Autonomy: a review and a reappraisal

Tom Froese, Nathaniel Virgo, Eduardo Izquierdo

CSRP 591

June 2007

ISSN 1350-3162

University of Sussex

Cognitive Science
Research Papers

Autonomy: a review and a reappraisal

Tom Froese, Nathaniel Virgo and Eduardo Izquierdo

Centre for Computational Neuroscience and Robotics (CCNR)
Centre for Research in Cognitive Science (COGS)
University of Sussex, Brighton BN1 9QH, UK

{t.froese, n.d.virgo, e.j.izquierdo}@sussex.ac.uk

Abstract

In the field of artificial life there is no agreement on what defines 'autonomy'. This makes it difficult to measure progress made towards understanding as well as engineering autonomous systems. Here, we review the diversity of approaches and categorize them by introducing a conceptual distinction between *behavioral* and *constitutive* autonomy. Differences in the autonomy of artificial and biological agents tend to be marginalized for the former and treated as absolute for the latter. We argue that with this distinction the apparent opposition can be resolved.

To appear in: F. Almeida e Costa *et al.* (eds.), *Proceedings of the 9th European Conference on Artificial Life,* Berlin, Germany: Springer-Verlag, 2007

1. Introduction

Two major research goals of artificial life are to 1) synthesize autonomous agents, and 2) through this process gain a better understanding of the generative mechanisms underlying autonomy in general. But what do we mean when we say that a system is *autonomous*? There seems to be no commonly accepted definition in the artificial life community or the cognitive sciences. For example, in engineering and robotics the notion of autonomy is often used to refer to the self-sufficiency of a machine to achieve a certain task (e.g. Brooks 1991; Pfeifer 1996), in artificial life the term 'autonomy' is commonly used to characterize self-organizing systems (e.g. Wheeler 1997; Nolfi & Floreano 2000, p. 117), Kauffman (2000) uses the term "autonomous agent" to refer to a life cycle constituted by thermodynamic work, and in the autopoietic and enactive tradition it is used to refer to the self-constitution of an identity in living systems (e.g. Weber & Varela 2002). Still, in spite of the evident definitional ambiguity there is arguably a sense in which most uses of the term 'autonomy' are united by a common concern with self-governance, a notion which is already implied by the term's etymology (*auto* [self] *nomos* [law])[1].

Nevertheless, the particular kind of self-governance which these authors have in mind can vary considerably. Indeed, due to the lack of a coherent conceptual framework which connects the different uses of the term, it is hard to measure the progress that has been made in the artificial synthesis of such systems. Are today's systems more autonomous than those presented at the first ECAL over 10 years ago? If this is the case, then what are the significant challenges that remain? And are current research methodologies appropriate for addressing them? In order to provide answers to these questions an understanding of autonomy is needed which enables the different uses of the term in artificial life and the cognitive sciences to be systematically related to each other. The aim of this paper is to provide a first step towards this necessary conceptual clarification.

2. Autonomy: a review

In this section the various uses of the term 'autonomy' are categorized into two main classes of approaches according to whether the focus is on the agent's 1) external behavior, or 2) internal organization. We introduce a conceptual distinction between *behavioral* and *constitutive* autonomy in order to differentiate between the type of autonomy referred to by 1) and 2), respectively.

2.1 Behavioral autonomy

For this class of approaches, it is generally a necessary condition that the behavior of an autonomous system is characterized by some capacity for stable and/or flexible interaction with its environment. The system's identity can be self-constituted (as is the case for all organisms), but it is sufficient for it to be externally imposed by some designer (e.g. the unit of selection in evolutionary robotics), or even explicitly represented by a particular component of the system (e.g. the central controller in

---

[1] The word 'autonomy' can also appear in an unrelated mathematical sense of meaning a dynamical system with no time dependence, which is another potential source of confusion.

GOFAI). Thus, this category includes all of those approaches which do not treat the autonomy of living beings as qualitatively (though, perhaps, quantitatively) different from the autonomy of most artificial agents. Three sub-categories can be distinguished:

1) The broadest use of the term ‘autonomy’ can be found in the context of engineering where the study of “autonomous systems” is basically equated with a concern for building robots (e.g. Smithers 1992). Thus, there is a sense in which even remotely controlled mobile robots (e.g. a Mars explorer) can be referred to as “autonomous agents” (e.g. Franklin 1995, p. 37). However, more commonly the notion is used to designate that the robot is engineered so as to be able to interact with its environment without requiring ongoing human intervention (e.g. Nolfi & Floreano 2000, p. 67). Brooks (1991), for example, uses the notion of autonomy to refer to tether-free robots, where all the energy and computational requirements are stored on board. Note that using the term ‘autonomy’ in this broad manner does not exclude agents whose behavior has been completely pre-specified. As such it can be criticized on the basis that the “agent can hardly be said to be autonomous because its behavior is largely dictated by the experimenter” (Nolfi & Floreano 2000, p. 148). A more restrictive notion is used by Pfeifer (1996) who proposes as the first design principle of autonomous agents that “they have to be able to function without human intervention, supervision, or instruction”. Nevertheless, it is clear that these requirements for autonomy are almost trivially fulfilled by many artificial agents and all organisms.

2) It is also often claimed that an autonomous system must be capable of satisfying some goal (or even of generating its own goals). For example, Beer (1995, p. 173) uses the term “autonomous agent” to mean “any embodied system designed to satisfy internal or external goals by its own actions while in continuous long-term interaction with the environment in which it is situated”. Similarly, Nolfi and Floreano (2000, p. 25) hold that “autonomous systems are expected to survive in unknown and partially unpredictable environments by devising their own goals and finding out solutions to challenges that may arise”. The way in which teleological concepts such as purpose, agenda, concern, or goal are used in this kind of approaches should generally be interpreted as rather loose metaphors. As a point in case, consider Franklin’s (1995, p. 233) use of these terms when he invites us to “think of an autonomous agent as a creature that senses its environment and acts on it so as to further its own agenda”, and then continues by claiming that “any such agent, be it a human or a thermostat, has a single, overriding concern – what to do next”. Following Beer (1995), we can say that in this context “the class of autonomous agents is thus a fairly broad one, encompassing at the very least all animals and autonomous robots”.

3) Another common approach is to relate autonomy to the robustness and flexibility of behavior. Smithers (1992), for example, claims that “autonomous systems” are those that “engage in specific kinds of task achieving behavior in particular real environments, and which do so reliably and robustly”. This view often relates autonomy to notions of self-organization (e.g. Wheeler 1997) and emergence (e.g. Nolfi & Floreano 2000, p. 117). While this sometimes implies some philosophical commitment (e.g. Bourgine & Varela 1992), it primarily manifests itself as a pragmatic response to the practical difficulties faced by the GOFAI tradition. For example, the approach for designing autonomous systems proposed by Pfeifer and Verschure (1992) “promises to resolve a number of fundamental problems of AI in

natural ways (such as situatedness and robustness), others will not need to be solved since they are artifacts of the traditional approach (e.g. symbol grounding)".

2.2 Constitutive autonomy

This category includes all approaches to autonomy which can be traced to the autopoietic tradition, a movement which originated in theoretical biology in the 1970's (e.g. Varela, Maturana & Uribe 1974; Maturana & Varela 1980), and/or which are generally related to metabolism (e.g. Moreno & Ruiz-Mirazo 1999; Ruiz-Mirazo & Moreno 2000). It is generally claimed that autonomy in living systems is a feature of self-production or *autopoiesis*[2]. However, this restriction of autonomy to living systems is unsatisfactory because we also want to refer to some systems as autonomous even though they are not characterized by metabolic self-production, for example artificial and social systems (Luisi 2003).

Thus, the original account was followed by an attempt to conceptually separate the notion of autonomy from that of autopoiesis. In 1979 Varela published his *Principles of Biological Autonomy*, a book that continues to be an important reference for many researchers (e.g. Di Paolo 2005; Beer 2004; Bourgine & Stewart 2004; McMullin 2004; Ruiz-Mirazo & Moreno 2000), and in which he formulated the 'Closure Thesis' which states that "every autonomous system is organizationally closed" (Varela 1979, p. 58)[3]. Accordingly, autopoietic systems are reinterpreted as one rather prominent member of a broader class of autonomous systems. Weber and Varela (2002) neatly summarize this position by proposing that we should identify the "'constitution of an identity' as the governing of an autonomy principle". The idea is that this principle should make it possible to "take the lessons offered by the autonomy of living systems and convert them into an operational characterization of *autonomy in general*, living or otherwise" (Varela 1979, p. 55).

This conception of autonomy clearly poses a significant difficulty for many common methodologies in artificial life research. For if we accept the general claim that an autonomous system is a self-defining or self-constituting system, then it follows that all current robots and most (if not all) artificial agents are "by constitution non-autonomous insofar as their realization and permanence as unities is not related to their operation" (Varela, Maturana & Uribe 1974). However, it is worth pointing out that while the question of "whether or not one may want to make an autopoietic system is, of course, an ethical problem" it is still the case that "if our characterization of living systems is adequate, it is apparent that they could be made at will" (Varela 1979, p. 44), at least in principle. Indeed, there is research in artificial life which tries to understand the generative mechanisms underlying such constitutive autonomy.

---

[2] One recent definition of autopoiesis as the minimal organization of living systems is: "An autopoietic system is organized (defined as unity) as a network of processes of production (synthesis and destruction) of components such that these components: 1) continuously regenerate the network that is producing them, and 2) constitute the system as a distinguishable unity in the domain in which they exist" (e.g. Varela 1997; Weber & Varela 2002; Di Paolo 2005).

[3] An autonomous system can be defined in operational terms as a system with an organization that is characterized by processes such that "(1) the processes are related as a network, so that they recursively depend on each other in the generation and realization of the processes themselves, and (2) they constitute the system as a unity recognizable in the space (domain) in which the processes exist" (Varela 1979, p. 55). This is essentially the definition of autopoiesis but without the implication that the processes necessarily involve physical synthesis and destruction.

Two main approaches can be distinguished according to whether their target is the 1) computational or 2) chemical domain.

1) The field of computational autopoiesis (McMullin 2004) attempts to explore the nature of living systems with the use of simulations. This research program originated over a decade in advance of the first Santa Fe Workshop on Artificial Life with the publication of a seminal paper by Varela, Maturana and Uribe (1974) in which the authors outline the first model of an autopoietic entity. It has subsequently given rise to a whole tradition of simulating autopoiesis (McMullin 2004). However, the question of whether such research can generate genuine autopoietic systems is still the subject of debate, with some researchers claiming for various reasons that computational entities can not be autopoietic in principle (e.g. Letelier, Marin & Mpodozis 2003; Thompson 2004; Rosen 1991; Varela 1997). Nevertheless it is clear that such modelling research has the potential to clarify some of the key ideas underlying autopoiesis and draw attention to some of the central questions which still remain open (e.g. Beer 2004).

2) The field of chemical autopoiesis has been investigating the "creation of chemical models of cellular life that can be constructed in the laboratory" since the early 1990's (see Luisi (2003) for a recent overview). In this manner some of the problems of the computational medium are avoided, but there are other challenges which derive from working with the chemical domain. Nevertheless, this approach has the advantage that it allows theoretical questions to be addressed on the basis of concrete experimental phenomena (e.g. Bitbol & Luisi 2004).

It is worth pointing out that, as computational models are becoming increasingly realistic, it is possible to relate them with actual chemical realizations in a mutually informative manner (e.g. Mavelli & Ruiz-Mirazo 2007). Moreover, in contrast to most of the current work on behavioral autonomy, this kind of research has the potential to discover the conditions under which autonomous systems emerge spontaneously (rather than having their identity pre-defined by the experimenter), and, since it is well grounded in the actual laws of physics and chemistry, it could thereby provide the basis for a proper *naturalization* of the concept of autonomy (e.g. Ruiz-Mirazo & Moreno 2000).

3. Autonomy: a reappraisal

In the previous section we identified two main approaches to autonomy. The advantage of the behavioral approach is that it can generally accommodate both artificial and biological agents. At the same time, however, it has difficulties in specifying exactly what makes such systems autonomous. Consequently, the requirements are often trivially met in many cases. As an ambiguous and inclusive approach, it threatens to make the concept of autonomy meaningless. In contrast, the constitutive approach can provide a more precise definition in operational terms, but this has the undesirable consequence that its applicability is mainly restricted to actual organisms. It thus excludes most artificial life research from potentially contributing to our understanding of the generative mechanisms underlying autonomy in general. These considerations make it evident that there is a pressing need of finding a principled way of integrating these two approaches into one coherent framework of autonomous systems research.

Accordingly, in this section it is proposed that one useful way of clarifying this issue is to 1) conceptualize autonomy as a continuum that includes both behavioral and constitutive autonomy as two distinct dimensions[4], and 2) relate these dual dimensions of autonomy such that they appear as two interrelated aspects of one unifying concept (i.e. life).

3.1 Autonomy as a continuum

Following Boden (1996), we agree that "autonomy is not an all-or-nothing property. It has several dimensions, and many gradations" (see also Franklin (1995), p. 266), and propose that these dimensions are best captured by behavioral and constitutive autonomy. Boden (1996) also addresses these two distinct aspects when she claims that "an individual's autonomy is the greater, the more its behaviour is directed by self-generated (and idiosyncratic) inner mechanisms, nicely responsive to the specific problem-situation, yet reflexively modifiable by wider concerns". This is a good guideline, but we are still faced by the considerable challenge of devising the precise operational criteria for measuring these gradations. In particular, there are two main issues that need to be addressed: 1) how to operationalize the criteria for behavioral autonomy, and 2) whether the dimension of constitutive autonomy is best conceived of as continuous or binary.

1) It is evident that the behavioral dimension of autonomy is best conceived of as continuous, but it is not exactly clear how. This is largely due to the fact that important behavioral criteria are often undefined (e.g. the requirement of 'stability' and 'flexibility') or phrased in ambiguous terms (e.g. the requirement of 'goal generation'). Fortunately, the ongoing development of the dynamical approach in cognitive science is ensuring that better tools for characterizing the dynamics of behavior are being appropriated from mathematics (van Gelder & Port 1995). For example, Kelso (1995, p. 45) points out that in the mathematical theory of dynamical systems the "measurement of the time it takes to return to some observed state -- local relaxation time -- is an important index of stability", and that "instabilities are hypothesized to be one of the generic mechanisms for flexible switching among multiple attractive states." Furthermore, it has been shown that the evolutionary robotics framework (Harvey *et al.* 2005) can help to investigate the dynamics underlying the behavioral autonomy associated with stability and flexibility (e.g. Di Paolo 2003, Iizuka & Di Paolo submitted).

2) Constitutive autonomy, as captured by the notion of autopoiesis, is strictly speaking an all-or-nothing systemic property (Di Paolo 2005). Varela (1979, p. 27), for example, notes that "the establishment of an autopoietic system cannot be a gradual process: Either a system is an autopoietic system or it is not. [...] Accordingly, there are not and cannot be intermediate systems". Even if we follow Varela (1979, p. 55) in extending the class of autonomous systems to include all systems which constitute their own identity, it still seems to be the case that either a system is constitutively

---

[4] Also useful, but out of the scope of this paper, would be to include substrate requirements as a third dimension of autonomy. Some authors require autonomous systems to be real physical/chemical systems, whereas others will allow simulated entities to be autonomous within a computational world.

autonomous or it is not. Nevertheless, there might be ways of treating the constitutive dimension as continuous. Bickhard (2000), for example, holds that an autonomous system is one which actively contributes to its own persistence and that "autonomy in this sense is a graded concept: there are differing kinds and degrees of such 'active contributions'". Barandiaran and Moreno (2006) outline another promising approach when they write that "while self-organization appears when the (microscopic) activity of a system generates at least a single (macroscopic) constraint, autonomy implies an open process of self-determination where an increasing number of constraints are self-generated".

Another possibility would be to measure the dimensions of autonomy along an increase in organizational requirements. For example, one could go from negative feedback, to homeostasis, and finally to autopoiesis[5]. This might make it possible to trace behavioral and constitutive autonomy from what might be called a 'weaker' sense to a 'stronger' sense, a continuum which roughly coincides with a transition from a more technological to a more biological usage of the term, and which finally culminates in a complete restriction of the term's applicability to actual living organisms. However, if this hierarchy of organizational requirements is to be actually useful in measuring autonomy, further work needs to be done to define the terms and their relationships more precisely.

3.2 Life as constitutive and behavioral autonomy

After conceptually teasing the constitutive and behavioral domain of autonomy apart, it is nevertheless quite clear that they do somehow relate in living systems. Varela (1997), for example, relates constitutive autonomy to the behavioral domain: "To highlight autonomy means essentially to put at center stage two interlinked propositions: Proposition 1: Organisms are fundamentally the process of constitution of an identity. [...] Proposition 2: The organism's emergent identity gives, logically and mechanistically, the point of reference for a domain of interactions"[6]. However, it is a non-trivial question as to exactly how the organism distinguished in the constitutive domain relates to its behavior distinguished in the behavioral domain. Moreover, this connection only works for some conceptions of behavioral autonomy, and a more precise definition of how such autonomy relates to living systems is needed before the relationship can be stated more formally.

While such further conceptual clarification is important for the development of a coherent theory of autonomy, it is also of practical interest for current artificial life research. Bourgine and Stewart (2004), for example, conceptualize autopoiesis and cognition as distinct aspects of living systems in such a way that it allows them to refer to artificial agents as 'cognitive' without them having to be autopoietic. This view is clearly a useful theoretical justification for using evolutionary robotics as a methodology for studying behavioral autonomy in the form of cognition (e.g. Harvey *et al.* 2005) without having to address the problem of constitutive autonomy.

---

[5] Thanks to Barry McMullin for pointing this out. This hierarchy is enhanced when we consider that "an autopoietic machine is an homeostatic (or rather a relations-static) system which has its own organization (defining network of relations) as the fundamental variable which it maintains constant" (Maturana & Varela 1980, p. 79). See also Varela (1979, p. 13).

[6] This was clearly also a part of his vision for ECAL, as is evident in Bourgine and Varela (1992).

Similarly, Beer's (2004) approach to cognition follows directly from an autopoietic perspective on life when two key abstractions are made:

> 1) Focus on an agent's behavioral dynamics. An agent's behavior takes place within its cognitive domain, which is a highly structured subset of its total domain of interaction.
>
> 2) Abstract the sets of destructive perturbations that an agent can undergo as a viability constraint on its behavioral dynamics.

Thus, we assume the existence of a constitutively autonomous agent, but model only its behavior and not the constitutive aspects of its autonomy. In other words, the agent is constitutively autonomous by definition only.

However, there are reasons for holding that in living systems autopoiesis and cognition are more tightly interlinked than the possibility of strict conceptual separation seems to indicate (Bitbol & Luisi 2004). Thus, as Beer (1997) himself makes clear, some of the abstractions made in artificial life research are not completely satisfactory:

> "[T]his explicit separation between an animal's behavioral dynamics and its viability constraint is fundamentally somewhat artificial. An animal's behavioral dynamics is deeply intertwined with the particular way in which its autopoiesis is realized. Unfortunately, a complete account of this situation would require a theory of biological organization, and the theoretical situation here is even less well developed than it is for adaptive behavior. [...] However, if we are willing to take the existence of an animal for granted, at least provisionally, then we can assume that its viability constraint is given a priori, and focus instead on the behavioral dynamics necessary to maintain that existence" (Beer 1997, p. 265).

It is clear from these considerations that, while the general aim of evolutionary robotics is not to study the mechanisms underlying constitutive autonomy, more thought needs to be given as to how natural cognition is constrained by the constitutive processes which give rise to living systems. In this regard it might be helpful to introduce more biologically inspired mechanisms into the controllers of the artificial systems being evolved, for example homeostasis (e.g. Di Paolo 2003; Harvey 2004; Iizuka & Di Paolo submitted). However, in general more work needs to be done in order for us to better understand what kind of methodology is best suited for studying autonomous artificial systems which actually self-constitute an identity at some level of description. Only when we are able to investigate both constitutive and behavioral autonomy via synthetic means can the field of artificial life claim to provide one coherent framework of autonomous systems research.

4. Conclusion

Are today's artificial agents more autonomous? By distinguishing between behavioral and constitutive autonomy, we can see that this question actually demands two distinct responses. It seems safe to say that today's systems are indeed more behaviorally autonomous (than at the start of ECAL, for example). Most of the work that is done in the artificial sciences under the banner of autonomous systems research is providing a wealth of tools of analysis and ways of understanding of how externally defined constraints can be successfully satisfied by increasingly complex artificial agents. However, the vast majority of this kind of research is not tackling the question of how such viability constraints (and, more importantly, an agent's identity) can emerge from the internal operations of those autonomous systems while coupled to their environments, though more work is starting to be done in this area.

Finally, it is important to note that the widespread disregard of the dimension of constitutive autonomy is a serious shortcoming not only for scientific research, but also in terms of our own understanding of what it means to be human. As Boden (1996) points out: "what science tells us about human autonomy is practically important, because it affects the way in which ordinary people see themselves – which includes the way in which they believe it is possible to behave". The field of artificial life is therefore also faced by an ethical imperative to invest more effort into improving our understanding of constitutive autonomy. Only then can we ground our understanding of human freedom – not only in terms of the behavior involved in mere external constraint satisfaction, but also in terms of the creativity involved in dynamic and open-ended self-realization.

Acknowledgments

Many thanks to all the participants of the 'Modeling Autonomy' workshop which was held in San Sebastian, Basque Country, during the 22nd and 23rd March 2007, as well as those involved in the Life and Mind seminars at Sussex University, for their helpful discussions.

References

Barandiaran, X. & Moreno, A. (2006), “On what makes certain dynamical systems cognitive: A minimally cognitive organization program”, *Adaptive Behavior*, **14**(2), pp. 171-185

Beer, R.D. (1995), “A dynamical systems perspective on agent-environment interaction”, *Artificial Intelligence*, **72**(1-2), pp. 173-215

Beer, R.D. (1997), “The dynamics of adaptive behavior: A research program”, *Robotics and Autonomous Systems*, **20**(2-4), pp. 257-289

Beer, R.D. (2004), “Autopoiesis and Cognition in the Game of Life”, *Artificial Life*, **10**(3), pp. 309-326

Bickhard, M.H. (2000), “Autonomy, Function, and Representation”, *Communication and Cognition – Artificial Intelligence*, **17**(3-4), pp. 111-131

Bitbol, M. & Luisi, P.L. (2004), “Autopoiesis with or without cognition: defining life at its edge”, *Journal of the Royal Society Interface*, **1**(1), pp. 99-107

Boden, M.A. (1996), “Autonomy and Artificiality”, in: M.A. Boden (ed.), *The Philosophy of Artificial Life*, New York, NY: Oxford University Press, pp. 95-108

Bourgine, P. & Varela, F.J. (1992), “Introduction: Towards a Practice of Autonomous Systems”, in: F.J. Varela & P. Bourgine (eds.), *Proc. of the 1st Euro. Conf. on Artificial Life*, Cambridge, MA: The MIT Press, pp. xi-3

Bourgine, P., & Stewart, J. (2004), “Autopoiesis and Cognition”, *Artificial Life*, **10**(3), pp. 327-345

Brooks, R.A. (1991), “Intelligence without reason”, in: J. Myopoulos & R. Reiter (eds.), *Proc. of the 12th Int. Joint Conf. on Artificial Intelligence*, San Mateo, CA: Morgan Kaufmann, pp. 569-595

Di Paolo, E.A. (2003), “Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop”, in: K. Murase & T. Asakura (eds.), *Dynamical Systems Approach to Embodiment and Sociality*, Adelaide, Australia: Advanced Knowledge International, pp. 19-42

Di Paolo, E.A. (2005), “Autopoiesis, adaptivity, teleology, agency”, *Phenomenology and the Cognitive Sciences*, **4**(4), pp. 429-452

Franklin, S. (1995), *Artificial Minds*, Cambridge, MA: The MIT Press

Harvey, I. (2004), “Homeostasis and Rein Control: From Daisyworld to Active Perception”, in: J. Pollack *et al.* (eds.), *Proc. of the 9th Int. Conf. on the Simulation and Synthesis of Living Systems*, Cambridge, MA: The MIT Press, pp. 309-314

Harvey, I., Di Paolo, E.A., Wood, R., Quinn, M. & Tuci, E. A. (2005), "Evolutionary Robotics: A new scientific tool for studying cognition", *Artificial Life*, **11**(1-2), pp. 79-98

Iizuka, H. & Di Paolo, E.A. (submitted), "Toward Spinozist robotics: Exploring the minimal dynamics of behavioural preference", *Adaptive Behavior*

Kauffman, S. (2000), *Investigations*, New York, NY: Oxford University Press

Kelso, J.A.S. (1995), *Dynamic Patterns: The Self-Organization of Brain and Behavior*, Cambridge, MA: The MIT Press

Letelier, J.C., Marin, G. & Mpodozis, J. (2003), "Autopoietic and (M, R) systems", *Journal of Theoretical Biology*, **222**(2), pp. 261-272

Luisi, P.L. (2003), "Autopoiesis: a review and reappraisal", *Naturwissenschaften*, **90**, pp. 49-59

Maturana, H.R. & Varela, F.J. (1980), *Autopoiesis and Cognition: The Realization of the Living*, Dordrecht, Holland: Kluwer Academic Publishers

Mavelli, F. & Ruiz-Mirazo, K. (2007), "Stochastic simulations of minimal self-reproducing cellular systems", *Phil. Trans. R. Soc. B*, in press

McMullin, B. (2004), "Thirty Years of Computational Autopoiesis: A Review", *Artificial Life*, **10**(3), pp. 277-295

Moreno, A. & Ruiz-Mirazo, K. (1999), "Metabolism and the problem of its universalization", *BioSystems*, **49**(1), pp. 45-61

Nolfi, S. & Floreano, D. (2000), *Evolutionary Robotics: The biology, intelligence, and technology of self-organizing machines*, Cambridge, MA: The MIT Press

Pfeifer, R. (1996), "Building 'Fungus Eaters': Design Principles of Autonomous Agents", in: P. Maes *et al.* (eds.), *Proc. of the 4th Int. Conf. on the Simulation of Adaptive Behavior*, Cambridge, MA: The MIT Press, p. 3-12

Pfeifer, R. & Verschure, P. (1992), "Distributed Adaptive Control: A Paradigm for Designing Autonomous Agents", in: F.J. Varela & P. Bourgine (eds.), *Proc. of the 1st Euro. Conf. on Artificial Life*, Cambridge, MA: The MIT Press, pp. 21-30

Rosen, R. (1991), *Life Itself: A Comprehensive Inquiry into the Nature, Origin and Fabrication of Life*, New York, NY: Columbia University Press

Ruiz-Mirazo, K. & Moreno, A. (2000), "Searching for the Roots of Autonomy: The natural and artificial paradigms revisited", *Communication and Cognition – Artificial Intelligence*, **17**(3-4), pp. 209-228

Smithers, T. (1992), “Taking Eliminative Materialism Seriously: A Methodology for Autonomous Systems Research”, in: F.J. Varela & P. Bourgine (eds.), *Proc. of the 1st Euro. Conf. on Artificial Life*, Cambridge, MA: The MIT Press, pp. 31-40

Thompson, E. (2004), “Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela”, *Phenomenology and the Cognitive Sciences*, **3**(4), pp. 381-398

van Gelder, T. & Port, R.F. (1995), “It’s About Time: An Overview of the Dynamical Approach to Cognition”, in: R.F. Port. & T. van Gelder (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge, MA: The MIT Press, pp. 1-43

Varela, F.J. (1979), *Principles of Biological Autonomy*, New York, NY: Elsevier North Holland

Varela, F.J. (1997), “Patterns of Life: Intertwining Identity and Cognition”, *Brain and Cognition*, **34**(1), pp. 72-87

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

Weber, A. & Varela, F.J. (2002), “Life after Kant: Natural purposes and the autopoietic foundations of biological individuality”, *Phenomenology and the Cognitive Sciences*, **1**, pp. 97-125

Wheeler, M. (1997), “Cognition’s Coming Home: the Reunion of Life and Mind”, in: P. Husbands & I. Harvey (eds.), *Proc. of the 4th Euro. Conf. on Artificial Life*, Cambridge, MA: MIT Press, pp. 10-19