Autopoietic and (*M*,*R*) systems ☆

Juan Carlos Letelier*, Gonzalo Marín, Jorge Mpodozis

Facultad de Ciencias, Departamento de Biología, Universidad de Chile, Casilla 653, Santiago, Chile

Abstract

From the many attempts to produce a conceptual framework for the organization of living systems, the notions of (*M*,*R*) systems and Autopoiesis stand out for their rigor, their presupposition of the circularity of metabolism, and the new epistemologies that they imply. From their inceptions, these two notions have been essentially disconnected because each has defined its own language and tools. Here we demonstrate the existence of a deep conceptual link between (*M*,*R*) systems and Autopoietic systems. This relationship permits us to posit that Autopoietic systems, which have been advanced as capturing the central aspects of living systems, are a subset of (*M*,*R*) systems. This result, in conjunction with previous theorems proved by Rosen, can be used to outline a demonstration that the operation of Autopoietic systems cannot be simulated by Turing machines. This powerful result shows the potential of linking these two models. Finally, we suggest that the formalism of (*M*,*R*) systems could be used to model the circularity of metabolism.

1. Introduction

Biology is an experimental science in which global theoretical principles concerning the intrinsic peculiarity of living systems have had great difficulties in opening avenues of research. In the sense of the interrelation between theory and experimentation, Biology has revealed itself to be a conceptual domain very different from physics. This difference is not surprising as the properties of inanimate matter and living systems appear so distinct to all observers. With perhaps the sole exceptions of the notions of Evolution and Mendelian genetics, *theories* in biology refer not to aspects central to biological organization but rather to applications of results from physics to subsets of (phenomena in) the biological world. An excellent example of such theories can be found in molecular biology, a field that has populated cellular metabolism[1] with imported concepts including: *signals*, *transducers*, *information*, *encoding* and *decoding*.

In spite of the difficulty of generating a theoretical framework about the central phenomenology of living systems, the 20th century witnessed the creation of such theories. In the context of this paper, it is important to mention three theoretical bodies that have served as general scaffolding or metaphors to describe biological phenomena. These include General System Theory (von Bertalanffy, 1950), Cybernetics (Rosenbluth et al., 1943) and the modern, eclectic field of Artificial Life (Langton, 1989).

In parallel with these three well-known viewpoints, the second half of the century saw the appearance of many other theories about living systems that have had less impact. These ideas encompass a wide variety of models, from the structurally rich notions of *Hypercycles* (Eigen, 1971), or *Autocatalytic Sets* (Kauffman, 1993) to the intriguing idea that "life equals cognition"

[1] The terms "cellular organization" and "metabolism" would be used interchangeably to denote the complete network of biochemical reactions (biosynthesis, degradation, catabolism, active transport, genome duplication, etc.) occurring inside the cell.

of *Evolutionary Epistemology* (Heschl, 1990), or the class of *Component-systems* defined by the impossibility of defining states or equations of motions (Kampis, 1991) along with attempts to join concepts from biology, chemistry, physics and mathematics in a single theoretical construct, like the *Generalized Theory of Life* (Kalmykov, 1998)) or the field of *Biosemiotics* that considers semantic communication to be the essence of living systems (Sharov, 1991).

In this crowded field, two theories stand out for their rigor, their central focus on the circular causality proper to living systems, the new epistemologies that they imply, their initial focus on cellular metabolism and their detachment from structural details. These two theories are (*M,R*) *systems* created by Rosen (1958a), and the notion of *Autopoiesis* set forth by Humberto Maturana and Francisco Varela in the early 1970s.

The purpose of this paper is to show the existence of a deep connection, or correlation, between (*M,R*) systems and Autopoiesis and to explore the consequences of such a connection. Because the notions of (*M,R*) systems and Autopoiesis are not generally well known, we will first present an overview of both theories, with an analysis of the impact that these concepts have had. Second, we will demonstrate that both theories, although originally stated in very different conceptual frames and languages, share a common structure. Third, we will demonstrate that Autopoietic systems are a strict subset of (*M,R*) systems, and finally we will use this inclusion to prove that living Autopoietic systems are not Turing computable.

2. An algebraic approach to circular organization: (M,R) systems

2.1. Background

In the 1930s, Nicolas Rashevsky,[2] a physicist by training, championed the biophysical approach to understanding living systems. Rashevsky and his students created a systematic theoretical effort that consisted of applying theories from physics to explain biological phenomena like cell division and neural processing (Rashevsky, 1938). Around 1950, Rashevsky became convinced that his intense and novel "biophysical" approach was fundamentally limited for understanding living systems as a whole. He realized that his previous work had dealt only with bit parts of the phenomena of living systems, without considering their peculiar organization. Thus, Rashevsky coined the

term *Metric Biology* to refer to all aspects where a reductionist approach to biology was valid and the term *Relational Biology* to aspects that depended on the organization of living systems rather than the matter found inside them (Rashevsky, 1954).

In 1958–1959, as a graduate student of Nicolas Rashevsky, Robert Rosen published three papers (Rosen, 1958a, b, 1959) that were a rigorous attempt to formalize the intuitive notions of relational biology. His formalism (known as (*M,R*) *systems*) used mathematical language based on a modern and abstract branch of mathematics (*Theory of Categories* (Eilenberg and MacLane, 1945)). Since not many biologists are well-enough versed in algebraic theory to evaluate its utility, (*M,R*) systems has not had the wide impact it may deserve. Despite the limited audience Rosen could capture with his ideas, Rosen continued to develop the theory of (*M,R*) systems and the use of the theory of categories in Biology for 40 years until his death in 1998.

2.2. The formal model

To model cellular metabolism, Rosen initially distinguished two types of entities: *Components* and *input/output materials* (i.e. transformable materials) (Rosen, 1958a). A component transforms a set of input materials into output materials. Enzymes represent components as they transform reactants into products without being changed themselves. With this metaphor in mind, it is possible to formalize the notion of a metabolic network of interconnected biochemical reactions by rewriting a metabolic map in terms of components and metabolic processes. Because components represent the action of metabolism, their action is termed *metabolic* (the *M* in (*M,R*)). A real metabolic network contains thousands of these components, forming a family of metabolic components (M_i) and many thousands of transformable materials. Thus any metabolic reaction, like the phosphorylation of glucose, could be thought of as the action of a component upon a set of input materials to produce output materials (Fig. 1A). Cellular metabolism could then be represented by a graph where the nodes are the M_i and the links show that the output of one component is the input for another (Fig. 1B).

In real living systems, each component M_i has a finite lifespan. Thus to have a stable system operating in a steady state, for every component M_i, there must be a mechanism (or subsystem) that produces M_i and maintains a functional concentration in the metabolic system. Hence in metabolic networks, associated with each component M_i, there must be a subsystem R_i (the *R* in (*M,R*)) that *repairs* each component M_i (Fig. 2). The collection of subsystems $\{R_i\}$ entails an enormous complexity, as these subsystems are also made of components that transform input materials into output materials (in this case, the outputs are the components

---

[2] Rashevsky was the founder of the Bulletin of Mathematical Biophysics (in 1938), as well as the creator of the first graduate program of theoretical biology (Program on Theoretical Biology at the University of Chicago).
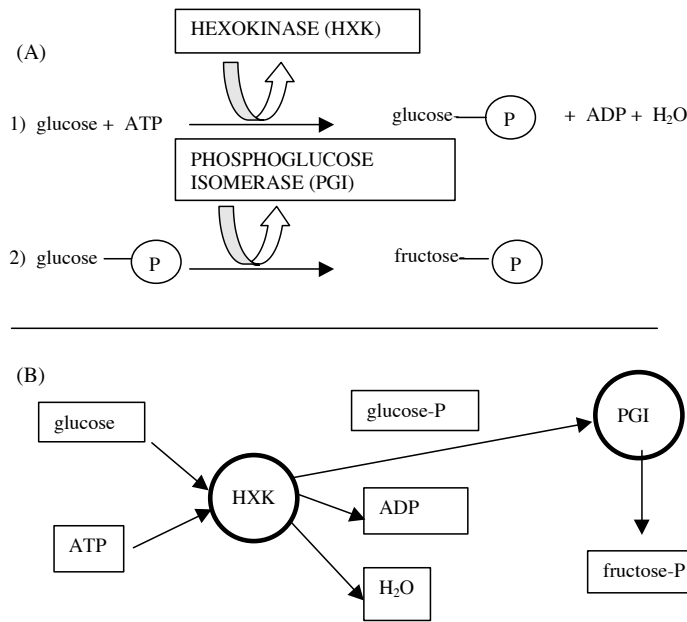
Fig. 1. The metabolic network. (A) Every biochemical reaction (like the initial step of glycolysis: phosphorylation of glucose), can be viewed as the action of a component (in this case the enzyme hexokinase) upon a set of inputs (glucose and ATP) to produce a set of outputs (glucose-6-phosphate, ADP and water). The second reaction of the glycolytic pathway (glucose isomerization) can be similarly viewed: glucose-6-phosphate is isomerized to fructose-6-phosphate by the enzyme phosphoglucose isomerase. (B) Dual view of the 2 previous biochemical steps. In this representation the enzymes, or "components" in Rosen's language, are the nodes of a graph. The vertices are labeled by the materials that are transformed. Two nodes are connected by a directed arrow if the output of one is the input for the other.
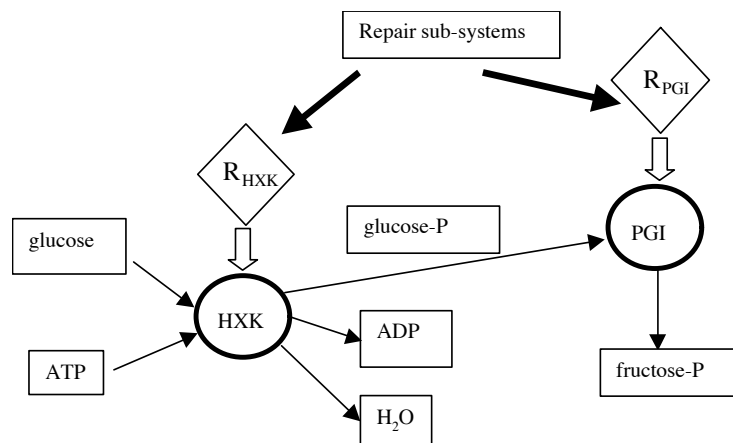

Fig. 2. R subsystems. The components in a metabolic network (i.e. like enzymes HXK and PGI) have a finite lifespan. Thus, in order to have a network operating in steady state, these components must be repaired by special subsystems: R_{HXK} and R_{PGI}. Repair subsystems are not simple; they can contain many components and transformable materials.

M_i). Furthermore, the input materials for each R_i are the byproduct of the collective action of the M_i such that an (M,R) system could have incredible resiliency as every component is dynamically maintained by a network of processes (the R_i subsystems) whose input materials are derived from the same metabolic network fueled by the action of the set of components M_i

A crucial question, first addressed by Rosen (1959), concerns the infinite regress implied by the existence of subsystems R_i. In effect, as is the case for every M_i, every

R_i is a physical entity with a finite lifespan. Which systems or collection of subsystems, then, repair R_i? The invocation of special subsystems that would repair each R_i is obviously an inadequate answer. Rosen suggested that the infinite regress could be avoided if the set $\{R_i\}$ had the capability of self-replication by the (M,R) system.

Rosen's main result is the demonstration that the synergy of metabolic and repair actions can imply, under some circumstances, self-replication in the sense
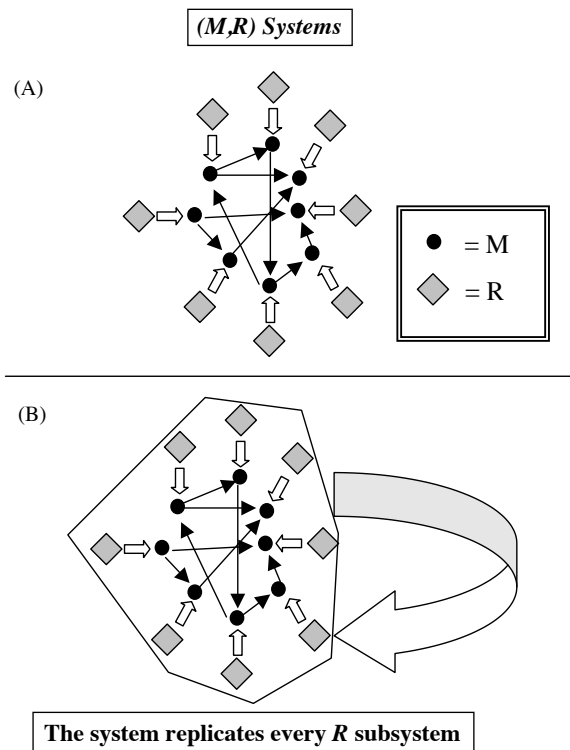
(M,R) Systems

(A)

● = M
◇ = R

(B)

The system replicates every R subsystem

Fig. 3. Metabolism and repair imply replication. In an (**M,R**) system *every* component **M** must have an associate **R** subsystem that maintains its concentration (A). But **R** subsystems are themselves physical entities that need to be repaired. Instead of invoking an infinite descend, or neglecting the problem, Rosen proved that the whole (**M,R**) system can, under very general circumstances, replicate each **R** subsystem. This systemic replication assures the continuous operation of the system (B).

of self-production (or self-maintenance) of the complete metabolic network. In other words, in some (**M,R**) systems, it should be possible that the metabolic and repair actions induce the replication of each subsystem R_i (Fig. 3).[3] Incredibly the demonstration of this important result concerning the closure of metabolic systems does not rely on the explicit connectivity of the graph between all the M_i (components) and R_i (repair subsystems) inside the metabolic network, but rather on a universal, but not obvious, property between sets and sets of mappings (Rosen, 1959).

2.3. The Φ formulation

In a second phase, Rosen used the very general framework of the theory of categories to convert the replication result of 1959 into a theorem about *closure* (Rosen, 1972). From a mathematical viewpoint, the theorem consists of a procedure for selecting functions from a set of functions, using those functions with their

---

[3] These systems are called replicative (**M,R**) systems, in this paper, we only deal with this class of (**M,R**) systems.

ranges and *domains* as the only elements needed to build the selecting procedure.

Rosen's first step was to obtain a new representation of the overall action of metabolism (i.e. the collective action of all M_i) as a mapping f between the set of possible input materials (A) and the set of possible outputs materials (B). Thus instead of a complex graph like the one depicted in Figs. 2 and 3A, which would be unmanageable for real metabolisms involving thousands of enzymatically controlled steps, metabolism (i.e. f interpreted as one instance of all possible instances of functions connecting (in the sense of set theoretic functions) set A with set B (this set of functions is denoted by $H(A,B)$). But f is indeed a very special function as it must embody the properties of metabolism *and* it must have the property of closure enunciated above (Metabolism + Repair → Replication). Rosen theorized that only some of elements of $H(A,B)$ would exhibit the closure (or circular organization) proper of biological systems, and the vast majority of the mappings between A and B would be "uninteresting" as they would not have circular organization.

Thus, Rosen's central theoretical problem was to find a mechanism (an operator in the language of functions) for any $b \in B$ (where b *must* be the image of a certain a, that is: $b = f(a)$) to *select* from among all the elements of $H(A,B)$ the only one (f) that is biologically relevant as it represents *circular* metabolism in a given (**M,R**) system. Rosen's interpretation was to assume the existence of an operator that, using elements b from set B as input, selects f from $H(A,B)$. This operator is denoted by Φ and must have the property that it uses an instance (or realization) of output materials (b) to define metabolism (f):

$$\Phi(b) = f \tag{1}$$

Φ implements the *repair* function in an (**M,R**) system, and thus plays the role of subsystems $\{R_i\}$. In effect, starting with an instance of output materials ($b \in B$), Φ obtains f (which represents metabolism) from $H(A,B)$ (Fig. 4).

Rosen did not view Φ as an imaginative theoretical device, or a metaphor for *repair*, but rather as a real object acting in the world. Accordingly, Rosen named these kinds of entities as *functional components*. Thus from a theoretical standpoint, a metabolic network is partially represented by Fig. 5.

In Fig. 5, a mapping, represented by f, acts upon a molecular configuration of inputs materials ($a \in A$) to produce a configuration of output materials ($b \in B$). While the operator, Φ (a mapping between B and $H(A,B)$), uses b to specify f. Superimposing the formalism of the four Aristotelian causes (*material*, *formal*, *efficient* and *final*), it can be said that f is the efficient cause that acts upon a material cause ($a \in A$) to
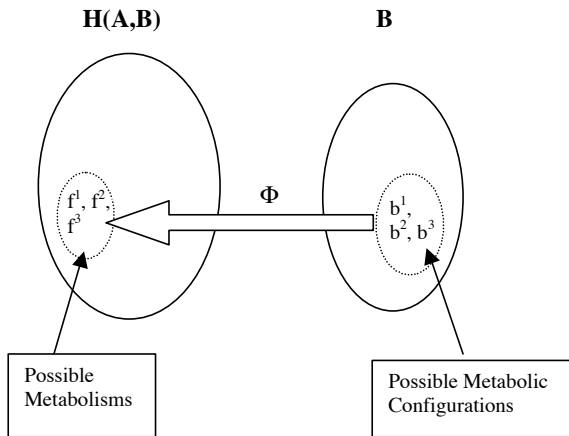
Fig. 4. The Φ formulation. Rosen reframed his 1958–1959 result concerning the systemic replication of **R** subsystems in pure functional terms. The operator Φ selects from a family of possible metabolisms (f^1, f^2,...) the function f that realizes a concrete metabolic network. The selection process uses molecular configurations (b^1, b^2,...) as its input. Most of the possible networks of processes (i.e. most of the elements of the set $H(A,B)$) do not have the property of circular organization.
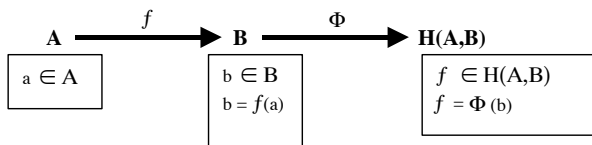

Fig. 5. Relation between f and Φ. Metabolism can be thought of as a function (f) that transforms a set of materials (the instance $a \in A$) into another instance ($b \in B$). The instance b is then used to select f from the set $H(A,B)$.

produce an effect ($b \in B$). Aristotelian causes play an important role in Rosen's (1991) analysis.

The next step concerns the mechanism by which Φ is specified. This question is the mathematical counterpart of "Which system produces every R_i?". In the language of closure and Aristotelian causes, this question demands that an "object" be found inside Fig. 5 that is the efficient cause of Φ. Rosen was able to prove that, in some cases, a mapping β, which Rosen called a *replication map*, could exist with the property $\beta(f) = \Phi$ (Fig. 6A) The crux of the formalism used by Rosen is to identify β, which is an element of the following set of functions $H(H(A,B),H(B,H(A,B)))$, with a (molecular) configuration in B.[4] This identification produces the diagram in Fig. 6B which summarized the complete research program of Rosen. The power of Rosen's approach is that he treated the sequence of mathematical objects (A,B,f,Φ) as an (**M,R**) system and as a
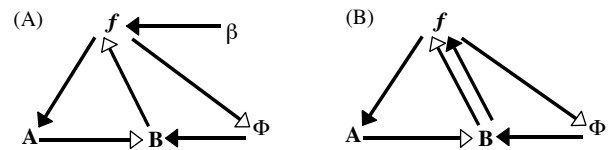

Fig. 6. Rosen's diagram. (A) The first step consists of finding a formal entity, β, (which plays the role of replication) that acts as the efficient cause for the production of Φ. In this diagram, β appears as independent of metabolism (f), repair (Φ), and the states of the metabolic network (A & B). (B) This diagram illustrates the fundamental result of Robert Rosen: β is identified with a metabolic configuration. Thus all important biological functions, metabolism (f), repair (Φ) and replication (β), are mutually dependent on each other or in Rosen's language "entailed". Metabolism (f), repair (Φ) and molecular configurations (β) define each other. These diagrams distinguish between material causation (open arrowhead) and efficient causation (solid arrowhead). Thus, B is the end result of efficient cause (f) acting upon a material cause (A) in order to produce its effect (B). B plays a dual role; it is the material cause associated with the efficient cause (Φ), and, when interpreted as β, it is the efficient cause that produces Φ.

category C.[5] Thus extending the class of (**M,R**) systems to include, beside models of metabolism, pure formal systems. Fig. 6B shows that three important biological functions (metabolism (f), repair (Φ) and self-replication (the diagram itself)) are entailed by another function inside the diagram and nothing else: thus proving the closure (circular causality) of metabolism. This diagram was summarized by Rosen with the dictum, "Organisms are different from machines because they are closed to efficient causes." Thus in an organism, all efficient causes are produced inside the organism. In this respect, organisms are very different from man-made artifacts, where every component (from a line of code, a transistor or a humble wooden handle) is produced by mechanisms generated outside the artifact.

Finally, Rosen proved that (**M,R**) systems encompass a wide spectrum of systems, many of them purely formal systems, and that cells, as they are usually investigated by biologists, were an example of a "molecular realization" of a (**M,R**) system (Rosen, 1972). Using the formal characterization of (**M,R**) systems, Rosen explored some of their properties. One of his results was the unexpected conclusion that (**M,R**) systems cannot be simulated by Turing machines (Rosen, 1964, 1966, 1991). Without any doubt Rosen's extremely coherent theoretical viewpoint cannot be summarized in few pages. Our aim is to present a glimpse of this complex theoretical approach and take the minimal steps necessary to explore the relationship between (**M,R**) and Autopoietic systems.

<hr>

[4] A more extensive description of these results is found in Rosen (1972). The term replication map is somewhat unfortunate as it evokes the idea of reproduction.

[5] The proof of the existence of β, which constitutes the very kernel of Rosen's approach, is a delicate demonstration because it requires the *invertibility of evaluations*, a situation that is seldom encountered.

3. A systemic approach to circular organization: autopoiesis (Maturana and Varela, 1972)

3.1. Background and model

In 1972, in the middle of a cataclysmic political turmoil, two Chilean biologists introduced the concept of Autopoietic systems[6] ("auto" = self and "poiesis" = generating or producing) as a theoretical construction on the nature of living systems centering on two main notions: the circular organization of metabolism and a redefinition of the systemic concepts of *structure* and *organization*. Maturana and Varela's starting point was that any system can be decomposed into *processes* and *components*. Components interact through processes to generate other components.

The notion of *circular organization* is given in Autopoiesis, and it is immediately clarified in the theory by the very definition of an *Autopoietic system*:

"an Autopoietic system is organized as a bounded network of processes of production, transformation and destruction of components which:

(i) through their interactions and transformations continuously regenerate and realize the network of processes that produced them

(ii) constitute the system as a concrete entity in the space in which the components exist by specifying the topological realization of the system as such a network" (Varela et al., 1974; Maturana and Varela, 1975, 1980).

In an Autopoietic system, the result of any given process is the production of components that eventually would be transformed by other processes in the network into the components of the first process. This property, termed *operational closure*, is an organizational property that perfectly coexists with the fact that living systems are, from a physical point of view, energetically and materially open systems. The molecules that enter the system determine the system's organization, which generates pathways whose operation produces molecular structures that determine the physical system and the system's organization (Fig. 7) (Fleischaker, 1990). Thus an Autopoietic system does not have *inputs* or *outputs*, instead it creates a web of molecular processes that result in the maintenance of the autopoietic organization. Because an Autopoietic system's internal dynamics are self-determined, there is no need to refer any operational (or organizational) aspect to the outside. Thus the environment does not *inform*, *instruct* or otherwise *define* the internal dynamics, it only *perturbs* the system's dynamics. This does not mean that an Autopoietic system is completely independent from its

---

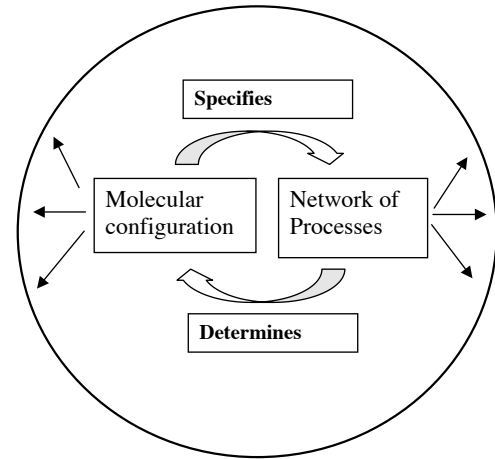[6] This theoretical body is also known as *Autopoiesis* or *Autopoietic Theory*.

Fig. 7. Partial representation of an Autopoietic system. The set of molecules found inside an Autopoietic system specifies the metabolic processes that determine the type and arrangement of these molecules. Thus, the molecular configuration and the network of processes define each other in a recurrent mode, and the boundary of the system is actively created by this interplay. An Autopoietic system produces a unity that is topographically and functionally segregated from its background.

medium. Instead it means that the system specifies its own internal states and the domain of its changes. In this context, external events act as perturbations that only trigger internal changes. But the *magnitude* and *direction* of these changes are defined by the internal dynamics of the system and not by the external perturbations (Maturana and Mpodozis, 2000).

The second clause demands that an Autopoietic system has "sufficiently complex" dynamics to self-produce the boundaries that separate the systems from the "non-system". This apparently trivial clause has profound implications as it touches upon the problem of autonomy and also serves to weed out from the Autopoietic forest some pure formal systems. Thus Autopoietic systems are not simple relational devices that connect components with components via complex graphs. Autopoietic systems must conform to an important topological property: their boundary (in the space where their components exist) is actively produced by the network of processes that define the system's identity. This property of Autopoietic systems couples a purely relational property (operational closure) with a topological property and it demands that an Autopoietic system must be an autonomous unity, topographically and functionally segregated from its medium, but yet dependent from this medium (Weber, 2001). In the realm of molecules, the coupling of these two conditions necessarily implies that the minimal metabolism must be rather more complex than the spatial coupling of a direct chemical reaction with its reverse reaction.

The theory of Autopoietic systems uses the concepts of organization and structure with a new viewpoint and,

in order to understand the theory, it is relevant to understand how these notions are used. Thus the *organization* of a system is defined as the pattern or configuration of processes between components of the system that define such a system as a member of a particular *class* of systems. The *structure* is the specific embodiment of these processes into specific material entities. When a process (i.e. elongation by condensation) is embodied into a specific physical entity (i.e. a polymerase), it also automatically defines the physical properties or characteristics (i.e. nucleotides) of the components transformed by the initial process. According to this definition, organization is a subset (in the mathematical sense) of structure. The organization of a system defines its systemic identity and its structure only specifies one instance of the system's organization. Furthermore for a given system, the ontogenies of organization and structure are not coupled, as some structural changes do not imply any organizational change.

Autopoiesis, as originally described by Maturana and Varela (1972, 1980), is an extremely coherent and formal theory formulated outside any mathematical framework. Many attempts have been made to formalize and simulate Autopoiesis. The first tessellation computer models, initially done in an IBM 360 (Varela et al., 1974; Zeleny, 1981)] and recently re-done in Swarm (McMullin and Varela, 1997) have been a direct translation of a minimal Autopoietic system into a small bi-dimensional lattice. Varela used an Indicational Calculus (Varela, 1979) to model autonomous systems. But Indicational Calculus, developed by Spencer-Brown (1969), is a difficult tool to master and the progress, aside from the effort's of Varela, has also been limited. Other mathematical formulations have included the use of differential equations to model feedback (Limone, 1977). None of these models has generated clear-cut, satisfactory results.

Implied in the early concepts about Autopoietic systems is the idea that an observer is not naive and transparent as science usually supposes. On the contrary, the process of cognition is embodied, not only in logical and inferential rules, but in a specific neurophysiological substrate with specific cognitive consequences, where the nervous system cannot distinguish illusion from perception (Maturana, 1970a, 1970b; Fleischaker, 1988).

4. The current impact of both theories

The formalism of (*M,R*) systems has had limited impact on biology. Excepting the direct and massive work of Rosen himself, his ideas have remained little explored (Rosen, 1991, 2000). One exception was the sustained theoretical work of Leguizamon and coworkers who used Rosen's ideas about the functorial representation of systems to embark upon a research program concerning the physical-chemistry of living and non-living systems (Zaretzky, 2000). Another important line of work is due to Casti (1988, 1997) who explicitily modeled linear (*M,R*) systems. Recently, the implications of Rosen's epistemology in understanding complex systems (Casti, 2002) or in Bio-informatics (Wolkenhauer, 2001) have been published.

The development of Autopoiesis has been very different from that of (*M,R*) systems. First, Autopoiesis has been an extremely successful idea in various arenas outside of Biology ranging from law (Luhmann, 1982) to business administration (Mingers, 1995) and even psychotherapy (Snyder, 1999). Second, because of Autopoiesis' epistemological foundations concerning the process of cognition, it has become a central paradigm of "second order cybernetics" (i.e. the observer is considered, at least, as part of the feedback loops defining the system) (Zeleny and Hufford, 1992). But in biology, apart from new versions of the original computer simulations (McMullin and Varela, 1997), some applications to the problem of the origin of life (Fleischaker, 1990; Mavelli and Luisi, 1996), approximations to the origin of higher brain functions (Mpodozis et al., 1995), its use in image processing (Ruiz-del-Solar and Köppen, 1999), and producing a new formalization of Evolution based on natural drift rather that natural selection (Maturana and Mpodozis, 1992, 1999, 2000), the notion has had limited advance. Autopoiesis, acclaimed by theorists in many disciplines (Mingers, 1995) has not penetrated the daily life of biologists.

In summary, neither Autopoietic systems nor (*M,R*) systems have been used to explain any experimental findings or to predict new biological phenomena in an unambiguous way. It is not surprising then that these theoretical models have been neglected by the vast majority of experimental biologists. This neglect may reflect the fact that both theories are incomplete in the fundamental aspect of how to map their theoretical concepts (structure, organization, Φ, circularity, etc.) with experimental entities. However, we feel an important step in theoretical analysis could be achieved by first finding a common link or correlation between them.

5. Relations between autopoietic and (M, P) systems. Autopoietic systems are (M, P) systems, but not vice versa

Autopoietic theory and (*M,R*) systems have been two disconnected models with a similar primary objective: to define circular causality as the core of biological organization. Are (*M,R*) systems a subset of Autopoietic systems, vice versa or is there any relationship between
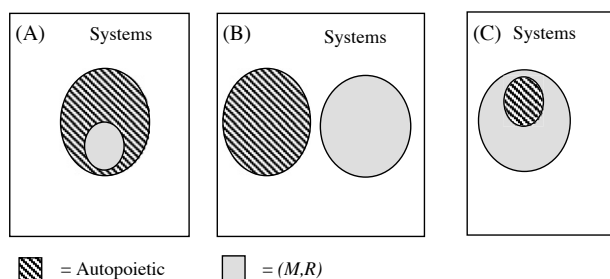
Fig. 8. Possible relations between Autopoietic and (*M,R*) systems. In the universe of systems defined by current theories about living systems (i.e. component-systems, Autocatalitic networks, Cybernetics loops, Semantically closed systems, etc.) it is important to envisage the possible relationship between Autopoietic and (*M,R*) systems. Perhaps Autopoietic systems contain (*M,R*) systems (A), viceversa (C), or they cannot be compared (B).

these two types of systems? Establishing this relationship could be a first step towards producing a synthesis between these two theories (Fig. 8).

Here we submit that every Autopoietic system is, at least conceptually, operationally equivalent to an (*M,R*) system, but not conversely as most (*M,R*) systems are not Autopoietic systems (Fig. 8C). As an initial point we need to decide how to transform or translate these two seemingly different formalisms into a common language. Also, it is important to realize that an Autopoietic system has a more general, and hence less restricted, formal structure because it only contains a single type of object: components. These *components interact through processes*. On the other hand, in an (*M,R*) system at least three types of objects are distinguished: components, transformable materials and *R* subsystems. Furthermore, these objects have different functions: *Materials are transformed*; *components transform input materials into output materials*; and *R subsystems repair components*. Thus from a purely formal viewpoint, the structure and the organization of an Autopoietic system and an (*M,R*) system are rather different. This suggests that it is easier, in the sense of having less restrictions, to work with the generalized formalism of (*M,R*) systems embedded in Fig. 6, where only metabolic functions *f* and *Φ* are used, than with Rosen's early formulations of 1958 and 1959 (Figs. 1 and 2).

The operational closure of Autopoietic systems implies that every process in an Autopoietic network is the direct consequence of the interplay of components produced in other parts of the network. Thus, the components of an Autopoietic systems are (in Rosen terms) the material causes, and the configuration of the network is the efficient cause for the existence of any given process in the network. This argument shows that every Autopoietic system has the property of being closed to efficient causes and thus of being an (*M,R*) system.

Conversely, we can prove that every (*M,R*) system exhibits the operational closure proper to Autopoietic systems. The demonstration flows directly from Fig. 6, where it is shown that *f* (which represents the overall metabolic transformation of set *A* into set *B*) is produced by the action of *Φ* upon *b* (which can be interpreted as a molecular configuration), but *Φ* itself is produced by *f*. This mutual interdependence between *f* and the dual role of the molecular instances (or metabolic realizations), *b* and *β*, found inside set *B* (i.e. $b = f(a)$ and $f = \Phi(b)$ with the functional restriction $\Phi = \beta(f)$) shows, in the language of categories and efficient causes, that an organism (from the point of view of Rosen) exhibits operational closure. In effect, $\Phi = \beta(f)$, can be rephrased as: the *action* of metabolic self-maintenance (*Φ*) is the product (via *β*) of the *action* of metabolism itself (*f*).

In conclusion, (*M,R*) systems exhibit the circular organization (operational closure) of Autopoietic systems. This shows that every (*M,R*) system fulfills the first part of an Autopoietic system in that it has circular organization with operational closure. But nothing in the formalism of (*M,R*) systems, not in the early formulation (1958, 1959) nor in the latter formulations (Rosen, 1972, 1991) indicates how a generic (*M,R*) system can generate a distinguishable unity (i.e. the second property of an Autopoietic system). The reason is that the pure algebraic structure of a (*M,R*) system precludes the second, and topological characteristic of Autopoietic systems. In effect it is possible to think that a mathematical structure (i.e. a purely logical construction) could fulfill the conditions embodied in Fig. 6, but this pure logical construct will not be a living system. Thus, we have the strict or proper inclusion of the class of Autopoietic systems inside the class of (*M,R*) systems. In essence, the property that a generalized (*M,R*) system lacks, in order to be Autopoietic, is the generation of its own border and the internal topology that Autopoiesis implies. Perhaps (*M,R*) systems cannot be realized, in the domain of molecules, without being Autopoietic. In other words: in the domain of systems (*M,R*) systems are a superset of Autopoietic systems, but in the domain of real and concrete systems, made of molecules, both sets are equal.

There are two other important similarities between Autopoiesis and (*M,R*) systems. First, both theories demand a drastic epistemological shift in defining the question "What is a living system?". While Autopoiesis was the direct offspring of a constructivistic theory of knowledge based on neurobiological results in visual perception (Maturana, 1970a, 1970b), using (*M,R*) systems as a starting point, Rosen built a theory that he termed "modeling relation" based on the relation between the process of measurement and the formal models that capture such measurements (Rosen, 1985). These epistemological considerations only reflect

that circular causality, as it is implied by living systems, cannot be analysed without changing the normal reductionist approach used in experimental science. An excellent example can be found in the work of Heschl and his notion of evolutionary epistemology where he states that a theory about living organization is at the same time a theory about cognition (Heschl, 1990).

Second, Autopoiesis and (*M,R*) systems do not consider time as a relevant parameter in the description of living organization. Neither theory uses evolution equations to characterize living systems. Both are essentially atemporal and intrinsically relational, which sets them apart from theories that consider time and thus frame biological phenomena in the language of differential equations. This characteristic should not be construed as a weak theoretical point as any algebraic structure, like a group or a vector space, is defined without using the concept of time. These structures simply exist and have properties that depend on other properties of the structure but not on any time-evolution equation or concepts. This atemporal characteristic definition does not negate that, in certain circumstances, as it has been done by Casti (1988) for (*M,R*) systems, time could be incorporate for certain models or formulations.

Finally, it is important to differentiate (*M,R*) systems and Autopoietic systems from Autocatalytic sets (also known as collective autocatalysis) (Kauffman, 1993). The main difference is that collective autocatalysis requires some form of spatial confinement for effective operation, a condition which is not produced by the set of chemical reactions making up the collective set. Thus, a realizable collective autocatalytic set requires spatial or topological properties imposed from the outside (McMullin, 1999). On the other hand (*M,R*) and Autopoietic systems produce all the efficient causes needed for their realization.

5.1. An easily obtained, but highly significant deduction

Because Autopoietic systems are included in the class of (*M,R*) systems, the important facet of their *computational abilities* can be addressed using the crucial result that (*M,R*) systems are not Turing computable in the sense that the extended Turing–Church hypothesis does not apply to them (Rosen, 1964, 1991). A simple application of this fundamental result in conjunction with the inclusion of Autopoietic systems in (*M,R*) systems shows the impossibility of constructing a Turing machine whose sequence of transitions is isomorphic to the states of an Autopoietic system. This latter result seems difficult to prove using only the elements of Autopoietic theory (Maturana and Varela, 1972, 1975), but it trivially flows from the inclusion of Autopoietic systems in (*M,R*) systems.

6. Discussion

Autopoietic and (*M,R*) systems define the problem of circular organization as the core of living systems but approach this circularity from two different perspectives. While Rosen tried, with (*M,R*) systems, to prove that circular organization can arise spontaneously from the functions of metabolism and repair inside a metabolic network with a time-invariant organization, Autopoiesis explored the biological consequences of a living system characterized by circular organization acting (behaving or interacting) with its environment through structural coupling (see below). (*M,R*) systems builds a theory to describe and manipulate metabolic networks, specially to understand, how an invariant metabolic organization arises through the interplay of metabolism, repair and replication. Both theories posit that the core of biological phenomena arises from circular organization, and not from information processing, reproduction, the generation of "correct" responses to outside stimuli or optimizing metabolic fluxes by minimizing energy use. Superficially, both theories are similar to the notion of *self-organization* but their focus on epistemological aspects, algebraic reasoning, the relation with computability theory (Rosen, 1991, 2000), emphasis on autonomy (Varela, 1987) or in its relation with cognition (Maturana, 1987) clearly show that they encompass a wider (and richer) field that the word self-organization denotes.

The idea of using the formalism of (*M,R*) systems as a possible framework of Autopoietic systems was previously advanced in the context of *Quasi-Autopoietic systems* (QAP) (Nomura, 1997). Despite their name, QAP systems are far more similar to (*M,R*) systems than to Autopoietic systems. Instead of having a replication map Φ QAPs have an iteration map that defines the internal dynamics through a recursive procedure. QAPs also lack a specific metabolism and, more importantly, they do not consider the existence of the boundary. The mathematical efforts behind QAPs are interesting but, in our opinion, they fail to grasp the fundamental biological relationship between (*M,R*) and Autopoietic systems. The theory of categories has also been used to model the internal organization of a living system, its evolution and its levels of organization (Ehresmann and Vanbremeersch, 1987). This model, known as *Systemes Evolutifs avec Memoire* (SEM), is a technically complex formalization of biological systems that crucially ignores operational closure or circular organization and, in this sense, is a theoretical effort centered on a categorical representation of living systems but very different from (*M,R*) systems.

By virtue of Autopoietic systems' essential turnover of components, as well as the destruction and creation of whole classes of molecules during ontogeny, these systems cannot be characterized within the scope of

traditional Dynamical Systems Theory. As their structure can change, without changing the organization, Autopoietic systems cannot be described with a fixed-state space (Kampis, 1991). The challenge is to use the categorical formalism of (*M,R*) systems to develop a categorical representation of Autopoietic systems. This specific work should be focused on finding a categorical representation of a system that is circularly closed and *produces its own boundary*. In summary the categorical representation used by Rosen (1972) must be refined to include the (extra) properties that make an (*M,R*) system an Autopoietic system. The relevance of the system's boundary revealed by our theoretical analysis is currently paralleled by the experimental efforts to synthesize proto-cells and the importance that metabolic compartmentalization acquires in such experiments (Szostak et al., 2001).

Autopoietic systems' non-computability by Turing Machines has many important theoretical consequences. First, it limits the validity of mimesis (i.e. simulation) as a means to understand living systems. In effect, this result shows that the phenomenology that arises from the circularity of metabolism cannot be simulated with current computer architectures based on the Von-Neumann implementation of Turing machines. Using different approaches this result has been hinted at on at least two occasions in the last decade. Using formal arguments, Boden argued for the impossibility of designing a living system without a real metabolism, thereby raising serious doubts about the conceptual program of Strong Artificial-Life (Boden, 1999). On the other hand, Kampis has developed a concept of living systems, which he calls "Component-systems," and he shows that equations of state, equations of motion or evolution equations cannot be applied to Component-systems (Kampis, 1991). The non-computability of Autopoietic systems, as advanced here, apparently collides with the simulation results involving tessellation automatas (Varela et al., 1974). But new versions of this simulation show that the original report of computational Autopoiesis was flawed, as it used a non-documented feature involving chain-based bond inhibition (McMullin and Varela, 1997). Thus the closure exhibited by tessellation automatas is not a consequence of the "network" of simulated processes, but rather an artifact of coding procedures (McMullin and Varela, 1997). Thus our point concerning the non-computability of Autopoietic systems appears supported by the more modern simulations. In any case, as our analysis has shown, the failure of closure in these computational models cannot be construed, in any way, as a conceptual failure of Autopoiesis, instead it reflects the non-computability of Autopoietic systems.

The non-computability of Autopoietic systems could initially appear as an incredibly (or suspiciously) strong result, but even in the restricted field of pure

Mathematics it has been possible to prove the existence of simple, but non-computable functions like the *busy beaver* problem (Rado, 1962). Thus, Turing non-computability is a property that does not require the complexities of circular organization to be apparent, as it is already demonstrable in simpler systems or problems. The failure of the Turing–Church hypothesis with respect to Autopoietic systems opens some important new questions. The first challenge would be to analyse whether an Autopoietic system can implement a Turing machine. The second, and far more interesting question, is to consider whether some Turing non-computable problems, like the busy beaver, can be computed by Autopoietic systems. These considerations belong to the new field of Emergent Computation or bio-computing. To tackle these problems, it is essential to expand the tools developed by Rosen, essentially the use of the theory of categories to represent Autopoietic systems and to understand and manipulate the operational closure of metabolism.

Autopoietic systems do not simply behave or exist passively in an environment. A central aspect of Autopoieis is the idea of *structural coupling*, a mechanism by which the living system and its environment determine, in a mutual way, some of their properties. This idea could be the basis of a new type of biologically oriented computation, which would not be program based. Because Autopoietic systems do not have inputs or outputs, only a circular dynamic which is perturbed but not defined by external agents, it is not possible to encode outside concepts into Autopoietic states, nor to control a trajectory of states (like Turing machines). Thus an external observer can only define a computation for an Autopoietic system as the particular ontogeny for that system. During the system's ontogeny, a relation between it and its medium is *selected* or *stabilized*. This relation has meaning, in the sense of the *Umwelt*, for the Autopoietic system, which is structurally coupled to its medium, but not for external observers. Thus external observers, if they wish to use Autopoietic systems to perform computations, must find a procedure to attach meaning to particular moments and properties of the system's ontogeny (Letelier et al., 2002).

Considering Autopoietic systems as a subset of (*M,R*) systems raises questions such as: (a) how to map the concept of Φ onto the formalism of Autopoietic systems; (b) how the notions of structure and organization of an Autopoietic system map onto (*M,R*) systems. Furthermore, Φ is a conflicting notion, as it is a functional component that has the function of selecting one realization of metabolism among many possible instances. Φ is a functional component (i.e. a relational entity) that does not map 1–1 onto physical entities. As Autopoiesis begins from the notion of circular organization (it does not try to prove how it can arise), Φ

appears, in a first approximation, as an alien concept to Autopoiesis. But a consideration of the very definition of an Autopoietic system ("...[an Autopoietic system] is a network of processes configured in a....") suggests a possible identification for Φ. From the point of view of Autopoiesis, Φ could be thought of as the configuration of processes that establishes a circular, time-invariant network. It is exactly this configuration that moment by moment drives the metabolism along the lines of circularity and stability. Thus a profound identification can be found between Φ and the "metabolic network or mesh" that defines an Autopoietic system. The clarification of this link is, without a doubt, one of the issues that must be addressed first.

The conclusion that Autopoietic systems are a subset of (M,R) *systems* could be rather surprising for some. As Autopoiesis is the direct offspring of a very general viewpoint of living systems, where the nature of the act of observing is the fundamental step, its viewpoint seems more general than the objectivistic vantage of early system analysis, which is the (implicit) foundation of (M,R) *systems*. Thus it could appear that (M,R) *systems* are more specific, and less general, than Autopoietic systems. We claim that, in this respect, it is important to distinguish between the *epistemological frameworks* of the models developed by Rosen, on one hand, and Maturana and Varela, on the other, and the *systems* (Autopoietic or (M,R)) that these two epistemologies bring forth. In this paper, we avoided the interesting point of comparing these epistemologies an effort that should be undertaken. Instead we focused on the more restricted domain of how these two classes of systems are related. Because Autopoietic systems incorporate, in the notion of separation from the environment, the discrete nature of living systems, we conclude that Autopoietic systems, although less general than (M,R) *systems*, capture the essential points of living organization as they couple circularity to discreteness. Thus, we have a dual situation in which *the epistemological framework* of Autopoiesis is more general than (M,R) *systems* (for example only one type of object exists in Autopoietic systems (called components) while transformable materials, components and repair subsystems compose (M,R) systems), but operationally an (M,R) system is more general than an Autopoietic system as it has only the property of circular cellular organization and lacks spatial confinement.

We consider that our principal contribution is connecting the two, until now, disconnected theories and using this link to prove that Autopoietic systems are not Turing-computable. The Autopoietic model has gained the possibility of using the theory of categories to describe the complex networks of processes that constitute a real metabolism. On the other hand (M,R) systems, along with QAPs and *SEM*, can benefit or be complemented by the more biologically oriented *Weltanschauung* of Autopoiesis revealed through its application to problems like the origin of life (Mavelli and Luisi, 1996), evolution (Maturana and Mpodozis, 2000) or neurobiology (Mpodozis et al., 1995). The demonstration of Autopoiesis' inclusion in (M,R) systems presented here is only the first step in a future synthesis. Besides the necessity of establishing a map between Φ and Autopoietic concepts, other very difficult points need to be addressed, specially how to incorporate the notion of the observer in defining "objects" that are not independent from the observing act.

Acknowledgements

References

Boden, M., 1999. Is metabolism necessary? Br. J. Philos. Sci. 50 (2), 231–249.

Casti, J., 1988. The theory of metabolism–repair systems. Appl. Math. Comput. 28, 113–154.

Casti, J., 1997. Reality Rules: II. Wiley-Interscience, Wiley, New York.

Casti, J., 2002. "Biologizing" control theory: how to make a control system come alive. Complexity 7, 10–13.

Ehresmann, A.C., Vanbremeersch, J.P., 1987. Hierarchical evolutive systems. Bull. Math. Biol. 49 (1), 15–50.

Eigen, M., 1971. Self-organization of matter and the evolution of biological macromolecules. Naturwissenschaften 10.

Eilenberg, S., MacLane, S., 1945. General theory of natural equivalences. Trans. Am. Math. Soc. 58, 231–294.

Fleischaker, G.R., 1988. Autopoiesis: system logic and origins of life. Ph.D. Thesis, Boston University.

Fleischaker, G.R., 1990. Origins of life an operational definition. Orig. Life Evol. Biosp. 20, 127–137.

Heschl, A., 1990. L=C, A simple equation with astonishing consequences. J. Theor. Biol. 145, 13–40.

Kalmykov, V.L., 1998. The Generalized Theory of Life. Third German Workshop on Artificial Life, Verlag Harri Deutsch.

Kampis, G., 1991. Self-modifying Systems in Biology and Cognitive Science. Pergamon Press, Oxford.

Kauffman, S., 1993. The Origins of Order: Self Organization and Selection in Evolution. Oxford University Press, Oxford.

Langton, C., 1989. Artificial Life: SFI Studies in the Science of Complexity. Addison-Wesley, California.

Letelier, J.C., Marín, G., Mpodozis, J., 2002. Computing with autopoietic systems. In: Roy, R., Koppen, M., Ovaska, S., Furuhashi, T., Hoffmann, F. (Eds.), Soft Computing and Industry Applications. Springer, London.

Limone, A., 1977. L'autopoiese dans les organisations. Doctorat de Troisieme cycle. Université de Paris, Paris.

Luhmann, N., 1982. The world society as a social system. Int. J. Gen. Systems 8, 131–138.

Maturana, H., 1970a. Biology of cognition. Biological Computer Laboratory, BCL Report 9, University of Illinois, Urbana.

Maturana, H., 1970b. The neurophysiology of cognition. In: Garvin, P. (Ed.), Cognition, A Multiple View. Spartan Books, New York.

Maturana, H., 1987. Everything is said by an observer. In: Thompson, W. (Ed.), Gaia: A Way of Knowing. Lindisfarne Press, San Francisco.

Maturana, H., Mpodozis, J., 1992. Origen de las especies por medio de la deriva natural. Dirección de Archivos y Museos, Santiago.

Maturana, H., Mpodozis, J., 1999. De l'origine des especes par voie de la dérive naturelle. Presses Universitaires de Lyon, Lyon.

Maturana, H., Mpodozis, J., 2000. The origin of species by means of natural drift. Rev. Chil. Hist. Nat. 73, 261–310.

Maturana, H., Varela, F., 1972. De Máquinas y Seres vivos. Editorial Universitaria. Universidad de Chile, Santiago.

Maturana, H., Varela, F., 1975. Autopoietic system. A characterization of the living organization. Biological Computer Laboratory, Urbana.

Maturana, H., Varela, F., 1980. Autopoiesis and Cognition: The Realization of the Living. Reidel, Dordrecht.

Mavelli, F., Luisi, P.L., 1996. Autopoietic self-reproducing versicles: a simplified kinetic model. J. Phys. Chem. 100, 16600–16607.

McMullin, B., 1999. Some remarks on autocatalysis and autopoiesis. Emergent Organizations and their Dynamics. University of Ghent, Belgium.

McMullin, B., Varela, F., 1997. Rediscovering computational autopoiesis. European Conference Artificial Life, Brighton.

Mingers, J., 1995. Self-producing Systems: Implications and Applications of Autopoiesis. Plenum Press, New York.

Mpodozis, J., Letelier, J.C., Maturana, H., 1995. Nervous System as a Closed Neuronal Network: Behavioral and Cognitive Consequences, Lecture Notes in Computer Science, Vol. 930, pp. 130–136.

Nomura, T., 1997. An attempt for description of quasi-autopoietic systems using metabolism-repair systems. Fourth European Conference on Artificial Life. MIT Press, Cambridge, MA.

Rado, T., 1962. On non-computable functions. Bell Syst. Tech. J. 41 (3), 877–884.

Rashevsky, N., 1938. Mathematical Biophysics. University of Chicago Press, Chicago.

Rashevsky, N., 1954. Topology and life: in search of general mathematical principles in biology and sociology. Bull. Math. Biophys. 16, 317–348.

Rosen, R., 1958a. A relational theory of biological systems. Bull. Math. Biophys. 20, 245–341.

Rosen, R., 1958b. The representation of biological systems from the standpoint of the theory of categories. Bull. Math. Biophys. 20, 317–341.

Rosen, R., 1959. A relational theory of biological systems II. Bull Math. Biophys. 21, 109–128.

Rosen, R., 1964. Abstract biological systems as sequential machines II: strong connectedness and reversibility. Bull. Math. Biophys. 26, 239–246.

Rosen, R., 1966. Abstract biological systems as sequential machines: III. Some algebraic aspects. Bull. Math. Biophys. 28, 141–148.

Rosen, R., 1972. Some relational cell models: the metabolism-repair system. In: Rosen, R. (Ed.), Foundations of Mathematical Biology. Academic Press, New York.

Rosen, R., 1985. Anticipatory Systems. Pergamon, Oxford.

Rosen, R., 1991. Life Itself. Columbia University Press, New York.

Rosen, R., 2000. Essays on Life Itself. Columbia University Press, New York.

Rosenbluth, A., Wiener, N., Bigelow, J., 1943. Behavior, purpose and teleology. Philos. Sci. 10, 18–24.

Ruiz-del-Solar, J., Köppen, M., 1999. Autopoiesis and Image Processing. In: Mohammadian, M. (Ed.), New Frontiers in Computational Intelligence and its Applications. IOS Press, Vienna, Austria.

Sharov, A.A., 1991. Self-reproducing systems: structure, niche relations and evolution. Biosystems 25, 237–249.

Snyder, E., 1999. Is freud's model of the mind Autopoietic? Ger. Rev. 74, 67–78.

Spencer-Brown, G., 1969. Laws of Form. George Allen and Unwin, London.

Szostak, J.W., Bartel, D.P., Luisi, P.L., 2001. Synthesizing life. Nature 409 (6918), 387–390.

Varela, F., 1979. Principles of Biological Autonomy. Elsevier, North-Holland, New York.

Varela, F., 1987. Laying down a path in walking. In: Thompson, W. (Ed.), Gaia: A Way of Knowing. Lindisfarne Press, San Francisco.

Varela, F., Maturana, H., Uribe, R., 1974. Autopoiesis: the organization of living systems, its characterization and a model. Biosystems 5 (4), 187–196.

von Bertalanffy, L., 1950. An outline of general system theory. Br. J. Philos. Sci. 1, 134–165.

Weber, A., 2001. The 'surplus' of meaning. Biosemiotics aspects in Francisco J. Varela philosophy of cognition. Cybernet. Hum. Knowing 9, 11–29.

Wolkenhauer, O., 2001. Systems biology: the reincarnation of systems theory applied to biology? Briefings Bioinform. 2, 258–270.

Zaretzky, A.N., 2000. Review: the algebraic relational theory and its applications. J. Biol. Systems 8 (3), 279–317.

Zeleny, M., 1981. Autopoiesis: A Theory of Living Organization. Elsevier, North-Holland, New York.

Zeleny, M., Hufford, K., 1992. The application of autopoiesis in systems analysis: are autopoietic systems also social systems? Int. J. Gen. Systems 21, 145–160.