Some Reflexions on Consciousness

Eric Schwarz
Autogenesis – Centre d'étude des systèmes autonomes
Université de Neuchâtel, CH-2000 Neuchâtel, Switzerland
eric.schwarz@unine.ch www.autogenesis.ch

Abstract
Consciousness, in particular its nature, is a very elusive notion. It can be felt when one tries to understand the correlation between the objective material aspect of the brain and the subjective mental aspect of consciousness (the explanatory gap). In this contribution we propose an interpretation based on a general metamodel we have developed for complex autonomous systems. We show that self-knowledge (i.e. consciousness) depends on the degree of self-reference of the system

Keywords: consciousness, explanatory gap, systems dynamics, ontology

1 Introduction and Context

1.1 Context: The Main Problems met in Trying to Understand Consciousness

1.1.1 What is Consciousness ?

Consciousness is one of the most elusive notions met in the scientific saga to understand the world around us and within us. Indeed, the usual configuration of the problems in science consists of two items: an observed object or process in space and time and a model or theory to explain what has been observed. Normally the challenge is to make an adequate model, but there is agreement about the thing that has to be understood: a piece of matter, the movement of an object, the organization of a network, the emergence of a new structure, etc.

With consciousness, even the "object" that has to be understood is not evident. Therefore here we have to make a model of something we do not know what it is !

Nevertheless, redactors of dictionaries and cognitive scientists have tried to propose definitions and functionalities that can be associated with consciousness. The Oxford Dictionary describes consciousness as " internal knowledge or conviction; knowledge as to which one has the testimony within oneself, especially of one's innocence, guilt deficiencies, etc. "

In the cognitive sciences, the following functionalities are often associated with the presence of consciousness:

- the ability to discriminate, categorize, and react to environmental stimuli
- the integration of information by a cognitive system
- the reportability of mental states
- the ability for a system to access to its own internal states
- the focus of attention

International Journal of Computing Anticipatory Systems, Volume 21, 2008
Edited by D. M. Dubois, CHAOS, Liège, Belgium, ISSN 1373-5411 ISBN 2-930396-08-3

- the deliberate control of behavior
- the difference between wakefulness and sleep

1.1.2 The Easy Problems and the Hard Problem

D. Chalmers (Chalmers, 1996) qualifies the problems connected to these aspects of the mental activities – the cognitive consciousness - as the "easy problems". The "hard problem", on the other hand – the phenomenal consciousness, is the problem of subjective experience. There is general agreement that experience arises from a material (biological) basis but numerous questions remain: why and how does it arise ? Why do some physical processes give rise to experience and others do not ? What is their nature ? What is the nature of the phenomenal conscious state ? What is the difference between quantifiable physical states and processes, and qualitative non measurable subjective experiences (qualia) ? All these questions are part of the "hard problem".

1.1.3 Present Debate: Cognitive Consciousness and Phenomenal Consciousness

In the framework of the cognitive sciences, it is now frequent to distinguish between cognitive – or access – consciousness and phenomenal – or subjective - consciousness.

- Cognitive consciousness is characterized by intentionality, which means that I am conscious of something, of an external reference like an object in front of me, or a somatic state or even a psychological drive or intention. It can be verbalized.
- Phenomenal consciousness is the phenomenon of subjective personal experience, which is private, non-transmissible and qualitative (qualia).

Cognitive consciousness can be organized according to different levels:

1. Primary, which corresponds to a representation of external things or somatic features.
2. Reflexive, which results from the introspection of our mental state; we are conscious of being conscious.
3. Self-conscious, which is consciousness of myself as a conscious subject

These distinctions are the object of many philosophical controversies.

Phenomenal consciousness can arise when we hear music, smell some odor, or see some image. It is usually independent of our will but may depend on the general context in which the experience happens: in some circumstances a given piece of music can trigger a deep pleasure or leave us untouched in other cases.

1.2 Some Models Proposed for Consciousness

1.2.1 Typology of the Main Approaches

Before summarizing some important models of consciousness proposed in the literature, let us list their main epistemological and ontological foundations, as discussed in the philosophy of mind.

1. Physicalist reductionism. Models or theories where the ultimate explicative power lies in the substance, in the detailed material structure of the system considered. Physical, biological or neurophysiological approaches are of this type.
2. Functionalism. Models or theories where the behavior of the system is not due explicitly to the material in the system but to the relational aspects inside the system: its organization, i.e. the way the parts are interconnected, the networks of causality between operating parts, the sequence of computation in a computer. Networks, flow charts, organigrams, list of instructions are tools used in the functionalist approach.
3. Holism. This epistemology does not consider only matter and relations but also – rather, above all - the existential aspect of a system as a whole. These existential dimensions include qualitative, subjective, or experienced states of being (qualia). We will go more into the details of this epistemology below in the presentation of the non-dualist framework proposed in this paper.

One can also devise a typology for the ontology of the models of consciousness - or of any model – depending on how they consider the fundamental substance of the things of the world.

1. Materialist monism. In this ontology, everything is matter and there is nothing else than matter.
2. Idealist monism: Here only non-material features are real, for instance ideas.
3. Substantial dualism: Some models consider that reality is made up of two substances, one material (actual objects) and the other immaterial (relations, laws of movement, ideas or other entities). This position should not be mistaken with the next one.
4. Existential meta-dualism. In this ontology, what exists is beyond matter (or objects) and ideas (or relations). In this meta-dualist view, objects and relations are only the two inseparable and complementary aspects of existing wholes (for example systems). We will come back to this ontology in the next section.

The vast majority of models belong to what can be called physicalist functionalism. This means they adopt a materialist ontology and a functionalist epistemology. The structures and processes within the brain are analyzed in terms of physiological

processes whereas the behavior of the brain and of its functionalities is connected to its internal organization, to the way its different parts are interconnected.

1.2.2 Brief Survey of the Main Models of Consciousness

The models of Francis Crick (Crick, 1994) and of Daniel Dennett (Dennett, 1993) are in this category. For Crick, the mind sits in the brain and its functioning depends on the internal organization of the brain. The unitary experience of consciousness would arise from a 40 Hz synchronous oscillation of neurons. Crick's philosophical attitude is between eliminative reductionism (consciousness does not exist) and explicative reductionism (one item (consciousness) is explained by another (physiological processes). Dennett belongs to the behaviorist tradition. For him, the brain is a massively parallel network, similar to a huge multilevel computer.

For Gerard Edelman (Edelman, 1989) in his *Biology of Consciousness,* the functioning of the brain derives from cerebral structures produced by a process similar to Darwinian selection, a sort of brain self-organization. This rather sophisticated model nevertheless does not give any indication on the phenomenal consciousness (qualia, subjectivity, experience). Israel Rosenfield (Rosenfield, 1993) proposes that self-consciousness emerges through a kind of self-referential recursive dialogue between actualized past experiences and present experiences.

Roger Penrose's very speculative model (Penrose, 1993) seeks the source of the brain's global coherence by quantum effects and by some non-algorithmic – Gödelian - effects which forbid consciousness to be analyzed by numbers, by mathematical equations. This holistic functionalism shows that the consciousness' nature is not accessible by dualist approaches

After this very brief survey of the main types of approaches to consciousness, we will summarize our holistic metamodel and will then apply it to the very elusive concept of consciousness. But first let us recall the crucial challenge brought about by the explanatory gap.

1.3 The Explanatory Gap - The Hard Problem

The existence of an explanatory gap is an indication of our incomplete understanding of how consciousness might depend upon a non-conscious substrate, especially a physical substrate. The basic gap admits variations in generality and thus in strength. In perhaps its weakest form, it asserts a practical limit on our present explanatory abilities: given our current theories and models we cannot now articulate an intelligible link. A stronger version makes an in principle claim about our human capacities and thus asserts that given our human cognitive limits we will never be able to bridge the gap.

In this context, David Chalmers (Chalmers, 1996) has proposed a distinction between the easy problems – i.e. the correlations between the neurophysiolgical processes and the cognitive performance - and the hard problem – i.e. the correlation between the

neurophysiological processes and the subjective phenomenal experience. He developed a model of consciousness that should satisfy two apparent contradictory conditions:

1) to rest on the brain and its internal physical organization (physicalist functionalism) and

2) to exhibit the non-reducible character of the existential subjective experience.

Some call this approach non-reducing functionalism, others call it neo-dualism or even a kind of monism, given the fundamental, holistic and universal character of experienced consciousness. Epistemologically, Chalmers describes cognition by functionalism, i.e. by the organization of the brain, the networks of interconnections between its parts, but recognizes, on the ontological level, that consciousness cannot be reduced to physical processes.

To conclude this section on the explanatory gap, let us mention Dennet's claim that the explanatory gap reflects the limits of our current theorizing rather than an unbridgeable in principle barrier.

This is precisely what we propose in what follows: a new non-mechanist onto-epistemology developed to interpret complex systems on their way toward autonomy, like living and conscious systems.

2 A Proposal for an Epistemology to Interpret Complex Autonomous Systems

The essential novelty of our work is the distinction between two worlds within nature: the material world of actual objects (like in physics) and the immaterial world of virtual relations (like in cybernetics). Indeed, in the mainstream Newtonian mechanist view, there is only one world, the world we apprehend through our senses, the so-called centimeter-gram-second (cgs) world, the place where everything happens. For many ordinary situations, this framework is largely sufficient. But when it comes to study complex systems with several kinds of closed loops in their organization, like in living organisms or in conscious beings, the mechanist epistemology is totally inadequate. As we shall see, our (apparently more complicated) way of interpreting complex systems simplifies the understanding

a) of the logic of life and

b) of the nature of consciousness.

2.1 The Main Features of the Proposed Onto-Epistemology

The purpose of our metamodel, whose details can be found elsewhere (Schwarz, 1997, 2004) is not only to describe things like in mechanical sciences, i.e. pre-existing objects (atoms in physics, individuals in social sciences); but to describe systems, i.e. more or less complex entities defined as sets of several (at least two) interacting parts. Therefore our starting point consists of the three inseparable primal categories present in all systems: objects, relations and wholes; these three types of initial ingredients are on

equal footing – in particular (non-material) relations which are as "real" as objects. Our metamodel goes beyond the mechanist paradigm where objects have a privileged ontological status, to a paradigm where the ultimate "reality", i.e. what exists, has two complementary, inseparable and irreducible aspects: objects and relations giving rise to a third level, that of the existential whole (the system).

Our metamodel is a general epistemological framework through which detailed models can be built for particular complex situations, as can be met in ecology, in biology, in social sciences or in cognitive sciences. These systems are not only characterized by dense networks of interactions, feedback loops, emergence of new structures (chaotic non linear systems) and temporary high sensitivity to noise; more fundamentally, we suspect that, in principle, they cannot be understood in the mechanist paradigm where it is supposed that the changes can be computed by a permanent set of invariant equations as can be done in astronomy for example. In complex systems, the equations themselves can change with the changes in the concrete system. In these cases we propose that a completely different approach be used, which goes beyond the Cartesian dualist pair (res extensa and res cogitans) and reaches the holistic level of existence.

As we can observe, the notion of "real" immaterial relation is hard to assimilate for people accustomed to the materialist point of view of the mechanist framework where only things are real. Even more difficult to apprehend scientifically are the concepts of whole, of existence or of being, which are traditionally associated with religion and philosophy, or, in the best case, with the "soft" sciences. Whatever their names, we are convinced that science will need meta-physical and meta-cybernetical notions that refer to a system as a whole and to its holistic, unitary and existential characteristics. We hope our metamodel is a useful step in this direction.

We think that consciousness cannot be understood clearly without a deep change in the onto-epistemological foundations of the framework used to interpret it. In our view, the "hard problem" is not a neurophysiological or a computationist scientific problem, but an ontological and epistemological problem. The accumulation of unsuccessful research in the fields of biology and of information science to clarify the nature of consciousness is an indication of the need of new fundamental dimensions.

As mentioned earlier we will not go into the details of our metamodel which has been exposed elsewhere but recall here only the points useful for its application to the problem of consciousness. Additional figures can be found in (Schwarz, 2004).

2.1.1 The Initial Ingredients of the Metamodel

Searching for the most general and simple configuration of things when we observe nature, we start with a very common system made up of two components in relation (see left of fig.1). It can represent either any pair of interacting objects or a subject observing an object. Drawing the conclusions from this trivial starting point, we propose that any existing situation is given by couples of interacting components, which constitute an existential whole, a "system". As can be seen in the prototypical system on the left of fig.1., we distinguish the actual physical interactions between the two parts

and the potential relations that may or may not be actualized. To be more specific, we distinguish the interactions between things, which belong to the physical world (like forces, exchange of photons, etc.) and the correlations between the state of an object and the state of another, which belong to the abstract world of relations.

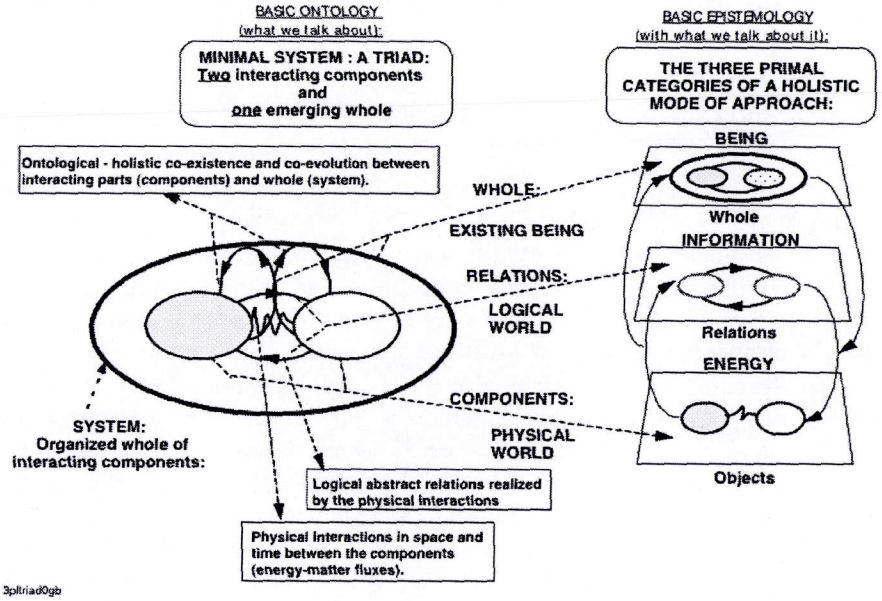

Figure 1. The basic entity (left side) described in our metamodel is the minimal system: a triad, i.e. a non-separable whole of two interacting components (ontology). The corresponding epistemology (right side) has therefore three primal categories: the physical world of objects (components), the abstract world of relations (images of interactions), and the existing world of the whole that is, i.e. the system.

As already mentioned, the usual Cartesian-Newtonian dualist view of an objective "reality" whose evolution is determined by some eternal "laws", is replaced here by a holistic approach where what happens emerges from a deep ontological dialogue between two inseparable and nevertheless irreducible aspect:s

- the physical world of the things, which we can perceive by our senses and which corresponds to the usual world of physics (Energy plane), and
- the cybernetical world of the potential relations immanent in the system (Information plane),

Such a potential relation can be later actualized as a physical interaction during the continuation of the dynamics of the system. This potential field is usually symbolized, in the framework of a theory, by symbols or algorithms, like numbers, parameters, differential equations, logical constraints or geometrical figures. But one should not

confuse the symbols of a theory, which are human artifacts, and the immanent potential relations existing in the system, which are part of nature. The permanent ontological dialogue between the real physical aspect of the system and its virtual potentialities is represented on the right side of fig.1 by the loop connecting the physical plane and the information plane and its integration in the system as an existing whole (plane of being). The usual mechanist dualist approach with the movements on one hand and the invariant laws on the other, is a particularly trivial case of the general case depicted here, where the laws can change when the movements proceed.

2.1.2 The Structure-Organization of Viable Systems

Starting from the primordial system and the basic epistemology indicated on fig. 1., we were led to a detailed general structure–organization for complex viable systems (see also Schwarz 1997b on the Web).

Broadly speaking, two types of systems exist in nature:

- Systems whose dynamics obeys the second principle of thermodynamics, i.e. the drive toward increase of entropy; these systems evolve from order to disorder or uniformity. Most inorganic systems are of this type.
- Systems (that, of course, also follow the second principle) whose dynamic is mainly driven by their internal organization; these (negentropic) systems' survival is due to the presence of several kinds of closed loops which give them the ability to escape the curse of the entropic trend toward disorganization, destructuration or destruction.

We will now discuss the main characteristics of these viable systems whose examples are living organisms and conscious cognitive beings. The structure-organization of this type of systems can be seen in fig. 2. Systems, that have non-entropic dynamic features, like self-organization, self-regulation, self-production (autopoiesis) or self-reference cannot be understood if we reduce them to their manifestation in the usual space-time as it is done in mainstream science. They have to be studied in both the physical plane (plane of energy) and the cybernetical plane (plane of information) as well as their form in the existential plane (plane of totality).

We now make some comments on the main loops in the organization of complex systems on their way to autonomy:

- Vortices, self-organization and feedback loops are well known processes
- Autopoiesis is a kind of ontological loop proposed by Maturana and Varela (Zeleny, 1981) between the physical plane and the relational plane; it corresponds to the logic of life. It means that the physical organism contains a relational network (the laws followed by the physical processes in the organism) whose product is precisely this organism: there is mutual production of the network by the organism and of the organism by the immanent network. The signature of life is therefore self-production.
- Autogenesis. The ultimate cycle represents the impact of the system as a whole on its own producing (= autopoietic) dialogue; in other words, autogenesis, or self-creation, is what makes a system autonomous: an

autonomous system is able to create its own laws. Autogenesis is pictured in fig.2. as a loop that connects the system as a whole in the plane of totality, and its own self-producing (autopoietic) process.

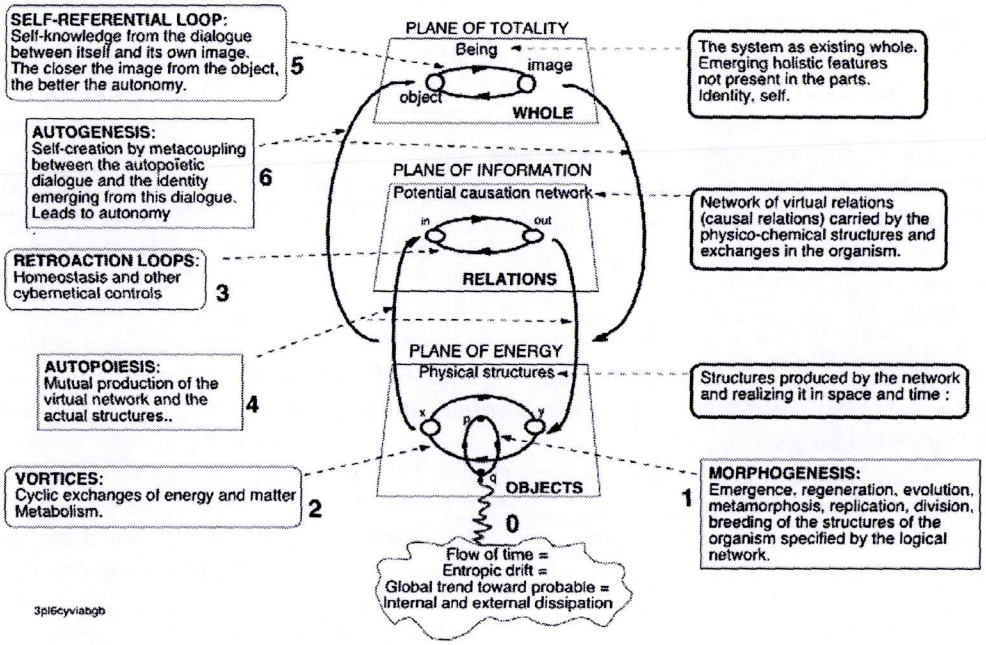

Figure 2. In this metamodel, the viability of complex systems on their way to autonomy depends on the presence of six cycles; three cycles are responsible for the stability (regulation) of the system and three cycles are responsible for the changes (adaptation to a changing environment).

The perennity comes from three cycles, one in each of the three planes:

1. recycling of matter in the physical plane (2, vortices),
2. feedback loops in the cybernetical plane (3),
3. self-reference loop between object and image in the totality plane.

The three "vertical" loops give plasticity to the system:

1. morphogenesis (1, self-organization) between the physical environment and the physical plane of the system,
2. autopoiesis (4, self-production) between the physical plane and the network organization within the system,
3. autogenesis (self-creation) between the autopoietic loop and the holistic plane of the system (the system's identity).

3 Consciousness and Self-Reference

After this presentation of the main points of our onto-epistemological language, we can now turn to the problem of consciousness, more specifically to the problem of the explanatory gap: how can consciousness emerge from a piece of non-conscious matter ? Why are there billions of pieces of matter in the world and only some are conscious ? Are there indicators pointing to conscious or potentially conscious material ?

To deal with this question, we have to concentrate on the third plane of the metamodel (fig. 3), the totality or existential plane. In this plane, we find a kind of onto-epistemological loop connecting the "object" and the "image". The "object" is here the existential dimension of the material system (which is in the lower physical plane). The "image" is the existential dimension of the immaterial network that belongs to the relational plane. This network is the mesh of potentialities and constraints, which drive the system's dynamics.

In the Newtonian mechanist dualist paradigm, the laws of movement belong to the world of "ideas", the world of the Cartesian *res cogitans*, and the physical objects belong to the sensible world of the *res extensa*. Therefore objects and laws of movement are ontologically different. Let us now consider living organisms of different levels of complexity and autonomy. For rather simple organisms like plants or primitive animals, the "laws of movement" of the organism must obviously satisfy the basic laws of physics and chemistry, but not only. The functioning of an organism also depends on the particular internal instantiation of the general laws of nature in that organism. All shrimps do not behave strictly in the same manner as watches do. This means there is some co-dependence between the laws and the movements.

Simple inorganic objects are completely determined by the general laws of physics in their textbook form. Complex, partially autonomous systems also obey the general laws of physics; but each individual system depends on the particular instantiation these laws take within that system: its history, its organization and other accidental features. In other words, one part of the laws come from "outside" the system (the universal laws of physics) and another part depends on the specific potentialities and constraints of that individual system, given its internal organization and own history.

The more complex a system is, the more important is the influence of its internal organization. The human brain, for example, indeed respects the laws of thermodynamics and the laws of gravitation, but its functioning, i.e. the transition from its state at time t_o to its state at time t_1 depends more on its internal potentialities and constraints as well as its history. The ontological dialogue between its physical structure and its relational organization (its self-reference) becomes ever more closed, i.e. autonomous. The brain-mind couple exists more and more as a unitary whole: it becomes self-knowing: in other words it becomes conscious.

We have tried to represent graphically this situation on fig.3. On the left side the picture represents an autonomous system with a rather low degree of self-reference; the objet and the image are well separated. The object and the image do not have much in common; the system is not conscious.

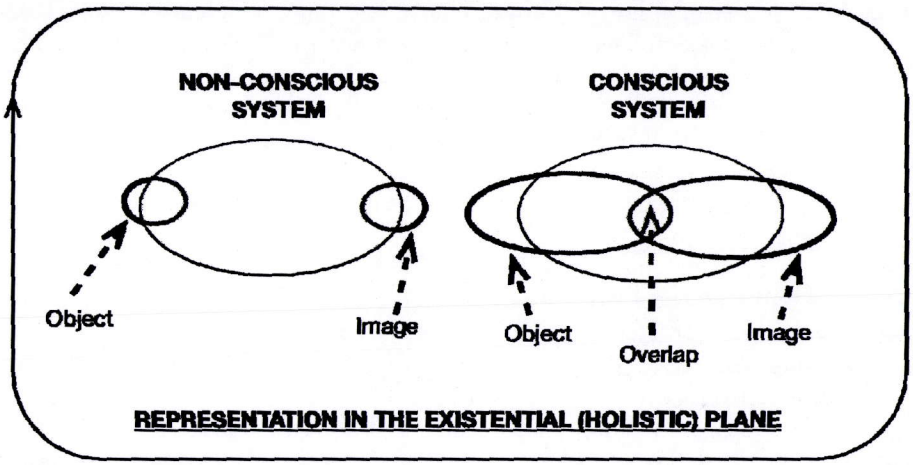

Figure 3. Graphical representation of the self-reference of autonomous systems in the existential plane (Plane of Totality in fig.2.). The self-reference loop links the symbol of the object, (the material aspects of the system) to the symbol of the image (the immaterial network of potentialities and constraints of its dynamics, the "laws of movement"). The word *symbol* used here refers to the existential (or holistic) dimension of the physical and relational aspects respectively.

On the right side of fig.3., on the contrary, the picture represents the situation for a conscious system: for example in the brain, the neurophysiological objects and fluxes are very similar to the organization of its internal network of connections. This similarity is represented graphically by the overlap between the symbol of the object and the symbol of the image. In the extreme case of a completely autonomous system, the overlap would be complete, the brain would be its network and the network would be the brain. There would be fusion of the two ontological dimensions, objectal and relational; the system would exist by itself as a whole, as a unity.

The human brain has not (yet?) reached this degree of existence, but its self-reference manifests itself by some degree of self-knowledge that is called consciousness.

4 Conclusive Remarks

Explaining the nature of consciousness is one of the most important and perplexing areas not only of philosophy, but also of science. Broadly speaking, in the philosophy of mind, theories follow two traditions: materialism and dualism. The former holds that all mental activities can be reduced to physical processes in the brain; the latter admits the possibility of some non-physical ingredient. In the present work we propose another framework - inspired by the systems paradigm - which is not materialist neither dualist but could be called existential meta-dualism. The foundations of our systems

interpretation consist of three primordial categories: objects, relations and wholes or systems. The onto-epistemology adopted here has a dualist flavor in the sense that it has both a materialist aspect (the objects) and an immaterial aspect (the relations). However, these two categories are not simply added but integrated in a third category, the holistic category of existence. That is the reason why we call this approach "existential meta-dualist". The absence of this third category in mainstream science is the reason, in our opinion, why consciousness (which is a manifestation of existence) is so difficult to interpret in usual space-time science.

Contrary to the usual mechanist physicalist approach, every system in nature is characterized by its physical structure and its relational organization. In simple (mainly inorganic) systems, structure (actual anatomy) and organization (the "laws of movement") are somewhat decoupled: the laws of physics do not depend on the system's characteristics.

For complex and partly autonomous systems, like living and conscious organisms, physical structure and processes, on one hand, relational organization (the potential relations that define its dynamics), on the other hand, are much more dependent on one another and therefore more similar than is the case in simpler systems. This ontological proximity between structure and organization is the source of the high degree of self-reference, of autonomy and of unitary existence. Life and consciousness are the manifest consequences of these abstract fundamental topological circularities.

References

Chalmers, David J. (1996). The Conscious Mind: In Search of a Fundamental Theory. Oxford University Press.

Crick, Francis (1994). The Astonishing Hypothesis: The Scientific Search for the Soul. Scribners.

Dennett, Daniel C. (1993), Consciousness Explained. Penguin Books Ltd (June 24, 1993)

Edelman, Gerald M. (1989). The Remembered Present: A Biological Theory of Consciousness. Basic Books

Penrose, Roger (1994). Shadows of the Mind. Oxford University Press

Rosenfield, Israel (1993) Strange, Familiar, and Forgotten: An Anatomy of Consciousness. Vintage

Schwarz, E., (1997a), Toward a Holistic Cybernetics. From Science through Epistemology to Being. Cybernetics and Human Knowing, Aalborg (DK), Vol.4. pp 17-49.

Schwarz, E., (1997b) http://www.autogenesis.ch/Res1997.html

Schwarz, E., (2004). Is Consciousness Reality or Illusion? A Non-Dualist Interpretation of Consciousness. Computing Anticipatory Systems: CASYS'03 – Sixth International Conference, edited by D. M. Dubois, published by The American Institute of Physics, AIP Conference Proceedings 718, pp. 425-434

Zeleny, M. edit (1981). Autopoiesis. A Theory of Living Organization. North Holland,