# **Quantifying Normative Behaviour and Precariousness in Adaptive Agency**

Matthew D. Egbert<sup>1\*</sup>, Xabier E. Barandiaran<sup>2</sup>

<sup>1</sup> Center for Computational Neuroscience and Robotics, University of Sussex, Brighton, U.K.
<sup>2</sup> Centre de Recherche en Epistemologie Appliquée, CNRS / Polytechnique, Paris, France. mde@matthewegbert.com

#### Abstract

An essential feature of autonomous adaptive agency is that a system behaves according to an intrinsic norm. In this paper, we illustrate and clarify this notion of "behavior according to an intrinsic norm" with a minimalistic model of agency. We present a minimal metabolic system whose auto-catalytic dynamics define a viability region for different concentrations of available resource or 'food' molecules. We initially consider the availability of food as a control parameter for metabolic dynamics. A bifurcation diagram shows that for fixed values of available food, there exists a viability region. This region has an non-zero stable equilibrium and a lower boundary that takes the form of an unstable equilibrium-below which, the tendency of the system is towards "death", a stable equilibrium with a zero concentration of metabolites. We define the viability region as that in which the system tends toward the "living" stable-equilibrium. Outside of this region, in the precarious region, the system may live for some time but will eventually die if the food concentration does not change. With a precise definition of system-determined death, living, precarious and viable regions we move on to reconsider the available concentration of resources ([F]), not as a free parameter of the system but as modulated by organismic behaviour. By coupling the metabolism to a behavioural mechanism, we simulate a stochastic, up-resource gradient climbing behaviour. As a result, the effect of behaviour on the viability space can be mapped and quantified. This lets us move closer to defining adaptive action more precisely as that course of behaviour whose effect is in accordance with an intrinsic normative field.

### Introduction

The way in which living systems (from bacteria to humans) actively regulate their relationship with their environments strongly contrasts with inanimate objects. This *agency* is widespread in nature and it continues to capture the attention of philosophers, theoretical biologists, psychologists and roboticist alike, for it has proven to be a difficult property to define, model or synthesise.

The notion of agency often carries with it closely related and traditionally problematic notions such as normativity, adaptivity, individuality, teleology, intentionality, goaldirectedness or free-will. Artificial life modelling techniques are well suited to provide a bottom-up approach capable of conceptually clarifying the systemic character of the properties associated with agency, its origins and nature.

After reviewing a wide variety of definitions and uses of the term 'agency' ranging from biology to robotics, Barandiaran et al. (2009) define agency as follows:

"an agent is an autonomous organization capable of adaptively regulating its coupling with the environment according to the norms established by its own viability conditions." (p.376)

In this paper we attempt to make more explicit what is meant by the expression "according to the norms established by its own viability conditions". Similar expressions have been used by Di Paolo (2005); Barandiaran and Moreno (2008); Skewes and Hooker (2009) but no model has yet been developed to illustrate and describe in detail the meaning of this expression (and others closely associated with it). The goal of this paper is to make progress in this direction using a minimalist model that can help understand and scientifically articulate a formal and quantitative definition of agency. To this end, we present a model that exemplifies the key concepts of "normative behaviour" in the context of agency. To further contextualize the model and its interpretation, in the next section we introduce the conceptual (i. e. philosophical) and theoretical problem and two contemporary approaches to it. We then introduce the design specifications of the model and analyze its dynamics and their interpretation in terms of normativity, precariousness, adaptivity and viability.

## Autonomous agency and normativity: some dynamic requirements

The issue of natural agency and norms is attracting increasing attention (Frankfurt, 1978; Burge, 2009; Di Paolo, 2005; Skewes and Hooker, 2009; Barandiaran et al., 2009; Silberstein and Chemero, 2011) and Artificial Life is very well suited to make some conceptual progress on key aspects of agency and its origins. In fact, minimal models of agency have been a recurring topic in the field (from protocellular models to robotics).



Copyright 2011, Matthew Egbert, Xabier Barandiaran Licensed under Creative Commons Attribution 3.0 Unported (http://creativecommons.org/licenses/by/3.0)

Figure 1: Classical picture of a Viability region defined for two essential variables (food and water), outside the region the system will die. Viable trajectories are those that remain within the boundaries of the viability region. Model generally focus on designing a control system that generates the appropriate trajectories inside the viability region but the region itself is given (e. g. arbitrarilly designed or experimentally determined).

Arguably, minimal forms of agency (like chemotaxis) encapsulate some of the most important properties of "higher" levels of agency (such as human agency). One such property is normativity: i.e. the dimension of behaviour in which value comes into play-in which actions are good or bad, adaptive or maladaptive, appropriate or inappropriate(Christensen and Bickhard, 2002; Barandiaran and Moreno, 2008; Burge, 2009). While artificial systems can be judged to operate in relation to norms, these norms have (thus far) always been defined by the designer of the artificial system or interpreted by an external observer or user. In other words, what is good or bad functioning for a robot, a car or a coffee machine has been a matter of the design specifications which are largely independent of the structure and organization of the artifact. This is unlike biological organisms that respond to norms that are more closely related to the organization of the organism itself and what is (or is not) conducive to its ongoing operation.

Philosophers and scientists have tried to justify this normative dimension of natural agency in two ways. The most popular is the evolutionary (Millikan, 1989) approach in which a behaviour is considered to be normative or adaptive if it has been selected by evolution. In this view adaptation is ultimately a result of natural selection and it is only as a result of a process of selection that a character or process (e. g. a pattern of behaviour) can be said to be adaptive or maladaptive. This evolutionary approach to etiology, faces nu-



Figure 2: The viability boundary is an unstable equilibrium between living and dead states.

merous problems. One of them is how to categorise the first instance of a particular adaptive (i.e. norm following) behaviour? If a norm depends on an evolutionary selective history, then the first case of a "norm following" behaviour does not qualify as norm-following until it has been selected. This is clearly unsatisfactory. A criteria that is independent from history and is instead grounded on the very organization of the system and its ongoing dynamics seems better suitedindeed required-if we are to derive a consistent definition of adaptivity and normative behavior. This is precisely the motivation underlying the main alternative approach to normativity and adaptation. The organizational approach (as it might be called) puts at its center the idea of autonomy; from the Greek *autos* = self, and *nomos* = norm (Varela, 1979; Ruiz-Mirazo and Moreno, 2004; Di Paolo, 2004; Kauffman and Clayton, 2006). Although the origins of this approach can be traced back to the works of Aristotle and Kant (his Critique of Judgement), it was through the relatively modern development of theoretical biology and the physics and chemistry of far-from-equilibrium systems that it entered the scientific discourse. The contemporary conception of the organisational approach contends that norms are to be found as conditions of viability of the system, sometimes depicted in adaptive behaviour literature as a viability region (see Figure 1) or discussed as 'viability constraints' (Ashby, 1952; McFarland, 1999; Aubin et al., 2011). A closely related term is that of precariousness (Jonas, 1966, 1968; Weber and Varela, 2002; Di Paolo, 2005; Barandiaran et al., 2009), related, but not identical to the notion of "being far from thermodynamic equilibrium" when the system is a chemical or metabolic system (Ruiz-Mirazo and Moreno, 2004). The idea is that natural agents are organisms (i.e. living systems) that stand always in precarious conditions: if they don't actively regulate their interaction with their environment (e.g. find food or a lower temperature) they will perish, since they exist in a continuous need of thermodynamic exchange with their environment. This precariousness is meant to form the basis of the normative character of behaviour: the system must actively seek to compensate its inherently decaying organization.

In a key paper where the theory of autonomy (in particu-

lar the autopoietic tradition) is complemented and expanded with Ashby's framework for adaptive behaviour, Di Paolo (2005) defined adaptivity (in relation to agency) as:

"a system's capacity, in some circumstances, to regulate its states and its relation to the environment with the result that, if the states are sufficiently close to the boundary of viability, 1. tendencies are distinguished and acted upon depending on whether the states will approach or recede from the boundary and, as a consequence, 2. tendencies of the first kind are moved closer to or transformed into tendencies of the second and so future states are prevented from reaching the boundary with an outward velocity."

Di Paolo's definition of adaptive agency could be explicitly modelled and formalized. However, most of the models that have been developed with similar approaches have failed to address two blind spots: (1) viability boundaries appear as given or defined from without and the models focus on how to shape adaptive dynamics to maintain the trajectories of essential variables within those boundaries; (2) as a consequence, the relationship between the organismic dynamics that define the boundaries and the dynamics that control adaptive behaviour remain decoupled. In previous work (Egbert et al., 2009, 2010b) we have explored the relations between the viability boundary determining metabolic dynamics and the dynamics that drive organismic behaviour A further problem remained however: although the boundaries of viability were directly linked to the modelled system, they were only defined by the system in a relatively trivial way. The boundaries of our models and similar efforts by others (see e.g. Ruiz-Mirazo and Mavelli, 2008) were the result of rough physical magnitudes: disappearance of the protocell due to complete lack of catalysts or bursting disintegration of the protocell marked by the upper limit of the tension of the membrane. The boundaries were not emergent from interactions between system processes in the holistic system-interdependent manner that characterizes integrity and systemic identity in real organism. In our previous models viability boundaries equated to absence of the system (i. e. total disintegration or zero quantity of its constituent elements). But, in natural systems, the limits of viability do not map with the physical disintegration of a system (Figure 2A), but rather with the loss of the capacity of the system to sustain itself. To lose viability is not to disappear altogether but to cross a much more subtle boundary after which the maintenance of life becomes impossible (Figure 2B). This boundary is the result of the dynamic organization of the system and, as we shall see, it defines a norm that behavioural patterns need to satisfy in order to be adaptive.

In this paper we model a minimal protocell-like system whose metabolic dynamics define an emergent viability boundary. For fixed concentrations of available resources, we can plot a bifurcation diagram of the chemodynamics



Figure 3: A) The influences of the forward and backwards flow of the autocatalytic reaction and degradation upon the concentration of A. B) The combined influence of the chemical reactions and degradation upon the concentration of A given a fixed concentration of [F] = 1.4.

that indicates the intrinsic boundaries of viability of the system. Different viability regions can be identified and the adaptive *norms* of the system clearly defined and quantified.

We then couple a gradient climbing behavioural mechanism to the metabolic dynamics. We show that in this metabolism-behaviour coupled system, the behaviour of the system can be directly mapped into the viability space of the simulated agent and it is possible to explicitly show and quantify how the system is adaptive for and by itself.

#### Model

### Minimal metabolism

The metabolic organisation of self-production is one of the most fundamental properties of living systems and has been studied as such by many (Kauffman and Farmer, 1986; Kauffman, 2003; Varela, 1979; Ruiz-Mirazo and Moreno, 2004). In creating and maintaining themselves, living systems define their own viability constraints-the necessary and sufficient conditions for their continued existence. Thus, for the present work, metabolism is particularly relevant because it captures precisely what we wish to study. In its minimal and essential form it suffices to model metabolism as the self-production of a chemical network through the transformation (by the network) of available resources into constituents of the network. In previous work (Egbert et al., 2010a,b, 2009) we have modelled these kinds of systems in more detail, but here we abstract the system into two categories of components that we use to approximate a more



Figure 4: The system is has a single stability at [A] = 0.0for low concentrations of F. At  $[F] \approx 1.1$  the system bifurcates, and for concentrations of F greater than this value, the system has two stable equilibria—the "living stable equilibrium" (where [A] > 0.0) and "dead" (where [A] = 0.0) and one unstable equilibrium, the viability boundary.

complicated metabolic system. These categories are 'food' or resource reactant(s) 'F' and metabolites, i. e. members of the autocatalytic set, 'A'. Note that this approximation can be read as a higher order description of a more complicated system where A might capture an order parameter of a complex network of reactions among multiple molecules. In fact a recent and more complex model by Piedrafita et al. (2010) can be taken as dynamically similar to the present one, although it has a higher number of metabolites and catalysts and addresses other theoretically relevant properties (such as catalytic closure—which despite its relevance for the overall project of defining life and agency we have decided to leave aside for the specific purpose of the present paper).

We approximate the global dynamics of a more complex network according to the following reaction in which two members of the autocatalytic set interact with F to produce a third member of the autocatalytic set.

$$2A + F \leftrightarrow 3A$$

Note that the arrow is bidirectional, meaning that the reaction can occur in either direction, as is the case for all chemical reactions. A rate constant is associated with each direction (forward and backward) of the reaction  $k_b = 0.45$ ,  $k_f = 1.0$ . In addition to this autocatalytic reaction, A is subject to degradation into lower energy chemicals that are assumed to have no subsequent effect on the the reaction and are therefore not modelled. The combined influence of the forward and backward autocatalytic chemical reaction and the degradation are simulated by the following differential equation in which the degradation constant  $k_d = 1.0$ .

$$[\dot{A}] = \frac{-k_b[A]^3}{6} + \frac{k_f[F][A]^2}{2} - k_d[A]$$

### Metabolism-based chemotaxis

To study how behaviour can be sensitive to the viability boundary, we couple the metabolism to a simple stochastic gradient-climbing behavioural mechanism known as "run or tumble". The run and tumble behaviour is inspired by the behaviour observed in *Escherichia coli* and other bacteria, that achieve chemotaxis through probabilistic modulation of two behaviours, "running" where the organism moves in a roughly straight line and "tumbling" where the organism chooses a new orientation at random. The mechanism modelled here is a form of *metabolism-based-chemotaxis*, meaning that no specific sensor nor chemical pathway is required to modulate behaviour; instead metabolism itself affects the behavioural probabilities so as to modulate the probability of tumbling (see Egbert et al., 2010b).

We have employed this coupling of metabolism and behaviour in previous papers to study the adaptability that such a coupling provides (Egbert et al., 2010b) and the possibility that an interaction between metabolism, behaviour and evolution can facilitate adaptive evolution of populations of protocells (Egbert et al., 2010a, 2011). Here we study how such a behavioural mechanism influences trajectories along the viability space.

In this case, the simulation of metabolism-based behaviour works as follows. The agent is considered to always be in a default state of running (moving in a straight line)  $\dot{x} = k\cos(\alpha), \dot{y} = k\sin(\alpha)$ . Tumbling occurs probabilistically with a likelihood that is modulated by the change in the concentration of A. If, since the previous iteration, [A] has decreased, the organism will tumble—i. e. a new orientation will be chosen from a flat distribution ( $\alpha = \operatorname{rnd}[0..2\pi]$ ). Otherwise, the agent will continue running. A tumble inhibits any further tumbling for 5 iterations.

This particular form of metabolism-based run tumble mechanism is a highly simplified approximation of the "derivative" method used by *E. coli* and that simulated in (Egbert et al., 2010a, 2011) that compares the current metabolic rate to its rate a few moments previous. A decrease in metabolic rate indicates a worsening situation and increases the chance of a reorientation of the organism. In this way, the organism performs a simple but highly effective and surprisingly adaptive (Egbert et al., 2010b) behavioural strategy that can be captured by the anthropocentrism "If things are going well, I'll keep going in this direction that I've been heading, otherwise, I'll go somewhere else."



Figure 5: Paths taken by a successful chemotactic agent (top plots) and an unsuccessful agent (bottom plots). The left plots indicate the path of the agents in space plotted against [A]. The surface at the bottom of the image indicates the concentration of F in the environment. The right images show the path taken by the agents through viability space (see Figure 4). Initial oscillations around the viability boundary are eventually replaced by a trend up to the "living stable" equilibrium, thanks to the chemotactic motion.

# Simulation results: metabolic and behavioural dynamics

# Metabolic dynamics: bifurcation line as viability boundary

We first consider the metabolic system independently of behaviour and study its dynamic for fixed concentrations of F. Intuitively, it is clear that with no food, [F] = 0, the system should be unable to maintain itself in the face of degradation. This is also the case for low concentrations of F. As we start to increase [F] however, the combined effect of its progressive disintegration and the forward and backward metabolic reactions of A leads to a bistable dynamic regime. The dynamic tendency of the three reactions and their combined effect for a fixed value of [F] = 1.4 can be seen in Figure 3A. It is clear from Figure 3B that this system has two stable equilibria, "death" at [A] = 0.0 and "living stable" at  $[A] \approx 7.5$ , with an unstable equilibrium, the viability boundary at  $[A] \approx 1.8$ .

Analysis of the metabolic dynamics for different, fixed values of [F] gives us the bifurcation diagram in Figure 4. For [F] > 1.1, there is enough food to maintain a non-zero concentration of A. In this area of the parameter space, the system has two stable equilibria: "living stable" (where A > 0.0) and "dead" (where [A] = 0.0) and one unstable equilibrium, the *viability boundary* (the dashed line in Figure 4). Below the viability boundary, the system tends towards the "dead" equilibrium.



Figure 6: A summary of the regions of viability space: living, death, viable and precarious. See main text for further explanation.

### **Chemotactic behaviour**

Figure 5 shows the trajectories of two different agents using the metabolism-based mechanism. The left-hand figures shows motion in space plotted against [A], the concentration of the autocatalyst. The right-hand figure shows the trajectory of the agent in "viability space" i.e. the same space as shown in the bifurcation diagram in Figure 4. The top images are for an agent that succeeded at performing chemotaxis. The lower images are the same, but for an agent that has had "bad luck" and the stochastic gradient climbing mechanism has failed.

# Model interpretation and discussion: Agency, precariousness, norms and adaptivity

This simple model suffices to satisfy a minimal requirement of normative behaviour, in that it generates a viability space where, living, viable, precarious, irreversible-terminal and death regions can be clearly identified. These are highlighted in Figure 6. The "dead region" can be clearly seen as the zero concentration of the required metabolites (a complete disintegration of the system). A viable region is identified where, given a fixed supply of resources, the system will maintain itself, growing or shrinking until it reaches the "living stable" equilibrium. The arrows indicate the tendency of metabolic dynamics for different regions of the viability space. The viable region can be precisely defined for a range of the parameter [F] and a range of initial conditions [A] as the subregion of the living space where for each point the evolution of the system will tend toward the stable living equilibrium. The unstable equilibrium at the bottom of the



Copyright 2011, Matthew Egbert, Xabier Barandiaran Licensed under Creative Commons Attribution 3.0 Unported (http://creativecommons.org/licenses/by/3.0)

Figure 7: The idea of the 'normative field' in the precarious region – the effects of behaviour as efforts to move the system into the region of viability.

viable region defines a lower boundary of viability below which, the system tends toward death. For small values of [A] and [F] we can distinguish a *precarious* region (medium grey area in Figure 6), where the system is still alive but will tend to die if the parametric condition [F] is kept constant, but could still recover if [F] is appropriately modulated. Underneath the precarious region a terminal-irreversible region can also be distinguished (dark grey area in Figure 6). If [A]falls in this region the system will be "alive" for some time, but will irreversibly die (given a certain limit of [F] increase, defined e. g. by diffusion).

We can now introduce the notion of a normative vector *field* defined by the minimal constant increase of [F] that is required at each point of the precarious region in order to move the state of the system into the viable region before the system reaches the terminal-irreversible region. Figure 7 is meant to illustrate this field: if the values of [F] and [A] are low (bottom-left side of the figure) the required increase of [F] is very big since the tendency of [A] will soon push the system to the terminal-irreversible region. If the concentration of F is low but there is a lot of A the required constant increase in [F] is low because the system has sufficient time to reach the viability boundary before the tendency to die becomes irreversible. Since [F] can be modulated by behaviour (provided that the environment displays a gradient of [F]) a sense of *normative agency* can be precisely defined for every state of the system in the precarious region: the amount of increase of [F] that behaviour should achieve to compensate for its precariousness, that is the required movement in space that increases available [F] in accordance with



Copyright 2011, Matthew Egbert, Xabier Barandiaran Licensed under Creative Commons Attribution 3.0 Unported (http://creativecommons.org/licenses/by/3.0)

Figure 8: The effect of agency, idealised in this figure, but also seen in Figure 5. See main text for further explanation.

the normative field. Note that the system can fail to meet the norm, i.e. to adapt, for a variety of causes (e.g. because there is not enough [F] in the environment, because it cannot move sufficiently fast or does not manage to move up the gradient —like the case of the experiment shown in Figure 5-bottom). And yet the action can be said to be in accordance with the norm if it positively correlates with the normative field.

Agency can thus be clearly defined as the behavioural modulation that positively correlates with the normative field (which shall, given the appropriate environmental conditions, bring the system to its viable region). Figure 8 illustrates this point. To further illustrate this idea we examined an agent with a "perfect" gradient climbing mechanism that always moves directly up-gradient with a constant velocity. (Removing stochasticity from the behaviour in the model makes some of the dynamics easier to visualize.) Figure 9 plots one such "perfect gradient climber" with the same initial values of [A] but different distances from the peak of F gradient. We can see how the agent repeatedly moves from the precarious region back into the viability region, except for very low values of [F] for which the system, despite its behavioural modulation of [F] fails to reach the viability region and perishes – as the behavioural mechanism is insufficient to compensate.

### Conclusions

To conclude, we state that *for autonomous agency* (that is agency in relation to self-generated norms) *to take place the overall global constitutive dynamics of the system* (its selfmaintaining organizational dynamics) *should at least* (that



Copyright 2011, Matthew Egbert, Xabier Barandiaran Licensed under Creative Commons Attribution 3.0 Unported (http://creativecommons.org/licenses/by/3.0)

Figure 9: Trajectories of the system in the viability space as a function of its gradient climbing behaviour. Solid lines indicate trajectories that lead to the living stable equilibrium, dotted trajectories (despite behavioural influence) tend to death.

is minimally) *display an intrinsic topology with a viability boundary* (with the form of a bifurcation) *that defines a precarious region where behaviour can compensate for a death-ward tendency*. Arguably, it is only in relation to the intrinsically determined normative field that behaviour can be properly be identified as adaptive and constitute a clear instance of natural normative agency.

The present model benefits from its low dimensionality in that it is easier to understand, but is also suffers, perhaps, from being over simplified in that there is really only two ways that the system can vary. Real organisms are of course much more complex and would display a multi-dimensional normative field and viability boundaries or surfaces. We are working on a more detailed model of a system similar to that described here in which the metabolism and behavioural mechanisms are more explicitly modelled (using more reactants and reactions). This will allow us to explore a greater variety of perturbations to system "health" as well as ways for the system to be sensitive (and therefore responsive) to its own viability. Another expansion of this work that we are considering is the quantification of a normative vector field and the formalization of the notion of positive correlation with it.

### Acknowledgements

Xabier E. Barandiaran holds a Postdoc with the FECYT foundation (www.fecyt.es), and he is funded by Programa Nacional de Movilidad de Recursos Humanos del MEC-MICINN (www.micinn.es), Plan I-D+I 2008-2011, Spain. XEB also acknowledges funding from "Subvencion General a Grupos de Investigacion del sistema universitario vasco. Grupo Filosofia de la Biologia" from Gobierno Vasco IT 505-10.

#### References

- Ashby, W. R. (1952). *Design for a Brain: the origin of adaptative behaviour*. J. Wiley, London, 2nd edition.
- Aubin, J., Bayen, A., and Saint-Pierre, P. (2011). Viability Theory: New Directions. Springer, 2nd edition. edition.
- Barandiaran, X. and Moreno, A. (2008). Adaptivity: From metabolism to behavior. *Adaptive Behavior*, 16(5):325–344.
- Barandiaran, X. E., Di Paolo, E. A., and Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5):367– 386.
- Burge, T. (2009). Primitive agency and natural norms. *Philosophy* and Phenomenological Research, 79(2):251–278.
- Christensen, W. D. and Bickhard, M. H. (2002). The process dynamics of normative function. *The Monist*, 85(1):329.
- Di Paolo, E. A. (2004). Unbinding biological autonomy: Francisco varela's contributions to artificial life. *Artificial Life*, 10(3):231233.
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4):429–452.
- Egbert, M. D., Barandiaran, X. E., and Di Paolo, E. A. (2010a). Behavioral metabolution: Metabolism based behavior enables new forms of adaptation and evolution. In Artificial Life XII Proceedings of the Twelfth International Conference on the Simulation and Synthesis of Living Systems, pages 213—220. MIT Press.
- Egbert, M. D., Barandiaran, X. E., and Paolo, E. A. D. (2010b). A minimal model of Metabolism-Based chemotaxis. *PLoS Comput Biol*, 6(12):e1001004.
- Egbert, M. D., Barandiaran, X. E., and Paolo, E. A. D. (2011). Behavioral metabolution: The adaptive and evolutionary potential of metabolism-based chemotaxis. Submitted, March 2011.
- Egbert, M. D., Di Paolo, E. A., and Barandiaran, X. E. (2009). Chemo-ethology of an adaptive protocell: Sensorless sensitivity to implicit viability conditions. In Advances in Artificial Life, Proceedings of the 10th European Conference on Artificial Life, ECAL, pages 242—250. Springer.
- Frankfurt, H. G. (1978). The problem of action. American Philosophical Quarterly, 15(2):157–162. ArticleType: primary\_article / Full publication date: Apr., 1978 / Copyright 1978 North American Philosophical Publications.
- Jonas, H. (1966). The Phenomenon of Life. Toward a Philosophy of Biology. Chicago-London.
- Jonas, H. (1968). Biological foundations of individuality. International Philosophical Quarterly, 8(2):231–251.

- Kauffman, S. (2003). Molecular autonomous agents. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 361(1807):1089–1099.
- Kauffman, S. and Clayton, P. (2006). On emergence, agency, and organization. *Biology and Philosophy*, 21(4):501–521.
- Kauffman, S. and Farmer, J. (1986). Autocatalytic sets of proteins. Origins of Life and Evolution of Biospheres, 16(3):446–447.
- McFarland, D. D. (1999). Animal Behaviour: Psychobiology, Ethology and Evolution. Longman, 3 edition.
- Millikan, R. G. (1989). In defense of proper functions. *Philoso-phy of Science*, 56(2):288–302. ArticleType: research-article / Full publication date: Jun., 1989 / Copyright 1989 Philoso-phy of Science Association.
- Piedrafita, G., Montero, F., Morn, F., Crdenas, M. L., and Cornish-Bowden, A. (2010). A simple Self-Maintaining metabolic system: Robustness, autocatalysis, bistability. *PLoS Comput Biol*, 6(8):e1000872.
- Ruiz-Mirazo, K. and Mavelli, F. (2008). On the way towards 'basic autonomous agents': Stochastic simulations of minimal lipidpeptide cells. *Biosystems*, 91(2):374–387.
- Ruiz-Mirazo, K. and Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life*, 10(3):235–259.
- Silberstein, M. and Chemero, A. (2011). Dynamics, agency and intentional action. *Humana Mente*, 15:1–19.
- Skewes, J. and Hooker, C. (2009). Bio-agency and the problem of action. *Biology and Philosophy*, 24(3):283–300.
- Varela, F. J. (1979). Principles of biological autonomy. North Holland New York.
- Weber, A. and Varela, F. J. (2002). Life after kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2):97–125.