*Adaptive* **Behavior**

# In Defense of the Abstracted Animat

Seth Bullock
*School of Electronics and Computer Science, University of Southampton*

It is healthy, I believe, to periodically examine the way in which a field's activity is supposed to be valuable, both internally, to those within the field itself, but also externally to the wider scientific community in which it is embedded, and ultimately society at large. Otherwise, there may be a tendency for modes or styles of work to lock in and be perpetuated uncritically. However, it is clear that such a process of critique must take care with how it defines what counts as valuable. If the definition is cast too close to the community being examined, there is a risk of swallowing a motivating story that may perhaps be internally consistent and widely accepted within the community, but does not survive scrutiny from a more critical perspective. Eugenic science might be cited as an example. Its value relies on subscribing to a set of background assumptions that should themselves be part of any critique. On the other hand, the value of many scientific activities would be difficult to discern from too remote a vantage point. What exactly is the value to future generations of investing in the study of a particular mollusk, for instance?

Here, Webb's sights are trained on the "synthetic methodology," and in particular a specific sub-community involved in building systems that are intended to shed light on biological questions. Consequently the sense of value that Webb is concerned with is largely derived from the aims of biological science, which seems fair enough. However, I believe that there is a problem in the scoping here, which I will return to below.

In large part, however, I am in agreement with Webb's line of argument and the useful distinctions that she makes between different motivating stories that can be found throughout the ALife and adaptive behavior literatures. I share her sense that the things being built should be thought of as models, and that they are therefore models of something. This mindset should always make us wary of systems that are presented as artificial worlds worthy of study for their own sake—free-floating models that have been disconnected from the target systems that they were originally motivated by. I agree that the animats built by the community under consideration, whether simulations or fabrications, do instantiate or represent hypotheses about nature, and that, as such, their value is intimately tied up in which hypotheses are represented and to what end. Where animat behavior challenges our ideas in a way that can be empirically tested, it is, or should be, taking part in a scientific cycle of theory–hypothesis–experiment–theory (Di Paolo, Noble, & Bullock, 2000).

However, I do not agree with the repeated assertion running through the article that the biological relevance of a model must be cashed out in terms of whether or not it speaks to the properties of a specific animal or species. This position seems to reflect Webb's own interests in biological mechanisms. How does it manifest itself? In the abstract: "Animat simulations ... represent hypotheses about ... animals". Later: "But if the results of animat models are to be relevant to biology, then ultimately *some* claim about the physical mechanisms of real animals must be asserted." (sec. 4.2). This idea leads Webb to assume that because Beer claims that his work "has important implications for our *understanding* of perception" it must also be the case that "empirical claims *are* being made about the world on the basis of the model results" (sec. 3.5).

I think this move is made too quickly and is overly narrow.

*Correspondence to*: Seth Bullock, School of Electronics and Computer Science, University of Southampton, UK.
*E-mail*: sgb@ecs.soton.ac.uk


(2009), Vol 00(0): 000–000.
DOI: 10.1177/1059712309340845

Consider the Hawk–Dove game (Maynard Smith, 1982). This is presented by Maynard Smith as a mathematical model, but could also be instantiated as a simulation or even perhaps with robots. It is also, indisputably, a model with biological relevance. But it does not achieve this relevance by being a model of hawks, doves, or any particular animal or species of animal. In its simplest form, the model presents an evolutionary game in which two strategies for resolving conflict may be adopted: Hawk or Dove. Hawks are aggressive and will fight to resolve a conflict. Doves are pacifist and will back down when attacked. Game-theoretic analysis demonstrates that neither strategy is evolutionarily stable. In particular, although a population of Doves settles disputes most efficiently (in terms of achieving the minimum average net cost per dispute), such a population can be invaded by Hawks that bring conflict and its attendant high costs. While, like the work of Beer and Webb, the Hawk–Dove model features idealizations of animal behavior, it is cast at the level of evolutionary adaptation. The lessons that it delivers concern our understanding of the biological world and the species that make it up, but do not speak to the properties of any particular species or animal (rather they shed light on the vulnerability of cooperation to exploitation and the way in which group-level benefits can be evolutionarily undermined by short-sighted individualism). As such, the Hawk–Dove game is an example of a kind of model that has value without making claims about "the physical mechanisms of real animals" and is therefore outside the scope of Webb's analysis.

Might pieces of modeling work within ALife or adaptive behavior also have value without shedding light on particular biological mechanisms within animal species? At first glance it might appear that models of evolutionary adaptation such as the Hawk–Dove game are a special case in this respect since evolutionary adaptation is an overarching process that organizes biological phenomena, whereas the majority of biology is concerned with understanding and explaining these phenomena, for example, cricket phonotaxis, arthropod digestion, insect cognition, and so forth. However, closer inspection reveals that even the most species-specific mechanism can be the subject of biological enquiry that does not seek to shed light on the causal mechanisms themselves, but on, say, their functional role within an organism (Tinbergen, 1963). This explains how biologists could study some behavioral phenomenon in ignorance of any mechanistic account of how it is achieved (e.g., Tinbergen's own work on the functional value of imprinting behaviors in birds, or, say, the evolutionary theory of the sex ratio).

Consequently, might not Beer's work on cognition and perception, or other ALife/adaptive behavior work on the origins of life, action selection, metabolism, group selection, and so forth, have "important implications for our *understanding"* of these phenomena without engaging with the particularities of how they are brought about by specific biological mechanisms? How are such models relevant to biology if they do not directly shed light on biological mechanisms? They are relevant in the same way that all well-formed biological models are relevant: they engage with extant biological ideas.

In fact models must always be resident within the "world of ideas." They do not take part in empirical science in the same way that an experiment does, providing evidence about the world. Instead they take part indirectly by re-presenting the theoretical ideas that were stimulated by empirical discoveries and steering further exploration. As such, the opposition between animats (models) and animals (target systems) presented in the title of Webb's article is spurious: Webb is not advocating that we build animals instead of animats. She would like us to build models with a clear role to play within biology, rather than animats that elide the difference between technological achievement and scientific value. At the same time as supporting this position, we should reserve a valid role for models that aim at exploring phenomena abstracted from the biological mechanisms that implement them in nature.

This is not to say that the process of abstraction is unproblematic. Maynard Smith has claimed that in specifying an ESS model, rather than ignoring or sidestepping the problem of engaging with the mechanistic, developmental or biochemical basis of behavior, the modeler is being forced to make explicit her position on what is within the remit of the model: Hawk behavior and Dove behavior, but not time-travelling behavior or magic behavior or the use of advanced logics, brute computation, and so on. An ESS model that posits behavioral strategies that do not make sense in the light of empirical observations of relevant living systems is vulnerable to being dismissed as a consequence. Famously, abstracting the properties required in order to count as a living creature from empirical observations of known living creatures is hamstrung by an inability to identify just such properties. Closer to the work being critiqued by Webb, Beer's notion of a *minimally cognitive task* cannot be evaluated outside the context of a particular set of biological mechanisms (Fine, Di Paolo, & Izquierdo, 2007). Consider that a foraging task that requires "memory" in one agent might be solved without it in another agent that has access to real-time satellite photography of the foraging patch.

However, in appearing to favor animals over animats, Webb's title does (I believe accidentally) evoke a concerning tendency to begin to believe that we can build real instances of target systems, and thereby achieve direct empirical biological significance—that we can do "experiments" within simulations, that there can be a virtual laboratory, an in silico empiricism. This is to take Webb's approach too far (something that I do not believe she has ever done herself)—to claim that the synthetic methodology is or should be a route to empirical data on systems that are difficult or expensive to assay in vivo. While we should strive for models that have *relevance* to animals (or plants, fungi, etc.), either mechanistically or, as briefly argued here, in a more abstract manner, we should not confuse this with some practice of building models that *are* animals.

## References

Di Paolo, E. A., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. In M. A. Bedau, J. S. McCaskill, N. Packard, & S. Rasmussen (Eds.) *Proceedings ALIFE VII: The Seventh International Conference on the Simulation and Synthesis of Living Systems* (pp. 497–506). Cambridge MA: MIT Press.

Fine, P., Di Paolo, E. A., & Izquierdo, E. (2007). Adapting to your body. In F. Almeida e Costa et al. (Eds.), *Proceedings of the Ninth European Conference on Artificial Life* (pp. 203–211). Heidelberg: Springer.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.

Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie*, *20*, 410–433.