Grounding Synthetic Knowledge:

An epistemological framework and criteria of relevance for the scientific exploration of life, affect and social cognition

Luisa Damiano[1], Antoine Hiolle[1] and Lola Cañamero[1]

[1]Adaptive Systems Research Group, School of Computer Science & STRI, University of Hertfordshire, College Lane, Hatfield, Herts, AL10 9AB,UK
luisa.damiano@gmail.com

Abstract

In what ways can artificial life contribute to the scientific exploration of cognitive, affective and social processes? In what sense can synthetic models be relevant for the advancement of behavioral and cognitive sciences? This article addresses these questions by way of a case study — an interdisciplinary cooperation between developmental robotics and developmental psychology in the exploration of attachment bonds. Its main aim is to show how the synthetic study of cognition, as well as the synthetic study of life, can find in autopoietic cognitive biology more than a theory useful to inspire the synthetic modelling of the processes under inquiry. We argue that autopoiesis offers, not only to artificial life, but also to the behavioural and social sciences, an epistemological framework able to generate general criteria of relevance for synthetic models of living and cognitive processes. By "criteria of relevance" we mean criteria (a) valuable for the three main branches of artificial life (soft, hard, and wet) and (b) useful for determining the significance of the models each branch produces for the scientific exploration of life and cognition. On the basis of these criteria and their application to the case study presented, this article defines a range of different ways that synthetic, and particularly autopoiesis-based models, can be relevant to the inquiries of biological, behavioural and cognitive sciences.

Introduction

In his seminal article of 1989, Christopher Langton introduces the "synthetic approach" (SA) as the methodology proper to artificial life (AL) — to "put living things together", "rather than take [them] apart" (Langton 1989, p.40). The methodological agenda that he proposes extends the focus of biological research to what is missing in the analytic approach traditionally applied to living systems. However, the plan does not merely consist in extending the focus from individual to relational components' properties; from matter to organization; from centralised mechanisms to distributed dynamics of self-organization, as scheduled by other 20th century biological research programs. Distinctively, AL's SA further aims to enlarge biology's perspective from terrestrial to alternative "made by man" forms of life, and to actually include in biological heuristics, besides the question "how does it work?", the question "why this and not that?" The intent is to investigate the essential principles of life, and to deal with the main issues about life, by attempting to "recreate" living things and their phenomenology. In other words: building artificially embodied and situated models of living systems and phenomena in order to explore, through experimental manipulation, aspects of life usually not accessible in natural systems and scenarios.

This foundational methodological plan still unifies the three main branches of AL developed over the last two decades, since "soft", "hard", and "wet" AL (Bedau, 2003, p. 505), in spite of their divergent methods, all continue to refer to the SA as their basic methodology, and tend to agree in its characterisation. In addition, all of them emphasise the genuinely scientific aspiration of this methodology, as opposed to the mainly technological purposes of other research programs within computer science, robotics and synthetic biology. Moreover, they tend to attribute to this methodology the same heuristic features,[1] which, very schematically, can be listed as follows:

(a) The programmatic inversion of the established order between analysis of behaviour and construction of models — the SA directing researchers to first embed their basic hypothesis on life and cognition in working artificial systems, then examine the behaviours they produce.

(b) The theoretical hypothesis that distinguishes between organisation and physical-chemical realisation of living and cognitive systems, and claims that these systems and their phenomenology can be recreated by implementing the former in new physical media i.e. artificial "embodiments" and "embeddedments".

(c) The emergentist framework which grounds living and cognitive behaviours not within the systems displaying them, but in the interplay between three basic organisational levels of life and cognition: the systems, their elemental components, and the environment(s) with which the systems interact.

(d) The correlated production of simple and generative models of living and cognitive systems, that is, models able to generate complex and unexpected behaviours through rather simple internal mechanisms. The latter models are designed to create complexity, not by themselves, but by participating in systems-components-environment(s) interactive dynamics.

Over the last few years considerable work has been done to extend the applicability of the SA within the domain of cognition (e.g. Pfeifer and Scheier 2000; Lungarella et al. 2003; Bedau 2003; Cañamero 2005; Dawson 2004; Froese and Ziemke 2009). These developments pertain to cognitive

[1] Cf. e.g. Luisi 2010, Pfeifer and Scheier 2000; Dawson 2002, 2004; Damiano and Cañamero 2010.

processes *lato sensu* (i.e. including affective and social processes)[2] and re-propose, in new terms, the epistemological question of how relevant this methodology is for the scientific study of natural complex behaviours.

At the origin of AL, this issue was mainly focused on living processes; as Langton put it, "the notion of studying biology via the study of patently non-biological things is an idea that is hard for the traditional biological community to accept" (Langton 1989, p. 52). Today this has significantly changed. Contemporary academic biological departments integrate areas of research such as bio-computation or chemical synthetic biology. This expresses the wide acceptance, within biology, of interdisciplinary collaboration grounded in SA. The SA is also being widely applied to the study of cognitive processes within communities such as behaviour-based robotics and embodied artificial intelligence and artificial life. The situation is still different for the study of affective processes, where embodied synthetic models are still a minority. The application of the SA to this area is, to a large extent, hindered by the lack of principled reflection regarding a number of unanswered epistemological questions, and this constitute an obstacle to the integration of the SA among the explorative practices accepted by the scientific community as sources of valuable insights for cognitive and behavioural sciences. These include, for example, questions such as: in what sense can systems endowed with artificial "embodiments" and "embeddedments" generate effective models of natural cognitive, affective, and/or social processes? In which ways and in what sense can the synthetic study of these processes provide significant advancements with respect to other models? Which are the criteria that permit to define the relevance of synthetic models for the inquiries of cognitive and behavioural sciences?

This article addresses these questions with the intent of taking a first step towards providing epistemological groundings to the application of the SA to the affective (and more generally other cognitive and behavioural) sciences. More concretely, our aim is twofold: (1) to define an epistemological framework (that is, a set of principles of knowledge) able to ground SA as a relevant methodology that also encompasses affective processes, in addition to other biological and cognitive processes already studied by embodied AI and artificial life; and (2) to derive from this framework a set of criteria of relevance for the synthetic modelling of all these processes, that is, criteria (i) valuable for all three main branches of AL (*soft*, *hard*, and *wet* AL) and (ii) able to define the relevance of their models for the scientific exploration of natural living affective and cognitive phenomena. This article pursues these objectives not through general and speculative dissertation, but by discussing a concrete case study: an interdisciplinary model of the development of attachment (and more generally affective) bonds in the area of developmental robotics.

Through the presentation of this case study, Section 1 introduces the epistemological issue we intend to face, as well as the epistemological approach we adopt. Section 2 describes in detail two principles of knowledge that we propose as an epistemological framework able to ground the application of the SA to affective processes. These principles are extracted from autopoietic biology founded by Humberto Maturana and Francisco Varela in the 70s (Maturana and Varela 1987). It is worth noticing that our use of autopoiesis is different from the usual one. We do not don't refer to Maturana and Varela's theory of life and cognition, as is often done in AL, to take inspiration for producing specific models of living and cognitive processes. Instead, we refer to Maturana and Varela's *theory of scientific knowledge*, and draw on some of its elements. This is in order to provide the synthetic study of life and the synthetic study of cognition with a shared epistemological framework, able to offer them common criteria of relevance for the models they produce. In Section 3 we formulate and discuss the meaning of these criteria. In Section 4 we apply them to the developmental robotics model of attachment bonds, and discuss its contribution to developmental psychology. In section 5 we present the range of different forms of relevance that synthetic models can have with regard to biological, cognitive and behavioural sciences.

[2] Cf. e.g. Nunez and Freeman 1999. In this article, when we refer to cognitive processes, we always refer to them *lato sensu*.

1. An interdisciplinary exploration of attachment bonds

Developmental robotics is a relatively recent area of research, located at the intersection of robotics and developmental sciences, within which the SA plays a crucial role. This area uses studies from developmental sciences not simply to construct "more autonomous, adaptable and sociable robotic systems" (Lungarella et al. 2003), but also to gain a deeper understanding of developmental processes. One of the programmatic goals of developmental robotics is to employ robots as tools to investigate, test and possibly further elaborate, in an interdisciplinary way, theories of development proposed by these sciences (Pfeifer and Scheier 1999, Pfeifer 2002, Sporns 2003, Cañamero 2005). This expresses the originality of this emerging area, which intends to use developmental theories not only for engineering purposes (to create better adaptive robotic systems), but also for genuinely scientific purposes, through the extension of the SA to cognitive developmental processes *lato sensu*.

Within this framework, our research on the development of attachment bonds provides a good example of research that intends to explicitly address these two goals (Cañamero *et al*. 2006), and this for the following main reasons.

Firstly, this kind of inquiry allows researchers to face one of the central issues that need to be successfully addressed to advance in the development of social robotics, that is, how to design robots that could learn from us, be accepted by us as social partners, and be able to adapt to our ever-changing social environments. The developmental approach deals with this issue on the basis of the idea that the most successful example of adaptation into our social and technological environment, without much prior knowledge, is given by infants. Following this approach, researchers have, for example, successfully managed to design robots that use algorithms to learn and adapt to new sensorimotor pairings (Berthouze and Lungarella 2004; Blanchard and Cañamero 2005; Andry *et al*. 2009; Hiolle *et al*. 2007). Other contributors have focused more closely on how developmental psychology describes infant development, investigating how infants explore and discover new features of the environment, particularly through drives like curiosity (Oudeyer *et al*. 2007) and seeking wellbeing through affect-driven interactions with objects and people (Blanchard and Cañamero 2005, 2006; Cañamero *et al*. 2006). Indeed, the latter contributions are addressing the issue of how positive affect, such as providing comfort, can promote an efficient and consistent learning experience, depending on the environment and especially the behaviour of the social partner.

Developmental robotics research on attachment bonds arose within this context, and, in accordance with the general orientation of the area, has a second goal: it also aims at contributing to the advancement of developmental psychology through the design of adaptive social robots modelled using scenarios, parameters and metrics that are also relevant to, and used by developmental comparative psychologists for, the study of attachment bonds in (human and non-human primate) infants. This modelling approach gives rise to robots that behave and interact with humans in ways that are comparable to young infants (in the specific variables of the phenomenon under investigation), and therefore could be used as tools to investigate and possibly further develop theoretical models about attachment bonds.

In the reminder of this section we briefly summarize the developmental robotics research on attachment bonds that we undertook within the Feelix Growing Project (http://www.feelix-growing.org/) in collaboration with developmental and comparative psychologists,[3] and are continuing within the ALIZ-E Project (http://www.aliz-e.org/). This work focuses on the mechanisms underlying the establishment and development of attachment bonds in the first two years of age, which has implications for all phases of affective development. As this paper is directed to introduce the epistemological issues related to the extension of SA to cognitive development, we do not provide here the technical details presented in other articles; for technical details, we refer the reader, for example, to (Cañamero *et al*. 2006; Hiolle *et al*. 2006, 2007; Hiolle and Cañamero 2007).

1.1 Development of attachment bonds

Human infants grow and discover their new environment most often accompanied by (or not far from) their mothers or primary caregivers. The skills they learn, and the objects and agents they encounter, are surely presented and assimilated within their cognitive and emotional experience with the constant help and assistance of these adult human beings alongside them. Attachment was originally defined (Bowlby 1969) as the affective tie between the infant and its primary caregiver which offers security and comfort when needed. In the last decades, developmental psychology has been trying to study how this affective tie influences cognitive and affective development of young children. This research has produced critical and revised versions of Bowlby's theory, which point out the complexity of attachment processes, as well as the dynamical and inter-individual character of the "dyad" child-caretaker (e.g. Tronick 2007; Keller 2008). They tend to describe this dyad as an inter-individual system whose components are involved in a dynamic co-determination which shapes the way the child interacts with his/her (social) environment, and re-shapes the way in which the care-giver(s) interact with the child. These critical developments strongly converge with the SA's assumption about the generation of complex behaviours,[4] and constitute the body of work that we took inspiration from. The remainder of this section briefly introduces the properties of attachment bonds we used and the robot model we produced.

1.2 Attachment bonds in infants

One of the main roles of attachment bonds is to provide mechanisms that permit to regulate ("negative") affective state, and particularly arousal, setting the grounds for the development of emotion regulation later in life. We therefore focused on arousal and its regulation in our model. This essential variable was designed to relate to the notion of excitement as defined in (Sroufe 1996), which, in the early months of life, is neither a positive nor a negative emotion or affect, but refers to the level of internal activity and external stimulations experienced by the infant. A high and sustained level would be too demanding and challenging, while a low level would not give rise to fruitful behaviour. Thus, maintaining a good level of this variable is desirable. This internal variable is close to the concept of arousal (Berlyne 1960), relating to the theory of optimal arousal, and the inverted U-shape hypothesis (Anderson 1990), where mammals try to maintain on average their arousal at a middle level where their physiology is optimal. In our investigation of infant development, the notion of arousal is very appropriate, as it is used in developmental psychology to assess emotional intelligence in newborns and its development (Brazelton and Nugent 1995). However, the notion of arousal is often used as a dimension of the two or sometimes three dimensions usually adopted in models based on the circumflex model of emotions (Russel 1980), such as in (Breazeal and Scassellati 2002). In this kind of model, arousal is an orthogonal dimension to the valence of percepts and behaviours, and the model offers a one-to-one mapping from a two dimensional vector from the arousal/valence space to a predefined emotion. In our work, however, we do not use the notion of arousal in the same way as these models. Instead, we see arousal as a variable related to internal activity, in terms of learning experience, which is implicitly tied to external perceptions, some being more stimulating then others, according to familiarity and complexity.

The robotics model that we designed is based on the notion of arousal, which we associated with the learning experience of the robot and how stimulating or familiar the experienced environment is — namely, the current sensorimotor state. To this end, our model asseses whether the current percepts are being correctly memorised and recalled, and this directly influences the arousal level of the robot: novelty increases arousal, familiarity decreases it. The robot does not have explicit drives or motivations beyond exploring the environment, and its behaviour is a function of the level of arousal. The human, playing the role of a "caregiver", also has an impact on the arousal of the robot, in accordance with the secure base paradigm: the arousal level is decreased when the human provides comfort to the robot, either via direct tactile contact or by being present on its visual field. This robot-caregiver system is a dynamical system that present the essential elements needed to reflect and test the hypothesis concerning the attachment bonds and caregiving influences on it: unfamiliar events and stimuli increase the arousal and provoke distress, and the attachment figure can then relieve this distress with comfort. Whenever the arousal is low, the infant-

3 Within the FEELIX GROWING project we worked in collaboration with Kim Bard and Jacqueline Nadel.

4 Cf. the Introduction of this paper, point (*d*) and (Damiano and Cañamero 2009).

robot would keep exploring its environment as long as there are unknown features, in order to further its learning experience.

1.3 Robot Model

The robotic system is based on a few simple hypotheses, as stated above. Firstly, the robot's only "motivation" is to learn the features it can perceive in its environment. The level of aroual of the robot is calculated as a function of the familiarity and novelty of these features. The arousal rises when the robot is stimulated, and decreases when the attachment figure provides comfort visually or via tactile contact. When the level of arousal is low, the robot will seek stimulation and carry on exploring. The learning system of the robot uses two different well-known neural networks, a Kohonen Map (Kohonen 1997) and a Hopfield-like associative memory (Davey and Adams 2004, Hopfield 1982). The arousal level depends directly on the variation of the weights of the Kohonen Map, and on the accuracy of the associative memory. Indeed, a high variation of the weights is consequential of the robot discovering new features, and a mismatch between the output of the associative memory and the current perception is proportional to the novelty thereof. The arousal level is calculated as the exponential average of these two contributions over a predefined time window. When the caregiver touches the sensors on the back of the robot, the arousal level decreases exponentially and faster than it could increase whilst being over-stimulated. According to the arousal level and predefined thresholds, the robot behaves as follows. When the level is in a medium range, the robot remains still and attends to the current stimuli. Finally, when the level is high, due to too many and familiar stimuli, the robot is "distressed" and it will seek comfort from its caregiver.
Using this robot model, we undertook a series of studies focused on studying different aspects of the development of attachment bonds with one or more human caregivers.. These studies rest on the interdisciplinary design of experimental scenarios, with the aim to provide insight and feedback to all the different disciplines involved. Several crucial aspects of the development of attachment bonds were under inquiry. These included the development of different attachment profiles; the influence of these different profiles on exploratory behaviors; the role of attachment bonds in the development of sensorimotor associations; and the development of attachment bonds in the presence of multiple caregivers (Cañamero *et al*., 2003; Hiolle et al. 2006, 2007; Hiolle and Cañamero, 2007).

1.4 The epistemological issue

The relevance of these studies for the construction of robots able to develop their skills and behaviours dependently on the interactions with their users is quite evident. However, considering the level of abstraction and simplification characterising the robotic architecture described above, the restricted possibilities of interaction of the "baby" robot with its human partner(s) and environment, the limited aspects of development of attachment bonds taken into account by the robot model, can we say that these studies are able to produce relevant feedback for human developmental psychology? Are there ways and conditions in which this "baby" robot can be fruitfully used to model and explore attachment phenomenology in humans?

If we refer to the widespread, and in our opinion incorrect, idea that a model system should "represent" the target system in all its aspects and behaviours, we have to answer in the negative. This notion, grounded in the classical "representationalist" scientific epistemology is also strongly questioned by the epistemological debate about synthetic modelling. The latter points out not only that necessarily scientific modelling fails in representing everything about target systems, but also that this is not its main goal. The basic purpose of scientific modelling is not to replicate the target system, but to investigate what are its relevant features. As in the case of synthetic modelling, the goal is to embed, in a model system, the scientific hypotheses about these features, and to test these hypotheses.
This argument has been proposed by both representationalist and non-representationalist epistemological approaches to AL. While the former use it to propose weak versions of the classical idea that scientific representation should be an exhaustive reflection of nature (e.g. Webb 2001), the latter use this argument to express the thesis that representationalism is not an appropriate epistemology for AL (e.g. Riegler 1992). According to the latter view, representationalist epistemological notions, based on the ideal of a science exploring objects independent from the observer, cannot orient the scientific practice grounded in the SA. Indeed, this methodology promotes a form of scientific knowledge which actively creates, and does not passively reflect, the phenomena explored. It refers to an observer who is the constructor, and not the old-fashion representationalist "spectator", of the systems he investigates. Moreover, the SA discards the representationalist dichotomy subjective/objective, as it proposes a way of doing science in which *facts* converge with *artefacts*, *discovery* corresponds to *invention*, *objective evidence* is not separable from *subjective construction*, and spontaneous manifestations of nature can be explored in the behaviour of artificial systems. As we showed in detail elsewhere (Damiano and Cañamero 2010), these remarks tend to lead non-representationalist approaches to AL that take inspiration from constructivist theories of scientific knowledge. These characterise science, in all its forms, as an activity of construction of objects of research (cf. e.g. Glasersfeld, 1995), and therefore propose to science epistemological notions and principles of knowledge which can be considered particularly appropriate for grounding, and supporting the scientific practice grounded in, the SA.
Our approach to the epistemological grounding of the SA belongs to this non-represationalist orientation. To address the issue of the relevance of synthetic models, like the robotic model presented above, for the inquiries of cognitive and behavioural sciences, we choose to refer to *autopoietic epistemology* (Maturana and Varela 1987). This is one of the best expressions of the constructivist epistemology developed across AL's scientific genealogy (Damiano and Cañamero 2010), and, as we argue, can provide shared groundings and common criteria of relevance to the synthetic study of life as well as to the synthetic study of affect and more generally cognition.

2. Grounding the SA in autopoietic biology

The connection between autopoietic cognitive biology and AL is strong. As already mentioned, Maturana and Varela provided more than an emergentist theory of life and an emergentist theory of cognition which are useful to inspire the production of synthetic models of biological and cognitive processes. As often neglected (even by AL researchers who take inspiration from these theories), Maturana and Varela elaborated an explicit constructivist theory of scientific knowledge, which proposes the SA as the proper methodology to investigate both life and cognition, at the theoretical level as well as the experimental one. This autopoietic version of the SA, formulated more than a decade before the Langtonian one, is based on epistemological notions and principles which give expression to the intuition at the basis of Langton's AL program, and, in this sense, can be considered as appropriate epistemological groundings for AL's SA, both in the domain of life and in that of cognition.

On the basis of these considerations, we saw in Maturana and Varela's production a source of epistemological elements useful to provide a shared epistemological framework to the synthetic study of life and that of cognition. In particular, we extracted from autopoiesis two principles which, as we try to show in the remainder of this section, are particularly significant with regard to this goal. Below we summarize these (well-known, to a large part of the AL community) principles, to make the paper self-contained and help readers who might not be totally familiar with this approach.

2.1 Principle 1: Explaining = Constructing

The first principle proposes an operational definition of scientific explanation, according to which explaining a phenomenon amounts to proposing a mechanism able to produce it (cf. e.g. Maturana and Varela 1987, chapter 1). Visibly, the aim of this postulate is to extend the classic view of scientific explanation. It juxtaposes the traditional notion "explaining = predicting" to a constructivist one, which, proposing the equation "explaining = constructing", can be applied to systems exceeding scientific capabilities of calculation and prevision. Requiring models able not to *predict*, but to *generate* the natural processes under inquiry, the principle locates the focus of scientific explanation not on *actual*, but on *possible* behaviours of the systems explored. That is, it grounds a category of scientific descriptions which is particularly appropriate for living and cognitive systems, since the kind of characterization it proposes cannot be affected by these systems unpredictability. AL's SA can be legitimately included within this category, for it presents the basic distinctive features characterizing the paradigmatic constructive description of nature grounded by autopoiesis in this principle. In particular, it shares the distinctive features of the main example of constructive explanation provided by Maturana and Varela, namely, the autopoietic explanation of life. Very schematically: *raison d'être* (the natural phenomena it intends to describe are untreatable through the classical predictive modelling), epistemological grounding (the constructivist postulate according to which knowing scientifically means to build objects of research), heuristic gender (operational characterizations of the natural processes explored), procedure (definition of a generator for the phenomenology to be described, and exploration of the phenomenology it produces), and, finally, the appellation "synthetic".

In Maturana and Varela's literature, the introduction of this kind of scientific characterization is described as implying a long series of shifts in classical scientific epistemology, which produce a new emphasis not only on *construction* instead of *representation*, on *generation* instead of *prediction*, on *possibility* instead of *actuality*, but also on *synthesis* instead of *analysis*. Indeed, in Maturana and Varela's production, the term *synthesis* defines the methodological orientation of autopoietic biology's theoretical program, just like, in Langton's literature, it defines the methodological orientation of AL's program.

On the basis of its principle of scientific explanation, autopoietic biology plans to formulate a procedurally new definition of life, which, instead of listing the main features of living systems, specifies a dynamical mechanism able to produce their phenomenology. Maturana and Varela call this kind of definition "synthetic", to distinguish it from the traditional "analytic" definitions of life presenting detailed lists of properties. The condition that this synthetic definition has to satisfy to be considered an appropriate explanation of life is expressed in terms of its theoretical productivity. The mechanism that it specifies has to show the capability of creating, from a set of elemental components, an entire biological domain. That is, it has to manifest the ability of generating, by the dynamical coordination of a set of elements, a minimal cellular system with its characteristic phenomenology. That is: not only cellular self-production, but, through this, also reproduction and evolution, to the extent to be able of producing, step by step, a differentiated living domain, as complex and populated as the terrestrial one.

This is the kind of scientific modelling of the living that Langton's characterization of AL intends to implement too, not at the level of a purely theoretical construction, but at that of an empirical one: the synthesis of "any and all biological phenomena, from viral self-assembly to the evolution of the entire biosphere", without restriction to carbon-chain chemistry. As in the case of Maturana and Varela, Langton's program is that of a constructive and universal biology, which converges with autopoietic biology not only on the basic epistemological principle of scientific explanation, but also on the principle of biology's universalisation. "Life is (…) a result of organization of matter, rather than something that inheres in the matter itself" (Langton 1989, p. 53).

2.2 Principle 2: Organization ≠ Structure

The second autopoietic principle we consider pertinent for the epistemological grounding of the SA is a theoretical postulate with a significant epistemological value. Its basic content is the distinction between two notions — *organization* and *structure*. Simplifying the original autopoietic formulation (Maturana and Varela 1987, chapter 2), we can put it as follows: the *organization* of a living system is its relational frame, that is, the network of relations which define the system as a unity of components; the *structure* of a living system is its materialization, given by the actual components and their interconnections.

This distinction is not a theoretical novelty introduced by Maturana and Varela. A first complete formulation can be attributed to Jean Piaget (1967, chapter 4), who proposed this conceptual distinction as the theoretical key to comprehend biological systems as dynamical, since it corresponds to the distinction between the invariant and the variant aspects of their dynamics. Piaget remarked that living systems can be considered dynamical systems endowed with a peculiarity: all their elementary components permanently change, while systems, as relational unities of components, remain. This, as

Piaget pointed out, can be affirmed at both the ontogenetic and the phylogenetic levels. The relational unity remains unchanged not only in the permanent flux of physical-chemical components typical of biological organisms, but also during the ontogenetic transformations which can make a living system unrecognisable from one observation to the next. Moreover, this relational unity is transmitted through reproduction and remains unchanged generation after generation. Indeed, this relational unity is the invariant of the biological dynamics and therefore the lowest common denominator of living systems. Distinguishing this invariant relational frame from the changeable materializations of living systems, and determining its configuration, amounts to isolating an element which can be used to define the class of biological systems.

These remarks point out the epistemological relevance of the distinction between organization and structure, which is at least two-fold. Firstly, this distinction allows biological research to hypothesize a defined mechanism for living dynamics (i.e. a mechanism creating organizational invariance through permanent structural variation), and therefore opens the possibility of a constructive explanation of life. Secondly, it generates significant insights about the SA's relevance to the study of natural living and cognitive processes, as it implies that: (a) in principle the materialization (structure) of living systems can be manifold; (b) artificial systems displaying the same organisation as living systems, and realising it in a different structure, have to be considered as belonging to the class of living systems.[5]

Thus, the autopoietic distinction between organisation and structure offers a theoretical ground to the thesis — "the big claim" — through which Langton expresses AL's aspiration: "a properly organized set of artificial primitives carrying out the same functional roles as the bio-molecules in natural living systems will support a process that will be 'alive' in the same way that natural organisms are alive. AL will therefore be genuine life — it will simply be made of different stuff than the life that has evolved here on Earth." (Langton 1989, p. 69)

2.3 Autopoiesis and the extension of the SA to the domain of cognition

Autopoietic biology does not limit itself to formulating principles that support Langton's initial AL program. It also supports the extension of this program to cognitive processes. The intent of "naturalising cognition" led Maturana and Varela to identify living systems as cognitive systems, since their general process of self-production (i.e. autopoiesis) corresponds to a permanent process of interaction with the environment and other systems (structural coupling) that allows living systems to survive. The conceptualisation of this process as a process of cognition is at the basis of the cognitive biology that Maturana and Varela developed as an extension of their theory of life, fathering the nascent "embodied cognitive science" (Clark 1999).According to this view, the phenomenology that has to be produced by the autopoietic synthetic definition of life includes not only all the biological, but also all the cognitive phenomenology *lato sensu* (Maturana and Varela 1987). In this sense, the autopoietic principle of the constructive explanation and the autopoietic distinction between organisation and structure offer a grounding framework not only to the synthetic study of life, but also to the synthetic study of cognition *lato sensu*.

3. Two criteria of relevance for the SA

The two autopoietic principles presented above can be transformed into two criteria for use in determining the relevance of the SA's implementations to the study of life and cognition.

3.1 – "Explaining = Constructing": phenomenological relevance [6]

From the principle of scientific explanation extracted from autopoietic biology's production (*P1: To explain scientifically is to provide a mechanism able to produce the phenomenology to be explained)* can be derived a criterion of "phenomenological relevance" for synthetic models of natural living and cognitive phenomena, according to which:

(C1) A synthetic model is relevant at a phenomenological level if it provides a mechanism which produces (according to explicit parameters) the same phenomenology as the living or cognitive phenomenology under inquiry.

The appellation "phenomenological relevance" expresses the fact that this criterion requires only a relation of identity (defined by explicit parameters) between the phenomenology produced synthetically and the natural phenomenology under inquiry. This means that (C1) does not impose any constraints on the biological plausibility of the synthetic mechanism by which the phenomenology under exploration is produced. Therefore, if (C1) is not correlated to a criterion which requires the biological plausibility of synthetic models, and specifies what this plausibility consists of, then (C1) cannot warrant that these models offer a biologically plausible explanation of the target processes, and that they do not simply imitate the phenomenology under inquiry.

However, from autopoietic biology we can also extract a principle to differentiate phenomenologically relevant models on the basis of their respective operational explanatory powers – that is, on the basis of their capability of providing an operational explanation of the phenomena under inquiry. This principle, belonging to the autopoietic theory of scientific explanation (Maturana and Varela 1987, chapter1; Maturana 1988), associates the operational explanatory power of a model to its "progressive" character, that is, its capability of producing, besides the phenomenology under inquiry, also other phenomena belonging to the same domain.[7] In this sense, autopoietic epistemology provides a principle of evolution to the phenomenologically relevant synthetic modeling of living and cognitive processes, which can orient the choice between different models referred to the same phenomenological domain. That is, on the basis of the quantity of supplementary phenomena, that they are able to produce, these models can be considered more or less progressive (that is, more or less operationally explanatory) than another models.

5 The thesis of the multiple material realization of organization implies a convergence between autopoiesis and functionalism. There is no room here for a detailed comparison. However it is worth noticing that autopoiesis and functionalism have different views about the implications of this thesis. For example, autopoiesis, differently from functionalism, emphasizes the dependence of cognition on the agent's embodiment.

6 Here the adjective 'phenomenological' has the meaning of 'relative to the phenomenology under inquiry'.

7 We use the term 'progressive' in accordance with Lakatosian philosophy.

According to this principle, we have to distinguish two basic kinds of phenomenologically relevant models. They can be respectively defined as follows:
(1) *minimal phenomenological models*, which produce only the phenomenology under inquiry, and therefore have a minimal operational explanatory power;
(2) *progressive phenomenological models*, which produce, besides the phenomenology under inquiry, other phenomena belonging to the same domain, and have an operational explanatory power proportioned to the quantity of supplementary phenomena produced.
Evolution towards better phenomenologically relevant models corresponds to evolution towards models endowed with a higher operational explanatory power, but not necessarily towards biologically plausible models. Indeed, even if a greater operational explanatory power could be considered as a clue of greater biological plausibility, this last remains uncertain in absence of a criterion which specifies what this plausibility consists of. A synthetic progressive model, in itself, could be useful for the traditional scientific exploration of living and cognitive processes as it could offer not a biologically plausible explanation of these processes, but a source of inspiration for the production of hypotheses about the mechanisms underlying them.

3.2 "Organization ≠ Structure": relevance in the strong sense

As pointed out before, the autopoietic distinction between organization and structure implies that *(P2.i) All living and cognitive systems share the same organisation, but not necessarily the same structure,* and therefore that (*P2.ii) Artificial systems which display a different structure, but the same organisation as living and cognitive systems, have to be considered legitimately belonging to the class of living systems.*
Thus, the autopoietic distinction between organization and structure produces a criterion of relevance for synthetic models of living and cognitive systems. In fact, in accordance with (P2.ii), the former can be considered *strong models* of the latter if they share the same organisation, since, in this case, they constitute *specimens* of the class of living and cognitive systems.
We can refer to this criterion as to a criterion of *organisational relevance*, which warrants the biological plausibility of synthetic models. Associated to the criterion of *phenomenological relevance*, it produces the criterion of "*relevance in the strong sense*":
(C2) Synthetic models are relevant in the strong sense if, besides providing mechanisms which generate the phenomenology under inquiry (phenomenological relevance), they present (according to some explicit theory of living and/or cognitive organisation) the same organisation as living and cognitive systems.
Satisfying this criterion is indeed a hard challenge for AL, which, of course, always has to be faced referencing one or more theories of biological and/or cognitive organisation, and always in an approximate way due to the intrinsic limits of these theories, the varieties of their interpretations, and the limited possibilities of their implementation. In this sense, relevance in the proper sense has to be considered for artificial life more a regulative ideal than a concretely attainable goal.

4. Interactive phenomenological relevance

If attachment phenomenology is defined as the closed set of phenomena normally used to exemplify it (e.g. seeking the proximity of the caregiver, developing stress in situation of separation and developing different attachment profiles depending on caregiver behaviour), then the robot model can be considered to (roughly) satisfy (C1). But, as far as we tested it, this model does not have a progressive character, and cannot be considered biologically plausible according to (C2). Therefore we are led to attribute it a *minimal phenomenological relevance* with regard to attachment behaviours, and to consider it as simply imitating them.
However the robot model does more than this when the system under consideration is the *human-robot interacting dyad*. Using this model in experiments involving humans in the role of caregivers, the resulting evidence suggests that it has further scientific potential, related not to its operational explanatory power or its biological plausibility, but to its capability of dynamically interacting with human agents. In fact, the "baby" robot appears able to engage humans in *interactive dynamics* which can be of scientific interest for the developmental psychology inquiry on attachment bonds (Hiolle *et al*. 2008). That is, it offers to developmental psychology the possibility of experimentally manipulating and exploring, in human agents, aspects of the attachment phenomenology that can be difficulty accessible in the classical psychological scenarios of research. An example of these processes can be found in human caregivers' reactions to different attachment profiles. This is an aspect of the attachment phenomenology that developmental psychology could study through robot models like the one presented above, as emerged from our interdisciplinary exploration of attachment bonds.
These remarks lead us to introduce a new kind of minimal phenomenological model. These can be defined as *interactive phenomenological models*: models able to synthetically produce the phenomenology under inquiry, and, through the expression of this phenomenology, to engage natural biological and/or cognitive systems in interactive dynamics which (according to some explicit parameter) prove interesting for the scientific exploration of the natural phenomenology under inquiry.[8] As such, this kind of minimal phenomenological model can concretely contribute to biological, behavioural and cognitive science's inquiries on natural living and/or cognitive processes, as synthetic tools that, through their capability of interacting with natural living and/or cognitive systems, can support the experimental manipulation and investigation of their processes.

5. Conclusions

This article proposes a constructivist solution to the issue of providing epistemological groundings for the application of the SA to affective and social processes. This solution consists of an epistemological framework extracted by autopoietic epistemology, able to provide to both the synthetic study of life and the synthetic study of cognition with (1) shared grounding principles of knowledge, and (2) shared general criteria useful to define the relevance of (soft, hard, and wetware) synthetic

[8] Soft AL's production is rich of examples of virtual agents interacting with human agents, and able to engage them in interactive dynamics which could be interesting from a scientific point of view. Examples of synthetic systems able to interact with natural systems are emerging also in wet AL (cfr. e.g. Kaneda *et al*. 2009 about interactions between synthetic models of minimal cells and cultured cells).

models for the exploration of life and cognition. We showed how these criteria open a space of relevance for the synthetic modelling of life and cognition defined by two extremes. The "lower" extreme is *minimal phenomenological relevance*, which characterises models that, by reproducing synthetically the natural phenomenology under inquiry, offer an operational explanation of these process, but, as they do not have biological plausibility, have to be considered synthetic imitations of them. The "upper" extreme, which has to be considered more a regulative ideal, than a concretely attainable goal, is (2) *relevance in the strong sense.* It characterises models that reproduce synthetically the natural phenomenology under inquiry, and, as they display the same organisation as living and/or cognitive systems, can be considered to belong to the class of living and/or cognitive systems. We argued that, within this space, AL can produce two kinds of synthetic models that could be of interest for its interdisciplinary cooperation with biological, behavioural and cognitive science. The first is given by models characterised by a *progressive phenomenological relevance*, that is, the capability of producing not only the phenomenology under inquiry, but also other phenomena belonging to the same domain. These models have a significant operational explanatory power, and, dependent on their biological plausibility, can prove useful for biological, behavioural or cognitive sciences as a source of inspiration for the definition of the mechanism underlying the phenomenology under inquiry. The second kind is given by models characterised by a an *interactive phenomenological relevance*, that is, the capability of producing synthetically the phenomena under inquiry, and, through this, engaging natural systems in interactive dynamics that prove useful to experimentally investigate, in natural systems, at least some aspects of the phenomenology under inquiry.

Acknowledgments
This work has received support from the EC-funded Projects FEELIX GROWING (FP6-IST-i) and ALIZ-E (FP7-ICT-248116).

REFERENCES

Anderson, K. J. (1990). Arousal and the inverted-U hypothesis. *Psychological Bulletin* 107, 1:96 – 100.
Andry, P., Garnault, N. and Gaussier, P. (2009). Using the interaction rhythm to build an internal reinforcement signal. In *Proceedings of the Ninth Int. Conf.on Epigenetic Robotics*. Lund University Cognitive Studies.
Bedau, M. (2003). Artificial life*:* organization, adaptation and complexity from the bottom up. *Trends in cognitive science*, 7:505-512.
Berlyne, D. E (1960). *Conflict, Arousal and Curiosity*. Mc Graw-Hill Book.
Berthouze, L. and Lungarella, M. (2004). Motor skill acquisition under environmental perturbations. *Adaptive Behavior* 12: 47–63.
Blanchard, A. and Cañamero, L. (2005). Using visual velocity detection to achieve synchronization in imitation. In *Third Intl. Symposium on Imitation in Animals and Artifacts*, SSAISB Press, 26–29.
Bowlby, J. (1969). *Attachment and loss. Vol 1: Attachment.* Basic Books.
Breazeal, C. and Scassellati, B. (2002). Challenges in building robots that imitate people. In: Dautenhahn, K. and Nehaniv, C. editors, *Imitation in Animals and Artifacts*. MIT Press, Cambridge, MA. 14, 363–390.
Brazelton ,T. B. and Nugent, J. (1995). *Neonatal Behavioral Assessment Scale*. CambridgeUniversity Press.
Cañamero,, L. (2005). Emotion understanding from the perspective of autonomous robots research. *Neural networks*, 18:445-455.
Cañamero L. (2002). Playing the emotion game with Feelix. What can a Lego robot tell us about emotion? In: Dautenhahn, K., Bond, A.H., Cañamero, L., Edmonds, B., *Socially Intelligent Agents*. Kluwer
Cañamero, L. and R. Aylett R. editors (2008). *Expressive characters for social interaction*. John Benjamins, London.
Cañamero, L., Blanchard A., and Nadel J. (2006). Attachment bonds for human-like robots. *International Journal of Humanoid Robotics*, 3:301-320.
Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Science*, 3, 9:345-351.
Davey, N. and Adams R. (2004). High capacity associative memories and connection constraints. *Connection Science* 16, 1, 47–65.
Damiano, L. and Cañamero L. (2010). Constructing Emotions. *Proceedings of AIIB 2010*.
Dawson, M. R. W. (2002). From embodied cognitive science to synthetic psychology. In: *Proceedings of the First IEEE International*.
Dawson, M. R. W. (2004). *Minds and machines*. Blackwell, London.
Froese, T and Ziemke T. (2009). Enactive artificial intelligence. *Artificial Intelligence*, 173: 466–500.
Glasersfeld, von, E. (1995). *Radical constructivism*. The Falmer Press. London.
Hiolle, A., Bard, K. and Cañamero, L. (2006). Assessing human responses to different robot attachment profiles. *Proceedings of the 18th International Symposium RHIC*, pp. 251-257.
Hiolle, A. and Cañamero L. (2008). Why should you care? An arousal-based model of exploratory behavior for autonomous robots. In *Proceedings of the 11th International Conference on Artificial Life*, 242-248, MIT.
Hiolle, A. and Cañamero, L. (2007). Developing sensorimotor associations through attachment bonds. *In Proceedings of the 7th International Workshop on Epigenetic Robotics,* 45–52.
Hiolle, A., Cañamero, L. and Blanchard, A. (2007). Learning to interact with the caretaker: a developemental approach, *In Proceedings of the 2nd ICACII*, 425-–436.
Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America* 79, 8, 2554.
Kohonen, T. (1997). *Self-Organizating Maps*. Springer-Verlag.
Kaneda, M. et el. (2009). Direct formation of proteo-liposomes by in vitro synthesis and cellular cytosolic delivery with connexin-expressing liposomes, *Biomaterials*, 30: 3971–3977.
Keller, H. (2008). Attachment-Past and present. But what about the future? *Integr Psych Beav* 42: 406-415.
Langton, C. (1989). Artificial life. In: Boden M editor (1996), *The Philosophy of Artifcial Life*. Oxford University Press, 39-94.
Luisi, P. L. (2010). *The synthetic approach in biology*. Manuscript.
Lungarella, M., Metta, G., Pfeifer, R. and Sandini, G. (2004). Developmental robotics: a survey. *Connection Science*, 1–40.
Maturana, H. and Varela, F. (1987). *The Three of Knowledge*. Shimbhala.
Nunez, R. and Freeman, W. J. editors (1999). *Reclaming cognition*. Imprint.
Oudeyer, P.-Y., Kaplan, F. and Hafner, V. (2007). Intrinsic motivation systems for autonomous mental development. IEEE Transactions on Evolutionary Computation 2, 11.
Piaget, J. (1967). *Biologie et connaissance*. Gallimard, Paris.
Pfeifer, R., Lungarella, M. and Sporns, O. (2008). The synthetic approach to embodied cognition. In Calvo P. and Gomila T. editors. *Handbook of cognitive science. An embodied approach.* Elsevier.
Pfeifer, R. and Scheier, C. (2000). *Understanding intelligence*. MIT Press, Cambridge, MA.
Riegler, A. (1992). Constructivist Artificial Life, and Beyond. In: McMullin, B. editor. Proceedings of the Workshop on Autopoiesis and Perception, *Dublin City University: Dublin, pp. 121–136.*
Russel, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology* 39: 1161–1178.
Sroufe, L. A. (1996). *Emotional Development: The Organization of Emotional Life in the EarlyYears*. Cambridge University Press.
Tronick, E. (2007). *The neurobehavioural and socio-emotional development of infants and children*. Northon & Company.
Webb, B. (2001). Can robots make good models of biological behaviour? *Behavioral and Brain Sciences*, 24: 1033-1050.