

Adaptivity: from metabolism to behavior

version 1.1

Xabier Barandiaran# <barandi@sc.ehu.es>

Alvaro Moreno <ylpmobea@sf.ehu.es>

*Department of Logic and Philosophy of Science,
University of the Basque Country
Post Box 1249 / 20080 San Sebastian-Donostia, Spain
<http://www.ehu.es/ias-research>*

Abstract: Our aim is to clarify certain issues around the fundamental requirements for the appearance of adaptivity, specially in relation to life, agency, body, neural dynamics and autonomy as a constitutive kind of organization. We proceed from the bottom-up in order to provide a naturalized and universal characterization of adaptive behavior. In the first section we introduce the notion of basic autonomy and the origin of teleology and normative functionality as it appears in minimal living systems. We then distinguish those interactive processes that give rise to agency and specifically to adaptive agency as the capacity to recursively interact with the environment to modulate the viability of the organism. The paper follows explaining how an open ended complexity growth of adaptive agency requires the appearance of the nervous system. The nervous systems is conceptualized as a hierarchically decoupled subsystem that minimizes the interference with local metabolic constraints to sustain an open domain of sensorimotor interactions with the environment, giving rise to adaptive behavior.

Keywords: biological naturalism, autonomy, adaptive agency, adaptive behavior, hierarchical decoupling, nervous system.

Table of Contents

1.Introduction.....	2
2.Autonomy and intrinsic teleology.....	3
2.1.The self-maintaining and far from equilibrium organization of minimal autonomous systems	3
2.2.Normative functionality and intrinsic teleology naturalized.....	5
3.Adaptivity and agency.....	7
4.The Nervous System and Behavioral Agency.....	14
4.1.The limitations of agential mechanisms at the metabolic level.....	14
4.2.The Hierarchical Decoupling of the Nervous System.....	15
4.3.Behavioral Agency.....	17
4.4.Neurodynamic constraints, self-organization and adaptive behavior.....	17
5.The situated and embodied nature of the sensorimotor world.....	19
6.Conclusions.....	20

Author for correspondence.

1. Introduction

Adaptivity is the capacity that certain systems possess to modify themselves in order to adjust to changes in the environment. As is shown in many kinds of systems this capacity is obviously an essential property of living beings, which not only exhibit an enormous functional plasticity at the phylogenetic scale, but also show a rich repertoire of responses to environmental changes in somatic time. However, many ecological and social organizations are also considered adaptive. Last but not least, a certain number of artificial systems –going from virtual computer networks to physical “autonomous” robots—are also designed to show adaptive capacities.

This generic idea of adaptivity hides two important philosophical and scientific questions, namely: on the one hand, the presupposed distinction between system and environment and on the other hand, the fact that the idea of “adjustment” (and consequently of adaptivity) implies an irreducible normative dimension.

The first problem we can name it as the *problem of identity*; i.e. from the set of possible and arbitrary separations between system and environment, which are the ones we choose and why. For, if we are to attribute adaptive capacities to a system we must specify what is going to be adjusted to what else, and in what sense it is to be described as a system that adjust to the environment. In addition, when talking about adaptivity, we are not just referring to a mere structural adjustment between a system and its environment, but to a *functional* one; which brings us to the second problem.

The second problem is the *problem of normativity*. Normativity refers to the value attribution that is given to a process or object; e.g. adaptive or maladaptive to an interaction or structure in an organism, true or false to a cognitive state or belief, beautiful or ugly to a work of art, etc. Normativity challenges physicalist scientific approaches to the understanding of our world because it introduces a value asymmetry (good/bad, true/false, adapted/maladapted) in the description of nature, an asymmetry that is not present in any of the fundamental laws of physics. But, although alien to fundamental physics, normativity is an essential component of biology: whether an structure or interaction is adaptive or maladaptive for an organism is a value judgment that a scientist engaged in the analysis and synthesis of adaptive systems must do.

According to these considerations, adaptivity is a capacity that a system possesses to maintain its identity by compensating, through the use of certain internal mechanisms, those environmental changes that jeopardize its identity. However, a system that triggers functional interactions to compensate environmental dangerous perturbations only because it is externally guided to do so (as, suppose, a tele-guided robot) will never be a true adaptive system (since the adaptivity lies in the intelligent system that guided it). Thus, a true adaptive capacity can only come about through systems that define their own goals and their own normative adjustment to the satisfaction of these goals. It is precisely in living systems where the processes triggered are not divorced from its internal organization; they are causally connected to its maintenance. Thus, living systems are true adaptive agents because they generate actions in order to achieve the goal themselves generate. As Kant acknowledged more than two centuries ago, organisms seem to have internal goals.

Thus, a full understanding of adaptivity is strongly related to the naturalization of concepts like goal, normativity, identity and functionality. To do so we have to look at the basic organization of living systems. But even the simplest organism is a highly complex system. If we try to understand from a naturalistic perspective what is adaptivity, we have to search for the minimal

organization that might support autonomous, plastic interactions with its environment. Explaining adaptivity requires that we specify under which basic or minimal organizational conditions does it appear, given the best available law-like understanding of the universe and the biological constraints that are observed upon them. Though it is obvious that the appearance of adaptivity requires the existence of certain *level of complexity*, we do not know which *kind of complexity* is required for displaying adaptive capacities. In other words... how do we draw the boundaries between adaptive and non-adaptive complex systems? If we are not to believe in rigid boundaries still... what makes some systems to be more adaptive than others? In addition... what kind of transitions determine the increase in complexity of adaptive mechanisms? How is normativity expressed through these transitions?

A naturalistic account of adaptivity should address all these questions from two complementary perspectives: a) an *evolutionary* perspective, which should account for the diachronic emergence of adaptivity (what kind of evolutionary transitions permit the appearance of adaptive capacities) and; b) an *organizational* perspective, which should account for the synchronic emergence of adaptivity from the bottom-up (how is adaptivity sustained and enabled by underlying, more fundamental, processes). At the same time the answer should be grounded on the available scientific knowledge and be capable to provide productive feedback to science both at empirical-analytic and constructive-synthetic (biorobotics and simulation of adaptive behavior) levels.

The structure of this paper is as follows. In the first section we introduce the notion of basic autonomous systems to characterize living beings, i.e. systems capable to define and maintain their own identity. Autonomy provides the naturalized grounding for normative functionality, teleology, agency, and other fundamental concepts for the understanding of adaptivity. The next section focuses on the appearance of the nervous system within basic autonomous systems, the transformations it requires and induces in biological organization leading to the notion of a hierarchically decoupled dynamic domain in control of adaptive behavior. We conclude with some fundamental properties of adaptive behavior that have been intensively studied in recent adaptive science: embodiment, situated and emergent functionality.

2. Autonomy and intrinsic teleology

It looks intuitive that adaptivity implies the distinction and specific relationship between a system doing something by itself –an agent—and an environment. This fundamental dichotomy is very often taken for granted and the characterization of adaptivity is reduced to the establishment of the kind of relationship between agent and world. But if we are to proceed from the bottom up, not taking any distinction for granted, we shall first ask what kind of natural processes can constitute a system that produces its own identity, separating itself as different from its environment and establishing some functional interactions with it.

2.1. The self-maintaining and far from equilibrium organization of minimal autonomous systems

A quick look at the kind of processes that surround us shows that the universe has evolved producing at some places forms of order (like rocks or galaxies), while, at some others, matter shows no cohesion at all (such is the case of gases, for instance). The ordered matter takes two different forms: in some cases basic components appear lumped together constituting conservative static structures and in some others they constitute dissipative dynamic structures.

The first type refers to spatially ordered forms of assemblage of material sub-units, where this order is temporally instantaneous, like in rocks or crystals or temporally unfolded, like atoms or planetary systems. In both cases the form exhibited is just an expression of the intrinsic nature of a set of components which interact under certain conditions, something that will stay indefinitely once created: i.e., energetically all these systems are conservative.

The other form of order is dissipative, namely, it appears in far-from-equilibrium (FEE) conditions. This form of order is found among the so-called “dissipative structures” (Nicolis & Prigogine, 1977). A dissipative structure¹ is a set of interacting elements that generate a cohesive dynamical pattern in FFE conditions. Examples of this type of system are whirls, hurricanes, oscillatory chemical reactions or living beings. All these different systems have in common the fact that a huge amount of microscopic elements adopt a global, macroscopic ordered pattern in the presence of a specific flow of matter and energy: i.e., under certain boundary conditions. Interestingly, their internal dynamic cohesion is not only a consequence of the material features of their components, but also of the achievement and maintenance (in far away from equilibrium conditions) of some type of circular causality, since the very macroscopic pattern contributes to the maintenance of the dynamical cohesion at the microscopic level. These systems are able to generate and maintain, through recursive dynamics, a new way of correlation among their constitutive elements that otherwise would remain disconnected. This recursivity is precisely what provides a minimal form of self-created identity.

Now, it is obvious that these systems are too simple to produce any interesting form of interactive process with the environment. Thus, in order to search the origin of adaptivity, we have to look for forms of (self)organization capable to evolve and generate complex and diverse ways of self-maintenance. Among the wide set of self-maintaining organizations, those based on chemical processes are of particular interest, because they allow the construction of complex recurrent organizations through the creation of local and selective constraints. This kind of specific constraints is possible at the chemical level due to the action of the shape and reactive capacities of its constituent components, the molecules. This is the organizational framework of the early prebiotic evolution.

During the process that gave rise to life a fundamental step was the appearance of systems whose productive activity included the construction of a selective and functionally active membrane. This change led to a progressive take over of (at least part of) those boundary conditions ensuring the maintenance of the system. In other words, a form of organization appeared in which the conditions of its maintenance were actively controlled by the very organization, thus creating a first and self-sustained separation between system and environment. We call them autonomous (from the Greek *auto-nomos*, self-law) because (some of) the constraints that define the dynamics of the systems are the result of the very organization of the system; in other words what determines the behavior of the system is not just physical laws and a set of externally defined boundary conditions but, somehow, the capacity of the system to redefine (part of) its boundary conditions.

Actually, this concept of (minimal) autonomy is very similar to the idea of autopoiesis developed by Maturana and Varela (1980). Our emphasis is, however, focused on the FFE and thermodynamically open nature of these systems, from which a crucial implication follows:

¹ A dissipative structure is in fact a self-organizing system. By self-organization we mean that local non-linear interactions between components generate a global behavior, which is maintained through a certain number of constraints of which at least one is a product of *the global pattern* (Ruiz-Mirazo, 2001). Note that the global pattern is not instructed (dynamically specified) from outside, nor can it be reduced to or predicted from the activity of any of its local components.

interactive dynamics are constitutive of the system. Thus, interactions with the environment is essential for the existence and definition of the very system and not something that comes additionally to its constitution, as the original definition of Maturana and Varela makes understand.

Two **constitutive processes** can be distinguished within this concept of minimal autonomous organization² (see figure 1A):

Constructive processes: constructive processes are those that participate on the continuous production of the system (e.g. chemical reactions). Constitutive processes appear networked creating a closed organization we call constructive cycle. The **constructive cycle** is defined by the production of a set of constraints that recursively regenerate the conditions of their production. The network of reactions that produces the components of the network itself (metabolism) constitutes the most basic example of a constructive cycle in a minimal autonomous system. However, given that this network only exists as a thermodynamically dissipative organization, its maintenance requires interactive processes.

Interactive processes: they are those processes generated by the constraining action exerted by the constructive cycle to the flow of matter and energy between the system's boundary and the environment³ so as to ensure the system's maintenance. At the molecular level interactive processes require systems endowed with a physical separation (a membrane) between their constitutive organization and the environment. An example of a minimal interactive process is active transport through the membrane, breathing, or adaptive behavior.

Practically all forms of present-day life are capable to perform a wide range of processes in their environments in order to ensure their maintenance. Living organisms at large define their own identity and distinguish themselves from their environments. The way in which they do it is through their metabolic organization: a self-producing network of chemical reactions that controls some of its boundary conditions. Thus, the basic organization of all present-day living beings is essentially what we have called *minimal autonomy*⁴.

2.2. Normative functionality and intrinsic teleology naturalized

Minimal Autonomy thus described already provides a naturalized criterion for functionality and normativity in natural systems. The issue of the attribution of normative functionality to natural systems is of mayor importance for the task of naturalizing adaptivity. As observers, we can define systems arbitrarily. For instance, if we are to use differential equations to build a mathematical model of a system we shall a) choose a set of observables, b) create some variables representing those observables in our model and c) perform operations on the system to establish how variations on an observable produces variations on other observables so that d) we can abstract a set of differential equations that specify the rate of change of the variables in the model. If the model

² Although conceptually separable interactive and constructive processes might be (and originally are) instantiated by the same organization.

³ When using the term environment we are not referring only to a physical environment but also to an ecological and social environment that is the result of evolutionary-ecological relationships, intraspecies social relationships and the very recursive effect of the agent on these socio-ecological environments.

⁴ Although necessary, this kind of organization is not sufficient for a minimal characterization of living beings. Life requires that a set of hereditary and non-reactive components be coupled with the dynamic internal medium of the basic autonomous organization (i.e., metabolism). These hereditary components support an open-ended capacity to conserve, instruct, modify and reproduce organizational complexity; in other words, inserting basic autonomous organizations in a historical and collective process of evolution (Ruiz-Mirazo et al. 2004).

adequately predicts the functioning of the system we have an adequate model. At this point we can attribute a “way of functioning” to the system; corresponding to the functions of our model. Nevertheless we are faced with the two problems presented in the introduction, namely: a) how to justify the selection of observables belonging to the system (attribution of *identity*) and b) how to justify that the system not only functions in a certain way but that, in addition, it “must function” that way and not in another (attribution of *normative functionality*). For artificial systems, we, as designers or users, can attribute a certain goal to the system according to *our* intentionality. Under the attributed goal or purpose we claim that the machine works *properly* or that it is broken and *malfunctions*; although the good or bad functioning of the system (the machine) is completely extraneous to the structure of the machine itself. Equally, what belongs to the system and what should be left out (as irrelevant) is determined by the goal we project upon the system; i.e. what is an essential part of the system (as distinct from its environment) is defined in relation to the desired functionality that we as designers or users expect to achieve in a set of contexts.

On the contrary, living systems in general and cognitive systems in particular are capable to define themselves (as we explained above) and, more specifically, to determine their own normative functionality: i.e. what is good or bad (right or wrong) for them does not depend on an external observer, designer or user but lays on their own organization. More specifically, in autonomous systems a process (constructive or interactive) is **functional** if it contributes to its self-maintenance (Bickhard 1993, Collier 1999). A process, in turn, becomes **normative** if it is dynamically presupposed by other processes in their contribution to the overall self-maintenance of an autonomous system (Christensen & Bickhard 2002): i.e. a behavior or internal component is dynamically coupled with the rest of components so that the overall maintenance of the whole organization depends on it. Normativity⁵ 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. 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 to sustain the network of reactions..., that in turn produces the membrane, and so on in a circular and interdependent manner. As a higher-level example, the normative function of the kidney is to filter blood, because the dynamic-metabolic organization of the rest of the organism relies on this blood filtering for their functioning and existence. 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 an autonomous system, identity and normative functionality is not observer dependent but intrinsically causal: 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

⁵ We distinguish between normativity, regulation and normative regulation. From an organizational and naturalist perspective **normativity** is a somehow transcendental property defined by the organizational conditions of possibility of a dissipative organization. **Regulation**, on the other hand, refers to the control or active compensation of perturbations according to a given goal state or rule. An organism might or might-not be able to regulate itself according to its autonomously defined norms. For instance it could be regulating its temperature to 42°C maintaining this temperature invariant in the face of perturbations although this temperature might be harmful. Thus, it could happen (and it often happens) that an organism is regulating itself badly, but still regulating. So there must be a principle by which, independently of the actual regulatory functioning of the organism, we can justify the claim that the organism is doing it wrong. This well/wrong, good/bad, value attribution must be naturalized in the normativity that the FFE organization that an autonomous system brings forth. We call **normative regulation** the regulation that is carried out according to the normativity of the organism.

dependency between processes). In other words, in autonomous systems what-the-system-does (the way it functions) and what-the-system-is (its structure) are highly intertwined, they merge together on its organization.

The holistic, integrated and self-maintaining organization of autonomous systems 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). Autonomous systems 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 autonomous systems the goal state of the system and the organization that instantiates it are one and the same thing. Autonomous systems have an implicit teleology since their internal causal circularity makes each process of the system a contribution to its global self-maintenance.

Thus, the basic kind of autonomy just described is the lower level, most fundamental, kind of autonomy: that of material and thermodynamic self-construction and self-maintenance, constitutive of all living beings, upon which higher levels of autonomy appear⁶. Basic autonomy generates a cascade of emergent properties such as identity formation, normative functionality and implicit teleology. By expanding the above analysis and properties to the interactive cycle of autonomous systems we will be able to naturalize a set of characteristics that are necessary to address cognitive phenomena. These are agency, adaptivity and explicit teleology.

3. Adaptivity and agency

From a dynamic point of view we can abstract a set of boundary conditions and an essential parameter value region to be necessary for the maintenance of a FFE system. Following Ashby (1952) we call these parameters and boundary conditions *essential variables*, and the range within which the system's organization can be maintained *viability constraints*. The FFE nature of autonomous systems makes at least one of their essential variables have an intrinsic inertia towards outside the viability constraints. Some of the essential variables are also non-controlled variables, in the sense that no-change of internal variable of the system can directly control its state. As a consequence only the coupled system-environment can maintain the essential variables within viability constraints (thus the importance of interaction processes in autonomous systems).

⁶ When referring to basic autonomy in living beings (unicellular, multicellular, with or without specialized organs, etc.) we mean the network of processes that constitute the self-maintaining and self-constructing organization of the system (excluding other biological functions such as reproduction, immune defense, etc.). Minimal autonomy on the contrary, refers to the minimal organization capable of recursive self-maintenance and construction. In the most simple autonomous systems interactive processes do not themselves perform cycles, since they are strongly coupled/embedded in the whole constructive cycle. For example, the consequence of the active transport on the system is modification of the rates of constructive processes.

In general among the system-environment relationships, and *from the point of view of the effect* of the process on the system, some processes are **functional** for the system when they contribute to maintain essential variables within their boundaries of viability. Some others can be **dysfunctional** when they “push” essential variables outside the boundaries of viability. And many others are **neutral** (they have no effect on the state of the essential variables). However, *from the point of view of the cause* of these processes we can also classify them as **active** (if the processes are triggered by the system as a whole) or **passive** (when the interaction is induced from outside or the process is the result of physico-chemical laws without any internally generated constraint acting upon the process—like osmosis). We call **agents** those systems that interact with their environments so that the changes produced between the system and its environment contribute to its self-maintenance (see figure 2 for an inclusive classification of the different kinds of processes). That a coupled process be active, rather than passive, is of fundamental importance for the task of naturalizing agency and deserves a more precise definition. By active we mean, in negative terms, that the interaction is neither produced by some external source nor by means of unconstrained physical laws (i.e. spontaneously and independently of the particular organization of the system). Positively we define active as the process that uses the organization of an autonomous system to produce a constraining effect in order to ensure its own self-maintenance. In addition, another crucial factor that characterizes active processes is the complexity asymmetry of the coupled processes, laden to the side of the autonomous system: i.e. the interactive mechanism of the system is, due to its holistic organization, more complex than the coupled process it sustains. More complex means that the set of variables internal to the agent are more functionally integrated and differentiated (Tononi, et al. 1998) than those of the environment involved in the dynamic coupling (for a quantitative comparison between internal and behavioral complexity see Seth & Edelman 2004).

Robustness is the capacity of a system to maintain its organization in the face of internal and external perturbations. This capacity can be achieved by a number of increasingly complex mechanisms: a simple buffering mechanism, a distributed homeostatic mechanism or an adaptive regulatory subsystem. In any case, a sense of margin of viability is required; otherwise even the smallest perturbations will destroy the system's organization. The simplest forms of self-maintaining systems, like hurricanes or candles, cannot really “do” anything in order to compensate new environmental conditions, they only have a margin of maintenance provided by a buffering or a flexible structure. But more complex forms of self-maintaining systems, like hypothetical cellular protometabolisms, achieve robustness only by performing functional interactions on their environment. The main difference between the simplest self-maintaining systems and these ones (autonomous systems) is that the increase in complexity is *used* to constrain some of their boundary conditions; in other words, the increase in complexity is retained because it contributes to its own maintenance. In order to keep self-maintenance under different environmental conditions they must “recruit” their internal organization to regulate the coupled processes, which now become interactive processes. Interactions are thus *active* and *functional* processes between the system and its environment. Active ion pumping by the membrane could be an example (Moreno & Barandiaran 2005, Moreno & Etxeberria 2005). This interaction becomes functional because, as a result of the ion-pumping, it produces an internal change that is necessary for the ongoing activity of the constructive cycle (reduction of ion concentration in the interior of the cell).

The most simple forms of autonomy are homeostatic, in the sense that they can compensate internal and external perturbations by means of a self-regulating mechanism that is integrated and distributed over their constitutive organization: i.e. the regulatory mechanisms are embedded on

constructive and interactive processes so that regulatory and regulated processes cannot be distinguished. The property of homeostasis corresponds with the idea of conservation of autopoiesis as originally formulated by Maturana and Varela in their original work (Maturana & Varela 1980). However, when FFE systems increase their complexity⁷ they become more fragile: noise and environmental perturbations effect their organization, which, given its holistic nature, gets easily disintegrated. Thus, as complexity increases even more, self-maintenance under relatively wider conditions requires special mechanisms capable to ensure adequate coupled processes. As it has been pointed out by W. Christensen (2006) in the absence of regulatory mechanisms, the capacity that a given system has to maintain its identity against perturbations depends on its capacity to find the appropriate collective pattern. Now, problems arise when the system becomes more complex: if the self-organization process is robust against variations in specific conditions the process will be reliable, but it will be difficult for the system to generate multiple finely differentiated global states. Alternatively, if the dynamics are sensitive to specific conditions it will be easy for the system to generate multiple finely differentiated global states, but difficult to reliably reach a specific state. Homeostatic mechanisms cannot ensure the maintenance of complex systems under different, variable environmental conditions. Hence, in earlier living systems self-maintenance requires an active control on its interactive and constructive processes, measuring different conditions and *monitoring* its own constitutive processes so as to avoid or *prevent* dysfunctional situations. In other words, in addition to performing a constructive or interactive process that contributes to self-maintenance, the system is also capable of switching between different alternatives, tuning them, etc. according to external changes. This is the essence of **adaptivity** recently defined by Ezequiel Di Paolo 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, a) tendencies are distinguished and acted upon depending on whether the states will approach or recede from the boundary and, as a consequence, b) 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, 2006)

Adaptivity requires the establishment of an explicit normative regulation. As Di Paolo sustains, in non-adaptive self-maintaining systems the natural distinction between self-maintenance and disintegration is not yet accessible to the system, unless it is also able to regulate itself with respect to a norm. Whereas in pre-adaptive systems self-maintenance depends on the range of values that the essential variables can take, adaptive systems have the capacity to modulate (internally and interactively) the trajectories of the essential variables of the constitutive processes (unlike homeostatic systems —see figure 1B). Interestingly, adaptivity requires a regulatory control over the basic functioning of the system. Therefore, in order to be adaptive, a system must be organized such that it exists a relative decoupling between the dynamics of a regulatory subsystem and that of its basic constitutive organization. This has important consequences, since some of the properties traditionally assigned to autopoietic systems cannot be derived from autopoiesis alone but presuppose adaptive capacities. In addition, in most cases, adaptive regulation takes place not just transforming outward tendencies into inward trajectories (i.e. not just avoiding negative tendencies) but actively seeking to improve the state of essential variables so that regulation takes place not just

⁷ In terms of number and variety of components and interrelations, as well as in the degree of integration of the different parts of the system.

in reference to the boundary of viability but as graded and directed by a “sense of well-being”.

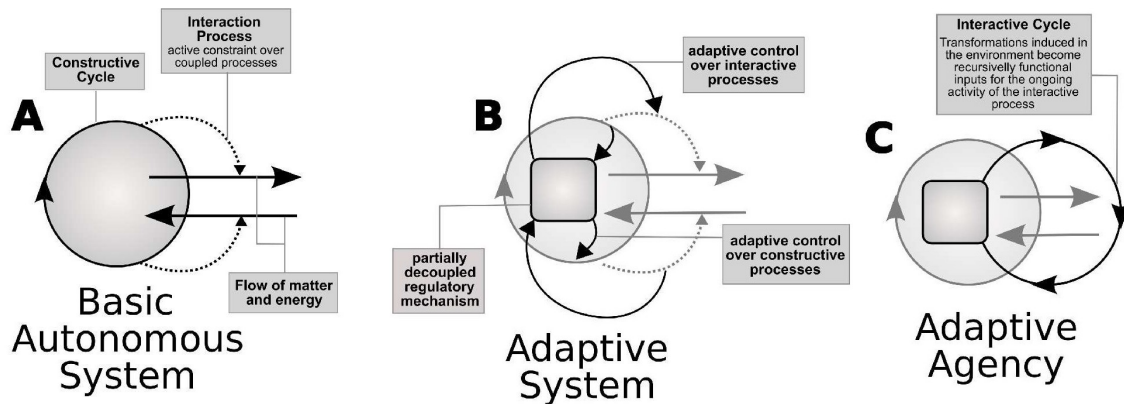


Figure 1: Conceptual diagrams of increasingly complex organizations. Diagram of constitutive processes in **Basic Autonomous Systems**: the constructive cycle requires a flow of matter and energy that is actively constrained by the system giving rise to interactive processes thus becoming a minimal form of agency (A). **Adaptive Systems** show a partially decoupled mechanism that regulates interactive and constructive processes adjusting and switching between different alternatives according to external perturbations and conditions (B). Finally **Adaptive Agents** appear when the interactive processes become a cycle (C).

Adaptivity is a capacity that all present-day living organisms have⁸. The simplest mechanisms of adaptive regulation fall into two different kinds. One is exemplified in the Operon activation and deactivation of genes as a switch between metabolic pathways according to certain environmental conditions. The other one is constituted by a whole subsystem of biochemical pathways not directly involved in the basic self-constructing metabolic network (as is the case of chemotactic agency in E-coli —see examples at the end of this section). But the common characteristic of both cases is that some degree of dynamic decoupling from the basic constitutive processes is required. In the first case metabolically off-line gene-strings act as instructive switches between different metabolic pathways. In the second case chemical pathways that are independent of the basic metabolic-constructive cycle sustain the interactive loop. This decoupling of genetic regulatory mechanisms from the basic metabolic network allows a selective choice among a large amount of not-yet functional dynamical states of the constitutive self-maintaining metabolic network. These decoupled systems open the possibility to consistently speak in terms of an internally generated mechanism for **normative regulation**. The capacity to distinguish and compensate tendencies requires that whatever makes a distinction and generates a compensation be dynamically differentiated from what it distinguishes and acts upon (which presupposes operational mechanisms to distinguish the different implications of equally viable paths of encounters with the environment). In other words, this implies a kind of measurement and control mechanism by which the system distinguishes and compensates tendencies that (if no compensation is carried out) would bring the system outside viability boundaries.

Thus, we are talking about two dynamical “levels” in the system: the constitutive level,

⁸ Homeostatic autonomous systems could have appeared on earth before the invention of organisms endowed with an instructed metabolism (a genetic code). But it is difficult to imagine how an adaptive control mechanism could spontaneously arise and adequately function in relation to metabolic needs without natural selection assuring that the coupling between the control mechanism and the controlled processes be functional for the system as a whole.

which ensures the ongoing self-construction of the system; and the (now decoupled) interactive subsystem, which regulates the dynamical rate of the former. This fact allows the regulatory subsystem to explore, in somatic time, many possible alternatives until finding a functional interactive outcome. This way, certain states of the low-level network will be stabilized when contributing to the creation of new forms of self-maintenance.

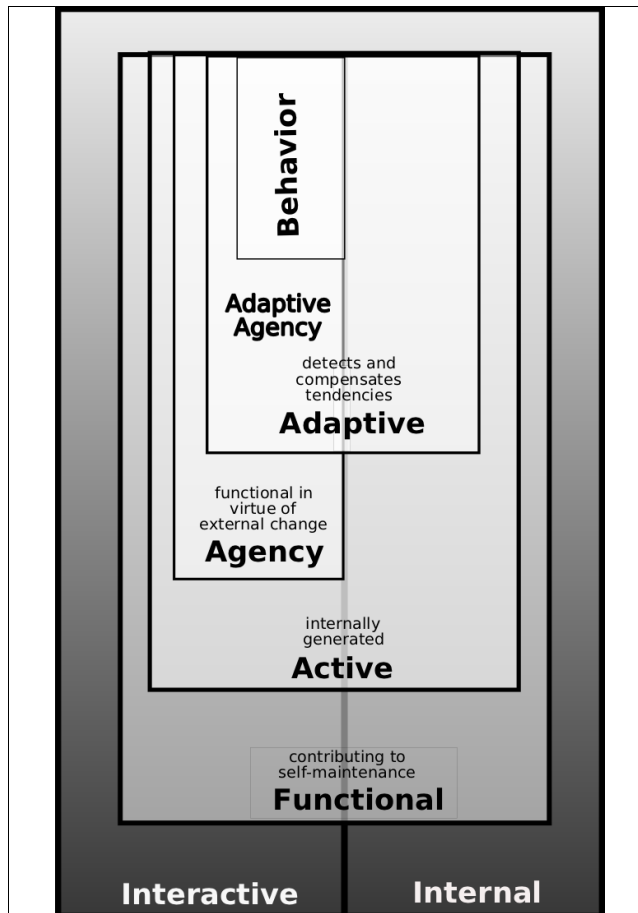


Figure 2: Inclusive classification of different processes in autonomous systems. Functional processes are those that contribute to self-maintenance, they can be both internal to the organism or interactive with the environment. Active processes are those generated by constraints produced by the system. Within active processes interactive ones give rise to agency: are those interactive processes which are functional in virtue of the changes induced on the environment. Adaptive processes are those active and functional processes which are carried out by a specific mechanism that detects and compensates tendencies of essential variables when they are sufficiently close to the boundaries of viability. Adaptive processes can be internal or coupled. Interactive and adaptive processes give rise to adaptive agency. Finally behavior appears as a subclass of adaptive agency.

We can naturalize the claim that some interaction or process is perceived as bad or good *by* and *for* the very system (and not only *by* and *for* the external observer); i.e. this well or bad functioning *for* the system is objective because it is detected and compensated *by* the system, in an effective, functionally integrated way. Thus, adaptive 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 meta-regulatory process emerges.

Adaptivity takes two basic forms depending on whether the mechanisms of regulation take place at the constructive level or at the interactive one. In the first case, internal or external perturbations are compensated by adjusting or transforming constructive processes (such is the case of the Lac Operon mechanism). The second form of adaptivity turns out to be of particular interest because it gives rise to **adaptive agency**: adaptation to perturbations is achieved through recursive interactions with the environment so that interactive processes become a cycle (see figure 1C). Thus, interactions become functional in virtue of the changes induced *outside* the system or, more specifically, on the relationship between the system and its environment (e.g.: motion). The transformations induced on the system-environment relationship become also functional inputs for achieving the adaptive regulation: hence the term cycle. For instance, when a bacteria swims in order to change its current environment towards another one which contains more nutrients. This set of interactive processes performed by the bacteria requires adaptivity because it is realized modulating effector processes according to the detected conditions of the environment.

The appearance of adaptive agency implies the emergence of detection and effector mechanisms by which the adaptive regulation is linked with the environment and the constitutive organization of the autonomous systems. In addition two regulatory processes appear linked to each other: the adaptive regulation of the essential variables through recursive interactions with the environment on the one hand, and the regulation of this interactive cycle according to its effects on the essential variables on the other hand. This intertwined regulation along with the detection-response coupling with the environment brings forth a genuine kind of purposeful agency, in which explicit teleology is extended into environmental interactions, creating a “meaningful” environment (a world) for the organism.

Last, but not least, we should introduce in this frame the concept of motility. **Motility** is the capacity an agent has of moving by its own means, so that it is able to exert fast (relative to its size) directional movements to change the environment looking for preferred conditions. Detection of and functional response to environmental relevant changes becomes in the case of adaptive motility a “sensorimotor” cycle, whose viability is strongly affected by size-time constraints. It is this high size-time (speed) constraint what characterizes sensorimotor adaptability from other forms of adaptability.

Let us summarize the proposed distinctions by providing some well known biological examples.

- 1 **Simplified bacteria**: let us imagine a bacterium living in a very homogeneous environment in which the necessary flow of matter and energy is assured by an external source of energy, so that the organism need not modulate its own functioning according to different environmental conditions. These bacteria could still be considered alive (since the machinery of self-construction and energetic and material exchange with the environment is at work), but should neither be considered an adaptive nor an interactive agent. An example of this case could be given by the sulfur bacteria, which lives next to vents (in deep ocean floors) that create a continuous gradient of sulfur from which the bacteria receives the necessary energetic and chemical input for self-maintenance (heat and sulfur). Due to the high homogeneity of its medium adaptive mechanisms in sulfur bacteria are highly degenerated, these bacteria could be

considered to be close to a homeostatic minimal-autonomous system.

- 2 **Lac-operon mechanism in E-coli bacteria:** The normal metabolism of E-coli depends on the presence of glucose in their environment. But when the levels of glucose on the environment become very low and another sugar (lactose) is abundant in the environment a mechanism called lac-operon is activated: the detection of lactose triggers the expression of certain dormant genes that in turn instruct a new metabolic pathway which metabolizes lactose. This metabolic mechanism is adaptive because it implies a (meta)regulation of the internal constructive processes according to the detection of a certain environmental condition (the presence of lactose and the absence of glucose) which jeopardizes the self-maintenance of the system. Nonetheless this form of adaptivity does not imply agency since the changes produced for self-maintenance are mainly internal.
- 3 **Chemotaxis in E-coli bacteria:** When the bacterium is close to a sugar gradient, certain proteins on its membrane detect sugar molecules and trigger metabolic paths which change the movement of its flagellum, so that the bacterium swims up the sugar gradient (instead of the usual tumbling movement). Swimming and tumbling are two different interactions that are appropriate in the sense of contributing to self-maintenance in differing conditions, but the bacterium can switch between them appropriately usually as the conditions change. This form of adaptivity should be characterized as *interactive agency* since the system interacts functionally with its environment (moves up-gradient until the necessary level of sugar is encountered). The interaction is functional in virtue of the transformations induced in the environment (in relation to the system the concentration of sugar increases) and the system operates recursively on these functional interactions (sugar-detection and motion are correlated). Among an open set of environmental variables the sugar concentration gradient shapes a functional world of interactions for the bacteria's self-maintaining capacity.

In short, some fundamental elements such as functionality, normativity and teleology are the consequence of the autonomous organization present in the simplest biological systems. Adaptive agents (as those autonomous systems capable of interacting with their environments detecting and compensating tendencies of their essential variables) extend an explicit teleology to their environments. As we shall see in the next section, biological organizations that support this kind of agency through biochemical mechanisms are severely limited on their capacity for *open-ended agential complexity*. Thus, without a qualitatively different internal organization of autonomous systems, adaptive agency reaches a dead end on its capacity for complexity growth. The appearance of the NS will solve this problem leading to a whole set of bodily changes and to a qualitatively different organization of adaptive agency.

4. The Nervous System and Behavioral Agency

4.1. *The limitations of agential mechanisms at the metabolic level.*

As we have seen, the appearance of adaptivity implies certain degree of *decoupling* within the organization of the agent: there should be mechanisms that support the interactive processes whose functioning is relatively independent from those involved in the constructive processes. This internal organization of adaptive agency requires therefore a relatively complex metabolism: instructed and regulated in somatic time by an off-line conservative structure (DNA) or by a ç different and relatively independent subsystem of chemical reactions. These early adaptive mechanisms are context-specific regulatory mechanisms and mostly genetically specified.

However, the use of biochemical mechanisms supporting interactive tasks severely limits the capacity to achieve increasingly complex forms of agency. The reason is that there is a serious bottleneck on the evolution of movement based agency supported in biochemical mechanisms. On the one hand, the bottleneck appears because the level of complexity that the adaptive subsystem can achieve (within the biochemical medium) without severe interference with metabolic processes, is very limited. On the other hand, as the size of the organism increases, the fast and plastic correlation between sensor and effector surfaces becomes harder (or even impossible in multicellular organisms) due to the slow velocity of diffusion processes. In eucariotic unicellular organisms fast adaptive responses requires chemically controlled (re)combination of structural aggregates such as microtubules⁹. But this new organization conveys a conflict between simultaneous movement and reproduction (Buss 1987) and further increases in size restrict seriously the development of biochemically supported motility.

The appearance of multicellularity posed an important challenge in the evolution of agency, since at this size the metabolic organization cannot support fast and versatile motility. There are two causes of this problem: the enlarged internal distance between parts of the body, which needs to be connected in short delays (so that the organism can move fast and coordinately); and the need to selectively modulate the organization of connections (to get the adequate sensorimotor correlations) for versatile, plastic agency. Hence, if metabolic network plasticity were the only mechanism for accomplishing adaptive interaction and self-maintenance, the forms of movement based agency would probably be very limited at the multicellular size¹⁰. However, when in the development of some metazoans a new kind of cell (the neuron) started to differentiate, this limitation could be overcome. Neurons differentiated as cells capable of forming branches, interconnected through plastic electrochemical pathways and capable to propagate and modulate electric potential

⁹ In paramecium, for example, when the organism encounters an obstacle (or a toxic chemical stimulus) it changes the sense of the movement of its cilia and goes back and/or alters its speed. These reactions are triggered by differences in the electrical potential of its membrane (induced by the stimulus). Fast adaptive responses are organized by means of microtubules, which also participate in several internal tasks like chemical channeling and plastic reorganization of selected parts of the structure of the cell (cytoskeleton). Interestingly, sensorimotor coordination is supported by a continuous reorganization of the cell's cytoskeleton (i.e., altering the internal topography and the membrane morphology).

¹⁰ The case for plants with quick movement capacity is obviously a marginal case. Quick plant movements can be considered a consequence of coordination mechanisms, which are highly dependent on the metabolic-constructive infrastructure of the organism. For example, in flytrap plants, rapid closing movements of the leaves involve changes in water pressure controlled by metabolic mechanisms (Simons 1981). As a consequence, this form of motility lacks flexibility and plasticity, and therefore there is no possible evolutionary increase in the complexity of sensorimotor coordination.

variability¹¹. In fact, these interconnected cells led to the establishment (about 600 million years ago) of a dynamic network capable of managing an efficient coordination between sensor and motor/effector structures in multicellular organisms (Llinás 2001).

4.2. The Hierarchical Decoupling of the Nervous System

Since the very beginning of its evolution, the neural organization appeared as an extended network capable of producing a *recurrent* dynamics of specific patterns independently of the underlying metabolic transformations that the organism undergoes. Unlike chemical signals circulating within the body, which directly interact with metabolic processes due to their diffusive nature, the electrochemical interactions between neurons make open-ended recurrent interactions within the very NS possible. The NS constitutes a cellular infrastructure that converts metabolic energy into finely modulable electrodynamic processes thus creating a new dynamic level free from the thermodynamic constraints characteristic of the biochemical level of metabolic-constructive processes. What makes neural interconnections so special is that they create an incredibly rich and plastic internal world of patterns of fast connections, *hierarchically decoupled* from the metabolic processes.

As we have elsewhere pointed out (Moreno & Lasa 2003, Barandiaran 2004, Moreno & Etxeberria 2005, Barandiaran & Moreno 2006) the hierarchical decoupling of the NS from metabolism means that metabolism generates and sustains a 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 providing the necessary energy to feed its dynamics. 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. Operationally speaking if we are to predict the state of the NS, hierarchical decoupling means that neither local states of cell metabolism nor the state of metabolic organs 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. In other words: the biophysical specificity, high connectivity, embodiment and situatedness of neural electrochemical dynamics make it irreducible to the metabolic substrate of its constituent components (the neurons) and the organismic processes of self-construction and repair (state of other body organs and processes involved)¹².

Hierarchical decoupling permits the specification of a set of operational primitives (the lowest level dynamical observables necessary and sufficient to model nervous activity) and their relationships, that constitute the control mechanisms of animal agency. It is commonly accepted that the primary operational primitives are the change of *membrane action potentials of neurons* over time (spikes) which conserve dynamic variability in terms of spike frequencies and time distance between spikes. *Synaptic connections*, on the other hand, specify a connectivity matrix (the transformation functions between primary operational primitives) while *neural modulators* (local and global synaptic modulators and action potential threshold modulators) become secondary

¹¹ The neuron is a cell specialized in connecting sensorimotor surfaces in a plastic, fast, and (metabolically speaking) cheap way.

¹² Nonetheless it must be noted that to the hierarchical decoupling of the NS follows a global coupling to some metabolic states of the organism in order to satisfy its adaptive needs.

operational primitives (since they become operational primitives in virtue of their effect on the spikes). The research for these dynamical primitives and its functional higher level causal organization constitutes the search for a neural “code”: the set of primitive variables and relationships that constitutes the dynamic domain capable of modeling adaptive behavior; in other words the kind of local operational differences that can make a systematic global difference in behavior (spike rates, interspike intervals, time of arrival, gas-net modulation, etc.).

The decoupling of neural processes from the underlying metabolic processes raises the question of how to characterize this specific dynamic domain. As we have seen, the active electrochemical conductivity of the components of the NS (the neurons) is organized in spikes or action potentials, thus generating a new dynamic domain that is built up tangentially to the metabolic processes of the organism, although realized and sustained on it. At the same time the all or nothing characteristic of neural spikes allows for a stable combination of them which, added to the network structure of the NS and the action of neural modulators, generates a 1) highly dimensional, 2) non-linear, 3) recurrent and 4) recursive domain. Nonlinearity allows for distinctiveness of states, while recurrency, provided by the structure of the network, allows circularity or reentry (Edelman 1987). Recursivity, on the other hand, happens because spikes can affect themselves through the neural modulators they activate. As a result the effective dimensionality of the system is constantly being redefined by its own activity. Thus, unlike the aforementioned forms of decoupling within the biochemical organization allowing primitive forms of adaptive agency, the hierarchical decoupling of the NS permits an open-ended growth of complexity in the forms of agency.

In addition, there is a causal link between the neural domain and certain external processes which belong to other dynamical levels (e.g. metabolic processes in the muscles). This causal connection is largely independent of energetic or material aspects, since neural states produce changes in body states by *formal* rather than energetic means in a kind of lock-and-key causality (what we shall call *formal* causality). This causal link is established through the pattern of spikes and not through the energetically determined causality through which these very patterns of spikes propagate. Thus, for example, the motor action caused by neural spikes is not determined by the electrochemical energy that constitute action potentials but by their form or pattern, which muscle cells “interpret” i.e., the process by which the neurotransmitters that neurons generate act selecting metabolic energy to produce movement. This process is similar to the electric patterns traveling along wires connecting two different computers: these patterns produce changes on the terminal not in virtue of the electric energy they convey but in virtue of the sequence of changes in amplitude and frequency. In other words the neurotransmitters that neurons generate (when a given pattern of spikes from further neurons arrive) trigger a cascade of chemical processes in the muscles that convert patterns of spikes into mechanical work.

Thus, the NS, hierarchically decoupled and endowed with the capacity for formal causation justifies the characterization of the neural domain as properly informational. Neural primitives can be considered as non-dynamical elements in relation to the underlying metabolic processes in the sense that, from the point of view of the modeling of metabolic-constructive dynamics of the organism, the NS appears decoupled. And from the point of view of musculoskeletal dynamics the NS acts as a formal control system, independently of the particular energetic details of how movement is achieved on the muscles. Thus we will hereafter use terms like “signals” to refer to the *informational* nature of the neural patterns in their functioning as integrated processes within living organisms.

4.3. Behavioral Agency

Interestingly, the appearance of multicellular organisms endowed with a (sub)system allowing fast, efficient and plastic agency, was necessarily accompanied with other important changes in their internal structure. The appearance and evolution of the nervous system brought along with it changes in the organization of internal circulation, the system of fixation, and even the body shape. Thus the unfolding of multicellular organisms whose way of life is based on motility requires radically different internal organization, namely a new bodyplan which allows a whole set of transformations both at the internal and interactive levels (Moreno & Lasa 2003). First, as we have already mentioned, the appearance of a decoupled non-biochemically based adaptive subsystem: the NS. Second, specific tissues and body structures capable of channeling metabolic energy into efficient mechanic energy (muscles, skeleton, etc.). *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 transformations that the organism undergoes by means of cell growth and reproduction.* We shall distinguish the new form of adaptive agency based on motility of those multicellulars endowed with a NS controlling a mechanical body as properly **behavioral agency**¹³. Thus, behavior is fast adaptive motility decoupled from morphological and, in general, metabolic-constructive processes.

It is precisely the hierarchical decoupling of the NS and its sensorimotor coupling with the environment what permits to study adaptive behavior in terms of sensorimotor dynamics (as it is the case in several fields such as robotics, cognitive neuroscience or embodied psychology) and qualifies behavior as a specific phenomenon distinct from generic biology. In contrast, the explanation of the interactions of plants with their environments would require the introduction of additional (non sensorimotor constraints) such as the rate of growth through cell replication according to exposure to the light, availability of water in the immediate surroundings and a host of alike agency-metabolism interdependencies. Nonetheless hierarchical decoupling from metabolism and sensorimotor coupling with the environment does not mean that the metabolic substrate of behaving is irrelevant. On the contrary, as we have seen, to the local decoupling from metabolic processes a global coupling follows, so that metabolic needs act as external constraints in neural dynamics.

4.4. Neurodynamic constraints, self-organization and adaptive behavior

The function in the overall organization of the organism of the neural domain is behavioral adaptivity, i.e.: adaptive maintenance of essential variables under viability constraints through the neural sensorimotor control of the interactive coupling with the environment (Barandiaran 2004). But the fact that the global dynamics of NS-body-environment produce an adaptive maintenance of essential variables under viability constraints does not specify how this functionality is achieved: i.e. what the dynamic organization of the NS is like and how it is related to behavioral adaptivity. Since it is impossible for metabolic needs alone to instruct functionally such a high dimensional space, the understanding of behavioral agency requires that we make explicit what kind of constraints act on the NS generating functional order.

Dynamically speaking the activity of the NS is defined by internal and external constraints. External constraints are those which are not the result of the activity of the NS itself. We can

¹³ What one of the authors has formerly called *Neural Agency* (Moreno & Etxeberria 2005).

distinguish two main kinds of external or innate¹⁴ constraints on the NS. *Architectural constraints* look like a first candidate for the “instruction” and generation of order in the NS, and can, in fact, generate highly constrained structures and dynamics in some primitive NSs¹⁵. These constraints are the result of genetically triggered anatomical-developmental processes selected in relation to the adaptive sensorimotor interactions they produced when coupled to body-environment systems. *Body signals*, on the other hand, act as strong perturbations of neural dynamics through specific signals (pain, hunger, pleasure, etc.) that operate modulating the overall activity of the NS (generally through specific neural modulator pathways such as the dopaminergic system). Thus, basic constraints are generally originated on metabolic needs and can be considered innate from the point of view of the early activity of the NS¹⁶. The role of body signals is to regulate the behavior sustained by neural activity in relation to the adaptive needs of the organism. These signals shape neural dynamics towards the satisfaction of certain goals, *basic* refers to the origin of these goals on the basic autonomous organization of behavioral agents.

So far, thus, neural dynamics can be captured through the specification of a) neural primitives, b) a set of innate (architectural and body signal) constraints, and c) environmental sensorimotor interactions. But in most of the known neural systems the complexity of the possible neural dynamics appears subdetermined by these constraints. Nonetheless we still find functional order. Therefore, different (non-external) principles of order are required to explain it. More specifically, the dynamics of the NS enter a process of local and interactive self-organization through the recursive activity of neural dynamics and sensorimotor interactions. The hierarchical decoupling achieved through the electrochemical functioning of neural interactions and their capacity to establish a highly connected and nonlinear network of interactions provides a dynamic domain with open-ended potentialities, not limited by the possibility of interference with basic metabolic processes (unlike diffusion processes in unicellular systems and plants). It is precisely the open-ended capacity of this highly dimensional domain that opens the door for spatial and temporal self-organization in neural dynamics and generates an extremely rich dynamic domain mediating the interactive cycle; overcoming some limitations of previous sensorimotor control systems.

A paradigmatic example of self-organizing patterns in the NS is given by CPGs (Central Pattern Generators) where the interaction between neurons on a local circuit generates robust oscillatory patterns. But self-organization also appears at the level of the coupling between NS, body and environment. In fact, the effect of sensory perturbations propagates recurrently through the network generating muscle contractions, which in turn feeds back to sensory neurons both through the changes that movement induces in the immediate sensory environment and through proprioceptive feed-back. The recurrent embodied coupling of the NS to the environment results in adaptive behavioral patterns whose functional stability is the result of the dynamic integration of neural, body and environmental features. Examples of this kind of interactive self-organization are given by active categorical perception, optic flow navigation, swimming, and a set of well known and studied phenomena¹⁷.

¹⁴ By innate we do not mean here that there is a genetically determined architecture of neural pathways in the NS but that, given an evolutionarily stable environment, a developmental process triggered by environmental and genetic factors certain anatomical structures are stabilized as a result of recurrent interactions in the developmental process.

¹⁵ For instance the nematode *Caenorhabditis Elegans* contains precisely 302 neurons and about 5000 synapses with highly stereotyped connections whose complete wiring diagram is already well known and equivalent among individuals of the same specie.

¹⁶ Through different processes of learning basic body signals develop into a wide range of signals.

¹⁷ Situated and autonomous robotics (Maes 1990, Brooks 1991, Clark 1997, Pfeifer & Scheier 1999 and Mataric' 2002)

So let us recapitulate. In big-size (multicellular) living systems sensorimotor agency requires a dynamic domain decoupled from local metabolic constraints, namely, the NS. This system, embedded on a sensorimotor architecture, is organized in terms of internal and interactive self-organized processes constrained by innate constraints and body signals. The functionality of the system is defined by the satisfaction of these basic and innate constraints giving rise to **adaptive behavior**.

5. The situated and embodied nature of the sensorimotor world

At this level of sensorimotor adaptive behavior we can make explicit some more fundamental properties of natural behavioral processes that are found to be at the basis of all cognitive processes. Some of these properties might already appear on earlier (even unicellular) agents, specially in those based on motility and endowed with sensory and motor mechanisms, but their full fledged significance shows up in neurally guided behaving systems.

Behaving systems are **situated systems**, their relation with the environment is relative to their situation on it. The behaving organism is not coupled to the environment as a Watt Governor might be coupled to a water flow. Sensory input is not only a function of the environment and the transformations that the agent induces on it but a function of the *controlled relative position* of the agent in its environment. This is a fundamental property that happens to be ignored in most of the Artificial Intelligence literature. For instance, when functional behavior is taken to be the result of extracting statistical properties or patterns from a string of predefined inputs. The consequences are non-trivial: for example, a non-situated system that is reactive (i.e., whose output is determined by the instantaneous input and a historically non-modifiable internal structure) cannot solve a non-markovian task, i.e., cannot successfully classify an environmental condition if its detection requires to extract a sequential order. On the contrary, a situated system with a reactive controller can transform non-markovian tasks into markovian tasks only by means of exploiting its relative position in its environment (Izquierdo-Torres and Di Paolo, 2005).

In addition, functional behavior is defined by the body in two different dimensions what we shall call sensorimotor and biological embodiment respectively. Adaptive Robotics has taken more attention on **sensorimotor embodiment**, which is a function of bodily properties in relation to the environmental sensorimotor coupling of the agent. On the one hand, body's physical and mechanical properties shape possible interactions and relative positions through enabling biomechanical constraints. The space of motor outputs to be instructed by the organism is not a uniform multidimensional space defined by a number of degrees of freedom. On the contrary, motor embodiment defines a biased "landscape" within that space determined by the shape, elasticity of joints, relative orientation and a host of alike body constraints. An extreme case of motor embodiment is given by dynamic walking models (McGeer 1990) in which, even in the absence of neural control, a mechanical system (a couple of legs) determines a well structured environmental coupling with its environmental surface giving rise to coherent and robust walking behavior. Instead, in the same situation, a disembodied approach would have required an exhaustive control motor output anticipating trajectories and a host of feed back control mechanism. On the other hand, embodied sensory surfaces define a range of sensory inputs and particular transformations and filtering of them. Finally, both sensory and motor embodied surfaces appear highly intertwined due

has provided a set of insightful models of embodied and interactively self-organized behavior: obstacle avoidance (Brooks 1990), wall following (Steels 1991), behavioral categorization (Cliff et al. 1993) and a number of other interactive behavioral phenomena that exploit recurrent interactions with the environment.

to the circular and recursive nature of sensorimotor interactions that have evolved together. We can call these *enabling constraints* because they bias the potential dimensionality of the sensorimotor coupling so as to enable or facilitate self-organized developmental and adaptive interactions (cross-modal sensorimotor spaces, developmental scaffolding by bodily changes, structural adaptation with certain object size and shape, etc.). In addition, and from a computational perspective, embodiment also means that much of the cognitive processing is carried out as embedded on the structure and mechanical functioning of sensorimotor processes.

A less commonly emphasized kind of embodiment is **biological embodiment**, which defines the ecological network of interactive necessities of the agent in order to satisfy its basic biological conditions of possibility¹⁸. Biological embodiment is in continuous feed-back with the sensorimotor flow. In fact the world of a behaving organism is not so much an independent, physical world but the coupling of this external world with the “internal” world, bringing forth what von Uexküll (1940:1982) called the *umwelt* of each organism: the dynamics of the constructive cycle “expressed” through body signals harnessing neural dynamics so as to achieve metabolic needs. The primary function of sensorimotor dynamics is therefore to maintain the essential variables under viability constraints. So the primary sensorimotor correlations in the organization of behavior are defined by the effect of the sensorimotor coupling on the dynamics of its biological embodiment. In fact, the world that comes about through biological embodiment is a mapping between the sensorimotor environment and the basic autonomous viability conditions. This is, properly speaking, the adaptive environment that a natural agent defines.

Artificial agents built according to the so-called autonomous situated robotics show emergent behavior arisen from real (or realistically simulated) perception-action cycles: they are (or tend to be) able to measure the relevant parameters of the environment to control certain degrees of freedom of the system from the very situation in which the system finds itself (and not from the point of view of an external observer), and to physically act in it. But what is lacking in most (if not all) allegedly adaptive or autonomous artificial systems is, as Ezequiel Di Paolo has pointed out (Di Paolo, 2003), a self-concern of the processes they undergo “because the desired goal is not designed by the robot but by the designer”. In other words, these robots lack their own normativity, which is a consequence of their lack of biological embodiment. This means that artificially created adaptive agents are not fully embodied.

6. Conclusions

We have argued that adaptivity requires autonomy. Artificial adaptive systems are, in fact, possible only because normative criteria are externally imposed by human beings, who, as living organisms, are autonomous systems. Autonomous systems create their own identity and distinguish by themselves from the environment. However, autonomy is a necessary but not sufficient requirement for adaptivity. As we have seen, the appearance of adaptivity also requires a regulatory subsystem within an autonomous system, partially decoupled from the dynamical organization of the constitutive processes of the system.

Although minimal forms of adaptivity can be achieved by relatively slow and very limited regulatory mechanisms (i.e., gene expression-inhibition) adaptive motility does require a biochemical metabolic-like regulatory subsystem. Even in very small size organisms (prokaryotic organisms) movement-based adaptability is achieved by a fast metabolic-like subsystem.

¹⁸ By this term we mean not only metabolic self-maintenance but a number of organizational constraints derived from the evolutionary dimension of living beings (reproduction, kin caring, sexual selection, etc.).

The most interesting point related with motility is the appearance of a new kind of interactive processes not directly involved in the maintenance of the constitutive (metabolic) organization of the system. These new interactive processes will give rise to the sensorimotor domain. Therefore, sensorimotor processes are neither mere interactive processes performed by an autonomous system, nor just adaptive agency, but specific, metabolically quasi-independent interactive dynamics of movement-based adaptivity.

However, biochemically supported adaptive motility becomes a serious organizational problem as size increases. Thus, full-fledged behavior only appears when adaptive motility is supported by a non-biochemical, metabolically decoupled regulatory subsystem (the NS), allowing fast and plastic functional sensorimotor coordination at the multicellular scale. Adaptive behavior is therefore body movement through neutrally controlled sensorimotor interactions that satisfy biological constraints: behavior oriented towards survival and reproduction. Although self-organizing patterns appear in the neural domain, as a result of internal and external recurrent interactions, metabolic signals and innate architectural constraints still define the normative framework and teleology of adaptive behavior. The organism as a whole is an autonomous agent but from the point of view of sensorimotor dynamics adaptive constraints are externally defined (by biological or basic autonomous needs). In this sense, adaptive behavior lies within the biological domain.

Finally, the creation by the NS of a highly rich sensorimotor domain decoupled from the functioning of the metabolic organization not only allows an open growth of the complexity of adaptive processes, but it has also a collateral consequence: the progressive take over by the neural system of both the organization of behavior and of the body itself, opening the way to the appearance of the cognitive domain. The study of how this new domain has appeared, its relation with the biological domain and its own nature deserves another paper¹⁹.

References

- Ashby, W.R. (1952) *Design for a Brain. The origin of adaptive behavior*. Chapman & Hall, 1978 edition.
- Barandiaran, X. & Moreno, A. (2006) On What Makes Certain Dynamical Systems Cognitive. *Journal of Adaptive Behavior*, **14**(2), in print.
- Barandiaran, X. (2004) Behavioral Adaptive Autonomy. A milestone in the Alife route to AI? *Proceedings of the 9th International Conference on Artificial Life*. MIT Press, Boston, Massachusetts, pages 514—521.
- Bechara, A. (2004) The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition* **55**: 30—40.
- Berntson, G.G. & Cacioppo, J.T. (2000) From Homeostatic to Allodynamic Regulation. In Cacioppo, J.T., Tassinary, L.G. & Berntson, G.G. (eds.) *Handbook of Pshychophysiology*, Cambridge University Press, pages 459—481.
- Bickhard, M.H. (1993) Representational Content in Humans and Machines. *Journal of Experimental and Theoretical Artificial Intelligence* **5**:285—333.
- Bickhard, M. H. (2000). Information and Representation in Autonomous Agents. *Journal of Cognitive Systems Research*, **1**(2):65—75.
- Brembs, B., Lorenzetti, F.D., Reyes, F.D., Baxter, D.A. & Byrne, J.H. (2002) Operant Reward Learning in *Aplysia*: Neuronal Correlates and Mechanisms. *Science* **296**:1706—1709.

¹⁹ The relationship between adaptive behavior and the origin of cognitive phenomena is the object of another paper that can be read as continuous and complementary to this one (Barandiaran and Moreno 2006).

- Brooks, R.A. (1990) Elephants Don't Play Chess. in P.Maes, ed., *Designing Autonomous Agents*, MIT press, pages 3—15.
- Brooks, R.A. (1991) Intelligence without representation. *Artificial Intelligence Journal* **47**: 139—160.
- Buss, L. (1987). *The Evolution of Individuality*. Princeton, NJ: Princeton University Press.
- Christensen, W. D. (forthcoming). The evolutionary origins of volition. In D. Spurrett, H. Kincaid, D. Ross & L. Stephens (Eds.), *Distributed Cognition and the Will: Individual Volition and Social Context*.
- Christensen, W.D. & Bickhard, M.H. (2002) The Process Dynamics of Normative Function. *Monist* **85** (1): 3—28.
- Christensen, W.D. & Hooker, C.A. (2000a) Autonomy and the emergence of intelligence: Organised interactive construction. *Communication and Cognition - Artificial Intelligence* **17**(3-4): 133—157.
- Christensen, W.D. & Hooker, C.A. (2000b) An interactivist-constructivist approach to intelligence: self-directed anticipative learning. *Philosophical Psychology* **13**(1): 5—45.
- Cliff, D., Husbands, P. & Harvey, I. (1993): Evolving visually guided robots. In J.-A. Meyer, H. Roitblat and S. Wilson (eds.), *From Animals to Animats 2: Proc. of the Second Intl. Conf. on Simulation of Adaptive Behavior*, (SAB92), MIT Press/Bradford Books, Cambridge MA, pp. 374—383.
- Collier, J. (1999). Autonomy and Process Closure as the Basis for Functionality. In Chandler, J.L.R./van de Vijver, G., (Eds.), *Closure: Emergent Organizations and their Dynamics*. Volume 901 of the New York Academy of Sciences.
- Damasio, A.R. (1994) *Descartes' Error. Emotion, Reason and the Human Brain*. G.P. Putnam's Sons, NY.
- Damasio, A.R. (1999) *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. Heinemann, London.
- Di Paolo, E. (2000). Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. *Proc. of SAB'2000*, MIT Press.
- Di Paolo, E. (2003). Organismically inspired robotics. In Murase, K. and Asakura, T., (Eds.), *Dynamical Systems Approach to Embodiment and Sociality*, pages 19—42. Advanced Knowledge International, Adelaide, Australia.
- Di Paolo, E. (2006). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*. In press.
- Edelman, G.M. (1987) *Neural Darwinism: The Theory of Group Neuronal Selection*. Basic Books, NY.
- Edelman, G.M. (1992) *Brilliant Air, Brilliant Fire: On the Matter of Mind*. Basic Books, New York.
- Edelman, G.M. & Tononi, G. (2000) *Consciousness: How Matter Becomes Imagination*. Penguin Books, London.
- Elman, J.L., Bates, E., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press.
- Farmer, J., Kauffman, S. & Packard, N. (1986) Autocatalytic replication of polymers. *Physica D* **22**: 50—67.
- Friston, K.J. (2000) The labile brain (I, II & III). *Phil. Trans. R. Soc. Lond. B* **355**: 215—265.
- Funahashi, K., & Nakamura, Y. (1993) Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks* **6**:801—806.
- Harvey, I., Di Paolo, E., Wood, R., Quinn, M., & Tuci, E.A. (2005) Evolutionary Robotics: A new scientific tool for studying cognition. *Artificial Life*, **11**(1-2): 79—98.
- Harvey, I., Husbands, P., Cliff, D., Thompson, A., & Jakobi, N. (1997) Evolutionary Robotics: the Sussex Approach. In *Robotics and Autonomous Systems* **20**:205--224.
- Kauffman, S.A. (1986) Autocatalytic Sets of Proteins. *Journal of Theoretical Biology* **119**: 1—24.
- Kauffman, S.A. (1993) *The Origins of Order: self-organisation and selection in evolution*. Oxford University

Press.

- Kohler, I. (1962) Experiments with goggles. *Scientific American* **206** (5): 62—72.
- Le van Quyen, M. (2003) Disentangling the dynamic core: a research program for a neurodynamics at the large-scale. *Biol. Res.* **36** (1):67—88.
- Luisi, P.L. & Varela F.J. (1989) Self-replicating micelles: A chemical version of a minimal autopoietic system. *Origins of Life & Evolution of the Biosphere* **19**: 633—643.
- Maes, P. (Ed.). (1990) *Designing Autonomous Agents. Theory and Practice from Biology to Engineering and back*. MIT Press, Cambridge, MA.
- Mataric, M.J. (2002) Situated Robotics. *Encyclopedia of Cognitive Science*, Nature Publishing Group, Macmillan Reference Limited.
- Maturana, H. & Varela, F. (1980) *Autopoiesis and Cognition: The realization of the living*. D. Reidel.
- McGeer, T. (1990) Passive dynamic walking, *International Journal of Robotics Research*, 9 (2): 62-82
- McGregor, S. & Fernando, C. (2005) Levels of Description. A novel approach to Dynamical Hierarchies *Artificial Life Journal* 11. 459--472.
- Millikan, R.G. (1984) *Language, Thought and Other Biological Categories*. MIT press, Cambridge MA.
- Moreno, A. & Barandiaran, X. (2005) A Naturalized Account of the Inside-Outside Dichotomy. *Philosophica* 73: 11—26.
- Moreno, A. & Etxeberria, A. (2005) Agency in natural and artificial systems. *Artificial Life* **11**(1-2): 161—176.
- Moreno, A. & Lasa, A. (2003) From Basic Adaptivity to Early Mind *Evolution and Cognition* **9** (1):12—30.
- Moreno, A., Umerez, J & Ibañez, J (1997) Cognition and Life. The Autonomy of Cognition *Brain & Cognition* **34**(1):107—129.
- Nicolis, G. & Prigogine, Y. 1977) *Self-organization in Non-equilibrium Systems*. Wiley, New York.
- Pfeifer, R. & Scheier, C. (1999). *Understanding Intelligence*. MIT Press.
- Ruiz-Mirazo, K. & Moreno, A. (2000) Metabolism and the problem of its universalization. *Biosystems* **49** (1), pp. 45—61.
- Ruiz-Mirazo, K. (2001): *Physical conditions for the appearance of autonomous systems with open-ended evolutionary capacities*. PhD dissertation (complete version only available in Spanish – <http://www.sc.ehu.es/PhBioGroup>).
- Ruiz-Mirazo, K., Peretó, J., & Moreno, A. (2004) A universal definition of life: Autonomy and open-ended evolution. *Origins of Life and Evolution of the Biosphere* **34**: 323—346.
- Schore, A.N. (2003) *Affect Dysregulation & Disorders of the Self*. W.W. Norton & Company, New York, London.
- Seth, A.K. & Edelman, G.M. (2004) Environment and behavior influence the complexity of evolved neural networks. *Adaptive Behavior* **12**(1):5—21.
- Simons, P.J. (1981) The role of electricity in plant movements. *New Phytologist* **87**:11—37.
- Skarda, C.A., & Freeman, W.J., (1987) How brains make chaos in order to make sense of the world. *BBS* **10**:161–195.
- Steels, L. (1991) Towards a theory of emergent functionality in Meyer, J. Wilson, R. (eds.) *Simulation of Adaptive Behavior*. MIT Press, Cambridge MA, pages 451–461.
- Stewart, J.(1996) Cognition=Life. Implication for higher-level cognition. *Behavioral Processes* **35**:311–326.

- Tononi, G., Edelman, G.M. & Sporns, O. (1998) Complexity and coherency: integrating information in the brain. *Trends in Cognitive Science* **2** (12): 474—484.
- Turrigiano, G.G. (1999) Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. *Trends in Neuroscience* **22**:221—228.
- van der Vijver, G. & Chandler, R. (eds.) (2000) Closure: Emergent Organizations and Their Dynamics. *Volume 901 of the New York: Annals of the New York Academy of Sciences*, pages 112—121.
- van Duijn, M., Keijzer, F. A., & Franken, D. (2005) Principles of Minimal Cognition. *ECAL 2005 - VIIIth European Conference on Artificial Life*. Submitted.
- van Gelder, T. (1998) The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* **21**(5):615–628.
- Varela, F., Lachaux, J.P., Rodriguez, E. & Marinerie, J. (2001) The Brain Web: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience* **2**: 229—239.
- von Uexküll, J. (1940). *Bedeutungslehre*. (Bios 10. Johann Ambrosius Barth, Leipzig), [translated by Thure von Uexküll, 1982: The theory of meaning. *Semiotica* 42(1): 25-82; glossary p. 83-87 by T v U).