MASTERS AND SLAVES OF INFORMATION

Solomon Marcus

Simion Stoilow Institute of Mathematics Romanian Academy <u>solomarcus@gmail.com</u> Ten Statements to Be Considered for Information and Communication in Science

- At least in a first step, take information (*inf*) and communication (*com*) as primitive terms. Everybody understands them.
- Inf and com are self-referential operators; it is meaningful to refer to information about information and to communication about communication.
 From second order information we move step by step to *n*-th order information and similarly for communication.
- 3. Information of (*n*+1)-th order is of a higher complexity than information of *n*-th order and similarly for communication.

But just the increasing value of *n* in these processes characterizes our times.

4. There is no way to improve information and communication in all respects.

An improvement In some respects is obtained at the expense of deterioration in some other respects. So, no optimization for them is possible.

 Inf and com have both quantitative and qualitative aspects, but they cannot be captured simultaneously.
Each of these two aspects has its specific history, any attempt to bridge them failed or lead to derisory results.

 Information and communication cannot be strong In both their syntactic and semantic aspects.
But there is the third way, that of semantics build by means of syntactic-contextual procedures and just this way prevails in contemporary science and culture.

- 7. As human beings, our competence in coping with information and communication is limited to the macroscopic universe, but the today challenge coming from the universe of the infinitely small such as quantum information theory and its semantic impact on the macroscopic aspects of information cannot be ignored (see in this respect the works of Cristian Calude, one of them a joint contribution at this Congress).
- 8. All fields of knowledge and creativity, be they natural or social sciences, exact sciences and the humanities, computational or non-computational approaches, engineering or artistic, philosophical or theological are involved in the understanding of information and communication processes, while the possibility to monitor and to keep under control this huge information proves to be beyond the capacity of the today international instances.

- 9. The mistakes accumulated along the history in focusing the learning process on atomistic, descriptive, static aspects, at the expense of the global, dynamic, interactive aspects are now exploding, proving the increasing incapacity to face the requirements of the informationcommunication-computation paradigm, culminating with the Internet revolution.
- 10. The whole development of the today scientific enterprise, where the syntactic-operational aspects are more and more marginalizing the meaning and the sense makes plausible the hypothesis that the so-called double bind complex to which the Palo Alto school of psychotherapy is referring could be valid for the whole community of the scientific research.

Masters of information

Because the globalizations of all kinds and mainly the Internet revolution increased tremendously the available information and the possibilities of communication

But



Because we are less and less able to monitor, to aggregate, to bridge and to understand the increasing diversity and complexity of what human intelligence is producing

Trees at the Expense of the Forest

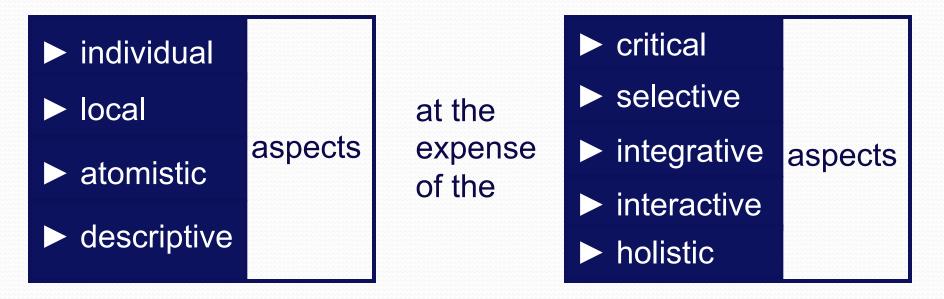
Do to the increasing number of trees, more and more trees remain ignored

and

we are less ad less able to move from their individual perception to the perception of the forest to which they belong

A Traditional Educational Mistake Is Now Exploding

A long, traditional trend of education, at all levels, to focus on



is now exploding

Are We Still Able to Monitor and to Survey the Information Related to Hot Research Topics? It happened that, according to my variety of scientific interests, I became aware of and sometimes directly involved in several directions of research related to the biological cell and coming from mathematics, computer science, linguistics, physics, chemistry, semiotics, philosophy, sociology and obviously biology, all starting with approximately the same claim:

Our aim is to understand the functioning of the biological cell

But in their next steps you hardly recognize that they have a common aim.

Each of them adopts a specific terminology, a specific jargon, and has specific bibliographic references with specific journals where the respective studies are published.

You expect that these different directions need to interact, but this expectation is not satisfied.

In most cases, cross references are very poor and it happens frequently that they ignore each other.

A Quarrel With Mathematics?

At a first glance we could believe that this lack of communication is due to the traditional quarrel between mathematics and non-mathematics, between social-human fields and exact sciences, mainly mathematics and computer science.

This fact could explain why some biologists and semioticians are reluctant to pay attention to the interaction of biology with mathematics and computer science, why bio-computing and bio-semiotics, for instance, are in a weak interaction. However, the reciprocal lack of communication can be observed between genomics, on the one hand, and DNA computing and membrane computing, on the other hand.

Between genomic linguistics and symmetry in molecular genetics, all of them impregnated with mathematics, to refer only to directions presented in the following.

Many other such situations can be observed.

It Is Scandalous! Let Us Describe Some of Them! Genetic Linguistics Already towards the middle of the past century emerged the interest for a linguistic reading of molecular biology, stimulated to a large extent by Roman Jakobson.

Gradually, this interest, having at its roots the believe in a deep parallelism between human language and the functioning of the biological cell lead to syntagms such that *molecular linguistics, protein linguistics* and culminated in the syntagm *genetic linguistics*, which is a chapter in some Encyclopaedias of Linguistics, for instance in the section *Comparative and Historical Linguistics* (by Ranko Matasovic) *Encyclopaedia of Life Support Systems* edited by UNESCO.

The Semiotics of the Biological Cell

I have mainly in view the work done by the Copenhagen-Tartu school, with scholars such as Jesper Hoffmeyer, Claus Emmeche and Kalevi Kull, on the one hand, and the cyber-semiotic trend followed by Søren Brier, on the other hand. **DNA Computing and Membrane Computing** DNA computing, called also biomolecular computing is a branch of computing which uses DNA hardware instead of the traditional siliconbased computer technologies.

But it is also a branch of molecular genetics.

However, we should try to look at such topics in a way different from the classificatory mentality ("branch of"), by supplementing it with an interactive, dynamic, transdisciplinary mentality. The experimental side was inaugurated by Leonard Adleman in 1994, while the theoretical side began a little earlier and it is described by Gheorghe Păun, Grzegorz Rozenberg and Arto Salomaa in their book *DNA computing; new computing paradigms* (Springer, 1998).

Membrane computing, aiming to argue in favour of the computational capacities of the biological cell, was proposed by Gheorghe Păun in the same year 1998.

Hoffmeyer Paying Attention to the Biological Membrane

The same year 1998 is the moment when Jesper Hoffmeyer called attention on the relevance of membrane in heredity, he is referring to

living systems as consisting of surfaces inside surfaces which turns inside exterior and outside interior

> Jesper Hoffmeyer, "Surfaces inside surfaces". *Cybernetics and Human Knowing* 5 (1), 1998, 33-42

Can we bridge Hofmeyer's surfaces inside surfaces with Păun's membranes?

A Wonderful Similarity With Levi-Strauss's Canonical Formula of Myth

Indeed, in Claus Emeche, Kalevi Kull, Frederik Stjernfelt's *Reading Hoffmeyer rethinking biology* (Tartu University Press, 2002), at page 17, reference is made to "the double twist of inside and outside, made possible by the membrane strictly governing the traffic between them [...]."

On the other hand, a collective book about Claude Levi–Strauss's canonical formula of myth, edited by Pierre Maranda (Toronto University Press, 2001) has the title *The double twist: From ethnography to morphodynamics*.

The double twist giving the architecture of both the biological cell and of ancient myths deserves attention.

The Emergence of a New Science: *Genomics*

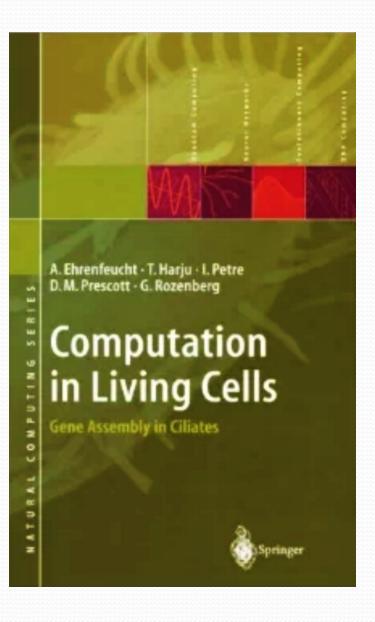
The success of the huge *Human Genome Project* towards the end of the 20th century gave birth to the new science called *Genomics*.

One of the journals reflecting this line of research is *Journal* of *Computational Biology - A Journal of Computational Molecular Cell Biology* aiming to produce, in continuation of the respective Project, a comprehensive genetic and physical map of the human genome.

I became specially interested in the mathematical aspects of the Human Genome Project, as they were revealed by Richard Karp.

Ciliates, the Simplest Living Organisms

They lead to a specific direction of research in cell biology, to which a collective work was devoted: Andrzej Ehrenfeucht, Teero Harju, Ion Petre, David M.Prescot, Grzegorz Rozenberg: Computation in living cell: Gene assembly in *ciliates*. Natural Computing, Springer, 2004.



Genetic Information Through the Glasses of Symmetry

This line of investigation is for many years very active in the journal *Symmetry: Culture and Science* of the International Association for Symmetry Studies.

Key words: genetics, golden section, symmetric matrices, Hadamard matrices, hydrogen bonds, molecular genetic systems and musical harmony, algebraic biology.

Main author: Sergey V. Petoukhov.

In 2011, Matthew He and Sergey Petoukhov published the book *Mathematics and Bioinformatcs* (Wiley, New York), where *knot theory, geometry, topology, dissipative structures, cognitive computing* and *fractals* play an important role in the study of molecular genetics.

A Relevant Preface

First lines of the Preface to He & Petoukhov's book:

- "Recent progress in the determination of genomic sequences has yielded many millions of gene sequences.
- But what do these sequences tell us, and what generalities and rules are governed by them?
- There is more to life than the genomic blueprint of each organism.
- Life functions within the natural laws that we know and those we do not know.
- It appears that we understand very little about genetic contexts required to «read» these sequences."

From Genes to Memes

A *gene* is a biological replicator that transmits hereditary characteristics.

A *meme* (Richard Dawkins, *The selfish gene*, 1976) is the cultural equivalent of a gene, "a bit of useful imitative information that passes from one person to another, but that can evolve in the process [...]."

Genes cannot provide children with all the information they will need to survive in a complex, interdependent, constantly shifting environment. Humans thus developed a learned meme system to replicate ad transmit useful imitative cultural information (Robert Sylwester, *From genes to memes,* Part I, 7 May 2003).

Memes do exist in our brain, they have a physical reality, as it is claimed by Robert Aunger: *The electric meme: A new theory of how we think,* Free Press, 2003.

I enjoy such analogies, they stimulate us, irrespective their veracity.

Bioengineering, Genetic Engineering I meet at various meetings concerning the biological cell many engineers and physicists, but in most cases their language is not mine.

Some Bold Metaphorical Slogans

A living being is a universal Turing machine (Stephen Wolfram, A new kind of science. Wolfram Media, October, 2001). DNA is essentially a digital software.

Human beings have much more DNA then viruses and bacteria.

We are universal Turing machines and we are surrounded by such machines.

But they differ in their program size complexity.

Life is a collection of universal Turing machines whose software evolves in complexity (Gregory Chaitin, Bulletin of the European Association of Theoretical Computer Science, 2002).

Such slogans are challenging us to try to bridge all the above approaches.

My Project: To Bridge This Diversity But in this respect I realized only some small steps, partly due to the almost total lack of communication between different approaches.

My choice was subjective: I refer just to those approaches that happened to arrive in my attention, according to the evolution of my scientific interests.

The Hot Summer of the Year 1971

In the summer of the year 1971, following the invitation received from David Hays, a leader in the field of Computational Linguistics, I organized within the framework of the *Linguistic Institute of America*, SUNY at Buffalo, a *Research Seminar at the crossroad of Molecular Genetics, Linguistics, Mathematics and Computer Science*.

As a product of this Seminar I published the first paper in the next list of my publications about the biological cell.

But it came too early to benefit of enough attention at that moment.

It had to wait about two decades.

My Publications Related to the Biological Cell

- a) "Linguistic structures and generative devices in molecular genetics". *Cahiers de Linguistique Theorique et Appliquees* 11, 1974, 2, 77-104;
- b) "Internal and external symmetries in genetic information". Symmetry: Culture and Science 12 (3/2), 2001, 395-400;
- c) "Membrane versus DNA". Fundamenta Informaticae 49, 1/3, 2002, 223-227;
- d) "An emergent triangle: semiotics, genomics, computation". *Proc. of the International Congress of the German Semiotic Society*, Kassel, 2002.
 CD–ROM, 2003;

- e) "Bridging P systems and genomics". In *Membrane Computing* (eds. G. Păun, G. Rozenberg, A. Salomaa, C. Zandron). LNCS 2597, Springer, Berlin, 2003, 371-378;
- f) "The duality of patterning in molecular genetics". In *Aspects of Molecular Computing* (eds. N.Jonoska, G. Păun, G. Rozenberg), LNCS 2950, Springer, Berlin, 2004, 318-321;

- g) "The semiotics of the infinitely small; molecular computing and quantum computing" in Semiotic Systems and Communication-Action-Interaction-Situation-Change. Proc. of the 6th National Congress of the Hellenic Semiotic Society (eds. K. Tsoukala et al.), Thessaloniki 2004, 15-22;
- h) "Semiotic perspectives in the study of cell", in *Proc.* of the Workshop on Computational Models for Cell Processes (eds. R.J. Back, I. Petre). TUCS General Publications no. 47, 2008, Turku, Finland, 2008, 63-68.

My Slogan: Life is DNA Software + Membrane Software

I suppose that the same anarchic scenario is valid for brain studies and for the field of information (inf) and communication (comm) we will have in attention in the following. To Be or Not to Be Self-Referential Various disciplines can be classified in two classes, according to their possible self-referential capacity.

It is meaningless to refer to "the physics of physics" or to "the chemistry of chemistry," unless we have in view a metaphorical utilisation.

By contrast, it is perfectly meaningful and very important to refer to "the philosophy of philosophy", "the literature about literature", "the inf about inf", "the comm about comm".

But just the iteration of these operators characterizes our time and so, instead to get inf abut something, we get inf about...inf.

Examples in exploring this self-referential operators are the French philosopher Edgar Morin and the German sociologist Niklas Luhmann.

Information: Quantitative and Qualitative

In contrast with *matter* and energy, located in some sciences of nature, *inf challenges the segmentation of knowledge in disciplines and the science / humanities opposition*.

It emerged concomitantly, in the second half of the 19th century, from thermodynamics (its quantitative version), associated with *entropy* (Clausius, Boltzmann) and from Darwinian biology (its qualitative version), associated with *form*, which is another self-referential operator, it is meaningful to refer to "the form of form." The Etymology Favours Information as Form *Inf* comes from the Latin *informatio*, while the verb *informare* means "to give a form."

The Greek *morph* became (by distortion?) the Latin *form.*

Plato, with his *Theory of Forms,* George Boole, with his *algebras* and C. S. Peirce, with his *signs* should be placed in this order of ideas.

So, inf as form is much older than inf as a measure of order.

9th Decade of the 19th Century: Peirce, Dedekind, Peano

The emergence of *recursiveness* as fundamental *form of thinking* is associated with Charles Sanders Peirce (1881), Richard Dedekind (1888) and Giuseppe Peano (1889) in connection with the axiomatization of natural numbers.

10th Decade of the 19th Century: Cantor, Hilbert, Weismann, Planck

- Georg Cantor's "diagonal argument" for the existence of uncountable sets (1891)
- The final form of its theory of cardinal and ordinal transfinite numbers (1895, 1897)
- And the eponymous paradox of the cardinal number of the set of all sets (1899)
- David Hilbert's new axiomatics of geometry (1899), a fundamental step challenging Euclid's way to understand the axiomatic-deductive thinking.

The evolutionary biologist August Weismann observes that problems related to heredity cannot be explained and understood exclusively in terms of matter and energy.

Something more is needed, he calls information.

With Max Planck, the quantum paradigm of discontinuity begins its great adventure.

First Decade of the 20th Century: Brouwer and Thue L.J. Brouwer's *intuitionism*, as a first form of effectiveness, against the use of Ernst Zermelo's *choice axiom*.

Semi-Thue combinatorial systems (due to the Norwegian Axel Thue), as a step towards what will be called later a *rewriting system*.

Second Decade of the 20th Century: Hilbert, D'Arcy Thompson, F. De Saussure, A. Einstein The emergence of form with Hilbert's *formal systems*, D'Arcy Thompson's *On growth and form*, Ferdinand de Saussure's *structural linguistics*, Albert Einstein's *relativity*.

Third Decade of the 20th Century: Kleene, Gödel, Bohr, Heisenberg, Nyquist, Hartley With S.C. Kleene and Kurt Gödel, the theory of recursive functions becomes a basic variant of the algorithmic thinking. With Niels Bohr's complementarity principle and Werner Heisenberg's uncertainty principle the quantum revolution challenges classical logic.

Harry Nyquist (1924) proposes to evaluate the speed V of transmission of a telegraphic message by the product between a constant k (depending of the number of modulations that can be transmitted in a unit of time and the logarithm of the number M of existing signs:

 $V = k \log M$

Ralph Hartley (1928) proposes a measure *m*(*s*) of the quantity of information transmitted by a signal *s*:

m(s) = log (1 / p(s))

where p(s) is the probability of appearance of s.

Fourth Decade of the 20th Century: Gödel, Turing, Shannon, Popper, Bertalanfy The Frege-Russell-Whitehead-Hilbert program is invalidated by Gödel's *incompleteness theorem*.

Alan M. Turing succeeds to extend the idea of computing from numbers to abstract symbols, realizing in this way the dream of Leibniz and the theoretical background for the future electronic computers.

Claude Shannon points out the similarity between electrical circuits and the Aristotle-Leibniz-Boole' s binary logic, bridging in this way two worlds, electrical engineering and human logic, which seemed to be far away each other. Karl Popper (*Logik der Forschung*, 1934, p.83) observes that a statement says about the empirical reality just what it puts on interdiction for the respective reality.

This negative way to look at information is convergent with that conceived later by Shannon.

The systemic thinking is emerging with the biologist Ludwig von Bertalanfy (1934).

Fifth Decade of the 20th Century: von Neumann-Morgenstern, McCulloch-Pitts, Wiener, Shannon, Hamming, Cherry Emergence of the theory of strategic games, with John von Neumann and Oskar Morgenstern

- Automata theory, starting with the simulation of the nervous system, with McCulloch and Pitts
- Cybernetics (Norbert Wiener)
- Computer science: the first programmed electronic computer, built by von Neumann and his team, a culminating moment after a long history including the abacus, Pascal's calculator, Babbage's engine, Hollerith, Alken, Eckert's punch card machines
- Mathematical information and communication theory (Claude Shannon); coding theory (R. Hamming)
- Engineering communication theory (Colin Cherry)

Tens of Information Fields, Increasing Difficulty to Bridge Them

In the previous section, seven information sciences were pointed out, all born in the fifth decade of the past century.

Each next decade brought in attention new information sciences, we will display in the following.

Their location, decade by decade, should be considered with approximation.

The spectacle of this succession of new and new information fields gives an idea of the richness and high complexity of the information paradigm.

Sixth Decade: Minsky, Carnap - Bar-Hillel, Watson - Crick, Brillouin, Chomsky

Marvin Minsky initiates in 1951 the field of *AI* (artificial intelligence) in a joint paper with Seymour Papert.

First attempt to capture semantic information by means of Shannon's approach belongs to Rudolf Carnap and Y. Bar-Hillel (1952).

The discovery in 1953 by James Watson and Francis Crick of the three-dimensional double helix structure of the *DNA* shows exactly in what sense molecular genetics is an information field. In exactly the same year Leon Brillouin publishes his *Science and Information Theory*, pointing out how thermodynamics is a special chapter of information theory.

Information is always obtained by production of entropy, so his proposal to call information *negentropy*.

In 1956 Noam Chomsky proposes his generative hierarchy of languages, transforming linguistics in a branch of cognitive psychology.

Concomitantly in Europe the *analytic* approach of *mathematical linguistics* is born.

Seventh Decade: Ginsburg, Rice, Floyd, Kolmogorov, Chaitin, Hintikka With Ginsburg, Rice an Floyd, Chomsky's formal generative grammars became the syntax of the computer programming languages, their common denominator being Hilbert's formal systems.

The semiotic triad syntax-semantics-pragmatics is thus transferred in computer science.

In contrast with Shannon's information theory, where the information parameters are related to the global, statistical aspect of a system, in A. N. Kolmogorov's algorithmic information theory (1965) and in Gregory Chaitin's approach (1966) the interest is focused on the local, individual aspect, with reference to the algorithmic-information complexity of a message, as it is given by the dimension of the shortest computer program permitting the identification of the respective message.

J. Hintikka (1968) tries to capture semantic information by extension of Shannon's approach.

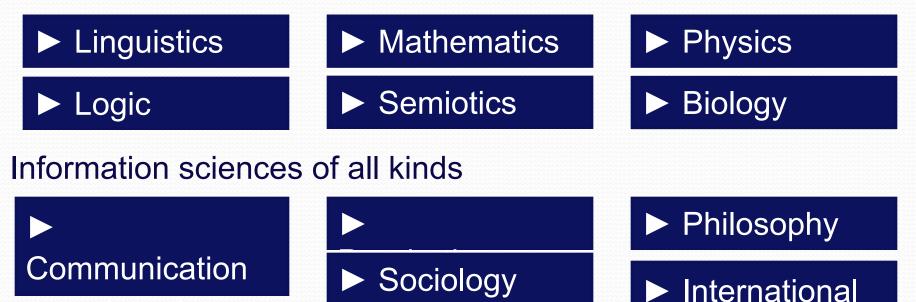
Seventh Decade: Zadeh, Peirce, Bakhtin, Lotman, Greimas

L. Zadeh starts (1965) his theory of fuzzy sets.

Charles Sanders Peirce's semiotics begins its explicit and systematic emergence, trying to impose the sign paradigm as a competitive one with respect to the information paradigm.

Other semiotic approaches, by Greimas, Bakhtin, Lotman, etc., give their contribution in this respect.

Trying Desperately to Bridge all Faces of Communication Processes This competition is visible and active in the way communication processes are represented.



All perspectives show their relevance, but in this respect the failure is visible.

Psychology

Poetics

relations

The seventies of the 20th century: Blum, Hartmanis, von Förster, Bateson, Thom, Mandelbrot, Maturana, Varela, Nauta

We have in view

- Complexity theory (Blum, Hartmanis)
- Second order cybernetics (Heinz von Föster, Gregory Bateson)
- Catastrophe theory (Rene Thom)
- ► The fractal geometry of nature (B. Mandelbrot)
- Chaos science in the line initiated in the 19th century by Henri Poincare
- Autopoietic systems by Umberto Maturana and Francesco Varela.

D. Nauta (1972) tries to bridge Shannon and Morris, information and sign.

The eighties: Bohm, Barwise-Perry, Pawlak

- David Bohm (1983: Wholeness and the implicate order (the hidden order of the quantum universe)
- I Barwise and J. Perry (1986) propose in Situations and attitudes a new face of information: situational semantics
- Z. Pawlak proposes a new approach to systems with incomplete information: rough sets

The nineties: Brier, Luhmann, Bennett-Shor, Stonier

- Søren Brier (1992) Information and consciousness
- Niklas Luhmann (1997) The society of society
- Tom Stonier (1997) proposes a general theory of information, starting from Wiener and Schrödinger
- Quantum information theory emerges as an extension of classical information theory to quantum world: Charles H. Bennett and Peter W. Shor (1998)

The emergence of the Internet in the last 25 years led to a considerable improvement of our access to information of all kinds, whose richness and variety made impossible to bring all of them under a common relatively simple and short definition.

As it happened with other fundamental ideas, such as *time* or *game*, no definition can be provided to cover all situations, so we can collect tens of definitions of information and tens of alternative quasi-equivalent terms.

In contrast with matter and energy, whose understanding was correlated to a relatively simple, small number of disciplines and contexts, the idea of information has been from the beginning related to a huge variety of situations, claiming for very difficult bridging processes, for which we are not at all prepared.

- Science and the humanities
- Nature and culture
- Macroscopic, quantum and cosmic
- Theoretical and applied
- Organic and inorganic
- Objective and subjective
- Natural and social
- Science and engineering
- Science and art

All these distinctions to be brought simultaneously in consideration.

So, in a world in which, against history, the bureaucracy of segmentation in disciplines and of science/ humanities opposition is still strong, the whole development of the information paradigm challenged the disciplinary borders and, to a large extent, ignored them.

But, in its dominant trend, the world of researchers was and it is still not prepared to cope adequately with this novelty.

So, we can understand why researchers in the field of biological cell or of information and communication, were not trained to face the today situation of explosion from all directions of the literature related to their problems of interest. Instead to challenge the complexity of the new situation, most of them reduced it to the dimensions of their disciplinary vision. There is a tension between information and sign, between information and meaning, between qualitative and quantitative information and this tension cannot be completely cancelled, but it can be attenuated.

At a first glance, each of them seems to reject the other, as it happened with other conflictual pairs such as

<position, momentum=""></position,>	<consistency, completeness=""></consistency,>
<rigor, meaning=""></rigor,>	<sensibility, clarity=""></sensibility,>

in well known specific contexts.

However, in logic, linguistics, mathematics, computer science the past century promoted *the meaning generated by syntactic means, by contextual behaviour*, where rigor is at home.

On the other hand, information and communication are often under the action of what G. Bateson called *the double bind constraint*.

One cannot improve at once both the emotional and the coding capacity of a communication process.

Some times, Grice's conversational principle does not work; you cannot be short and at the same time avoid ambiguity.

The school life, the social life in general often creates double bind situations.

To the extent to which we learn more and more, we increase our chance to keep under control information and communication; but to some extent, larger for some of us, smaller for others, we remain slaves of information and of communication, manipulated by them.

We are witnessing now the proliferation of publications related to information and communication.

Exactly like in the field of biological cell, the international monitor system is less and less capable to face this tsunami of information, to keep it under control.

A major obstacle in coping with information comes from the genuine limits of human semiosis, blocked as soon as we want to understand what happens beyond the macroscopic world. In Front of Us Fundamental Open Questions:

To bridge:

- Syntactic and semantic information
- Bio-semiotics and bio-computing
- Macro and quantum information (in this respect, see Calude et al. in this session)
- Algorithmic information theory and biological information
- Kolmogorov and Bateson
- DNA computing and Hoffmeyer Emmeche's code duality
- Hoffmeyer's surfaces inside surfaces and Paun's membrane computing.