

Animats in the Modeling Ecosystem

Xabier E. Barandiaran¹, Anthony Chemero²

¹ *Autonomous Systems Laboratory, Universidad Politécnica de Madrid, Spain*

² *Scientific and Philosophical Studies of Mind Program, Franklin and Marshall College
Lancaster*

1 Discovery and Diversity in the Modeling Ecosystem

There are many different kinds of model and scientists do all kind of things with them. This diversity of model type and model use is a good thing for science. Indeed, it is crucial especially for the biological and cognitive sciences, which have to solve many different problems at many different scales, ranging from the most concrete of the structural details of a DNA molecule to the most abstract and generic principles of self-organization in networks. Getting a grip (or more likely many separate grips) on this range of topics calls for a teeming forest of techniques, including many different modeling techniques. Barbara Webb's target article strikes us as a proposal for clear-cutting the forest. We think clear-cutting here would be as good for science as it is for non-metaphorical forests. Our argument for this is primarily a recitation of a few of the ways that diversity has been useful.

Recently, looking at the actual practice of artificial life modelers, one of us distinguished four uses of simulation models classified in terms of the position the models take up between theory and data (see Figure 1). The classification is not exhaustive, and the barriers between kinds are not absolute. Rather, the purpose of the taxonomy is to open up the view for an epistemic ecology of modeling practices. First, and closest to the empirical domain, there are *mechanistic* models, in which there is an almost one-to-one correspondence between variables in the model and observables in the target system and its environment. Webb's

cricket robot is a paradigmatic example of this type. Second, there are *functional* models, which aim for a behavioral or functional rather than a variable-to-variable correspondence between the model system and its target. Many models in cognitive psychology are of this type, as are many in biology when the underlying mechanisms are not accessible to modeling correspondence (see Vickerstaff & Di Paolo, 2005, for a good example of this). Third, there are *generic* models, which cover a wide spectrum of phenomena in search for generic principles of complex systems. Cellular automata, random Boolean networks, and the like belong to this class. Finally, there are *conceptual* models, which do not target any particular natural system nor a wide spectrum of them. Instead conceptual models are built from theories, for which they embody assumptions, illustrate concepts, simulate theoretical principles, and so forth. Beer's model and many other animats are of this type (for more details see Barandiaran, 2008, chap. 2; Barandiaran & Moreno, 2006).

Modeling is a relational activity, it involves the template or construct (usually referred as the *model*) and also an interpretative framework (made of assumptions, generalizations, definitions, etc.) that puts the model in connection with other models, theories or objects. It is this interpretative framework and the modelers' intentions that situate a given model into one of the categories of models. But only mechanistic models are used as Webb suggests all models should be: "the cricket robot is a mixed physical and computational implementation of a particular hypothesis about an observed phenomenon of sensorimotor behavior in

Correspondence to: Xabier E. Barandiaran, Autonomous Systems Laboratory, Universidad Politécnica de Madrid, Spain.
E-mail: barandi@sf.chu.es

The Modeling Ecosystem

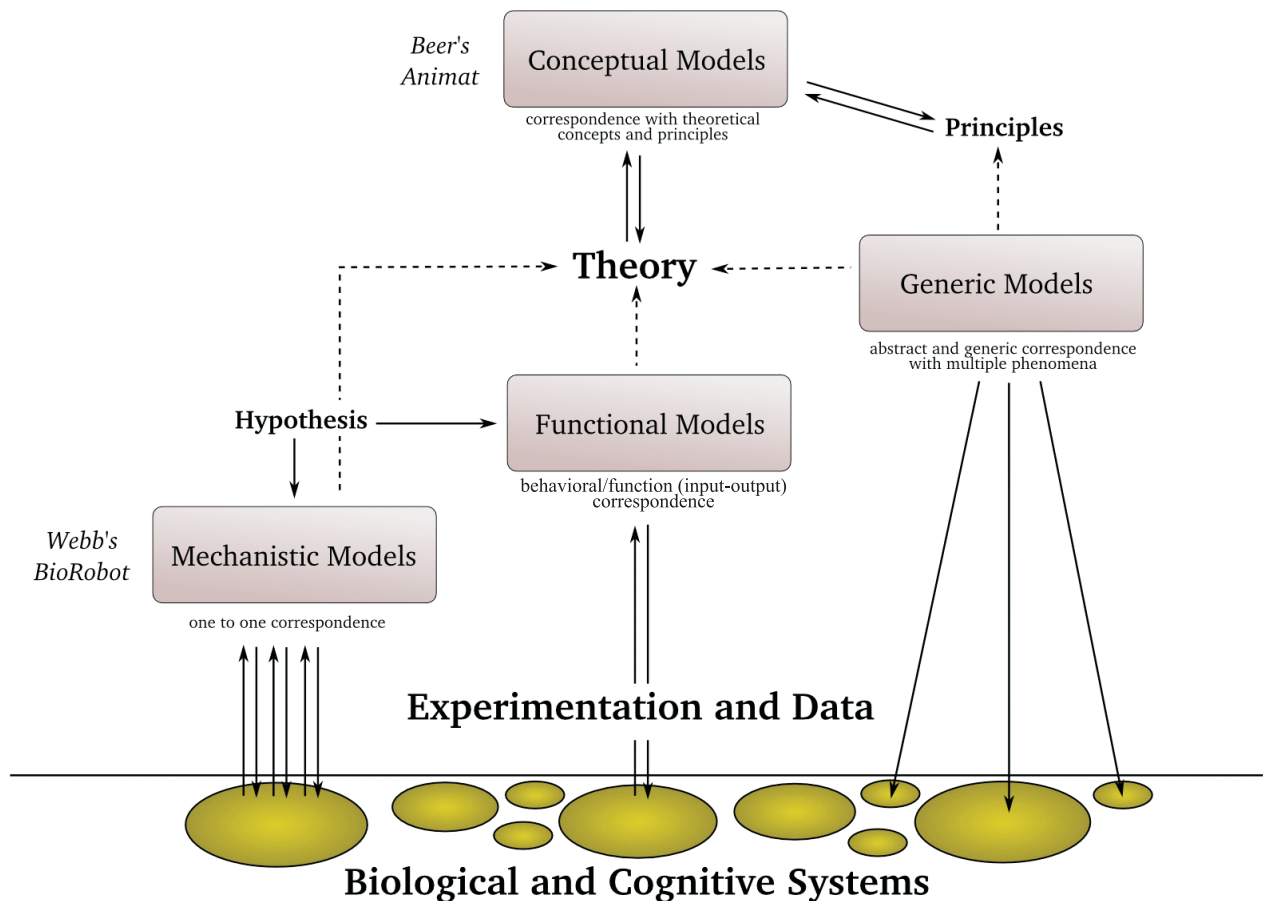


Figure 1 The modeling ecosystem. Copyright (c) 2009 Xavier E. Barandiaran under a Creative Commons Attribution Share-Alike licence, freedom to copy, modify and distribution provided that this notice is preserved.

a certain animal. It is relevant to biology to the extent that its components and behavior can be directly compared with that animal” (sec. 3). The validity of this type of model depends on its correspondence with empirical data. However, this is not the only way in which a model can become relevant to biology, as Webb herself points out.

It is worth noting at this juncture that this work has not simply been a case of using the robot model to verify or falsify existing hypotheses in the biological literature. It has involved substantial integration of disparate information about auditory mechanisms, neural data, and behavioral observations. It has led to the proposal of several novel hypotheses about the function, some of which have been supported, and some contradicted, in subsequent investi-

gations. It has also made much more apparent the areas in which biological data is most critically lacking. (sec. 2.1).

The point Webb is making here is that her models are acting as *guides to discovery* (Chemero, 2000, 2009). A guide to discovery is some means for a scientific research program to advance by making predictions for future experimentation, or extending the reach of the program to new phenomena, or solving conceptual problems within the program, or casting empirical findings in a new light, and so forth. A scientific research program, in other words, needs a guide to discovery in order to be progressive in the Lakatosian sense (Lakatos, 1970). Models, because they are poised between theory and data, are in an ideal position to act as guides to discovery. So, although she

does not use this language, Webb is correct that detailed mechanistic modeling can act as a guide to discovery for a research program in the biological and cognitive sciences. But so too can conceptual models, the type that most animat models fit into. Webb seems to miss this point and tries to shoehorn animat research into the category of mechanistic models:

However, in the general area of animat models, and specifically in work such as that of Beer, the systems ... are constructed to represent (however loosely or abstractly) some mechanisms taken, by hypothesis, to have causal relevance to biology.... establishing the relevance requires some explicit specification of how the two systems—artificial and biological—are supposed to correspond.” (sec. 3.4).

She latter adds “the question of its relevance to biology remains dependent on the existence and extent of the mapping between his agent and real cognitive systems.” (sec. 4.1).

As we will see in detail in what follows, conceptual models are used to illustrate or embody theoretical concepts, and their validity depends on how fruitful they are in terms of disclosing contradictions within theories, forcing precise definitions, reorganizing conceptual systems, comparing explanatory paradigms, and exploring their potential interaction among separate theories. All of which is to say that conceptual and generic models can also serve as guides to discovery, but not in virtue of close correspondence with any particular animal. Their relevance does not depend “on the existence and extent of the mapping between agent and real cognitive system”, as Webb claims.

We will detail a few of the ways in which conceptual models act as guides to discovery in the biological and cognitive sciences in what follows, but before doing so we simply state our contention that any type of model (or mathematical tool or experimental methodology) that acts as a guide to discovery in the biological and cognitive sciences is legitimate. This includes animat research, Beer’s in particular.

2 Validity and Correspondence of Conceptual Models

Conceptual models involve the simulation or realization of processes which are, in virtue of some dynamic

or structural analogy with theoretical notions, conceptualized under a certain theory of the living, cognitive, social or, more generally, complex systems. Conceptual models can be very abstract or very specific depending on the theory under which they are interpreted/constructed. For an instance of the former case, the model could work to illustrate, formalize, or compare one or more theories of reduction and emergence using, let us say, cellular automata patterns (e.g., Bersini, 2004; Crutchfield, 1994). For an instance of the latter, a domain-specific conceptual model can be exemplified by a simulation of active perception in situated agents (such as Beer’s animat). In either case, conceptual models are not abstractions from specific sets of biological data, but from the biological theories themselves: for example, from generalization and abstractions of other models, from the theoretical assumptions required to interpret a family of models, or from idealized and artificial applications of generic principles. Conceptual models are not attempts to capture some specific set of worldly facts, but are virtual worlds in their own right (e.g., Tierra; Ray, 1992). Webb is not the first to criticize animat research along these lines. Maynard Smith called artificial life into question arguing that it is a “science without facts” (quoted in Horgan, 1995), and asking how to assess a set of computational models whose (potential) empirical references are imprecise or non-existent. However, it would be an error to evaluate conceptual simulation models by traditional empiricist or observational standards (as Webb’s repeated appeal to “validity” suggests). The main interest (and methodological novelty) of conceptual simulation models lies in their capacity to develop experimental research on the internal conceptual relationships within theories of biological or cognitive organization. This computational research allows what Dennett (1994) calls the realization of highly rigorous and far-reaching thought experiments, which the naked human mind could never perform on its own.

Bedau (1998) and, particularly, Di Paolo, Noble, and Bullock (2000) have elaborated a more detailed account of the role and methodology of artificial life as *opaque thought experiments*. The opacity of the thought experiment lies on the complexity of the model. The unfolding of properties and patterns from a set of premises (local rules or differential equations) are not always predictable in the absence of a computer simulation that performs recursive calculations of

non-linear functions, integrates random perturbations, visualizes the results, and so on. As is the case with traditional thought experiments, the epistemic value of conceptual simulation models does not lie on their adequacy to some empirical phenomena (since the thought experiment involves hypothetical and idealized situations). On the contrary, the model operates on the hidden assumptions of the theories used to design and interpret the model and on the conceptual relationships between these assumptions.

When relationships among the concepts of a theory cannot be derived on logical grounds, computer simulations and robotic artifacts become cognitive tools for theoretical development (Casti, 1997; Harvey, 2000) and necessary guides to discovery. For instance, learning and ontogenetic or phenotypic plasticity have intricate effects on evolution. The interaction between these scales is difficult to study through natural fossil records or other empirical means. Furthermore, it turns out to be extremely difficult to theorize about the interplay between dynamics that occur at two radically different time scales and with consequences that only appear through repeated interactions and infinitesimal (and often non-linear) cumulative changes. An alternative is to develop artificial worlds (whose local rules are abstractions of the generic principles that evolutionary theory takes to be essential for natural evolution) where simplified forms of evolution and learning can be studied. The Baldwin effect (Baldwin, 1896), for example, was nicely demonstrated by a computer model by Hinton and Nowlan (1987) and gave rise to a revival of the subject (Weber & Depew, 2003). Subsequent artificial life research has made explicit many other properties and dynamic relationships between learning and evolution (Ackley & Littman, 1992; Mills & Watson, 2005; Suzuki & Arita, 2004) that remained opaque or hidden to naked human thinking and analytic mathematical techniques. For instance, Mills and Watson (2006), making use of a simulation model, argued that genetic assimilation is, at least, theoretically sufficient for the Baldwin effect to occur and that canalization is not necessary. These conceptual models act as guides to discovery primarily forcing biologists to reconsider the phenomena they observe, a reconsideration that would not have occurred had the models been ruled out as irrelevant.

3 Webb and Beer, Not Webb Versus Beer

It is not at all clear why conceptual models are not valid for Webb, but it is clear that targeting real animals is not required for conceptual models to become relevant to biology. In fact, some undeniably useful conceptual models could not possibly target any specific, empirically addressable phenomenon. Consider for instance simulation models of the Baldwin effect just discussed. There is no doubt that this type of model is relevant to biology. Biology as a science is not only about concrete real animals but also, and perhaps more importantly, about general principles of biological organization, development and evolution. It is this type of general principle that models of the Baldwin effect are trying to capture, improve upon, reconceptualize or illustrate for theoretical purposes. How could the targeting of a specific species be of help or even possible for this purpose? Or, to take an example from Webb's target article, Bongard and Paul (2000) explore the correlation between bilateral symmetry and locomotive efficiency. What exactly would have their model gained if it had targeted jellyfish radial body plans versus that of early bilaterians (which are anyway lost and only theoretically reconstructed)?

It is obvious to us that there are a variety of modeling practices in biology, each of which is necessary and none of which is sufficient, on its own, to solve all the real problems of biology. Some models (particularly most of the good animat models) are valid and useful for biological and cognitive theories without having to target existing animals. These conceptual models, and Beer's in particular, act as guides to discovery. Because of this, Webb's insistence that one needs a mapping between the simulated agent and the real animal for a model to be of relevance to biology is mistaken. It is hard for us to see why Webb fails to see the importance of modeling like Beer's. One possibility is that it is so different from what she does. Indeed, what Beer is doing is just the opposite of Webb's biorobotic implementation of specific mechanisms. Beer is raising the theoretical question of the importance of embodiment and situatedness for cognitive behavior and illustrating epistemological and methodological principles for cognitive explanations. He is doing so in the (theoretical) context of representation-hungry problems which, according to computational

theories, *in principle* required internal representational tokens. Beer's model shows this not to be the case and challenges the explanatory role of representations in cognitive science. Perhaps more importantly, Beer provides one of the first (and certainly the most complete) dynamicist explanations of full brain–body–environment system in the context of a minimally cognitive task. This is not to say that a model of a real animal could not have done the job and there is no problem in using a model originally developed as a mechanistic model to raise conceptual and methodological issues. But even in such a case, the fact that the model could correspond to a specific target organism would be irrelevant for the theoretical claim made; it could even obscure the claim by including unnecessary and *irrelevant* details that only apply to the specific organism that it models.

We want to be clear that we are not to be interpreted as criticizing mechanistic models, or Webb's models in particular. We are fans of Webb's work on cricket phonotaxis. It is clear to us, though, that Webb holds too narrow a conception of models, their validity and how they might serve as guides to discovery. Science is not only about directly and empirically addressable hypotheses and their instantiation in models. It is also about theoretical articulation of hypothesis, about principles, about explanations of the possible (not only about the contingently real in our world). And scientific models can be adapted to a variety of uses. By making a model represent a real existing animal, there is a gain in specific explanatory usefulness, potential applicability and one can experimentally enrich the model with feedback from its target. There is no doubt about this. Yet, Beer's animat is a perfectly valid model without having to target any real existing organism. And like Beer's many other theoretical and conceptual models are of great relevance in biology and cognitive science. They act as guides to discovery when science deals with principles of complex systems whose theoretical articulation often requires the aid of computer simulations. Animats are but one species of such models and preserving diversity is vitally important in the modeling ecosystem, just as it is in real ecosystems. We should not be forced to choose between animats and animals.

References

- Ackley, D., & Littman, M. (1992). Interactions between learning and evolution. In C. G. Langton et al. (Eds.), *Proceedings of Artificial Life II* (pp. 487–509). Redwood City, CA: Addison-Wesley.
- Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, 30, 441–451, 536–553.
- Barandiaran, X. (2008). *Mental life. A naturalized approach to the autonomy of cognitive agents*. Unpublished doctoral dissertation, University of the Basque Country, Spain. <http://barandiaran.net/phdthesis/>
- Barandiaran, X., & Moreno, A. (2006). Alife models as epistemic artefacts. In L. M. Rocha, L. S. Yaeger, M. A. Bedau, D. Floreano, R. L. Goldstone, & A. Vespignani (Eds.), *Artificial Life X: 10th International Conference on the Simulation and Synthesis of Living Systems* (pp. 513–519). Cambridge, MA: MIT Press.
- Bedau, M. A. (1998). Philosophical content and method of artificial life. In T. W. Bynum & J. H. Moor (Eds.), *The digital phoenix: How computers are changing philosophy* (pp. 135–152). Oxford, UK: Basil Blackwell.
- Bersini, H. (2004). Whatever emerges should be intrinsically useful. In *Artificial Life 9* (pp. 226–231). Cambridge, MA: MIT Press.
- Bongard, J., & Paul, C. (2000). Investigating morphological symmetry and locomotive efficiency using virtual embodied evolution. In J.-A. Meyer et al. (Eds.), *From Animals to Animats: The Sixth International Conference on the Simulation of Adaptive Behavior*.
- Casti, J. (1997). *Would-be worlds*. New York: John Wiley.
- Chemero, A. (2000). Anti-representationalism and the dynamical stance. *Philosophy of Science*, 67(4), 625–647.
- Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- Crutchfield, J. P. (1994). The calculi of emergence: Computation, dynamics, and induction, *Physica D*, 75, 11–54.
- Dennett, D. (1994). Artificial life as philosophy. *Artificial Life*, 1(3), 291–292.
- Di Paolo, E. A., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. In M. A. Bedau, J. S. MacCaskill, N. H. Packard, & S. Rasmussen (Eds.), *Artificial Life VII: The Seventh International Conference on the Simulation and Synthesis of Living Systems*, (pp. 497–506). Cambridge, MA: MIT Press.
- Harvey, I. (2000). Robotics: Philosophy of mind using a screwdriver. In T. Gomi (Ed.), *Evolutionary Robotics: From Intelligent Robots to Artificial Life, Vol. III* (pp. 207–230). Ontario, Canada: AAI Books.
- Hinton, G. E., & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1, 495–502.
- Horgan, J. (1995). From complexity to perplexity. *Scientific American*, 272, 104–109.

- Mills, R., & Watson, R. A. (2005). Genetic assimilation and canalization in the Baldwin effect. In M. Capcarrere et al. (Eds.), *Proceedings of the Eighth European Conference on Artificial Life* (pp. 353–362). Berlin: Springer Verlag.
- Ray, T. (1992). An approach to the synthesis of life. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II. Proceedings of the Workshop on Artificial Life* (pp. 325–371). Redwood City, CA: Addison-Wesley.
- Suzuki, R., & Arita, T. (2004). Drastic changes in roles of learning in the course of evolution. In J. B. Pollack, M. Bedau, P. Husbands, T. Ikegami, & R. A. Watson (Eds.), *Proceedings of Artificial Life IX* (pp. 369–374). Cambridge, MA: MIT press.
- Vickerstaff, R. J., & Di Paolo, E. A., (2005). Evolving neural models of path integration. *Journal of Experimental Biology*, 208, 3349–3366.
- Weber, B. H., & Depew, D. J. (Eds.) (2003). *Evolution and learning. The Baldwin effect reconsidered*. Cambridge, MA: MIT Press.