Check for updates

Article

Adaptive Behavior

Adaptive Behavior
1–11
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1059712318823723
journals.sagepub.com/home/adb
SAGE

Are living beings extended autopoietic systems? An embodied reply

Mario Villalobos[1,2] and Pablo Razeto-Barry[2,3]

Abstract

Building on the original formulation of the autopoietic theory (AT), extended enactivism argues that living beings are autopoietic systems that extend beyond the spatial boundaries of the organism. In this article, we argue that extended enactivism, despite having some basis in AT's original formulation, mistakes AT's definition of living beings as autopoietic entities. We offer, as a reply to this interpretation, a more embodied reformulation of autopoiesis, which we think is necessary to counterbalance the (excessively) disembodied spirit of AT's original formulation. The article aims to clarify and correct what we take to be a misinterpretation of AT as a research program. AT, contrary to what some enactivists seem to believe, did not (and does not) intend to motivate an extended conception of living beings. AT's primary purpose, we argue, was (and is) to provide a universal individuation criterion for living beings, these understood as discrete bodies that are embedded in, but not constituted by, the environment that surrounds them. However, by giving a more explicitly embodied definition of living beings, AT can rectify and accommodate, so we argue, the enactive extended interpretation of autopoiesis, showing that although living beings do not extend beyond their boundaries as autopoietic unities, they do form part, in normal conditions, of broader autopoietic systems that include the environment.

Keywords

Autopoiesis, extended autopoietic systems, enactivism, body, multiple realizability

Handling Editor: Tom Froese, Faculty Member at the UNAM, México, Mexico

1. Introduction

When we study living beings, we can consider many of their different aspects and properties: their anatomy, their physiology, their metabolic balance, their behavior, and so on. We may fill a whole handbook with the details of these and many other data about living beings. But, beyond all the information we can get about living beings, can we tell what a living being essentially is? Can we provide a definition of "living being"? One that gives us the necessary, not the contingent, features of all and only living beings? Even more, can we tell not only the necessary but also the sufficient conditions to qualify, identify, and recognize a given entity as a living being? Answering this latter question is the primary purpose of the autopoietic theory (AT): to give a universal definition of the living being, providing the necessary and sufficient conditions to recognize an entity as a living being.

AT looks quite simple at first glance. It simply asserts that living beings are systems that produce themselves (i.e. what "autopoiesis" literally means). This simplicity, however, is only apparent. Many different (and sometimes opposed) interpretations have been elaborated regarding the precise sense of the autopoietic conception of living beings (Razeto-Barry, 2012; Razeto-Barry & Ramos-Jiliberto, 2013).

One of these interpretations, advanced by enactivist authors such as Virgo, Egbert, and Froese (2011; but see also Colombetti, 2015; Di Paolo, 2009; McGregor & Virgo, 2011), views living beings as autopoietic systems that may extend beyond their own physical boundaries as organisms, including elements and processes of the environment (both abiotic and biotic). Under this view, a living being, say a worm, may prove to be composed, as an autopoietic unity, of elements and processes external to

[1]Escuela de Psicología y Filosofía, Universidad de Tarapacá, Arica, Chile
[2]Instituto de Filosofía y Ciencias de la Complejidad, Santiago, Chile
[3]Universidad Diego Portales, Vicerrectoría Académica, Santiago, Chile

Corresponding author:
Mario Villalobos, Escuela de Psicología y Filosofía, Universidad de Tarapacá, 18 de septiembre 2222, Arica, Arica y Parinacota 1010069, Chile.
Email: mario.kirmayr@gmail.com

the discrete organism that we usually identify as a worm (e.g. it could be composed, in part, of soil chemical processes). This "extended" interpretation of living beings as autopoietic systems, counterintuitive as it might sound, is not ungrounded. It has its possibility in the very AT's original formulation. AT, as we will see, conceives of living beings as multiple realizable systems, which, for reasons that will be examined in this article, may open the door to extended interpretations.

In this article, we will argue that the extended interpretation of living beings, despite having some basis in AT's original formulation, mistakes the definition of living beings as autopoietic entities. We will offer, as a reply to this interpretation, a more embodied reformulation of living beings' autopoiesis, which we think is in line with how AT (really) conceives of living beings. AT, as we will show, conceives of living beings as physically discrete unities, but fails to establish this point in a sufficiently clear manner in its definition of autopoiesis. The reformulation we propose in this article tries to fix this failure.

An embodied reformulation of AT will also help us to distinguish and rescue, so we hope, what we take to be the seed of truth behind the enactive extended view of autopoiesis. We will reject the view that living beings extend into the environment as autopoietic unities, but recognize that there are indeed autopoietic systems that exceed the boundaries of living beings. The upshot of our analysis is that AT's original formulation and its enactive extended interpretation, once rectified, can be fruitfully articulated.

The structure of the article is as follows. First, we will briefly show how the thesis of multiple realizability brings with it the logical possibility of extended interpretations. Second, we will demonstrate that AT's original formulation is committed to the multiple realizability of living beings, and therefore, implicitly, to the possibility of extended interpretations of living beings. Then we will illustrate how the extended enactivist view[1] exploits this feature of AT, drawing the conclusion, incorrect to our eyes, that the autopoietic unities that constitute living beings extend beyond the physical boundaries of the organisms. Finally, we will offer an embodied reformulation of AT and discuss how it may help to prevent the extended interpretation of living beings while making good sense of the idea that autopoietic systems may extend beyond the boundaries of living beings.

To be clear, the aim of this philosophical exercise is not to persuade the reader that, say, processes that take place in the soil that she sees surrounding a worm are not really parts of the worm. We think the reader is already firmly and correctly convinced that those environmental processes are not parts of the worm (unless the reader subscribes to some version of the extended view, of course).[2] The aim is rather to clarify and correct what we take to be a misinterpretation of AT as a research program. AT, contrary to what some enactivists seem to believe, did not (and does not) intend to motivate an extended conception of living beings. AT's primary purpose was (and is) to provide a universal individuation criterion for living beings, based on the manner that we typically distinguish living beings, that is, as discrete bodies that are embedded in, but not constituted by, the environment that surrounds them.

This latter point does not mean, as we will see, that AT cannot (or should not) broaden its theoretical scope and go beyond the autopoiesis of living beings. Quite the contrary. We hope to show that AT, after a subtle but important rectification in its original formulation, can, in turn, rectify and accommodate the enactive interpretation of extended autopoiesis. We think that an embodied reformulation of living beings' autopoiesis will help us to see that, although enactivism is wrong in thinking that living beings extend into the environment (and also wrong in assuming that AT supports such an interpretation), its basic intuition that there are autopoietic systems other (and larger) than living beings makes much sense within AT if formulated in the right way.

2. Multiple realizability and extended systems

When we see a system from the functional point of view, we focus on the causal relations, patterns of activity, processes, or mechanisms that can be distinguished as constituting the system, abstracting away from the specific materiality of the components, parts, or pieces that realize those patterns, processes, or mechanisms. Given a system, the functional approach asks how it works, not what is it made of. From a functional viewpoint, what defines the identity of a system is its functional organization, not its material realization. For example, drainage systems may be made of metal pieces, wooden pieces, or plastic pieces. If all of them work as drainage systems, then all of them are drainage systems, no matter their different material composition. The idea that a system may have, in principle, multiple material realizations, is known as the multiple realizability thesis.

The multiple realizability thesis expresses strict neutrality among the many possible material realizations of a system. But this neutrality, importantly, can also be about the spatial location of the system's constituents. The reason is that, from a functional viewpoint, a system corresponds, ultimately, to a concatenation of causes and effects, processes, and mechanisms that bring about a specific result. And such mechanisms need not always coincide with the system one tends to distinguish and take as a unity of analysis from the material point of view. A simple example to illustrate this idea, and that will be especially useful later in our analysis, is that of the thermostat (and its associated thermostatic system).

If you buy a thermostat in a shop, what you get is a physical apparatus packed in a box. Once set up and working in your office, what you have, basically, is a homeostatic system (specifically a thermostatic system) organized as a negative feedback loop, which runs through the physical apparatus you purchased in the shop *and* the physical medium to which it is functionally coupled, that is, the air inside your office. The material system (the thermostat) and the functional system (the thermostatic system) constitute two systems with different boundaries and compositions. The thermostat always remains the same, whether packaged in a box or working in a room. The thermostatic system goes beyond the boundaries of the thermostat, including any physical entity or medium that happens to be functionally coupled to the latter (e.g. the air of your office).

Interestingly, in some cases, the processes or mechanisms that constitute a functional system can be said to expand or extend, so to speak, beyond their habitual borders. This is the basic intuition behind the much-discussed hypothesis of the extended mind (Clark & Chalmers, 1998; Menary, 2010). To briefly illustrate, if we assume a strictly functional viewpoint, we can identify a cognitive system as the set of processes whose concatenation brings about the achievement of a certain cognitive task, such as the calculation of a mathematical operation. Suppose a trained mathematician can calculate mentally a complex arithmetic task that I can only estimate by writing down mathematical symbols on a piece of paper. If we take the functional characterization of the cognitive system seriously, we are led to recognize that while in the case of the mathematician the cognitive processes were realized through neurons alone, in my case they were realized through neurons, pen, and paper. That is, in one case the cognitive system runs inside the head, whereas in the other it extends beyond the head, including elements and states of the external world. If, and this is a big "if," the *unique* criterion to individuate a cognitive system is the functional one, then we are led to admit that cognitive systems, in certain occasions, may extend beyond the physical boundaries of the biological person.

It is not the purpose of this section to take any position regarding the hypothesis of the extended mind. Rather, what interests us is to show that when the unique criterion to individuate a system is the functional one, and strict neutrality regarding the physical realization of the system is respected, the possibility (though not necessarily the actuality) of extended systems is logically secured.

The point we want to emphasize with these examples is that the multiple realizability thesis brings with it the logical possibility of extended systems (Wheeler, 2017). But just the possibility. The multiple realizability thesis, as we said before, expresses strict neutrality about two different aspects: one related to the material nature of the components of a system, the other related to their spatial location. Both aspects are equally entailed by the thesis *when no restriction or specification is provided.* A restricted version of the thesis, however, may entail just one of them, blocking the other. For example, we may conceive of a system as multiply realizable concerning the materiality of its components, but not concerning the spatial location of said components (or vice versa). To this purpose, importantly, we need to explicitly specify the reach of the multiple realizability thesis, specifying at the same time the way we are blocking the aspect we want to block.

As we will see in the following sections, AT commits to the multiple realizability of living beings concerning their material composition, but not regarding the spatial location of their components. This is because AT, as we will show, conceives of living beings as physically discrete bounded systems. However, as we will see, AT fails to explicitly specify this latter aspect in its definition of living beings as autopoietic unities, leaving room for extended interpretations.

3. AT and the multiple realizability of living beings

AT offers a characterization of living beings that focuses on the *organization* of processes that, according to the theory, defines the class of living beings. To characterize a system regarding its organization, according to AT, is to characterize it regarding the pattern of relations that defines it, abstracting away from the specific material nature of the components that realize those relations. For example,

> [T]he same organization may be realized in different systems with different kinds of components as long as these components have the properties which realize the required relations. It is obvious that with respect to their organization such systems are members of the same class, even though with respect to the nature of their components they may be distinct. (Varela, Maturana, & Uribe, 1974, p. 188)

When approaching living beings, it is their organization as systems, and not the specific nature of the components that realize said organization, which matters for AT: "It is our assumption that there is an *organization* that is common to all living systems, *whichever the nature of their components*" (Maturana & Varela, 1980, p. 76, emphasis added). Living beings, according to AT, are multiple realizable systems *with respect to the material nature of their components* (cf. Maturana & Varela, 1998, p. 49).

This AT's feature is also visible in its "mechanistic" approach:

> Our approach will be mechanistic: … our problem is the living organization and therefore our interest will not be in properties of components, but in processes and relations

> between processes realized through components. (Maturana & Varela, 1980, p. 75).

Revealingly, this mechanistic approach to living beings is compared with the way we approach and explain a control plant (a typical instance of a functional kind) in contrast with the way we explain a physical phenomenon:

> [T]o explain the movement of a falling body one resorts to properties of matter, and to laws that describe the conduct of material bodies according to these properties … while to explain the organization of a control plant one resorts to relations and laws that describe the conduct of relations. In the first case, the elements used in the explanations are bodies and their properties; in the second case, they are relations and their relations, *independently of the nature of the bodies that satisfy them*… This mode of thinking is not new, and is explicitly related to the very name of mechanicism. (Maturana & Varela, 1980, pp. 75–76, emphasis added)

A control plant is a neat example of a functional kind, like mousetraps, filter valves, and the thermostatic system analyzed in the previous section. These kinds of systems are characterized strictly in terms of relations (of processes, mechanisms), independently of the nature of the bodies that satisfy said relations.

It seems relatively clear that AT commits to the multiple realizability of living beings concerning their material composition. The question is "Does this commitment imply that AT is also neutral about the spatial location of the components that constitute a living being?" In the next section, we will answer this question in the negative, but before that, let us review the way AT formulates the notion of an autopoietic system.

In line with its declared mechanistic (functional) spirit, AT defines autopoietic systems as follows.

> *An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces the components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in the space in which they (the components) exist by specifying the topological domain of its realization as such a network.* (Maturana & Varela, 1980, pp. 78–79, original emphasis)

In this definition, as we can see, there is no reference to the physical nature of the components produced by the network. Their physical nature might, in principle, be any, as long as they satisfy the autopoietic organization. The only requirements of an autopoietic organization are (1) the presence of a circularity (or recursivity) between the network of processes of production and its products and (2) the specification of a topology for such a network. This latter point, the topology, refers to the spatial constitution of the system, and it might be read as restricting the reach of the multiple realizability thesis only to the material aspect of the system, blocking an extended interpretation. However, for reasons that will be exposed in Section 5, this topological specification does not manage to prevent such an interpretation. In fact, as we will see now, enactivists have built an extended interpretation of living beings as autopoietic unities respecting, and being consistent with, AT's original formulation.

4. Extended living beings?

In the previous section, we said that AT's definition of living beings as autopoietic systems opens up the possibility of an extended interpretation. In this section, we will see how extended enactivism exploits this possibility. Enactivists have exploited this possibility mainly as a way of showing that their view of cognition as relational and world involving is not compromised by the (alleged) internalism that some authors ascribe to AT (Di Paolo, 2009; Virgo et al., 2011; Wheeler, 2010). Interesting as this discussion may be, however, it is not our target in this article. Our concern here is with AT as a theory of living beings, not as a theory of cognition.

There are two important claims enactivists make regarding AT, and that should be distinguished in this analysis. They claim the following:

1. That the constitutive processes of living beings as autopoietic unities extend beyond their spatial boundaries as organisms, and
2. That this extended view of living beings as autopoietic unities was the intended one, although not explicitly developed, in AT's original formulation.

Both claims can be found, for example, in this passage:

> we defend our claim that the operationally closed network that constitutes an autopoietic unity can include processes that occur outside of its spatial boundary by showing that this was the interpretation intended in one of the earliest pieces of literature on the subject, Maturana and Varela's *Autopoiesis and Cognition*. (Virgo et al., 2011, p. 245)

Let us start by focusing on Claim 2, to then reconstruct the way enactivists reach Claim 1. Why do enactivists think that the extended interpretation of living beings' autopoiesis was the intended one in AT's original formulation? Here is the reason. In *Autopoiesis and Cognition*, Maturana and Varela (1980) start by introducing the notion of an homeostatic system, claiming that in such a system every feedback loop, included those that run through the environment, must be considered as "internal" to the system.

> There are machines which maintain constant, or within a limited range of values, some of their variables. The way this is expressed in the organization of these machines must be such as to define the process as occurring completely within the boundaries of the machine which the very same organization specifies. Such machines are homeostatic machines and all feedback is internal to them. If one says that there is a machine M, in which there is a feedback loop through the environment so that the effects of its output affect its input, one is in fact talking about a larger machine M' which includes the environment and the feedback loop in its defining organization. (Maturana & Varela, 1980, p. 78)

In line with the example we reviewed in Section 2, Virgo et al. (2011) take the case of the thermostat to illustrate Maturana and Varela's point:

> according to [Maturana and Varela] it is not correct … to think of the thermostat as being the box on the wall that is connected to a heater and contains a thermocouple, because this machine (machine *M*) has a feedback loop that runs through the environment … Since the thermostat relies on [a] feedback loop for its operation, we should actually define the thermostat as a larger machine (machine *M'*) which includes the heater, the air in the room, and the feedback loop that passes through them. (p. 246)

This is, recall, basically the same point we made in Section 2 when illustrating the distinction between the thermostat as a material object and the thermostatic system as functional kind. Enactivists claim, correctly, that the physical boundaries of the former do not coincide (and are not to be confused) with the operational limits of the latter.

Enactivists then call our attention to what they take to be a crucial point: the fact that AT defines autopoietic systems as a subclass of homeostatic systems. In *Autopoiesis and Cognition*, Maturana and Varela (1980) indeed claim that "an autopoietic machine is an homeostatic … system which has its own organization … as the fundamental variable which it maintains constant" (p. 79). Building on this point, the enactivist reasoning is that if "[a]utopoietic machines are homeostatic machines" (Maturana & Varela, 1980, p. 78), then "[i]t follows that their definition must be expanded in the same way [that the definition of homeostatic systems is expanded] if they rely on a feedback loop that runs through their environment" (Virgo et al., 2011, p. 246).

Based on this reasoning, enactivists take the example of the earthworm that secretes to its environment a substance which helps to digest its food, and invite us to see the worm as an extended autopoietic system:

> We can try to see the worm as an autopoietic system (and hence an homeostatic system) whose operational limits are defined by its physical boundary (its skin). However, the worm relies on the effects of its secretions; this is a feedback loop which runs through its environment. The above quoted paragraph [Maturana and Varela's paragraph about homeostatic systems] thus compels us to redefine the system so that it includes not only the worm itself but also the secretions and their effects. On this view *the autopoietic system that constitutes the worm* is not coextensive with the unity that we refer to as "the worm," it is much bigger. (Virgo et al., 2011, p. 246, emphases added)

According to enactivists, "the organism, as an autopoietic system, includes processes that are not occurring within its spatial boundary" (Virgo et al., 2011, p. 245). This extended condition, goes on the argument, would be the general condition for most living beings, "since most organisms rely not only on sensory-motor loops that run through their environment but also on nutrients that are recycled externally to them" (Virgo et al., 2011, p. 246).

We have arrived then at the idea that most living beings are constituted as autopoietic systems that extend beyond the physical unities that we usually identify as living beings. One thing is the material organism as such, and another is the organism as an autopoietic system, the latter being much bigger than the former. This extended view of autopoiesis, according to extended enactivism, should not represent a problem, since it would be (allegedly) just an exegetical clarification of what was intended (but not fully developed) in AT's original formulation.

We think differently. AT, we argue, did not (and does not) intend an extended conception of living beings, despite leaving room for such a conception in its original formulation. AT, as we will see, conceives of living beings as discrete physical entities embedded in, but not extended into, the environment they inhabit. Although this conception is not explicitly stated in AT's original formulation (hence the room for extended interpretations), it is clearly expressed in several passages where the theory is explicated. For instance, early in 1974, Varela, Maturana, and Uribe (1974), after giving the formal definition of an autopoietic system, add a concrete exemplification which makes explicit the bounded and separable nature of the living being:

> Consider for example the case of a cell: it is a network of chemical reactions which produce molecules such that (i) through their interactions generate and participate recursively in the same network of reactions which produced them, and (ii) realize the cell as a material unity. Thus the cell as a physical unity, *topographically and operationally separable from the background*, remains as such only insofar as this organization is continuously realized under permanent turnover of matter. (p. 188, emphasis added)

Later, Maturana and Varela (1998) explain the notion of an autopoietic system, making explicit reference to the physical boundary of the biological cell, that is, the cell membrane. The cell membrane, when present, is, according to the authors, a key component of the operationally closed network that constitutes the cell as an autopoietic system. Specifically, "this membrane …*limits the extension* of the transformation network that produce[s] its own components" (Maturana and Varela, 1998, p. 46, emphasis added).[3] For AT, indeed, "the most striking feature of an autopoietic system is that it pulls itself up by its own bootstraps and becomes *distinct from its environment* through its own dynamics" (Maturana and Varela, 1998, pp. 46–47, emphasis added).

Now, the question is as follows: If AT conceives of living beings as physically discrete entities, and not as systems that extend beyond their physical boundaries, then what is wrong with its formulation such that an extended interpretation of life, as we have seen with enactivism, is *in fact possible*? Is there something missing from AT's definition of living beings whose inclusion might help to prevent extended interpretations? In the next section, we will argue that what is missing in AT's original formulation is the bodily dimension of living beings. AT, we argue, should explicitly take into account the trivial but crucial observation that living beings are physical *bodies*.

5. Living beings as autopoietic *bodies*

In this section, we will argue that living beings are, specifically, autopoietic *bodies* (and not merely autopoietic *systems*). In this discussion, by "body", we will understand a collection of matter that, thanks to internal forces (cohesive, adhesive, gravitational, etc.), constitutes a unitary whole endowed with an identifiable boundary that separates it from its surrounding. All bodies, being collections of matter in physical space, have mass and volume (though not every mass and volume, as we will see, constitutes a body in physical space). Also, to the extent that bodies have an identifiable extension in physical space, they all have a defined topology, although having a topology does not necessarily imply having a body (more on this and its implications soon). This notion of a body, although somehow sketchy, entails some aspects that are important to highlight.

First, the boundary of the body is defined by the properties of the material that constitutes the physical unity itself, and not merely by some distinction (though or stipulated) made by the observer. For example, an imaginary sphere of marble within a larger block of marble counts as a collection of matter with high internal cohesion but does not constitute a body, since the material of that imagined sphere does not establish itself any discontinuity with the rest of the block. A gold stone encased in a larger piece of marble, on the contrary, constitutes a body, since its extension as a collection of matter can be determined on the base of its material properties.

Second, a body is an aggregation of matter whose unity as an object is established by the physical interactions of its own components, not by some demarcation (perceived or thought) made by an observer, or by the physical action of external agents.

For example, the asterisms that we can distinguish in the night sky constitute objects in our perception (we can see the "Southern Cross" or the "Summer Triangle"), but not bodies, since their unity as objects is not given by physical links among their components but by optical effects in our observation. They have, so to speak, perceptual unity, but not physical unity.

Likewise, a mass of air trapped in a cavity inside a block of marble constitutes a distinguishable collection of matter, but not a body. Although the mass of air has a different material composition from the marble that surrounds it, its unity as a collection of matter is not given by its own constitution but by the action of the marble. Unlike the gold stone we mentioned before, the mass of air does not have physical unity by itself. If the surrounding marble is removed, the gold stone will remain as a distinguishable collection of matter, while the mass of air will not. The point is not, of course, that a mass of gas cannot form a body (stars do the trick through massive gravitational forces) but rather that the discreteness and unitary character of the object, to count as a body, must come as a function of the object itself.

This latter aspect is important for our discussion. A body is a material aggregation that, due to the internal forces that keep the proximity of its components, behaves as a unitary whole; that is, its material components are constrained by their own interactions to move as one object. In our previous example, the set of gas molecules trapped inside the marble cavity are in proximity, but not thanks to their own interactions. It is the spatial constraint imposed by the marble cavity which keeps the proximity of the gas molecules.

Another way to consider this aspect is to notice that the material components of a body tend to resist separation or disaggregation. If a given aggregation of matter offers null (or near to null) resistance to disaggregation, then that aggregation is not a body. Bodies can be partitioned and separated into pieces, of course, but not without resistance.

Bodies, it should be clear, need not have immutable boundaries or fixed components. The boundaries of a body may be deformable, and its components renewable over time. The critical point is that, despite deformation and renovation, the object remains as an identifiable physical unity.

Surely more detailed characterizations can be given for a technical notion of a body, and it is beyond the reach of this article to provide for such a notion a formal (neat, absolute) definition.[4] However, as we will see, the key distinction we need to establish in the present analysis is between the notion of a body and the notion of a system (machine or network), for which the idea of the body provided here should be enough.

Living beings appear to us as discrete physical bodies. Microorganisms, plants, insects, mammals, and so on, are all discrete bodies delimited by certain physical boundaries: membrane (in unicellulars), exoskeleton (in arthropods), skin (in many vertebrates and invertebrates), and so on. These and other similar biological structures contribute to maintaining the physical unity of living beings through cohesive/adhesive forces and are the product of the poietic (synthetic) activity of the living beings (McMullin, 2000; Razeto-Barry, 2012). Arguably, we might say that being a discrete physical body is a universal condition of living beings. Or, put another way, that it is hard to think of a living being that is not at the same time a discrete physical body.[5]

This bodily aspect of living beings might sound trivial for most readers. However, in the context of our discussion, it is important to note that, though perhaps trivial, the bodily character of living beings is not sufficiently established in AT's original formulation. Recall the definition given by Maturana and Varela (1980):

> *An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces the components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in the space in which they (the components) exist by specifying the topological domain of its realization as such a network.* (pp. 78–79, original emphasis)

We can distinguish (at least) three aspects in this formulation. First, the definition, loyal to AT's mechanistic spirit (see Section 2), presents living beings as multiple realizable autopoietic systems with respect to their material constitution. We are presented with a machine (system) that is organized or defined as a network of processes of production, without specifying what kind of products the machine produces or what kind of components are involved in the production. The unity of analysis is strictly functional, that is, a network of processes of production.

Second, there is the aspect of a circularity (or recursivity) in the working of the network (point (i) in the definition), such that it can be said that the network produces *itself*. This aspect is expressed in the prefix "auto" in the notion of autopoiesis.

Third, there is a topological aspect (point (ii) in the definition). It is said that the components produced by the network "specify" the topological domain of the network. This topological aspect, presumably, aims to capture the discrete and spatially bounded nature of living beings we have just highlighted (McMullin, 2000; Thompson, 2007). However, as we shall see, the topological specification is by itself too weak to establish the discrete bodily nature of living beings.

AT, notice, speaks in terms of autopoietic *machines* or *systems*. Even when AT brings the formal notion of autopoiesis to the physical domain, stating that living beings exist and are constituted in the molecular domain, it insists in talking of living beings as physical (molecular) autopoietic *machines* or *systems*. The problem with the notions of "machine" and "system" is that while they allow for topological specifications, they fail to unambiguously refer to a discrete bodily entity. That is, functional machines and systems can be said to have a defined topology despite lacking a discrete bodily constitution. This is because all *real* functional systems have a physical base or implementation, and to the extent that this base or implementation is physical, these systems have a determined localization in physical space, that is, a topology. However, not every physical topology constitutes or corresponds to a physical body. An engine that performs combustion is a physical body that performs combustion. A radio wave communication system has a physical implementation too, and therefore a topology, but it does not constitute a physical body. The thermostat, to take again our first example, considered as a material apparatus, constitutes a physical body with a certain spatial configuration delimited by physical boundaries. It has a defined topology as a body, such that it can be moved here or there, packed in a box, installed on a wall, and remain as a discrete cohesive physical entity. When working, the thermostat becomes a part of the functional system that we call the thermostatic system, which has its own topology. The physical implementation of the thermostatic system runs through the thermostat and the air to which this is coupled. However, although the thermostatic system has a defined topology, said topology does not constitute a discrete physical body. If you uninstall the thermostat, move it, and set it up in another house, the air that was coupled to it and that constituted a part of its physical implementation as an homeostatic system will not remain as a part of the system. The air and the thermostat do not behave as a unitary whole; they do not move as one object. Or, put another way, the air and the thermostat offer no resistance to their separation. The thermostat and the air are functionally coupled and define the topology of the thermostatic system but do not constitute a body.

When AT defines the autopoietic machine as a network of processes of production of components which constitute the machine as a concrete unity "by specifying the topological domain of its realization as … a network" (Maturana & Varela, 1980, p. 79), it says something too weak and general to establish and unambiguously denote the discrete bodily character of living beings. For the operationally closed system of processes that enable the existence of the living being, which extended enactivism correctly identifies as a network that goes beyond the boundaries of the living being, meets this definition but does not constitute a living body.

Take the case of extracellular or extraorganismic digestion alluded to by enactivists (Virgo et al., 2011). The worm produces and secretes enzymes that catalyze the decomposition of large molecules in the external medium, whose products then cross back within the organism's boundary and enter into metabolic reactions. Does this feedback network constitute an autopoietic *system* or *machine*? Yes, it does. It is a network of production processes (it is *poietic*). It is circularly or recursively organized, such that the products of some parts of the network participate in the production of the components of other parts of the network, and vice versa (it is *auto*poietic). It has a defined spatial distribution, to the extent that the molecular components and the chemical reactions that constitute the network have a defined spatial location, some of them inside the organism, some of the outside (it has a *topology* specified by the components of the network). Fair enough.

Now, does the feedback network of extraorganismic digestion constitute an autopoietic *body*? We think it does not. Similar to the case of the thermostat, while the organism, say a worm, remains as a discrete cohesive body, the extended circuits and networks associated to it do not. You can pick the worm up and locate it in a different place, and the chain of soil chemical reactions that constitute its extraorganismic digestion will not move with it. The components of those chemical reactions will offer null resistance to the separation (in comparison, at the right scale of analysis, to the resistance the worm's flesh would offer to be separated into pieces).

To live, organisms certainly need to engage with the environment in a set of feedback loops and operationally closed circuits. However, these circuits, essential as they may be, do not constitute discrete cohesive bodies, which is, we argue, one of the most recognizable features of living beings.[6] Or at least, it is one of the features that AT does not (and should not) want to overlook when defining living beings.

Having reached this point in our discussion, it is time to offer our proposal. AT, we think, should be reformulated by making explicit reference to the bodily dimension of living beings as follows:

Every living being is an autopoietic body

We offer this statement not as an exhaustive definition of the living being, since a full definition in terms of AT should unpack the very notion of autopoiesis in the definition,[7] but as an abbreviated formulation that makes explicit the bodily character of living beings. Under this formulation, we think, it becomes clear that while all living beings are, by implication, autopoietic machines or systems, not all autopoietic machines or systems are living beings. AT, according to our view, should define living beings as autopoietic bodies and not simply as autopoietic machines or systems. Interestingly, it also becomes clear that living beings, as autopoietic bodies, can and usually do form part of larger autopoietic systems, machines or networks, which is, we want to believe, the point that enactivists really wanted and want to make.

With the distinction between autopoietic bodies and autopoietic systems, we think, the enactive interpretation of autopoiesis as something that can exceed the boundaries of living beings gets corrected and earns a more clear formulation. Enactivists may say now, correctly, that some autopoietic systems extend beyond the physical boundaries of living beings, without having to say that living beings themselves get extended. To take again the enactivists' example, we may consider that the worm is a part of a larger autopoietic system or network that includes, for example, soil chemical reactions (and perhaps other processes), while recognizing that in this system it is the worm, and only the worm, the part that constitutes an autopoietic body, and therefore, a living being.

Our proposal presents the body and the autopoiesis as necessary conditions to individuate a living being.[8] An autopoietic system, machine or network that does not constitute a body, is not a living being. It has the right kind of organization as a productive network but does not generate the right kind of physical unity (in the next section we will argue that entities such as autocatalytic networks and candle flames fall into this category). A body whose physical constitution as a body is not generated by its own poietic activity is not a living being either. It has the right kind of physical constitution, but not the right kind of generative process for such a constitution (any cohesive body, such as a stone or a chair, falls into this category). This reading requires, additionally, a strong causal link (or metaphysical dependence) between body and autopoiesis, and not a mere conjunction of conditions. For instance, a body that hosts some autopoietic process but whose condition as a body is not generated by this process is not a living being (in the next section we will argue that our planetary ecological system, which some see as a superorganism, might fall into this category).

An autopoietic body is, essentially, a body that produces itself, in the sense that it produces its own material components *as well as* the bodily physical unity that characterizes it. This aspect can be made explicit by introducing our notion of a body in (a simplified version of) the original formulation of autopoiesis:

> An autopoietic body (i.e. a living being) is a body constituted as a network of processes of production of

> components that produces the components which, through their interactions and transformations continuously regenerate and realize (i) the network of processes that produced them; and (ii) the very body that they constitute in the physical space.

Suppose, for the sake of the discussion, that we provisionally accept the idea that living beings are autopoietic bodies, and also the idea that said bodies are typically embedded in (larger) autopoietic systems or networks. What would be the implications of such a view? In the next and final section, we will say some few words about this.

6. Conclusion

We have argued that living beings are a subclass of autopoietic systems, namely, those that constitute autopoietic bodies. We have also argued that living beings, in normal conditions, form part of larger autopoietic systems that include environmental networks. Such autopoietic systems, we have said, are part of what maintains the autopoiesis of living beings, but do not constitute living beings.

Although the logical possibility of extended interpretations of living beings' autopoiesis was opened by the disembodied nature of the original definition, such an interpretation was not (and is not), we have argued, the intended one in AT as a research program. With a simple respecification focused on the identity of the autopoietic body of living beings, AT, we think, is able to account for the manner in which biology typically distinguishes living beings from other natural systems, and also, to rectify and accommodate the extended enactive view of autopoiesis. In what follows, we will briefly illustrate these points through some examples.

The embodied formulation of AT we have proposed here might help, we think, to decide between some instances that are usually problematic when we try to distinguish, *in autopoietic terms*, living beings from other natural systems. For example, systems in which some sort of autopoietic process takes place are usually presented as problematic cases for AT. Candle flames contain sets of chemical reactions that deliver products (poietic processes), some of which seemingly loop back in some further reactions (autopoietic reactions). Autocatalytic networks not only contain but they *are constituted* as autopoietic networks. Why do not we and biology consider these systems as living beings? We answer that these systems are not living beings because, although they exhibit autopoietic processes, they do not constitute bodies. Candle flames and autocatalytic networks do not constitute bodies in the relevant sense we have introduced in this article. Candle flames appear as objects in our observation with a more or less defined shape, but their material components do not conserve proximity. On the contrary, they constitute a flow of exothermic reactions whose elements are in constant spatial dissipation (see Razeto-Barry, 2012; Razeto-Barry & Ramos-Jiliberto, 2013). The components of the autocatalytic network exhibit proximity (otherwise they could not react), but such proximity is provided by the physical container of the chemical solution in which the network takes place (e.g. the test tube), not by the components of the autocatalytic set itself.

Notice that in our analysis, candle flames and autocatalytic sets are discarded as living beings not because they lack a specific structure that acts as a boundary. Some authors, when analyzing minimal cases of autopoiesis, consider that the presence of a specific structure serving as a boundary (like the cell membrane) is a necessary condition to qualify a system as a living being (Fleischacker, 1988; McMullin & Varela, 1997). This strategy, we think, is not entirely correct, since there are cases of living cells that manage to survive without a cell membrane (Kim, Klotchkova, & Kang, 2001). Living cells that lack a cell membrane are autopoietic bodies because despite lacking a dedicated structure acting as a boundary, they still manage to keep the proximity of their components constituting a discrete and distinguishable unity in the physical space, while maintaining their autopoietic activity (for the interesting details of this phenomenon, see Kim et al., 2001). Candle flames and autocatalytic networks lack boundary structures, true, but that is not a fundamental reason to discard them as living beings. The reason is, rather, that these kinds of systems do not keep their material components in proximity, failing to constitute the physical unity that we identify as a body.

Neither our analysis appeals to the fact that two or more autocatalytic sets, or candle flames, merge when put together. Some have argued that a distinctive feature of living beings, unlike the systems mentioned above, is that they do not merge when put together (McMullin, 2000). This argument, we think, is not correct, since there are known cases of living beings, such as the amoebas, that fuse and form a single organism (a Plasmodium). Two or more autocatalytic networks (or candle flames) merge when put together, true, but that is not a fundamental reason to discard them as living beings. The reason, again, is that these kinds of systems do not keep their material components in proximity, failing to constitute the physical unity that we identify as a body. In contrast, to follow with the example, either as a unicellular amoeba or as a plasmodium, the living being is always a discrete physical body.

If the embodied reformulation we have provided here is sound, it might also lead us to reconsider the way we understand living beings' ecological niches. Ecological niches, under this view, might turn out to be, at least partially, autopoietic systems of a certain scale. Even the Gaia theory might get a more clear (and less contentious) formulation (Lovelock, 1979, 1988; Margulis, 1998). We might be open to seeing Gaia, that is, the whole planetary ecological system (including biota, atmosphere, oceans, and soils) as a massive autopoietic system, without this leading us to consider, in any serious sense, such a system as a super living being. Why? Some might argue that Gaia, as a whole, constitutes a kind of body (at the right scale of observation), and that, since it contains autopoietic networks, should be identified as a living being. In Section 5, however, we specified the way we read the claim that living beings are autopoietic bodies, saying that the claim does not mean mere co-occurrence of conditions but a causal link or metaphysical dependence between autopoiesis and body, such that it can be said that it is the body itself that is autopoietic. The components of Gaia (biota, atmosphere, oceans, soils) are kept in proximity, but such a condition is due to the gravitational force exerted by the Earth's mass, for which the autopoietic loops that may take place in Gaia are not essential. It is a certain quantity of matter what is critical to keep the proximity of the components (the only requirement to exert gravitational force), not a specific organization of processes of production. In a living being, by contrast, the proximity of its components is not merely a function of a certain quantity of matter but of a certain dynamic organization of components, which, through its poietic activity, maintains the physical unity of the system as a body.

Interestingly, and going deeper in this (admittedly highly speculative) comment, we might think of Gaia as an autopoietic system for which living beings are, ultimately, a facultative component. According to some interpretations, before the appearance of living beings on Earth, and rather as a prerequisite for such an event, the Earth surface, oceans, and atmosphere constituted already a complex system of chemical reactions and proto ecological cycles (Morowitz, 1992). If that were the case, the way we picture the origins of life on Earth might give us the view of an early planetary autopoietic system in whose space, gradually and perhaps through millions of years, some subproducts in the form of autopoietic bodies (or somethings similar enough to them) started to emerge, giving origin to life as we know it.

Acknowledgements

Mario Villalobos wants to thank Simon McGregor for helpful discussions. He also wants to thank Nathaniel Virgo and Alexandra Penn for their detailed and lucid comments on an earlier version of this paper. Finally, both authors want to thank the anonymous referees for their very constructive observations.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the Universidad de Tarapacá, grant UTA Mayor 3737-16.

Notes

1. The extended interpretation we address in this article is a particular version of enactivism and does not (necessarily) represent the general view of the enactive community. Evan Thompson, for example, a prominent enactivist, does not seem to subscribe to this extended view (see Thompson, 2007). In the specific context of this article, however, for the sake of simplicity, we will use the term "enactivists" to refer to the enactive authors who subscribe to the extended view.
2. Enactivists are neither the only nor the first theorists to promote an extended view of biological systems. An important precursor of this view is Turner (2000), who argues that the physiology of (some) living beings includes the external environment. He, however, does not link this argument to the autopoietic theory, which is our concern here.
3. The same idea is developed in Thompson (2007, p. 107), with the insight, correct in our view, that what matters is the fact that the autopoietic network produces its own physical demarcation, its own boundary, not the structural specificities of such a boundary (see also Bourgine & Stewart, 2004; Razeto-Barry, 2012).
4. We admit that many borderline cases might be presented as counterexamples or problematic instances.
5. Discreteness, notice, does not necessarily mean compactness. The roots of trees and plants do not form compact (agglutinated) entities but are spatially very distributed bodies. These roots, however, though very scattered, are distinguishable from the subsoil that nourishes them, and therefore constitute, in the sense that is relevant for our argument here, discrete bodies. In this line, for example, the so-called "superorganisms" such as the Pando aspen grove (Utah), or the humongous fungus (Michigan), are not exceptions to this rule. In both cases, there is a single massive root system that, though spatially very scattered, remains as a distinguishable and discrete body with respect to the subsoil.
6. This criterion, notice, establishes the necessity of a discrete body, not the invariance of the body itself. We comment on this point in Section 6.
7. And should also give, probably, some specifications about the thermodynamic context in which the formula applies (see Razeto-Barry & Ramos-Jiliberto, 2013).
8. We are not sure, for the moment, whether they are also jointly sufficient conditions to unambiguously individuate a living being. However, the purpose of the present work is not to come up with a final definition of living beings, but to improve the AT's original definition of living beings.

ORCID iD

Mario Villalobos https://orcid.org/0000-0003-0567-2888

References

Bourgine, P., & Stewart, J. (2004). Autopoiesis and cognition. *Artificial Life*, *10*, 327–345.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*, 7–19.

Colombetti, G. (2015). Enactive affectivity, extended. *Topoi*, *36*, 445–455.

Di Paolo, E. A. (2009). Extended life. *Topoi*, *28*, 9–21.

Fleischacker, G. (1988). Autopoiesis: The status of its system logic. *Biosystems*, *22*, 39–47.

Kim, G. H., Klotchkova, T. A., & Kang, Y.-M. (2001). Life without a cell membrane: Regeneration of protoplasts from disintegrated cells of the marine green alga Bryopsis plumosa. *Journal of Cell Science*, *114*, 2009–2014.

Lovelock, J. E. (1979). *Gaia: A new look at life on earth*. New York, NY: Oxford University Press.

Lovelock, J. E. (1988). *The ages of Gaia*. New York, NY: W. W. Norton.

Margulis, L. (1998). *Symbiotic planet: A new look at evolution*. New York, NY: Basic Books.

Maturana, H., & Varela, F. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Maturana, H., & Varela, F. (1998). *The three of knowledge* (Rev. ed.). Boston, MA: Shambhala.

McGregor, S., & Virgo, N. (2011). Life and its close relatives. In G. Kampis, I. Karsai, & E. Szathmáry (Eds.), *Advances in artificial life: Darwin meets von Neumann* (ECAL 2009. Lecture notes in computer science; Vol. 5778, pp. 230–237). Berlin, Germany: Springer.

McMullin, B. (2000). Remarks on autocatalysis and autopoiesis. *Annals of the New York Academy of Science*, *901*, 163–174.

McMullin, B., & Varela, F. (1997). Rediscovering computational autopoiesis. In P. Husbands, & I. Harvey (Eds.), *Proceedings of the 4th European conference on artificial life* (pp. 38–47). Cambridge, MA: The MIT Press.

Menary, R (Ed.). (2010). *The extended mind*. Cambridge, MA: The MIT Press.

Morowitz, H. J. (1992). *Beginnings of cellular life: Metabolism recapitulates biogenesis*. New Haven, CT: Yale University Press.

Razeto-Barry, P. (2012). Autopoiesis 40 years later: A review and a reformulation. *Origins of Life and Evolution of Biospheres*, *42*, 543–567.

Razeto-Barry, P., & Ramos-Jiliberto, R. (2013). ¿Qué es autopoiesis? [What is autopoiesis?] In P. Razeto-Barry, & R. Ramos-Jiliberto (Eds.), *Autopoiesis: Un concepto vivo* [Autopoiesis: A living concept] (pp. 27–57). Santiago, Chile: Editorial Nueva Civilización.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge, MA: Harvard University Press.

Turner, J. S. (2000). *The extended organism: The physiology of animal-built structures*. London, England: Harvard University Press.

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

Virgo, N., Egbert, M. D., & Froese, T. (2011). The role of the spatial boundary in autopoiesis. In G. Kampis, I. Karsai, & E. Szathmáry (Eds.), *Advances in artificial life: Darwin meets von Neumann* (ECAL 2009. Lecture notes in computer science; Vol. 5777, pp. 240–247). Berlin, Germany: Springer.

Wheeler, M. (2010). Mind, things, and materiality. In C. Renfrew, & L. Malafouris (Eds.), *The cognitive life of things: Recasting the boundaries of the mind* (pp. 29–38). Cambridge, UK: McDonald Institute for Archaeological Research Publications.

Wheeler, M. (2017). The revolution will not be optimised: Radical enactivism, extended functionalism and the extensive mind. *Topoi*, *36*, 457–472.

About the authors

Mario Villalobos is an associate professor of Philosophy at the University of Tarapaca, Arica, Chile, and an associate researcher at the Institute of Philosophy and Complexity Sciences, Santiago, Chile. His area of research is the philosophy of mind and cognitive science, focusing mainly on the autopoietic theory of life and cognition.

Pablo Razeto-Barry is the director of the Institute of Philosophy and Complexity Sciences, Santiago, Chile. His research lies in the intersection of philosophy of science and theoretical biology, focusing especially on evolutionary theory and autopoiesis.