Will Computers Ever Think ? On the Difference of Nature Between Machines and Living Organisms

Eric Schwarz
Autogenesis - Centre d'études sur les systèmes autonomes
Université de Neuchâtel, CH-2000 Neuchâtel, Switzerland.
e-mail: Eric.Schwarz@unine.ch

Abstract

One should not forget that the initial aim of the first builders of computers and scientists in Artificial Intelligence was to make "thinking machines". Today, people are more cautious about the nature of intelligence and the difference between natural intelligence and artificial intelligence. Numerous definitions of thinking and of intelligence have been proposed. Nowadays, even consciousness, which has been taboo for a long time in the AI circles, begins to be discussed. In this paper, after a brief introduction about the different paradigms of the cognitive sciences in the last 50 years, we present the main features of a systemic holistic metamodel that has been proposed recently to interpret complex self-organizing systems evolving toward autonomy, like those found in biology, in the social and cognitive sciences. We then use this graphical language to show the specificity of living systems which is autopoiesis or self-production; we then propose an interpretation of consciousness which is related to self-reference. Using these concepts, we then comment on the question raised in the title.

Keywords: nature of life, autopoiesis, intelligence, self-reference, consciousness.

1. Introduction

In this paper we would like to propose some reflections about the correlation between the structure and organization of a system, on one hand, and its capacity to think and to be conscious, on the other.

Obviously, the first point to clarify is to define what we mean by thinking, and also to discriminate between the close notions of thinking, mind, intelligence, consciousness. It is not the first time that the apparently trivial question "what is intelligence?" is raised. In fact the definition of thinking is so tortuous that a scientist as deep as Alan Turing could not propose a more sophisticated test than hiding a human being and a potentially thinking device behind a curtain; if an observer is not able to tell the difference from the answers of the two systems, they are both declared to be thinking.

The ideas about the nature and processes of intelligence, the outcome of thinking, have changed dramatically over the last half-century. Before the 1940's, the questions about intelligence were left to psychologists, since thinking was reserved to humans. But with the fabrication of computers, "intelligence" was not a human exclusivity anymore, it could be artificial.

International Journal of Computing Anticipatory Systems, Volume 8, 2001
Edited by D. M. Dubois, CHAOS, Liège, Belgium, ISSN 1373-5411 ISBN 2-9600262-1-7

According to the first approach in the cognitive sciences in the 1950's, the brain was a seen as computer, an information processing machine manipulating symbols whose logical combinations were supposed to represent real-world phenomena. Despite their good performance to solve abstract logical problems, computers were, and still are, unable to solve simple daily life problems.

The connectionist approach was then proposed, according to which a cognitive system, like the brain, is seen as a self-organizing network of elements able to learn by sensory-active loops with its environment. The quality of the performance of the neural networks is then measured by the adequacy of the behavior of the system in its environment. Although more pertinent than the symbolic approach, connectionism, with a better embodiment of the cognitive functions, still leaves many features of real living systems hard to interpret. For instance, the capacity to learn in the connectionist robots must be prepared by the designer.

The central concept of the third approach to cognition is autonomy. An important feature of intelligence is the capacity to survive in one's Umwelt. Such an entity must first have the right functionalities to respond to the challenges of its environment. Its viability is improved if it is, in addition, able to learn from past experiences. The capacity to evolve increases further its viability on a larger time scale. Builders of robots try to design devices having these three different capacities: functioning, learning and evolving.

An important concept to help designing so called autonomous robots is that of autopoiesis or self-production, which has been proposed as the basic logic of life (Zeleny M., 1981). A system is said to be autopoietic when a closed loop exists between the physical processes in that system and its logical organization. In other words, in a living system, the physical structures and fluxes are such that they constitute an abstract network of causality whose effect is to re-produce the physical structure and fluxes that generated it. The system is operationally closed, it produces itself; but it is so in the environment in which it has grown. F. Varela has proposed to name enactionism (Varela F., 1989) this third approach where cognition is seen as a continuous process of co-creation of a world by the interactions of agents in an environment. The logic (and even the meaning) of that whole system cannot be grasped if one tries to separate the agents and their environment.

In this view it is not only impossible to separate the structure (the body) from the functioning (the mind) of the agents, but also it is also impossible to understand the agents in isolation from their environment and their singular, unique history.

In this paper we do not present another phenomenological model for cognition. We propose instead a new framework, that is a new epistemology and a new ontology to interpret complex and partly autonomous systems, like living, cognitive and conscious systems. Mainstream materialist and dualist mecanist science is very efficient to understand simple heteronomous physical systems ruled by permanent laws like Newton laws of movement. But for complex systems with dense and changing interdependence networks between their components, a more adequate framework should provide a more meaningful interpretation.

We think that "reality" is not reduced to material objects moving in space and time as it is the case in the mechanist view, but we propose that what is, the existing, is the whole that emerges from the conjunction of two irreducible and inseparable aspects: actual substance, which corresponds to the usual material objects studied by physics, on one hand, and the immaterial immanent network of relations which creates and constraints the possible future states of the system, on the other hand.

In this new context, we will show how the structure and the organization of systems of increasing complexity like:

1) a simple dissipative object like an electric heater or a von Neumann computer,
2) a self-producing (autopoietic) system like a living organism, and
3) a self-referential entity like a human being,

generate qualitatively different properties, like the simple trend toward disorder, uniformity and disappearance for the first one, a capacity to be viable, that is to last, for the second, and a property of self-knowledge, which we identify to consciousness, for the third one.

Before describing the main features of our metamodel, we will recall some steps in the history of understanding of what thinking means.

2. Historical Background: What does "Think" Mean ?

2.1. The Mind-Body Problem

We can suppose that very early in history, men were puzzled by the obvious difference between a living person and a dead body. The missing entity in the dead corpse was given a name, like "soul", whose features were discussed with the words of mythology, philosophy, theology, ethics, and other fields. In most views, the whole reality was not restricted to the perceived material world. Some gave indeed priority to the material world, others to the world of ideas (idealism, spiritualism), others were dualist like Descartes who envisioned two kinds of things: res extensa, the thing of the usual space and res cogitans, the thing of the world of thought. Later, phenomenologists like Merleau-Ponty insisted on the non-separability of the mind and the body.

In the 20th century, several models were proposed about the mind / body problematics.

- Physicalist behaviorism holds that mental states do not exist, only physical processes between stimuli and responses are real.
- The identity theory claims that brain processes and mental states are the same thing; they are only different descriptions of the same phenomenon.
- In the functionalist view, the processes in an organism do not depend so much on the material substance involved but mainly on its structure and organization; the functioning depends on the way the parts are connected.
- About the processes involved, the computationists think that the mind works by manipulating symbols according to rules, like a computer.
- The connectionists see the brain as a self-organizing system interacting with its environment, whereby emergent features can appear.

Let us mention finally two more exotic views. The strongly dualist model of Eccles and Popper, where the brain and the mind are two independent and interacting entities: the conscious mind has a causal effect on the brain. In the panpsychism-epiphenomenalism view, every matter has a sort of inherent internal "mental" or protopsychic state; since this state is inherent to matter it has no causal effect on it.

2.2. What is Intelligence ?

Thinking has been variously associated with the ability

- to solve problems,
- to manipulate symbols according to rules, i.e. to compute,
- to survive, that is to insure the viability of the organism, or
- to have a purpose, therefore to have a teleonomic logical organization.

The outcome of thinking is called intelligence. Descriptions of intelligence are also very numerous: the aptitude to understand, to grasp with the mind, to adapt, to choose, or to give meaning. Intelligence is often said to imply the sequence to collect information, to process it, to make decisions, to act, to learn, to have objectives.

2.3. Artificial Intelligence versus Natural Intelligence

With the advent of computing machines in the late 1940's, the pressure became stronger to have more precise, general and concrete methods to define and measure mental activity and intelligence of human and non-human devices (see details in Churchland, 1990). In 1950 Alan Turing proposed his famous criterion to qualify mental functions: if a human observer cannot distinguish the answers given by a hidden machine and a hidden human being to all sorts of questions, the machine is said to be able to think.

It was shown that a Universal Turing Machine (UTM), a machine able to compute any input/output function, should be able to pass the Turing test for intelligence. Consequently, in the 1950's the answer to the question raised in the title was a clear yes. In the 1970's and 80's doubts began to appear. In the 1980's, workers in AI and especially in pattern recognition met problems of computing capacity and of time consumption, despite the increasing speed of computers. At the same time, Searle questioned the assumption that intelligence can be reduced to manipulating symbols according to rules. At the same time, P. M. and P. S. Churchland, observing the large difference of structure between computers and brains, became skeptical about the capacities of digital computers to think as living brains. But they did not exclude this possibility for massively parallel machines with a more brainlike structure.

A device that passes the Turing test is considered as intelligent because, from its answers, it cannot be distinguished from a human being. But is it conscious ?

Now, let us consider a zombie, defined as an entity that behaves like a human being in every respect, but has no consciousness: "there is nobody there". It does not experience its own existence; it has no unitary, singular identity, no phenomenal consciousness, this eminent quality of a human mind. Obviously, by definition, a zombie would pass the Turing test with success but is not an intelligent conscious thinking being. The conclusion is that the Turing test cannot tell the difference between thinking in the human sense, that is conscious thinking, and simulation of thinking.

The UTM, which is able to compute any input/output function, simulates the behavior of a thinking mind, but lacks the existential dimension of a conscious thinking human being. We conclude here that the computationist-cognitivist paradigm is not able to interpret this important feature of the human mind that is consciousness. What about the connectionist paradigm ?

2.4. The Computationist Model versus the Connectionist Model

In the connectionist paradigm, the brain is considered as a self-organizing system, that is a nonlinear dynamical system interacting with an environment. The emergence of new structures through this interaction corresponds to the function of "learning". Cognition corresponds to the emergence of global states in a network of simple components. Such a system is said to work in an appropriate fashion when an adequate solution is found to a given task.

From this brief description it can be seen that both the cognitivist and the connectionist approaches belong to the mechanist paradigm of mainstream science: both are realist, in the sense that there is a given reality out there; both are materialist, they hold that reality is made out of matter only; they are essentially dualist in the sense that the movements of matter in physical space and time can be expressed in the language of logic and by the laws of physics. They follow the objectivist view according to which the objects of reality have properties that do not depend on the interpretation or the presence of an observing subject.

We conclude at this point that the problem of consciousness is not only a scientific challenge asking for some new theory, but above all, questions the validity of the epistemological foundation of science, and the pertinence of its primordial categories like space, time, energy, and the interactions of physics.

Physics and the physicalist sciences have demonstrated their great efficiency in relatively simple situations where energy and entropy play a dominant rôle, like in terrestrial and celestial mechanics, in thermodynamics, and more generally in inorganic contexts. But in the case of complex systems, like social, living, or cognitive systems, the internal organization, the networks of cybernetical feedback loops, the emergent properties, are more pertinent than the purely physical aspects.

3. Proposal for a Holistic Metamodel to Interpret Complex Self-Organizing Systems Evolving Toward Autonomy.

We have recently proposed a new basic framework from which one can build models to interpret real life complex systems having some degree of autonomy, or operational closure, like self-organization, self-regulation, self-production or self-reference. As the details of this language have been published elsewhere (Schwarz, 1997a, 1997b), we will present here only its main characteristics before applying it to the case of conscious systems.

3.1. Primordial Categories and the Prototypical System

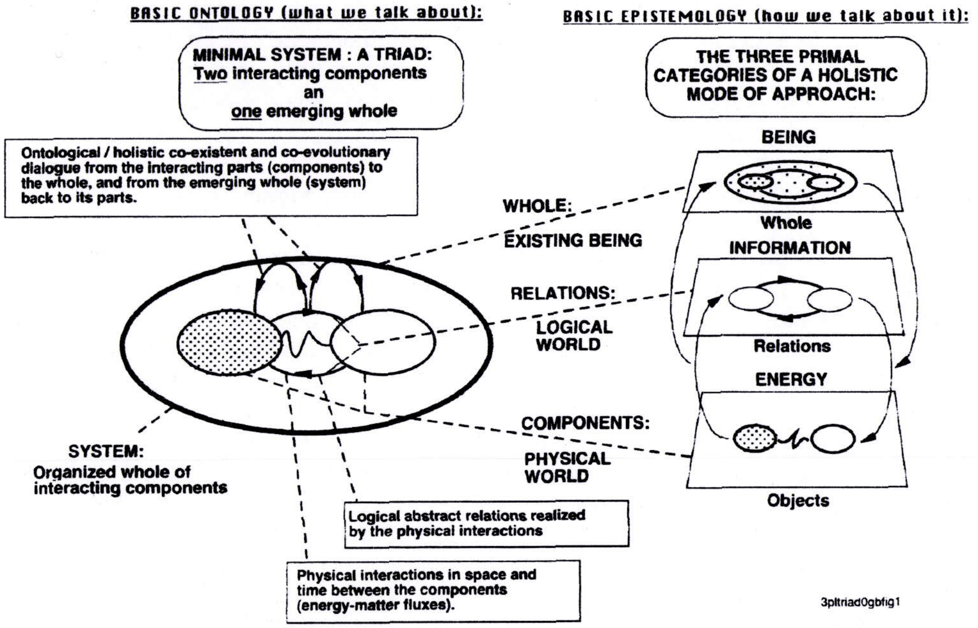

Fig.1. The basic entity which is the object of the description proposed in this paper is the minimal system: a triad (non-separable whole) of two interacting components (ontology). The corresponding epistemology has therefore three primal categories: the physical world of objects (components), the abstract world of relations (images of interactions), and the existing world of the whole which is (system).

Looking for the most general configuration of things when we observe nature, we propose a most simple and general system made up of two components in relation (see fig.1). It can represent either any pair of interacting objects (for something to happen you need to be two!) or a subject observing an object. Drawing the epistemological and ontological conclusions from this trivial starting point, we propose that any existing situation, is given by couples of interacting components, and constitutes an existential whole, a "system" emerging from the ontological confrontation, at all scales, between "the objects" constituting the physical aspect of the system, and "the relations", the virtual network of causality immanent in the system, which represent the possible subsequent states of the system and which manifests itself by the interactions between objects.

In other words, the usual Cartesian-Newtonian dualist view of an imperial "reality" whose evolution is determined by some eternal "laws", is replaced by a holistic approach where what happens emerges from a deep ontological dialogue between two inseparable and nevertheless irreducible "phase spaces" or worlds, the physical world of the things, which we can perceive by our senses and which corresponds to the usual world of physics, and the cybernetical world of the potential relations, one of which will be realized in the next stage of the evolution of the system. This potential field can be symbolized in the framework of a theory by symbols or algorithms, like numbers, parameters, differential equations, logical reasoning or geometrical figures. One should not confuse the symbols of a theory, which are human artifacts, and the potential relations, which are part of nature.

The main difference with usual physics is that the laws of physics are invariant and represented by the equations of the theory, whereas here, in the general case, the relations change each time the structure of the system changes and, furthermore, the relations do not belong to some theory but to the system as a whole. On fig.1. are represented, on the left side the prototypical simplest system, made of two interacting components (basic ontology), and on the right the corresponding three primal categories: objects (for example energy-matter), relations (basis for information), and wholes (systems), which are used in our metamodel to describe the world (basic epistemology).

3.2. The Six Cycles of Viable Systems.

In our metamodel for complex self-organizing systems, a viable system, which is typically a living systems, or more generally any system that is able to survive thanks to its internal logic, is characterized by the presence of six cycles, which can help interpret the evolution of a system toward complexity and autonomy as well as the functioning of any viable system. Let us make some more comments on these six cycles and/or steps (or seven steps if we include the entropic drift) with the help of fig.2.

0) The entropic drift of the medium is the natural trend of the preceding (parent) system, which may drive it far from its stable point ("far from equilibrium"), where a fluctuation can be amplified and start a catastrophic cascade of changes. This natural drift corresponds to the trend toward the more probable formalized by the increase of entropy for the most simple systems; for more complex cases this same drift can be more adequately called actualization of potentialities or Popperian propensions.

1) Morphogenesis. The first of the six cycles can be visualized as a positive feedback loop between two (or several) mutually produced variables or parameters of the medium far from equilibrium, with the effect of differentiating the medium (dissipative structures, cancerous cells or demographic proliferation for example).

2) Vortices. The second cycle is a physical cycle in space and time, like vortices in a moving fluid, ecological recycling of matter, or oscillations like heartbeats. A valid relation must be circular; it is the first necessary condition for perennity.

3) Feedback, Homeostasis. The next step in the development of a viable system is the possibility of being stable. This feature requires the compatibility between the fluxes and exchanges in the physical plane (vortices, physiology) and the corresponding network of causality, that can be seen as an abstract image of the concrete processes. The regulating feedback loop belongs to the relational, or cybernetical plane.

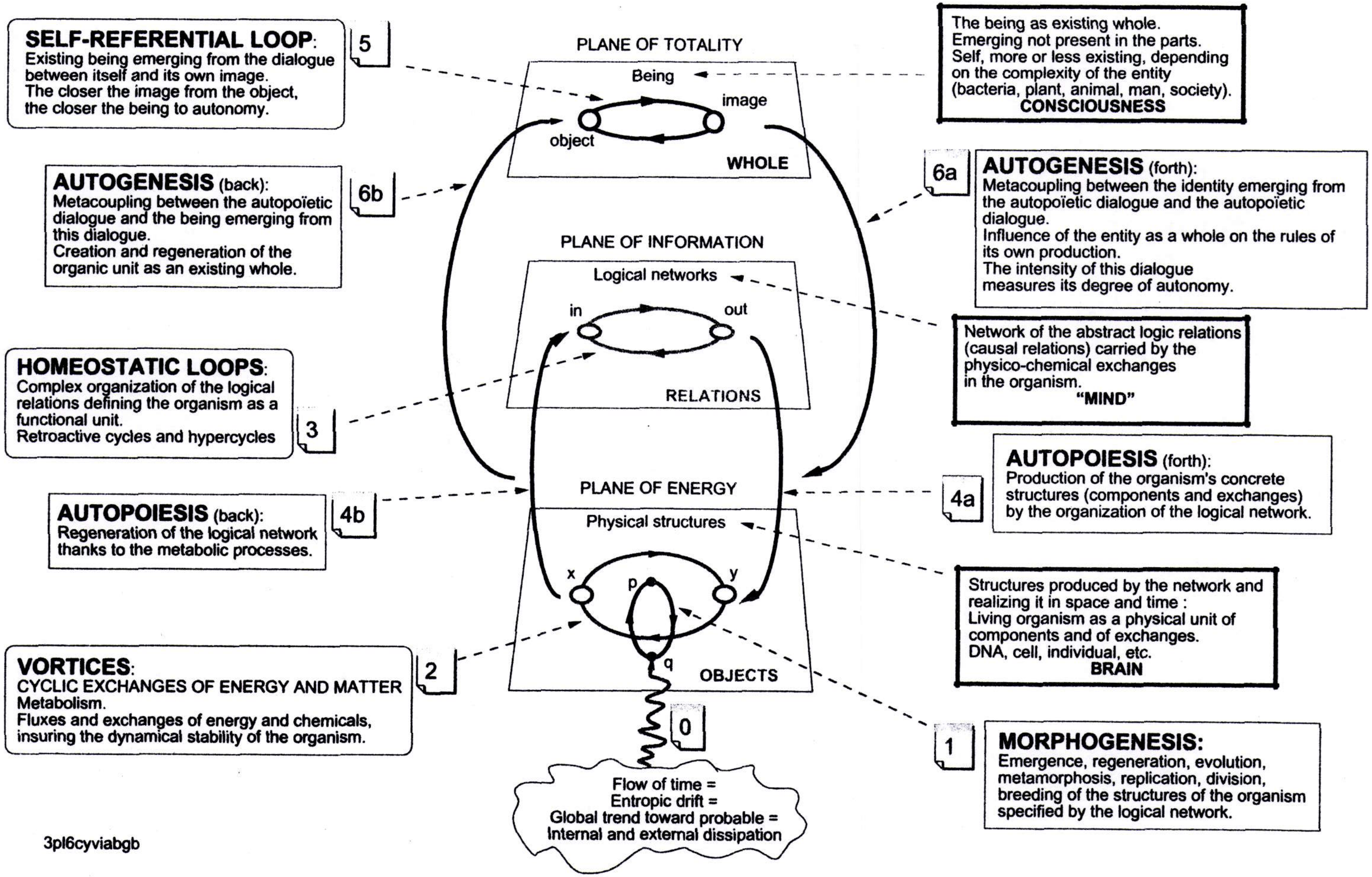

fig.2. The six cycles defining viable natural systems, in the three physical, relational and existential planes. The three "horizontal" cycles (vortices (2), homeostasis (3), and self-reference (5)) inside the three planes are responsible for the stability of the system; the three "vertical" cycles between the planes (morphogenesis (1), autopoïesis (4) and autogenesis(6)) are responsible for the changes. For human beings, the respective places of the brain, the mind and consciousness are also found.

4) Autopoiesis. When a homeostatic system complexifies for billions of years like it was the case for the prebiotic evolution, it may reach a level where there is not only compatibility between the physical structure and the logical organization, but also self-production: the organism incarnates a causality network which produces the organism that incarnated it. This new super-circularity, called autopoïesis and proposed by Maturana and Varela (see ref. 3) is pictured here as a loop that connects thephysical plane and the relational plane. A self-producing (= autopoïetic) system is an entity that, as a whole, produces itself by an adequate dialogue between its organic structure (and material fluxes) and its own network of causality. This step corresponds to the logic of life.

5) Self-reference. Autopoïesis is the beginning of self-reference: the system is its own reference. The system is operationally closed; a completely autopoïetic system does not need any logical connection with the outside. In the picture, self-reference is symbolized by the overlapping between the object and the image, the two terms in relation in the holistic plane. The object can be seen as the organism (the brain, for example) and the image as the immaterial network ("the mind" in traditional parlance). In this metamodel, the degree of self-reference of a system is interpreted as its level of self-knowledge, which means its level of consciousness.

6) Autogenesis. The ultimate cycle represents the impact of the system as a whole on its producing 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 existential plane and its own self-producing (autopoietic) process. A strictly autonomous system is operationnally closed: it has absolutely no logical connection with the outside world. The actual systems and sub-systems forming the Earth living system are only partially autonomous systems.

Let us notice that three cycles contribute to the stability of the system: vortices (recycling of matter), self-regulation and self-reference; the other three cycles, morphogenesis, autopoiesis and autogenesis insure the capacity to change that also contributes to the perennity of the system as an identity.

4. Life and Consciousness

We are convinced that the metamodel presented in this paper is a more pertinent framework than the usual physicalist approach to interpret these two daughters of complexity that are life and consciousness. As we have said above, it is based on three primordial categories: objects, relations and existing whole. This ternary starting point replaces the usual and apparent evidence of matter as sole aspect of what exists, which is assumed in mainstream science. If this ontological-epistemological hypothesis is adopted, then the notions of life and of consciousness are easier to grasp.

4.1. Life

Starting from the verbal explanation of Maturana and Varela about the concept of autopoiesis, which is necessary and sufficient to define the organization of the living systems, we translate it into the graphical pattern of our model.

Let us first quote the definition of the concept of autopoiesis given by Varela (Varela 1979) (our translation): "An autopoietic system is organized as a network of processes of production of components which:

1) through their interactions and transformations continuously regenerate the network of processes that produced them; and:

2) constitute the system as a concrete unity in the space where it exists by specifying the topological domain of its realization as a network".

As can be seen in fig. 2 in our scheme for viable systems, autopoiesis is a sort of ontological loop that connects the organism as a physical structure and a set of physiological processes in the usual space (energy actual plane), on the one hand, with the immaterial causal network immanent in that organism and which contains its possible future states (information virtual plane), on the other hand. The ascending left side of the autopoietic loop represents the effect of the structures and fluxes in the organism on the network of causality that opens and limits the possible future states of the system, whereas the descending right side of the loop represents the effect of that field of possibilities on the next actual state of the system.

This loop is called ontological because it binds two irreducible and inseparable basic aspects of any existent entity. In a common non-living object like a stone or a car, the set of laws that prescribe its movements, like the laws of mechanics or of thermodynamics, is invariant and does not depend on the movements taking place. This situation gives rise to the regular reversible movements of mechanical systems (planets for example) or to the irreversible entropic trend toward disorder, uniformity or equilibrium typical of classical thermodynamical systems (trend toward thermal death).

The extraordinary specificity of living systems or more generally of self-producing complex systems, is not only the presence of the autopoietic loop which symbolizes the fact that the set of laws (physical, chemical, biological, etc) that drives it, changes all the time with the changes of the material structures and fluxes in the organism, but – this is the extraordinary part – the fact that the virtual causality network generated by the material structures and fluxes re-generates these same kind of structures and fluxes (with possible minor changes at each turn). A living system produces itself: the blueprint implicit inside it, explicits the system, which implies again a regenerating network, and so on; there is a sort of ontological turning wheel or oscillation between the structural-objectal aspect and the organizational-relational aspect of living - or more generally viable – systems. A living system is a (partially) autonomous system: it has its own law inside itself; an inorganic system is heteronomous: the laws to which it obeys are outside: Newton law, 2nd principle of thermodynamics, etc.

4.2. Consciousness

Our pattern for viable systems (fig.2.) also points to a possible interpretation of consciousness and its nature.

Autopoietic systems are not only a collection of parts like a pile of sand or even a watch: they exist as wholes, they are holistic existential systems, they have an identity, they exist as a unity. They have emergent features not present in the parts or in the relations. The presence of the autopoietic loop that connects the objectal (energy) world to the relational (information) world means that the system can be reduced nor to its physical structure neither to its relational organization, but exists as a whole emerging from this ontological connection between the actual objects and the potential relations.

These holistic and unitarian dimensions are difficult to interpret in the framework of the physicalist and reductionist natural sciences, which have built their reputation since the Renaissance by going away from the non refutable speeches of the philosophers and of the theologians. But we think that science today has almost exhausted the potentialities of hard atomistic materialism and will have to open itself to other dimensions in order to interpret complex autonomous systems. The ontology and epistemology of science will have to be widened.

In this perspective, let us go back to fig.2. and particularly to the third plane, the existential plane, where is found the self-reference loop between object and image. What does it mean ? Let us first remind that the existential plane symbolizes the holistic aspect of the system, its existence as a whole, as a unity. The "object" indicated in the self-referential loop is the contribution of the physical-structural aspect of the system (the visible organism) to its existence, whereas the "image" is the contribution of its (implicit) relational-organizational network, which can be thought of as the distribution of the probabilities of its future states, like the probability wave function in quantum theory. In other words, *the complete present state of a system is not only its actual* material state (the object in the energy plane) but also contains implicitly its virtual future state (the image in the information plane) in the form of the distribution of probabilities following from the general laws of nature applied to the particular configuration of that system.

The self-referential loop in the existential-holistic plane is a sort of meta-image of holistic dimension of the autopoietic loop that connects the organism – the object – with the causality network – the image - that will regenerate it or slightly modify it. The more self-referential a system is, the more its network - the image - is adequate as compared to its state - the object. Adequate means here that the "image" allows the system to survive, to last, to continue to exist.

A kindred idea was suggested some time ago by Robert Rosen who proposed a cybernetical model of anticipative systems according to which such a system has a model of itself and of its environment enabling it to make predictions and act accordingly (Rosen 1979). But in the same article this author confessed that "I continue to believe that the properties of anticipatory systems raise new questions for the scientific enterprise of the most basic and fundamental kind." Indeed our model is not a phenomenological model built to solve some specific concrete problems, but is the result of a fundamental questioning of the ontological and epistemological basis of science. It is encouraging to see that critical thinking about the foundations of science can lead to former intuitions of renown scientists and also opens roads to very concrete questions, for examples questions related to making thinking machines.

In summary, self-reference is the consequence of autopoiesis or self-production, which is the production of the physical structures and processes of a system by the immanent set of natural laws that are activated by those same structures and processes. But self-reference is also the threshold of a kind of new super-circularity, autogenesis, the intervention of the system as the entity produced by the autopoietic process, in that autopoietic process. When a system, political or other, is able to intervene in its own laws of production, it is said to be autonomous. In other words, autogenesis is the process within a system that enables it to become autonomous or operationally closed; a strictly autonomous system is logically closed, it contains all what it needs to exist.

The last step is now to interpret self-reference and to relate it to some feature of real life living organisms. As we have shown, in our perspective (and in Varela's and others') autopoiesis is the logic of life, or more generally of viability. The different living organisms, depending on their complexity, have varying degrees of self-reference. The more self-referential they are, the more autonomous they are, and therefore able to face the challenges of their environment. Furthermore, as we mentioned above, the more self-referential they are the better the compatibility between their actual state and their "image", which can also be described as their knowledge of themselves. We propose that the degree of overlap between the object and the image in the existential plane (fig.3.) corresponds to the degree of self-knowledge, or more precisely to the degree of existentiality of self-knowledge, which is what consciousness is. Consciousness is the experience of being; its nature is not material, nor logical, it is existential.

In summary, consciousness needs an organism (an object) to manifest itself; it needs more precisely a very complex and self-referential organ. High self-reference means that the network of laws (the image) that drives its changes in time must be very closely connected to the structural configuration of the components. The brain is obviously a good candidate for complexity. Even if more studies should be done, we can assume that the brain is a very operationally closed system: its state at time t2 is more dependent on its own state at time t1 (t2>t1) than on Newton's law or Maxwell's equations (although they are certainly respected!). Thirdly, our metamodel shows that consciousness cannot be reduced to the brain or to the "mind", if we use this word to refer to the immaterial relational organization of the brain. It certainly needs both carriers, but its nature is essentially existential. Much ingenuity should therefore been shown by those who want to "see" consciousness in the framework of the neuro-sciences alone.

In the last section we will draw some conclusions from this work about the initial question: will machines ever think ?

5. Conclusive Remarks about Machines and Men

Using the framework and the concepts we have just described, we will compare the structure and organization of an inorganic system, like a planet, a watch or a digital

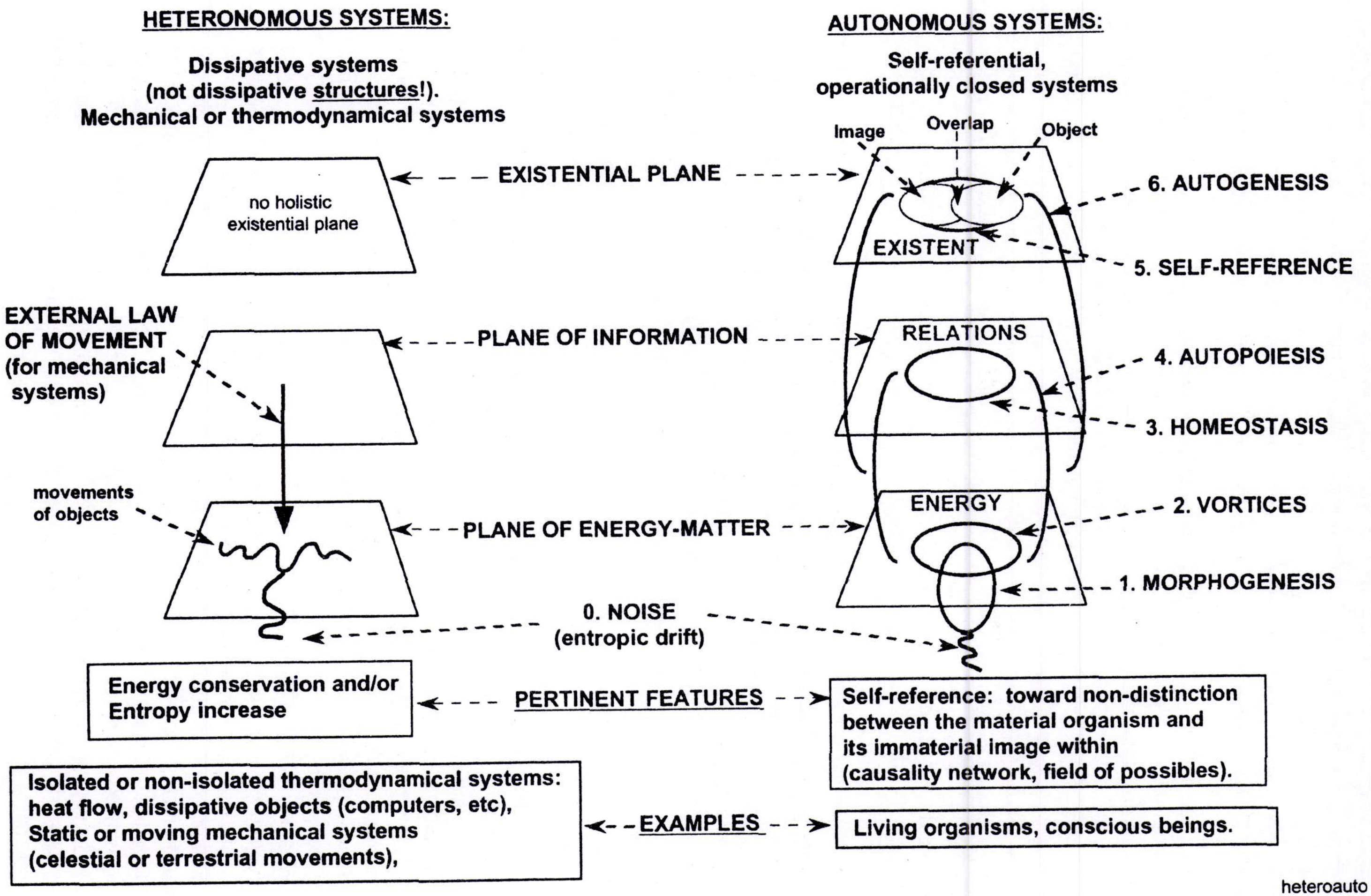

fig.3. Comparison between a simple heteronomous system *and a complex autonomous* system.
Or: the fundamental difference of nature between a physical object or machine ruled mainly by the general external laws of physics or chemistry and an living and cognitive organism ruled mainly by specific immanent laws resulting from historical evolution.

computer, and of a living – and thinking conscious – organism, like an animal or a human being, with the help of fig.3. Due to the limited space available here and in order to make our message as clear as possible, we have to simplify the situations presented and cannot go into the details of intermediate cases. On the left side of fig.3., we have applied our scheme for viable systems to the ordinary simple inorganic objects or systems that are studied in physics or chemistry like static or moving mechanical, thermodynamical, or electromagnetic systems. On the right side we have applied our scheme to living and cognitive systems. We have chosen two extreme cases to make the point more clear. The purpose of this comparison is to draw the attention on the deep structural and organizational differences between a nonliving system and a living system as shown by our holistic metamodel.

The main characteristic of the nonliving systems is that they are heteronomous: they follow the usual general invariant laws of the basic sciences, physics, chemistry, themodynamics, etc.; the laws are imposed to them from the outside, their own movements do not change the laws. This feature is represented graphically on the right side of fig.4. by the simple top-down static arrow that symbolizes the external laws followed by the system.

On the other side, living systems are (partially) autonomous: they contain within themselves, in their organization, the essential part of the laws that drive their dynamics. Furthermore, these laws are such that they insure the viability of the system: the laws re-produce the structures and processes that produced them; the system is therefore self-producing. Autopoiesis includes and relies on the lower level loops: morphogenesis, vortices and cybernetical regulations. In addition autopoiesis generates itself two upper level holistic loops: self-reference, that we interpret as self-knowledge, which is consciousness, and autogenesis that leads to autonomy.

Our holistic metamodel allows us to interpret the main stages of the 3 billion years of evolution of life on this planet. The successive turning on of the six cycles gave rise to the processes of self-organization (Prigogine's dissipative structures), matter recycling, self-regulation, self-production, self-reference, and autonomization. With self-production – and the earlier cycles of self-organization, recycling, and self-regulation – appeared life and thinking, seen as the communication between the organism and its environment in the purpose of surviving. With the increase of self-reference, this communication process was later accompanied with the development of consciousness, the organism's increasing knowledge of its own state. The most recent and complex species, like mammals and modern human beings are the heirs of this 3 billion years evolution.

What about digital computers ?

As can be seen on fig.3., in our interpretation computers are simple dissipative objects (not dissipative structures !). Their natural evolution is driven on the long run by the second principle of thermodynamics – the increase of entropy: dissipation of electric energy into heat, wearing of the components, etc. They have no manifestation of spontaneous circularity (operational closure), like feedbacks, self-organization, autopoiesis, self-reference. They are nevertheless able to compute thanks to the ingenuity of engineers and programmers who have placed artificial temporary obstacles to the entropic drift. By manipulating electrical currents, computers simulate computation (in the sense of symbol manipulation), and by computing they simulate thinking.

In conclusion of the present analysis, today's digital computers can be said to be able to think only within the cognitivist paradigm where thinking is precisely synonym of computing or manipulating symbols. But within the much broader context of our holistic systemic framework, developed to interpret the specificities of complex autonomous systems, digital computers are only simple dissipative objects like any natural inorganic system close to thermodynamical equilibrium. Their ability to temporarily process electrical currents is used to simulate the manipulation of symbols and therefore to perform calculations. But there is no track of circularity, of autonomy, of holistic aspect. Their capacity to compute and their physical behavior are totally disconnected, whereas the basic characteristic of living systems is the fundamental (not to say ontological) integration of the physical processes and the corresponding logic, of which autopoiesis is the most brilliant example. The absence of this integrative level (level symbolized by the existential plane in our model) means that digital computers have no holistic feature, like self-reference, and therefore no track of consciousness. In that sense, digital computers do not think in the way human beings think: consciously.

Will computers ever think ? If we mean think consciously like human beings, our metamodel indicates that such devices should be able to self-organize their structures, to recycle matter, to self-regulate, to self-produce, and, above all, to be self-referential.

Much remains to be done...

References

Churchland, Paul M. and Patricia Smith, 1990. Could a Machine Think ?
Scientific American, January 1990.

Rosen, R., 1979, Anticipatory Systems in Retrospect and Prospect.
General Systems Yearbook 24, 11. Society for General Systems Research.

Schwarz, Eric, 1997a. Toward a Holistic Cybernetics. From Science through Epistemology to Being. Cybernetics and Human Knowing, Vol.4. No 1. p. 17-49.

Schwarz, Eric, 1997b. The Evolution of Anticipation. A Systemic View.
Casys 1997. International Journal of Computing Anticipatory System. Vol 2. p. 88-101.

Varela Francisco, 1989. Connaître les sciences cognitives. Seuil, Paris

Zeleny, M. ed., 1981. Autopoiesis. A Theory of Living Organization. North Holland, New York.