# Artificial Evolution: creativity and the possible

**Arantza Etxeberria**
Logika eta Zientziaren Filosofia Saila (Dept. of Logic and Philosophy of Science)
Euskal Herriko Unibertsitatea (University of the Basque Country)
1249 Posta Kutxa, 20.080 Donostia-San Sebastian, Spain
ylpetaga@sf.ehu.es

### Abstract
In this paper the aims and goals of artificial evolution are discussed in relation to two of the founding features of Alife: how to characterize the domain of the possible and the criterion of lifelikeness. It is argued that artificial evolution should aim to understand the evolution of organizations and that this will bring about a better understanding of possible evolutions.

## Introduction

Artificial Life (Alife) emerged as a research programme to study living phenomena in artificial media so as to exempt the notion of life from its dependence on a single example. Three main features characterize the goals of this science as it was conceived by Langton (1989). First, its object is to explore the domain between the existent life-as-we-know-it and the possible life-as-it could-be. In relation to biology, which studies life on Earth, Alife is a science of universal living phenomena. Second, this universality will be achieved by studying artificially generated invariants of form or organization, and not of matter. The systems under study are synthetized, instead of (or, ideally, before) analysing them, and it is assumed that the relevant phenomena have more to do with the organization of the parts than with the nature of the components. Third, the extension of the systems under inquiry is open, but characterized by the notion of lifelikeness; thus, the question of defining life is left aside and substituted by a criterion based on experience and intuition. Research may emulate different aspects: complex organization, stability in a given medium, capacity to absorb, transform and use energy, self-reproduction, evolution, development and growth, adaptation, etc. and those aspects will be interesting according to how lifelike they are. All the three –the domain of the possible, universal organizing principles independent of matter, and lifelikeness– are interesting start points, but deserve careful discussion.

In previous work several authors have pointed out the problems of analyzing universality by means of organizations that do not emerge from a realistic material dynamics (Pattee 1989, 1995; Cariani 1992, Moreno et al 1994). In this paper I intend to discuss artificial evolution in relation to the other two: how to characterize the domain of the possible and the criterion of lifelikeness.

The distance between existent life and possible life is manifest in that different possible evolutions could have taken place on Earth starting from the same or similar initial conditions, because the history of evolution contains both contingent, or fortuitous, events and necessary ones, determined by the properties of evolving matter. An interesting line of research for a discipline interested on the universal aspects of possible life is then to tell the necessary and the contingent of the history of evolution. In this sense, and presuming that laws are universal, the domain of the possible has, at least, two orders of magnitude: starting with the same initial conditions would give us the scope of non-realized possibilities of life on Earth, whereas different initial conditions would generally describe the scope of other possible lives. This would be the counterfactual approach to evolution.

However, another way of considering the domains of possibility that artificial evolution may unfold has to do with evolutionary theory itself and the way it conceives creativity or production of novelty. The theory of evolution is epistemologically challenging because it introduces creativity in the realm of science. Unlike other fields, like physics, that try to discover the laws of nature underlying the behavior of all systems, evolutionary theory may be read as saying that almost anything is possible, because it describes a procedure by which novelty appears and develops in nature.

It is in this sense that it is important to find a criterion of lifelikeness as a foundation for research in artificial evolution. The creativity and open-ended nature of life has been characterized as "supple adaptation", as a hypothesis that life can be defined in terms of a system that exhibits "lifelike" evolution (Bedau 1998). This is an interesting idea, but somehow developed at the cost of renouncing to consider lifelikeness at the level of individual living beings and preferring, instead, to focus on the whole of life as a process. In this paper my interest is directed to the problem of how to conceive the evolution of embodied agents, a perspective that obliges to take into account the nature of living organization and its relation to evolution. An adequate understanding of the evolution of organizations is still lacking: evolution and organization are difficult terms or perspectives to bring together. This is probably the reason why researchers who have worked on the problem of biological organization, like Varela and Rosen, have somehow left the problem of evolution aside as secondary. Rosen, for example says: "We cannot answer the question (...) "Why is a machine alive? with the answer "Because its ancestors were alive". Pedigrees, lineages, genealogies and the like, are quite irrelevant to the basic question. Ever more insistently over the past

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022

century, and never more so than today, we hear the argument that biology *is* evolution; that living systems instantiate evolutionary processes rather than life; and ironically, that these processes are devoid of entailment, immune to natural law, and hence outside of science completely. To me it is easy to conceive of life, and hence biology, without evolution." (Rosen, 1991, pp. 254-55). A similar feeling may be found in Varela (1979): "I maintain that evolutionary thought, through its emphasis on diversity, reproduction, and the species in order to explain the dynamics of change, has obscured the necessity of looking at the autonomous nature of living units for the understanding of biological phenomenology. Also I think that the maintenance of identity and the invariance of defining relations in the living unities are at the base of all possible ontogenic and evolutionary transformation in biological systems" (p. 5). Although these positions might sound extreme, they reveal that there is a real problem for artificial evolution to be a process of evolving organizations, given the use of evolutionary theory done by many researchers. Thus, a perspective like the one I want to suggest requires to work on what I call diachronic embodiment. This would constitute an original contribution of Alife beyond actual biology.

## Three perspectives of artificial evolution

Artificial Evolution has been explored according to different interests by the participants in this interdisciplinary research area. Very roughly we could distinguish three main tendencies of research:

- In a first one evolutionary theory is used as a problem solving strategy, convenient for certain purposes, but which does not require any biological plausibility as it is not being applied to generate life or lifelike phenomena. As the main goal of this case is the efficiency of the method used to achieve certain practical goals, both theoretical and empirical knowledge can be violated (lamarckian evolution or very high mutation rates) in order to achieve certain goals such as optimization.
- A second group uses artificial models to study biological phenomena for which no adequate model or theory has been developed so far. The aim in this case is to develop modelling strategies to enhance our capabilities of explaining biological phenomena, by developing theory or analytical tools based on artificial systems.
- A third division is constituted by those artificial systems in which the frontiers or boundaries between the artificial and the biological are diffuse, and instances of alternative universes are constructed that, although do not correspond literally with the facts of real life, provide an intuitive grasp of lifelikeness.

These three categories correspond roughly to the production of three different *things*: tools, models and instantiations (Etxeberria 1995). *Tools* are clearly methodological innovations that may or may not apply to life, although inspired by it; whereas models and instantiations are distinguished by their respective different relation to a target real phenomenon to be explained or grasped. The term *model* is reserved for the case usually developed by the natural sciences: as a way of representing certain relevant aspects of an empirical domain by means of data collected through measurements. In this case the natural systems are, so to say, the referents of the artificial models built to explain or emulate them. This reference relation is looser in *instantiations*, more related to the way models are conceived in formal sciences: as an instance or a system built after a set of axioms. As they do not have a concrete empirical referent, their value resides in their being new creations of ontology that, although inspired by natural phenomena, provide as an end result something different from life on Earth, but with a strong lifelikeness.

The three kinds of artificial systems have different goals. Tools enhance our epistemological capacities in the sense that they provide information processing ways that, although inspired by the biological theory, do not operate in a transparent way for the human mind. They are directed to fulfil human or technological goals. Models may allow us to study the distance between realized and non-realized life, the conditions under which other possibilities of our life could have happened and counterfactuals based on different parameters. They may provide theoretical contributions if the data were set to run the appropriate experiments. Instantiations may be much more abstract and their only connection with reality depends on the way we apply the theory we have. Although it may be the case that instantiations will provide the most radical ways of exploring life-as-it-could-be, present day work is strongly limited by its dependence on difficult to overcome theoretical assumptions.

## Intuitive notions of evolution and evolutionary theory

Models and instantiations are, in principle, more likely to produce a notion of lifelike artificial evolution than tools. At the moment most of the research in artificial evolution is based either on an intuition or on a theory, or on both. The intuition is that evolution, whatever it is, must reveal itself as the growth of something: complexity, adaptation, optimization, or some other global property. Even if this view of evolution as growth tends to be rejected by the latest evolutionary theory, where the usual perspective is neutral with respect to this, it remains an important guiding notion for a science that has to provide artificial examples of something so poorly understood as evolution. Evolutionary Biology considers that evolution takes place when there is a change in the gene frequencies of a population (the Hardy-Weinberg Law), change driven by different "evolutionary forces": natural selection, sexual selection, genetic drift, etc. This notion of evolution assumes a fairly neutral situation with respect to what is expected from evolution, whether there is or not an increase in complexity in evolution, or whether evolution has a direction. Most artificial evolution models are not that neutral.

The motivation of researchers trying to study or to

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022

implement artificial evolution might in fact preclude that neutrality. For instance, when von Neumann (1966) studied the problem of artificial self-reproduction, he took into account the possibility of an increase of complexity of the reproduced system. In his view, the capacity living systems have to evolve derives from the possibility that they, unlike machines, have to produce other systems which are of equal or superior complexity; and many of the constraints he imposed to the self-reproducing logic, like the necessity of a universal constructor, where implied by this conception of open-ended evolution. Similarly, evolutionary computation, as a general problem solving tool, assumes that this methodology will bring along some sort of "improvement" or optimisation of the proposed solutions.

Thus, Alife experimenters trying to develop artificial evolution face the disjunctive option of either providing an external goal to their systems, or to have no natural "goal" at all. Many artificial systems are indeed supplied with an external goal: this is the case of most tools, and also of experiments of artificial evolution that explore a kind of creativity which is in fact evaluated by humans (an example could be art created by this procedure). However the case of an artificial evolution system that wants to be true to nature or lifelikeness is more tricky: on the one hand, the evolving system is not supposed to have any intrinsic direction and, on the other, it is assumed that under the appropriate evolutionary dynamics (and the appropriate encoding of the problem) anything is possible. The reason for considering that anything is possible is that the kind of abstraction of natural evolution used by artificial evolution must be able to derive any interesting system or organization as its evolution under the appropriate variety and the selection pressure. This is why in general a procedure of artificial evolution, such as the genetic algorithm, gives the impression that anything is possible. Yet, a situation in which anything is possible is equivalent to one in which nothing interesting happens. The problem is that this is not the case of natural evolution and the reason for it is that there are intrinsic trends that have not been sufficiently investigated by artificial systems.

This general idea can be expressed in another way by saying that the evolution should be based on synchronically complex or organized systems. The study of complex systems presents a dichotomy of methodologies: the approach taken is either synchronic or diachronic. The synchronic one (sometimes also called *vertical or emergent*) studies the relation of components with the aggregates or totalities they form, the target is complexity or interesting behaviours at a global level, starting from simple components or interactions at a lower level. It is a bottom-up perspective, largely inspired by Thermodynamics. Examples of this approach are auto-catalytic sets, swarm organisations, neural nets and embodied robotics. The diachronic one (also *horizontal or transformational*) studies how complex systems arise and develop in time, as a substitution of subsequent generations. The system is considered in terms of the appearance of novelty through time, and not in relation to levels, selection being the main mechanism or force acting in the system. Work in artificial evolution, like the genetic algorithm, is an example of this.

Hence, in many cases artificial evolution seems to rely in an intuitive idea of evolution, instead of trying to develop a new, productive notion to explore what evolutionary change would be like for artificial systems. This kind of new notion will only be developed by an appropriate notion of the organizations –as opposed to bit strings– able to evolve.

## Organization, functionality, and design

The lack of a well developed notion of diachronic embodiment is evident in the treatment of functionality. The functional analysis of complex systems is a hard issue for biology and Alife and this can be explained, even if in a sketchy way, by referring to the difference between two traditions of thought, the Kantian and the Darwinian, through the way each of them compares the living organism with a watch.

For the Darwinian tradition this comparison poses the problem of the "argument from design" (developed, among others, by Aquinas, Hume and Paley). Paley said that if in crossing a heath I found a watch I will not think that (like a stone) it had lain there for ever, but the inference will be that the watch must have had a maker: "Arrangement, disposition of parts, subserviency of means to an end, relation of instruments to a use, imply the presence of intelligence and mind" (Paley, in Ruse 1998, p. 38). In the same way, the design of organisms takes us to necessarily accept the existence of a creator: "every manifestation of design which existed in the watch, exists in the works of nature, with the difference on the side of nature of being greater and more, and that in a degree that exceeds computation" (idem, p. 39). When the Darwinian tradition responds to this argument, a natural explanation of design is produced that requires no divine intervention –the principle of natural selection–, but it lets a likening of the watch and the organism in, maybe accepting it as a good one.

Kant had already used the same comparison in the *Critique of judgement*, but in a rather different way. He notices a fundamental difference between the two of them, whereas the watch is formed by fixed components, fabricated before hand and later ensembled, in the organism the parts are formed for the others, some parts produce the others. Kant accepts an internal teleology in the living system.

Many authors (for example, Mayr) have said that the later development of the theory of evolution corrects the Kantian summon to teleology and makes it possible to explain this in another way. Maybe this is true, but Kant points to a problem that the Darwinian tradition has not collected: the relation among the parts to form an organization. Actually, like in the case of the watch, for the Darwinian tradition the assimilation of watch and organism is not problematic, whereas the Kantian tradition feels that distinction must be set and explained.

This distinction implies that there is a difference between explaining the function of components in relation to organization or to adaptation, either in the historical or ahistorical way, like it is usual in the

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022

evolutionary biology. However, both traditions coincide in considering that the problem of organization must be approached in an analytical way that derives the whole from the properties –decomposable or not– of the parts. The question is whether this should continue being this way, whether it would not be much better to produce a complementary "synthetic" way of looking at this problem.

This is the reason why embodiment –both in a synchronic and a diachronic perspective– is such an important issue for Alife. Now the field of "embodied cognition" seeks to understand how physical properties contribute to the processing of information needed for the interactions involved in the autonomous behaviour of robots or artificial agents in general (Mataric 1997). The hint is that when the physical properties of bodies (shape, orientation, degrees of freedom, etc.) are taken into account, a great deal of the explicit information processing becomes superfluous. Probably taking into account more aspects, like metabolism, would induce more radical changes (Moreno & Ruiz-Mirazo 1999). Traditionally this phenomenon could be conceived as an imperfection, as some sort of handicap or constraint that limits information processing capabilities, but embodiment precludes the search for "clean" functions and this has serious consequences for the task of designing embodied agents. In other words, if evolution were a process of creation of functionalities (that is to say, a process in which matter is shaped to produce –and reproduce– certain functions), in most cases these would be non-detachable properties of the physical substratum.

The problem of functions brings us directly to the one of design. Every theory of evolution attempts to provide natural mechanisms for the origins and diversification of living forms, so that no designing agent has to be proposed. Evolution is intuitively different from design in its non-intervention, evolutionary explanations are naturalistic because form or function emerge from matter and interactions. Naturalistic explanations of design are, nevertheless, difficult; even if it is agreed that the process of evolution is not designed and, moreover, that it does not itself design. For machines, the case is easier, but the aim of Alife is to produce lifelikeness. Let's turn for a moment to Polanyi´s ideas about machines. He described them as working under two distinct principles: a higher one of design, and a lower one harnessed by the former, which consists in the physical-chemical processes on which the machine relies (Polanyi 1968). This analysis may hold for living beings as well, but, for them, the higher level of design (at any scalar level) is not easy to characterise. Then the attempt to substitute design for artificial evolution in the construction of complex machines, such as adaptive robots, should imply to start with no functions, to let functions emerge.

## Artificial evolution and natural selection

In Alife artificial evolution has been often used to build artificial creatures with features adapted to their artificial environments by a procedure similar to evolution by natural selection (for example, the genetic algorithm). The methodologies developed with this inspiration generate (at random or hand crafted) a population of bit strings which are interpreted as possible solutions for the problem; the "fitness" of each of them is scored using an evaluation function; there is a process of selection of individuals according to these scores (usually the best are taken to be the "parents" of the subsequent generation, but, sometimes, samples of medium or even bad individuals are also preserved); and individuals are modified using the genetic operators (mutation and/or recombination) to produce a new generation. This is basically the genetic algorithm; by letting the system "evolve" in this way, the population usually converges into a situation of mutual environment/creatures adaptation. In "artificial worlds" this kind of artificial evolution governs the change of certain features of simulated agents and environment, it introduces a higher scale temporal change than the one of processes taking place at the scale of the life time of individuals, for example, learning. In the most interesting cases, the interaction between both scales produces interesting phenomena. In "evolutionary robotics" this method has been conceived either as an alternative to the explicit design of the morphology and/or the cognitive systems of artificial creatures, or as a design methodology (there is some ambiguity about which is the case).

In the first kind of work, the artificial world is usually a two-dimensional array of cells (though it can be n-dimensional, and, sometimes, continuous) where a population of agents coexists with similar ones and other living (predators, parasites, etc.) or/and non-living systems (food, shelter, geographical barriers, etc.). Agents are typically represented by a control system for behaviour specification (such as a neural network, a finite state machine, or an abstract grammar), and, very seldom, with a simulated physical body. They are endowed with a sensorimotor apparatus to perceive the relevant features of the world and with a set of behaviours to act in it. The sensorimotor system is, in general, as complex as the world: sometimes both are relatively realistic, while, in others, perception consists merely in a detection of the state of the nearby cells. The metabolism of agents is usually represented as an energy storage, whose level determines whether they are in good shape to perform several actions (mate, escape, etc.) or close to death. The generational replacement will result in a modification of the control system (and/or the body) of the agents: like in evolution by natural selection, agents will be able to survive and reproduce according to their abilities in the world they inhabit.

Most of these models try to simulate the relevant features considered by evolutionary models (morphological, functional or behavioural traits that influence the fitness of individuals, size of populations, selection pressures, etc.) to be able to set up experiments. These are rather close to the phenomena studied by population genetics. The problem is that the change (that is to say, the evolution) likely to arise in simulations is, in general, already known from the theoretical results and the creativity or exploration of potential novelty expected from an evolutionary process is generally lacking.

These works bring about little new about interesting

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022

ways in which evolution can influence complex (vertical) organizations. Within the field of Artificial Life the main criticisms they have received have to do with their being simulations. As such, agents cannot present the physical organization that makes it possible to study embodied cognition in the simulated domain, therefore research should be directed towards physical realizations. I think this criticism is correct, but insufficient. The problem of the kind of artificial evolution has been overlooked, and it is as important as realism at the time of implementing the dynamics of cognitive interactions.

In the work on robot evolution, the evaluation function depends on how well the robot behaves in the real world, instead of in a simulation (Cliff, Harvey, & Husbands 1993). Because this methodology requires a combination of simulations (encoding of the evolving structure, genetic operators applied to selected individuals) and physical building (performance, test and scores), different technical problems arise in transferring the evolved individuals to the robot for evaluation. The rationale for adopting the methodology of artificial evolution is related to the problems encountered with an external functional decomposition and the complexity and big number of interactions among sub-systems, which sometimes are even mediated through the environment. Thus, agents whose interaction with the environment is very sophisticated need also a very complex sensorimotor system that is difficult for hand design. Artificial evolution will be used as the search method and also the heuristic to handle the intractable complexity that up to certain levels arises in other hand design incremental procedures (Brooks 1986, Harvey 1996). There is an in principle interest in avoiding functional decomposition and to evolve whole architectures that can adopt non decomposable functions within the system. This is, also, linked with the ideal goal of obtaining via evolution an incremental complexification of the behavioural capacities of the robots in a non-foreseen way. Nevertheless, the separation of the diachronic and synchronic aspects of the process (performed respectively in simulation and realisation), and the correspondent diachronic disembodiment, demand further investigation of artificial evolution.

If the artificial evolution procedure applied in these systems is analysed, then it is clear that getting rid of the task of designing is not that easy after all, many problems arise as a difficulty not to intervene in the process from the outside. Questions like how the genotypes of the population are encoded (and decoded), the relation between evolving populations and their environments, and how the artificial models exhibit "natural" selection make apparent that there is a fundamental lack of understanding of how a spontaneous, not externally directed evolutionary process takes place.

The first problem, how to encode the evolving traits of agents in the genotype (only the encoded ones are affected by the GA) is still poorly understood in most work. The genotype-phenotype relation tends to be very simple (in most cases, a direct mapping) and, normally, only a few genes are allowed to evolve and in a fixed-length genotype. Besides, genes evolve individually, there is no epistatic interaction between them, and this reveals what we may call a "diachronic disembodiment" of the evolved structures. There is a growing awareness that the process of development has consequences for the way genotypes evolve, and there have been interesting attempts to overcome the difficulties (for example, using variable length genotypes and, more important, developing more complex ways to represent the genotype-phenotype mapping, so that some aspects of development are taken into account). Nevertheless important background questions are still underdeveloped. Von Neumann, in his work on self-reproduction, already said that only systems endowed with a self-description were capable of open-ended evolution. It was important to establish this threshold, but it still requires deeper understanding in the light of new research. In fact, the way von Neumann conceived the relation between the description and the constructed system is not satisfactory (Etxeberria & Ibañez 1999) and should be a matter of concern and study for artificial evolution.

The second problem, the relation between the agents and the environment as a common history, is also difficult to implement, in such a way that conditions appear in which agents and environment inform one another and intervene in the change of the other (Lewontin 1983). Simulated environments are too simple, fixed and external, and this makes it difficult to observe the action of agents to produce their own environments. Yet, even the real environments in which robots operate are usually very controlled and not at all constructed by the creatures themselves.

Finally, it is very difficult to model the conditions in which selection can arise from the characteristics of the whole system, that is to say, to define the possible adaptive landscapes and intrinsic evaluation or fitness functions according to them. In fact, some models use straightforward (human) artificial selection, using theoretical or aesthetic criteria for it and others define conditions that will be automatically evaluated, but with an absolute criterion, that is to say, they pre-define a "perfection" to be achieved. Natural selection in artificial evolution is external and designed, when a simple evaluation function scores individuals of the population. But it is not a lot more "natural" when creatures live, reproduce or die in this world according to their abilities, because the poor genotype-phenotype encoding and the simple agent-environment relation makes it rather obvious which of the parameters will be that define the best "fitness" in that world. An interesting solution for this has been sought in co-evolution, so trying that the fitness of a population becomes evaluated in relation to changing conditions (Hillis 1992, Sims 1994).

For evolutionary biology, artificial selection is the process of change induced in a population by a human agent selecting something for some purpose. Natural selection, in turn, is blind, and even if some can consider it a process that produces adaptations, or even further, design, these are not foreseen in advance: they depend on many changing conditions that shape the interaction between the evolving entity and the environment. For artificial evolution the situation is confusing because the concept of natural selection is, in some ways,

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022

epistemologically very close to that of artificial selection. In fact, the metaphor of artificial selection in breeding was used by Darwin himself to propose natural selection as a explanatory mechanism for evolution. Perhaps the word "selection" cannot but suggest a selective agent, or it might be that natural selection needs further examples (as suggested, for example, by Depew & Weber 1995).

## Artificial Evolution´s contribution to science

Some issues presented in the last section are a warning about the achievements of artificial methodologies whose main criterion is biological plausibility reduced to a standard textbook theory. This is an important problem for Artificial Life because, since organisms are a clear example of complex systems and the phenomena under study are complex, the obvious way to tackle it appears to appeal for biological plausibility. However, an excessive tribute to the biological might be a handicap, because to obtain insight of some phenomena being studied it is important to acknowledge the characteristics of the medium in which the model is produced. For example, embodiment, emergence, or complexity itself, are difficult notions to understand and copying biology is not necessarily the best way. In fact, the strategies of biology and cognitive science are often reductionist and if research in the artificial domain is expected to overcome some of the problems encountered by these approaches, then it should look for other ways.

In biology there are theoretical problems to integrate two apparently opposed perspectives on organisms, the main being the functional and the structural. The notion of self-organisation is problematic for the standard theory of evolution which has problems to elaborate systemic points of view. Self-organisation, unlike the notion of natural selection, can be characterised as a *systemic* property, in the sense that the entity that organises itself is composed of parts whose configuration and interaction determine the whole they form (which cannot be reduced to those parts). It is also a property that has to do with the production of *spontaneous order* in the system, which is a result of the dynamics of interaction among the components of the system and has nothing to do with an external organising agent or with external design. Finally, it is a capacity that expands the explanatory domain of classical Physics and Chemistry; hence, even if it is not a uniquely biological (it also appears in inanimate systems), self-organization provides a view of nature that suggests a *continuity between the inanimate and the animate*. These characteristics contrast with the atomistic, externalist and non-physicalist perspective of evolution by natural selection.

Then, one possibility for artificial evolution is to explore the integration of both perspectives. This way, it would do more than provide computational versions of the analytical models already developed, it would contribute to the unsolved problems of the field. In fact, artificial models of biological systems have already contributed before; for example auto-catalytic sets (Farmer et al. 1986, Kauffman 1986), Turing's reaction-diffusion model of morphogenesis, or models of Random Boolean Automata have been relevant to reconsider the biological theory or at least to open interesting debates within it. An application of the new ideas proposed in evolutionary biology to the domain of the artificial is important in order to expand the often too narrow application of evolutionary principles found in some artificial models of evolution, like for example, in genetic algorithms.

An idea that is particularly relevant is the notion of *developmental constraint* as a "bias on the production of variant phenotypes or a limitation on phenotypic variability caused by the structure, characteristics, composition, or dynamics of the developmental system" (Maynard Smith et al 1985, p. 266). The possible sources of those constraints are varied, but they all introduce intrinsic limits to the action of natural selection, both universal or local (only in certain taxa). If developmental constraints can be characterised, then they contrast with ideals of perfection: natural selection is not a mechanism of unlimited optimisation, evolution does not produce "perfect design". Thus, developmental constraints are material limits to perfection.

Among the universal developmental constraints, those studied by Kauffman (1993) are conceived as generic properties of organised matter which do not depend upon natural selection, even if they could influence the conditions in which it takes place. These generic properties are based on the capacity for spontaneous order upon which natural selection acts. Self-organisation is previous to the constitution of the system itself, it prepares the conditions in which natural selection can take place. Kauffman has constructed mathematical models of genetic regulatory systems, as logical networks of connections. By changing the connectivity parameters, it is possible to study the conditions in which attractors and limit cycles appear. Intermediate systems, those found between order and chaos, have the most propitious landscape to evolve. This set of conditions define a new null hypothesis to determine whether there is evolution, it may act as a substitute for the Hardy-Weinberg equation. (Burian & Richardson 1996)

Yet, evolution can also be understood in terms of stages, like the model of *generative entrenchment* suggests (Shank & Wimsatt 1986). The starting standpoint is systemic: nature is divided into systems which are only partially decomposable and form several levels and unities. The idea is that natural systems are locked in stable ontogenetic paths which, once formed, can not be reshaped again, except when there is a general reorganisation of the great phylogenetic taxa. Development, then, makes evolution a quasi irreversible process in which each stage strongly determines subsequent evolution.

Also, from a thermodynamic perspective, Brooks & Wiley (1986) present what they call a Unified Theory of Evolution. A consideration of this may be useful to compare whether the generic properties are best expressed in terms of dynamics or of thermodynamics. Dynamic descriptions are deterministic, reversible and require a detailed knowledge of the initial conditions, while thermodynamic descriptions are stochastic, irreversible and require a selective description of initial data. A thermodynamic approach of evolution can bring

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022

together the diachronic and the synchronic perspectives.

If artificial evolution is elaborated in this way (in which if the evolving system is conceived as a complex material entity) both the change of embodied systems and the study of the generic properties of evolutionary change may be reunited. Many of the works where Artificial Evolution is used to design robots or artificial autonomous agents recognise and try to advance in the study of synchronic embodiment, but diachronic embodiment is not questioned. They combine a self-organising perspective of the structure and behaviour of the agent, with an limited selective perspective. Hence, elements which are disembodied in the diachronic sense are automatically introduced (adaptationism, optimality models and evolutionary stable strategies, externalism, poor genotype-phenotype mappings). Because of these reasons, a natural expansion of the notion of embodiment is to develop new forms of artificial evolution that can find a unified perspective.

The unified perspective contributes to embodiment because it can develop a different way of understanding functions as emergent and embodied constraints. Moreover, the philosophical discussion of whether they have ontological reality may be converted into more concrete ones. For some, functional or symbolic explanations only consider secondary properties, largely related to human observation and categorization, whereas others think that some natural systems (for example, the cell) produce their own functional or symbolic structures, that "stand for" longer reactions (or information processing) and can be transmitted as such in evolution (Pattee 1977). The search of generic properties is not an alternative between these positions, what is actually contended is that, underlying the property, there will be a certain specific relation between levels and it might be a non-detachable unit of evolution for the system. Thus, the investigation of functions as generic properties can integrate the diachronic and synchronic perspectives of embodiment, because it involves a further freeing from design (or a naturalistic understanding of it).

Counterfactual evolutions become also interesting. Instead of focusing on higher order functional properties as such, these are attached to the evolution of matter in the universe. The kind of properties and questions proposed for research in this sense have been of the type of why there are only two sexes in most taxa. Probably it is insufficient to investigate this by looking at different selection pressures that could have produced a variety of sexes. However, there might be insights to find if the question is posed in terms of self-organising properties and cohesion.

## Conclusion: evolution, creativity and the possible

Artificial Evolution and research in autonomous systems inherit ideas from sources of a very different epistemological style. The first has mainly followed standard evolutionary biology, while the second is based on self-organisation. An effort to expand the theoretical basis of artificial evolution should be made so as to achieve a better understanding of the evolution of organizations. This would have interesting consequences for the way we understand evolutionary creativity and the domain of the possible opened by it. When creativity is conceived in a combinatorial way, almost anything is possible, but no interesting phenomena occur. The reason for this is that organizations are based on emergent properties that confer some kind of physical cohesion to the system that makes it something different from a conjunction of parts or elements (Collier & Muller 1998).

The different purposes driving artificial evolution seek creativity in different ways but all of them –tools, models and instantiations– would benefit from a notion of constrained creativity that the evolution of organizations suggests. Non-intervening dynamical models are important to advance towards a notion of natural selection that is indeed natural, especially in the case of models of evolutionary phenomena. It is not so important to preclude conscious intervention in art, where perhaps the emphasis is placed more in using the evolutionary inspiration as a way to explore new forms of creativity, rather than in making the process biological. Yet, probably in both cases, biology and art, there is a similarity in the effort to understand the sources of creativity, which, also in the two of them, are not completely free or combinatorial, but constrained in ways we would like to understand better. A constrained creativity not only limits variety, it also enables novelty to appear in the system.

## Acknowledgement

I thank two Alife7 referees for comments. Funding for this work was provided by the DGICYT (MEC, Madrid) Grant PB95-0502, Basque Government Grants HU-1998-142 and EX-1998-146, and University of the Basque Country (UPV) Grant 003.230-HA079/99.

## References

Bedau, M. A. 1998. Four Puzzles about Life. *Artificial Life* 4 (2): 125-140

Brooks, R. A. 1986. A Robust layered control system for a mobile robot, *IEEE J. Rob. Autom.* 2: 14-23.

Brooks, D. & Wiley, E.O. 1986 *Evolution as entropy: Toward a unified Theory of Biology*. Chicago: Chicago University Press, 2nd edition.

Burian, R. & Richardson, R. 1996. Form and order in Evolutionary Biology. In M. Boden Ed. *The Philosophy of Artificial Life*. Oxford University Press.

Cariani, P. 1992. Emergence and Artificial Life. In C.G. Langton, C. Taylor, J.D. Farmer & S. Rasmussen Eds. *Artificial Life II*, Redwood City CA: Addison Wesley, 775-797.

Cliff, D., Harvey, I. & Husbands, P. 1993. Explorations in Evolutionary Robotics. *Adaptive Behavior* 2 (1): 73-110.

Collier, J. & Muller, S. 1998. The Dynamical Basis of Emergence in Natural Hierarchies. In George Farre and Tarkko Oksala (eds) *Emergence, Complexity,*

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022

*Hierarchy and Organization*, Selected and Edited Papers from the ECHO III Conference, Acta Polytechnica Scandinavica, MA91 (Finish Academy of Technology, Espoo, 1998).

Depew, D. & Weber, B. (1995) *Darwinism Evolving. Systems Dynamics and the Genealogy of Natural Selection.* Cambridge, MA & London: MIT Press.

Etxeberria, A. 1995. Vida Artificial. Bases teóricas y metodología. In I. Ramos, M. A. Fernandez, P. Gonzalez, G. Moreno & T. Rojo (eds) *Vida Artificial*, Albacete: Ed. Universidad de Albacete, Colección Estudios, pp. 15-35.

Etxeberria, A. 1998. Embodiment of Natural and artificial Agents. In G. van der Vijver, S. Salthe & M. Delpos (eds.) *Evolutionary Systems*, Kluwer, pp. 397-412.

Etxeberria, A. & Ibañez, J. 1999. Semiotics of the Artificial: The self of self-reproducing systems in cellular automata. *Semiotica* 127 (1/4): 295-320.

Farmer, J.D., Kauffman, S.A. & Packard, N.H. 1986. "Autocatalytic Replication of Polymers", *Physica D*, **22**: 50-67.

Gould, S.J. & Lewontin, R. 1979. The spandrels of San Marco and the Panglossian Paradigm: A critique of the Adaptationist Program, *Proceedings of the Royal Society of London*, 205: 581-98.

Harvey, I. 1996. Artificial Evolution and Real Robots. In Proceedings of International Symposium on Artificial Life and Robotics (AROB) Beppu, Japan Feb 18-20 1996. Masanori Sugisaka (ed.). ISBN4-9900462-6-9 (1996), pp. 138--141.

Hillis, W. D. 1992. Co-evolving parasites improve simulated evolution as an optimization procedure, En Langton, C.G., Taylor, C., Farmer, D. & Rassmusen, S. Eds. *Artificial Life II*, Addison-Wesley, Redwood City CA, 313-324.

Husbands, P., Harvey, I., Cliff, D. & Miller, G. 1997. Artificial Evolution: A new Path for Artificial Intelligence? *Brain and Cognition*, 34 (1): 130-159.

Kauffman, S.A. 1993. *The Origins of Order: Self-organization and Selection in Evolution.*, Oxford: Oxford University Press.

Langton, C.G. 1989. Artificial Life. In Langton, C.G. Ed. *Artificial Life*, Addison-Wesley, Redwood City CA, pp. 1-47.

Lewontin, R. 1983. The organism as the subject and object of evolution, *Scientia*, 118, pp 65-82.

Mataric, M. 1997. Studying the Role of Embodiment in Cognition, *Cybernetics & Systems*, 28 (6). Special issue on Epistemological Aspects of Embodied AI

Maynard Smith, J., Burian, R., Kauffman, S. Alberch, P., Campbell, J, Goodwin, B. Laude, R., Raup, D. & Wolpert, L. 1985. Developmental constraints and evolution: A perspective from the Mountain Lake Conference on development and evolution, *Quaterly Reviewsof Biology*, 60: 265-287.

Moreno, A., Etxeberria, A. & Umerez, J. 1994. Universality without matter? In R. Brooks, R.& P. Maes Eds. *Artificial Life IV*, Cambridge (MA): MIT Press, 406-410.

Moreno, A. & Ruiz-Mirazo, K. 1999. Metabolism and the Problem of its Universalization, *Biosystems*. 49 (1): 45-61

Pattee, H.H. 1977. Dynamic and Linguistic Modes of Complex Systems, *Int. J. General Systems*, **3**: 259-266.

Pattee, H.H. 1989. Simulations, Realizations and Theories of Life, In C. Langton Ed. *Artificial Life*, Redwood City CA: Addison Wesley, 63-77.

Pattee, H. H. 1995. Artificial life needs a real epistemology. In F. Moran, et al Eds. *Advances in Artificial Life*, Springer-Verlag. pp 23-38.

Polanyi, M. 1968. Life's Irreducible Structure, *Science*, **160**: 1308-1312.

Ray, T. 1992. An approach to the synthesis of life, En Langton, C.G., Taylor, C., Farmer, D. & Rassmusen, S. Eds. *Artificial Life II*, Addison-Wesley, Redwood City CA, 211-254.

Rosen, R. (1991) *Life itself*, New York: Columbia University Press.

Ruse, M. 1998. *Philosophy of Biology*, Prometheus Books.

Shank, J.C. & Wimsatt, W. C. 1986. Generative Entrenchement and Evolution, PSA 1986, Vol. 2, pags. 33-60.

Sims, K. 1994. Evolving 3D morphology and Behavior by Competition, En Brooks, R. & Maes, P. Eds. *Artificial Life IV*, Cambridge, MA: MIT Press, 28-39.

Varela, F. (1979) *Principles of Biological Autonomy*, New York: Elsevier North Holland.

von Neumann, J. 1966. *Theory of Self-reproducing Automata*, (A. Burks Ed.) Urbana: University of Illinois Press.

Umerez, J. 1998. The Evolution of the Symbolic Domain in Living Systems and Artificial Life. In G. Van der Vijver, S. Salthe & M. Delpos Eds. *Evolutionary Systems*, Dordrecht: Kluwer, pp. 377-396.

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268231/9780262291071_cco.pdf by guest on 19 August 2022