Artificial Life as a bridge between Science and Philosophy

Alvaro Moreno

Department of Logic and Ph. of Science
University of the Basque Country
Post Box 1249
20080 San Sebastian -- Donostia
ylpmobea@sf.ehu.es

Abstract
Artificial Life is developing into a peculiar type of discipline, claiming the principle of computational construction as the main avenue to explore and produce a new science of life "as it could be". In this research program, the generation of complex virtual systems may become the actual object of the theories, somehow substituting the usual empirical domain. This brings along not only a deep change in the traditional relationship between the ontological, epistemological and methodological levels, but also the appearance of a new relationship between the theoretical and technical realms, in what constitutes a relevant epistemic problem. Such a state of affairs forces us to reconsider the solid differences apparently established between science and philosophy.

Introduction
Artificial Life (AL) was originally conceived as a generalization of biology, since it was meant to become the study of "all possible life" (Langton, 89). However, like in the case of Columbus, who discovered the American continent when trying to reach India, in the course of its brief history as an independent discipline, AL has developed into something rather different from that foundational claim. This is partly because it is at stake whether it may actually reach that goal, and partly because it has generated a new set of problems and a new way of approaching the living phenomenon.

AL can be defined as the study of "lifelike" systems created by human action. But there are several elements in this definition that need to be clarified and stated more precisely. First, the very sense of the term "study", for AL is not only an epistemic process but also a technical activity; in AL the objects of study are literally created through technological action. Second, the actual meaning of "lifelike systems" (or systems that show "lifelike" behavior) becomes much complex (and controversial) than in traditional biology, since this concept has been understood in a much wider sense than that of biologically real biological systems. Thus, as C. Emmeche (94, p. 161) has pointed out, AL is founded as a modal discipline, establishing as its own objective the study of life "as it could be" and not simply "as we know it" (even if we include here extraterrestrial forms of life that might be discovered in the future). Last but not least, the idea of "artificiality" should be subject to examination. In addition to its generic sense of human construction, the term artificial has a double-sided peculiarity in the context of AL. On the one hand, it has a paradoxical meaning that comes from the idea that such humanly constructed systems be capable -like the actual natural living beings- of exhibiting creativity. The so-called "emergence" (in the behavior, capabilities, morphology, etc.) must be understood precisely as a form of indirect human creation (Boden, 96; Risan, 97) what in other words Langton describes as "getting the humans out of the loop", and thus endowing the actual machine with creativity. That is to say, one of the essential features of AL is that the artificially created system displays some type of agency, which allows us to speak (without falling into contradiction, although somewhat paradoxically) about autonomy in those cases. And on the other hand, artificial has the peculiar meaning (not exclusive but prevailing), like in AL, of virtual "construction" as opposed to physical realization.

The complex combination of all these elements has provided AL with a proper identity as a discipline, distinct not only from traditional biology, but also from the whole set of traditional empirical sciences. The idea that the living entity subject to study will be generated by a human agent -and, on top of that, in a computational universe- is truly suggestive; but, nevertheless, it also brings about novel issues and challenges in many respects.

AL as a computational research project
As it was mentioned above, the most usual meaning of the term artificial in AL research does not denote that the systems under study be the result of human action generically (i.e., as material realizations), but specifically refers to the generation of virtual systems in the computational universe. It is in fact necessary to take into consideration the strength that, since half a century ago, the tendency to separate organizational and informational
aspects from material and energetic ones has acquired in the study of systems, if we are to understand even the development (and specificity) of a discipline like AL at the end of this 20th century. Both IA and AL (following in this regard the precedent set up by cybernetics) share a common approach based on the idea that all of the material and energetic aspects of an organization do not affect its logical essence. Therefore, although a wide range of research enterprises (like robotics or some new branches of bioengineering) may be included within the field of AL, the main interest and area of activity is focused on the study of virtual systems generated in a computational environment.

Philosophically speaking, the computational versions of both IA and AL have been associated to functionalism. This position assumes and defends that the specific materiality that sustains a certain capability (mental, biological or of some other nature) is not relevant (Block, 80; Sober, 92). In the specific case of AL, it is claimed that biological phenomenology is exclusively the result of some organizational arrangement, rather than of a particular material implementation of it (Langton, 89). In fact, the question of whether those organizational arrangements are sustained by carbon or silicon molecules or by patterns of electrons in a computer is considered as completely irrelevant. Accordingly, and given the enormous possibilities to create and manipulate formal processes in the computational realm, it is obvious that the major line of research in AL be the purely computational one.

This huge potential for exploration of virtual organizations becomes in fact a way of "experimenting" with formal entities in formal environments, which should be therefore empirically interpreted. Hence, it is only at the end of the whole process that the analysis of the result can proceed; a result that will probably be difficult to compare or extrapolate to the real domain, since it is the consequence of pure abstractions of processes taking place in empirical environments. Nevertheless, even if this new strategy is highly promising with regard to the problem of the universalization of biology, it poses also new intriguing questions.

In particular, we want to refer here to three issues that we consider most interesting. The first one, which is basically epistemological (although it has ontological implications, as well), refers to the confusion between object and model; the second one, of methodological nature, is the problem of evaluating hypotheses (and, thereafter, research programs, too); and the third one concerns the interrelation between science, philosophy and technology.

1 This is the reason why Langton has defended the idea of universalizing classical biology through abstracting the materiality of biological phenomena and studying such phenomena assuming it may take place in a purely formal organizational domain.

The confusion between “object” and “model”

In biology (like in all the other traditional empirical sciences) conceptual models are elaborated to represent operatively a certain type of empirical systems, natural living systems, in this case. Such empirical systems thus constitute the objects of reference of the models. Regardless of whether these models bring about computational simulations or not, the latter involve the preexistence of a reference system, whose behavior is trying to be, total or partially, reproduced. However, computational “models” of AL are elaborated without direct and precise reference to the empirical biological reality. C. Emmeche (op. cit., p. 163) regards them as “second-order simulacra”, that is, copies of the copies themselves, generated not as abstractions on empirical biological systems (like in the case of concepts and theories of biology, which would be “first-order simulacra”) but on the theories themselves. Their main goal is to allow a new way of “computational experimentation” that enables us to “discover” the universal principles of living systems.

Sometimes, even the “model” or “simulacrum” is literally considered as a realization, that is to say, as an object whose phenomenology would make it equivalent to any other natural system of the corresponding empirical domain. This occurs when, due to the functionalist theses predominant in AL, the actual systems created in the computer are conceived not as representations of natural biological systems, but literally as artificial creations of the same type. Nevertheless, among those authors that regard these systems as true realizations, the use of the term “model” is quite widespread (even being aware of the differences with the classical meaning), probably because in the course of research and experimentation subsequent virtual tentative structures are generated (that is, for methodological reasons).

As a matter of fact, then, depending on circumstances, sometimes a computational system is interpreted as an

1 Indeed, there are also some intermediate cases of computational models based on standard AL methodologies. However, the design of such models aims only at the study of certain phenomenologies of living beings. In my view, the epistemological status of these models, rather than constituting a special differentiated case, is somewhere between models in traditional biological sciences and in AL. Usually these models are closely related to certain empirical biological systems, and typically elaborated and developed according to the synthetic and 'bottom-up' criteria of AL. This is why they potentially aim to find principles that not only govern the structure of those known systems taken as a reference, but of all possible realizations of the same type.

2 For instance, in the so-called "strong AL" it is claimed that computational simulations of living systems may really come to be living systems. Whereas "weak" AL considers that models represent certain aspects of living phenomena, strong AL would be ready to defend that the phenomenology that takes place in the actual computational environment is life in a proper sense.
object (ultimately, what is interpreted as an object is the model associated to the physical structure of the machine that sustains the execution of such a model), some other times as a proper model, and some other even as a tool or methodological technique mediating between the "true object" (natural biological systems) and the "true theory" (theories of empirical life sciences). There is another source of confusion, which must be located in the tendency to regard the computational system as a representation of a particular type of empirical biological system, i.e., as the realization of a universal that would include all possible systems of its own kind; or in other words, as they could be and not only as we happen to know them. In fact, there is a broad range of systems where the known phenomenology only constitutes a subset of all possible empirical systems of that type.

In short, the solution to these dilemmas involves determining which may be, ultimately, the ontological status of such artificial systems. Are they real or virtual? Which is the empirical reference of a computational "simulacrum"? Whichever answer may be given to these questions, it will certainly have methodological implications, since all those problems are tightly related to that of establishing the criteria for evaluation of the hypotheses that the very design of such systems attempts to test.

The problem of “empirical” evaluation

The fundamental problem in the methodology of AL research programs is that the ways of evaluating models are not empirically conclusive, for, by definition, the hypothetical empirical references of these models belong to a domain broader than the already known and even than the effectively existent. Despite the fact that “virtual experimentation” provides some formal rigor to the methodology, there is always a difficult problem of global empirical interpretation (Casti, 97; 99). This is one of the main differences with the classical scientific methodology, where experiments and measurements, no matter how idealized, are always performed or stated within the general framework of models that can be given a rigorous empirical interpretation.

Typically, an AL model is designed taking as a starting point some basic principles inspired in the general theories of a certain area of empirical biology. Then, these principles are introduced in the design of a computer program in which only low level instructions (local rules) are made explicit, so that new patterns (which play the role of new rules at higher levels) may appear. The characteristics of these patterns, however, are not previously known by the designer. In this process, the emergence in the computer of new structures (clearly distinguishable in spatial and/or temporal terms) is fundamental. Some of these new patterns get organized hierarchical and functionally so that, eventually, they become subject to empirical interpretation. As a result, the methodology consists in a continuous revision of the values of the parameters (and even, of some of the actual principles according to which the original design of the model was made), depending on the degree of coherence that the results being produced keep with the phenomenology of the empirical domain under study.

Therefore, the problem is manifest firstly because in AL computational systems evaluation is not discernible from the difficulties of interpreting “emergent” patterns or processes generated in such systems, and neither from the establishment of the epistemological criteria to select the primitives that define the models. So we find the initial problem of how to set up the main features of the model. Here, apart from the basic criterion for simplicity, the usual way to proceed is to search for (sub)models that, somehow, implicitly ensure that the assumed abstraction be “valid”; i.e., that the new provided information be coherent with what happens in well-known biological systems. The limits on simplification are set so as to avoid that the results produced appear trivial in the selected phenomenological domain. The aim is, thus, double-sided: simplicity is pursued, but provided that, at the same time, the model is able to generate enough complexity in the course of the simulation.

Thus, the design of models ought to fulfill two basic requirements. On the one hand, it is important that emergent processes or structures, which may be interpreted as new pieces of knowledge in the empirical domain under investigation, are obtained. But on the other hand (and in order to avoid the problem of how to determine when the empirical interpretation of these new structures and/or

---

4 E.T. Olson (97), for instance, has argued against the objections of the “immateriality” of those “virtual organizations”, defending that one could actually take them as real physical systems if, instead of looking at the computer program in abstract terms, one considers the patterns of electrons which materially support the execution of that program. Yet, this argument misses the point, because the material structures which support the operational level of computer simulations are entirely passive, since their intrinsic dynamics is completely constrained (Moreno & Ruiz-Mirazo, 99).

5 H. Pattee (89) has criticized this confusion between computational simulations and realizations of material biological systems, arguing that the former are mere symbolic systems, whereas in the latter their capacities derive precisely from the symbolic constraint (autonomously generated) in material systems.

6 B. Smith (96), for instance, claims that the computational world, in fact, does not constitute a proper subject matter of study; rather, if should be considered as a complex social practice that involves the design, construction, maintenance and use of intentional artifacts (pp. 75 and 359).

7 In fact, one of the main objectives of a model in AL is to show something new not only in that phenomenological domain but also in the domain of the internal correlations of the systems under study.
processes is objectively justified), it is necessary to design the computational experiment so that at least the type of expected results can be forecasted, and these are specific enough within each particular domain. If only the first condition is met, the problem then becomes how to justify its interpretation; and if it is only the second one, the risk is that the result becomes trivial. Thus, depending on the results given by running the program it will be assessed whether the model is valid or not, although there are some other elements that take part in the evaluation, like in traditional sciences (for instance: auxiliary theories, the adequacy of the computational tools employed, or of the set-up values for the initial parameters, etc.). Nevertheless, the problem of the evaluation of models goes further beyond this, since, depending on the epistemological status given to computational systems, the actual research program may vary quite radically. Should we try to design “plausible” models (understanding plausible in the sense that computational models of AL ought to “resemble” the behavior of the corresponding natural systems) or, on the contrary, the main goal should be oriented in some other direction, like in the search for models that fulfil primarily certain formal criteria (for example, the capacity to generate computational complexity, or similar ones).

The problem with a purely formalist conception of AL (i.e., that the task of AL models is mainly to “unfold” and make explicit the logic consequences of the starting premises) is that AL moves into the risk of becoming a discipline closer to mathematics than to any empirical science. Several authors have formulated serious doubts and wariness about this issue. Maynard Smith, for instance, has objected that AL is a “science without facts”, referring to the problem of how to assess a set of computational models whose (potential) empirical references are imprecise and generic (quoted in Horgan, 95). However, this criticism is too strong, for most Alifers, both in the design of the models and in the evaluation of their results, usually take quite seriously into account the theories as well as the behaviors present in the respective empirical domain. At the same time, the worries to escape mathematization might well put at stake the actual autonomy of AL as a discipline.

The attempt to answer these questions takes us well beyond the strictly methodological realm, as it requires determining the actual nature of that which the model discerns (sometimes, rather than a particular phenomenon in a particular empirical domain, what the model addresses and hopefully elucidates are certain computational issues...). Nevertheless, this is related at the same time to the role that meta-theoretical conceptual elements play in the design of models, as well as to questions that have to do with the nature of the technical objects employed.

8 Miller (95), for example, has claimed that AL can only become a fully scientific discipline through its integration in the field of theoretical biology.

The problem of the interrelation between science, philosophy and technology

AL constitutes a discipline that, due to the nature of the problems it addresses, has progressively moved into an intermediate area between philosophy and science. AL allows to face, analyze and sometimes solve (or at least re-formulate) in an entirely new way several important philosophical issues in biology, such as the debates between reductionism and vitalism, the problem of emergence, or the matter-form relationship. This is so as a result of the deep entanglement between the AL methodology and the meta-theoretical problems of biology. Regarding this question, J. Casti (97, pp. 187-8) and D. Lane (95) hold that the construction of “models” in AL (and other computational sciences) requires theories about particular empirical domains; namely, about the kind of objects that are found in each of them, and about how these objects relate to each other, or the type of processes that change their nature (including creation and destruction processes). The relevance of computational models with regard to their corresponding empirical domains would be mediated by these theories, which constrain the interpretation of the former (i.e., the models). In turn, such models have a significant advantage, for they allow the analysis of these theories with the aid of more powerful tools than ordinary language.

Yet, this idea poses some problems. Firstly, as it has been broadly argued by the post-positivist philosophy of science, the concepts on which models in traditional empirical sciences are constructed, are also mediated by theories, in a similar sense to that pointed by Lane (95). This is why the differences between one and the other type of models are not stated in such terms, although it is quite true that the use of this kind of virtual models does indeed involve, and rather directly, the aforementioned meta-theoretical problems. The question to address is why it does so.

Secondly, these models are, in principle, purely formal, even though the formulation of its basic principles is inspired in a given biological domain. As we mentioned above, the difficulties of interpreting and testing AL models in terms of the current biological knowledge lie on the meta-theoretical implications that are found implicit in the epistemological status of the entities and relations that constitute such models. Nevertheless, the main criterion that we will use to evaluate the usefulness of these models is their capacity to improve our understanding of the theories about the empirical biological world. Apparently, this is a vicious circle.

However, according to Lane (95), these problems could be progressively solved. On the one hand, the models of AL and other computational sciences can improve our theories about the corresponding empirical domains; and on the other hand, these theories become essential for testing the relevance of virtual models. In this way, a kind of “hermeneutic” circularity is established between models and theories (Kleindorfer et al, 98), since the models are used for the elucidation of the theories (about the empirical

---

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268203/9780262291071_ccc.pdf by guest on 12 July 2021
world) and these for the interpretation of the models. A highly significant feature of this idea is that the interaction is not established merely between a philosophical or meta-scientific level and a proper scientific one, but a new technological level (constituted by the computational tools, with their double-sided nature, as software and hardware) plays also a crucial role in this hermeneutic process. This is because computers\textsuperscript{9} make possible the exploration and visualization of a given set of work premises through the creation of virtual worlds of indefinite complexity.

D. Dennett (94) holds a similar position when he points out that AL can be conceived as a special sort of philosophy\textsuperscript{10} which allows the creation and testing of complex thought experiments, “kept honest by requirements that could never be imposed on the naked mind of a human thinker alone”. That is why this author considers that the research program in AL consists basically in the creation of prosthetically controlled thought experiments of indefinite complexity. By increasing the capacity and precision of the human mind through these prostheses that are the computers, we would be ready to increase indefinitely the complexity of such experiments, as well.

Anyway, it is not just a question of the computer acting as an extension of our conceptual world, as Dennett asserts; the novelty lies on the fact that the blending between concepts and computational technology creates a new domain with its own ontological claims. Somehow, in AL a set of systems generated as a result of putting together human concepts and artifacts, gets constituted as a set of objects. In other words, an ontology is founded from an epistemological standpoint. Yet, this happens in a way that such concepts, once embodied in the machine, are transformed (since they have acquired some ontological dimension) and establish, in turn, a new epistemological relationship with their creators and users. This is why AL, on the one hand, constitutes an engineering activity whose starting point is located at a set of basic biological intuitions about the empirical biological domain; it attempts to build systems capable of producing by themselves emergent and functional processes. In order to do so, it proceeds by integrating creatively different technological resources, together with theoretical developments in computer science and physics. And on the other hand, it involves a process of interpretation of the behavior of such systems with the aim of broadening the phenomenology of the empirical biological domain and of developing new concepts coherent with the body of knowledge of traditional biology, which properly studies that domain. Both processes -construction and interpretation- are deeply entangled, since the AL research program is a continuous concatenation of constructions, interpretations, modifications, new interpretations and reconstructions.

**Conclusion**

If things are displayed and understood in this way, AL would come to be a bridge between empirical science and philosophy of science, as each of them has been traditionally conceived. The methodological realm appears deeply intertwined with the properly theoretical one, so as to make it very difficult to discriminate among the possible contributions of a certain model and state neatly which are the heuristic, the epistemological, the theoretical and even the “empirical” aspects involved. Whereas in traditional philosophy of science the interaction under analysis is that between the empirical domain, the empirical theories and the meta-theories (a relationship in which technology only plays a significant role in the interactions between the first two), in AL a new way of interrelating empirical theories and meta-theories arises, a way that involves the mediation of technological devices. This new connection is, on the one hand, more limited than the established in the traditional philosophy of biology, since AL must be restricted to deal with conceptual issues that come out in the context of implemented (computational) systems; but on the other hand, AL is more rigorous and powerful, because it uses quasi-experimental methods of validation (computational experimentation) and also because it can, in principle, make the traditional range of problems of biological meta-theories even wider. Hence, research in AL is opening up radically new perspectives, not only in the sense of bringing about profound changes in the meaning of concepts like experimentation, model or evaluation (of theories); but also because the status of the philosophy of biology is clearly modified and deeply interwoven with technology. All this puts at stake the classical differences made between science and philosophy and demands a more elaborate framework to give proper accounts of the more and more complex and dialectical bonds established between them.

**Acknowledgments.**

The author acknowledges funding from the Research Project Number PB95-0502 from the DGICYT-MEC, the EX-1998-146 and HU-1998-142 from the Basque Government, and the UPV 003.230-HA079/99 from the University of Basque Country. K. Ruiz-Mirazo read a first draft of this paper and contributed to make complete its final version.

**References**


