On What Makes Certain Dynamical Systems Cognitive: a minimally cognitive organization program

On What Makes Certain Dynamical Systems Cognitive

Xabier Barandiaran[1,*]
Alvaro Moreno[1]
[1]University of the Basque Country,
Department of Logic and Philosophy of Science

[*] PHONE: +34 943015516
FAX: +34 943015470
EMAIL: xabier@barandiaran.net
http://www.ehu.es/ias-research
Department of Logic and Philosophy of Science, FICE,
UPV-EHU (University of the Basque Country)
PO BOX 1249 / 20080 Donostia - San Sebastian / Spain

Abstract

Dynamicism has provided cognitive science with important tools to understand some aspects of "how cognitive agents work" but the issue of "what makes something cognitive" has not been sufficiently addressed yet, and, we argue, the former will never be complete without the later. Behavioristic characterizations of cognitive properties are criticized in favor of an organizational approach focused on the internal dynamic relationships that constitute cognitive systems. A definition of cognition as adaptive-autonomy in the embodied and situated neurodynamic domain is provided: the compensatory regulation of a web of stability dependencies between sensorimotor structures, is created and preserved during a historical/developmental process. We highlight the functional role of emotional embodiment: internal bioregulatory processes coupled to the formation and adaptive regulation of neurodynamic autonomy. Finally, we discuss a "minimally cognitive behavior program" in evolutionary simulation modelling suggesting that much is to be learned from a complementary "minimally cognitive organization program".

Keywords: Minimal cognition, adaptive autonomy, neurodynamic organization, cognition-as-it-could-be, emotional embodiment.

1 Introduction

The dynamical hypothesis van Gelder (1998) proposes an ontological claim: cognitive systems are instances of a dynamical causal organization. In addition there is a methodological claim: cognitive systems are better understood with Dynamical System Theory (DST). But neither the dynamical hypothesis nor DST itself offers any criteria to distinguish cognitive from non-cognitive causal organization. As van Gelder puts it (and the question is similarly addressed or ignored by many other dynamicists): "This paper simply takes an intuitive grasp of the issue for granted. Crudely put, the question here is not what makes something cognitive, but how cognitive agents work" (van Gelder, 1998, p.619). But... can we really understand how cognitive agents work without knowing what makes them cognitive?

There are many dynamic phenomena that happen within the cognitive domain and that we do not understand sufficiently well yet. Research on these phenomena is an important part of cognitive science, but the same dynamic causal structure can also be found in non-cognitive systems (for instance the bifurcation phenomenon that can be seen in a model of a decision making process can equally be found in models of chemical, economic or physical systems). Although blind on itself to address the generic question of "what makes something cognitive" DST is a good starting point for this task since: a) it allows for modeling of underlying mechanisms (bridging psychological and biological concepts —Lewis, 2005), b) it does not presuppose a distinction between mind and world; i.e. it crosses over brain, body and world so that it appears specially well suited for situated and embodied

approaches to cognition (Clark, 1997; Chiel and Beer, 1997) and c) dynamicism implies no compromise with representational theoretical primitives to model cognitive systems.

Certainly "how cognitive agents work" requires that we better understand and model some of those specific phenomena which are part of cognition (categorization, learning, memory, etc.) and cognitive science has learned important lessons from dynamicist explanations: that cognition implies the effect of coupled neural, body and environmental systems (Chiel and Beer, 1997), that cognition has an irreducible temporal dimension (Port and van Gelder, 1995), that behavioral and neural forms of order can appear spontaneously by means of self-organization (Kelso, 1995; Skarda and Freeman, 1987), that the body's mechanical dynamics are essential for cognitive developmental processes (Thelen and Smith, 1994), that global neural dynamic synchronization is at the core of cognitive processes (Edelman and Tononi, 2000; Varela et al., 2001; Skarda and Freeman, 1987; Bressler and Kelso, 2001) and so on. But a *full explanation* of "how cognitive agents work" cannot be reduced to the dynamic modeling of particular and isolated features of cognitive behavior. Putting a set of partial models together will not produce a genuinely cognitive system. As it happens in complex systems global properties cannot be reduced to the addition of its component parts. Some fundamental issues regarding meaning, purpose, identity, development, etc. require that we understand "what makes something cognitive" since they rely on integrated properties that distinguish cognitive agents from other dynamical systems.

Traditional cognitivist approaches have solved the problem of "what makes something cognitive" by claiming that the mark of the cognitive is its representational nature. But representationalism has been put into question by many dynamicists (Brooks, 1991; Beer, 2003; Keijzer, 2001; Varela et al., 1991; Freeman and Skarda, 1990). Some dynamicists consider that, once a full mechanistic-dynamicist explanation is available, representationalism appears as a noisy, unnecessary and even mistaken way to explore "how cognitive agents work". Whatever the result of the representationalist-antirepresentationalist debate turns out to be, it looks worth trying to define "what makes something cognitive" (what the specificity of cognition is) from a dynamicist point of view; without an *a priori* appeal to representations as the defining feature of cognition. The question, then, could be addressed as follows: How do we draw the boundaries between cognitive and non-cognitive dynamical systems? If we are not to believe in rigid boundaries still... what specifies the gradient towards the cognitive? How can we specify cognition as a natural phenomenon which is distinct from those that surround it, underlay it and precede it?

A proviso is worth taking at this point. By dynamical system we do not mean fundamental physics but differential equation modelling over observable systems. The possibility of a full understanding of living and cognitive organization with DST is a controversial issue for some authors like Rosen (1991), Pattee (1977) or Kampis (1991). Certainly some aspects of the proposal presented here (for instance the interpretation of the model and its relation to the object that is being

modelled) are not fully reducible to DST itself. Nevertheless a full discussion of the limitations of DST to characterize cognitive organization exceeds the scope of this paper. Our claims and use of DST terminology, thus, should not be interpreted as a defense of DST as being sufficient in its own to explain cognition. The goal of this paper is to shift the focus of attention from behavior to organization within the dynamicist paradigm and to *squeeze the DST framework for a maximal approximation to minimally cognitive organization.*

2 From behavior to organization. In search of intrinsic teleology

2.1 A critique to behavioral cybernetics

One of the most influential attempts to characterize some fundamental properties of cognition (such as purpose and teleology) in purely dynamicist and behavioral terms can be found within the early cyberneticians (Rosenblueth et al., 1943). In this seminal paper, Rosenblueth, Wiener and Bigelow defined behavior as "(...) any change of an entity with respect to its surroundings". Purposeful behavior is taken to be a subdivision of active behavior, where "the term purposeful is meant to denote that the act or behavior may be interpreted as directed to the attainment of a goal —i.e., to a final condition in which the behaving object reaches a definite correlation in time or in space with respect to another object or event" (p. 18). Teleology, in turn, is purposeful active behavior controlled by negative feedback.

Finally, teleological behavior can be predictive in an increasing number of orders. At this point, the cyberneticians could argue, we already have a dynamicist criteria to establish a gradient towards cognitive processes. Within the set of all possible dynamical systems, cognitive systems are those whose behavior is active, purposeful, teleological and predictive, all of which is claimed to be specifiable on purely behavioristic terms.

In *The Phenomenon of Life* (2001) the phenomenologist Hans Jonas strongly criticized the cyberneticist, purely behavioristic, approach. He argued by *reductio ad absurdum*: If something is purposeful, as far as it can be interpreted as directed to the attainment of a final condition of "definite correlation (...) with respect to another object", then, Jonas argues, all objects have the purpose of "running down" or decreasing (minimizing) entropy, since the only objectively sustainable final state of a machine is its disintegration. This leads to the unavoidable and uncomfortable conclusion that disintegration is the main and ultimate purpose of all machines and organisms.

The root of this failure to account for purpose is the fact that behavior (and thus purpose, teleology and, we could follow, cognition) is defined exclusively in terms of external relationships: "change of an entity with respect to its surroundings". According to Jonas, a behavioral account of purpose cannot capture the intrinsic properties that define it. Unlike the early cybernetic machines, genuine purposefull entities (organisms) have intrinsic, rather than extrinsic, ends (which shows up on the awareness of failure that "natural" purpose is embedded with):

This *end*, motivating the action from the start, is intrinsic in all the part-motions, providing the reference by which they are in themselves failures and make the whole undertaking a failure. Presumably the patient finds his inability to perform distressing. But the machine, for all we know, may just as well be said, instead of being distressed, to abandon itself with relish to its wild oscillations, and instead of suffering the frustration of failure, to enjoy the unchecked fulfillment of its impulses. "Just as well" amounts of course to "neither". (p.112)

A possible counterargument defending that reference of failure is provided by a design principle, Jonas argues, will not solve the problem since "[w]hat a mechanism is designed for, by its maker, is of course entirely irrelevant, because extraneous, to the description of its working". The argument works equally well against those defending teleology on the grounds of natural or evolutionary design (Millikan, 1989). Whatever could have been selected by evolution (presumably genes) is not in itself an explanation of intentionality (nor the conditions in which it was originally selected) if not accompanied by an explanation of how the selected variation is inserted on the causal organization of the system that gives rise to purposefull behavior. For instance we could artificially select the color of a car, but that would not justify the claim that the car *holds the purpose* to please us; that its intention is to satisfy the owner by having a certain color.

Another problem that a behavioral approach should solve, in order to provie a necessary and sufficient condition to define cognitive systems, is the problem of

mimicking. If the principle we use to classify a system as cognitive is purely behavioristic, one is left out with no criteria to distinguish between genuinely cognitive behavior and imitation. John Searle's Chinese Room argument (1980) is a well known objection to the behavioristic approach to intelligent system discrimination —specifically for the Turing Test (Turing, 1988).

2.2 Autopoiesis or the *conatus* of living organization

A very similar and often overlapped question to that of the origin of cognitive intentionality and purpose is found at the level of the origin of life and biological function. Not surprisingly all the systems that we take for genuinely cognitive around us are living beings. In fact, and although he didn't provide any explicit and clear account of the mechanisms leading to purposefull behavior, Jonas' intuition was that life has intrinsic purpose by means of its self-sustaining nature; which is the result of its internal far-from-equilibrium (FFE hereafter) metabolic organization. Jonas was not the first to notice that living system's tendency for self-preservation is at the core of intentionality. 400 years earlier philosopher Baruch Spinoza (recently rescued by neurobiologist Antonio Damasio to address some of these very same issues) had a similar intuition regarding the origin of intentionality, what he termed the *conatus*: the driving force for perseverance that living beings are embedded with.

It was not until more recently that all these intuitions started to be explicitly addressed in terms more amenable to science, systemic approaches and, in partic-

ular, DST. One of these attempts was Maturana and Varela's theory of *autopoiesis* (Maturana and Varela, 1980) which had (and still has) a big influence on closely related issues such as adaptive behavior (Beer, 1997), cognitive science (Varela et al., 1991) and theories of the origins and definition of life (Ruiz-Mirazo and Moreno, 2004). Unlike the behavioristic approach, which "omits the specific structure and the intrinsic organization of the object"(Rosenblueth et al., 1943, p.18), the autopoietic theory focuses on the networked internal relationships of living beings that makes them self-sustaining and self-constructing systems while differentiating themselves from the environment. According to Varela "[a]utonomous systems are mechanistic (dynamic) systems defined as a unity by their organization. *We shall say that autonomous systems are organizationally closed. That is, their organization 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).

Similar accounts of the basic (or minimal) living organization are also found among a number of authors —Rosen (1991), Kauffman (2001), Bickhard (2000), Ruiz-Mirazo and Moreno (2004), Collier (1999), among others. They all share in common the emphasis on a specific kind of closure or circular causality that characterizes living beings: that of a network of processes whose stability is both the effect of the interaction between its component parts and the interactions that they

altogether maintain with their environment. Thus, the activity that the organization generates is both cause and effect of the very organization; i.e. the organization can only be maintained through the constraints that it generates. DST alone might not be able to accommodate such strong requirements (as, for instance, Rosen's *closure to efficient causation* and other formal accounts of circular causality in complex systems). But the simulation of differential equations with strongly interacting variables is a promising framework to *approximate* such systems. In Varela's words, dynamical systems and computer simulations are "*one* way in which properties of systems, autonomous *or* allonomous, can be expressed. Differentiable dynamics represent, in practice, the most workable framework in which these two points of view can actually coexist and be seen as complementary in an effective way" (Varela, 1979, p.264).

We shall call *Basic Autonomy* (Ruiz-Mirazo and Moreno, 2000, 2004) the organization by which FEE and thermodynamically open systems adaptively generate internal and interactive constraints to modulate the flow of matter and energy required for their self-maintenance. Now, autopoiesis, or basic autonomy (BA hereafter) refers to the material self-construction upon which other levels of autonomy can appear[1].

2.3 Normativity and explicit teleology in autonomous systems

From a different point of view, more directly related to the etymological origin of the term (*auto* = self and *nomos* = law, norm), autonomous system are those

capable to generate their own norms. Autonomous systems allow to speak in terms of a strong sense of norm or normativity, where the nature of the norm (what is good or bad for the system) is not externally interpreted or derived from an adaptive history but defined intrinsically by the very organization of the system.

Normativity refers to the fact that a set of processes that constitute the system *must* happen as they do in order for the very system to exist[2]. A basic example of normative (proper, necessary) functionality is given by the active transportation through the membrane of cells. This process becomes normative because the level of chemical concentrations that the membrane's active transport keeps within the cell is necessary for some metabolic reactions to maintain the appropriate rate necessary to sustains the network of reactions..., that in turn produces the membrane, etc., etc. This kind of circularity is characteristic of autonomous systems: a set of networked component processes that depend recursively on each other, so that the system, as a whole, is cause and effect of itself. Due to this circularity in Basic Autonomous Systems (BAS), identity and normative functionality is not observer dependent but intrinsically causal, since the whole network (the very system) will not exist in the absence or malfunctioning of the component processes (given its FFE nature and the circular dependency between processes). In other words, in BASs what-the-system-does (the way it functions) and what-the-system-is (its structure) are highly intertwined, they merge together on its organization.[3]

The holistic, integrated and self-maintaining organization of BASs has some important consequences on the way they are described. For instance, the use of

teleological terms to characterize their functioning can be naturalized; unlike its use to describe some artifacts that perform a goal-seeking behavior such as the thermostat or target-seeking missiles. These are artifacts that have been designed to correct their behavior (usually by a negative feed-back mechanism) according to an externally defined goal state. Expressions such as "the purpose of the thermostat is to maintain the room temperature at 23°C" are used as metaphorical shortcuts to describe the behavior of such systems. But what the goal state is remains completely extraneous to the mechanism that achieves it, the system is independent of the goal state or set of parameters it controls (which are externally imposed). BASs are different. Their existence depends on the FFE stability they produce. The stability point (or set of points) through which the system can exist are the goal states of the system. This goal state is not just a goal state because the system compensates deviations from it, but because the goal state is the condition of possibility of the very system. In other words, in BASs the goal state of the system and the organization that instantiates it are one and the same thing. BASs have a genuine purpose of self-maintenance, an intrinsic teleology is implicit in the system.

As Di Paolo has recently argued (2005), autopoiesis is not enough to account for *explicit* teleology or sense-making: adaptivity is also required. In fact, the existence of a FFE and self-sustaining network of processes does not necessarily (at least at the conceptual level) imply that the system also behaves so as to actively compensate internal and external perturbations. In other words, following Di Paolo, minimal BAS have intrinsic (rather than extrinsic) but implicit (rather

than explicit) teleology. Adaptive systems have the capacity of distinguishing and transforming the trajectories of their essential variables (those defining the viability constraints of the system) so that tendencies toward outside the boundaries of viability are compensated. The good or bad functioning *of* the system is objective (since the very system would cease to exist otherwise), and it is objective *for* the system because it is detected and compensated *by* the system, in an effective, functionally integrated way. Thus, adaptive autonomous systems are an instance of explicit teleology since in addition to *having* an intrinsic goal (due to their basic autonomous organization) they also *act* according to this goal generating global constraints, over their minimal basic organization, so that a metaregulatory process emerges. It follows that BASs with adaptive capacities constitute agents, whose interactive processes with their environment become inevitably meaningful. This meaning is ultimately nothing else than the functional effect that a given interaction with the environment will have on its self-maintaining organization, as it is evaluated and sustained by the metaregulatory subsystem that controls behavior. Thus, the environment becomes, as von Uexküll (1982) has pointed out, a "umwelt", namely, an enacted world made of the structural couplings (interactive processes) between the agent and its environment, as they are integrated on the ongoing processes of self-preservation and construction.

2.4 But the mind has a life of its own

Certain authors have argued that adaptive material self-construction is not only necessary but *sufficient* for cognition (Maturana and Varela, 1980; Stewart, 1996), that this kind of self-constructing closure is the organizational principle we are looking for to characterize cognitive systems. But as well as somewhat awkward for cognitive robotics (since it would imply that no genuine cognitive behavior can be expected from non-self-constructing artifacts) this thesis is also conceptually uncomfortable: it is somehow intuitive that cognition relates to sensorimotor interactions rather than to material self-constructing processes. Failure to satisfy a cognitive purpose does not *necessarily* imply failure of material self-maintenance. The opposite of cognition and cognitive success is not death or biological illness but (as it would become more evident later) some kind of coma, "madness" or loss of behavioral coherence. The slogan could be that "the mind has a life of its own", i.e. that a new level of normativity and adaptive autonomy must be in place in order to explain genuine cognitive systems. The main intuition here is that what is characteristic of cognitive systems is not the satisfaction of biologically defined constraints (by evolutionary of metabolic adaptive needs) but the creation of a new domain with its own internal coherency.

From a phenomenological point of view, a cognitive world is not (just) a world of biological (or basic autonomous) constraint satisfaction and metabolically derived intentionality but a meaningfull world in itself, i.e. in relation to sensorimotor behavior and to its internalized history of interactions. Since living behavior can-

not violate basic autonomous constraints it is difficult to find examples of natural behavior that explicitly illustrate what we have just mentioned (specially without referring to anthropomorphic examples). Nevertheless, there are clear examples where behavior responds to non-metabolic and non-adaptive teleology. Games are an instance of such cases but cognition is, we defend, another kind of non-purely biological teleology. Let us take the example of Coggy, a dog that is barking to its owner (Phil) when coming back from work. Coggy usually responds friendly to strangers and specially to Phil. The difference today is that Phil smells differently. At this point it seems difficult (if not impossible) to defend that Coggy's barking to Phil is an evolutionary adaptation, or responds to metabolic values or even that it be a learned behavior. On the contrary Coggy's barking seems to respond to some kind of cognitive dissonance regarding a breakdown of the habit of identifying Phil in a certain way. A principle of preservation of experiential coherency (non reducible to evolutionary or metabolic values) seems to be in place.

So, before we throw the baby out with the bath water we can retain the notion of adaptive autonomy (and the naturalized foundations of teleology and normativity that it permits) in order to explore how it could be applied to the sensorimotor domain (rather than to the material self-construction of the most fundamental biological domain). We could then sketch a minimal form of behavioral adaptive autonomy that could provide us with an organizational definition of "what makes something cognitive".

3 The autonomous organization of behavior

3.1 The appearance of the nervous system

Those organisms that are distinctively cognitive appear endowed with a specific subsystem that is *hierarchically decoupled* from metabolic-constructive processes: the nervous system (NS hereafter). There is nothing particularly revealing on neurons, electrochemical interactions, and other features of the NS of animals in our planet other than the fact that nervous interactions generate a set of dynamic relationships not governed by the general metabolic organization: it constitutes a new dynamic domain.

By the term hierarchical decoupling of the NS from metabolism we mean, thus, that metabolism generates and sustains a new dynamical system (the NS) minimizing its local interference with it. The term *hierarchical* refers to the fact that metabolism produces and maintains the architecture of the NS. On the other hand, the term *decoupling* means both a) that neurons act as minimizing the interference of their local metabolic processes with their ion-channeling capacities and b) that the metabolic-constructive organization of the organism *underdetermines the activity of the NS, which depends on its internal dynamics and its embodied sensorimotor coupling with the environment*[4]. Operationally speaking if we are to predict the dynamic behavior of the NS, hierarchical decoupling means that neither local states of cell metabolism nor the state of metabolic organs (kidney, viscera, heart, etc.) alone are going to be very useful; on the contrary, the electrochemical

states of other neurons and their embodied sensorimotor coupling with the environment might provide a much better model for prediction. Insofar as, in organisms endowed with NS, behavior is controlled by neural dynamics, the level of modelling of behavioral interactions is not that of metabolic cycles and material self-maintenance and construction, but that of a different dynamic domain constituted by neural activity and its sensorimotor embodied coupling with the environment.

While it is fairly evident that the components of a minimal metabolic organization are biochemical reactions and, in multicellular organism cells, tissues and organs, making neurodynamic autonomy explicit requires (once the domain in which it appears has been delimited) that we specify what the components of such organization are and how their stability depends on their mutual interactions and on the behavior they sustain.

3.2 Dynamic structures as organizational components of behavior

In principle, the appearance of the NS conveys an open expansion of interactive possibilities, which can become indefinitely complex (sustained by the potentially unlimited richness of neural dynamics). There is increasing evidence to sustain that cognitive brain activity is the result of large scale integration of the activity of distributed neural ensembles —see Tononi et al. (1998), Varela et al. (2001), Friston (2000), Bressler and Kelso (2001). Some classical views on brain functioning, take the brain to operate in functionally specific modules so that behavior is the result of intra-modular processing and point to point information transfer between modules.

On the contrary, large scale integration models defend that behavior is the result of global patterns of oscillations that emerge from reciprocal dynamic interactions between highly distributed structures of the brain. Thus, if we are somehow to specify the internal dynamics of cognitive behavior we can't take anatomical components as organizational primitives. That is why we shall introduce the concept of *dynamic structure*. By dynamic structure we mean the subset of internal variables and their relationships involved in a certain sensorimotor coupling. A dynamic structure emerges when (for a given time window) we can systematically reduce the dimensionality of the internal operational organization of the NS to predict the behavior of the system.

A dynamic structure is dynamic in two senses. First, it is a structure of change, a region of dynamic trajectories in the phase space (e.g. a local chaotic attractor). Thus by structure we do not refer to any physical or anatomical component or ensemble but to a cohesive dynamic form that emerges in a particular interactive coupling with the environment. Second, by dynamic we mean that the structure can also change over time: the "shape" of the phase space region is subject to different stability conditions that might alter it[5]. We can thus, for a given time window, decompose the whole activity of the NS into a set of dynamic structures that predicts its behavior in different couplings with the environment.

Once dynamic structures have been defined as basic components of neurodynamic organization we are ready to see how their stability might depend on behavior and on the nested relationships that are established between them.

3.3 Dynamic structures and behavior: interactively dependent stability

In the late 80s and early 90s Rodney Brooks proposed the subsumption architecture as a design principle for autonomous/situated robotics (Brooks, 1991). Brooks criticizes GOFAI sense-plan-action architecture (receiving an input, processing it and generating a motor command) and proposes, instead, to build robots on the basis of behavioral layers. A behavioral layer acts as a circuit that, closed through environmental interaction, generates a stable behavior (such as obstacle avoidance, phototaxis, random search, etc.). The overall behavior of the robot is the result of internal and environmental interactions between different behavioral layers: under particular sensorimotor conditions different layers are activated and deactivated, combined or inhibited so that a repertoire of behaviors is generated in continuous embodied interaction with the environment. Let us now imagine a subsumption architecture where the structure of the behavioral layers depends both a) on the stability of the generated behavior and b) on the interaction and stability of other behavioral layers. The crucial feature of dynamic structures is that the stability of the component structures of the control architecture and the structure of behavior depend recursively on each other.

The main idea is that dynamic structures (unlike conservative behavioral layers) are subject to stability conditions that depend on the activity of the whole or big part of the activity of the NS *and* on the particular way in which the dynamic structure is coupled to certain sensorimotor correlations. The structure of behavior

depends on the way in which internal dynamic structures are coupled (through the body) with the environment; but behavior sustains dynamic structures too: i.e. the stability of a dynamic structure depends on the particular sensorimotor correlations that the coupling it sustains generates. If this coupling is lost structural transitions happens, or the system undergoes structural change: it enters a region of structural instabilities until a new structure is stabilized/created or it is disintegrated.

The stability condition might also depend on future sensorimotor correlations, as is the case of operant conditioning and reinforcement learning. We could define expectancies as dynamic counterfactuals (conditionals): if a certain interactive condition is not met during or after a certain behavioral coupling the dynamic structure involved on that coupling dissolves. This way dynamic structures might depend on the satisfaction of expectations, not just immediate sensorimotor correlations.

3.4 The emergence of neurodynamic autonomy

So far the relationship between the stability of a dynamic structure and the behavior it generates provides a first form of self-maintenance (distinct from the basic metabolic-constructive one): dynamically speaking the stability dependencies created between dynamic structures and behavior act so that dynamic structures maintain themselves (as far as they are capable to generate the behavior that recursively satisfies their stability conditions).

At this point we are forced to explore possible hypothesis. There exists no empirical research on the global organization of brain dynamics at the scale required

to characterize cognition as a particular kind of, holistic, dynamic organization.

Taking dynamic structures to be the organizational primitives of neurally controlled behavior, we can now imagine a nested web of dynamic structures as constituting the dynamic organization of neurally guided behavior. Given the integrated nature of the NS and the non-localized nature of some dynamic structures it is highly probable that the stability of certain dynamic structures be dependent on the activity of the whole system. It is likely that the most primitive NSs worked essentially driven by innate constraints. But, once a certain level of neural complexity is attained, a web of stability dependencies can be created between dynamic structures too (in addition to behavioral stability dependencies) so that a global structural interdependence emerges (similar to those that are found between different reactions in auto-catalytic networks). In other words, the stability of a given dynamic structure will depend on the existence of a network of other dynamic structures which, in turn, depend on the expectancies that are recurrently satisfied through behavior. Inversely, the instability of a particular dynamic structure could generate a cascade of instabilities and transformations propagating along the web of dynamic structures. This nested web can be considered to be the mechanistic counterpart of a phenomenological world of believes, habits, assumptions, etc. with their internal coherency and interdependencies.

Now, given the far-from-equilibrium condition of this web, we could hypothesize that an autonomous level of normativity emerges when neural dynamics have a self-maintaining organization; i.e. when the web is homeostatic and behavior is

directed towards the self-maintenance of the global stability conditions of the web (and not only of a unique dynamic structure). This would lead us to a new level of autonomy and self-maintenance embedded on, but distinct from, that of BA or autopoiesis.

In this sense, neurally guided behavior can be characterized as that of an autonomous system in the hierarchically decoupled, embodied and situated sensorimotor domain. We can, more explicitly, define an autonomous system in this domain through two necessary and sufficient principles:

Identity: A cohesive web of dynamic structures can be distinguished, a cluster of highly integrated correlations appears where *the correlations are the recursive effect of the activity of the set of variables*. This cluster we call *identity* because it defines the system as a self-sustaining unity.

Agency: The system (defined by its identity) establishes a *functional control* over its relationship with the rest of the variables not belonging to the system (the environment).

Although conceptually distinguishable, the situatedness of neurodynamic processes makes it difficult to establish a clearcut distinction between agential and identity processes. In fact we have emphasized that the neurodynamic identity is maintained through continuous interaction with the environment. Can we, thus, distinguish the agent from its environment? We can. The key notion of the agency principle is that of functional control. *Functional* means that the result of the in-

teraction is the maintenance of the identity (see the definition of functionality in autonomous systems —section 2.3), in other words, what remains invariant across the set of interactions that the system establishes with its environment is the system's identity, to which the interactions contribute. *Control* implies that the structure of behavior is asymmetrically determined by the identity of the system. Since there is no possible linear decomposition of causality on the agent-environment interactions, agency requires a complexity asymmetry on the production of behavior which appears laden to the side of the agent[6]. For instance a system approaching a light source is a agent only if the light source is less complex on defining the distance between the system and the light than the system itself. Thus, autonomy implies a *causal asymmetry* on the continuous system-environment dynamics so that a cluster of variables (the agent) can be distinguished as the 'origin' of behavior.

Autonomy is, thus, achieved through situated and embodied self-organizing processes. The main difference between self-organization and autonomy is 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. It is nonetheless impossible that all the constraints be self-generated. In fact, embodied and innate constraints (those present as initial conditions and parameters of the sensorimotor control system) are essential, but do not completely specify the dynamics of the NS. Starting with these innate constraints and through its sensori-

motor coupling with the environment, the autonomy of the NS is an open historical process of self-determination. In fact the autonomy of the NS is not an absolute term but a gradual *becoming* (unlike Maturana and Varela's notion of operational closure). We could say that the organism (through the hierarchical decoupling of the NS) generates a dynamical domain of a much higher variability than what its metabolic regulatory capacity and genetically specified constraints can control, so that an increasing number of historical/developmental constraints are interactively created/stabilized. This new autonomous level is, as we shall see, the organizational ground for the cognitive domain.

4 The role of emotional embodiment on the formation and regulation of increasing neurodynamic complexity

The potential increase of neural complexity (understood in terms of the number of dynamic structures and their interdependecies required to achieve an autonomous organization) is in practice bounded by several factors. One type of factors is related with the bodyplan of the organism, since the increase in the number of neurons in a concentrated area requires certain organizational conditions, like a fine-tuned circulatory system in order to provide their necessary metabolic maintenance (Moreno and Lasa, 2003). The other type of factors is related with the necessity of some sort of mechanism for the formation and functional regulation of the increasing complexity of the dynamics generated by bigger neural concentrations. At this

point a number of important questions can be asked: a) Since the web of dynamic structures is not formed at the beginning of the developmental process... what kind of processes participate on its formation? b) Since the stability of the emergent dynamic structures is dissipative (namely, there is a FFE network of structures, which are recursively and interactively maintained by their own activity)... how is it maintained in the face of internal and external perturbations that could destroy the stability dependencies? and; c) Where does the selective drive to retain new dynamic structures come from?

At this point we need to introduce a different function of the NS, namely the regulation of metabolic organization. In fact, as the size of the organism increases, and the organism's bodyplan permits the formation of large encephalized brains, the regulation of metabolism itself becomes increasingly complex; to the point where the NS becomes more and more important for its regulation. The neuroendocrine system, the autonomic NS, the limbic system and a set of related neural structures form what neurobiologist Gerald Edelman (1989) calls the Nervous System of the Interior (INS hereafter), that can be distinguished from the Sensory-Motor Nervous System (SMNS). According to several neuroscientists (Damasio, 1994; Lewis, 2005) the complex interplay between the INS and the SMNS gives rise to the emotional world.

As for the fist question it is precisely this INS that will originally retain dynamic structures. Thus, originally (ontogenetically speaking) dynamic structures are stabilized through the modulatory effect of the INS, according to the effect (on

the body) of the behavior they sustain. At the same time correlations, or associations, between dynamic structures and body reactions starts to form, so that the modulatory effect can take place before the emergent dynamic coupling produces a certain behavior: destabilizing the dynamic structure before the body is damaged or accelerating its formation when correlated with positive effects on body regulation. Nonetheless this "evaluation" of the dynamic structure that emerges on the SMNS still depends on the direct correlation between dynamic structures and the interactive effects they produced on the body in the past.

Thus, at an early stage of neurodynamic development, body regulation (biological constraints) and sensorimotor regulation are directly correlated. But, as the web of dynamic structures starts to form, the original modulatory correlation between INS's internal regulation and SMNS's regulation becomes blurred on the web of dynamic structures. If we visualize the correlations between single SMNS dynamic structures and INS modulatory capacity as vertical links, and correlations between SMNS dynamic structures as horizontal links, the developmental process involves the formation of an increasingly higher number of horizontal links; to the extent that dynamic structures become more dependent between each other than with the direct modulatory effect they "originally" triggered in relation to body regulation. In other words; 'horizontals links' appear nested and sustained as a whole by the 'vertical links'. Thus early stability dependencies with biological constraints (those that gave rise to the formation of dynamic structures in the first place) become integrated on the internal relationships between SMNS dynamic

structures. At this point we can properly speak of *neurodynamic autonomy*. Now, the maintenance of increasing levels of neural complexity is necessarily linked to the their functional regulation, otherwise the achieved level of autonomy will be lost in the face of internal and interactive perturbations.

What research on the interplay between the INS and cognitive evaluation has shown (Damasio, 1994; Bechara, 2004) is that emotional dynamics are not only activated when the basic viability conditions are in danger but also when cognitive evaluation takes place. We hypothesize that the modulatory capacity of emotional dynamics in SMNS organization is recruited to adaptively modify SMNS in order to preserve global coherency or stability when a cognitive dissonance takes place (i.e. when internal or interactive perturbations of neurodynamic organization threatens its viability as a neurodynamic identity, not just as a metabolic body). Thus the INS acts as *adaptively* regulating neural activity for the self-maintenance of neural organization itself, which appears highly intertwined with bioregulatory organization. In fact one of the major roles that some neuroscientists attribute to emotional and appraisal dynamics is precisely the modulation of SMNS dynamics for global reinforcement and coherency: attention (amplification of sensory perturbations in order to satisfy a certain goal stability condition), stress (generation of global instabilities for dynamic rearrangement, like in spin glass models in Boltzmann machines), satisfaction (reinforcement of the dynamic structure whose produced behavior satisfies a global stability condition), evaluation, and, in general, shaping the neural dynamics in the direction of the global FFE organization

that has been created through its history of interactions.

Emotional embodiment (involving different anatomical components, coupled to body regulation and functioning at different timescales) implies the formation of a neurodynamic processes that operate "off-line" the ongoing sensorimotor activity, so that a hierarchical control/regulation becomes feasible. Emotional embodiment is an often neglected aspect of organismic embodiment, which is different to the, generally, more emphasized physical, sensorimotor or other types of embodiment (Ziemke, 2003). The interaction between the INS and the SMNS becomes, thus, of fundamental importance for neural and behavioral organization to the extent that the adaptive regulatory capacity of the INS over the SMNS will be recruited by the later to regulate its own autonomy. What we at this point is *adaptive autonomy* at the embodied and situated neurodynamic domain; which allow us to speak in terms of genuine cognitive teleology and normativity.

5   What makes something cognitive

We can now provide an organizational account of cognition and cognitive process. The adaptive web of dynamic dependencies that are created within and between the NS and its coupling with the metabolic-body and with the environment is what we shall call *cognitive organization*, (unlike generic adaptive behavior which "only" needs to satisfy externally defined constraints. By externally we mean that the viability constraints that adaptive behavior must satisfy are external to the behavior and the sensorimotor organization that generates it. For instance metabolic needs

are not defined by behavior itself but by the metabolic organization that supports it. Thus, *it is not until the adaptive preservation of the internal organization of neural dynamics becomes the mayor source of neurodynamic regulation that cognition appears*. Our main hypothesis is, thus, that the specificity of cognitive dynamics (what makes it different to other dynamical systems) is given by a particular kind of dynamic organization within the NS and between the NS and the internal and external environment: the adaptive preservation of a web of dynamic sensorimotor structures sustained by continuous interactions with the environment and the body (specially through the interaction between SMNS and INS). And when we use the term NS we are not restricting our definition to cognition-as-we-know-it in animals of our planet but, on the contrary, we are referring to any hierarchically decoupled dynamic domain embedded and embodied in a basic autonomous system, so that our definition permits a universal characterization of cognition (cognition-as-it-could-be).

6   A synthetic approach to minimal cognition

The hypothesis presented through this paper still requires a further elaboration in order to be introduced on properly scientific practice. But our current understanding of the global neurodynamic organization of, even the simplest, natural cognitive systems is still far from complete (not to speak of embodied and situated dynamics or developmental scales). In addition, most of current neurobiological details under investigation are not amenable to an abstract model of a universalist under-

standing of cognition (cognition-as-it-could-be). How could, then, it be possible to advance on the theoretical framework sketched here?

One of the most fruitful methodologies that integrates brain, body and environments in an attempt to rethink the foundations of cognitive science in dynamic terms is Randall Beer's "minimally cognitive behavior program" (Beer, 1996, 2003) and a host of alike evolutionary simulation models of embodied robotics, also called evolutionary robotics (Husbands et al., 1997; Nolfi and Floreano, 2000). This approach stands for one of the best available tools for a deeper understanding of minimal cognitive organization; since artificial evolution can explore the high dimensional parameter space continuous time recurrent neural networks — i.e. potentially universal smooth dynamical systems (Funahashi and Nakamura, 1993). The potential of evolutionary robotics to explore the foundations of cognitive science does not only lie on its capacity to explore dynamic organization, but also on the way in which brain, body and world can be analyzed together in an integrated way (unlike most of current neurobiological experiments that are forced to isolate neural and environmental processes to accommodate them to the available experimental techniques). In addition, far from the intricate complexities and biological constraints that natural systems are subject to, this kind of synthetic approach permits to systematically explore and manipulate minimal models of cognitively relevant behavior, generating what Randall Beer has called a *frictionless brain* experimental framework: i.e. the necessary simplification and minimalism that manipulation of experimental setups requires in order to initiate a systematic

exploration into a new scientific domain. In Beer's words:

> these model agents must be simple enough to be computationally and
> analytically tractable, so that we have some hope of evolving and ana-
> lyzing them using techniques that are at most an incremental step be-
> yond what is currently known to be feasible. The term "minimally
> cognitive behavior" is meant to connote the simplest behavior that
> raises cognitively interesting issues.

The limitation of this approach is that artificial evolutionary design of these agents has been directed towards "the simplest behavior that raises cognitively interesting issues", without considering in depth the kind of organization that makes agents cognitive. *Minimally "cognitive" behavior* is conceptually (not always technically) straight forward in evolutionary robotics. It "just" requires to set up the fitness function to match the behavior that "raises cognitively interesting issues". We have previously argued that a behavioral achievement (or mimicking) is not enough to account for cognition. Recovering our example in section 2.4 we can imagine an evolutionary robotic simulation model that imitates Coggy's behavior (barking to its owner 'because' he smells differently). In the evolutionary robotic simulation model, a target object might be presented to the robot associated with a given smell (any additional sensory modality will do for this purpose). Whenever the target object is present with a different 'smell' the robot is required to react differently. In addition the robot's environment should also include other objects with varying associated 'smells'. This task could be achieved with current artificial

evolution and continuous time recurrent neural networks and, at an abstract behavioral level, the successfully evolved robotic behavior would be equivalent to that of Coggy's. But... are the robot's "minimally 'cognitive' behavior" and Coggy's reaction cognitively equivalent? And, more importantly, can the distinction be made on purely behavioristic terms? The imitation of cognitive behavior is not equivalent to cognitive behavior. We can imitate somebody else's movements (or even speech) without having any intention or understanding of what we are doing. Reference to internal dynamic and organizational differences is required in order to distinguish cognitive from non-cognitive behavior.

Paradoxically, in spite of (or, properly speaking, thanks to) the emphasis on behavioral selection (abstracted from internal cognitivist presuppositions on control architectures) some of the most important contributions of the minimally cognitive behavior program are closely related to organizational issues. The use of artificial evolutionary techniques applied to shape the space of dynamic control architectures has permitted to overcome robotic design principles based on cognitivist assumptions, thus leading to theoretical discussions over the necessary organizational requirements to achieve certain cognitive behaviors. Other models in evolutionary robotics have more directly addressed the effect of organizational constraints on behavior [7]. A promising research avenue has recently being opened by Ezequiel Di Paolo's *organismically inspired robotic approach* which addresses the interplay between internal homeostasis, metabolic values and behavioral stability (Di Paolo, 2003).

In fact, time is ready, we believe, to start modelling adaptive-autonomous robotic architectures. The holistic, emergent and situated nature of neurodynamic autonomy makes its hand design unsuitable. Nevertheless a set of design constraints could be envisioned for evolutionary simulation models. For instance, a fitness function, for artificial selection, based on the stability of a FFE variable that can only be interactively maintained would satisfy behavioral, internal and implicit tness function criteria; which, according to Floreano and Urzelai (2000), shall produce highly self-organized control systems. This could, ultimately, lead to an autonomous organization of behavior. A principle that links internal stability to behavioral stability might also be required, Di Paolo's (2000) homeostatically adaptive robot is a good example. On the other hand, the modelling of an internal environment on the robot has just started to be explored (Parisi, 2004). In particular modelling of internal bioregulatory processes coupled to sensorimotor control is a promising research avenue. Finally, in addition to synthetic methodologies development of analytic tools would be required. A formalized version of minimally cognitive organization could be achieved using recent proposals for descriptions of dynamical hierarchies (McGregor and Fernando, 2005). In addition, complexity measures to understand functional integration in neural processes (Tononi et al., 1998) are producing interesting results, as exemplified by its application to embodied simulation models (Seth and Edelman, 2004). Seth and Edelman made a quantitative comparison between environmental, behavioral and neural complexity and robustness, leading to interesting results on the relationships between them.

This kind of analytic tools should also be usefull for a quantitative study of autonomy and agency as described in this paper.

7   Conclusion

Certainly by studying certain sensorimotor correlations in living beings we can acquire some knowledge about their operating sensorimotor structures (including some aspects of their long term change in learning) and we can simulate agents that satisfy cognitive-like behavioral constraints (associative learning, categorization, etc.). But if we are to create and study genuinely cognitive systems the approach presented here would require that cognitive organization be explicitly modeled. As we have seen adaptive-autonomy is a kind of organization that can account for intrinsic and explicit teleology and normativity. From the autopoietic tradition in robotics the issue of autonomy and autopoiesis has often been neglected as a concept that is only relevant for the more fundamental metabolic level. Researchers have focused, instead, on other autopoietic concepts such as structural coupling, structural determinism or the anti-representational implications of autopoietic theory, proposing a minimally cognitive behavior program with no reference to the agent's internal dynamic organization. We have argued that minimal cognition requires more than cognitive-like behavior, that it requires a particular kind of dynamic organization that adaptively sustains a behavioral repertoire and, most importantly, a capacity to reorganize and generate new dynamic structures in order to preserve its organization. How "what the agent does" (behavior) relates

to "what the agent is" (organization) and *vice-versa* becomes a key issue. At the same time *emotional embodiment* (the coupling between internal regulatory function of the NS with its sensorimotor control function) might become an essential component for the formation and adaptive regulation of neurodynamic autonomy. Thus, the focus of the research on minimal cognition should not only be the reproduction and mimicking of the behavioral repertoire that cognitive psychologist and philosophers take for "cognitively interesting" —for instance "representation hungry problems" Clark and Toribio (1994)— but the exploration of dynamical systems capable of adaptive autonomous self-sustaining interactive organization. This paper was meant to contribute a first step toward a shift from a "minimally cognitive behavior program" to a "minimally cognitive organization program" from "how cognitive agents work" to "what makes them cognitive", since the former, we argued, will never be complete without the later.

Notes

[1] Basic Autonomy is a far more elaborated concept than material self-construction, but for the purpose of this argument we need not specify it fully (for a proper definition of the concept see Ruiz-Mirazo and Moreno (2004)). Along the paper material self-construction, metabolic constructive processes and basic autonomy will be used as synonyms.

[2] We use the term *regulation* to refer to the way in which the system or any of its parts is actually behaving, while *normativity* will only be used to describe how the system *must* or *should* work. *Normative regulation*, in turn, is the system's regulation in relation to the satisfaction of its normative requirements.

[3] The issue of normative functionality in autonomous systems has been extensively treated by Christensen and Bickhard (2002).

[4] In this sense it is important to note that fast movement, in multicellular organisms, is only possible through specialized organs which directly convert metabolic energy into mechanic energy independently of the continuous process of metabolic self-maintenance and morphological processes that the organism undergoes by means of cell replication and growth, unlike other kinds of organism-environment interactions like those found mainly in plants.

[5] The distinction between these two kinds of change is a matter of the time window and the level of description we consider. If we take the lower level differential equation model of the NS-body-environment system then a dynamic structure appears as a higher level description of the model, that remains stable for a certain time window. This higher level description of the model is also a a differential equation model, thus a dynamic model. Under certain circumstances (for instance after learning occurs) the higher level model might not be predictive any more and requires to be changed. The two senses of dynamic in the term "dynamic structure" thus refer to the trajectories within the higher level model and to the change of the higher level model itself. In fact dynamic structures can be considered to be hyperdescriptions (McGregor and Fernando, 2005) of a functional subset of the dynamic relationships of the NS.

[6] By complexity we mean that there is a tradeoff between functional segregation and integration as mathematically defined by Tononi et al. (1998).

[7] For instance Husbands et al. (1998) have explored the evolvability of control architectures endowed with gas neuromodulators; i.e. where behavioral performance is achieved through the interplay between fast variability (neural spikes) and its slowly and geometrically distributed modulation (gas neuromodulatory networks)

Acknowledgments

References

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition*, 55:30–40.

Beer, R. D. (1996). Toward the evolution of dynamical neural networks for minimally cognitive behaviour. In Maes, P., Mataric, M., Meyer, J. A., Pollack, J., and Wilson, S., editors, *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behaviour*, pages 421–429. Harvard, MA: MIT Press.

Beer, R. D. (1997). The Dynamics of Adaptive Behavior: A research program. *Robotics and Autonomous Systems*, 20:257–289.

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behaviour*, 11(4):209–243.

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

Bressler, S. and Kelso, J. (2001). Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5:26–36.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence Journal*, 47:139–160.

Chiel, H. and Beer, R. (1997). The brain has a body. *Trends in Neurosciences*, 20:553–557.

Christensen, W. and Bickhard, M. (2002). The process dynamics of normative function. *Monist*, 85 (1):3–28.

Clark, A. (1997). *Being There: putting, body and world together again*. MIT, Cambridge, MA.

Clark, A. and Toribio, J. (1994). Doing without representing? *Synthese*, 101:401–431.

Collier, J. (1999). Autonomy and Process Closure as the Basis for Functionality. In Chandler, J.L.R./van de Vijver, G., editor, *Closure: Emergent Organizations and their Dynamics. Volume 901 of the New York Academy of Sciences.*

Damasio, A. (1994). *Descartes' Error. Emotion, Reason and the Human Brain*. Putnam's Sons, New York.

Di Paolo, E. (2000). Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. In Meyer, J.-A., Berthoz, A., Floreano, D., Roitblat, H., and Wilson, S., editors, *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior*, pages 440–449. Harvard, MA: MIT Press.

Di Paolo, E. (2003). Organismically inspired robotics. In Murase, K. and Asakura, T., editors, *Dynamical Systems Approach to Embodiment and Sociality*, pages 19–42. Advanced Knowledge International, Adelaide, Australia.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*.

Edelman, G. and Tononi, G. (2000). *Consciousness: How Matter Becomes Imagination*. Penguin Books.

Edelman, J. (1989). *The remembered present*. Basic Books.

Floreano, D. and Urzelai, J. (2000). Evolutionary robots with online self-organization and behavioural fitness. *Robotics and Autonomous Systems*, 13:431–443.

Freeman, J. and Skarda, C. (1990). Representations: Who needs them? In McGaugh, J., Weinberger, N., and Lynch, G., editors, *Brain Organization and Memory Cells, Systems, and Circuits*, pages 375–380.

Friston, K. (2000). The labile brain. i. neuronal transients and nonlinear coupling. *Proceedings of the Royal Society London B*, 355:215–236.

Funahashi, K. and Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6:1–64.

Husbands, P., Harvey, I., Cliff, D., and Miller, G. (1997). Artificial Evolution: A New Path for Artificial Intelligence? *Brain and Cognition*, 34:130–159.

Husbands, P., Smith, T., Jakobi, N., and OShea, M. (1998). Better living through chemistry: Evolving GasNets for robot control. *Connection Science*, 10(3–4):185–210.

Jonas, H. (1966, 2001). *The Phenomenon of Life: Toward a Philosophical Biology*. Northwestern University Press, Evanston, IL.

Kampis, G. (1991). *Self-Modifying Systems in Biology and Cognitive Science: A New Framework for Dynamics, Information and Complexity*. Pergamon Press, Oxford.

Kauffman, S. (2001). *Investigations*. Oxford University Press.

Keijzer, F. (2001). *Representation and behavior*. MIT Press, Cambridge, MA.

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

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences*, pages 169–194.

Maturana, H. and Varela, F. (1980). Autopoiesis. The realization of the living. In Maturana, H. and Varela, F., editors, *Autopoiesis and Cognition. The realization of the living*, pages 73–138. D. Reidel Publishing Company, Dordrecht, Holland.

McGregor, S. and Fernando, C. (2005). Levels of description: A novel approach to dynamical hierarchies. *Artificial Life*, 11(4):459–472.

Millikan, R. G. (1989). Biosemantics. *Journal of Philosophy*, 86 Issue 6 (June):281–297.

Moreno, A. and Lasa, A. (2003). From Basic Adaptivity to Early Mind. *Evolution and Cognition*, 9(1):12–24.

Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence and Technology of Self-Organizing Machines*. MIT Press.

Parisi, D. (2004). Internal robotics. *Connection Science*, 16(4):325–338.

Pattee, H. (1977). Dynamic and linguistic modes of complex systems. *International Journal of General Systems*, 3:259–266.

Port, R. and van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. MIT Press.

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

Rosenblueth, A., Wiener, N., and Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of Science*, 10(1):18–24.

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

Ruiz-Mirazo, K. and Moreno, A. (2004). Basic Autonomy as a Fundamental Step in the Synthesis of Life. *Artificial Life*, 10:235–259.

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3:417–458.

Seth, A. and Edelman, G. (2004). Environment and behavior influence the complexity of evolved neural networks. *Adaptive Behavior*, 12(1):5–21.

Skarda, C. and Freeman, W. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10:161–195.

Stewart, J. (1996). Cognition=life. implication for higher-level cognition. *Behavioral Processes*, 35:311–326.

Thelen, E. and Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Tononi, G., Edelman, G., and Sporns, O. (1998). Complexity and coherency: integrating information in the brain. *Trends in Cognitive Sciences*, 2(12):474–484.

Turing, A. M. (1988). Computing machinery and intelligence. In Collins, A. and Smith, E. E., editors, *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence*, pages 6–19. Kaufmann, San Mateo, CA.

van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioural and Brain Sciences*, 21:615–665.

Varela, F. (1979). *Principles of Biologicall Autonomy*. North-Holland, New York.

Varela, F., Lachaux, J., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews on Neuroscience*, 2:229–239.

Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind. Cognitive science and human experience.* Cambridge MA, MIT Press.

von Uexküll, J. (1982). The theory of meaning. *Semiotica*, 42(1):25–82.

Ziemke, T. (2003). What's that thing called embodiment? In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*.

Biographies

Xabier Barandiaran

Xabier Barandiaran is a PhD student at the University of the Basque Country. He is a graduate student in Philosophy from the University of Deusto (Bilbao, Spain) and obtained a Msc in Evolutionary and Adaptive Systems at the University of Sussex (Brighton, UK). His main research areas are the philosophy of Artificial Life, neurophilosophy, enactive cognitive science and naturalized epistemology.

Xabier Barandiaran

Departamento de Logica y Filosofia de la Ciencia

Universidad del Pais Vasco UPV/EHU

Apartado 1249 / 20080 San Sebastian / Spain

Email: xabier@barandiaran.net

Alvaro Moreno

Alvaro Moreno is Professor of Philosophy of Science at the University of the Basque country, in Spain, where he has created a research group specialized in Complex Systems, Philosophy of Biology and Artificial Life. He is author of over a hundred publications and organizer of several international workshops on the relationship between Artificial Life and AI.

Alvaro Moreno

Departamento de Logica y Filosofia de la Ciencia

Universidad del Pais Vasco UPV/EHU

Apartado 1249 / 20080 San Sebastian / Spain

Email: alvaro.moreno@ehu.es