Autopoietic Systems: A Generalized Explanatory Approach – Part 1

Hugo Urrestarazu • Independent Researcher • hugo.urrestarazu/at/wanadoo.fr

> Context • This paper is intended for readers familiar with Humberto Maturana's theory of autopoietic systems and with the still unresolved debate concerning the existence of non-biological autopoietic systems. Because the seminal work of the Chilean biologist has not yet been fully and correctly understood in other disciplines, I consider that it is necessary to offer a more generalized concept of the autopoietic system, derived by implication from Maturana's grounding definition. **> Problem** • The above-mentioned debate is rooted in a deficient application of some rigorous distinctions, definitions, and epistemological considerations introduced by Maturana when he coined the term "autopoiesis." Some researchers think that social or economic organizations could be considered as autopoietic systems of a higher order because they appear to behave autonomously and be self-organized and self-producing. However, in practice some precise distinctions would need to be verified through observation in order to claim properly their autopoietic nature. These distinctions were defined by Varela, Maturana and Uribe in 1974 as a set of six decisional rules ("MV&U rules") whereby an observer may possibly justify this stand. My aim is to pinpoint clearly the basic cognitive tasks that an observer should perform in order to ascertain such a claim. **> Method** • I accomplish this with a thorough analysis of the entailments derived from each *rule* when applied to the most general case – when the observational domain where the system manifests itself is not specified. A bottom-up approach is used to avoid referring to "apparent autopoietic behavior" as a starting point distinction (top-down approach): the aim is to distinguish "autopoietic behavior" as an *outcome* of more basic distinctions, not as a *premise* for these. These may be used as abstract tools to facilitate a rigorous description of observations and lead to precise explanations of the emergence of complex self-generated dynamic systems. In theory, this conceptual frame is not limited to the macro-molecular domain and may therefore be applicable to non-biological systems. **> Results** • According to *MV&U rules*, the most important distinctions are those that refer to *intra-boundary phenomenology*: this focus is necessary to explain how the key processes involved in the emergence of autopoiesis actually manifest themselves. These explanations are crucial to validating a claim about the "autopoietic nature" of an observed system. **> Implications** • This work could help multidisciplinary researchers to apply properly the theory of autopoietic systems beyond the realm of biology and to settle ongoing debates. It could also help investigations related to the specifications of software simulation processes for modeling a minimal artificial autopoietic system. However, the rigorous focus on the role of *intra-boundary phenomenology* and *self-production of components* reveals that our chances of detecting "natural" meta-molecular autopoietic systems are scarce.
> Key words • Autopoiesis, self-production, autonomy, causation structure, artificial life, organizational knowledge.

Introduction

This paper is intended for readers familiar with Humberto Maturana's theory of autopoietic systems[1] and with the still unresolved debate concerning the existence of non-biological autopoietic systems.

The original authors of the concept of autopoiesis addressed the biological domain exclusively, and they described the concept using rigorous distinctions, definitions, and epistemological considerations. In my opinion, the continuing debate derives from poor and incomplete application of these rigorous elements by those who seek to apply the concept to non-biological domains. Non-biological domains[2] could be the domains of existence of social systems, economic organizations, or artificial life simulation systems created by means of software engineering or robotics developments, for example. Some researchers in various disciplines feel that autopoiesis should be domain-free rather than domain-specific – regardless of Maturana's and Varela's opinions – and that only then could the concept of autopoiesis serve its specifics in different domains.

In order to refer to a generalized notion of a realm in which autopoietic systems could be observed and identified as such, I use the term "observational domain." Throughout this paper, the term "observational domain,"[3] refers to any domain of perceived phenomena in which we, as hu-

1| The label "theory of autopoietic systems" is meant to refer to Maturana's theoretical work on the organization of living systems. For a brief overview of the basic concepts see Maturana's historical account (2002: 5–10) and the introductory sections 1.1 and 1.2 of Varela's (1992) essay.

2| What we call "domains" are social human constructs. In nature we observe the manifestation of an undifferentiated continuum of interlayered molecular, chemical, physical, biological, behavioral, and social phenomena.

3| The meaning ascribed in this paper to the term "observational," as derived from terms such as "observing" and "observer," is meant to corman observers, can perform sensorial and operational experiences through interactions with the observable entities under consideration. It is assumed that any observational domain is a physical domain and that ordinary physical laws are applicable in it. These interactions should occur as phenomena pertaining to our own biological domain (i.e., phenomena of our own direct sensory-motor perception capabilities) or as indirect biological effects resulting from the use of suitable transducers (measurement apparatus). Hence, the attention is focused on the occurrence of humanly observable and describable events occurring in what appears to us as the ordinary physical space.

The identification of autopoietic systems by observers of phenomena occurring in a given *observational domain* is an abstract exercise that needs to refer to the notions and explanations introduced by Maturana when he coined the term *autopoiesis*. In the following I assume that the reader is familiar with the general concept of autopoiesis as an approach to explaining what distinguishes living beings from other dynamic systems observed in nature. My aim is to go beyond the original purpose of explaining autopoiesis for biological systems by broaching a generalization of the concept in order to make explanations applicable to different kinds of observational domains.

This is a debatable subject and for years Maturana has been extremely cautious of pronouncing himself on the matter of deciding whether other observational domains do exist in which autopoietic systems could be observable. More recently, he concluded in a negative way by stating: "The molecular space is peculiar [...] I claim that neither the elements of the sub-molecular nor the elements of the supra-molecular domains can by themselves give rise to autopoietic systems," Maturana (2002: 14). Nevertheless, the concept lends itself to an explicit generalization, provided that we avoid explanations relying on the nature of specific interaction mechanisms between components. I will show that even the peculiarities of the molecular domain can be stated in abstract general terms, regardless of their possible applicability. My basic claim is that it is possible to explain *what autopoiesis is about* without explaining *how* this emergent property results from a particular interaction mechanism prevailing in a given observational domain. To do this I focus on the most common aspect of all possible types of dynamic interactions, namely on their *causation structure*, but not on explicit descriptions of particular physical interactions that link system components together.

The problem

The theory of autopoietic systems has been stated in terms that can be interpreted as encompassing a general *class of autopoietic systems* in which biological organisms may be considered as a particular case. In theory, an autopoietic system constitutes a particular case of dynamic system in which some observational distinctions need to be verified in order to claim properly its autopoietic nature. These distinctions were defined by Maturana, Varela and Uribe as a set of decisional rules (Varela et al. 1974; hereafter "VM&U rules"), whereby an observer may provide the answers to six basic questions and possibly justify a positive claim by means of descriptions of observational experiences performed in the intersection of an observational domain with the system's *domain of existence*.[4] Inspired from observed biological macromolecular phenomenology, these rules were intended to characterize living beings in terms of described phenomena occurring in molecular dynamic systems. In fact, these proposed distinctions overtake the scope of the molecular domain and are theoretically applicable to other "meta-molecular"[5] domains (i.e., where the observable entities need not be described at their molecular level of composition).

However, many difficulties have been encountered in practice when applying these rules to other particular domains in which the existence of autopoietic systems has been suspected. This is especially the case for the social or economical domains, for example. Discussions often focus on the question of identifying the boundaries of purported autopoietic systems as a way of discriminating the observed composite unities from their background and deciding if they are indeed members of a particular class of entities showing autopoietic behavior. I show that the identification of boundaries is a necessary but not sufficient step.

By proposing a bottom-up explanatory approach based on some elementary distinctions, definitions, and conceptual tools, I intend to encourage the reader to consider the entailments of imposing compliance with all six decisional rules, taken one by one, on the most general case, where the nature of the observational domain is not specified. By "not specified" I mean that the physical mechanisms involved in the emergence of dynamic systems in an observational domain are either not known, or are voluntarily not referred to (for practical or theoretical reasons) although they are assumed to exist.

Contents

Due to space limitations this paper has been organized into three parts that are not completely independent but are published separately. Parts II and III refer to

respond to the notions explained by Maturana (1988a) in his paper "Ontology of Observing."

4| The "domain of existence" of an observed dynamic entity does not necessarily coincide with the observational domain in which the observer interacts with it because a system may operate through activities that do not involve biological interactions with the observer at a given moment. In this sense, the observational domain is restricted to *phenomena that the observer has already experienced by interacting with the observed entity.* The observer may "see" only a portion of the entity and of its operations because it could well happen that the observational domain that the observer is able to deploy could not be extended to the whole domain of existence of an observed entity. Some entities would thus remain only partially observable.

5| By *"meta-molecular" domain* I mean any physical domain in which the dynamic entities being considered by the observer are not distinguished or described as individual molecules (taken in the usual sense employed in physics, chemistry, and/or biochemistry). The term "meta-molecular" connotes the notion of "more than simply molecular" and is used to signify that the observed interactions occurring *between* dynamical entities may be explained without referring explicitly to a ground level molecular phenomenology.

terminology, concepts, and developments presented in the previous parts. Above all, Part I is essential to understanding the terminology used in the following parts and sections.

The present paper corresponds to Part I, in which I discuss briefly the applicability of the VM&U rules to generalized meta-molecular domains, present the basic definitions and conceptual tools used in a bottom-up explanatory approach to understanding dynamic systems, and present an overview of the six VM&U rules (viewed as a validation test to qualify a system as autopoietic), explained and interpreted from within this explanatory approach.

Part II is an expansion of the above-mentioned explanatory approach, applied to the explication of the natural emergence of composite self-organized dynamic systems endowed with self-produced embodied boundaries. This is followed by an analysis of the requirements imposed on the "intra-boundaries" phenomenology for compliance with the self-production capabilities that autopoietic systems should show, according to Varela, Maturana and Uribe. Special attention is given to the consequences derived from the application of the 5th and 6th component "production rules" with respect to the topology of the interaction structures resulting from the components' coupling activity.

Part III treats several issues: the "scale of description problem" posed when trying to identify a suspected autopoietic system; the relation between the "intra-boundaries" phenomenology and the fundamental notion of operational closure; and a discussion about the possibility of establishing a strict equivalence between the generalized notion of autopoietic system and the notion of a living being. These developments are followed by a "Self-critique" section, where I discuss the validity of my bottom-up explanatory approach, and a "New reflection perspectives" section with a discussion concerning the potential use of the proposed methodology to analyze existing complex systems, to design and build artificial self-organized and autonomous systems, and to attempt to overcome the debate about the purported existence of non-biological autopoietic systems.

Applicability of the VM&U rules

I will show that, given an observational domain in which a certain dynamic entity is being observed as a unity, if this entity belongs to the class of autopoietic systems it should be possible to describe it as *bounded, composite, mechanistic, self-referential, self-producing, and autonomous*. This is a compact way of expressing the *characterizing properties* that the observer should be able to distinguish and describe in order to ascertain that the observed unity is autopoietic, as implied from the "(…) six-point key for determining whether or not a given unity is autopoietic," (Varela et al. 1974: 192). I consider that the term "key" should be replaced by "validation test," because it expresses better *what the observer is expected to do* when trying to identify the nature of the observed entity. This test consists of the *six* VM&U rules that the observer should apply while observing. Each rule consists of one or more questions that should all be answered positively in order to validate a claim.

It is not my purpose to review and discuss these rules per se, but to use them as a guiding reference in the bottom-up explanatory approach to autopoiesis that I unfold in the next section and in parts II and III of this work. These rules represent in fact a very compact definition of the notion of the autopoietic machine, as it was initially conceived by Maturana four decades ago, and a full comprehension of their scope should be obtained by comparison with other developments on autopoiesis produced by Maturana and Varela ever since. According to Maturana (2002: 6–8), he developed the basic notion back in 1964, but did not coin the term "autopoiesis" until 1970. The rules sketched and published by Varela, Maturana and Uribe in 1974 were seldom explicitly cited by Maturana or Varela thereafter, even though they never stated that this definition, as implied from the rules, was flawed or inadequate. In my opinion, they just left to others the task of exploring the entailments that can be derived from their "six-point key."

The VM&U rules' "natural" domain of applicability is the biological domain, as they were purposefully devised to account for the observed phenomenon of the emergence of molecular entities showing the specific properties that characterize living beings. However, because they were formulated in general abstract terms, their applicability is in principle not restricted to a specific observational domain as the authors explicitly admitted a domain-free interpretation: "For the purpose of explaining and studying the notion of autopoiesis, however, one may take a more general view as we have done here […] where physical space is replaced by any space […] and molecules by entities endowed with some properties." (Varela et al. 1974: 191). Although in Maturana's and Varela's literature emphasis has been put mainly on physical spaces and molecular phenomenology, they nevertheless left the door open to the consideration of a "more general view," and this is precisely what I intend to do here. But I intend also to show that some of these rules impose severe restrictions on the identification of autopoietic systems in meta-molecular domains. This analysis may lead someone to modify some rules, and produce a broader version that could satisfy aspirations for a theory of autopoietic systems that was domain-free not only in principle, but also in practice. The question would arise then of whether we could still call that modified version a theory of "autopoiesis," or would need to qualify it with another name instead.

The reader should be aware that unfolding the operational meaning of the VM&U rules from within my explanatory approach inevitably constitutes a particular interpretation of them. In the final section of this paper, "The VM&U validation test," I present an overview of the rules and I explain how they relate, in my view, to the required characterizing properties mentioned above. More detailed discussions are also developed in Part II and Part III.

Bottom-up explanatory approach

In this paper I intend to avoid referring to the distinction of *autopoietic behavior* as a starting point of discussion. My aim is to introduce this distinction as an *outcome* of some preliminary explanations, *not as a premise* for them. I begin with the explanation of what I understand as those necessary *basic distinctions* that allow an observer to describe dynamic objects and their mutual interactions. From there I step into an explanation of more elaborated distinctions that allow us to describe a general composite dynamic system existing in a general non-specified observational domain.

The global aim is to explain through which basic distinctions we may bring forth more elaborated descriptions that are essential to identify a structure determined dynamic activity compliant with the rules that define an autopoietic machine in a composite system being so far unknown as a member of the class of autopoietic systems. In other words, I intend to pinpoint the necessary cognitive steps that an observer should perform in order to focus on pertinent observations that would allow him or her to describe a possible autopoietic behavior. By following this explanatory path and applying the VM&U rules to validate a positive claim, I intend to show how a very general composite dynamic system can be "recognized" by the observer – step by step – as an autopoietic system.

I impose the requirement that the resulting explanations should lead to the identification of three distinct sets of dynamic objects related to the suspected autopoietic system:

a | a "boundary" (among other possible boundaries);

b | the "body" (what is "inside this boundary"); and

c | the "medium" (what is "outside this boundary").

By following the entailments of Maturana's proposed distinctions concerning the nature of the dynamic components that constitute an autopoietic system, I show then that this task must be completed by focusing on what happens "inside the boundaries." I discuss the terms whereby observers can describe the intra-boundary phenomenology in order to explain how the key processes (such as self-organization, self-production, and self-adaptation to varying environmental circumstances) manifest themselves. These explanations are vital to revealing what is determinant in general and applicable in particular to an observed system so that observers can validate a possible claim that it is indeed autopoietic.

Ontological and epistemological considerations

We will see that in the more general case, when we tackle the problem of "detecting" or "identifying" an autopoietic system in a non-specified observational domain, the "autopoietic nature" of the system may not appear as a directly observable physical feature of the observed entity, but as a conceptual construct based on multiple observations and descriptions. Even in the case of living cells or organisms, we may encounter difficulties in physically circumscribing the unity and differentiating its structure from the background as a compact physical object.[6]

On the one hand, we need to know what it means to say that an "autopoietic system," its boundary, and its environment *exist* as physical objects and not as "views" of our minds. This is a typical ontological problem and we need to give an account of the terminology used to refer to the "real" observable entities that we call autopoietic systems and to refer to the abstract conceptual constructs that we may use as explanation tools. On the other hand, we need to clarify the relationships that the observer may establish with these observable entities and what he or she can learn from the act of observing. This is an epistemological problem that the above-mentioned conceptual constructs should provide as well.

The criteria needed to generate the above-mentioned basic distinctions are not trivial, but I claim that they can be specified in general for any observational domain, provided that the formal relations used in the *languaging domain*[7] to describe relations between dynamic objects existing in the observed domain express the occurrence of causal interactions between them, *whatever the causation mechanism may be*. I stress the issue of focusing on causality as it needs to be explicitly stated in order to choose the most appropriate abstract language capable of expressing the operations of distinction proposed in this explanatory path.

Causation and entailment relations

Before addressing the subject of the identification of a boundary, let me first discuss some basic distinctions concerning causality. I beg the reader to consider the following paragraphs as a non-trivial and necessary step to *render the notion of causality explicitly distinct from the notion of entailment* (or logical implication). This distinction is essential to avoid a common theoretical snare when appealing to abstract reasoning. The snare consists of confusing the phenomenological relationships that an observer can establish on observational grounds with logical inferences resulting from theoretical formalisms meant to model an observed phenomenology.

- *entailment* is a cognitive operation involving logical statements of the sort: "if *A* is true then *B* is true" (*A* implies *B*) for example;
- whereas *causation* is a description of observed phenomena involving statements such as: "I observe that whenever *X* happens, then *Y* happens consecutively within a finite time interval" (*X* causes *Y*).

Both previous statements can be used to define abstract relations such as $\mathcal{R}e$ (*A*, *B*) (for entailment) and $\mathcal{R}c$ (*X*,*Y*) (for causation), but the meaning of the word *then* is quite different in each case. In the first case it expresses an inference that has significance only in formal terms with respect to rules, axioms or theorems expressed within a formal mathematical system with a timeless *logical validity*. In the latter case, the word "then" can be replaced by the expression "it happens that," which

6 | A biological example is the identification of the immune system as an autopoietic unity by Varela (Varela et al. 1988); this is not spatially fixed and is not limited by a "surface" acting as boundary, such as the membrane in the cell. "The immune system is not spatially fixed, it's best understood as an emergent network" (Varela 1995: 213).

7 | The term "languaging domain" refers to the domain of human conversations in which language, conceived as a consensual coordination of consensual coordination of behaviors between humans, is the fundamental interaction used by scientific observers to express their descriptions of their experiences performed in the observational domain under consideration, to explain their operations of distinction, and to formulate their claims about phenomena observed within these domains.

refers to an *observed fact* totally unrelated to any observer's cognitive operation other than the act of observing/describing, and its validity is inter-subjectively established, circumstantial, and time dependent. This distinction between *entailment relations* and *causation relations* becomes crucial when describing reflexive and reciprocal relations between dynamical entities, because in the first case reflexivity or reciprocity relations are interpreted as timeless *recursive* functional applications leading to infinite regress, whereas in the second case, as the causation propagation time is explicitly taken into account, reflexive and reciprocal relations may be interpreted as *succession of causation events* with a restricted *time dependent phenomenological validity*, with no recursive entailment involved.

Time and causality

Causality refers to the notions of events occurring in time, of successions of such events and of observed regularities in those successions of events. In order to explicitly state what the word "time" connotes in this paper, I refer the reader to Maturana (1995), where he states that the notion of time is a notion that arises (for the observer) as an abstraction with no transcendental ontological status. For him, the word "time […] cannot refer to an entity that exists independently of what we do" as human observers.

The notion of a causal relationship linking objects that have arisen in our experience as observers refers to the observation of repeated regularities in the successions of events concerning those objects, so that we are allowed to claim that we can distinguish relations that couple those objects together with respect to our perception of their dynamic activity.

We humans act as observers of phenomena occurring in time by describing to other humans our experiences (by interacting with them in the languaging domain) while or after being involved in biological interactions with the outcomes of those observed phenomena. In order to make scientifically valid descriptions of our experiences we need to achieve what we could call "objectivity of description." This is an outcome of the coherency of our experiences as observers in which a described phenomenological experience leads to the actual realization of experiences of the same kind that can be described (in the languaging domain) in the same way by other observers. In this sense, objectivity of description does not refer to an observer independent reality, but to the coherence of the experiences of multiple observers as expressed by what they say about their own experiences. As such, the term "objectivity" connotes an "inter-subjective outcome" of our experience of living as biological systems and of our languaging activity as humans.

Now, what do we mean by phenomena occurring in time? In this paper, all that needs to be inter-subjectively said about this notion in order to explain the notion of causal relationship is synthesized by Einstein's famous (and provocative) phenomenological distinction: "time is what is measured by clocks."[8] As clocks are just extensions of our perceptive organs, conceived to help us to assign a value to the "size" of perceived intervals that "separate" observed events distinctly (for us as observers), it is clear that the word "time" cannot be invoked to connote anything independent from the observer's experience. I consider that this reference to the notion of time used by Einstein in physics is in no way problematic with respect to the ontological considerations expressed by Maturana (1995, 1988a, 1988b) concerning the notion of time. In my interpretation,[9] the expression "what is measured" could be better understood as "whatever is measured," to emphasize that in physics there is no point in asking if Einstein's "what" has a transcendental ontological status on its own or not. In this statement, Einstein purportedly bypassed this question and, independently from any personal metaphysical stand of his own, he equated "time = measure," by stating implicitly that in physics, it is irrelevant to imagine what else it could be. This measure refers to the result of a biological experience of the observer when interacting with a device (a clock). This result is expressed as a number that corresponds to a count of clock events (a cognitive action of the observer). This number is an abstraction that exists only in the languaging domain, not in the observational domain, in agreement with Maturana's considerations referred above. What we connote here by "time" arises from the observer's doing when counting repeated events, be it in language (announcing consecutive numbers, aloud or silently) or by relying on a counting device.

Phenomena occurring in time are, then, interactive events undergone by observers who are able to discriminate their occurrence in time by assigning them a measure with the help of a clock. The fact that our spontaneous perception of the occurrence of observable phenomena allows us to discriminate sequentially separated events without using an instrument and talk about what occurred before, simultaneously or after the occurrence of a referential event (because we are endowed with memory), is a cognitive ability of our species. This ability is efficient only within a biologically determined range of values. The use of a clock is a reinforcement needed to generate scientifically valid explanations in the languaging domain, especially for observational domains in which our natural perception is unable to discriminate successive events separated by extremely short time intervals.

8| Einstein's statement was completed by Lamaître: "[…] if a number of experiments are started and finished together, then the same experiments repeated elsewhere and started together must finish together. Any instrument that repeats endlessly the same experiments and counts their number is a clock" (Georges Lemaître, letter to Albert Einstein, 1947, quoted by Jean-Pierre Luminet (2004), my translation). The statement "must finish together" refers to an accepted "fact of nature" acknowledged by experiments performed by scientific observers. I mean by "fact of nature" an observable regularity agreed upon by scientific observers to be considered as a physical law: in the context of Lemaître's remark, he refers to an invariant of the general theory of relativity.

9| I acknowledge that Einstein also posed that time is a fourth dimension, mathematically similar to the other three spatial dimensions, proved to be relative in its quantity to the velocity of observers moving in physical space. My interpretation is based on the opinion that these notions of spatial/temporal dimension are all abstractions of the same kind, treated in an abstract mathematical model, and thereby they all pertain to the languaging domain of human observers. To discuss these considerations further is out of the scope of this paper.

Cause-effect relationships

The term "causally linked phenomena" refers to an *association established by observers* between events occurring regularly as a *consequence* of other events that occurred before and that do not occur if those *causing* events do not occur. This is a one-way "cause-effect relationship" with respect to the occurrence of "causing events" and of "effect events" as discriminated by an increasing counter of "clock events."

My purpose is to describe observable phenomena produced by the activity of dynamic entities, that is to say, of observable objects in which we can distinguish *changes* and consider those changes as *events* occurring in time. Those entities exist in an observational domain and are subject to changes as *a consequence of their mutual interactions*. These changes are considered as the causal outcome of those interactions.

The phenomenon under investigation is the "emergence of interaction structures" giving rise to the observation of groups of interacting entities that can be distinguished as being mutually associated by *cause-effect couplings* while other entities existing in the same observational domain are not associated in this way.

The observer's "point of view"

When observing a population of dynamical objects and differentiating a composite unity of such objects from the background, the observer needs to define some *inclusion criteria* in order to distinguish the said composition phenomenon. Of course, in any theoretical approach, the *inclusion criteria* – used by the observer to decide whether some entities are to be considered as components of a system or not – *define the boundaries of the referred system* independently from any observed "(objective) reality." Stated in this way, the operations of distinction should be understood as cognitive statements made *arbitrarily* by an observer. These proposed distinctions may facilitate or prevent the observation of an autopoietic system in action, and this is essential to judge on their *pertinence*. Furthermore, such inclusion criteria are not imposed *a priori* by the nature of a particular observational domain.

Each possible set of inclusion criteria constitutes a *hypothetical approach* giving rise to interpretations of observable events that can be more or less useful for "detecting" the existence of a suspected autopoietic unity in the observed observational domain. In this sense, *the* "boundary" *exists only in the languaging domain:* it is the "observer's point of view," allowing him or her to talk about the existence of an observable composite unity, seen as clearly "distinguishable" from its environment and showing a possible autopoietic behavior in the considered observational domain.

We need to keep in mind that some inclusion criteria may prevent observers from identifying such a composite unity as an autopoietic system at all. This is an important consideration because phenomenological descriptions depend strongly on those aspects of a phenomenology upon which we focus attention as observers. Therefore, the choice of *pertinent inclusion criteria* and their corresponding observation protocols is at the heart of the problem of identification of autopoietic behavior.

Nevertheless, when the observed phenomenology allows us to describe the *emergence of spontaneously generated structures of components involved in distinguishable mutual interactions* – distinguishable from the interactions with other dynamical objects – we can say that it is the phenomenon of emergence itself that "defines" the boundary (for an observer).

This is an outcome of considering the observed emerged unity as a structure determined system. The particular structure of the system arises (for an observer) only when the *attention of the observer is focused on a particular type of interaction* occurring as an observed phenomenon in the system's domain of existence.

This might sound in contradiction with the observer's freedom to choose any set of inclusion criteria leading to the identification of a composite unity, of its boundary, and of its environment. But the contradiction is only apparent because the observer's operations of distinction are just hypothetical approaches expressed *in language* and what the observer says about the observations performed on the system/environment interactions are descriptions of the observer's experiences *in the observational domain* in which the system behaves as a unity. When observations do not match with expectations derived from a particular "point of view," the observer may always change it in order to search for a better match.

Embodiment

It is the confrontation of the theoretical operations of distinctions with the description of observational experiences that allows the observer to claim that there is a formal correspondence between the proposed distinctions and the descriptions of actual observational experiences. If the observer is not able to establish such a correspondence, he or she is free to modify the proposed inclusion criteria until a better match can be claimed. When a formal correspondence can be claimed and other observers agree to say that the claim is valid, "objectivity in parenthesis" (Maturana 1988a) allows us to say that the observed system (objectively) defines its structurally determined emerged boundaries.

Briefly, I will show that in the most general case, "causal coupling relations" are the base level abstractions necessary to describe a proposed theoretical "boundary" for the system. However, in the practice of observing a particular domain, the consideration of a *particular type of causation mechanism* allows the observer to define specific observational protocols devised to verify that this theoretical "boundary" *coincides* with a particular set of physical components where each member is distinguished by the causal coupling relations in which it participates, based solely on the operation of the considered causation mechanism in the domain where the system arises as such (to the observer). When this is the case we can talk about an "embodied boundary."

Explanations in the biological domain

Let us see how these considerations apply in the domain for which Maturana and Varela developed their theory of autopoiesis, namely the biological domain. Here, living systems are observed as biochemical dynamic systems capable of conserving their organization in continuous interaction with their biochemical environment. Maturana proposed that:

> "[…] living systems are dynamic systems constituted as autonomous unities through being closed circular concatenations (closed networks) of mo-

lecular productions in which the different kinds of molecules that composed them participated in the production of each other, and in which everything can change except the closed circularity of the concatenation of molecular productions that constitutes them as unities [...] Francisco Varela and I expanded this characterization of living systems by saying: first, that a composite unity whose organization can be described as a closed network of productions of components that through their interactions constitute the network of productions that produce them and specify its extension by constituting its boundaries in their domain of existence, is an autopoietic system; and second, that a living system is an autopoietic system whose components are molecules. Or, in other words, we proposed that living systems are molecular autopoietic systems and that as such they exist in the molecular space as closed networks of molecular productions that specify their own limits.” (Maturana 1988a)

The type of interactions considered in the macromolecular domain and observed in the physical space are those established between complex molecules giving rise to more or less stable macromolecular structures participating in the dynamics of the system that they constitute. These interactions are explained in terms of chemical couplings in which the spatial closeness and chemical affinity of the interacting entities (macromolecules) play a fundamental role. In physical terms, interactions occur according to the manifestation of electrochemical attractive or repulsive forces subject to three-dimensional Euclidean constraints. These constraints may be expressed in terms of distances between the interacting entities and of the mutual three-dimensional orientation of their spatial structure.

The resulting structure of a composite system of interacting macromolecules is also a spatial structure occupying a volume in the *physical* space. Such structures have “natural boundaries” that can be described mathematically in terms of topological surfaces distinguished in an *abstract* three-dimensional “Euclidean space.” For example, the cell membrane is the embodied locus of the boundary for the most elementary structure of living beings: the living cell described as an autopoietic system. This boundary is (objectively) produced by the cell itself.

However, the distinctions made in the biochemical domain (i.e., the identification of the basic dynamic entities, called “macromolecules”) need to be stated with caution. For example, the scale of the observation is essential to bring forth the appropriate structures to be observed and described as components. When considering biochemical interactions, everything occurring in this observational domain is a succession of multiple biochemical couplings continuously linking molecules to other molecules. If the scale of observation is not chosen to be large enough to allow for the observation of the emergence of large scale structures, the identification of structure determined boundaries might become impossible. At a low scale of observation, a cellular membrane would be seen as a sieve and interacting molecules could not be classified as being part of an “inside” or of an “outside” volume. On the other hand, if the scale of observation is too large, the observer will not be able to observe interactions between basic components – only between large scale composite structures – nor explain the emergence of an autopoietic unity as a result of underlying interaction mechanisms.

These examples of operations of distinction that depend on the scale of observation show how an observer can miss the identification of an autopoietic system and its boundary altogether or be unable to claim an explanatory path for the emergence of such a composite unity.

Explanations in a non-specified observational domain

In order to test the applicability of Maturana’s general definition of autopoietic systems in other observational domains, many authors have discussed the existence and the identification of such composite unities. This is especially the case with sociologists or economists, who consider the idea that some observed organizations perceived as possessing some traits of self-rule could be described as autopoietic entities existing in a social, political or economical domain.[10] When discussing the problem of identifying the *boundary* of such systems, a preliminary question arises: How should we distinguish the agents that constitute the system’s components? In other words, what inclusion criteria should be applied in order to differentiate the system components from other agents existing in the same observational domain?

We will see that this question finds a natural answer when we focus our analysis on the natural emergence of causation networks, conceived as connected paths for the propagation of cause-effect events among a population of interacting dynamical entities. That autopoiesis is not an outcome of the internal properties of components but of the *dynamic relations* established between them is *implicit* in the definition of an autopoietic system compliant with the VM&U rules, though it has been repeatedly argued by Maturana and Varela. This consideration is essential to argue in favor of the need to stress the point *explicitly* in the *definition of the notion of relation* itself.

I intend to show that in the case of a general (i.e., non-specified) observational domain, any pertinent operation of distinction made by an observer in order to discriminate the boundary of a suspected autopoietic system should rely only on the notion of *causal propagation* occurring among dynamic objects interacting in that domain. I focus on *pure causation flow* because *it is the only common property left* when we generalize the observational domain. If autopoietic behavior should be distinguishable and explainable in *any* observational domain, *the specific prevailing causation mechanism should not be part of the explanation*. The description of the nature of the underlying mechanism responsible for such interactions is needed only to verify the pertinence of the operation of distinction by carrying out observations in a particular observational domain.

10| As my purpose is to explain autopoiesis in general terms and not to discuss the applicability of autopoietic theory to social/political/economical systems, I refer the reader to some authors that discuss the autonomous nature of these systems and the role played by some basic notions used in autopoietic theory in the attempts to describe them (Beer 1972, 1980, 1995; Schwember 1977; Dupuy 1982; Luhmann 1986; Robb 1989; Zeleny & Hufford 1992).

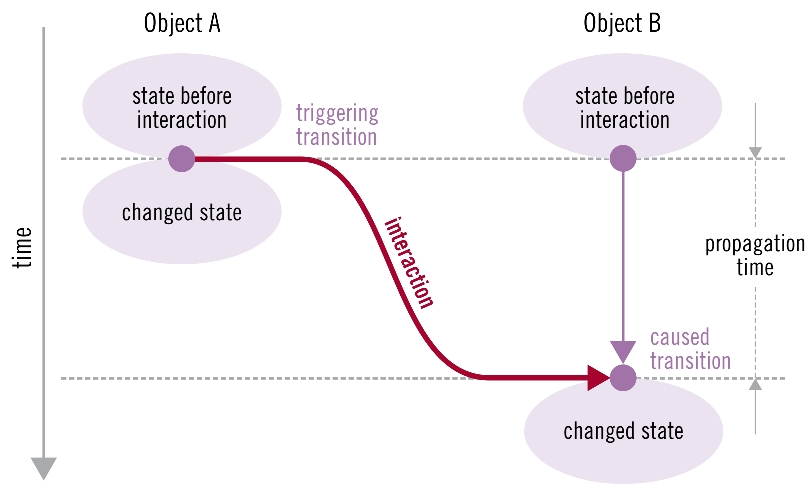


Figure 1: Observer's description of a cause-effect interaction. This figure is an abstract schematic representation of a succession of events occurring in time (increasing time is represented by the left top-down oriented vertical arrow) as described by an observer. These events involve two observed objects *A* and *B*. The interactions between the observer and the objects are not depicted, but are assumed to have already taken place. This graphical representation is to be taken as an observer's description of the occurrence of correlated changes in both objects. This correlation is represented by the arrow called "Interaction." Note that an interaction occurs within a non null time interval. The triggering transition in object *A* occurs within a non null time interval, too, but the figure depicts the event as occurring when the observer perceives it. The interaction "Propagation time" is the time needed for the effect to manifest itself on object *B*. The latter depends on the interaction propagation speed, the distance separating the objects in physical space, and the caused transition delay in object *B* (reaction time). The figure depicts the caused transition event in object *B* as occurring when the observer perceives it. All these time intervals depend on the specific underlying interaction mechanism and the nature of the dynamic objects under observation. The occurrence of a transition in object *A* and its resulting arrival state is supposed to be absolutely independent of the starting state of object B (we are dealing with a one-way interaction only). The arrival state of object *B* may or may not depend on the starting state of object *A*, on its arrival state, or on both.

Cause-effect coupling

The notion of cause-effect coupling between dynamic objects needs to be defined explicitly in order to be able to describe the emergence of interaction structures in general, i.e., without referring explicitly to the interaction mechanism that provokes the coupling.

Definitions and conceptual tools

In order to make this notion explicit, let me propose some brief definitions and statements that will be explained or justified later with more details:

1 | I define *dynamic objects* as entities, agents, components (or whatever we want to name the dynamic objects existing in an observational domain) that are capable of performing *changes of state.*

2 | I mean by the *state* of a dynamic object the *set of values ascribed to a number of variables*[11] or features *that the observer can associate* with the object at a given time. The chosen variables, called *state variables,* should express quantifiable features that are subject to changes in the object's dynamics. The values of the state variables should be describable and observable by the community of observers at any chosen time.

3 | A *state transition* (or change of state) of a dynamic object occurs when at least one state variable value changes in time.

4 | We distinguish a *cause-effect coupling* between two dynamic objects *A* and *B if a state transition in object A* (triggering or causing transition in *A*) *is the condition sine qua non*[12] *of a state transition in object B* (a triggered transition in *B*). An *interaction* between dynamic objects is a *cause-effect coupling* that links them for us as observers of the evolution of their history of state transitions in time (see Figure 1).

5 | By definition, in a non-specified observational domain, the *mechanism responsible for cause-effect coupling* between dynamic objects is not specified, but *it is assumed to exist as a causal influence of one object on another that the observer can distinguish as occurring in time* (i.e., in any given (specified) observational domain, it would correspond to an observable phenomenon).

6 | The latter means that: (a) in the absence of any other cause-effect coupling between *B* and any other object different from *A*, the caused transition in *B never*[13] *occurs before* the occurrence of the triggering transition in *A*; *and* (b) *whenever* objects *A* and *B* are in a specified

11 | A state does not need to be defined as an analytical function of the independent variables (in any case this is irrelevant for our explanatory purpose); it is just an identifier for a specific combination of values of the variables at a given time. The only feature retained here is that a state represented by the *set of values* of those variables at a given time should be distinguishable from a state represented by *another set of values for the same variables* at a later time. If this happens, the occurrence of this variation is called a state transition.

12 | It should be noted that in the most general case, to claim that an observed event *A* is the "condition sine qua non" (indispensable condition) for observing an event *B* is inevitably a hypothetical assertion, because observers could well be misled to associate their occurrences even if they are in fact unrelated.

13 | This is a strictly deterministic definition. We could have said "occurs with a nearly 0 probability" instead of "never occurs."

state and a specified triggering transition in *A* occurs, the *same* caused transition *always*[14] *occurs* in *B*, *within a finite time interval.*

7 | A cause-effect coupling can be *represented formally as an oriented relation* in a abstract multi-dimensional relational space (linking relational nodes with as many dimensions as the number of variables considered in the associated state variables' set)

8 | We can define a general dynamic system as a *set of dynamic objects* linked together via cause-effect coupling (i.e., causal interaction). This set can be represented by a *network of oriented relations* existing in the multi-dimensional relational space where the relational nodes' dimension is defined by the cardinal number of the state variables set associated to the linked objects.

Physical constraints

Up to this point it is important to note that these definitions are *applicable to any kind of interaction* because nothing is said about the mechanism responsible for the production of transition events allowing observers to attribute a cause-effect coupling between dynamic objects. Nothing is said either concerning the nature of the "influence" of an object upon another, nor about the nature of the dynamic objects themselves.

Dynamic objects could be embodied by physical particles, molecules, cells, organisms, individuals, social organizations or even software processes, for example. The interactions could be embodied by physical forces, physical signals or *any kind of mechanism capable of triggering changes of state in a dynamic object that is "sensitive" to it*. The only thing being said is that the interaction is causal, in the terms defined above. I voluntarily neglect other considerations concerning some physical constraints necessary for an interaction to occur (i.e., energy transfer, for example), not because I believe that a physical interaction could occur without being subjected to those constraints, but because this consideration is not relevant for the explanatory path developed below.

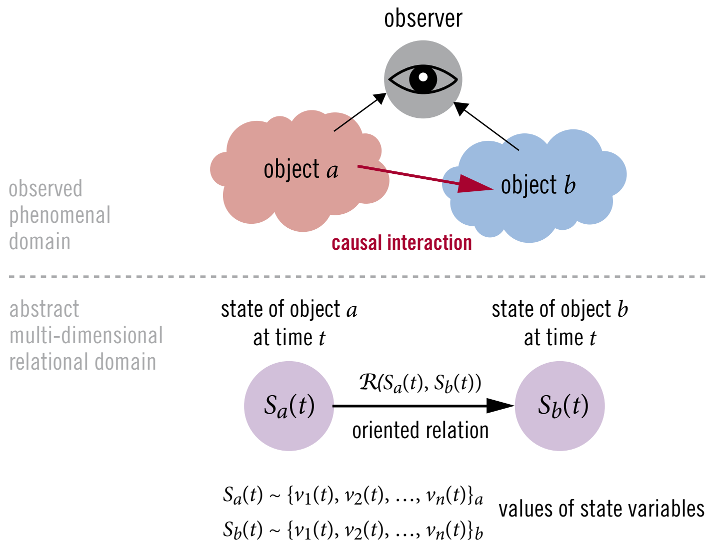


Figure 2: Causal interaction represented by an oriented relation. In this figure the ordered pair is denoted by the relation $\mathcal{R}(S_a(t), S_b(t))$, where $S_a(t)$ and $S_b(t)$ are the states of objects ***a*** and ***b*** at time ***t***. These states are labels for the n-tuples of values of their n ascribed state variables at time t:

$S_a(t) \sim \{v_1(t), v_2(t), \ldots, v_n(t)\}_a$
$S_b(t) \sim \{v_1(t), v_2(t), \ldots, v_n(t)\}_b$

These n-tuples represent two n-dimensional points ascribed to objects ***a*** and ***b*** in an abstract n-dimensional space. It is to be noted that this abstract representation corresponds to the description that an observer can make of an already performed observational experience in which he or she was able to determine the values of the state variables as observed (measured) at time ***t***, before actually observing the triggering and the caused transitions in objects ***a*** and ***b***. This means that the observer is able to say a posteriori, that ***a*** and ***b*** were related at time ***t*** by an oriented relation $\mathcal{R}$ that represents a cause-effect interaction started at that moment. The observer can say this because he or she observed the occurrence of a triggering transition in ***a*** and ***a*** caused transition in ***b*** within a finite time interval. The reference to time point ***t*** is important in this description because the states of both objects change during the interaction propagation time interval and in their new arrival states they may not be related in the same way (for the observer). In this sense, the "arrow" that represents the observed interaction "exists" only when the state transitions are about to be observed (at time ***t***). Afterwards, the situation is different: both objects may be either "connected" by an arrow or not connected at all.

The only constraint that should be foreseen concerns the conditions needed to assure that interaction activity actually takes place within a finite time interval. This means that a "poised stasis situation"[15] should not happen. A poised stasis situation occurs when:

a | cause-effect coupling *triggered by other objects* is the *only way* through which dynamic objects are observed to undergo state transitions (i.e., no spontaneous, internally generated transitions are observed), *and*

b | interactions are never observed to occur because *no triggered state transitions are observed on any object,* albeit they are all poised (i.e., in states that the observer may qualify as of "readiness to react"), according to observations performed in other situations.

14 | This is a strictly deterministic definition. We could say "occurs with a nearly 1 probability" instead of "always occurs," but let us stay simple (and classical) this time.

15 | The term "stasis" connotes here the notion of "stoppage of a normal flow" as in medical terminology. In the context it refers to the stoppage of the "causation flow."

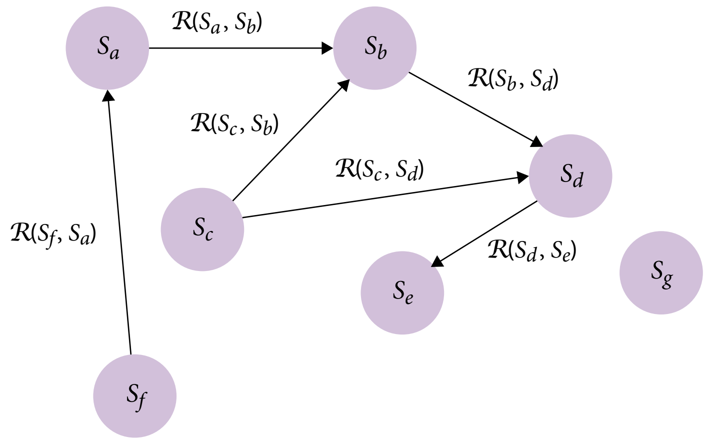


Figure 3: Example of a graph of nodes and edges representing a network of oriented relations.

In other words, if at a given time the values of the state variables of all existing objects are such that the underlying interaction mechanism does not produce any causal effect, a stasis with poised objects will occur and the interactive dynamics will stop (or never start).

A situation of poised stasis cannot be completely avoided, since nothing prevents dynamic objects that exist in an observational domain from never encountering each other through cause-effect couplings. The necessary conditions for interactions to occur depend on the range of values of the state variables that allows the manifestation of a specific cause-effect coupling mechanism.

To reduce this possibility we should admit that *dynamic objects are capable of performing state transitions that do not depend on interactions* with other dynamic objects. This could be the outcome of:

a | *global variations of the values of some state variables* due to the action of underlying *non-interactive causal mechanisms affecting synchronously or asynchronously all dynamic objects* existing in the observational domain being considered;[16] *or*

b | *spontaneous fluctuations in the values of some state variables* affecting individual dynamic objects[17] (internally triggered transitions).

Thus, we should always keep in mind that the phenomenology of interacting dynamic objects *might not be the only observable phenomenology* in a given observational domain. In this explanatory approach attention is focused on phenomena due only to interactions because I want to show that this is the domain in which autopoietic behavior manifests itself. The existence of other phenomena that allow for the manifestation of state transitions is assumed, but they are not taken explicitly into account.

Moreover, an observer may decide to discriminate *specific interactions* according to different underlying mechanisms responsible for cause-effect coupling occurring *between dynamic objects*. When *different types of interactions occur* in the same observational domain for the same dynamic objects, these may become components of distinct but intersecting interaction structures or unities with distinct boundaries and environments.[18] In this sense, an emerged unity becomes apparent for an observer depending on the type of interaction considered while observing, whereas other emerged unities may be composed partially by the same dynamic components and remain unobserved as totalities.

Abstract explanatory domain

As the notion of cause-effect interaction refers implicitly to a causation mechanism observed in a particular observational domain, when generalizing the domain we need to express our reasoning in more abstract terms.

An interaction between two dynamic objects can be represented as an oriented relation or, in other words, as an *ordered pair* of objects coupled by a non-specified causation mechanism. As each object is represented at a given time by its dynamic state, which is described by the instantaneous values of its associated state variables, it follows that a dynamic object can be represented by a *point in a multi-dimensional space.*

An *ordered pair of multi-dimensional points or oriented relation* is an element of the power set or Cartesian product of this multi-dimensional space with itself (hereafter called the "multi-dimensional relational space"). The order of the pair is defined by placing the object undergoing a triggering transition first, followed by the object undergoing the transition caused (see Figure 2).

Thus, a network of interactions between multiple objects can be described at a given time as a subset of the multi-dimensional relational space. This subset of oriented relations can be represented graphically as a graph made of nodes and edges (see Figure 3), in which the nodes represent the states of interacting dynamic objects at a given time and the edges represent the oriented relations established through observation.

The graph corresponds to a network of relations defined in an abstract multi-dimensional relational space that is portrayed graphically in the n-dimensional space where the nodes are defined. The topological configuration of the relational

16 | In the case of the bio-molecular domain, thermal agitation (Brownian motion) affects all molecules simultaneously, thus reducing the probability of a poised stasis situation at suitable temperature ranges.

17 | In the case of the bio-molecular domain, some macromolecules are subject to spontaneous changes of state: protein folding is one example.

18 | In the case of the biological domain, we can differentiate normal cells from neurons, for example: neurons interact by means of the specific interaction mechanism of synapses and constitute a self-organized nervous system that intersects with the whole organism, a system composed of normal cells plus the neurons themselves.

network represents the global state of all "connected" nodes at a given time and is isomorphic with the interaction network distinguished by the observer in a given observational domain. In my theoretical extension to a non-specified observational domain, the interaction network becomes the causation network.

In Figure 3, the multi-dimensional points are drawn as circles with their associated state labels and the oriented relations established between them are drawn as arrows. Any dynamic object describable with a chosen set of n state variables can be represented in this abstract n-dimensional space (in the figure it is drawn flat, for convenience). An isolated circle represents a dynamical object in a particular state (a particular set of values of the state variables) in which it is not related to other objects (it does not trigger transitions on others nor is affected by any transition occurred in other objects). In this particular case, at the moment considered, the n-dimensional point is not a node of any graph, but by convention it is nevertheless represented in the figure as it may become a node at another moment.

Formally, the oriented relations are not members of this space, but of its power set, the multi-dimensional *relational* space, but they are drawn conventionally in the same figure as one-way connectors relating the existing n-dimensional points *at a given time*.

In my explanations I use a language containing terms pertaining to two interlacing explanatory domains: one pointing to an observational domain (in which particular observations are made) and one pointing to the above-mentioned abstract explanatory domain based on mathematical concepts (in which generalized explanations are produced). Both domains intersect through the doings of the observer, who is involved both in biological experiences in the observational domain and in cognitive experiences of "distinguishing-describing-explaining" in the languaging domain.

For a wider understanding, I avoid unnecessary rigorous mathematical definitions. However, some correspondences between terms need to be highlighted (see Table 1 on next page).

Causation structures

To advance my explanation, I will need to propose further definitions necessary to account for the emergence of observable causation structures:

1 | A *chained cause-effect coupling* occurs when a change of state of object *A* unidirectionally triggers a change of state of object *B* via the change of state of an intermediate dynamic object *C* linked by direct cause-effect coupling with *A* and *B*. In this sense, the interaction between *A* and *B propagates* via *C*.

2 | A *dynamic structure* is a set of dynamic objects linked together by direct cause-effect coupling or chained cause-effect coupling. In this sense, a dynamic structure can be described as a *causation network* emerged from the activity of dynamical objects interacting in the observational domain; it is a propagation path for causal effects between dynamic objects.

3 | A *causation network is isomorphic*[19], *with the network of relations* (or graph) representing the dynamic structure in its associated multi-dimensional relational space. The linked *dynamic objects of a causation network are represented by the nodes of the corresponding network of relations*.

4 | A dynamic object *X* belongs to a dynamic structure *if there are one or more dynamic objects coupled directly or chained to X* and all of them are coupled in the same way to each other at a given time. A dynamic object that is member of a dynamic structure is called a *component* of the structure.

19 | In this explanatory approach I consider the network of interacting components (existing in the observational domain) and the network of oriented relations (existing in the multi-dimensional relational space) as isomorphic sets with respect to components and cause-effect couplings, that is to say: a) a relational network node always represents an existing component and every existing component can be represented by a node; b) an oriented relation linking two nodes always represents a cause-effect coupling between the corresponding components and every cause-effect coupling between components can be represented by an oriented relation linking the corresponding nodes.

5 | A dynamic object *Y* does not belong to an identified (observed) dynamic structure *if there is no direct cause-effect coupling linking Y to any member* of the identified structure at a given time. Such an object is said to belong (at a given time) to the *environment* in which the structure emerges. In Figure 3, this would be the case of the isolated point in state S_g

6 | *Dynamic structures* arise when *identifiable subsets of more than one dynamic object within the global dynamic system* emerge as distinct (observable) dynamic objects that can be represented as nodes of a network of oriented relations.

7 | We shall note that the *emergence of dynamic structures* is a "natural" outcome of the propagation of causal effects throughout the *history of state changes* experienced by all the dynamic objects existing in an observational domain. It is a phenomenon in itself and it does not depend on any observer criterion other than the choice of a particular type of interaction that the observer decides to observe and describe.

8 | *Components of a dynamic structure may become environmental objects* if they reach states in which they no longer interact with structure components (the case of S_g in Figure 3). Conversely, *environmental objects may become components of a given dynamic structure* if they reach states in which they do interact with at least one structure component. In other words, dynamic objects can enter and leave a dynamic structure in the history of state changes of its components.

9 | *Dynamic structures may emerge and disintegrate*, according to the propagation of causal effects through time.

A causation network as defined above is an observable entity evolving in time. Its representation in an abstract multi-dimensional relational space is a network of oriented relations (a graph) whose topology changes continuously according to the changes induced by the causation flow (which depends on the underlying interaction mechanism) on the states of each system component, transition after transition.

When considering different type of interactions in a given observational do-

Terms related to an observational domain	Corresponding terms related to an abstract explanatory domain
Dynamic object is an observable entity existing in an observational domain that *an observer can interact with* and associate to a set of well-defined state variables with values that can change in time.	**Multi-dimensional point** (abstraction for a dynamic object) is an element of a multi-dimensional space and is defined by an *n-tuple* (a set of coordinates) corresponding to the instantaneous values of the state variables of a dynamic object.
State of a dynamic object is the set of observable instantaneous values assigned to each variable associated to a dynamic object (an *n-tuple* of values). A state transition occurs at a given time whenever a value of a state variable changes.	**State of a multi-dimensional point** (abstraction for the state of an object) is an *identifier* assigned to a particular n-tuple that defines the state of a corresponding dynamic object. A state transition occurring in the latter is represented by a different identifier. We need only a changing identifier to express that a transition has taken place
Interaction is the cause-effect coupling between dynamic objects described as the occurrence of a state transition on a dynamic object that provokes a state transition on other dynamic objects by means of an unspecified underlying interaction mechanism.	**Oriented relation** (abstraction for an interaction) is defined as an ordered pair of multi-dimensional points; the order of the pair is established by the corresponding cause-effect timed sequence of triggering and caused state transitions observed in the corresponding interacting dynamic objects.
Observational domain is any domain of perceived phenomena occurring in time in which we, as human observers, can perform sensorial and operational experiences through interactions with dynamic objects. Within an observational domain, observed dynamic objects may enter into observable interactions with each other and constitute groups of interacting objects (for us, as observers). An observational domain is delimited by the actual observational capabilities of the community of observers, but it can also expand as new observational means become available (tools and methods), enlarging thus the intersection with the domain of existence of the observed dynamical entities.	**Multi-dimensional relational space** (abstraction for a general domain containing the observational domain) is the *space of all possible oriented relations* linking pairs of multi-dimensional points together: within this abstract space, any set of *topologically connected* oriented relations may be represented by a graph in which the relations are represented as edges (arrows) and the connecting multi-dimensional points as nodes (circles). These graphs are abstractions that represent groups of interacting dynamic objects, at a given time, in the corresponding observational domain. Its mathematical definition as the power set of an *n*-dimensional space entails that any pair of *n*-dimensional points would be a member. Among these, I consider only one-way related points (oriented relations) at a given time, and among the latter, only those oriented relations that correspond to observable interactions.
Interaction network is an observed set of dynamic objects acting as agents of interaction propagation within a dynamic system, where **Dynamic system** is the set of dynamic objects observed as a composite unity of mutually interacting dynamic objects.	**Network of oriented relations** (abstraction for interaction network) is a graph made of nodes and edges: it is the isomorphic representation of an observed interaction network at a given time.
Component is a dynamic object pertaining to an interaction network; it is seen as an observable totality that participates in the constitution of a dynamic system by acting as source and receptor of causation flow.	**Node** (abstraction of component) is a multi-dimensional point defined by the starting or arrival point of a relational arrow in a graph and is in isomorphic correspondence with a component of the interaction network represented by the graph.
Global state is the set of all values assigned to the variables of all components of a dynamic system at a given time.	**Instantaneous configuration of oriented relations** is the topological arrangement of the graph of a network of oriented relations at a given time depending on the state of each node at that time.

Table 1: Correspondence between terms of the observational domain and terms of the abstract explanatory domain.

main, each type of interaction may produce cause-effect coupling configurations that are causally independent from each other. Different types of interactions may depend each on different sets of variables, so that the corresponding cause-effect coupling overlapping configurations may be represented by distinct networks of oriented relations in multi-dimensional relational spaces with nodes of different dimensions. If the respective sets of state variables associated with each type of interaction do not have variables in common, the respective causation networks are independent and constitute different dynamic systems.

Causation networks may produce particular cause-effect coupling configurations involving multiple dynamic objects that are recognizable to the observer as being repetitive in time or in space. These special causation flows may be distinguished as patterns of network activity or *causation patterns* that do not necessarily involve a fixed set of dynamic objects, but rather substructures of connected nodes (local graphs) showing internal topological arrangements that are recognizable by certain characteristic features. These causation patterns may be perceived as processes that are identifiable through observing the chronological and/or spatial regularities of their occurrence.

Example in the macromolecular domain

In very schematic terms, macromolecules interact through the establishment of chemical bonds between some of the atoms they are made of. The necessary conditions for the establishment of such bonds, which are the outcome of electromagnetic forces that bind macromolecules components together at the atomic level (this is the underlying interaction mechanism), are essentially those of distance and spatial orientation.[20] If they are not close enough and/or not adequately positioned in a three-dimensional mutual orientation allowing for a match between mutually interacting atom structures to occur, the binding or unbinding interaction may not take place. To each molecule an observer may assign at least two state variables: say spatial *position* and spatial *orientation* (expressed in any suitable coordinate system).

Suppose that at a given time two macromolecules A and B are not interacting in this manner. Their states can be expressed by the values of their spatial *position* and spatial *orientation* at that moment. The physical constraints defined above in very general terms can be made explicit in this case: macromolecules are subject to thermal agitation in the substrate in which they exist. Within a suitable temperature range, thermal agitation implies global variations in the values of *position* and spatial *orientations* of all individual macromolecules. Macromolecules A and B will undergo non-interactive state transitions and the state variables *position* and *orientation* may reach values in which an interaction mechanism can take place.

Let me be even more explicit. Suppose that A moves and rotates with respect to a previous situation in which no interaction with B occurred. The change of state of A is expressed by changes in the values of its *position*, its global spatial *orientation*, and its constituent atoms' spatial configuration. That is to say, for example, that some values of its set of variables change in time in such a way that a) the distance between A and B allows the attractive or repulsive electromagnetic force to act by inducing a mutual movement, and b) the mutual spatial orientation of the nearest atoms produces a three-dimensional match of atom structures of A and B that facilitates a chemical interaction. The interaction mechanism produces a change of state in B: it may occur, for example, as B moves due to the action of attractive or repulsive electromagnetic forces (its *position* variable changes).

After an observable time interval a chemical bond may be established or broken and a binding or unbinding interaction can be accomplished. Dynamic interaction structures may arise in such an environment where macromolecules are free to move and rotate.

It is important to note that the *dynamic causation structure* that we are talking about is not to be confused with the *spatial structure* of the resulting bonded macromolecules. In fact, the dynamic causation structure is an *abstract concept* (expressed in the languaging domain) based on the notion of *described oriented relations* of cause-effect coupling between dynamic objects leading to the identification of a *network of oriented relations* where the dynamic objects are represented by the *nodes*: it is an entity *existing in a multi-dimensional relational space that can be mathematically defined* only within the languaging domain.

However, the resulting spatial structure of bonded macromolecules is an observable conglomerate of macromolecules *existing in the macromolecular (phenomenological) domain and represented as embedded in Euclidean space*. Euclidean space is a mathematical construct expressed in the languaging domain, although *isomorphic* with a system made of observable three-dimensional physical objects, separated by physical distances, oriented in the ordinary physical space of our experience as observers, and existing in an observational domain that we call the classical physical domain. For all practical purposes of observation, the macromolecular domain is usually considered as a particular case of the classical physical domain, even if quantum mechanical interactions are involved in the underlying interaction mechanisms.

A similar analysis could be applied to the neuronal domain, the human social domain, the world economical-financial system domain, the animal societies domain, the robotic agents domain or the concurrent software processes domain, etc. The problems left to the theoretician-observer are those of describing the underlying interaction mechanism and defining the adequate state variables sets that will account for the changes of state that are determinant in order to explain the occurrence of an interaction between the corresponding basic dynamic objects (neurons, individuals, enterprises, animals, robots, processor threads, etc.). In most cases, these tasks are not trivial at all and any choice made by the observers should be handled merely as working hypotheses meant to guide the deployment of adequate observational protocols.

20| For the sake of simplification I neglect other physical conditions that determine binding interactions between molecules such as the presence of a substrate environment (liquid water substrate) and of other catalytic molecules involved in complex macromolecular reactions.

Interaction conditions

The nature of the underlying interaction mechanism determines the conditions in which cause-effect couplings can take place. These conditions affect the range of values of the state variables of dynamical objects that allow for the manifestation of an "influence" of an object on another. The range of values of the state variables of dynamic objects for which an interaction is possible is determined by the "laws" that apply to a particular type of interaction in a particular observational domain.

It is important to highlight the fact that the set of variables associated with a dynamic object existing in a general observational domain may contain variables corresponding to any measurable feature that participates in the occurrence of an "influence" of an object upon another. In an ordinary physical domain *the most likely variable to be included is the position of objects* in three-dimensional Euclidean space, since the most common "influence" is the action of a *physical force* produced by a force field *that depends on the distance* separating two physical objects.

The generalization introduced by the abstract model proposed in this paper does not assume anything about the nature of the characterizing state variables. In particular, this means that ordinary physical distances may not appear as participating in the establishment of cause-effect couplings. Hence, the notion of *spatial vicinity* affecting dynamical objects may be considered as irrelevant for the existence of an interaction dynamics, *provided that other mechanisms that do not depend on the spatial separation between objects are available* to "transmit" the "influence" of a state transition of an object to other objects and provoke state transitions in them.

For example, consider a situation in which dynamical physical objects are separated by distances for which no physical force can act as a direct interaction "support." A simple "signal" carrying the information of the occurrence of a state transition in an object can trigger state transitions on other objects, provided that they are "sensitive" to the manifestation of a state change in the "signal emitting" object. The interaction thus established is a cause-effect coupling, even if the "signal emitting" object is light years away. The interaction is possible just because the underlying interaction mechanism is embodied almost completely by the nature of the dynamic objects that are capable of "reacting" to that kind of "signal."

This consideration implies that the notion of cause-effect coupling is so general that almost no conditions are to be imposed to assure that their occurrence is possible in physical terms. The minimal and unavoidable condition is that the "information"[21] concerning the state transition occurring in the causing dynamical object should be able to propagate via a suitable substrate to the affected dynamical objects. I emphasize that, by definition, any observational domain is a physical domain and that ordinary physical laws are applicable in it. In particular, this means that any cause-effect coupling is possible only if an energy transfer occurs between the causing and the affected dynamical objects (in the minimal case considered above it is the energy of the "information" carrying signal, or *transition triggering energy*). But we must keep in mind that *the total energy necessary for the accomplishment of the state transitions in the affected dynamical objects does not need to be provided entirely by the causing object itself*: Except for the transition triggering energy, it can be provided entirely by energy reservoirs embedded in the affected objects themselves or available as "local" sources of energy related only to the affected objects.[22]

The VM&U validation test

This is an introductory presentation of the VM&U rules and for the sake of readability I interpret them in my own wording. However, I think that their overall meaning is preserved. Most issues derived from their precise meaning are discussed in Part II and Part III of this work. In the endnotes I reproduce the original wording of each rule (Varela et al. 1974: 192–193).The rules are presented as related to the adjectives *bounded, composite, mechanistic, self-referential, self-producing, and autonomous*. These are meant to synthesize the meaning of the rules as qualifications of the *characterizing properties* that the observer should be able to distinguish and describe in order to ascertain that an observed unity is autopoietic in the sense defined by these authors.

The six VM&U rules constitute an operational validation test for any claim concerning the autopoietic nature of a dynamical system supposedly existing in a domain that intersects with a particular observational domain.

Bounded

The first question to answer is the following:[23]

> **Rule 1**: Is the observed unity circumscribable, which is to say, can we say where the unity ends and/or where its environment begins? If the answer is NO → the unity cannot be described and nothing can be said about it.

Hence, the first task is to identify the unity by distinguishing its *boundaries*. In the section, "The observer's "point of view," I discussed that, in order to differentiate a

21| Talking about "information" is a linguistic operation of the observer for commenting that there is a phenomenological connection or causation mechanism at work. It is just a suitable manner to name events in which the only observable "exchanges" can be described by means of ordinary physical concepts such as forces, energy transfers, etc. A dynamical object is not affected by any "informational contents" but by induced physical manifestations of phenomena occurring somewhere in its physical domain of existence.

22| In the case of static objects, the notion of "interaction" comprises the energy transfer necessary to produce the "action" on the affected object since the effect depends on the energy that the causing object can provide. In the case of dynamic objects, the caused "action" can be performed almost entirely by the affected object by using local energy resources. In other words, it is a situation in which there is no observable direct cause-effect relationship between the state transitions occurred in the causing dynamical object and the activation of the energy reservoirs that provide the total energy required to produce state transitions in the affected dynamical object. This fact entails important consequences for the thermodynamic analysis of the whole system, but this issue is outside the scope of this paper.

23| "1. Determine, through interactions, if the unity has identifiable boundaries. If the boundaries can be determined, proceed to 2. If not, the entity is indescribable and we can say nothing."

composite unity from the background, the observer needs to define some *inclusion criteria* based in its own interactions with the entities being observed. These criteria need to refer to some lower level distinctions related to the intrinsic nature of the unity that makes it appear as differentiated from any other entities observed in the same observational domain.

Composite

Assuming that the unity under examination is already described in terms of observational distinctions that allow observers to claim that there is a *distinguishable object* brought forth to our perception within the observational domain under consideration, the second question to answer is:[24]

Rule 2: Does the observed unity have constitutive elements, which is to say, can we distinguish components in the unity and see the whole as a describable set of parts?
NO → the unity can only be analyzed as an undifferentiated whole. It is not an autopoietic system.

From within my explanatory approach, this composite unity should not appear as the description of an arbitrary collection of objects among other objects, but of a set of components related to each other by an observable feature that makes them appear to the observer's perception as *mutually related by the occurrence of a phenomenon* in which they are dynamically involved. Even though Varela et al. 1974, talk just about "relations" and "mutual relations" without referring to dynamism, I consider, without loss of generality, that the relations between objects are always associations made by the observer that correspond to the observation of physical interactions occurring between objects *in time* (therefore I use the qualification of "dynamic" objects and "dynamic" relations). The distinction of a dynamic composite unity is related to the *distinction of a particular kind of activity* occurring *between* the dynamic objects that compose the unity. This activity can be stated as an *observable dynamic relation* between them.

In the practical case of a *particular observational domain*, the observer should distinguish *mutually related objects* from other dynamic objects that do not participate in the composition of the described unity. Again, from within my explanatory approach, the observable relation between dynamic objects should be describable as a *physical interaction* in which *the activity of an object affects the activity of other objects* by means of a particular *interaction mechanism* that the observer chooses to observe and describe.

However, when considering a *general non-specified observational domain*, the description of these interactions becomes an abstract exercise that should rely only on distinctions in which the nature of the interaction mechanism is not to be taken explicitly into account. Hence, the observable links should be described only with respect to the *characteristics of their occurrence within a causation propagation flow* and not to their specific physical nature. In other words, what matters is *how* the components are causally connected (i.e., related) to each other, and not by which physical means.

The theoretical problem of defining the boundary of a dynamic system is tantamount to specifying the preliminary distinctions necessary to bring forth a composite unity of dynamic objects or entities as described before. The task consists of specifying an observationally based operation of distinction that allows for the *identification of two distinct classes of dynamic entities*: those that *belong to the observed unity as such* (components) and those that *belong to the environment* in which the said composite unity supposedly behaves as an autopoietic system.

Now, what would be a *boundary between the unity and the environment*? One would be tempted to consider a boundary as an "abstract frontier" that separates the two sets of dynamic entities (a sort of separation criterion, a point of view, existing only in the mind of the observer). Instead, my explanatory approach requires that the observer should be able to distinguish the boundaries by describing observable interaction phenomena involving, especially, components and non-components. Obviously, this supposes that the observer has already chosen an inclusion criterion in order to distinguish components from non-components. Now the aim is to distinguish "edge" components from other components. In other words, the observer should focus attention on a particular "material subset" of components that could constitute a boundary, seen as a "material entity" "specialized" in being related directly to non-components. This subset would be composed of those components that are "in contact" with the environment. In this logic, a boundary is the set of those *components of the unity that are affected by observed direct interactions with non-components.* This statement may appear to be a trivial definition, especially when we consider the basic biological case, where we easily see that the boundary is materialized by the cell membrane, which is indeed part of the cell and is in contact with the environment. We shall see that in the more general case this identification of a boundary with a particular subset of components of the unity is not trivial at all.

Mechanistic

Once the unity has been described as a composite dynamic entity clearly differentiated from its environment, the question of the explanation of its overall behavior arises. What makes the system behave as it does? In this regard, there is a third subtle question to answer[25]:

Rule 3: Is the unity a mechanistic system, that is to say, are its properties the outcome of the relations between its components and not the expression of properties of the components themselves?
NO → the unity is not an autopoietic system.

Here we need to clarify what we understand by property. In this explanatory approach, emphasis is given to dynamics, so this rule states that the dynamics of the whole system should be described only in

24| "2. Determine if there are constitutive elements of the unity, that is, components of the unity. If these components can be described, proceed to 3. If not, the unity is an unanalyzable whole and therefore not an autopoietic system."

25| "3. Determine if the unity is a mechanistic system, that is, the component properties are capable of satisfying certain relations that determine in the unity the interactions and transformations of these components. If this is the case, proceed to 4. If not, the unity is not an autopoietic system."

terms of the interactions between *components seen as totalities* in which the *internal dynamics* of each component is irrelevant for explaining the dynamics of the composite unity as a whole.

Although there is no explicit claim by Maturana or Varela that the only properties to be considered are those related to a component's internal dynamics, they consider that mechanistic systems are "specifiable only in terms of *relations between processes generated* by the interactions of components, and *not by spatial relations* between these components." (Varela et al. 1974: 188, my emphasis). From our point of view this notion of "generated processes" argues in favor of considering the components' properties as features ascribed to their intrinsic dynamic capabilities.

In a particular observational domain, the actual interaction mechanism involved in the manifestation of the causal effects that constitute the relations between components should be considered only with respect to the results of this causation.

This means that the explanation of the system's history of state changes in time should rely entirely on the description of the mutual relational dynamics established between "atomic" objects that the observer does not need to describe in detail because they share a common and general behavior that accounts for the relational outcomes observed in the system's dynamics. Within this explanatory approach, components should be describable as entities of the "same kind" that behave in a specified similar manner. This similitude does not mean that they show similar properties, but that they show reactions to causing events impinged on them (by other components) that are describable as state transitions that can be observed by using a set of state variables (see "Definitions and conceptual tools" above) that is *common to all components*.

In a non-specified observational domain, as the interaction mechanism is not specified but only assumed to exist, *components can be described only as sources and receptors of causation*, period. Hence, components should be describable as generalized dynamical entities that participate in causation propagation effects only as a result of effects impinged on them by other similar dynamical objects. Thus, the system's dynamics should be *describable in terms of the evolution of "causation patterns"* established between components (see "Causation structures" above).

Self-referential

As said before, within this explanatory path, the boundary of the system should be describable as the subset of components that interact with non-components. Furthermore, the *origin and existence of the boundary* should be describable as an *outcome of the system's own dynamic*. For Varela et al. (1974), the boundary components should "constitute these boundaries through preferential neighborhood relations and interactions between themselves," meaning implicitly "neighborhood relations" in physical space. Their wording does not include an explicit reference to the *role of a boundary* with respect to the relations of the unity with the background elements. I intend to be more explicit on this matter.

This is expressed in the fourth question to be answered,[26] which introduces the role of the observer:

Rule 4: Do the components of the observer-described (apparent) boundary participate in that boundary as a result of their interrelationships with the other components of the system?
NO → the boundary is specified by the observer, not by the unity itself. The unity is not autopoietic.

This identification rule states that *the description of a boundary should be expressed in terms of the observed interaction activity between the dynamic objects that constitute the system and leads to the emergence of a special subset of components showing preferential neighborhood relations and interactions between themselves*. A boundary should be describable as a continuous production process of the system itself. In this sense, the identity of the system (distinction of what belongs to the system and what does not) appears to the observer as being independent from any prior observer-defined inclusion criterion used to *differentiate the unity from the background*. The unit differentiates on its own by *producing* a subset of boundary components as an outcome of its own activity, and in this sense we say that it is self-referential.

In a non-specified observational domain, this description could be termed only by appealing to *descriptions of distinct causation patterns* (as they imply special neighborhood relations between components) that allow us to differentiate dynamical objects that act as system components from other objects that do not. Furthermore, these distinguished causation patterns should account for the *identification* of particular sets of components that constitute the "borders" or "frontiers" between the system and the environment. As a dynamic system evolves through time, this identification procedure should be applicable (at any moment) in order to identify the components that participate in the "frontier" set (at any moment) as an outcome of observed past causation patterns.

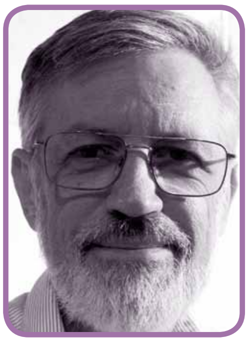
Self-producing

Once the self-referential property is established, the question of the origin of the system's components arises. The very nature of an autopoietic system resides in its *ability to produce its own components*.[27] The two last VM&U rules deal in fact with some requirements of the component properties concerning the production of system components themselves. These intra-boundary related aspects are most important and specific to the task of identifying an autopoietic system.

Rule 5: Are the components produced as a result of interrelations between components, either by transformation of previously produced components or by accreting non-components through the boundary?
NO → the unity is not an autopoietic system.

26| "4. Determine if the components that constitute the boundaries of the unity constitute these boundaries through preferential neighborhood relations and interactions between themselves, as determined by their properties in the space of their interactions. If this is not the case, you do not have an autopoietic unity because you are determining its boundaries, not the unity itself. If 4 is the case, however, proceed to 5."

27| The term "autopoiesis" has been coined from the Greek words *auto* = self and *poiesis* = creation.



HUGO URRESTARAZU
was a teacher-researcher for nuclear and solid state physics at the Universidad de Chile and conducted postgraduate Ph.D. research on semiconductors at Imperial College in England. He possesses 13 years of industrial experience in design and development of embedded real time operating systems and object oriented design and programming at Alcatel in France. He also worked as a consulting engineer in Central America and as a linguistic consultant in France. Back in 1972 he attended Humberto Maturana's Biology of Cognition lectures and was Francisco Varela's friend and colleague at the Facultad de Ciencias in Chile. Since 1990, he has been involved in independent interdisciplinary scientific extension activities concerning the theory of autopoietic systems.

This rule[28] states that the observer should be able to explain the production of new components in terms of the system's natural dynamics. In a particular observational domain this should be accounted for by *explaining how the interaction flow between components results in: (1) the acquisition of non-components as fully qualified components (coupled to the interaction structure) and/or (2) the transformation of existing components* according to the interaction mechanism prevailing in that particular domain.

In a general non-specified observational domain this should be expressed only in terms of causation patterns and the *components should be describable as entities capable of undergoing "productive transformations" under certain circumstances*. This means that components of autopoietic systems must comply with certain specific behavioral properties that are determinant for the self-production of the system as a whole. The *focus* shall be put only on the causal outcome of the *interaction patterns* (components' interaction activity) that result in component production, but not on the *particular internal properties* of components that allow them to act as "producers" of other components.

Autonomous

Furthermore, it is required that the self-production interaction dynamics should involve the participation of *all* system components. We shall see that this requirement leads to a very strong constraint on the components' interaction structure, namely the emergence of a highly complex and recursive causation network. This structural constraint results in a high level of system autonomy, which is the outcome of an internally closed causation dynamics where environmental interactions are to be considered just as external perturbations. The question to answer is:

Rule 6: Are all components produced by the interrelations between other components and do all components participate in the production of new components?
NO → there are components not being produced within the system and components not participating in the production of new components. The unity is not autopoietic.
YES → the unity is an autopoietic system in the space in which its components exist.

This rule[29] expresses a requirement related to the interaction configuration and to the *relational connectivity* between components whereby the participation in self-production is verified for all components of the system. In a particular observational domain this should be accounted for by explaining the way in which the interaction propagation flows involve any single component in component production events.

In a general non-specified observational domain this requirement should also be expressed in terms of *topological constraints* that are to be observed in the causation network, as required to assure that all system components are affected by component production events.

Acknowledgements

Above all, I wish to thank my late dear wife, Anne-Marie Lopez del Rio, for encouraging me through her immovable and loving support to transform the original manuscript into a publishable paper (done with love, in memoriam).

I especially thank William P. Hall, Ph.D. (Evolutionary Biology of Species and Organizations) for his valuable comments and suggestions expressed during several years of fruitful discussions. I also thank the anonymous *Constructivist Foundations* peer reviewers who helped me to improve the presentation of my views. Finally, I warm-heartedly thank my former professor Humberto Maturana, who, 40 years ago, managed to enthuse my mind with a life lasting and meaningful reflection process that has been going on ever since.

28| "5. Determine if the components of the boundaries of the unity are produced by the interactions of the components of the unity, either by transformation of previously produced components, or by transformations and/or coupling of non-component elements that enter the unity through its boundaries. If not, you do not have an autopoietic unity; if yes, proceed to 6."

29| "6. If all the other components of the unity are also produced by the interactions of its components as in 5, and if those which are not produced by the interactions of other components participate as necessary permanent constitutive components in the production of other components, you have an autopoietic unity in the space in which its components exist. If this is not the case and there are components in the unity not produced by components of the unity as in 5, or if there are components of the unity which do not participate in the production of other components, you do not have an autopoietic unity."

References

Beer S. (1972) The brain of the firm. The managerial cybernetics of organizations. Penguin Press, London.

Beer S. (1980) Foreword to autopoietic systems. In: Maturana H. R. & Varela F. J. Autopoiesis and cognition: The realization of the living. Reidel, Boston.

Beer S. (1995) Platform for change. Wiley, New York.

Dupuy J. P. (1982) Ordres et désordres. Editions du Seuil, Paris.

Luhmann N. (1986) The autopoiesis of social systems. In: Geyer F. & van der Zouwen J. (eds.) Sociocybernetic paradoxes. Sage, London.

Luminet J.-P. (2004) L'invention du big bang. Éditions du Seuil, Paris: 200–202.

Maturana H. R. (1988a) Ontology of observing, The biological foundations of self consciousness and the physical domain of existence. In: Conference workbook: Texts in cybernetics. American Society For Cybernetics Conference, Felton CA. Available at http://www.inteco.cl/biology/ontology/index.htm

Maturana H. R. (1988b) Reality: The search for objectivity or the quest for a compelling argument. The Irish Journal of Psychology 9(1): 25–82. Available at http://www.enolagaia.com/M88Reality.html

Maturana H. R. (1995) The nature of time. Instituto de Terapia Cognitiva, Santiago de Chile. Retrieved from http://www.inteco.cl/biology/nature.htm

Maturana H. (2002) Autopoiesis, structural coupling and cognition: A history of these and other notions in the biology of cognition. Cybernetics & Human Knowing 9(3–4): 5–34.

Robb F. (1989), Cybernetics and suprahuman autopoietic systems. Systems Practice 2(1): 47–74.

Schwember H. (1977) Cybernetics in government: Experience with new tools for management in Chile 1971–1973. In: Bossel H. (ed.) Concepts and tools of computer-assisted policy analysis. Volume 1. Birkhäuser Verlag, Basel: 79–138.

Varela F. J., Maturana H. R. & Uribe R. (1974) Autopoiesis: The organization of living systems, its characterization and a model. BioSystems 5: 187–196.

Varela F. J., Coutinho A., Dupire B. & Vaz N. (1988) Cognitive networks: Immune, neural, and otherwise. In: Perelson A. (ed.) Theoretical immunology Part II. Addison-Wesley, Reading MA: 359–375.

Varela F. (1992) Autopoiesis and a biology of intentionality. In: McMullin B. & Murphy N. (eds.) Proceedings of the workshop "Autopoiesis and perception." Dublin City University, Dublin. Available at ftp://ftp.eeng.dcu.ie/pub/alife/bmcm9401/varela.pdf

Varela F. J. (1995) The emergent self. In: Brockman J. (ed.) The third culture: Beyond the scientific revolution. Simon & Schuster, New York: 209–222. Retrieved from http://www.edge.org/documents/ThirdCulture/t-Ch.12.html

Zeleny M. & Hufford K. D. (1992) The application of autopoiesis in systems analysis: Are autopoietic systems also social systems? International Journal of General Systems 21: 145–160.

Received: 13 April 2011
Accepted: 5 July 2001