

EPISTEMIC IMPLICATIONS OF TWO BIOLOGICAL CONCEPTS

*Srdjan Lelas**Introduction*

Traditional foundationist epistemology and its modern successor, philosophy of science, with their ambition to ground knowledge before cognition and to justify science without presupposing it, have failed to provide us even with a modest comprehension of the totality of their subject. After the failure became evident in the late fifties we have witnessed several “turns”. “Historicist” and “sociologist” turns have pointed out that cognition and science are primarily historical and social phenomena, and that consequently epistemology should be replaced by history and sociology of cognition and science. Recently Rorty (Rorty 1980, 1982) advocated something that, keeping the same style, might be named the “literary-criticist turn”. He emphasizes, like the celebrated “linguistic turn”, that science express itself in writing and, therefore, it is essentially a literary phenomenon which should then be studied and interpreted by the hermeneutic method.

All these turns suffer from two grave shortcomings : they are external to the best possible knowledge we presently have — science; and they are incorrigibly incomplete. As to the first drawback Quine (Quine 1968) has proposed a new setting for epistemology. He has harbored it inside the body of natural science as a chapter of psychology. However, that appears to be just a modernized behavioristic version of something already experienced and abandoned. Similarly, the ideas that cognition and science are, besides their other qualifications, natural phenomena, that their development ought to be treated as natural process in the sense that “at no stage has there been any transfusion of knowledge from the outside (of nature), nor of mechanisms of knowing; nor of fundamental certainties” (Campbell, 1974, 413), and that epistemology should be “compa-

tible with the description of the world provided by contemporary science" (ibid.) got a cold reception among philosophers of science. According to a modest naturalism such as the one advocated by Campbell ("modest" in the sense that from the beginning it rejects any form of reductionism), epistemology and natural science are intermingled in such a way that natural science comprises epistemology as the part whose subject matter is the natural phenomenon called "cognition" or "science", while epistemology embraces natural science as a form of cognition whose character and value it seeks to understand. The apparent circle underlying this approach is, in spite of all neglect, at least an interesting alternative to the hermeneutic one.

As to the second objection to the fashionable approaches, what modern naturalism in its anti-reductionist vein has to offer so far is an encyclopedia of all the disciplines dealing with cognition or science, together with some hope that a unifying schema, like Campbell's "blind variation and selective retention", could possibly be found. After all, if cognition and science could be seen as natural phenomena, why not try the same with history and society? In the present situation, where historicist, sociologist, and hermeneutic approaches are tending towards extreme cultural relativism, and where scientific realism is not able to accommodate radical historical change, I think the modest naturalist option described above is worth investigating.

An implicit premise of contemporary naturalism is that cognition as natural phenomenon is tightly bound to living systems, or more strongly that cognition is the very nature of living systems. In other words, it is supposed that something similar to cognition could eventually be found outside the animate world only in man-made systems. The first aim of this paper is to examine the validity of the premise, and if it shows to be sound, to look closer for the reasons that make it valuable.

This issue becomes important when conjoined with another implicit premise of the naturalistic approach. Since no transmission of any cognitive element or process from outside nature is allowed, whatever is valid for all living systems as cognitive systems in general must also be valid for humans as cognitive systems in particular. In more cautious phrasing the premise requires that the features of human cognitive system must be compatible with general features of living systems as cognitive systems. Taking this premise for granted I will then consider possible epistemic implications of

different views on the nature of living systems.

The central concept of the biology of cognition is the concept of a nervous system. That a nervous system is *per se* a cognitive system seems obvious. That looks like a true claim even if a nervous system has to be considered as a part of an organism which is itself a cognitive system. How precisely the role of the nervous system and its cognitive function should be described is the most debatable issue. I will take it together with some of its epistemic implications in the second part of the paper.

THE LIVING SYSTEM

One non-standard view

Recently the neurophysiologists H.R. Maturana and F.J. Varela in the book entitled *Autopoiesis and Cognition* (1980), undertook a very similar endeavour, similar both in naturalistic style of thought and in focussing on the same two basic biological concepts. Since they have tried from the very beginning to find a unified treatment of the two systems so that their cognitive nature will be the same, they were forced to take an unorthodox position, particularly in regard to the concept of living system. Although expressed in rather difficult style, their position is well articulated and not too far from mine. So in order to save time and space I will try to accomplish my task by discussing their results.¹

From the outset the approach of Maturana and Varela is philosophical, or one may prefer to say theoretical. Throughout the book the distinction has been maintained between characterization of the system in terms of relations and organization of its components and description or realization of the system in particular components endowed with specific properties. It resembles very much the noumena/phenomena distinction, but also it is the distinction between the whole and its parts as well as between essence and its realizations. Related to this distinction there is an epistemic point. An observer, most obviously a human being, has a special privilege not given to the living system being observed. He always perceives a living system in a medium, and interacts independently with each of them. Furthermore the domain of interaction of the observer does not necessarily overlap completely with the domain of interaction of the living system, hence his position is not at all the position of the observed system. According to Maturana and Varela this is inevitable

at the level of description but ought to be taken into account at the level of characterization.

What characterizes living systems, according to Maturana and Varela is their organization, i.e. a set of relations that determines the dynamics of interactions and transformation the systems may undergo while maintaining their identities as living systems. The organization of living is

“a circular organization which secures the production or maintenance of the components that specify it in such a manner that the product of their functioning is the very same organization that produces them. Accordingly, a living system is an homeostatic system whose homeostatic organization has its own organization as the variable that it maintains constant through the production and functioning of the components that specify it” (Maturana and Varela 1980, p. 48).

In another phrasing living systems are autopoietic, self-made systems; “they transform matter into themselves in a manner such that the product of their operation is their own organization” (82). An autopoietic system “continuously generates and specifies its own organization through its operation as a system of production of its own components, and does this in an endless turnover of components under conditions of continuous perturbations and compensation of perturbation” (79).

Autopoiesis, as the central concept which embraces all different aspects of living systems, is opposed to a traditional concepts like adaptation and a more fashionable one like genetic program, which are alternative candidates for supposedly the same role in theoretical biology. But the source of unorthodoxy does not lie here. It is in the following line of further theoretical development.

Crucial in the concept of autopoietic system, according to Maturana and Varela, is the homeostatic nature of the system with its own organization as the variable which is kept constant. This means that the system subordinates all changes to the maintenance of its own organization and is, therefore, its own reference. Autopoietic systems are, therefore, self-referred systems in the sense that they can be characterized only with reference to themselves. This makes them also autonomous systems, their autonomy being “continuously revealed in the self-asserting capacity of living systems to maintain their identity through the active compensation of

deformations" (73).

The contrast to self-referred systems are allo-referred systems that can only be characterized with reference to a context or environment. Accordingly, in order to characterize an autopoietic system we do not need to refer to, or even to presuppose, any ambience whatsoever. Maturana suggests that we had better stop "looking at living systems as open systems defined in an environment" (xiii). We are advised to attain just the opposite view; a living system is "like the nervous system, a stable state-determined and strictly deterministic system closed on itself and modulated by interaction" (50).

In terms of their functional organization, living systems are, according to Maturana, closed systems not only in the sense of self-reference but also in the sense of "not having input and output." The role of effector surfaces is not to act upon an environment but to maintain constant the set of states of the receptor surfaces. And Maturana emphasizes that "a grasp of this is fundamental for the understanding of the organization of living systems" (51).

At this point it becomes obvious that external independent events can only disturb living systems, and that their behaviour is nothing else than sets of internal unessential structural changes which compensate these perturbations. In other words, all feedbacks are internal to them.

When one closes off conceptually or otherwise the living system, some side effects are inevitable. For example energetic and thermodynamic considerations become superfluous, and the notions like program, coding, and information do not even "enter in the realization of a concrete autopoietic system because they do not refer to actual processes in it" (90). All those considerations and notions belong to "the domain of description" of an observer and should be ruled out from proper characterization of the living system.

So much unorthodoxy. Now some intrinsic problems. First, how to introduce the environment into the scheme in which the living system functionally does not need it? The idea is to create the need by moving from the space of characterization (conceptual space !?) to physical space deriving the necessity of the latter. This is done in the following way: "from the mere fact that a physical autopoietic system is a dynamic system, realized through relations of productions of components that imply concrete physical interactions and transformations, it is a necessary consequence of the auto-

poietic organization of a living system that its ontogeny should take place in the physical space" (98). First, Maturana and Varela speak here of "a physical autopoietic system" not of conceptual one, so we are already in physical space. Furthermore, it is a typical argument from "is" to "ought to be". From something that is so thoroughly closed it is not possible to derive openness. Closed systems can tolerate environment, they do not need it, hence one can introduce the ambience in such theoretical schema only by God's *fiat*.

Once we are in the physical space, we are in a quite different situation. There living systems exist in an ambience and they become "units of interaction". "From a purely biological point of view", says Maturana, "they cannot be understood independently of that part of the ambience with which they interact : the niche; nor can the niche be defined independently of the living system that specify it" (9). However, even here "interaction" rather means "disturbance". "In the history of interactions of a composite unity in its medium, both unity and medium operate in each interaction as independent systems that, by triggering in each other a structural change, select in each other a structural change" (xx). For the living system, the existence of an environment is still a pure accident.

The next intrinsic problem is, of course, cognition. How an autonomous self-sufficient system can give rise to cognition; is it again a pure accident, or is there still something at least in the realization of the living system that makes its appearance a necessity? To illustrate the difficulty we are facing here and the internal necessity for different perspective let me take a somewhat longer quotation from Maturana.

"Living systems as units of interactions specified by their condition of being living systems cannot enter into interactions that are not specified by their organization. The circularity of their organization continuously brings them back to the same internal state (same with respect to the cyclic process). Each internal state *requires* that certain conditions (*interactions with the environment*) be satisfied in order to proceed to the next state. Thus, the circular organization implies the prediction that an interaction that took place once will take place again. If this does not happen, the system desintegrates; if the predicted interaction does take place, the system maintains its integrity and enters into a new prediction. In a continuously

changing environment these predictions can only be successful if the environment does not change in that which is predicted. Accordingly, the predictions implied in the organization of the living system are not prediction of particular events, but of class of interactions... This makes living systems *inferential* systems, and their domain of interaction a *cognitive domain*" (10). (Italics mine).

First we should note here that the circular organization is now surprisingly a cyclic process which does not bring the system back to the same internal state by homeostatic compensation of perturbation, but by the process that requires an interaction with the environment. This interaction might still be an event that only triggers an internal development but now it is a necessary one. Moreover, this external event must happen not only once but all over again in a regular way. The environment no more can be something accidental, since the repetition of its disturbing actions is absolute necessity for the maintainance of integrity.

Second, it must not escape our attention that Maturana, when defining the living system as an inferential (cognitive) system, does not refer to the realization of the system in physical space or to the description of its components. The definition fully remains in the space of characterization of the system. The same can be said for the following definition of the cognitive system.

"A cognitive system is a system whose organization defines a domain of interaction *in* which it can *act* with relevance to the maintainance of itself, and the process of cognition is *the actual* (inductive) *acting* or behaving *in* this domain. Living systems are cognitive systems, and *living* as a process is a process of cognition" (13). (Italics mine).

Although the term "acting in" still can be interpreted as "actions that compensate perturbation", inferential or inductive acting requires that perturbation necessarily repeats itself. Moreover, the perturbation must happen again not because it is the feature of the environment as deforming agent to re-act again and again, but because the living system needs an external agent to recursively trigger some processes in itself. This is an unavoidable move, for otherwise if the cognitive system was characterized only as a homeostatic system which compensates the external perturbations, any

drop that changes its shape according to the conditions in the environment would be a cognitive system.

Maturana's approach is, therefore, inconsistent. To make his biology of cognition coherent one has to make the choice either to maintain the interpretation of autopoietic system as essentially closed system and to give up ambition to demonstrate that "cognition and operation of living system are the same thing", or to reinterpret the notion of autopoietic system in such a way as to include in it "the necessary material openness of the system as it realizes the physical space". If the first option is taken and cognition placed together with other concepts in the domain of the observer, there is still the paradox that remains, the paradox which Maturana himself recognizes. "Living systems in general, and their nervous system in particular, are not made to handle a medium, although it has been through the evolution of their handling of their medium that they have become what they are" (56). With the second option I will deal in the next section.

There is an important epistemic ingredient in biology of cognition that Maturana clearly pointed out and that does not depend on the closed nature of living systems and the accorded interpretation of autopoiesis so strongly defended by him. The domain of interaction of a living system as the set of all interaction it may undergo without loss of autopoiesis is obviously determined by the particular mode through which its autopoiesis is realized. Since the domain of interaction is the upper limit of the cognitive domain of a living system, different living systems with different modes of autopoiesis have different cognitive domains and likely also different modes of cognition. From this Maturana draws the inescapable conclusion that "intrinsically, then, no absolute knowledge is possible, and the validation of all possible relative knowledge is attained through successful autopoiesis" (119). More about that in the conclusion of the paper.

The Standard View

Reinterpretation of autopoiesis along the line suggested by the meaning of the term, which is first of all self-production, brings us to the standard "purely biological point of view". It sees a living system as a discrete specific material system with definite boundary necessarily dependent on the continuous exchange of matter and energy with its environment, and capable to control this exchange

in order to maintain, over a certain period of time, its specific composition and ability for the exchange and reproduction. In it all the meaning of autopoiesis as circular organization is retained. The sole addition is "material openness", i.e. exchange of energy and matter with the environment. Could we even conceive a self-made system without this exchange ?

In physical space, Maturana does admit, autopoiesis is realized through peculiar chemical composition and organization of living which enables the living system to collect free energy from the environment bit by bit, to slow down its dissipation by storing it in specific organic compounds and to use it in a controlled way for creation and maintenance of its biological identity, i.e. of that peculiar chemical composition and organization which makes this performance possible. Endergonic chemical reactions, highly ordered structures needed to implement them, concentration of substances rich of free energy, controlled behaviour, all that points, however, to thermodynamically unprivileged highly improbable direction. How then is autopoiesis possible at all ?

There is a standard answer which, although unique, embraces three interdependent components : closure in space and time, openness to the environment, and selectivity of the interaction with the environment. These three aspects of a living system that make it capable to confront the Second Law of thermodynamics and are preconditions for successful autopoiesis, are quite general. They are not accidentally related to any particular feature of the components that realize autopoiesis. Therefore, they pertain to the "space" of characterization of living systems.

An autopoietic organization requires definite boundary, because the high concentration of energy, specific substances and highly ordered processes indispensable for self-production are possible only in "the bag", i.e. if a sharp separation and a clear distinction between "inside" and "outside", between that what belongs to the living system and that what belongs to the environment, are established. The living system then necessarily occupies a limited region of space well separated and closed off from the environment. If by the play of pure chance a high concentration is established in the medium and the boundary condition is not satisfied the "system" immediately starts to disintegrate due to the Second Law. So not only for an external observer but for the living system itself the world is inevitably split into two parts, the internal and the external one.

However, the closure of physical space alone cannot protect

the living system against the Second Law because the law operates inside as well as outside a closed space. What is needed is the constant renewal of the components. The living system cannot be just a homeo-*static* system; it must be a system with indefatigable dynamics. Dynamics of what? Of many things, of course, but first of all of matter and energy exchanged with the environment. What really circumvents the Second Law is the *selective* flow of matter and energy through the boundary. Again, this dynamics is not just an accidental way of realization of autopoiesis, it is a necessary precondition for it.

Selective openness, which implies handling of the environment, does not stand in opposition to circular organization and self-reference. As open systems, living systems can be seen as embodiments of their own labour done within an environment and on the materials taken from the environment; their existence is the result of their own work. That work is neither the work of the environment that shapes the system through its adaptation, nor it is the work done for the environment in order to become adapted, it is not even done *in* the environment and embodied there as in human artefacts. It is the work done by the living system, with material and energy supplied by the environment, and incorporated in the same system that performs the work. Circular organization and self-reference do not exclude openness. On the contrary, they presuppose the environment and interaction with it, they even demand them.

Before we come closer to cognition by considering the way selectivity is realized, let us point out that autopoiesis reinterpreted along this line can easily accommodate and even give the basis for notions like program, reproduction, and death. When autopoiesis is seen on the background of the Second Law, a living system appears as only *transient* exception due to the probabilistic character of the law and to the program that sets the closure-openness condition. Transient nature of the living systems is just the reflection of the fact that the Second Law as statistical law can be violated but only in limited space and for limited time. Reproduction secures its violation in another limited space and for another limited time. The repetition of violation, of course, being insured by the transfer of the program from one living system to another.

A living system, then, exists only because it is an open system, in other words, because it is essentially an incomplete, insufficient, unfinished, *needy* system. It completes itself through the environment. The circle of organization passes through the ambience and

self-reference presupposes the reference to something else through which it comes back to itself, refers again to itself.

It is this dramatic and fatal need of the living system for interaction with the environment that makes it a cognitive system, as will be shown shortly. The same need makes it the unit of *interaction* and specifies the domain of interaction or cognition.

Openness, however, should be partial, or better to say selective, for obvious reasons. Unselected flow of matter and energy through definite region of space evidently does not constitute the living system. Clear distinction between input and output is needed with some specific substances passing into the system and other specific substances leaving the system. The boundary then performs selection and through it unites two apparently conflicting requirements, for closure and for openness.

Selective openness or selective closure has to integrate two sorts of demands. The demands of one sort refer to the mode of autopoiesis of the living being. They express its biological identity. The other sort of demands is related to the forms in which matter and energy needed for autopoiesis are present in the environment. How does the living system realize this *double reference*? There are two answers: selective semi-permeability of biological membranes, and purposive behavior. The former we shall take now, the latter will be left for the next section, because it is the most conspicuous in the organisms with the nervous system.

The "back-bone" of the membrane is the hydrophobic fatty bilayer which repels charged molecules usually dissolved in water. The active transport of molecules through membrane is provided by asymmetric molecules of specific so-called "transport proteins" which are embedded in the fatty bilayer. In general terms their role could be compared with the role of Maxwell's demon. Like the demon they "open up" for these molecules rich in free energy or needed as the building blocks for autopoiesis, and "close off" for those that are harmful or superfluous. By performing this selection an improbable distribution of matter and energy is established in physical space; thus the first condition for successful autopoiesis is satisfied. The essence of this process is the "recognition" of possible entries or the processing of information in which information carriers and metabolic ingredients are not separated.

How the selection and the transport is done precisely is the most active field of study in modern biology, and we certainly may expect many interesting discoveries to come which will enable us

to understand transport phenomena in purely physical and chemical terms, but their thermodynamical and biological role will be the same. In order to describe that role one might be reluctant to use terms like "recognition" or "information processing", not to mention "cognition". However, selective openness or interaction as materialized in the active transport through biological membranes is endowed with some elements that look essential for something to be called the "cognitive process".

First of all we should repeat that the membrane divides the whole space into two sharply distinguished parts: the internal space of the living system and the external space of its environment. This distinction can be understood as the biological basis or the general framework in which the "fundamental finding of our thought existence", as Jaspers praised subject/object division, finds its place. Furthermore, the set of asymmetric molecules of transport proteins that pump into the living system the part of the environment externalizes and materializes the living system's need for the external world, the life-preserving reference to something outside the system. Finally, the specific composition of transport proteins makes the input and output highly selective, or in other words, it enables the living system to "recognize" the relevant part of the environment. "Recognize" is the proper word here, for in spite of the fact that the living system specifies its niche by the mode of its autopoiesis, it is the form of matter and energy available in the environment to which those molecules refer. Therefore, by refusing to use terms like "recognition", "information processing" or other similar words, one misses the essential aspect of autopoiesis and the only natural place where cognition can be introduced in nature.

If used in this context the terms "recognition", "information processing", and "cognition" obviously receive very broad, somewhat vague, highly unspecific, but still, I think, proper meaning and legitimate usage. The same goes for the term "representation" which I would like to introduce now. The type and distribution of the molecules of transport protein and the mechanism of transport determined by their structure are related to the type, distribution, and the structure of the molecules that have to be pumped in. Hence we may say not only that membrane structure *refers* to part of the environment, but also that it makes that part of environment in a sense *present again* in the membrane structure. It makes the relevant part of the ambience *re-present* not accidentally but in order to be recognized and *internalized*. Seen from the inside of the

living system the membrane structure represents the part of the environment which partakes of autopoiesis.

Remaining still in this general and vague vein we can say a bit more about the so frequently used term "cognition". In Maturana's interpretation of the living system as cognitive system, "cognition" means inference, expectation that what happens once will happen again. This certainly is not what we usually understand when we speak of inference. That living systems presuppose certain regularities, like repeated presence of certain forms of matter and energy in the environment, and that they make use of them is both obvious and trivial. It is also obvious that living systems can only deal with the classes of events or molecules, because they must operate above the level of thermal chaos. However, the proper issue for epistemology is how they do that, how living systems generate "inductive" behaviour and how they recognize classes.

Assuming that the term "representation" is accepted in this context we might give another less operational meaning to the term by saying that "cognition" is the re-presence of one system in another, or more precisely cognition is the state or process through which an autopoietic system makes a part of its ambience present again in itself. Obvious prerequisites for it are selective openness/closure, and interaction between the system and its environment. Whether the terms "representation" and "cognition" could be made more specific at the invertebrate level remains to be seen on the basis of future biological research and its conceptual analysis. What we can expect when the nervous system is present we shall try to find out in the next section.

The Nervous System

If some reservations regarding the application of concepts like "cognitive system", "information processing", "representation", and "recognition" to the simple unicellular organisms still remains, it seems as though we confront quite another situation when dealing with organisms having a nervous system. That the nervous system produces behavior through representational processes of a particular kind is a commonly held thesis. But when neurophysiologists like Maturana report that in their scientific practice they realized that "the mapping of the external world was an inadequate approach" (xv) to the functioning of the nervous system, and that the new approach required them "to treat seriously the activity of the

nervous system as determined by the nervous system itself, and not by the external world" (ibid), then one at least should be careful with any straightforward application of the aforementioned concepts. Let us look again at Maturana's arguments.

The first, and I think the main argument against any simple representational thesis is the discovery that the "languages" in which the distinctions among things and events are specified in the external world and in the nervous system are so different that any concept of mapping seems quite inappropriate. According to Maturana, the language of the nervous system is the language of geometric relations, space distribution, and afferent influences in which nevertheless even something so ungeometric as chromatic differences are expressed. The second argument relies on the finding that the generated behavior is the result of the activity of the nervous system as a whole, so that it is very difficult, if not impossible, to locate particular regions exclusively responsible for particular behavior. Namely, the nervous system is, according to Maturana and Varela "a closed network of interacting neurons such that a change of activity in a neuron always leads to a change of activity in other neurons" (127).

These arguments led Maturana again to his favorable conclusion that "one had to close off the nervous system to account for its operation". He found an additional support for this conclusion in the fact that "the objects that the animal sees are determined not by the quantity of light observed, but by the relations holding between the receptor-induced states of activity within retina, in a manner determined by the connectivity of its various types of cells" (21). Stated more generally "when any particular interaction takes place at the level of the sensors, the relations accessible to the nervous system are given at this level in a certain state of relative activity of the sensing elements not in the state of activity of any particular one" (22).

The argument goes further. When one passes the boundary of sensors the nervous system as closed network appears to begin at any point. "The closed nature of the functional organization of the nervous system (open only to modulations through interactions) is particularly evident in ... the subordination of conduct to the correlation of activity between the receptor and effector surfaces". This means "that the 'visual handling' of an environment is no handling of an environment, but establishment of a set of correlations between effector (muscular) and receptor (proprioceptor

and visual) surfaces, such that a particular state in the receptor surfaces may cause a particular state in the effector surfaces that brings forth a new state in the receptor surfaces... and so on" (26).

The argument ends up in a very strong conclusion that the nervous system neither has an input or output, nor the outside or inside exist for it.

What is puzzling in this whole picture is that the nervous system appears to be completely self-sufficient; so much so that not only its interactions with the external world of the organism seem to be just contingent perturbations, but its relation to the organism itself are contingent as well. The picture suggests that the nervous system stimulates itself through the actions of effector surface caused by the state of receptor surface and so on. The effect is that the distinction between receptor and effector surfaces, which Maturana maintains throughout the book, loses sense.

Besides being puzzling in itself, the picture contradicts the description of the anatomy and functioning of a neuron and of the architecture of the nervous system as well. In most types of neurons one can clearly differentiate collector area, distributive element, and effector area with the impulses going from the former to the latter. Furthermore, Maturana strongly emphasizes that the whole architecture of the nervous system is "subordinated to the order of the sensory and effector surfaces." "This subordination", he continues, "has two aspects: (i) the receptor and effector surfaces project to the central nervous system retaining their proper topological relations; (ii) the topological relations specified by the receptor and effector surfaces in their projection constitute the basis for all the architectural order of the central nervous system" (20).

If these claims about the architecture of the central nervous system are correct, than instead of the "uniform configuration" of a closed interacting network one ought to speak about essential polarity between receptor and effector surfaces and about anisotropy caused by the fact that inside the nervous system there is a privileged direction, i.e. from collector to effector area and from receptor to effector surface. In that case to close off the process, i.e. to secure that the change of effector surface does cause the change of sensory surface one has to step out of the nervous system and to pass through the environment. As with the living system, it appears again that the nervous system operates as an incomplete, as an open system that needs an environment to close itself off functionally. It is the action in the environment generated by the nervous system

that completes its function. This is the proper reason why Maturana may say that it is not a neuron or a group of neurons which is a functional unit of the nervous system, but the conduct produced by it.

The ways one needs the environment to properly characterize functioning of the organism and the nervous system are not of the same type. The functioning of the nervous system is subordinated to the mode of autopoiesis of the living system, and the nervous system performs a rather special role in the maintenance of autopoiesis. In reference to this special role there is a sense in which the nervous system is closed system working in the way described by Maturana.

In whatever way one describes the role of the nervous system it certainly cannot be perceived as the organ directly engaged in metabolic processes responsible for the maintenance of autopoiesis. Therefore, there is no need for the nervous system to be open for all those forms of matter and energy necessary for production and reproduction of the components (the nervous system included). Its input at the sensory surfaces has a special task — to trigger generation of certain behavior. The substances, momenta or energy forms affecting sensory surface need not to be taken into the organism and incorporated into its structures, they do not serve as the raw material for autopoiesis. This provides the basis for the claim that the nervous system is a closed network *modulated* by the interaction with the environment.

Let me make at this point one general remark about the way receptors operate, a remark extensively exploited by Campbell (1974) but not mentioned by Maturana. Firstly, because of specialization of inputs it is unnecessary to draw into the system almost any amount of substances or energy, so receptors are usually adjusted to react to a very small quantity of triggering agent. Secondly, in order to generate, together with effectors and the nervous system when present, purposive behavior subordinated to maintenance of autopoiesis receptors have to bring the organism above the level of random thermal movement dominant at the macroscopic level. Purposive behavior can be synthesized only over microscopic fluctuations and must be protected against the effects of individual microscopic particles. There is, therefore, an upper and a lower threshold of sensitivity of any receptor. The classes of substances or energy form that serve as triggering agents are, of course, determined by the structure of receptors that is subordinated to the mode of autopoiesis of the living system.

Now for our considerations, there is one very important aspect or even consequence of this limited range of sensitivity of receptors and of the closure (in the sense described) of the nervous system. Local events at the sensory surfaces, although immediate stimuli for the receptors, have no importance in themselves for the conduct which is generated. They stand for something else, serve as vicars (Campbell, 1974) signaling something which is really important for maintaining autopoiesis.

The example of paramecium is a paradigmatic one. With its chemoreceptor paramecium is able to differentiate nourishing from noxious substances. It does so on a small sample, for the entrance of large quantity of a noxious substance will kill the organism. But, if detected, the small harmless amount of poison will be taken as the vicar of a larger potentially harmful quantity and that will trigger the process resulting in the quick motion away from the place where the sample is taken.

The same principle of substitution of one thing for another or, if we want the word, the representation of one thing in another, one can find in the visually oriented behavior of the organism in the unhomogeneous environment (Campbell, *ibid.*). The eye as the part of the system that generates and conducts locomotion exploits one favorable circumstance in the unhomogenous environment, namely the coincidence between the regions transparent for visible light and the region penetrable for the organism. Transparency is the substitute for or representative of penetrability. This principle is quite general. It is the principle on which all receptors work.

The very nature of substitution or representation, where something having its own properties stands for something else, implies that the substitute or representative cannot be a perfect copy of the represented. There are always exceptions to the rule, marginal undecidable cases, false judgements etc. like glass and fog in the case of visually guided locomotion. Those unavoidable imperfections make the world of receptors a *phenomenal* world, the world where illusions are possible, because no direct contact with the substituted or represented things or events can be established in it.

On the effector side, however, the situation is different. Effectors must deal with the real things or processes going on in the environment, otherwise the conduct executed by them will not contribute to the maintenance of autopoiesis, the interaction will be out off, the organism will disintegrate. In a somewhat free philosophical spirit one might say that the effectors operate in the

noumenal world.

Coming back to the role of the nervous system as a part of the cognitive system or as a cognitive system in itself, we should be reminded that the fundamental schema of receptor — processor — effector is preserved throughout the nervous system from the nerve cells, through the architecture of the system, even to the neo-cortex which, according to Maturana, appears as “the center of internal anatomical projection” of the nervous system onto itself. Being a functionally incomplete system, the nervous system is best described as a system of *translations* from the “language” of collector areas to the “language” of effector areas, from the “language” of receptors to the “language” of effectors, and finally from the “language” of vicars or phenomena to the “language” of conduct or noumena. This is compatible with Maturana’s claim that the nervous system is the organ for the synthesis of behavior, not for the representation of the world.

But let us have a closer look and see why Maturana refuses the representational model of the functioning of the nervous system, if he really does reject it.

Due to the geometry of the collector area and to the fact that afferent influences coming from afferent cells, neuronal or others, do not superimpose linearly, the neuron, Maturana points out, reacts with a definite transfer function to relations, or better to say, to a configuration of afferent influences (precisely to *classes* of these influences), not to a singular one. Furthermore, the response of the neuron depends essentially on its own previous state, i.e. the state immediately after the previous response, and that is not necessarily always the same state.

Similarly, it is the configuration of relative states of activated neurons at the receptor surfaces which is translated to another configuration of relative neuronal states at the effector surfaces through the network of interacting neurons. Hence, says Maturana, for the nervous system “any interaction (presumably with the environment) is *represented in* the nervous system by the sequence of states of relative neuronal activity leading to the conduct which it generates” (18) (*italics mine*).

Because of relational and state-determined character of the functioning of the nervous system the participation of the particular neuron or group of neurons in the synthesis of behavior is not fixed. We have here one-many and many-many-one correspondences, i.e. the same conduct generated by

different neurons, and vice versa. "Thus, under no circumstance is it possible to associate the activity of any particular cell with any particular interaction of the living system" (21). Some localizations can be done, but only as "the areas where certain modalities of interactions converge, and not of localization of faculties or functions" (22).

Nevertheless, "if an external interaction takes place, the state of activity of the nervous system is modified by the change of relative activity of the neurons, which in close association with the sensory elements *embody* the relations given in the interaction" (23) (*italics mine*). And, moreover, "these relations are not those that the observer can describe as holding between component properties of the entity in his cognitive domain; there are relations generated in the interaction itself and depend on both the structural organization of the organism *and* the properties of the universe that match the domain of interaction that this organization defines" (23) (*italics mine*).

What remains constant is the flux of relative neuronal activity is the *connectivity* that defines the neuronal network. Here again it is not a particular connectivity which is constant, but the class of connectivity, whatever that might be, and the general pattern of organization of the nervous system. I assume that it is these somewhat undefined classes of connectivity that determine "the domain of possible states" of the nervous system in such a way that "the domain of possible states of the nervous system continuously become *commensurate* with the domain of the possible states of the ambience" (131) (*italics mine*).

However, "the continuous correspondence between conduct and ambience revealed during ontogeny is the result of the homeostatic nature of the autopoietic organization, and not of the existence of any representation of the ambience in it; nor is it at all necessary that the autopoietic system should obtain or develop such representation to persist in a changing ambience" (99). So we have explicit denial of the usefulness of concepts like "representation" for the description of the functioning of the nervous system and implicit use of it in concepts like "embodiment", "correspondence", "commensurate". The next quotation will show that there is no unavoidable contradiction here.

Speaking specifically about learning Maturana says :

"Learning is not a process of accumulation of representations

of the environment; it is continuous process of transformation of behavior through continuous change in the capacity of the nervous system to synthesize it. Recall does not depend on the infinite retention of *structural invariant that represents* an entity (an idea, image, or symbol), but of the functional ability of the system to create when certain recurrent conditions are given, a behavior that satisfies the recurrent demands” (45) (*italics mine*).

It seems that what Maturana denies is the possibility that the representation enters into the description of the nervous system as *a component*, as though there is a minuscule “picture” of the niche somewhere in the system to which the system has recourse whenever it needs it for the synthesis of behavior.

That one can speak here of two types of representation rather than of a negation, of representation is suggested by Maturana’s metaphor of two houses built by two groups of workers with different books of instructions. One book is done in the standard way and it contains all the plans of the house showing the layout of walls, windows, electric connection, water pipes etc., together with several perspectives of the finished house. This group of workers study the plans and guided by the leader construct the house. There is no leader in the second group, and the book of instructions does not contain the plans of the house but “only neighborhood instructions ... of what a worker should do in the different positions, in the different relations in which he finds himself as his positions and relations change” (54). Two books code the instruction differently; one codes the house, the other codes the process. “The first case is typical of the way in which the observer (a human being) codes the systems that he builds; the second corresponds to the way that the genome and nervous system constitute codes for the organism and for the behavior, respectively” (54).

The metaphor suggests that there is a sharp difference between the mode of operation of an autopoietic system (with or without a nervous system) and the mode of operation of men in artefact production (*allo-poiesis*). That these two modes of production cannot be identical seems pretty obvious. Still even if the concept of representation does not enter into the description of the autopoietic and nervous systems, if one cannot find there the leader endowed with the blueprint of the organism, its environment, and the interactions which goes in between, there still must be something

embodied. This is, let us say, the connectivity or the general pattern of the organization of the nervous system that ensures the congruence between the domain of states of the nervous system and the niche, since only that congruence can guarantee that the conduct generated by the nervous system will result in maintenance of the autopoiesis. So it is not necessary to abolish the concept of representation altogether, but only a certain type of representation which is incompatible with Maturana's description of the nervous system.

We may conclude, then, that what makes the nervous system a cognitive system is its participation in autopoiesis where its role is to synthesize a conduct that maintains the interactions with the environment which is necessary for autopoiesis. Its participation is successful if its connectivity and organization determine the space of possible states of the system congruent with the space of possible states of the environment. Nevertheless, one must be ready to give up any idea of simple mapping or isomorphism particularly at morphological level. One must allow the possibility that there is no particular site in the system where representations are embodied. Incompleteness of the nervous system, its architectural asymmetry and unisotropy, the congruence between the conduct generated by the nervous system and the ambience, all that suggests that the concept of reference to the environment and of re-presentation of the environment in the nervous system is reasonable to retain in the description. With one further reservation. Whatever will be the form of representation, it will not be possible to separate in it the part that pertains exclusively to the external from the part belonging exclusively to the internal world; the reference is always the double reference.

Conclusion

As regards the epistemic implications of the concepts of living and nervous system, we are in the following situation. The traditional view, which is (roughly speaking) centered round the concept of adaptation, harbors the idea that the organization and the structure of an organism reflect the organization and the structure of its niche. Pushing this idea to its extreme, organisms appear organic embodiments or another materialization of the features of the niche; they are alter-beings of the parts of physical environment. Consequently cognition is more or less accurate reflection or imaging the environ-

mental characteristics. In the fashionable view of today, the central concept of a genetic program which organisms are expressions of, can be approached either as a set of coded information of possible environments organisms will meet in ontogeny, or as set of coded instructions for internal development and appropriate modulations that use the environment in order to materialize their peculiar identities. In the former case the epistemic effect is the same as in the traditional view. In the latter one may find the Leibnizian language of pre-established harmony the appropriate one. Maturana could easily be a Leibnizian if he would allow for the pre-established program. By closing off living and nervous systems he forbids the organisms to have windows.

So it seems that so far we have strayed between extremes, the extreme openness where the environment freely enters the system to be reflected therein, and the extreme closure in self-pleasing self-development. Both are the consequences of organisms being perceived as thrown into the world by pure accident or God's will, caught there in strange environments that constantly threaten their existences. Epistemologically this treatment leads to establishment of the classical subject — object relation, with unavoidable polarization between subjectivism and objectivism.

The concept of autopoiesis brings a new element into the discussion, but only if it is interpreted as selective openness through interaction with the environment. It is shown, I hope, that the living system with or without the nervous system is autopoietic and cognitive system *because* it is *in need* for an environment, because it must *internalize* a part of the environment in order to produce and maintain itself. In order to satisfy its own internal and internally structured need for the environment the living system, its nervous system included when present, completes itself through selective interaction with the ambience. At the same time, through the same interaction, the living system externalizes itself, its biological identity, its mode of autopoiesis.

Externalization here is the expression of the program which codes the organism's mode of autopoiesis and constitutes one side of the selectivity of interaction. Internalization also means the building up of certain structures subordinated to autopoiesis but "referring" to the external reality, to the niche. These structures make a part of the environment re-present again in the closed space of an organism. Externalization and internalization together make cognition the component process of autopoiesis.

Selective closure/openness, double reference, and the dynamics of internalization and externalization rule out "mirror imagery", the idea of tiny pictures or maps of niches made of organic materials and wired in organism. Re-presence of the niche in the organism is an integral part of its autopoiesis and is subordinated to the basic, metabolic interaction which has to be maintained in all circumstances. So neither pure subjectivism nor pure objectivism nor simple subject-object relation do justice to the fundamental process of life.

This does not mean that we can announce the end of the subject/object cleavage and proclaim the establishment of subject-object identity. As was shown, I hope, the separation between inner and outer space, the living system and the environment, the system which accomodates re-presence of another system and the system which is re-presented, subject and object are essential for living. The crucial point is selective *interaction*. Interaction is the most important part of autopoiesis, it is the homeostatic parameter which has to be kept constant, and both the internal structure of the organism and the niche are subordinated to it. Organisms appear as units of interaction, niches as domains of interaction. The organic evolution may be seen as evolution of modes of interaction between autopoietic systems and their environments.

This suggests that maybe we need an *interactionist epistemology* which will replace the traditional one that is caught in the subject — object polarization and the mirror imagery. Instead of asking "How does the subject obtain information about its object?" it will try to answer the question "How does it happen that the subject has a structure that permits it to interact with the object in the way that maintains its identity?" It will then focus on different modes of interaction appearing during organic evolution and human history. Cognition could then be seen as indefatigable repetition of the experiment in design of variegated modes of interaction capable to support autopoietic systems which are nothing but embodiments of those same modes of interaction. The establishment of nervous systems and human artefact production does not require the change of perspective, they are just new modes of autopoietic interactions.

The naturalistic approach which an interactionist epistemology would embrace (with the results of biology of cognition taken seriously) is apparently committed to a sort of pragmatism and relativism, as Maturana also suggests. If the mode of interaction is

the central concept through which one can understand how the unit of interaction, which is the cognitive system, and the domain of interaction, which is the cognitive domain, are constituted; and if all aspects of life, and therefore of cognition, are subordinated to maintenance of interaction; then pragmatism of the sort *determined by the mode of interaction* is unavoidable. Similarly, if we cannot separate cognition from autopoiesis, than absolute, species-independent, God-like knowledge is impossible too. All cognition is then species-relative.

However, the question about the validity of naturalistic inference still remains: "Since humans are living systems, whatever is true for all living systems is true for humans as well". To answer the question positively one has to provide a naturalistic account of the specifically human mode of autopoiesis, of human society, and history. Maturana courageously enters this difficult terrain, and although I think the successful voyage can be done, I am not able to follow him there in this paper.

University of Zagreb

NOTE

¹ Maturana wrote more than half of the book himself, and the second part is extension of one chapter from the first part; so for the sake of brevity I will often refer only to Maturana.

REFERENCES

- CAMPBELL, D.T., (1974), "Evolutionary Epistemology", in: P.A. Schilpp (ed.), *The Philosophy of Karl Popper*, Vol. I, The Library of Living Philosophers, La Salle, Ill., Open Court Publ. Co, 413-463.
- MATURANA, M.R. & P.J. VARELA, (1980), *Autopoiesis and Cognition*, Dordrecht, Reidel.
- QUINE, W.V., (1968), "Epistemology Naturalized", in: *Akten des XIV' Internationales Kongress für Philosophie, Wien, 2-9 September 1968*, Wien, Hencler.
- RORTY, R., (1980), *Philosophy and the Mirror of Nature*, Oxford-Princeton, Basil Blackwell & Princeton Univ. Press.
- RORTY, R., (1982), *Consequences of Pragmatism*, Brighton, The Harvester Press.