions, while a discrete formulation, using global variables, avoids this mixing.

Let us consider a second class of equations, that of circuital laws. The prototype can be the Biot and Savart law of electromagnetism: the magnetomotive force along a closed loop that is the boundary of a surface is equal to the electric current passing though the surface. Also in this case the shape of the surface, and therefore of its boundary, is immaterial. Thus a circuital law is also a **topological** law.

A third class of equations is the following. Let us consider a physical variable associated with points, say temperature. Let us consider two points and a line connecting them. One can form the difference of the values of the variable at the two points and assign this difference to the line connecting them. Also in this case the line connecting the two points can have any shape and extension and, therefore, the process of forming the difference is a **topological** one.

These three topological equations relate global physical variables as follows:

1. **balance law**: a physical variable associated with the boundary of a volume (surface) is linked to a physical variable associated with the volume;
2. **circuital law**: a physical variable associated with the boundary of a surface (line) is linked to a physical variable associated with the surface;
3. **forming the differences**: a physical variable associated with the boundary of a line (points) is linked to a physical variable associated with the line.

In algebraic topology the process of assigning a physical variable which is associated with the boundary of a p-dimensional cell with the p-cell itself is called the *coboundary process*. This process is described by an operator, called *coboundary operator*, which corresponds to the *exterior differential* on differential forms.

In the classification diagram, topological equations link a physical variable contained in a box with another physical variable contained in the box which immediateley follows in the same column. These links are the vertical ones in the diagram. Topological equations are valid in the large as well as in the small, and are valid even if the region is filled with different materials.

All that we have said regarding space elements can also be said for time elements (instants and intervals) and for space-time elements. In Fig. 4 we show the space-time diagram of electromagnetism.

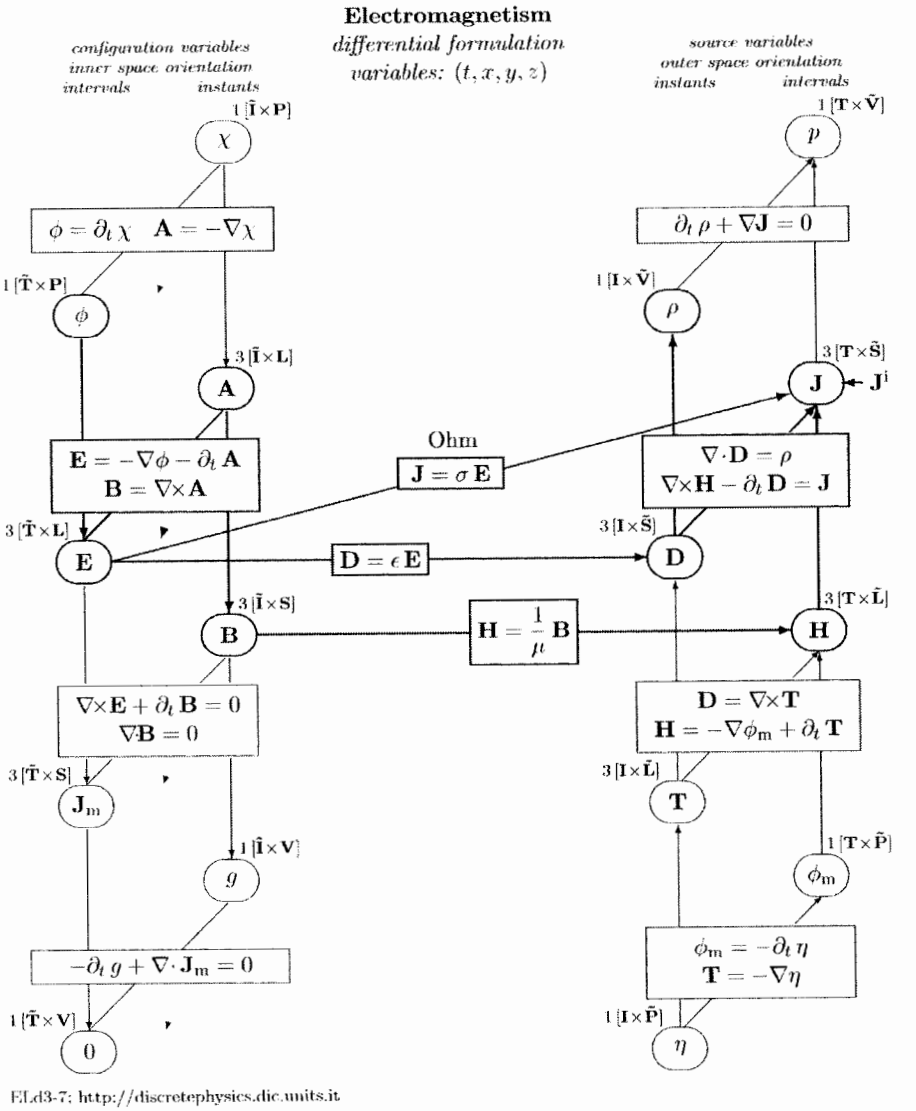


Figure 4. The space-time diagram of electromagnetism.

9. CONSTITUTIVE EQUATIONS

Another class of physical laws is that of constitutive laws which specify the behaviour of a material and hence contain material parameters. Constitutive laws link a configuration variable which refers to a space element with a source variable referring to the dual element. These links are the horizontal ones in the diagram. They contain metrical notions such as lengths and areas. They are experimented in regions in which the field is uniform, and hence are valid only in the small.

10. CONCLUSION

Analogies in physics follow from the fact that global physical variables are associated with space elements. Homologous variables are those which refer to the same space elements.

Once space elements have been organized into a diagram, the associated physical variables can be inserted in the same diagram. In this fashion many mathematical and physical properties are clearly exhibited. The diagram reveals a mathematical structure that is common to many physical fields.

REFERENCES

All papers quoted here can be downloaded from: <http://discretephysics.dic.units.it>. Diagrams of many physical fields can be downloaded from this web site. For a more detailed explanation of the diagrams and how to build them one can download the file Diagram Explained from the same web site.

Feynman R. P., Leighton R. B., and Sands M., 1964, *Lectures on Physics*, Addison Wesley, Reading, MA.

Tonti E., 1976, The Reason for analogies between physical theories, *Appl. Math. Modelling* **1**:37-50.

Tonti E., 2001, Finite formulation of the electromagnetic field, *Progress in Electromagnetics Research, PIER* 32 (Special Volume on Geometrical Methods for Comp. Electromagnetics), pp. 1-44.

Tonti E., 2001a, A direct discrete formulation of field laws: the cell method, *Computer Modelling in Engineering and Science, CMES* **2**(2):237-258.

PRISONER DILEMMA: A MODEL TAKING INTO ACCOUNT EXPECTANCIES

Natale S. Bonfiglio and Eliano Pessa
Dipartimento di Psicologia, Università di Pavia, Piazza Botta 6, 27100 Pavia, Italy

Abstract: This paper introduces a new neural network model of players' behavior in iterated Prisoner Dilemma Game. Differently from other models of this kind, but in accordance with theoretical framework of evolutionary game theory, it takes into account players' expectancies in computation of individual moves at every game step. Such a circumstance, however, led to an increase of the number of model free parameters. It was therefore necessary, to search for optimal parameter values granting for a satisfactory fitting of data obtained in an experiment performed on human subjects, to resort to a genetic algorithm.

Key words: prisoner dilemma; evolutionary game theory; neural network; genetic algorithm.

1. INTRODUCTION

Evolutionary game theory (May, 1974; Maynard Smith, 1982; Axelrod, 1984; Akiyama and Kaneko, 2000; Gintis, 2000) was introduced to account for the fact that in most interactions between individuals evidence how altruism be responsible for the emergence of evolutionarily stable cooperative behaviors. Within this framework the outcome of a game, or the strategies used by the players, cannot be predicted in advance on the only basis of a previous knowledge of individual characteristics of the players themselves. Namely the decisions about the move to be performed at a given game step depend essentially on past game history, on momentary players' goals, and on their expectancies. Within this context the concept itself of equilibrium loses its predictive value (see, for a simple example, Epstein and Hammond, 2002).

The attempts to test the validity of such ideas gave rise to a conspicuous theoretical and experimental work, dealing mostly with players' behaviors in incomplete information games, among which one of most studied has been iterated Prisoner Dilemma Game (IPDG). In this paper we will introduce a neural network model of the latter, proposed to account for data obtained in an experiment performed on human subject pairs engaged in a particular version of this game. In this context the idea of introducing a neural network modeling of players' cognitive system is not new (see, for instance, Cho, 1995; Taiji and Ikegami, 1999). What is new in the present model is that it takes into account in an explicit way players' expectancies in correspondence to every new move to be done. The price to pay for the introduction of such a feature has been the increase of number of free model parameters. Of course, this is a drawback when searching for optimal parameter values, in order to fit in the best way experimental data. Thus, we resorted to a genetic algorithm in order to find them. In this regard, we obtained that the latter gave a satisfactory result (that is it was able to find the searched values). Even if this could be considered as a proof of model suitability, we tested whether such a result was or not depending on the explicit presence of expectancies in our model. Thus, we applied the same genetic algorithm to a simplified version of previous model, without expectancy computation mechanism. In the latter case the genetic algorithm performed very poorly, indicating that the role of expectancy in accounting for human players' behavior in IPDG was essential.

2. THE EXPERIMENT WITH HUMAN SUBJECTS

A sample of 30 player pairs (whence 60 subjects) was engaged in an "economic" version of IPDG, characterized by the payoff matrix described in Table 1.

Table 1. Payoff matrix of our version of IPDG. The symbol C_i denotes cooperation by the i-th player, while D_i denotes defection by the i-th player. Every cell of this matrix contains two values, separated by a comma: the one on the left denotes the payoff of first player, while the one on the right denotes the payoff of second player. Both payoff are expressed in Italian money.

	C_2	D_2
C_1	5000 , 5000	- 25000 , 30000
D_1	30000 , -25000	0 , 0

Before the starting of experiment each player was individually tested on his/her motivational attitudes, to ascertain whether he/she was *cooperative* (aiming at maximizing the payoff of both players), *competitive* (aiming at

minimizing opponent's payoff), or *individualist* (aiming at maximizing his/her own payoff). We followed, in this regard, the procedure introduced by Kuhlmann and Marshello (1975).

During the experiment, a great deal of care was devoted to prevent from any physical contact between the two players of each pair, be it under a visual, auditory, or even olfactory form. Each IPDG was always consisting of 100 steps, even if both players were never informed in advance on total game duration. At any game step, after both players made separately their move, each player received, as feedback, the value of obtained payoff, as well as the information on the move of his/her opponent. As such communication was performed through a paper sheet, always the same in all game steps, every player had the opportunity (at least in principle) of observing, before every move decision, all past game history.

3. THE NEURAL NETWORK MODEL

Each player is modeled through a feedforward neural network architecture, whose overall structure is depicted in Figure 1. The player move depends on the output of Move unit (on the top of Figure 1), which is a standard threshold unit receiving four inputs, respectively from the output of module performing the comparison between expectations at time t and at time t-1, from memorized value of previous player move at time t-1, from memorized value of opponent's move at time t-1, and from player payoff at time t-1. The output of Move unit is computed according to the following prescription:

$$u = 1 \text{ if } P > 0, \quad u = -1 \text{ if } P < 0, \quad P = \sum_i p_i y_i - s$$

where y_i (i = 1,...,4) denotes the input signal coming to Move unit along the i-th input line and p_i is the connection weight of this line. The symbol s denotes a suitable threshold value. The weights p_i vary with time according to the following law:

$$p_i(t+1) = \sin \frac{t}{A_i - p_i(t)}$$

where A_i are suitable parameters. Moreover, even the threshold value varies with time according to the rule:

$$s(t+1) = s(t) + s_{\max} - \varepsilon s(t) + \eta G(t)$$

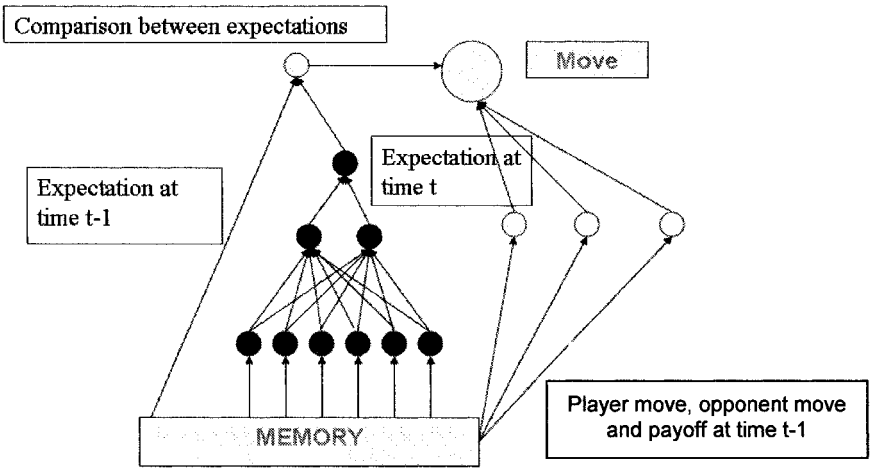


Figure 1. The overall architecture of neural network modelling player behaviour.

in which $s_{\max}$, ε, η are further parameters and $G(t)$ denotes the player gain at time t.

The module performing a comparison between expectations at different times produces an output given by the following rule:

$$y = 1 \text{ if } Q > 0; \quad y = -1 \text{ if } Q < 0; \quad Q = w_1 a(t) + w_2 a(t-1)$$

Here the two weights w_1 and w_2 are, in turn, varying as a function of performed moves and of obtained gain. Their variations obey the laws:

$$w_1(t+1) = w_1(t) - \alpha w_1(t) - \delta G(t)$$
$$w_2(t+1) = w_2(t) - \beta w_2(t) - \gamma G(t) m_1(t) m_2(t)$$

The symbols α, β, γ, δ still denote suitable parameters, while $m_1(t)$ and $m_2(t)$ are, respectively, the player move and the opponent move at time t. Finally, the module computing the expectation at time t consists of a three-

layered Perceptron, with one unit in the output layer, 4 units in the hidden layer and 10 units in the input layer receiving the memorized values of player and opponent moves at times $t - 1$, $t - 2$, $t - 3$, $t - 4$, $t - 5$ (we assumed a finite extension of memory for past moves). Output and hidden units had, as activation function, the hyperbolic tangent.

4. PARAMETER OPTIMIZATION

At first sight, it appears very difficult to find the optimal parameter values granting for a correct reproduction by this model of at least particular games involving human subjects. Namely the number of free parameters is very high: 66, of which 55 are given by connection weights of Perceptron described above. Besides, the parameter values of one player could differ from the ones of opponent player, a circumstance which raises the total number of free parameters up to 132. On the other hand, only the existence of a small difference between an IPDG played by two neural network models and an IPDG played by two specific human subjects can support the hypothesis that this model describe in a correct way the cognitive operation of players engaged in this sort of game and, therefore, the validity of evolutionary game theoretical framework on which the neural architecture itself was based.

Owing to the difficulty of parameter optimization task, we resorted to a genetic algorithm (see Goldberg, 1989), based on a real number coding, in which each individual was described by a genome with 132 genes, coding the different parameter values of each one of two player. As fitness function, to be maximized, we choose the expression $4\ N/(1 + mindist)$, where N is the total number of steps in a single game (100 in our experiment on human subjects) and *mindist* is the minimal distance between the game played by two neural networks, whose parameters are given by the individual genome taken into consideration, and the set of games really played by human subject in the experiment described above. Here the distance was computed as the sum of absolute differences between the values of corresponding moves in the two games. The choice of such a fitness function means that we were not searching for the best fitting of the average human player behaviour, but for the best fitting of at least *one* game played by human subjects. We recall, in this regard, that the use of genetic algorithms in investigating IPGD is not new (see, for instance, Miller, 1996).

We applied our genetic algorithm to two different kinds of models: 1) the complete neural network model described in previous section; 2) the same model as in 1), but without the modules computing expectations and comparison between expectations. The latter was investigated as, because

the key role played by expectations in a evolutionary game theoretical framework, we were interested in checking whether, even in absence of expectations, a simplified neural network architecture would, in principle, be able to reproduce correctly at least a particular game played by human subjects.

Computer simulations, in which we monitored, at each generation, the maximum fitness value within the population, evidenced how, only in the case of models of the kind 1) it was possible to reach a very high value of maximum fitness. On the contrary, in the case of models without expectation, the genetic algorithm was unable, even with high numbers of generations, to produce a satisfactorily high value of maximum fitness. In the Figure 2 we show an example of trend of maximum fitness vs generation number in the case of a model of type 1), while in the Figure 3 we show an example of the same trend in the case of a model of type 2), that is without expectations.

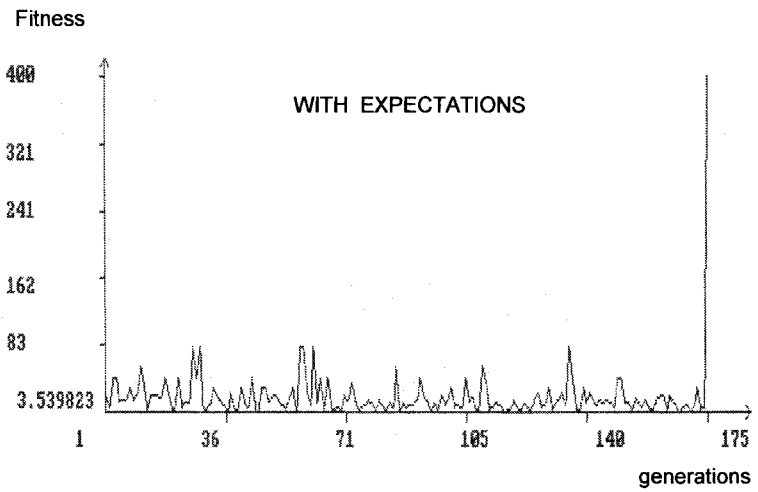


Figure 2. Maximum fitness vs number of generations for a model of type 1). It is to be remembered that maximum possible fitness value is 400, reached after 175 generations.

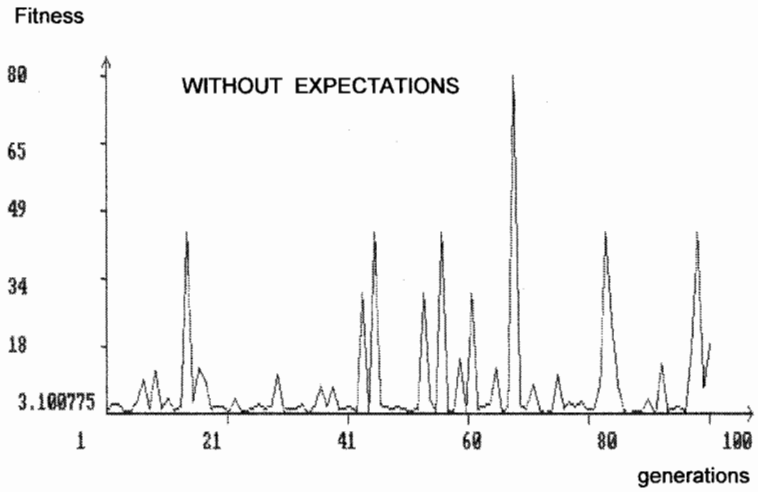


Figure 3. Maximum fitness vs number of generations for a model of type 2). The maximum fitness reached is 80, very far from maximum possible value of 400.

5. CONCLUSION

The application of genetic algorithm evidenced how the presence of expectations in our model neural network be essential in order to reach a parameter value optimization granting for the correct reproduction of at least one game played by human subjects. This seems to support our modelling choices and, therefore the adoption of an evolutionary game theoretical framework on which our model was based. However, further investigations will be needed to give a solid ground to evolutionary game theory which, despite its successes in describing many biological behaviours, is often viewed with a sort of scepticism when applied to account for human decision making, despite the drawbacks of rational decision theories (see, for a review, Busemeyer and Townsend, 1993).

REFERENCES

Akiyama, E., and Kaneko, K., 2000, Dynamical systems game theory and dynamics of games, *Physica D* **147**:221-258.

Axelrod, R., 1984, *The Evolution of Cooperation*, Basic Books, New York.

Busemeyer, J. R., and Townsend, J. T., 1993, Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment, *Psychological Review* **100**:423-459.

Cho, I.-K., 1995, Perceptrons play the repeated Prisoner's Dilemma, *Journal of Economic Theory* **67**:266-284.

Epstein, J. M., and Hammond, R. A., 2002, Non-explanatory equilibria: an extremely simple game with (mostly) unattainable fixed points, *Complexity* **7**:18-22.

Gintis, H., 2000, *Game Theory Evolving*, Princeton University Press, Princeton, NJ.

Goldberg, D., 1989, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA.

Kuhlmann, D. M., and Marshello, A. F. J., 1975, Individual differences in game motivation as moderators of preprogrammed strategy effects in Prisoner's Dilemma, *Journal of Personality and of Social Psychology* **32**:922-931.

May, R. M., 1974, *Stability and Complexity in Model Ecosystems*, Princeton University Press, Princeton, NJ.

Maynard Smith, J., 1982, *Evolution and the Theory of Games*, Cambridge University Press, Cambridge, UK.

Miller, J. H., 1996, The coevolution of automata in repeated prisoner's dilemma, *Journal of Economic Behavior and Organization* **29**:87-112.

Taiji, M., and Ikegami, T., 1999, Dynamics of internal models in game players, *Physica D* **134**:253-266.

THE THEORY OF LEVELS OF REALITY AND THE DIFFERENCE BETWEEN SIMPLE AND TANGLED HIERARCHIES

Roberto Poli
University of Trento and Mitteleuropa Foundation

Abstract: The main features of the theory of level of reality are presented. The conceptual framework according to which levels follow a linear, brick-like order is opposed to a more sophisticated, "tangled" framework.

Key words: level of reality; stratum; layer; hierarchy; dependence; autonomy.

1. INTRODUCTION

Most discussion about levels is focused on levels of description. The topic of the levels of description is obviously important, but I do claim that it should be kept as separate as possible from the problem of the levels of reality. Although confusion between the two planes is not infrequent, their names themselves indicate that they occupy different places in a well structured conceptual framework. The levels of reality have a strictly ontological significance, while those of description have an epistemological one. The presence of intermediate or ambiguous cases does not authorize one to confound categorical specificities. The distance that separates the two themes is therefore the same distance that separates epistemology from ontology. Whatever the relationships between them (of opposition, connection, inclusion, or anything else) may be, they are replicated in the difference between (levels of) description and (levels of) reality.

In what follows I shall restrict my discussion to only certain aspects of the problem of levels of reality. Consequently, I shall be concerned with ontological matters. I shall not address the question of the relationships between ontology and epistemology. Indeed, I shall take care not to slide

from one plane to the other (for an outline of my view on the relationship between epistemology and ontology see Poli, 2001b; for a general presentation of my views on levels see Poli, 1997, 2001a, 2001b; Heller, Herre and Poli, submitted; Gnoli and Poli, submitted).

An intuitive understanding of the basic problem of the theory of levels will facilitate subsequent analyses. The following section is based on my (2001a).

2. HOW MUCH INFORMATION IS THERE?

Let's consider the pen in front of me on my desk. What type of object is this pen? How should I model it? First of all, I may say that the pen is an object made in a certain way, with its own shape, colour and material. In saying this, I am using concepts which describe the physical world of things. The pen must also perform functions: it has been designed to write. This reference to function introduces a different dimension into the analysis: writing, in fact, is not something that I can model using only concepts describing the physical world. Writing is an activity typically performed by humans. By virtue of being constructed to fulfill the function of writing, the pen is in some way connected with this aspect of the world. But when I observe the pen, it tells me many other things. For example, that it has been constructed by somebody, and that this somebody is my contemporary: this pen is not an object from the Roman age or from ancient China. The material it is made, its manufacture, the way it works tell me that there must be somewhere an organization that produces things like pens. If we now shift our focus to this organization, the pen must be an object designed, manufactured and distributed so that it can be sold and end up on someone's desk. In their turn, the points of view of the designer, of the production department and of the distribution department are different, and they describe my pen using different concepts. For the designer the pen is essentially an aesthetic and functional object; for the production department it is the outcome of materials processed in a certain way, etc. For the company producing the pen it is all these things together. For the shopkeeper who displays the pen on his shelves and seeks to sell it to customers, it is again different. To return to myself, the pen is also an object of which I got especially fond because it reminds me of the person who gave it to me.

All these different descriptions are correct: each of them express a facet of the object. Yet they are all descriptions of the same object. Hence, one of the main tasks of information science is to find ways to integrate different descriptions of the same object. Some of these descriptions are biased toward the observer, some other are biased toward the object. Both cases

articulate the basic situation composed by an observing system and an observed systems (von Foster, 1984; Rosen, 1978; to my knowledge, the best development of a system-based ontology is Weissmann, 2000).

Ontologically, the example of the pen teaches us two important lessons: (1) reality is organized into strata (material, psychological, social); (2) these strata are organized into layers (the physical and chemical layers of the material stratum; the intentional and emotive layers of the psychological stratum; the productive, commercial and legal layers of the social stratum).

3. THEORIES OF LEVELS AND THEIR AUTHORS

Not many thinkers have systematically worked on the theory of levels of reality. We may conveniently distinguish the "English-writing" camp from the "German-writing" one. The former comprises, among many others, thinkers such as Spencer, Alexander, and Lloyd-Morgan (possibly the deepest figure among those quoted). Blitz (1992) provides a reliable synthesis of their main contributions. The "German-writing" camp comprises thinkers as relevant as Husserl, Ingarden, Plessner, and Hartmann. Even if some of them are very well known names, there is no academic work summarizing their contributions to ontology in general and to the theory of levels in particular. Unfortunately, thoroughgoing comparison between the "English" and the "German" camps is lacking.

4. WHAT IS A LEVEL OF REALITY?

No general consensus exists about how to define, describe or at least sketch the idea of level of reality. My own choice is to adopt a categorical criterion: the levels of reality are characterized (and therefore distinguished) by their categories. The main subsequent distinction is between universal categories (those that pertain to reality in its entirety – time, whole/part, substance/determination, etc.) and categories that pertain solely to one or some levels of reality.

Most authors prefer instead to adopt an objectual standpoint, rather than a categorical one. Arguing in favor of the objectual standpoint has the undoubted advantage that it yields an elementary definition of level: a level consists of a collection of units (Pattee, 1973, p. 75). From this point of view, the series of levels is a series of objects interacting at different degrees of granularity. A model of this kind is accepted by large part of the scientific community, because it depicts the widely held view of levels based on a reductionist approach. Higher-order groups of items may behave differently,

even to the point that it is impossible to calculate (predict) their specific behaviour, but in the end what matters is that they can all be reduced to their *atoms*.

If this were indeed the way matters stand, then the general neglect shown towards the problem of the levels would be justified.

In order to deal with the real complexity of the problem of the levels, it must be altered so that it becomes possible to study not only 'linear' hierarchies but 'tangled' ones as well. This conclusion bears out the approach which undertakes categorical analysis, compared to the one which studies items in iteration.

An argument in favor of the approach 'by objects' is the ease with which it is possible to pass from a substantialist description to a processualist one: if a level is defined by items in iteration (where the items can be canonically conceived as objects), then a level can be defined by a dynamics. A multiplicity of structurally stable dynamics, at diverse levels of granularity, may define a multiplicity of levels. However, if it turns out that the structuring in levels does not respect a universal principle of linearity, then one is forced to restrict the multidynamic frames to their linear fragments. Which is precisely the situation of current theories of dynamic systems. On careful consideration, in fact, the predominant opinion is that there is only one multi-dynamic (multi-layered) system: the one described by the natural sciences. Other forms of knowledge are scientific to the extent that they can be located in the progressive series of supraformations (groups of groups of groups of items, each with its specific kinds of interaction). Hence the alternative: a discipline is scientific to the extent that it can be located in the series of aggregation levels – if so it can be more or less easily reduced to the base level – or it cannot be thus located and is consequently not a science: it has no citizenship in the realm of knowledge and is scientifically stateless.

5. THE THREE MAIN STRATA OF REALITY

The distinction is widespread among three basic realms or regions (or strata, as I will call them) of reality. Even if the boundaries between them are differently placed, the distinction among the three realms of material, mental and social phenomena is essentially accepted by most thinkers and scientists. A major source of discussion is whether inanimate and animate beings should be placed in two different realms (this meaning that there are in fact four and not three realms) or within the same realm. The latter option defends the thesis that a phase transition or something like that connects inanimate and animate items.

From a categorical point of view, the problem about how many strata there are can be easily solved. Leaving apart universal categories (those that apply everywhere), two main categorical situations can be distinguished: (a) Types (Items) A and B are categorically different because the description / codification / modelling of one of them requires categories that are not needed by the description / codification / modelling of the other; (b) Types (Items) A and B are categorically different because their description / codification / modelling requires two entirely different groups of categories. Following Hartmann, I term the two relations respectively as over-forming and building-above. Strata or realm of reality are connected by building-above relations. That is to say, the main reason for distinguishing as clearly as possible the different strata of reality is that any of them is characterized by entirely different categorical series. The group of categories that are needed for analyzing the phenomena of the psychological stratum is essentially different from the group of categories needed for analyzing the social one, which in its turn is different from the one needed for analyzing the material stratum of reality.

Over-forming (the type (a) form of categorical dependence) is weaker than building-above and it is used for analyzing the internal structure of strata. Each of the three strata of reality has its specific structure. The case of the material stratum is the best known and the least problematic. Suffice it to consider the series atom-molecule-cell-organism (which can be extended at each of its two extremes to include sub-atomic particles and ecological communities, and also internally, as needed). In this case we have a clear example of a series that proceeds by levels of granularity. The basic distinction of the realm (stratum) into physical, chemical and biological components can be considerably refined (e.g., by distinguishing biology into genetics, cytology, physiology, ethology, ecology – a slightly more articulated picture is provided by Poli (2001b). Compared to the material realm, the psychological and social ones are characterized by an interruption in the material categorical series and by the onset of new ones (relative to the psychological and social items). More complex types of over-forming are instantiated by them. The basic situation is sketched in Poli (2001b). However, much work is still required.

A terminological note can be helpful. I use the term 'level' to refer in general to the levels of reality, restricting the term 'layer' to over-forming relationships, and the term 'stratum' to building-above relationships.

6. FORMS OF CONNECTION AMONG STRATA

The question now arises about how the material, psychological and social strata are connected together. The most obvious answer is that they have a linear structure like the one illustrated by Figure 1.

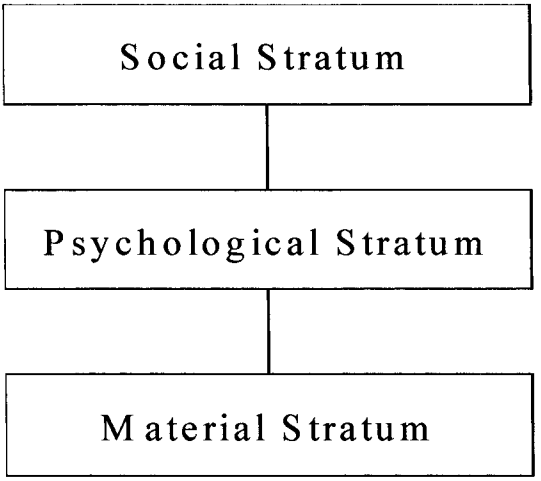


Figure 1. Linearly organized strata.

On this view, the social realm is founded on the psychological stratum, which in its turn is founded on the material one. Likewise, the material stratum is the bearer of the psychological stratum, which in its turn is the bearer of the social one.

The point of view illustrated by Figure 1 is part of the received wisdom. However, a different opinion is possible. Consider Figure 2.

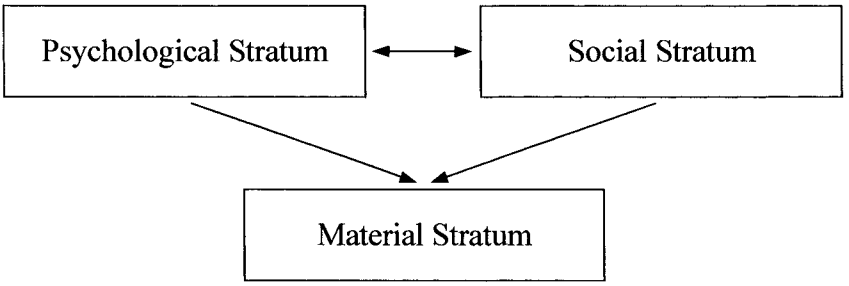


Figure 2. The architecture of strata with bilateral dependence.

Material phenomena act as bearers of *both* psychological *and* social phenomena. In their turn, psychological and social phenomena reciprocally determine each other. Psychological and social systems are formed through co-evolution: the one is the environmental prerequisite for the other (Luhmann, 1984).

7. CAUSATION

The theory of levels of reality is the natural setting for elaboration of an articulated theory of the forms of causal dependence. In fact, it smoothly grounds the hypothesis that any ontologically different level has its own form of causality (or family of forms of causality). Material, psychological and social forms of causality could therefore be distinguished (and compared) in a principled way.

The further distinction between causal dependence (between items) and categorical dependence (between levels) provides means for elaborating a stronger antireductionist vision.

The architecture of levels we have quickly sketched grounds one facet of the claim. As a matter of fact, it is much easier to advocate reductionism if the levels are structured in a serial, linear order. Reductionism will have an even worse currency as soon as the problem (not considered in this paper) of the internal organization of the strata is considered. I have shown elsewhere (e.g., in Poli, 2001b) that the internal organization of each stratum is structurally different. This contributes making reduction to the lower layer of the lower stratum simply unobtainable.

Beside the usual kinds of basic causality between phenomena of the same nature, the theory of levels enables us to single out upward forms of causality (from the lower level to the upper one). But this is not all. A theory of levels also enables us to address the problem of *downward* forms of causality (from the upper to the lower level). The point was first advanced by Donald Campbell some years ago (see e.g. 1974 and 1990). Andersen et al. (2000) collects a series of recent studies on the theme.

8. VIRTUOUS CIRCULARITY

The connection between the theory of levels and causality entails recognition that every level of reality may trigger its own causal chain. This may even be taken as a definition of level of reality: A level of reality is distinguished by its specific form of causality. As a consequence, we thus

have a criterion with which to distinguish among levels of reality and levels of description.

The acknowledgement also enables us to develop a theory able to accommodate different senses of causality (distinguishing at least among material, mental and social causality). However, if the downward option is also available, the direct or elementary forms of causality should have corresponding non-elementary situations.

REFERENCES

Andersen, P. B., Emmeche, C., Finnemann, N. O. and Christiansen, P. V., eds., 2000, *Downward Causation. Minds, Bodies and Matter*, Aarhus University Press, Aarhus.

Blitz, D., 1992, *Emergent Evolution*, Kluwer, Dordrecht.

Campbell, D. T., 1974, Downward causation in hierarchically organised biological systems, in: *Studies in the Philosophy of Biology*, F. J. Ayala and T. Dobzhansky, eds., Macmillan, London, pp. 179-186.

Campbell, D. T., 1990, Levels of organization, downward causation, and the selection-theory approach to evolutionary epistemology, in: *Theories of the Evolution of Knowing*, G. Greenberg and E. Tobach, eds., Erlbaum, pp. 1-17.

Gnoli, C., and Poli, R., (submitted), Levels of reality and levels of representation.

Heller, B., Herre, H. and Poli, R., (submitted), Formal ontology of levels of reality.

Luhmann, N., 1995, *Social Systems*, Stanford University Press, Stanford.

Pattee, H. H., 1973, *Hierarchy Theory*, Braziller, New York.

Poli, R., 1996, Ontology for knowledge organization, in: *Knowledge Organization and Change*, R. Green, ed., Indeks, Frankfurt, pp. 313-319.

Poli, R., 1998, Levels, *Axiomathes* **9**(1-2):197-211.

Poli, R., 2001a, *ALWIS. Ontology for Knowledge Engineers*, PhD Thesis, Utrecht.

Poli, R. 2001b, The basic problem of the theory of levels of reality, *Axiomathes* **12**(3-4):261-283.

Rosen, R., 1978, *Fundamentals of Measurement and Representation of Natural Systems*, Elsevier, Amstersam.

Von Foester, H., 1984, *Observing Systems*, 2nd ed., Intersystems Publications, Seaside, CA.

Weissmann, D., 2000, *A Social Ontology*, Yale University Press, New Haven and London.

GENERAL SYSTEM THEORY, LIKE-QUANTUM SEMANTICS AND FUZZY SETS

Ignazio Licata
Istituto di Cibernetica Non-Lineare per i Sistemi Complessi
Via Favorita, 9 - 91025 Marsala (TP), Italy
licata@neuroscienze.net

Abstract: It is outlined the possibility to extend the quantum formalism in relation to the requirements of the general systems theory. It can be done by using a quantum semantics arising from the deep logical structure of quantum theory. It is so possible taking into account the logical openness relationship between observer and system. We are going to show how considering the truth-values of quantum propositions within the context of the fuzzy sets is here more useful for systemics. In conclusion we propose an example of formal quantum coherence.

Key words: quantum theory; fuzzy sets; semantics; logical openness.

1. THE ROLE OF SYNTACTICS AND SEMANTICS IN GENERAL SYSTEM THEORY

The omologic element breaks specializations up, forces taking into account different things at the same time, stirs up the interdependent game of the separated sub-totalities, hints at a broader totality whose laws are not the ones of its components.
In other words, the omologic method is an anti-separatist and reconstructive one, which thing makes it unpleasant to specialists.
(F. Rossi-Landi, 1985)

The systemic-cybernetic approach (Wiener, 1961; von Bertalannfy, 1968; Klir, 1991) requires a careful evaluation of epistemology as the critical praxis internal to the building up of the scientific discourse. That is why the

usual referring to a "connective tissue" shared in common by different subjects could be misleading. As a matter of fact every scientific theory is the outcome of a complex conceptual construction aimed to the problem peculiar features, so what we are interested in is not a framework shaping an abstract super-scheme made by the "filtering" of the particular sciences, but a research focusing on the global and foundational characteristics of scientific activity in a trans-disciplinary perspective. According to such view, we can understand the General System Theory (GST) by the analogy to metalogic. It deals with the possibilities and boundaries of various formal systems to a more higher degree than any specific structure.

A scientific theory presupposes a certain set of relations between observer and system, so GST has the purpose to investigate the possibility of describing the multeity of system-observer relationships. The GST main goal is delineating a formal epistemology to study the scientific knowledge formation, a science able to speak about science. Succeeding to outline such panorama will make possible analysing those inter-disciplinary processes which are more and more important in studying complex systems and they will be guaranteed the "transportability" conditions of a modellistic set from a field to another one. For instance, during a theory developing, syntax gets more and more structured by putting univocal constraints on semantics according to the operative requirements of the problem. Sometimes it can be useful generalising a syntactic tool in a new semantic domain so as to formulate new problems. Such work, a typically trans-disciplinary one, can only be done by the tools of a GST able to discuss new relations between syntactics (formal model) and semantics (model usage). It is here useful to consider again the *omologic perspective*, which not only identifies analogies and isomorphisms in pre-defined structures, but aims to find out a structural and dynamical relation among theories to an higher level of analysis, so providing new use possibilities (Rossi-Landi, 1985). Which thing is particularly useful in studying complex systems, where the very essence of the problem itself makes a dynamic use of models necessary to describe the emergent features of the system (Minati and Brahms, 2002; Collen, 2002).

We want here to briefly discuss such GST acceptation, and then showing the possibility of modifying the semantics of Quantum Mechanics (QM) so to get a conceptual tool fit for the systemic requirements.

2. OBSERVER AS EMERGENCE SURVEYOR AND SEMANTIC AMBIGUITY SOLVER

What we look at is not Nature in itself,
but Nature unveiling to our questioning methods.
(W. Heisenberg, 1958)

A very important and interesting question in system theory can be stated as follows: given a set of measurement systems M and of theories T related to a system S, is it always possible to order them, such that $T_{i-1} \prec T_i$, where the partial order symbol $\prec$ is used to denote the relationship "physically weaker than"? We shall point out that, in this case, the ith theory of the chain contains more information than the preceding ones. This consequently leads to a second key question: can an unique final theory T_f describe exhaustively each and every aspect of system S ? From the informational and metrical side, this is equivalent to state that all of the information contained in a system S can be extracted, by means of adequate measurement processes.

The fundamental proposition for reductionism is, in fact, the idea that such a theory chain will be sufficient to give a *coherent* and *complete* description for a system S. Reductionism, in the light of our definitions, coincides therefore with the highest degree of semantic space "compression"; each object $D \in T_i$ in S has a definition in a theory T_i belonging to the theory chain, and the latter is - on its turn - related to the fundamental explanatory level of the "final" theory T_f. This implies that each aspect in a system S is unambiguously determined by the syntax described in T_f. Each system S can be described at a fundamental level, but also with many phenomenological descriptions, each of these descriptions can be considered an approximation of the "final" theory.

Anyway, most of the "interesting" systems we deal with cannot be included in this chained-theory syntax compatibility program: we have to consider this important aspect for a correct epistemic definition of systems "complexity". Let us illustrate this point with a simple reasoning, based upon the concepts of logical openness and intrinsic emergence (Minati, Pessa, Penna, 1998; Licata, 2003b).

Each measurement operation can be theoretically coded on a Turing machine. If a coherent and complete fundamental description T_f exists, then there will also exist a finite set - or, at most, countably infinite - of measurement operations M which can extract each and every single information that describes the system S. We shall call such a measurement set Turing-observer. We can easily imagine Turing-observer as a robot that executes a series of measurements on a system. The robot is guided by a program built upon rules belonging to the theory T. It can be proved, though,

that this is only possible for logically closed systems, or at most for systems with a very low degree of logical openness. When dealing with highly logically open systems, no recursive formal criterion exists that can be as selective as requested (i.e., automatically choose which information is relevant to describe and characterize the system, and which one is not), simply because it is not possible to isolate the system from the environment. This implies that the Turing-observer hypothesis does not hold for fundamental reasons, strongly related to Zermelo-Fraenkel's choice axiom and to classical Gödel's decision problems. In other words, our robot executes the measurements always following the same syntactics, whereas the scenario showing intrinsic emergence is semantically modified. *So it is impossible thinking to codify any possible measurement in a logically open system*!

The observer therefore plays a key rule, unavoidable as a *semantic ambiguity solver*: only the observer can and will single out intrinsic-observational emergence properties (Bass and Emmeche, 1997; Cariani, 1991), and subsequently plan adequate measurement processes to describe what – as a matter of fact- have turned in new systems. System complexity is structurally bound to logical openness and is, at the same time, both an expression of highly organized system behaviours (long-range correlations, hierarchical structure, and so on) and an observer's request for new explanatory models.

So, a GST has to allow - in the very same theoretical context – to deal with the observer as an emergence surveyor in a logical open system. In particular, it is clear that the observer itself is a logical open system.

Moreover, it has to be pointed out that the co-existence of many description levels – compatible but not each other deductible – leads to intrinsic uncertainty situations, linked to the different frameworks by which a system property can be defined.

3. LIKE-QUANTUM SEMANTICS

I'm not happy with all the analyses that go with just the classical theory, because nature isn't classical, damm it, and if you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy. Thank you. (R. P. Feyman, 1981)

When we modify and/or amplify a theory so as to being able to speak about different systems from the ones they were fitted for, it could be better to look at the theory deep structural features so as to get an abstract

perspective able to fulfil the omologic approach requirements, aiming to point out a non-banal conceptual convergence.

As everybody knows, the logic of classical physics is a dichotomic language (*tertium non datur*), relatively orthocomplemented and able to fulfil the weak distributivity relations by the logical connectives AND/OR. Such features are the core of the Boolean commutative elements of this logic because disjunctions and conjunctions are symmetrical and associative operations. We shall here dwell on the systemic consequences of these properties. A system *S* can get or not to get a given property *P*. Once we fix the P truth-value it is possible to keep on our research over a new preposition P subordinated to the previous one's truth-value. Going ahead, we add a new piece of information to our knowledge about the system. So the relative orthocomplementation axiom grants that we keep on following a successions of steps, each one making our uncertainty about the system to diminish or, in case of a finite amount of steps, to let us defining the state of the system by determining all its properties. Each system's property can be described by a countable infinity of atomic propositions. So, such axiom plays the role of a *describable axiom for classical systems*.

The unconstrained use of such kind of axiom tends to hide the conceptual problems spreading up from the fact that every description implies a context, as we have seen in the case of Turing-observer analysis, and it seems to imply that systemic properties are independent of the observer, it surely is a non-valid statement when we deal with open logical systems. In particular, the Boolean features point out that it is always possible carrying out exhaustively a synchronic description of the properties of a systems. In other words, every question about the system is not depending on the order we ask it and it is liable to a fixed answer we will indicate as 0-false / 1-true. It can be suddenly noticed that the emergent features otherwise get a diachronic nature and can easily make such characteristics not taken for granted. By using Venn diagrams it is possible providing a representation of the complete descriptiveness of a system ruled by classical logics. If the system's state is represented by a point and a property of its by a set of points, then it is always possible a complete "blanketing" of the universal set I, which means the always universally true proposition (see fig. 1).

The quantum logics shows deep differences which could be extremely useful for our goals (Birkhoff and von Neumann, 1936; Piron, 1964). At the beginning it was born to clarify some QM's counter-intuitive sides, later it has developed as an autonomous field greatly independent from the matters

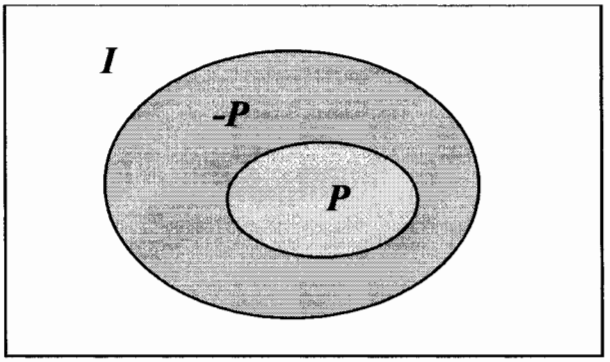


Figure 1. Complete blanketing of a classical Boolean system.

which gave birth to it. We will abridge here the formal references to an essential survey, focusing on some points of general interest in systemics.

The quantum language is a non-Boolean orthomodular structure, which is to say it is relatively orthocomplemented but *non-commutative*, for the crack down of the distributivity axiom. Such thing comes naturally from the *Heisenberg Indetermination Principle* and binds the truth-value of an assertion to the context and the order by which it has been investigated (Griffiths, 1995). A well-known example is the one of a particle's spin measurement along a given direction. In this case we deal with semantically well defined possibilities and yet intrinsically uncertain. Let put Ψ_x the spin measurement along the direction x. For the indetermination principle the value Ψ_y will be totally uncertain, yet the proposition $\Psi_y = 0 \vee \Psi_y = 1$ is necessarily true. In general, if P is a proposition, ($-P$) its negation and Q the property which does not commute with P, then we will get a situation that can be represented by a "patchy" blanketing of the set I (see fig.2).

Such configuration finds its essential meaning just in its relation with the observer. So we can state that when a situation can be described by a quantum logics, a system is never completely defined a priori. The measurement process by which the observer's action takes place is a choice fixing some system's characteristics and letting other ones undefined. It happens just for *the nature itself of the observer-system inter-relationship*. Each observation act gives birth to new descriptive possibilities. The proposition Q – in the above example – describes properties that cannot be defined by any implicational chain of propositions P. Since the intrinsic emergence cannot be regarded as a system property independent of the observer action - as in naïve classical emergentism -, Q can be formally considered the expression of an emergent property. *Now we are strongly tempted to define as emergent the undefined proposition of quantum-like anti-commutative language*. In particular, it can be showed that a non-

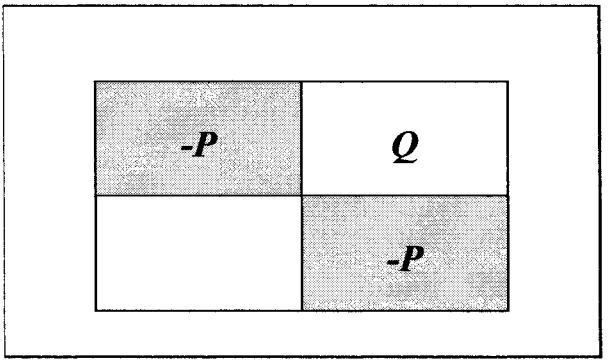


Figure 2. Patchy blanketing of non-Boolean quantum system.

Boolean and irreducible orthomodular language arises infinite propositions. It means that for each couple of propositions P_1 and P_2 such that none of them imply the other, there exists infinite propositions Q which imply $P_1 \vee P_2$ without necessarily implying the two of them separately: *tertium datur*. In a sense, the disjunction of the two propositions gets more information than their mere set-sum, that is the entirely opposite of what happens in the Boolean case. It is now easy to comprehend the deep relation binding the anti-commutativity, indetermination principles and system's holistic global structure. A system describable by a Boolean structure can be completely "solved" by analysing the sub-systems defined by a fit decomposition process (Heylighen, 1990; Abram, 2002). On the contrary, in the anti-commutative case studying any sub-system modifies the entire system in an irreversible and structural way and produces uncertainty correlated to the gained information, which think makes absolutely natural extending the indetermination principles to a big deal of spheres of strong interest for systemics (Volkenshtein, 1988).

A particularly key-matter is how to conceptually managing the infinite cardinality of emergent propositions in a like-quantum semantics. As everybody knows traditional QM refers to the frequentistic probability worked out within the Copenhagen Interpretation (CIQM). It is essentially a *sub specie probabilitatis* Boolean logics extension. The values between [0,1] - i.e. between the completely and always true proposition I and the always false one O – are meant as expectation values, or the probabilities associated to any measurable property. Without dwelling on the complex – and as for many questions still open – debate on QM interpretation, we can here ask if the probabilistic acception of truth-values is the fittest for system theory. As it usually happens when we deal with trans-disciplinary feels, it will bring us to add a new, and of remarkable interest for the "ordinary" QM too, step to our search.

4. A FUZZY INTERPRETATION OF QUANTUM LANGUAGES

A slight variation in the founding axioms of a theory
can give way to huge changings on the frontier.
(S. Gudder, 1988)

The study of the structural and logical facets of quantum semantics does not provide any necessary indications about the most suitable algebraic space to implement its own ideas. One of the thing which made a big merit of such researches has been to put under discussion the key role of Hilbert space. In our approach we have kept the QM "internal" problems and its extension to systemic questions well separated. Anyway, the last ones suggest an interpretative possibility bounded to fuzzy logic, which thing can considerably affect the traditional QM too. The *fuzzy set theory* is , in its essence, a formal tool created to deal with information characterized with vagueness and indeterminacy. The by-now classical paper of Lotfi Zadeh (Zadeh, 1965) brings to a conclusion an old tradition of logics, which counts Charles S. Peirce, Jan C. Smuts, Bertrand Russell, Max Black and Jan Lukasiewicz among its forerunners. At the core of the fuzzy theory lies the idea that an element can belong to a set to a variable degree of membership; the same goes for a proposition and its variable relation to the true and false logical constants. We underline here two aspects of particular interest for our aims. The fuzziness' definition concerns single elements and properties, but not a statistical ensemble, so it has to be considered a completely different concept from the probability one, it should – by now – be widely clarified (Mamdani, 1977; Kosko, 1990). A further essential – even maybe less evident – point is that fuzzy theory calls up a non-algorithmic "oracle", an observator (i.e. a logical open system and a semantic ambiguity solver) to make a choice as for the membership degree. In fact, the most part of the theory in its structure is free-model; no equation and no numerical value create constraints to the quantitative evaluation, being the last one the model builder's task. There consequently exists a deep bound between systemics and fuzziness successfully expressed by the *Zadeh's incompatibility principle* (Zadeh, 1972) which satisfies our requirement for a generalized indeterminacy principle. It states that by increasing the system complexity (i.e. its logical openness degree), it will decrease our ability to make exact statements and proved predictions about its behaviour. There already exists many examples of crossing between fuzzy theory and QM (Dalla Chiara and Giuntini, 1995; Cattaneo, Dalla Chiara and Giuntini 1993). We want here to delineate the utility of fuzzy polyvalence for systemic interpretation of quantum semantics.

Let us consider a complex system, such as a social group, a mind and a biological organism. Each of these cases show typical emergent features owed both to the interaction among its components and the inter-relations with the environment. An act of the observer will fix some properties and will let some others undetermined according to a non-Boolean logic. The recording of such properties will depend on the succession of the measurement acts and their very nature. The kind of complexity into play, on the other hand, prevents us by stating what the system state is so as to associate to the measurement of a property an expectation probabilistic value. In fact, just the above-mentioned examples are related to macroscopic systems for which the probabilistic interpretation of QM is patently not valid. Moreover, the traditional application of the probability concept implies the notion of "possible cases", and so it also implies a pre-defined knowledge of systems' properties. However, the non-commutative logical structure here outlined does not provide any cogent indication on probability usage.

Therefore, it would be proper to look at a fuzzy approach so to describe the measurement acts. We can state that *given a generic system endowed with high logical openness and an indefinite set of properties able of describing it, each of them will belong to the system in a variable degree.* Such viewpoint expressing the famous theorem of fuzzy "subsetness" – also known as "the whole into the part" principle – could seem to be too strong , indeed it is nothing else than the most natural expression of the actual scientific praxis facing intrinsic emergent systems. At the beginning, we have at our disposal indefinite information progressively structuring thanks to the feedback between models and measurements. It can be shown that any logically open model of degree n – where n is an integer – will let a wide range of properties and propositions indeterminate (the Q_s in fig. 2).The above-mentioned model is a "static" approximation of a process showing aspects of variable closeness and openness. The latter ones varies in time, intensity, different levels and context. It is remarkable pointing out how such systems are "flexible" and context-sensitive, change the rules and make use of "contradictions". This point has to be stressed to understand the link between fuzzy logic and quantum languages. By increasing the logical openness and the unsharp properties of a system, it will be less and less fit to be described by a Boolean logic. It brings as a consequence that for a complex system the intersection between a set (properties, propositions) and its complement is not equal to the empty set, but it includes they both in a fuzzy sense. So we get a polyvalent semantic situation which is well fitted for being described by a quantum language. As for our systemic goal it is the probabilistic interpretation to be useless, so we are going to build a fuzzy acception of the semantics of the formalism. In our case, *given a system S*

and a property Q, let be Ψ *a function which associates Q to S, the expression* $\Psi_S(Q) \in [0,1]$ *has not to be meant as a probability value, but as a degree of membership.* Such union between the non-commutative sides of quantum languages and fuzzy polyvalence appears to be the most suitable and fecund for systemics.

Let us consider the traditional expression of quantum coherence (the property expressing the QM global and non-local characteristics, i.e. superposition principle, uncertainty, interference of probabilities), $\Psi_S = a_1 \Psi_1 + a_2 \Psi_2$. In the fuzzy interpretation, it means that the properties Ψ_1 and Ψ_2 belong t Ψ with degrees of membership a_1 and a_2 respectively. In other words, *for complex systems the Schrödinger's cat can be simultaneously both alive and dead!* Indeed the recent experiments with SQUIDs and the other ones investigating the so-called macroscopic quantum states suggest a form of macro-realism quite close to our fuzzy acception (Leggett, 1980; Chiatti, Cini and Serva, 1995). It can provide *in nuce* an hint which could show up to be interesting for the QM old-questioned interpretative problems.

In general, let x be a position coordinate of a quantum object and Ψ its wave function, $|\Psi(x)|^2 dV$ is usually meant as the probability of finding the particle in a region dV of space. On the contrary, in the fuzzy interpretation we will be compelled to look at the Ψ square modulus as the degree of membership of the particle to the region dV of space. How unusual it may seem, such idea has not to be regarded thoughtlessly at. As a matter of fact, in Quantum Field Theory and in other more advanced quantum scenarios, a particle is not only a localized object in the space, but rather an event emerging from the non-local networks elementary quantum transition (Licata, 2003a). Thus, the measurement is a "defuzzification" process which, according to the stated, reduces the system ambiguity by limiting the semantic space and by defining a fixed information quantity.

If we agree with such interpretation we will easily and immediately realize that we will able to observate quantum coherence behaviours in non-quantum and quite far from the range of Plank's h constant situations. We reconsider here a situation owed to Yuri Orlov (Orlov, 1997).

Let us consider a Riemann's sphere built on an Argand's plane, where each vector represents a complex amplitude (Dirac, 1947) and let assume that each point on the sphere fixes a single interpretation of a given situation, i.e. the assigning of a coherent set of truth-values to a given proposition. Alternatively, we can consider the choosing of a vector $\boldsymbol{v}$ from the centre O to a point on the sphere as a logical definition of a world. If we choose a different direction, associated to a different vector $\boldsymbol{w}$, we can now set the problem about the meaning of the amplitude between the logical descriptions of the two worlds. It is known that such amplitude is expressed by

$$\frac{1}{2}(1+\cos\vartheta),$$

where ϑ is the angle between the two interpretations. The amplitude corresponds to a superposition of worlds, so producing the typical interference patterns which in vectorial terms are related to $|w|/|v|$. In this case, the traditional use of probability is not necessary because our knowledge of one of the two world with probability equal to $p = 1$ (certainty), say nothing us about the other one probability. An interpretation is not a quantum object in the proper sense, and yet we are forced to formally introduce a wave-function and interference terms whose role is very obscure a one. The fuzzy approach, instead, clarifies the quantum semantics of this situation by interpreting interference as a measurement where the properties of the world $v\,|\,\Psi_v\rangle + w\,|\,\Psi_w\rangle$ are owed to the global and indissoluble (non-local) contribution of the v and w overlapping.

In conclusion, the generalized using of quantum semantics associated to new interpretative possibilities gives to systemics a very powerful tool to describe the observator-environment relation and to convey the several, partial attempts - till now undertaken - of applying the quantum formalism to the study of complex systems into a comprehensive conceptual root.

ACKNOWLEDGEMENTS

A special thank to Prof. G. Minati for his kindness and his supporting during this paper drafting. I owe a lot to the useful discussing on structural Quantum Mechanics and logics with my good friends Prof. Renato Nobili and Daniele Lanzillo. Dedicated to M. V.

REFERENCES

Abram, M. R., 2002, Decomposition of systems, in: *Emergence in Complex, Cognitive, Social and Biological Systems*, G. Minati and E.Pessa, eds., Kluwer Academic, New York.

Baas, N. A., and Emmeche, C., 1997, On emergence and explanation, *SFI Working Paper*, Santa Fé Inst., 97-02-008.

Birkhoff, G., and von Neumann, J., 1936, The logic of quantum mechanics, *Annals of Math.* **37**.

Cariani, P., 1991, Adaptivity and emergence in organism and devices, *World Futures* **32**(111).

Cattaneo, G., Dalla Chiara, M. L., and Giuntini, R., 1993, Fuzzy-intuitionistic quantum logics, *Studia Logica* **52**.

Chiatti, L., Cini M., and Serva, M., 1995, Is macroscopic quantum coherence incompatible with macroscopic realism?, *Nuovo Cim.* **110B**(5-6).

Collen, A., 2002, Disciplinarity in the pursuit of knowledge, in: *Emergence in Complex, Cognitive, Social and Biological Systems*, G. Minati and E. Pessa, eds., Kluwer Academic, New York, 2002.

Dalla Chiara, M. L., and Giuntini, R., 1995, The logic of orthoalgebras, *Studia Logica*, **55**.

Dirac, P. A. M., 1947, *The Principles of Quantum Mechanics*, 3rd ed., Oxford University Press, Oxford.

Feynman, R. P., 1982, Simulating physics with computers, *Int. J. of Theor. Phys.* **21**(6/7).

Griffiths, R. B., 1995, Consistent quantum reasoning, in: arXiv:quant-ph/9505009 v1.

Gudder, S. P.,1988, *Quantum Probability*, Academic Press, New York.

Heisenberg, W., 1958, *Physics and Philosophy: The Revolution in Modern Science*, Harper and Row, New York, (Reprint edition 1999, Prometheus Books).

Heylighen, F., 1990, Classical and non-classical representations in physics: quantum mechanics, *Cybernetics and Systems* **21**.

Klir, J. G., ed., 1991, *Facets of Systems Science*, Plenum Press, New York.

Kosko, B., 1990, Fuzziness vs. probability, *Int. J. of General Systems* **17**(2).

Legget, A. J., 1980, Macroscopic quantum systems and the quantum theory of measurement, *Suppl. Prog. Theor. Phys.* **69**(80).

Licata, I., 2003a, *Osservando la Sfinge. La Realtà Virtuale della Fisica Quantistica*, Di Renzo, Roma.

Licata, I., 2003b, Mente & computazione, *Sistema Naturae, Annali di Biologia Teorica*, **5**.

Mamdani, E. H., 1977, Application of fuzzy logic to approximate reasoning using linguistic synthesis, *IEEE Trans. on Computers* **C26**.

Minati, G., and Brahms, S., 2002, The dynamic usage of models (DYSAM), in: *Emergence in Complex, Cognitive, Social and Biological Systems*, G. Minati and E. Pessa, eds., Kluwer Academic, New York.

Minati, G., Pessa, E., and Penna, M. P., 1998, Thermodynamical and logical openness, *Systems Research and Behavioral Science* **15**(3).

Orlov, Y. F., 1997, Quantum-type Coherence as a Combination of Symmetry and Semantics, in: arXiv:quant-ph/9705049 v1.

Piron, C., 1964, Axiomatique quantique, *Helvetica Physica Acta* **37**.

Rossi-Landi, F, 1985, *Metodica filosofica e scienza dei segni*, Bompiani, Milano.

Volkenshtein, M. V., 1988, Complementarity, physics and biology, *Soviet Phys. Uspekhi*, **31**.

Von Bertalanffy, L., 1968, *General System Theory*, Braziller, New York.

Zadeh, L. A., 1965, Fuzzy sets, *Information and Control* **8**.

Zadeh, L. A., 1987, Fuzzy Sets and applications, in: *Selected Papers by L. A. Zadeh*, R. R. Yager, R. M. Tong, S. Ovchnikov, and H. T. Nguyen, eds., Wiley, New York.

Wiener, N., 1961, *Cybernetics: or Control and Communication in the Animal and the Machine*, MIT Press, Cambridge.

ABOUT THE POSSIBILITY OF A CARTESIAN THEORY UPON SYSTEMS, INFORMATION AND CONTROL

Paolo Rocchi
IBM, Via Shangai 53, 00144 Roma, Italy, paolorocchi@it.ibm.com

Abstract: A variety of studies such as operational research, control theory, information theory, calculate relevant sides of system operations. They although cover small and separed areas, and provide a feeble support to engineers who need an ample theoretical framework. This paper illustrates an axiomatic theory that attempts to cover and integrate three ample topics: systems, information and control. We comment the reasons which steered this study and the significance of some formal results that have been achieved.

Key words: general system theory; information theory; ideal models.

1. PROLEGOMENA OF THE PROJECT

A personal experience triggered this study in the seventies. Once I entered the computer sector I observed a high number of professionals who were excellent specialists in hardware and/or software but found it exacting to place their works within the overall situation. We are still facing an evident imbalance: producers offer amazing hardware and software items, whereas technicians are in short supply of comprehensive illustrations. The scientific community lends its significant support to production but meets difficulties keeping the step with the manufactured items on the intellectual plane. Non-stop generation of practical remedies interferes with theorists' meditations, which need time to filter telling novelties from the background noise. As a consequence, researchers define small views or otherwise put forward qualitative studies that cannot support technology. A few remarks on investigations upon systems, information and control may lighten this landscape.

System theories are moving along different directions and include so many themes that they coveys an air which borders on being an all-inclusive discipline. The main trend takes its origins in the generalization of circuits theory and has generated mathematical findings. The Markov chains, invented in the early twentieth century, became one of the most popular and powerful mathematical tools. Electric techniques, their bodies of results proved to be useful in the field of complex aggregates. During the years 1940-60 calculations progressed and so generated a growing awareness on the part of many individuals that a "system theory" was born (Zadeh, 1962). From then onwards works more extensively addressed abstract algebra and tried to provide a formal theory on systems. Among the most recent developments that warrant mention, there are the investigations on "chaotic" behaviors of systems and the fuzzy systems (Wang, 1980). A second trend in system theory progressed in parallel with the first one. It has oriented toward sectorial arguments namely in economics, in organization, etc. As an eminent example, we find the "input/output analysis" that typically calculates multi-sectored planning models (Leontief, 1966).

Shannon's theory emerges as the most popular mathematical framework in the informational sector (Shannon, 1993). It introduced the statistical approach to information and has infiltrated various disciplines. The author deliberately excludes from his investigation the question of the meaning of the message and has been repeatedly attacked for this omission (Ritche, 1986). Solomonoff, Kolmogorov and others (Chaitin, 1966) focused on the complexity of messages. They worked around the algorithmic theory of information and put forward two kinds of measures for information. It is evident how these approaches neglect the possibility of any intellectual / human form of information. They overlook countless studies who assume meaning as the essence of information. I remind (Nauta, 1970) among the most acute thinkers. Some semiotic investigations take origin in linguistics, others in logic Bar-Hillel and Carnap, others are influenced by cognitive sciences (Bateson, 1951).

During the Second World War automatic control emerged as an important and distinct field. Norman Wiener who termed cybernetics in which the concepts of feedback plays a key role, was especially influential in motivating works in the area of automated control. The post-war industrialization provided the stimulus for additional investigations, which progressively addressed specialist questions. The mathematical study of control has drawn several topics such as feedback, networks and signals and may be seen as a branch of the system theory (Casti, 1982). For example, Kalman provided the concepts of controllability and observability as a systemic property.

In conclusion, there are a lot of theories that cover specific areas and/or do not harmonize. Current justifications provide answers to limited questions and dramatically fail to respond to the most complex. Theories are effective in accessory problems and flop in those that are important.

Deficiencies in engineering pressed me toward the definition of a rigorous framework capable of integrating the Wiener triad: systems, information and control. I found a solid reference in Ludwig von Bertalanffy who openly pursued the possibility of an exhaustive discipline.

"Modern science is characterized by its ever-increasing specialization, necessitated by the enormous amount of data, the complexity of techniques and of theoretical structures within every field. Thus, science is split into innumerable disciplines continually generating new sub-disciplines... It is necessary to study not only parts and processes in isolation (as in classical analytic scientific methods), but to solve the decisive problems found in the organization and order unifying them, resulting from dynamic interaction of parts, and making the behavior of parts different when studied in isolation or within the whole. ... These considerations lead to the postulate of a new scientific discipline which we call general system theory. Its subject matter is formulation of principles that are valid for "systems" in general, whatever the nature of the component elements and the relations or "forces" between them." (Bertalanffy, 1968)

I have also taken my cue from philosophers of science who constantly encouraged progress toward a comprehensive reasoning and an exhaustive understanding of computer science. Since the fifties, their attention especially addressed the possibilities and limits of computers with respect to the human brain. The progress in Artificial Intelligence in the eighties vividly relaunched philosophical debates.

I firmly believed in the possibility of a Cartesian theory and devoted resources in this direction.

2. ESSENTIALS OF SYSTEMS

An axiom, formally accepted without proof, is to be the cornerstone for the systematic derivation of a structured body of knowledge. I wonder: How can the science of systems, information and control be summarized? Which is the essence of such an endless variety?

Systems are natural and artificial, automatic and manual. They cover the planet and are so tiny as to be included in a chip; they are managed by political criteria and by rigid algorithms. Systems and their parts present opposing properties not only to superficial sensations but also to traditional disciplines, and stress the discovery of their most relevant qualities. I have

found a guide in Bertalanffy who grasped the essence of this apparently unconfined world. He conceived a system as "elements in standing relationship" (Bertalanffy, 1968) to wit system stands for a configuration of parts connected and joined by a web of relationships. Basically, a system is how things are working together; namely a family of relationships among the members interacts as a whole. This interpretation hinted the following postulate:

The idea of relating, connecting and linking is primitive.

This axiom sums up the fundamental features of systems and constitutes the starting point for inferences. As first we draw two algebraic elements *specialized in relating and in being related.* The entity ε and the relationship μ lead to the formal definition of the system in compliance with Bertalanffy's mind

$$S = (\varepsilon\,;\,\mu) \tag{1}$$

Current literature already calculates this expression, however it raises some remarks which Klir summarizes as follows:

"The definition is weak because it is too general and, consequently, of little pragmatic value. It is strong because it encompasses all other, more specific definitions of systems. Due to its full generality, this common-sense definition qualifies for a criterion by which we can determine whether any given object is a system or not: an object is a system if and only if it can be described in the form that conforms to $S = (\varepsilon;\ \mu)$. Once we have the capability of distinguishing objects that are systems from those that are not, it is natural to define system science ..." (Klir, 1991)

These objections are relevant and experience substantiates their impact. Practitioners sometimes cannot determine a system and, even if they realize the whole, they find its analysis hard. Tangible obstacles emerge, for instance, in the information system design. Graphs should aid specialists, instead the use of edges and vertexes, which respectively picture relationships and entities, appear doubtful.

The reason is the following. Current algebra assumes ε and μ as axiomatic notions. Their properties are not explicitly elucidated and they seem rather generic to the engineers who have to translate different elements of a physical system into ε and μ. Conversely, the present theory takes on an axiom, which is more general and derives the definition of the algebraic elements from this. The meaning of the entity and the relationship is wholly evident. The relationship has the property to connect hence it denotes the active components of a physical system; the entity the passive ones. The

above mentioned difficulties do not arise on the logical plane neither in the practice. This kind of theoretical refinement opposes the current trends in algebra to the extent that it has to be commented.

After the fascinating birth of the set theory by Cantor, algebraic investigations provided material for determining abstract structures that are ever more complexes. Specialists, in order to study intricate and awkward systems, are moving towards structures of structures such as the categories that attracted the mathematicians' attention over the past decades (Borceux, 1994). Instead, my efforts addressed the opposite direction. I searched for the essence of algebraic elements and dissected their fundamental attitudes. The axiom summarizes the essence of systems and contemporary is the solid foundation of the theoretical building which otherwise could not be built.

3. PERFECT ELEMENTS

A correct thinker, who puts forward a theory, critically scrutinizes the relation of his formulae with the physical reality and has to seek the "ideal cases" in the world. These inquiries confirm the winning qualities of the logical model or otherwise uncover its possible weak spots. Ideal cases pose the most tricky questions and some disputes last long. To exemplify, man is naturally acquainted with the concept of motion, but the debate upon the "ideal motion" covered several centuries. Aristotle, under the influence of Plato, sustained this was circular. This wrong stance obstructed the development of mechanics until Galilei proved that motion, constant in module, in direction and versus, is ideal whereas the circular movement is accelerated. The perfect use of a formula can draw astonishing consequences that subvert solid beliefs. For instance, the orbit of planets is not a smooth and perfect movement, as the ancients had believed up to then. Instead it is determined by the balance of opposing forces and can collapse due to their inconstant equilibrium.

The ideal cases comply thoroughly with the mathematical models by definition. They also extend the knowledge in the field, because they exploit the inner contents of the theory. They are not mere refinements.

We go deep into formula (1) and discuss how it conforms to the physical reality. In particular the specific properties of ε and μ yield the following three possibilities in the world.

Case 1: A machine (= μ) processes the product (= ε) which randomly is defective. On the theoretical plain we hold that the relationship μ links the entity ε but the algebraic scheme (1) generically fits the physical reality due to the failures of ε.

Case 2: A machine receives objects which are not scheduled for production. The algebraic element μ and ε do not suit the physical items that are out of use. The theory is inappropriate for the real case.

Case 3: The machine selects the input, namely it accepts the correct product ε and rejects the wrong entities ε^*. This is the ideal case for the theoretical model (1) as μ connects regularly the input. The ideal entity ε takes an exclusive relation with the process μ, and cannot be confused with any other item ε^*. We conclude that the ideal ε differs from any entity ε^* with reference to μ, and expresses the basic quality by this inequality

$$\varepsilon \neq \varepsilon^* \tag{2}$$

Concluding the theoretical statement (1) is unsuited to Case 2, it may be applied in Case 1 and perfectly calculate the physical system if (2) is true.

Note that we usually accredit this feature to a distinguished class of items. In fact, a communication, a message, a sound inform only if they are distinct. On the contrary, a confused symbol is ineffective and we get no news. Several authors recognize the idea that information is something distinct, see for instance the ontological principles of (Burgin, 2002). The author who best explains this notion, is probably Bateson.

"The simplest but the most profound is the fact that it takes at least two somethings to create a difference. To produce news of difference, i.e., information, there must be two entities (real or imagined) such that the difference between them can be immanent in their mutual relationship; and the whole affair must be such that news of their difference can be represented as a difference inside some information-processing entity, such as a brain or, perhaps, a computer." (Bateson, 1979).

We conclude that, if a physical item is perfect, namely distinct with respect to μ, it is information. The inquiry about the ideal application of the theoretical model (1) yields (2) to wit it provides the definition of information.

I am capable of inferring most formulas of analog and binary technologies from (2) although they go beyond the scopes of this paper. The significance of (2) is even more intriguing on the cultural plane.

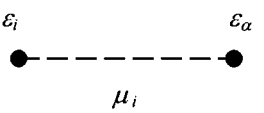
The inequality holds that anything in the world, if distinguishable through whatever criterion, is information. A river, a tree, the Sun are pieces of information. Any object in the world is potentially information and thanks to this quality any individual is capable of surviving. Definition (2) encompasses artifacts and spontaneous pieces of information, simple and complicated items. It expresses a property valid on the logical plane too.

Mental notions have to be distinct, otherwise they appear approximate, vague, fuzzy, uncertain, confused etc. (Marcus, 1998) and waste their capability of informing. This broad view overcomes technology and fits the most complete studies on information.

4. INTRICATE CONNECTIONS

Current literature shares the idea that primitive information is sensation, that is the passive capacity of perceiving an external object through the contact. Without any necessary exchange of matter, the perceiver takes on the form, or the color or any character of the object or even the object as the whole. The inequality (2) consists with the "sense-data theory" (which was treated by Russell, and recently by Moore, 1962) and matches with the thought of prominent empiricists such as Bacon, Locke and Hume. They claim that genuine information about the world has to be acquired, so that nothing can be thought without first being sensed. An uninterrupted production of information, namely a generation of pieces that verify the inequality, goes from the physical world to the inside of the brain and comes back through a different course. Before the discussion of this articulated process, we take an intermediate step.

Spontaneous information in the world (e.g. the Sun, a mountain, an event) works out defective, as it is a volatile image or vice versa unwieldy, it is imperceptible or intrusive. A natural item may present so many defects that man skirts these difficulties by substituting the original and artless ε_α with the artifact ε_i. The latter represents the former and we commonly call μ_i as meaning of information.



(3)

Authors are familiar with this phenomenon but investigate the relation (3) from a different and somewhat opposite perspective. The physical nature of ε_i is pivotal from the engineering stance and draws the keen interest of technicians. On the contrary, humanists focus on μi and the intricate mental world that has generated it. The entity ε_i plays an ancillary role within these studies and is a mere *information carrier* (Nauta, 1970). Terminology reflects opposed minds upon the unique scheme (3). These studies are incapable of distilling the contents of the simple statements (2) and (3), as we shall easily do. In fact:

a) Inequality (2) claims that an item of information is a material distinct item
b) Scheme (3) holds that an item of information stands for something else.

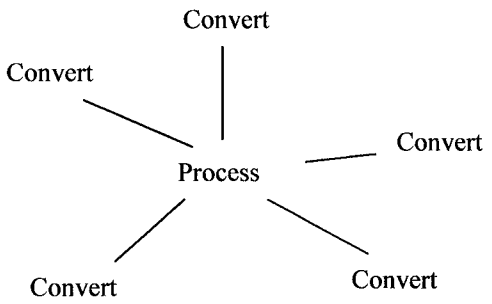
Consequently, operations on information can do nothing but modify the properties a) and b). They fall into the ensuing classes:

A) The operation converts the physical nature of the input. For instance, the eye changes visual information into nervous impulses. The printer transforms the electric information into ink. The keyboard converts mechanical information into electric bits.

B) The operation produces an item that represents something different from the input. Applied calculus provides an immediate example. Take this division that gets two values and brings about the speed, namely a novel model of the reality with respect to input

250 miles / 2 hour = 125 miles/hour
Space / Time = Speed (4)

Experience substantiates these theoretical issues. Computers and nervous systems handle items of information, which are physically different. Material co-ordination necessarily makes them homogeneous and the conversion units ring the informational processing which lies at the center.

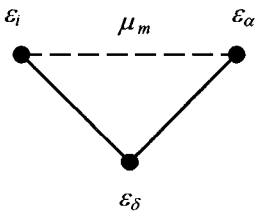


(5)

The computer peripherals execute physical transformations in support of the central unit. The five organs of sense and the receptors serve the brain in a similar way. The central unit and the brain produce pieces of information carrying different meanings with respect to the input. In short expressions (2) and (3) justify the structural features of computers and neural systems as well. The points A) and B) unify the biological and the mechanical manipulations of information.

Now we see how the results we have just achieved can give us an insight into the origin of meaning. In fact, the scheme (3) presumes that man

perceives ε_α in the world and creates the model ε_δ of ε_α in his mind before being capable of defining ε_i



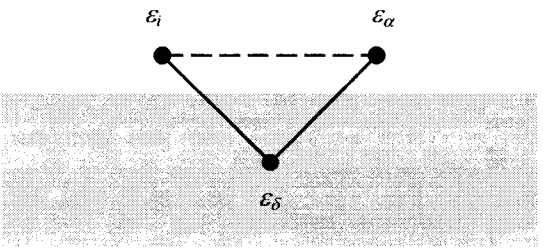
(6)

This semantic triangle summarizes the intricate role-play made by three items: natural information ε_α, artificial information ε_i and mental information ε_δ. A triangular relationship was firstly introduced by (Odgen and Richards, 1923) and overhauled by various authors through individual interpretations.

The triangle (6) holds that common information ε_i always stands for the mental thought ε_δ that, in turn, relates to the *reference* ε_α, according to the terminology of Gottlob Frege (Frege, 1892). I could detail the mental processes through the scheme (5), the full discussion although goes beyond the purposes of this article and I just trace a few lines.

The process $\varepsilon_\alpha - \varepsilon_\delta$ forms the thought, while the process $\varepsilon_i - \varepsilon_\alpha$ produces and interprets the language. Each edge of the triangle constitutes a system, which is potentially independent on the objects. They comply with the point B) hence even a microscopic change inside the observer's brain can drastically modify the conception ε_δ, whereas the world remains unchanged by this fact. Mental ideas and significance depend on the observer's position from which the projected universe is perceived; namely the physical reality ε_α is active and contemporarily passive as the psychological literature has amply illustrated.

Knowledge begins with the reference ε_α acquired through the senses, and the rational process $\varepsilon_\alpha\ \delta\ \varepsilon_\delta$, which is more or less extensive, achieves the results. From statement B) we argue that, if $\varepsilon_\alpha\ \delta\ \varepsilon_\delta$ is brief, the mental idea ε_δ is very similar to ε_α, to wit it is sense-driven, singular, sensitive. If the process is long, it brings about the sophisticate and abstract idea ε_δ which is far different from the input ε_α. For example, my dog Trick generates the immediate idea of "Trick" which is rather photographic. It also brings about the idea of "dog" which is elaborate, namely it is the outcome of a complex process by which the mind filters several dogs. This case offers an example of the reasoning $\varepsilon_\alpha\ \delta\ \varepsilon_\delta$ linear and poor, and the case of articulated mental creation.



(7)

I separate the interior (grey zone) from the external (white zone) of the human brain in the triangle in order to underline that cerebral processes cannot be accessed from the outside. The couple of areas proves that semantics is subjective and out of control. They corroborate the titanic efforts of men/women to communicate. People state the relationship $\varepsilon_i - \varepsilon_\alpha$, which simplifies the profound soul but, on the other hand, ensures the objective basis for comprehension. The unbiased side of semantics $\varepsilon_i - \varepsilon_\alpha$ is a rough simplification but can sustain measurements and tools due to its external physical position and to its linearity. Intricate and subjective significance may be hardly reduced to numbers. Technicians appreciate the objective edge which suggests effective solutions. For example, the double entries for digital encoding derive directly from $\varepsilon_i - \varepsilon_\alpha$. Conversely humanists, sociologists and psychologists find it superficial and amply reductive as they grasp the complete scenario. The zones clarify the conflicting concern and disparate approaches regarding information. Opposite interests arouse irreconcilable theories as far as now; whereas the present research tries to bridge these gaps and to reconcile the engineering and humanistic stances in between one exhaustive logical frame.

5. CONCLUSIVE REMARKS

This paper illustrates the layout of a theoretical investigation and comments a few results. They try to persuade the reader that the Cartesian approach to systems, information and control is a reasonable, and appears as an open way.

The present theory deduces a number of mathematical expressions, which go beyond the scopes of the present work. They reinforce our belief in the unifying framework.

The ruminations commented here have this secret: they have tackled all the questions, even those apparently obvious. I left no stone unturned. For example, writers introduce algebraic structures, instead I have scrutinized whether algebra yields telling and solid models. Authors calculate the system

performances assuming systems are capable of running, instead I have discussed *if systems are able to run.*

The primitive principle, introduced in these pages, aims at summarizing the essence of systems and offers a unique support for theoretical inferences.

Since the beginnings I saw the need to express the overall framework and have compiled a book for this purpose (Rocchi, 2000). The first ten chapters offer a theoretical account, the remaining formulate problems challenging professionals in the area of the software engineering. The book provides standard construction for technology that leads to the methods in the software application area.

BIBLIOGRAPHY

Bateson, G., 1979, *Mind and Nature: a Necessary Unity*, Bantam Books.

Bateson, G., and Ruesch, J., 1951, *Communication: The Social Matrix of Psychiatry*, W. W. Norton & Co., New York.

Bertalanffy von L., 1968, *General System Theory*, Brazziller, New York.

Borceux, F., 1994, *Handbook of Categorical Algebra*, Cambridge University Press.

Burgin M., 2002, The essence of information: paradoxes, contradictions and solutions, *Electronic Conf. on the Foundations of Information Science*, 6-10 May 2002; http://www.mdpi.net/fis2002/.

Casti J. L., 1982, Recent developments and future perspectives in nonlinear system theory, *SIAM Rev.* **24**(2).

Chaitin, G. J., 1977, Algorithmic information theory, *IBM Journal of Research and Development* **21**(4):350-359.

Frege, G., 1892, Über Sinn und bedeutung, *Zeitschrift für Philosophie und Philosophische Kritik* **100**:25-50, (Included in: *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black, ed., 1980, Blackwell, Oxford).

Klir, G., 1991, *Facets of Systems Science*, Plenum Press, New York.

Leontief, W. W., 1966, *Input-Output Economics*, Oxford Univ. Press, London.

Marcus, S., 1998, Imprecision between Variety and Uniformity, in: *Poznan Studies in the Philosophy of Sciences*, by J. J. Jadacki, **62**:59-62, Rodopi, Amsterdam.

Moore, G. E., 1962, *Some Main Problems of Philosophy*, Collier, New York.

Nauta, D. jr, 1970, *The Meaning of Information*, Mouton, Paris, Le Hague.

Odgen, C. K., and Richards, I. A., 1923, *The Meaning of Meaning*, Kegan Paul Trench Trubner, London.

Ritchie, L. D., 1986, Shannon and Weaver: unraveling the paradox of information, *Communication Research* **13**(2):278-298.

Rocchi, P., 2000, *Technology + Culture = Software*, IOS Press, Amsterdam.

Shannon, C. E., 1993, in: *Collected Papers*, N. J. A. Sloane and A. D. Wyner, ed., IEEE Computer Society Press, Los Alamos.

Wang, P. P, and Chang, S. K., eds., 1980, *Fuzzy Sets*, Plenum, New York.

Zadeh, L. A., 1962, From circuit theory to system theory, *Proc. IRE* **50**:56-63.