Function: embodied, embedded and autonomous

Xabier Barandiaran
xabier@barandiaran.net
http://barandiaran.net

18–04–02

Copyleft ⓒ

Versions

v.1.0.	18-04-02	Essay for the "History and Philosophy of Adaptive Systems" course by Proff. Margaret Boden at the Evolutionary and Adaptive Systems MSc at COGS, University of Sussex, Brighton, UK.

Format and Sources

html	sindominio.net/~xabier/textos/bdn/bdn.html
pdf	sindominio.net/~xabier/textos/bdn/bdn.pdf
ps	sindominio.net/~xabier/textos/bdn/bdn.ps
fonts	sindominio.net/~xabier/textos/bdn/bdn/

Citation

Xabier Barandiaran (2002) Function: embodied, embedded and autonomous **url**: sindominio.net/~xabier/textos/bdn/bdn.pdf

Contents

1 Introduction

According to a hermeneutic gadamerian attitude there is one single way to be objective and that is to make our prejudices, intentions and interests as explicit as possible in a cognitive quest (as every coursework and essay ought to be), and test them against the alterity of the text: be it the Bible, the episodes of life, the chapters of political history or the book of nature to whose pages we (as philosophers of adaptive systems), in this course, belong. We shall, thus, specify that it is not our concern in this essay to work upon a general theory of function, autonomy and adaption (which will go beyond the reasonable aims of an inmature coursework) but to detect and work (as a artisan would do) upon some problematic (and other hopeful) conceptual issues regarding the theoretica framework of the scientific-cognitive practice on complex adaptive systems. Our ultimate goal is to work towards a framework for the cognitive, so, although our discussion will be more grounded in biological phenomena, this will be so within the framework of a bottom-up continuity thesis by which we understand that cognition, as a late phenomena in the evolutionary history of organisms, must be understood and grounded in biological organization and history, in its biological conditions of posibility, biological trascendentals (*à la* 'naturalized' Kant). We tend to categorize complex realities in hierarchical levels (cells, organisms, communities, societies, etc.), and the idea that higher levels of complexity (and adaptation) can be explained form lower level 'dynamics' and components articulates an ambitious scientific program. But the theoretical and pragmatic demands of our scientific practice requires more than a program, more than an intuitive idea. Methods, frameworks in which experiments and theory are to be articulated and connections between different levels of analysis are required.

In this essay we explore the concept of function, specially its evolutionary version (whose significance is central in contemporary biology —and philosophy of the mind) and try to link it with the concept of autonomy, exploring how a normative notion of function can be sustained by present causal mechanisms. We try to explore how a next step in the hierarchy of complex systems could be achieved by redefining this terms in the cognitive domain.

2 Function

2.1 The function of functionalism

Functionalism stands for one of the most relevant theoretical frameworks in the history of biological and cognitive sciences. The relevance of the functionalist framework can be understood from a double requirement it fulfills when facing the understanding of adaptive systems:

1. *As the grounds for semantic, intentional (mental) and teleological explanations* whose presence in human understanding (be it philosophical, historical, or folk psychological) and self (conscious) comprehension and control, is essential. Some phenomena are seen as empirically untestable because they require a backwards causation (future outcome explains present trait or behaviour [1]): teleological and intentional phenomena belong to this class and a functional framework is required to ground the scientific explanations of such phenomena.

2. *As a higher order of description from purely physicalist or behavioral explanations* . When facing certain phenomenal spaces (specially those over the fields of physics and chemistry in a hierarchy of sciences) in which mechanistic explanations (in terms of the physical components and mechanisms involved in the phenomenon) do not reduce the complexity of the phenomena under study to epistemologically and pragmatically efficient grounds, functionalist explanations are seen as a necessary conceptual tool.

The functionalist framework, by decomposing adaptive systems into functional components, stating its causal and formal relations, and/or by specifying its selective advantages, can be said to be the common (or at least classical) ground to analyze and understand biological and cognitive systems. But what do we mean by functionalism? How is the concept of function articulated with this purpose?

2.2 Kinds of functionalism

Ned Block [1] identifies three senses of functionalism: decompositional, computational and metaphysical.

- *Decompositional* functionalism relies on the methodological strategy of decomposing a system into components and analysing its functional role.
- *Computational* functionalism (Putnam and Fodor), relying on the computer metaphor, proposes that functional states are equivalent to those achievable by digital computers or Turing machines.
- *Metaphysical* functionalism claims that mental states *are* functional states, independently of its physical realization. The main claim of functionalism is the claim that the essence of a function (and a functionally decomposable system) is not its physical realization but the causal structures (relations) involved among components.

[1]the bird is digging (now) *because* it wants to hide a swarm with the *purpose* of recovering it when inmediate food resources are unavailable (*future*)

Computational functionalism provides the epistemological grounding of a formalizable theory of the mind in terms of logical relations between states, inputs and outputs of a system. Since different physical devices can (in principle) instantiate the same formal (computational) relations, functionalism is often related to the so called *multiple realizability* hypothesis.

The behaviour of an ant can be taken as an example of functional explanation: if COLLECTING-FOOD, AVOID[low-quantity-sources], GO[to-big-energy-sources]. This is a functional (computational) explanation: the ant is in state COLLECTING-FOOD, and has to execute to outputs: AVOID and GO. But that is 'just' a functional interpretation both 'AVOID[low-quantity-sources]' and 'GO[to-big-enery-sources]' are human categories never present as such in any mechanism giving rise ant behaviour. What we interpret as GO[to-big-energy-sources] is a complex process that emerges from local interactions with other ants through trofolaxis (chemical 'communication') and path following behaviour [2] as well as from internal neural dynamics. This *independency* between the physical and the functional, although considered as the biggest strength of functional explanations (specially regarding the ontological *status* of mental events) introduces several dificulties when trying to cope with the complexity of the underlying physical mechanisms from wich functions are predicated. But before going into the analysis of such problems we shall point to an alternative conceptualization of function that trys to solve the following question (among others): How can we assure the interpretative adequacy (if our interpretation is correct or just a methaporical description far from what 'really is goin on' in the system under study) of a functional explanation[3]?

2.3 In Presence of Proper Functions

On the other side of the functional stand a more evolutionarily oriented concept of functionality has been developed by several authors (Wright, Millikan, Sober, Maynard Smith,...) We will focus on Millikan's account of 'proper' function, exposed in [15], as one of the most representative conceptualization of this notion of evolutionary function.

Evolutionary or Selected functions are defined in terms of history rather than "the item's present properties or dispositions" (p.288). The formal definition goes as follows (a simplified version of Millikan's definition is given here):

[2]Ants tend to follow the smell of other ants, if food is found in a place ants will return with the food to the colony, the paht will be travelled twice, the smell of the path increases, the probability of the path to be taken by other ants increases, when taken by other ants keeps increasing and a collective behaviour emerges: all the ants are following a path towards a succesfull energy source.

[3]This is doubtless one of the major problems of functionalism. In [5] Dennet emphasises that any trait or mechanism is 'in principle' subjected to an arbitrary functional interpretation.

The item A has the proper function F iff:

1. *A* originated as the reproduction of *A'* and *A* causally-historically exists because A' performed function F in the past; or
2. *A* originated as a product of some prior device *A'* that: a) had F as a proper function and b) that causes F to be performed by *means* of producing an item like *A*. (this will be a *derived* proper function).

In short: F is a proper function of A if the performance of F helped the proliferation of the ancestors of A. An example can be illustrative at this point: the function (F) of the shell (A) of a turtle is to protect the turtle because the shell has been reproduced over the generations precisely because of its protective function.

Millikan pretends to give a theoretical definition of function in the sense of proper purpose and not a universal definition of anything that can be said to perform a function (the rock functions as a paperweight). Nonthe-less she claims that, as a matter of fact, having (or performing) a function corresponds, in the mayority of cases, with having a proper function. For Millikan the approaches that try ground the concept of function in terms of structure or disposition fail to account for failure in purpose (nothing in the present structure can account for normative, the disposition is as it is). Historical analysis, on the other hand, does account for those cases.

Evolutionary functionalism provides the temptative conclusion of a happy coupling scenario between a functionalist theory of representations and an adaptationist evolutionary theory: the epistemological normativity of representations can be grounded in the fact that organisms must adapt to an objective environment. Thus adaptation and representation get married in the grounds of evolutionary function. Intentional explanations can be grounded in selection (the content of a representation is fixed by selection) and ants *do* AVOID[low-energy-sources] because it is evolutinarily advantageous.

3 Failure of function(alism)?

3.1 The origin of a function: in favour of organizational and not historical hermeneutics

But the 'happy' evolutionary functionalism is not out of theoretical and pragmatic problems. Millikan recognizes that the historical dependency of its definition of proper function runs out the possibility of (logically possible) accidental doubles to have a function [4] But she argues that such 'logically possible' cases are fallacious constructs that do not, as a matter of fact, happen in reality. Millikan follows that other definitions of function according to which accidental copies do have functions, are just descriptions of markers that do help us recognize functions but do not address the constitutive nature of function and purposiveness which is rightly addressed by her definition of proper function, from which the rest of the markers parasite. As a matter of fact, Millikan concludes, functions (and any non historical definition of them) depend on their being proper functions.

But the artificiality of accidental exact copies, twinearthian logical possibilities, and other analytic angloamerican scholastic affections do not run out the whole problematicity space of Millikan's definition [5].

The theoretical reason is that if selection (and consequent higher differential reproduction) is considered as a modeling factor of complex adaptive systems whose mechanisms (although selected by evolution) operate by means of their physical instantiations, selection cannot be considered as ontologically present in the system under study, selection is not a present cause of any behaviour or dynamic of the system, as pointed out by Maturana [9] [6].

Thus proper functions turn to be ontologically separated from the system in which they are performed. This turns to be specially problematic when considering the following sort of questions:

1. The question of the origin of a function. A function must perform as a function to be selected. Millikans account falls into a *petitio principii* because the function of a trait recursively refers to its historical function.

[4] Imagine that an accidental exact copy of a system with a proper function comes to existence (as if a copy of yourself was to appear suddenly behind you). In such case the exact copy will not have any proper function if it didn't have the history of the original system (your accidental copy's heart "does not, in fact, have circulating blood as a proper function [sic.]"p.292).

[5] It must be said that Millikan specifically considers her definition to be related to purposive function and not to "function as" (as the performance of an item within the context of a system –decompositional functionalism above). To quote an example from her own: "a diseased heart may not be capable of pumping, of functioning *as* a pump, although it is clearly its function, its biological purpose, to pump (...)"(p.294). What Millikan wants to have is a normative notion of function. "The problem [she tries to address, and solve] is how did the atypical members of the category that cannot perform its defining function *get* into the same function category as the things that actually can perform the function?" (p. 295). But categories and epistemological normativities do not belong to entities but to subject-object cognitive relations (are not ontological, but epistemological, categories), thus, it is, at least, problematic to ground the notion of function in such relational space and responding to epistemological problems that do not happen to be in the domain of the systems under study (but of its cognitive and normative categorization —in the domain of our understanding of them)

[6] According to Maturana functional explanations bring non present features to the explanation: the evolutionary history of an organism is not present in the mechanisms governing its behaviour. Functional explanations are semantic projections which could be metaphorically useful but ontologically misleading.

2. The question of functional innovation (new meanings, learning, new adaptive strategies, etc.). How do new functions arise remains unresolved in her proposal. Although Millikan does not pretend to address this issue in her account of 'proper' functions, it is of primary importance specially if we take into account that cognitive and adaptive systems often have to find innovative solutions to problems they never encoutered before.

3. The question of function determination when there is no access to history. How do we determine the function of a trait, or behaviour, when the history of the system it belongs to is unaccesible or nonexistent?

These questions can be summarized in the following one:

- How can we ground functionality to be ontologically tided to the system it belongs to and still have a 'proper' hermeneutic horizon in which evaluate different functional accounts?

Although Millikan's approach can be an interesting way to normativize the concept of function it leaves (as we have seen)some fundamental problems unresolved. In particular those related to organization (how function relates to its physical realization and other components in a system), origins and the nature itself of the concept of function. Her definition is circular in that sense (proper functions require previous functionality of ancestors, and so on *ad infinitum*).

3.2 Interaction, emergence and other problems of functionalism:

The problems pointed out on Millikan's account of functionalism are not the only problems that a functional explanation must face. As Emmeche points out [6] any functional account must face three major problems:

1. Formalization may not catch all the causally effective mechanisms

2. "The construction that implements the formal structure is still in need of our interpretation in order to give any meaning." Meaning is a projection of an observer if it cannot be linked with materiality (semantics are not intrinsic to syntax *á la* Searle).

3. Formally, functionally, equivalent systems can be absolutely non equivalent in real life where time and energy can be essential.

This problems point to the importance of the physical (energetical, spatial and temporal —dynamical) embodiment of functions. The embodiment and environmental embededness of adaptive (biological and cognitive) systems give rise to specific problems of functionalism that has been pointed out

by Andy Clark [2, 3] under the label of *interactive emergence*. Interactive emergence points to two major difficulties:

1. *Dissolution of functional accounts in highly interactive loops.* This is the agent-environment side of interactive emergence by wich behavioural functions are the outcome of complex interactive loops between agent and environment where the causal mechanisms are difficult to specify by classical computational functionalism since functional units/behaviours cannot be isolated from the agent-environment *metasystem*. In this sense Luc Steels [13] defines behaviour as an emerging from the direct interaction and control of different components and mechanisms with the environment and other agents. This way emergent functions cannot be understood as formal relations between components of an organism or adaptive mechanisms to certain environmental conditions. A dynamical definition of function is proposed instead with more emphasys on indirect control and the tendency of interactive asymptotic stability (the agent-environment system tends to a dynamic equilibrium)

2. *Dissolution of functional units in distributed causal networks.* Probably one of the major chalenges to functionalism comes from the connectionist approach, its main conceptual weapon being the focus on the importance of a distributed causal mechanisms where no specific component can be said to perform a specific function. The major problem of a distributed causal network is not that there are no directly specifiable physical components that perform independient functions (because functionalism does not require a physical decompositionality) but the fact that this distributed nature of the mechanisms constraints and modifies the behaviour or nature of a trait. Some things cannot be explained by (evolutionary) functional accounts because they are caused by complex underlying dynamics which are not subjected to functional analysis. Langton [8] points to the nonlinear nature of biological organism: their behaviour is the effect of the interactions between components (unlike linear systems where the global behaviour is the summ of the behaviour of the components analyzed in isolation) and an explanation of the system must study it as a hole, requiring a synthetic efford, rather than an analytic one.

4 Conclusion: Autonomy as an hermeneutic axis for functionalism

In this section we shall finally consider how the concept of autonomy can help solving some of the problems mentioned above (specially those concerning the mechanistic nature of a function and its normativity) and how autonomous robotics and computer simulations become necesary tools for this synthetic and holitic efford.

The concept of autonomy as a central notion for the study of biological, cognitive and adaptive systems in general has been largelly ignored by clasical funcitonalist. First introduced by Maturana and Varela in [10] under the concept of autopoiesis, *autonomy* becomes the key notion from which biological systems are understood from a systemic perspective.

We shall point out how the concept of autonomy can contribute to solve the problems mentioned above by considering some of the main characteristics of autonomy.

The basis of all autonomy is the active (internal and external) role of the system to contribute to its *self-maintenance*: the production of components and the active change of boundary conditions of the system to maintain the essential variables (energy, temperature, etc.) homeostatically. In fact autonomous systems wouldn't exist in nature without this central characteristic. Ruiz-Mirazo and Moreno [12] have developed the notion of *basic autonomy* by specifying the termodynamic and energetic requirements of a system to be autonomous in this sense. The physical realization of a function is analyzed by the authors dynamically integrating energetic and termodynamic constraints into the explanation. This way the function of a component can be explained not as its adaptation to an external feature of the environment but as its contribution to the self-maintenance of the system [?] (structure and function are thus integrated in the physical realization of the autonomy of a system [7]). The normativity of a function can be, thus, refered to its physical realization and not to its past history (the heart of your magic twin copy has a function becaus it is activelly contributing to the maintenance of the system). It should be noted that this reference to the self maintenance is compatible with the millikanian account but still solves some of its problematic holes and it is not constrained to its metabolic side. Metabolic reactions are sometimes too slow to maintain an internal stability under rapidly changing conditions, that is the role of the nervous system connecting sensory-motor surfaces to deal with environmental changes rapidly and without involving change in the purelly metabolic proceses. Thus the nervous system, metabolically decoupled but embeeded in the organism has to know how to evaluate the results of its body control. Affections and internal sensors become then major features to autonomously build an cognitive internal normativity [4]. A second level of functionality (cognitive functionality) can thus be described where nervous mechanisms anticipate the effect of environmental interactions for the self-maintenance of the system without having to produce the interaction itself (and subsequently compensate for the produced desequilibrium of essential variables).

[7] solving some (but not all) of the problems of the adaptationist program as critisized by Gould and Lewontin [7]

This way what grounds semantics is not an *absolute* epistemological stand which requires *absolute* observers but the intrinsic evaluative mechanisms of a cognitive agent.

The concept of function must be grounded in autonomy (from a bottom-up *basic autonomy approach*) if we want to develop appropriate tools to understand and conceptualize adaptive systems without falling into an hermeneutic relativism (everything can be interpreted as functional) but avoiding recursion to history as a causally effective mechanism. Autonomy becomes, thus, the hermeneutic axis of a bottom up explanation of adaptation and cognition. In this task computer simulations are to be found as necesary tools not only in the modeling of concrete phenomena but in the development of a proper conceptual frameworks, specially by the interactive-emergent nature of biological and congnitive functions. The synthetic enterprise but at work by the artificial life community is the first steep for a more sincere cognitive relation with our complex world.

References

[1] Block, N. (1980)"Introduction: what is functionalism?" in Block, N. (Ed.) *Readings in the philosophy of psychology*. Cambridge, MA, Harvard University Press. pp. 171-184.

[2] Clark, A. (1996) "Happy couplings: Emergence and explanatory interlock" Chapter 9. of Boden, M. *The Philosophy of Artificial Life*, Oxford University Press, 1996.

[3] Clark, A. *Being There. Putting Brain, Body, and World Together Again*. MIT Press, 1997.

[4] Christensen, W.D. and Hooker C.A., (in press) "Self-directed agents". Macintosh, J. (Ed.) *Contemporary Naturalist Theories of Evolutiona and intentionality*. Canadian Journal of Philospohy, 31, Special Supplementary Volume.

[5] Dennet, D. (1990b) "The interpretation of Texts, People and Other Artifacts" *Philosophy and Phenomenological Research* L, Supplement, 177-94, Fall 1990.

[6] Emmeche, C. (1992) "Life as an Abstract Phenomenon: Is Artificial Life Possible?" in Varela, F. and Bourgine, P. (eds) *Towards a Practice of Autonomous Systems. Proc. of the First European Conference in Artificial Life*. MIT Press, Cambridge, Mass., 1992. pp. 466-474. Online version was use: to be found at: http://www.nbi.dk/ emmeche/cePubl/92a.liabsphe.html

[7] Gould, S.J. and Lewontin R.C. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme" *Proc. of the Royal Society of London*, Series B, vol. 205, No. 1161 (1979) pp.581-598.

[8] Langton, C. (1991) "Artificial Life". Chapter 2 of Boden, M. *The Philosophy of Artificial Life*, Oxford University Press, New York, 1996.

[9] Maturana, H. (1978) "Biology of Language: The empistemology of reality" Chapter 2 in Miller, G. and Lenneberg, E. (eds.) *Psychology and Biology of Language and Thought: Essays in Honor of Eric Lenneberg*. Academic Press, 1978, pp.27-63.

[10] Maturana, H. and Varela, F. (1973) "Autopoiesis. The organization of the living". From Maturana, H. and Varela, F. *Autopoiesis and Cognition. The realization of the living*, D. Reidel Publishing Company, 1980. pp.63-135.

[11] Maturana, H. and Varela, F. (1980) "The Tree of Knowledge. The Biological Roots of Human Understanding" Shambala, 1992.

[12] Ruiz-Mirazo, K. and Moreno, A. "Autonomy in natural and artificial systems" *Artificial Intellingence* **117(3)**, 2001.

[13] Steels, L. (1991) "Towards a theory of emergent functionality" in Meyer, J. Wilson, R. (eds.) *Simulation of Adaptive Behavior*. MIT Press, Cambridge MA, 1991, 451-461.

[14] Varela, F. *Autopoiesis and a biology of intentionallity. Preceedings of a workshop on Autopoiesis and Perception, Dublin City University* 1991.

[15] Millikan, R.G. (1989)"In defense of Proper Functions" in *Philosophy of Science*, 56 (1989) pp.288-302.

This document was done using LaTeX 2_ε ; running on Debian GNU/Linux.

The use of free software is not 'just' a possible technological choice but the condition of possibility of a technological (cybernetic—control through applied knowledge) autonomy