Animals versus animats: or why not the real iguana?

Barbara Webb

Institute for Perception, Action and Behaviour, School of Informatics,

University of Edinburgh

JCMB, Kings Buildings, Mayfield Rd, Edinburgh EH9 3JZ

bwebb@inf.ed.ac.uk

Abstract

The field of Artificial Life (ALife) has focussed attention on bottom-up explanations of adaptive behaviour, and how this can result from relatively simple systems interacting dynamically with their environments. Two differing approaches in developing these ideas are: to build models of simple sensorimotor systems in specific animals, assessed within complete behaviour-environments loops; and to explore the behaviour of invented artificial animals (often called 'animats') under similar conditions. But how can we learn about real biology from simulation of a non-existent animal? I will argue that most animat research, to the extent that it is relevant to biology, should also be considered as model building. Many claims are made for the animat approach: that it involves instantiation and is thus not simulation; that it uses idealised rather than realistic models; that the aim is theory exploration and existence proof rather than accounting for biological data; that the abstraction of the models makes the conclusions drawn from them more general. Few of these claims hold up under analysis, and it is suggested that we will learn more by choosing real, and not made-up, targets for our models.

1. Introduction

Is Artificial Life biological science? Langton (1989), who named the field, said ALife "complements the traditional biological sciences concerned with analysis of living organisms by attempting to synthesize life-like behaviors within computers and other artificial media" (p.1). Work within this wide field ranges from the clearly biological (e.g. simulated ecosystems; research into the chemical origins of DNA; exploration of the neural circuits underlying adaptive behaviour in specific animals) to areas more on the fringe (e.g. cellular automata; complexity and chaos; theoretical capabilities of artificial neural networks). What unites these areas is an interest in explaining the phenomena of life in terms of high level properties emerging from simple low level interactions. As Langton writes: "[ALife is] strongly based on themes such as *bottom-up* rather than *top-down* modeling, *local* rather than *global* control, *simple* rather than *complex* specifications, *emergent* rather than *pre-specified* behavior, *population* rather than *individual* simulation, and so forth" (op. cit.). The other obvious feature that distinguishes it from standard biological sciences is the 'artificial', i.e.,the use of synthetic methodologies to investigate these issues. "The field [of ALife] as a whole represents an attempt to vastly increase the role of synthesis in the study of biological phenomena" (op. cit.). One of the main questions I wish to explore here is whether and how "the role of synthesis" differs from the role of modelling in biology. Is ALife really a new

methodology for understanding real life? What criteria – if not that of model-building – should be used to assess its success as an approach to explaining how biological systems work?

To focus this discussion, I will draw examples mostly from one strand of ALife: where the questions addressed are how organisms or agents (singly or collectively) can be given, acquire or evolve sensorimotor capabilities that support interesting, potentially intelligent behaviour. This field is also often referred to as 'Adaptive Behaviour' (Meyer & Guillot, 1991). Braitenberg's 1984 book 'Vehicles' is one of the key inspirations. He describes a series of thought experiments in which increasingly intelligent (and life-like) vehicles are constructed, evolve and learn. At each stage, it is argued, surprisingly complex external behaviour will result from apparently simple internal structures. Simon (1969) famously used the example of an ant's irregular path on a beach to make a similar point, i.e., it is possible that the apparent complexity of the behaviour is the result of a simple internal system interacting with a complex environment. Another important stimulus in the field was the robots built in the1980s by Brooks and his colleagues (reviewed in Brooks, 1990). They demonstrated, in real devices rather than mere thought experiments, how having the right physical interfaces and relatively simple control mechanisms (eschewing the internal world modelling approaches of traditional AI robotics) could produce successful behaviours in real environmental contexts. W. Grey Walter's robots, built in the early 1950s, are rightly seen as forerunners of this work. These approaches meshed naturally with existing ideas from ecological psychology (Gibson, 1979) that were re-emerging in AI and cognitive science (e.g. Clancy, 1997) as an interest in active, situated and embodied aspects of cognition (e.g. Clark, 1997; Pfeifer & Scheier, 2000). However, the new ALife influence could be seen in the suggestion that investigation of these issues should start with relatively simple, yet complete, systems (such as insect-like reactive behaviours) and approach more complex capabilities (such as human intelligence) only incrementally.[1]

One obvious way to proceed towards these goals is to study real insects (or other simple biological systems) and to try to simulate them in a complete way, i.e., encompassing sensing and action, bottom-up control, embodiment, environmental interaction, and so on. For convenience in what follows, I am going to call this the 'animal' approach – as it involves trying to simulate some actual, existing, and well-identified animal system. An early example, which anticipates many ALife principles, is Arbib's work on modelling prey capture in toads (Arbib, 1982). My own work on robotic implementations of cricket phonotaxis (e.g. Webb, 1995) is another example, which I will describe more thoroughly below. I have reviewed similar work by other researchers in Webb (2000) and have argued that this is, essentially, a methodology of model building; and can thus usefully be compared to other kinds of biological modelling on a number of dimensions (Webb, 2001).

However another approach, common in ALife research, follows instead the suggestion made by Dennett (1978): "one does not want to get bogged down with technical problems in modeling the cognitive eccentricities of turtles if the point of the exercise is to uncover *very* general, *very* abstract

---

[1] "In building artificial creatures we might well make progress by starting with low expectations and gradually building upon our experience…One approach then, is to aim initially for the intelligence level of lower animals (e.g. insects) and respect the constraints that biology seems to have worked under." (Brooks, 1989). Brooks in fact did not follow his own advice in this regard, jumping in one step from insect-like robots to a humanoid robot project (Brooks and Stein, 1994).

principles… So why not then *make up* a whole cognitive creature, a Martian three-wheeled iguana, say, and an environmental niche for it to cope with?" (p.104). I will refer to this as the 'animat' approach – the invention and study of an artificial creature that does not correspond to any specific real animal, but is (nevertheless) intended to provide some insight into real issues in biological or cognitive science. The term 'animat' was first coined by Wilson (1985) who used an agent moving on a 18x58 hexagonal grid, able to distinguish three cell states (labelled 'food','tree' and 'empty'), to explore how classifier systems might be used to learn strategies for efficient gathering of 'food' and avoidance of 'trees'. In Wilson (1991) he further describes this approach to "the scientific understanding of intelligence, or how mind arises from brain" as "simulating and understanding complete animal-like systems at a simple level" (p.15). A more recent example is the work of Randall Beer (e.g. Beer, 1996) on evolving and analysing "minimally cognitive creatures", which I will describe in more detail below. There is again much similar work; many relevant examples can be found in the proceedings of *ALife* and *Simulation of Adaptive Behaviour* conferences, and in the journals *Artificial Life* and *Adaptive Behaviour*.[2] Is this also a model-building methodology, and can the same dimensions be used to characterise it?

To answer this question I will compare and contrast the two specific examples mentioned above. Hence section 2 will provide a short summary of each of these, with a brief description of some other examples of research that fall in the 'animal' or 'animat' category. I will then, in section 3, discuss how each approach fits into a model-building account of explanation. This comparison will, I hope, illuminate a number of interesting issues related to scientific modelling – such as the meaning of abstraction, idealisation, accuracy, generality and so on – that are not always well handled in either the ALife literature, or the philosophy of science.

2. Description of research approaches

2.1 'Animals'

In research over more than ten years (e.g. Webb 1995, Lund et al 1997, 1998, Webb & Scutt 2000, Reeve and Webb 2003, Reeve et al. 2005) we have been modelling the phonotaxis behaviour of crickets. Female crickets can recognise and locate the calling songs of male crickets. This target system was chosen for study due to the extensive behavioural and physiological data (reviewed in e.g. Huber & Thorson 1985; Hennig et al 2004) that could be fed into a model, and used to evaluate it. It was also plausible to investigate this behaviour as a complete loop from sensing to action, addressed at the level of single neurons and their connectivity. The behaviour depends on a well-tuned physical auditory apparatus. The two ears, on the cricket's forelegs, are linked by a tracheal tube, thus acting as pressure difference receivers to detect the direction of sounds of a specific frequency (Michelsen, 1994). There are also interesting emergent properties such as apparent 'choice' between songs that may be a simple consequence of interaction of the localisation mechanism and the complex sound field produced by multiple males singing. Phonotaxis behaviour interacts in interesting ways with other sensorimotor loops, such as the optomotor reflex (Böhm et al, 1991; Webb & Harrison, 2000), allowing the model to be gradually extended towards more complex capabilities.

---

[2] 'Animat' research can sometimes overlap with the field of 'agent' research but I am referring here only to 'agent' simulations that claim to be addressing problems relevant to biology, in particular, about the mechanisms of adaptive behaviour in real animals.

The more recent models we have built simulate spiking neuron activity using a leaky-integrate-and-fire representation (based on the approach described by Koch, 1998). Synaptic input is modelled as a change in conductance with a specific time course; the synaptic 'weight' is a function of the battery potential (positive for excitation, negative for inhibition), the size of conductance change occuring for each spike, and the time constant of conductance decay. This is a more realistic model than the more common treatment in integrate-and-fire neural modelling of synaptic input as an instaneous injection of current. The synapses also have some temporal dynamic properties, depression (decreased conductance with repeated spikes) and facilitation (increased conductance with repeated spikes), each with their own time course. However there is no permanent adaptation of weights (no learning). The networks consist of 4-20 units, and are generally hand-designed and hand-tuned, although we have explored the use of genetic algorithms to tune some parameters. Parts of the network are based on identified neurons in the cricket, for example the thoracic auditory interneurons known as AN1 and ON1 (Wohlers and Huber, 1982). Other parts are more speculative, either proposing specific connectivity between neurons that have been identified but not anatomically linked e.g. the brain neurons BNC1 and BNC2 (Schildberger 1984) or proposing new neural elements e.g. a 'fast' neuron that – based on recent cricket behavioural data (Hedwig & Poulet, 2004; Poulet & Hedwig, 2005) – may be involved in a rapid reflex response to sound pulses that is modulated by a slower recognition process.

These neural simulations have in most cases been tested in the context of a robot implementation. The robots (several different platforms have been used) have an analogue electronic auditory system to represent the physical auditory receptor mechanism of the cricket. The auditory input is translated into spike trains for the neural simulator, and the output of the simulator is used to drive the wheels (or in one case, 'whegs' – a hybrid wheel-leg mechanism (Horchler et al 2004)) of the robot. Thus the system can be tested in real experiments, using the same stimulus that is used for real crickets, to evaluate whether comparable behaviour can be produced.

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 behavioural 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 investigations. It has also made much more apparent the areas in which biological data is most critically lacking. One example is that the detailed dynamics of the cricket's response (how quickly it turns, the time constant of integration of signals into the response) has only recently been studied (Hedwig & Poulet, 2005) and there is still almost no information about the precise leg movements during phonotactic tracking that would be needed for a walking robot implementation.

There is much similar work that I would classify as part of the 'animal' approach, some of which I have described in previous reviews (Webb 2000, 2001). To give a few recent examples:

- Human walking behaviour is a subject that has been widely investigated using simulated and real robots (though many 'humanoid' robots do not have human-like walking). One interesting strand of this research is the investigation of passive walking machines. In these, the mechanics of human walking are represented in physical devices that mimic the hip and knee joints, the kneecap as an endstop, the foot, and sometimes the counter

swing of the arms e.g. the work of Collins et al, 2005 who say "We study human locomotion by designing and testing walking machines that we compare to humans in terms of morphology, gait appearance, energy use, and control"(p.1082). It has been demonstrated that convincingly human-like walking gaits emerge from the dynamics of these simple physical systems.

- The 'Psikharpax' project described in Meyer et al. (2005) includes simulations of rat brain mechanisms for navigation (based on hippocampal place-cell data), integrated with action selection (based on basal ganglia-thalamus-cortex loops) and learning (based on dopamine reinforcement mechanisms) and aims to implement these on a robot with appropriate sensorimotor capabilities. However it should be noted that Meyer uses the term 'animat' in the way I am here using the term 'animal' model – *"animats*, i.e., simulated animals or real robots whose sensors, actuators and control architectures are as closely inspired from those of animals as possible" (p.211) with an explicit aim in this project to "better understand the control mechanisms of rats"(p.221).

- The larvae collecting behaviour of the ant *Leptothorax albipennis* has been modelled using simulated and robot agents, to see which minimal set of action rules both most efficiently and most accurately reproduce the results. This can be used to infer what variables the ants themselves may be using to perform the task (Scholes et al 2004).

2.2 'Animats'

Over a similar time period as the cricket robot modelling described above, Beer and colleagues have been carrying out an investigation into "minimally cognitive agents" using simulated animats (Beer, 1996, 1997, Slocum et al 2000, Beer 2003a, Goldenberg et al 2004). The methodology is discussed in detail in Beer (2003a, 2003b) with a particular focus on how the agent behaviour can be analysed using methods from dynamical systems theory. The main task explored is the ability of an agent to actively categorise different stimuli, by moving to intercept items of one shape and to avoid items of another shape. Beer provides an explicitly (though perhaps not exclusively) biological motivation for this work as a "scientific approach to the neural basis of behaviour" (p.239). He is interested in discovering how small neural circuits can produce adaptive sensorimotor functioning, in the context of complete sensorimotor loops through the environment, and how this can be explained using a dynamical framework rather than a computational one.

The agents in these simulations have small continuous time recurrent neural networks (CTRNNs). The neural units are modelled as continuously varying values (interpretable as equivalent to the firing rates of real neurons) that decay with a set time constant. Synaptic input is the sum of the weighted activity, after passing through a logistic function, of all connected neurons, plus any external input. For the system described in Beer (2003a) the network has a fixed architecture consisting of 7 input neurons projecting to 5 fully interconnected interneurons, which project to two motor neurons. The weights of all the connections are determined using an evolutionary algorithm, evaluating each agent over a set of trials by its ability to minimise its final distance to one shape, and maximise the distance to the other.

This neural system is embedded in a simulated agent, which has seven sensors, evenly distributed across a $\pi/6$ 'field of view'. If an object intersects the axis of a sensor, the sensory neuron is

activated by an amount inversely proportional to the distance of the intersection from the agent. The sum of the activity of the two motor neurons determines the horizontal velocity of the agent. Objects fall from above, with a fixed velocity, and a random horizontal offset from the agent's starting position. A trial ends when the object hits or passes the agent. Agents that have successfully evolved to perform the task can be tested in experiments (e.g. with varying stimuli) to evaluate exactly how the agent is able to produce the observed behaviour. An important conclusion drawn is the need to understand this behaviour as emerging from the continuous interaction of neural, bodily and environmental factors, rather than treating it as a computational transformation of sensory input to motor output. To this end, dynamical analyses are applied at several levels: to the overall behaviour, to specific stimuli interactions, and to the functional role of specific neurons (Beer, 2003a).

It should be evident that this integrates a number of issues in sensing, neural circuits, adaptation and motor control. It proposes a rather novel hypothesis about the nature of cognitive capabilities, i.e.,that they involve such a tight interlinking of neuron-neuron, neuron-body, and body-environment dynamics that the standard kinds of computational decompositions used in investigating cognitive systems may be simply false. It also points out where there are substantial gaps in the methodological tools available for analysing such systems, and attempts to fill some of these gaps.

Some other examples that illustrate the scope of the animat approach include:

- Investigation of the behavioural consequences of evolved morphologies, e.g. to determine if there is an adaptive advantage to body symmetry for locomotion (Bongard and Paul, 2000). In this case the agents are simulated using a physics engine and consist of identical spherical units that can be connected to each other by hinge joints in six cardinal directions. The behaviour is controlled by a recurrent network of sigmoidal neurons that is evolved in concert with the morphology. The higher efficiency of the symmetric systems is taken to support the possibility that efficiency was a factor in the evolution of bilateral symmetry in biological systems.

- Study of the origins and evolution of communication and language (e.g. MacLennan 1992, MacLennan and Burghardt 1993). For example, this can involve a population of simulated agents able to have simple linguistic interactions, genetic interactions, and sensorimotor interactions with an environment which contains various entities that they come to label in a common manner. These systems are then studied for effects such as population size and distribution on the development of vocabulary and syntax. MacLennan and Burghardt (1993) describe this work as "synthetic ethology" in which "we do not attempt to model any existing natural system. Instead we construct "synthetic worlds" which are simple, but complete, and which manifest the phenomena of interest" (p.162) where "the goal is to use the synthetic approach to understand the natural phenomena" (p.163).

- The 'synthetic epistemology' approach explored in the Distributed Adaptive Control (DAC) architecture by Verschure and colleagues (e.g. Verschure et al, 2003). This falls somewhat closer to the 'animal' end of the continuum than the previous examples, in that it utilises models of learning based more directly on biological examples (e.g. classical

conditioning using a Hebb rule), and some analogies to vertebrate brain structures are described. But the model "is not inspired by the anatomy, physiology or neuropsychology of hippocampus and cortex, but addresses the general problem of communication between different neural structures in the context of behaviorally realistic tasks and well-evaluated models of learning" (Voegtlin and Verschure, 1999 (reprint 34/55). The architecture is tested in an animat, which is not designed to represent any specific animal, that performs 'foraging tasks' to locate targets and avoid obstacles. The behaviour is assessed in terms of efficiency and stability of learning, as aspects of the architecture are varied, rather than compared directly to learning or foraging data from biology.

3. Comparing the two approaches

It should evident from the above descriptions that there are many ways in which the two main examples I have described – my 'cricket robot' and Beer's 'minimally cognitive agent' – are similar. The behavioural competence studied is closely comparable: the robot, like Beer's agent, has to discriminate the correct stimulus (cricket songs with the right temporal pattern) from the wrong stimulus (other sounds) and move towards the former. Both attempt to account for behaviour at the same level of description, that is, examining the functional roles of individual neurons in small circuits for an agent in a particular action-environment context, while not addressing lower level mechanisms of neural activity such as channel dynamics or dendritic branching. Many animat models are higher level, controlling the agent's behaviour with a set of rules rather than a neural network e.g. Tyrrell and Mayhew (1991); but then some animal models are described at a similar level, e.g. the robot model of lobster chemotaxis in Grasso et al (2000); so it is not the level of explanation *per se* that distinguishes the two approaches.

One seemingly conspicuous contrast is that the cricket model is embodied in a robot, whereas the minimally cognitive agent is embodied in a computer simulation. However, despite the energetic debates that have taken place over this issue, I do not regard it as important in the current instance. The 'animat' approach can be, and often is, pursued using real robots, and the 'animal' approach is often pursued using computer simulations. For example, in work closely related to that of Beer, Husbands et al (1995) evolved a small neural network to control an agent that would approach triangles and not squares, and carried out a dynamical analysis of the behaviour. But in this case the agent was embodied using a gantry robot system, required to make a real visual discrimination. Steels (2003) has investigated language evolution in agent systems, where the fact that the agents are real robots acting in the real world is considered critical to the methodology. The work by Verschure (2003) mentioned above has been implemented both in simulations and real robots. Similarly, in work that is closely comparable to the cricket model, Neumann & Bulthoff (2002) have carried out some intensive simulations, using virtual reality environments, of visual control of motion in the fly. In Neumann et al (2001) they argue that the experiments possible in simulation are in fact more realistic than any plausible robotic implementation could be. Sometimes the stimulus used in a robotic experiment is simpler than in a simulation. Compare the use of just four identical landmarks on the environmental boundary for an analogue robot implementing the average landmark vector model for homing based on desert ant behaviour (Möller, 2000), to the field of different size landmarks randomly distributed in depth used in simulation testing of the

same algorithm (Lambrinos et al 2000). Choice of a physical (robot) model, and the amount of environmental complexity experienced by the agent, are independent factors. And either real or simulated environments, and either highly simplified or richly complex environmental interactions, might be used in testing animal or animat models.

So what are the key distinguishing factors? The most obvious difference is that Beer's agent is described as an idealised model of a general cognitive agent; whereas the cricket robot attempts to be a realistic model of a specific animal. But what do 'idealised' and 'realistic' actually mean in this context? How much does an idealised model differ in detail or accuracy from a realistic one? What exactly is more 'general' about Beer's agent categorising shapes than the cricket robot recognising songs? Are idealisations models at all? Are animats better treated as alternative creatures that can be studied independently of their match to biology? I will try to explore these questions in what follows.

3.1 Are animats models?

The term 'model' is notoriously inexact[3] although most people agree it basically involves one thing being used to represent another. In Webb (2001) I approached the "model muddle" (Wartofsky, 1979) by describing the process of scientific reasoning into which modelling enters. This first involves selecting a *target* phenomenon in the world to be explained – where a specific target might be considered a 'model' for a class of phenomena. A hypothesis or causal account is then offered as an explanation of the phenomenon – this may be called a 'theoretical model', which represents the target in the same way that any language description represents a state of affairs (see also Callendar and Cohen, 2005) although the 'language' used may be more precise than natural language. The hypothesis may draw on an analogy to some other, presumably better understood, system – often called a 'source' model. The extent to which the hypothesis accounts for existing data and predicts new data from observations on the target phenomenon is taken to support its status as an explanation[4]. These observations may include both the overt behaviour of the system and any structural components and relations that are revealed by 'dissection' of the system. For a moderately complex hypothesis, demonstrating what data it can account for and will predict may require implementation in a mathematical, computational or even physical model, which represents (usually in some different medium) the mechanisms and causes described in the hypothesis, and can be solved, run or tested to see what their consequences are.

It is reasonably clear to see how 'animal' models such as the cricket robot fit into this schema – they are physical and/or computational implementations of particular hypotheses about the phenomenon of sensorimotor behaviours in certain animals, producing results that can be directly compared to observations on the animals. In Webb (2001) I described this as the 'relevance' of a model, with the intent to distinguish systems that might appear similar (e.g. my cricket robot and the sound localising robot described in Andersson et al 2004) but differ in their intent (to explain a biological system vs. solving a technical problem using some ideas from biology). 'Biological inspiration' differs from biological modelling because the animal is not the target of the model building but rather the source model or analogy. The target is instead some technical problem, and

---

[3] Winsberg (1999) cites Goodman, 1968 as noting that a model is "almost anything from a naked blonde to a quadratic equation".

[4] Other factors, such as simplicity, or coherence with other theories, may also enter into determining this status, in particular if there exist alternative hypotheses also able to account for the data.

the hypothesis is a proposed solution to that problem that draws on the biological system. Another example would be the use of 'ant colony' algorithms to solve communication routing problems (Dorigo and Stützle, 2004) compared to models of ant colony behaviour e.g. Calenbuhr and Deneubourg (1990). Note, however, that even when the target is no longer biological, engineering still often involves the implementation of models, using mathematics, computer simulations and scale models to make predictions about how a hypothetical system, if constructed, would behave; or indeed to predict the effects of manipulations on a system that has been constructed.

Although some research in ALife and animat modelling has an engineering purpose (many examples are reviewed in Kim and Cho, 2006), the basic aims as described by Langton, Beer and others are not technological but scientific. Several different attitudes to the intended 'relevance' of these systems can be discerned. One view is that animat and other ALife models are simply worth exploring in their own right, as analytical systems, loosely based on biological principles (e.g. replication with variation) that have interesting properties. Much work in evolutionary algorithms has this character. Such work may turn out to be relevant to biology, but interpretation of the systems as models that represent a particular biological problem is not considered necessary to justify the ongoing research programme. Whether such exploration is science is debatable (is pure mathematics science?) but it seems reasonable to say that until or unless these explorations are used to make empirical claims about real life they are not biological science. The relevance is at best undecided. Taylor describes similar models in ecology as "exploratory tools" that do not need to be justified by either fitting to data or plausibility of their assumptions:

"It can be explored systematically as a mathematical system, e.g. how does the system's behaviour change as its parameters change or become variables, as time lags are added, and so on? Such mathematical investigations may help us derive new questions to ask, new terms to employ, or different models to construct." (p.201)

However, he notes:

"Strictly speaking, without a quantitative analysis of correspondence [to observations], the insights from exploration are insights about a mathematical system. Their relevance to biology is yet-to-be-established; truth or falsity is a moot point…Qualitative insights might have misguided research. The categories of exploratory models, often chosen with an eye to mathematical tractability, may have obscured profound issues about biology."(op. cit.)

A second view, expressed by Miller (1995) and Noble (1997) for example, is that ALife could be considered a branch of theoretical biology, and therefore needs to be explicit about the biological targets being represented, the empirical predictions being made as a result of the modelling, and the results of comparing these predictions to data from biology. These authors are generally critical of the failure of much work within ALife to meet these standards, in particular, work that starts out by claiming to represent some biological target, but ends up exploring the resulting system for its own sake, and fails to make any substantive empirical claims or comparisons to biological data. Clearly, work that does meet these standards should be considered relevant, but then it also clearly fits the schema of scientific explanation in just the same way as the cricket robot does.

Langton's well-known view of ALife as an 'alternative' biology provides a different twist. Clearly he believes the systems to be relevant, but at the same time he does not regard them as straightforward simulations of existing biological systems, or implementations of specific

biological hypotheses. Rather, he appears to be arguing for them to be used as source models, i.e.,analogical systems that will help in constructing novel hypotheses for biology. This is intended to proceed through a comparative approach, between the artificial and natural examples of life, to discover general principles beyond the contingent details of existing biology. Similarly, Steels (2001) suggests:

> "we may want to compare the behaviour and mechanisms of artificial systems to that of natural systems achieving the same functionality. Comparing is not the same as mimicking or modelling the natural system ... it is a way to gain insight by confronting it with something that is different but still sufficiently similar to make the comparison interesting." (p.1078)

What normally characterises a 'source' model (or the comparative approach to constructing hypotheses) is that system being compared to the target exists, or was built, independently of the hypothesis; for example, pumps existed before they were proposed as a good model for the functioning of the heart. For some areas of ALife this might be the case, e.g. cellular automata were developed and have been explored independently of biology, but have become a popular analogy for describing how interesting global patterns might arise from local rules in biological systems, e.g. as discussed at length by Dennett (1998). Similarly, an ALife system built for an engineering purpose might subsequently be usefully compared to a biological system that carries out a similar function. However, in the general area of animat models, and in work such as that of Beer, the systems are not usually constructed independently, but are constructed to represent (however loosely or abstractly) some mechanisms taken, by hypothesis, to have causal relevance to biology. Consequently it seems more appropriate to consider them – to the extent that they are claimed to be relevant – as models that implement a hypothesis rather than as source models. Note that, either way, establishing the relevance requires some explicit specification of how the two systems – artificial and biological – are supposed to correspond.

Some 'animat' researchers explicitly claim that their systems are not simulations or models. For example MacLennan and Berghardt (1993) state:

> "The techniques of synthetic ethology must be carefully distinguished from simulation or mathematical modeling. In a simulation, an attempt is made to imitate in a computer or other modeling system the salient aspects of a complex situation that exists, at least potentially, in the real world." (p.161)

Yet for synthetic ethology:

> "The goal…is to create [in the computer] synthetic worlds that are like the natural world in relevant ways…" (MacLennan, 2002, p.152)

How 'imitating salient aspects' (simulation) is supposed to differ fundamentally from 'creating things that are like in relevant ways' (synthetic methodology) is not immediately obvious. The distinction sometimes rests on the argument that (at least some) ALife systems are *realisations* of life and therefore not models. Ray (1994) for example also contrasts "AL simulations" that represent some natural phenomenon with "AL instantiations" which are "living forms in their own right, and are not models of any natural lifeform"; Langton uses the phrase "simulate *or* synthesize"[emphasis added]. This has generated much philosophical speculation about whether these systems are really 'alive'; a debate that many recognise to be similar to the debate over 'strong' vs. 'weak' Artificial Intelligence. However, irrespective of the reality of Artificial Life, it is essential to note that

'simulation' and 'realisation' are not in fact exclusive categories. It is a mistake to argue that because a system is a simulation, it therefore cannot be an instantiation (Webb, 1991); and equally mistaken to argue that because a system is an instantiation, it is therefore not a model. An example is provided by the cricket robot –it does real phonotaxis, but it is also a simulation of phonotaxis in crickets.

Beer (2003b) himself is somewhat equivocal regarding the intended relevance of his model (though he does call it a model). For example he says "The intention here is not to propose a serious model of categorical perception" (p. 210) and yet "The analysis described in this paper also has important implications for our understanding of perception" (p.236), i.e., empirical claims *are* being made about the world on the basis of the model results. Arguing that "the intent of idealized models is not empirical prediction, but conceptual clarification" (p.240) and that "many of the details of an idealized model's analysis will be tied to the particular decisions made in its formulation, and thus be of questionable direct relevance to any real cognitive system" (p.241), it seems he regards the model more in the mode of potential 'source' analogies (similar to Langton's views as described above). Yet his system is not theory-independent, but has been specifically designed to support the assertion that "dynamical analyses of cognitive behavior…can…broaden and clarify our understanding of the mechanisms of cognition" (p.241). To draw any such conclusion *relies* on the fact that the agent and task have been intentionally built to represent real cognitive agents and tasks. Beer recognises this in arguing for the biological and cognitive relevance of his agent's design and behaviour, for example when he defends the choice of CTRNNs as follows:

> "from a scientific (as opposed to an engineering) point of view, we are not free to choose models based solely on the convenience of analysis. Rather, we a are faced with understanding the particular brain-body-environment systems that Nature has presented to us" (2003b, p.302).

That is, the task, environment and structure of the invented creature are taken to be representative of some class of real biological problem. Beer (2003) makes a direct connection between his work and idealised models such as 'frictionless planes' used in theoretical physics. There are some definitions of 'scientific model' that a frictionless plane would fail to satisfy, but this is usually regarded as a problem for those definitions, not for the status of frictionless planes as examples of model-use in physics (Suarez, 2003).

3.2 Does animat behaviour need to match biology?

If animat researchers are not making any claims of relevance, but are justifying their work either as engineering or formal analysis, then there is no reason to require empirical validation. However, insofar as the animat is being used as a model, i.e., to represent and draw conclusions about some aspect of biology, then it seems reasonable to ask how well its behaviour matches that of the target phenomenon. But if the 'target' is a Martian three-wheeled iguana, how are we supposed to proceed? It should be evident from the previous discussion that despite the 'invented' nature of animats, researchers usually have in mind some real phenomenon that they wish to account for. An example used by Beer (2003a) is human categorical perception, in particular, the sharp perceptual boundary between phonemes that differ on a continuous dimension (Ehret, 1987). Beer shows that his agent produces the typically observed sigmoidal curve for 'labelling' (in this case, for 'mean catch performance') of the stimulus as it varies smoothly from a circle to a triangle.

Yet there remains ambivalence in much animat research towards making such direct comparisons. Although Beer refers to a particular data set, he does not consider anomalies between that data set and his results to be important. Harvey et al (2005), describing the animat approach in evolutionary robotics, say "it is not necessary … to aim at modelling specific animal behaviour so as to validate the model with empirical data" (p. 84). This is justified by the suggestion that animat simulations are meant to provide existence proofs, or proofs of concept, rather than account for data. Harvey et al argue that "We will have demonstrated cognitive capacity or phenomenon X under clearly specified conditions and constraints, hence showing that these provide sufficient conditions for X… the production of proofs of concept has a different scientific function [from empirical validation]: that of catalysing theoretical re-conceptualizations and facilitating the production of novel hypotheses" (p. 84). A researcher could thus, for example, refute a claim that phenomenon X requires condition Y, e.g. that rats demonstrating confusion between similar corners of a space require an internal geometrical representation of that space, by showing that an animat makes the same confusion without having such a representation. Steels (2001) argues that "We may want to understand how a particular functionality can be achieved at all, and thus comprehend a mystery that nature somehow solved" (p.1078).

However, existence proofs clearly do require comparisons between model results and empirical data. One cannot evaluate the claim that phenomenon X requires condition Y unless one can show that phenomenon X actually is produced (with or without Y). And the claim or proof will be stronger or weaker depending on how well the simulated X matches the real X. Harvey et al (2005) indeed recognise that for 'proofs of concept' to be relevant *does* require an attempt at validation with empirical data – the "novel hypotheses…then need to be appropriately translated to domain-specific cases and tested empirically" (p. 84). Yet the animat conception of an invented animal often seems like a convenient way to put off this testing indefinitely; by contrast, an 'animal' model will already have one domain specific translation, and the corresponding data set, immediately to hand. Note that this does not mean a novel hypothesis produced in simulating a particular target animal behaviour cannot subsequently be translated into other domains, and thus also evaluated for its more general application. The Hodgkin-Huxley model of neural dynamics, first developed to account for data from the squid giant axon, is a notable example.

The issue here is not about requiring a perfect match between the model and the data, or having to reproduce all the "cognitive eccentricities" of a particular animal. The cricket robot, by and large, provides only a qualitative match to real cricket data. For example, the discrimination or preference curve for songs of different syllable rates for the robot has the same characteristic band-pass shape, but it is not value by value identical to the curve for the cricket. A criteria of exact value matching (as opposed to matching the distribution, or the point averages, or the overall pattern of the data – Carley 1997) would be one that few models in biology could meet. The match between the robot cricket and the real cricket data might be improved by very careful parameter tuning. Conceivably, the discrimination of Beer's agent might also be tuned to more precisely match some particular results on human phoneme classification, e.g. by using similarity to the human discrimination curve as the fitness function in the evolution process. But this is not necessarily informative: with enough free parameters and enough tuning, any data can be matched. The issue is that, lacking an *a priori*

target, animat modellers can choose *post hoc* to note any biological data that seems to match their results while ignoring any data that does not.

Moreover, for an existence proof to be interesting, there needs to have been some claim or general belief that phenomenon X was not possible without condition Y, for example that cricket song preference behaviour required separate sound recognition and sound localisation circuitry (Huber and Thorson, 1985) which the cricket robot demonstrated was not the case (Webb, 1995). But we do not lack for existence proofs that categorisation behaviour such as that of Beer's agent can be achieved using a simple neural network in an appropriately embodied agent interacting with the environment, without explicit representation. There were already a number of examples of simple robot systems, capable of approaching certain visual targets while avoiding other stimuli, some based specifically on the 'existence proofs' provided by insects, e.g. Srinivasan and Venkatesh, 1997. Indeed, the cricket robot itself has been used by several philosophers to illustrate exactly this point about the capabilities of such agents (e.g. Clark, 2001). And those who argue that these kinds of systems *are* in fact using representations (e.g. Bechtel, 1998; Mandik, 2002) do not find Beer's agent a convincing counterexample; the disagreement is not over how the task is actually accomplished in these simple systems but over the definition and application of the term 'representation' in describing such mechanisms.

Rather than necessarily claiming his agent is a novel existence proof for cognition without representation, Beer emphasises instead his analysis as an 'existence proof' of how a dynamical account of cognition can be made. The development and application of dynamical tools is indeed an interesting contribution, but the question of its relevance to biology remains dependent on there being some mapping between his agent's behaviour and the behaviour of real cognitive systems. Beer argues that he uses a minimalist invented animat because such an analysis is not yet feasible to apply to any real system, which would be far too complex – "Only in idealized model agents can we presently hope to study the dynamics of a complete agent-environment system and thus clarify the fundamental nature of such systems." (p. 240). But why could this not be attempted using a highly simplified or minimalist model of a real animal, rather than an invented one? The assumption seems to be that choosing a real target immediately implies one must build a complex, detailed, accurate, low level model; but this simply does not follow.

3.3 Inaccurate, abstract and ideal models

In the previous section I argued that unless, at some point, an animat model can be shown to account for or match some real and significant biological data, it is hard to argue convincingly that it is relevant to biology. But this is not the same as requiring the implemented model to be a detailed, structurally accurate representation of the mechanism producing this behaviour in the biological system. When Beer says he has not built a 'serious model', what he appears to mean is not that it fails to produce categorisation behaviour that can be compared to biological examples such as phoneme discrimination, but that his detailed analysis of how the agent produces the behaviour is not being offered as an explanation of phoneme perception in humans (or only in the very indirect sense of suggesting that a non-symbollic dynamical account of this aspect of perception is a possibility that should be taken seriously). Similarly, Harvey et al (2005) note that the claim of an animat model to have demonstrated that phenomena X can be produced under certain conditions is independent of any claim that "this gives us direct insights into the actual

physical mechanisms of real biological organisms"(p.84). Clearly it is going to be difficult to assess the accuracy of animat models in the same way as might be done for animal models, e.g. by looking at the mapping between neural units in the model and identified neurons in the animal, if the target is a Martian three-wheeled iguana. But as for match, the components and processes in animat models do have some intended resemblance to real biological components and processes, and can at least be assessed for plausibility. However, though such comparison is useful, I will argue that it is not actually necessary for a model to be accurate to be relevant – indeed, many animal models are no more accurate or detailed than their animat counterparts.

It is useful first to distinguish the issue of accuracy from that of abstraction. Perfect accuracy might be thought of as a one-to-one mapping or isomorphism between the elements and structure of the model and those of the target system. This mapping could be defined at different levels (e.g. a mapping of control laws, or of brain structures, or of individual neurons). A model that, at some particular level, includes less detail[5] could be thought of as a partial mapping (i.e.,not all elements in the system are mapped to elements of the model) or as a many-to-one mapping (e.g. several steps of a process are mapped to one process in the model). I would consider both of these to be changes in abstraction rather than accuracy: a decrease in accuracy should be associated with a *failure* of mapping. This may take the form of approximation, using a value that is known to be inaccurate but considered 'close enough', or pushing a value to limit known to be impossible in reality, such as zero friction. It may involve substitution of one process with another that has an input-output relation that is 'similar enough' e.g. using a linear function for a monotonic non-linear relation. It may be a case of having elements in the model that are not intended to map to anything in the system, but need to be there to make it run. It may also involve intentional decision to contradict some known facts about the system, either for the sake of creating a more tractable model, or indeed to directly test whether those particular facts are critical or not.

'Idealisation' is sometimes associated with abstraction, and sometimes with inaccuracy. Frigg and Hartmann (2006) distinguish 'Aristotelean' idealisation, in which properties are stripped away, from 'Galilean' idealisation, in which there is deliberate distortion. Pincock (2004) suggests an idealised model "satisfies two conditions i) there is no isomorphism relating it to the situation that it puportedly represents and ii) the relevant agents [meaning the model's builders or users] are aware of this". But models can, of course, be inaccurate without being idealised. Perhaps a more useful characterisation of idealisation is in how the model is arrived at, and how this differs from simplification. In a process of simplification, variables or details are gradually excluded, or combined. For example, in running an experiment on real crickets, the environment is usually highly simplified, having flat terrain, no visual stimuli, only one or two auditory cues, and the cues themselves having much of the natural variation removed. Moreover the crickets themselves are usually pre-selected to minimise variation, using one species, at a certain age, in good condition etc., and the reported behaviours are combined across individuals and trials. In this case, a

---

[5] I distinguish 'level' as corresponding to the general principles that apply to all systems of that type (e.g. the biochemistry of membrane permeability; the activation properties of synaptic connections) from 'detail' which corresponds to the specific structures at that level in a particular system (e.g. the types and distributions of ion channels found at certain synapse; the connectivity between neurons in a particular circuit). As is familiar to physicists from the three-body problem, or to geneticists trying to unravel the human genome, knowledge of the general principles of operation of a system (Newtonian physics, the protein coding system of DNA) is not the same as having a detailed model of it at that level.

simplified model of the 'normal' behavioural situation and response has been produced by processes of elimination and combination. By contrast, idealisation can be thought of as a process of construction: starting from nothing, variables are gradually introduced, preferably in a mathematically tractable form (e.g. linear response functions, gaussian noise) with the intention of creating the simplest system that might adequately represent the problem. Both processes are likely to result in models that are both abstract and inaccurate.

The idea that we can learn something from models that we know do not (or doubt, or would hesitate to claim) work the same way as the target system is not a new one. Wimsatt (1987) for example discusses the issue in terms of "False models as a means to truer theories", listing twelve uses for false models that include: acting as a starting point for developing better models; undercutting a preferred hypothesis by indicating an alternative line of explanation; as a neutral baseline for assessing causal claims in other models; as a phenomenological fitting of data that might suggest an underlying mechanism, or indicate which factors have most predictive power; by comparing which results hold across a family of models that have differing false assumptions. Frigg (2003) adds several points, such as how we may learn during the building process, forcing us to see the target system from a new point of view; and that features may be highlighted by isolation or exaggeration. Two uses for 'unrealistic models' that are often stressed in ALife are that we can search for the minimum set of components that still produces something like the phenomenon of interest, and that this might be found by some automatic search process, typically a genetic algorithm, to generate solutions that are relatively unbiased by the modeller's theoretical preconceptions (Bedau, 1999). Inaccurate and counterfactual models can often be very revealing about the scope of the problem to be explained. A model that incorporates all current structural knowledge (i.e., that we think is accurate) yet fails to match the data can be similarly informative, in suggesting critical gaps in understanding.

All kinds of abstractions and inaccuracies are commonly found in 'animal' models, and indeed in most (perhaps all) scientific models. It is misleading to characterise a 'realistic' approach to modelling as "[striving] for maximal fidelity to the details of particular natural systems" (Bedau, 1999, p.20). The cricket robot substantially simplifies and distorts in its representation of real crickets. This is perhaps most obvious in the motor output, where two wheels substitute for six legs. It also has one auditory transducer on each side, compared to the cricket's 50 or so receptor cells, and these transducers differ in a number of other respects. The neural circuit does not include a number of identified auditory neurons that may influence phonotaxis (such as the thoracic neuron pair labelled 'AN2', Schildberger and Horner, 1988). But there is a difference between being *aware* of inaccuracy, and being *unable to assess* the accuracy because the target for the model has been arbitrarily invented by the animat modeller. Steels (2001) in discussing 'formal' (idealised?) models suggests that they "do not necessarily describe a natural system" and approvingly quotes Milton Friedman: : "to be important. ... a hypothesis must be descriptively false in its assumptions". But an assumption about a non-existent system cannot be true or false: what Steels appears to mean is that formal models do not (necessarily) *accurately* describe a natural system – but nor do most other kinds of models.

3.4 Generality

The argument for the relevance of animat models seems to be closely tied to the claim that what is represented by the model is some 'general' problem of biology or cognition, and that the model will thus provide some fundamental conceptual insights. Indeed, this idea is explicit in Dennett's original (1978) suggestion to "*make up* a whole cognitive creature" in order "to uncover *very* general, *very* abstract principles". Generality is sometimes contrasted to being definite or precise, for example when we talk about something in 'general terms', meaning we are using rough approximations. But this meaning of general does not seem to apply to animat models which usually *are* precisely defined, with the exact properties of inputs and outputs and the equations of transformations fully described, as they need to be for the simulated or physically implemented system to be run. On the other hand, 'generality' can be contrasted to specificity; when we talk about the properties common to a range of systems, rather than those specific to only one system. Unlike the previous definition this explicitly invokes the idea of a description that covers multiple systems; you cannot talk about a 'general' description of only one system. This idea of generality is more obviously the one that is valued in science. To be able to represent multiple systems with one model implies unification of explanation, a widely accepted criteria for a good scientific theory.

There is an obvious way in which generality defined as what is common to a set of systems can be related to detail: the subset of properties common to a set of systems must necessarily be no larger than the set of properties of any one of those systems, and is most likely to be smaller; hence the general description will usually contain less detail. But although a general description might have fewer details, and thus be more abstract, the opposite is not automatically true: an abstract model, containing few details, is not necessarily general. It may still only describe one system, or indeed no existing system at all. However, it is extremely common in all fields of modelling to see the mistaken claim that abstraction directly implies generality, with all its virtues: e.g. Bedau (1999) in justifying simple ALife models says "the more abstract the model, the more broadly it applies, so the more unified the explanation it provides" (p.20).

This relationship of generality to detail (i.e.,more generality usually implies less detail, but not the reverse) also underlines the distinction between 'detail' and 'level'. Lower level principles in biology are more general: biochemistry encompasses a larger range of systems than neuroscience; and neuroscience a larger range than cognitive science. On the other hand, the details of a specific biochemical pathway may not be general to more than one species. So why is generality so often associated to 'high-level' descriptions? Partly this is because 'high-level' is sometimes used simply to mean 'less detailed' or 'approximate' – it is common to see the phrase 'level of abstraction' used in discussions of models, e.g. Bedau (1999). Lower level models often *are* elaborations of higher level ones and hence do contain more detail, but this is not always the case. Neural net models of perceptual classification were interesting precisely because they seemed much simpler than the higher-level rule-based models that had been previously proposed: the lower level was *not* just implementation detail but an entirely different, simpler, and more abstract, principle of operation. Much of the interest in ALife comes from just this kind of explanation in which some abstract low-level principle accounts for high-level complexity. Even if the lower level model adds detail, this does not necessarily make it less general, if these details are still common to all systems – e.g. the neural level cricket model applies to just as many cricket species as the algorithmic version did.

However, a more interesting association of 'high-level' and 'generality' is the idea that we can find general principles of operation turning up at higher levels of description *despite* the intervening levels having different principles of operation. We assume both robots and animals are governed at the lowest level under the same general laws of physics; general principles at the level of biochemistry do not apply to robots; yet we assume the principles governing networks of neuron-like units are the same for computer simulations and for real neurons. For example, we assume that there is nothing at the level of biochemistry that is essential to performing 'real' phonotaxis. This idea, of higher level capabilities that can be multiply realised in systems that differ in their lower level properties, is of course central to ALife (and to Artificial Intelligence and Cybernetics before it). "Of course, the principle assumption made in Artificial Life is that the 'logical form' of an organism can be separated from its material basis of construction, and that 'aliveness' will be found to be a property of the former, not of the latter." [Langton, 1989, page 11]. It also supplies a motivation for the animat approach. As spelt out by Langton, the idea is that by inventing or artificially constructing novel examples of the phenomenon of interest ('life' or 'adaptive behaviour') we discover the general (high-level) laws of life-as-it-could-be rather than the contingent (low-level) specifics of life-as-it-is. The problem remains, however, of how to decide that the invented system *is* an example of the same phenomenon, if we lack these general laws. This issue does not seem too problematic for phonotaxis, but is hotly debated for 'life' and 'cognition'. Moreover, as Keeley (2000) usefully points out, there seems to be an inherent circularity in which we decide the taxonomy of a phenomenon without reference to lower levels, and then take this as proof that the lower levels are irrelevant to the phenomenon. It remains a possibility that to replicate life – or cognition – requires the right artificial chemistry, for example. Discovering principles relevant to all life is a laudable aim, but these may turn out to be contingent after all. It should be noted, then, that it *is* possible to aim for general explanations (e.g. of all existing examples of life or cognition) without assuming multirealisability (i.e.,that life could also occur with some different material basis).

In fact, Beer does not explicitly claim his 'idealised' model is a general one. Rather, he says it is put forward as a specific example of a 'minimally cognitive system', and that close examination of this specific system will be a route by which general principles about cognition may emerge. Just the same motivation is given (e.g. Reeve and Webb, 2003) for the 'realistic' cricket model: phonotaxis is being studied in detail not simply to understand more about crickets, but because it represents many general problems of sensorimotor control. Indeed, these are largely the same problems that interest Beer – such as the interaction of body, task and environment, the role of individual neurons in behaviour, what minimal complexity of behaviour requires cognitive capabilities and so on. The only difference in generality of the two kinds of models, it seems, is that the animal model can be shown to apply to at least one class of real systems (crickets), whereas the animat model is not shown to explain any.

4 Conclusion

I have argued that Artificial Life simulations and animats, if they are to be relevant to biology need to be considered as models. That they are abstract, approximate and not necessarily accurate representations, useful for exploring principles, providing existence proofs and generating novel hypotheses, does not remove the requirement to demonstrate that they can usefully account for

observations made on real biological systems. Otherwise they can be considered only as explorations of artificial systems – perhaps interesting in their own right, or with potential technological applications – but not part of biology.

It is noticeable that many of the arguments for the animat approach are based on 'making the best the enemy of the good'. Thus, it is often argued that because it is impossible to provide detailed, accurate and complete models of any real animal system, the only alternative is to invent artificial creatures for study. But this ignores the option of building abstracted and inaccurate models of real animals. As I have argued here, there is little difference between my cricket robot and Beer's minimally cognitive agent on most of the dimensions of modelling. The level of description is much the same, i.e.,neural activity supporting discriminatory motor behaviour. The robot model abstracts and distorts many elements and processes in the real cricket, and only qualitatively matches cricket data. It is assumed by both models that general insights into sensorimotor control, relevant to biology, can be obtained from these specific examples. Although the medium (robot or computer) differs, it is not hard to imagine a purely computational version of the cricket model, or a robotic implementation of Beer's agent. The cricket robot provided a novel view of, and new hypotheses for, the perceptual task performed by the animal, and an existence proof that no explicitly representational processing was needed for this "minimally cognitive" task[6].

The only apparent advantage in not grounding one's model by reference to some real target system (i.e.,for using animat rather animal models) is that the model is thus immune to criticism of the validity of the representation. The price paid for this immunity, however, is to undermine the justification for its relevance. It is often forgotten in discussions of Braitenberg's (1984) thought experiments about 'hypothetical' vehicles that he spends the second half of his book explaining the biological relevance of each of the features with which they have been endowed. To do so requires reference to real iguanas, not the Martian three-wheeled variety.

References

Andersson, S.B., Handzel, A.A., Shah, V., and Krishnaprasad, P.S. 2004. Robot Phonotaxis with Dynamic Sound Localization, *Proceeding IEEE International Conference on Robotics and Automation (*pp. 4833-4838)

Arbib, M. 1982. Rana Computatrix: An Evolving Model of Visuo-motor Coordination in Frog and Toad. In J.E. Hayes, D. Michie, and Y.H. Pao (eds.) *Machine Intelligence 10* (pp. 501-517) Chichester: Ellis Horwood Publishers.

Bechtel, W. 1998. Representations and Cognitive Explanations: Assessing the Dynamicist's Challenge in Cognitive Science. *Cognitive Science*, 22:295-318.

Bedau, M. A. 1999. Can Unrealistic Computer Models Illuminate Theoretical Biology? In A. Wu, (ed.) P*roceedings of the 1999 Genetic and Evolutionary Computation Conference (*pp. 20-23). Orlando: GECCO.

---

[6]In fact, I do not wish to claim this task is 'cognitive', but it is no less 'cognitive' than the task performed by Beer's agent.

Beer, R.D. 1996. Toward the evolution of dynamical neural networks for minimally cognitive behavior. In P. Maes, M. Mataric, J.A. Meyer, J. Pollack and S. Wilson (eds.), *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior* (pp. 421-429). Cambridge MA: MIT Press.

Beer, R.D. 1997. The dynamics of adaptive behavior: A research program. *Robotics and Autonomous Systems,* 20:257-289.

Beer, R.D. 2003a. The dynamics of active categorical perception in an evolved model agent *Adaptive Behavior,* 11(4):209-243.

Beer, R.D. 2003b. Arches and stones in cognitive architecture. *Adaptive Behavior,* 11(4):299-305.

Böhm, H., Schildberger, K., Huber, F. 1991. Visual and acoustic course control in the cricket *Gryllus-bimaculatus*. *Journal of Experimental Biology,* 159:235-248.

Bongard, J. and 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.*

Braitenberg, V. 1984. *Vehicles: Experiments in Synthetic Psychology.* Cambridge MA*:* MIT Press.

Brooks, R. A. 1989. How To Build Complete Creatures Rather Than Isolated Cognitive Simulators. In K. VanLehn (ed.) *Architectures for Intelligence* (pp. 225–239), Erlbaum: Hillsdale, NJ.

Brooks, R. A., 1990. Elephants Don't Play Chess. *Robotics and Autonomous Systems,* 6:3–15.

Brooks, R. A., and Stein, L. A. 1994. Building Brains for Bodies, *Autonomous Robotics,* 1 (1):7-25
Calenbuhr, V. & Deneubourg, J.L. 1990. A model for trail following in ants: Individual and collective behaviour. In W. Alt and G. Hoffman (eds.) *Biological Motion,* Lecture Notes in Biomathematics, (pp.453-469) Berlin: Springer-Verlag.

Callender, C. and Cohen, J. (2005) There Is No Special Problem About Scientific Representation. http://philsci-archive.pitt.edu/archive/00002177/

Carley, K.M. (1997) Validating computational models. Working paper: social and decision sciences, Carnegie Mellon University, Pittsberg, PA.

Collins, S., Ruina, A., Tedrake, R., Wisse, M. 2005. Efficient Bipedal Robots Based on Passive Dynamic Walkers. *Science Magazine*, 307:1082-1085.

Clancy, W. J. 1997. *Situated cognition*. Cambridge: Cambridge University Press.

Clark, A. 1997. *Being There: Putting Brain, Body, and World Together.* Cambridge MA: MIT Press.

Clark, A. 2001. *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford: OUP.

Dennett, D.C. 1978. Why not the whole iguana? *Behavioral and Brain Sciences*, 1: 103—104.

Dennett, D.C.1998. *Brainchildren: Essays on Designing Minds.* Cambridge, MA: MIT Press Dorigo, M. and Stützle, T. 2004. *Ant colony optimization*. Cambridge, MA: MIT Press.

Ehret, G. 1987. Categorical Perception of Speech Signals: Facts and Hypotheses from Animal Studies. In S. Harnad (ed.) *Categorical perception: The groundwork of Cognition.* New York: Cambridge University Press.

Frigg, R. 2003. Self organised criticality—what it is and what it isn't. *Studies in the History and Philosophy of Science* 34:613–632.

Frigg, R. and Hartmann, S. 2006. Models in science. In Edward N. Zalta (ed.).*Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/spr2006/entries/models-science/

Gibson, J. J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Goldenberg, E., Garcowski, J. and Beer, R.D. 2004. May We Have Your Attention: Analysis of a Selective Attention Task. In *From Animals to Animats 8: Proceedings of the Eighth International Conference on Simulation of Adaptive Behavior*, July 13-17, 2004, Los Angeles, CA.

Grasso, F.W. Consi, T.R., Mountain, D.C., Atema, J. 2000. Biomimetic robot lobster performs chemo-orientation in turbulence using a pair of spatially separated sensors: Progress and challenges. Robotics and Autonomous Systems 30(1-2):115-131.

Harvey, I., Di Paolo, E. A., Tuci, E., Wood, R., Quinn, M., 2005. Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life*, 11:79 - 98.

Hedwig B., Poulet J.F.P. 2004. Complex auditory behaviour emerges from simple reactive steering. *Nature,* 430:781-785.

Hedwig B., Poulet J.F.P. 2005. Mechanisms Underlying Phonotactic Steering in the Cricket *Gryllus bimaculatus (de Geer)* Revealed with a Fast Trackball System. *Journal of Experimental Biology*, 208: 915-927.

Hennig R.M., Franz A., Stumpner A. 2004. Auditory processing in insects. Microscopy Research Techniques, 63:351 – 374.

Horchler, A., Reeve, R., Webb, B. & Quinn, R. 2004. Robot phonotaxis in the wild: A biologically inspired approach to outdoor sound localization. *Advanced Robotics,* 18(8):801-816.

Huber, F. and Thorson, J., 1985. Cricket Auditory Communication. *Scientific American,* 253:47-54.

Husbands, P., Harvey, I. and Cliff, D. 1995. Circle in the Round: State Space Attractors for Evolved Sighted Robots. *Robotics and Autonomous Systems* 15:83—106.

Keeley B. L. 2000. Shocking lessons from electric fish: The theory and practice of multiple realization. *Philosophy of Science*, 67:444-465 .

Kim, K-J. and Cho, S-B. 2006. A comprehensive overview of the applications of ALife. *Artificial Life,* 12:153-182.

Koch, C. 1999. *Biophysics of Computation*. Oxford: Oxford University Press.

Lambrinos D., Möller R., Labhart T., Pfeifer R., Wehner R. 2000. A mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems*, 30:39-64.

Langton, C. 1989. Artificial Life. In C. Langton (ed.) *Artificial Life* (pp. 1-47). Addison-Wesley.

Lund, H. H., Webb, B., & Hallam, J. 1997. A robot attracted to the cricket species *Gryllus Bimaculatus*. In *Fourth European Conference on Artificial Life* Cambridge MA: MIT Press

Lund, H. H., Webb, B., & Hallam, J. 1998. Physical and temporal scaling considerations in a robot model of cricket calling song preference. *Artificial Life*, 4(1):95-107.

MacLennan, B. 1992. Synthetic Ethology: An Approach to the Study of Communication. In C. G. Langton, C. Taylor, J.D. Farmer and S. Rasmussen (eds.) *Artificial Life II: The Second Workshop on the Synthesis and Simulation of Living Systems* (pp. 631-658) Santa Fe Institute Studies in the Sciences of Complexity, Proceedings Vol. X. Redwood City, CA: Addison-Wesley.

MacLennan, B. and Burghardt, G. M. 1993. Synthetic Ethology and the Evolution of Cooperative Communication. *Adaptive Behavior*, 2(2):161-188.

MacLennan, B. 2002. Synthetic Ethology: A New Tool for Investigating Animal Cognition. In Marc Bekoff, Colin Allen, and Gordon M. Burghardt (eds.), *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* (pp. 151-156). Cambridge MA: MIT Press.

Mandik, P. 2002. Synthetic Neuroethology. *Metaphilosophy*. 33 (1-2):11-29.

Meyer J.A., Guillot A. 1991. Simulation of Adaptive Behavior in animats: Review and Prospect. In Meyer J.A., Wilson S.W. (eds.) *From animals to animats. Proceedings of the First Conference on Simulation of Adaptive Behavior (pp.2-14)*. Cambridge MA: MIT Press.

Meyer, J.A., Guillot, A., Girard, B., Khamassi, M., Pirim, P. and Berthoz, A. 2005. The Psikharpax project: Towards building an artificial rat. *Robotics and Autonomous Systems,* 50(4):211-223

Michelsen, A., Popov, A.V., Lewis, B. 1994. Physics of directional hearing in the cricket *Gryllus bimaculatus*. *Journal of Comparative Physiology A*, 175:153-164.

Miller, G.F. 1995. Artificial life as theoretical biology: How to do real science with computer simulation. Cognitive Science Research Paper 378, School of Cognitive and Computing Sciences, University of Sussex.

Möller R. 2000. Insect Visual Homing Strategies in a Robot with Analog Processing. *Biological Cybernetics*, 83(3):231-243.

Neumann, T.R. and Bülthoff, H.H. 2002. Behavior-oriented vision for biomimetic flight control. P*roceedings of the EPSRC/BBSRC International Workshop on Biologically Inspired Robotics: The Legacy of W. Grey Walter* (pp.196-203), 14-16 August, HP Labs Bristol, UK.

Neumann, T. R., Huber, S. and Bülthoff, H. H. 2001. Artificial systems as models in biological cybernetics. *Behavioural and Brain Sciences,* 24 (6):1071-1072.

Noble, J. 1997. The Scientific Status of Artificial Life, poster presented at The Fourth European Conference on Artificial Life, Brighton, UK, 28--31 July.

Pfeifer, R. and Scheier, R. 2000. *Understanding Intelligence.* Cambridge MA: MIT Press.

Pincock, C. 2004. Overextending Partial Structures: Idealization and Abstraction. In *Proceedings Philosophy of Science Assoc. 19th Biennial Meeting – PSA2004.*

Poulet J.F.A., Hedwig, B. 2005. Auditory orientation in crickets: Pattern recognition controls reactive steering. *Proceedings of the National Academy of Science,* 102:15665-15669.

Ray, T.S. 1994. An evolutionary approach to synthetic biology: Zen and art of creating life *Artificial Life,* 1:179-209.

Reeve, R. & Webb, B. 2003. New neural circuits for robot phonotaxis. P*hilosophical Transactions of the Royal Society A*, 361:2245-2266.

Reeve, R., Webb, B., Indiveri, G., Horchler, A., & Quinn, R. 2005. New technologies for testing a model of cricket phonotaxis on an outdoor robot. *Robotics and Autonomous Systems,* 51(1):41-54.

Schildberger, K. 1984. Temporal selectivity of identified auditory interneurons in the cricket brain. *Journal of Comparative Physiology A*, 155:171-185.

Schildberger, K, Horner, M. 1988. The function of auditory neurons in cricket phonotaxis .1. Influence of hyperpolarization of identified neurons on sound localization. *Journal of Comparative Physiology A*, 163, 621-631.

Scholes, S., Wilson, M., Sendova-Franks, A.B. and Melhuish, C. 2004. Comparisons in Evolution and Engineering: The collective Intelligence of Sorting. *Adaptive Behaviour,* 12: 147-159.

Simon, H. A. 1969. *The Sciences of the Artificial*. Cambridge MA: MIT Press.

Slocum, A.C., Downey, D.C. and Beer, R.D. 2000. Further experiments in the evolution of minimally cognitive behavior: From perceiving affordances to selective attention. In J. Meyer, A. Berthoz, D. Floreano, H. Roitblat and S. Wilson (eds.), F*rom Animals to Animats6: Proceedings of the Sixth International Conference on the Simulation of Adaptive Behavior* (pp. 430-439), Cambridge MA: MIT Press.

Srinivasan, M. V. & Venkatesh, S. (1997). *From living eyes to seeing machines*. Oxford: Oxford University Press.

Steels, L. 2003. Evolving Grounded Communication for Robots. *Trends in Cognitive Sciences*, 7(7):308--312.

Steels, L. 2001. **The methodology of the artificial.** *Behavioral and Brain Sciences*, 24:1077-1078.

Suarez, M. 2003. Scientific Representation: Against Similarity and Isomorphism. *International Studies in the Philosophy of Science*, 17(3):225-244.

Tyrrell, T. and Mayhew, J.E.W. 1991. Computer simulation of an animal environment. *Proceedings of the First International Conference on Simulation of Adaptive Behavior*, (pp. 263-272).

Voegtlin, T. and Verschure, P.F.M.J. 1999. What can robots tell us about brains? A synthetic approach towards the study of learning and problem solving. *Reviews in the Neurosciences* 10(3-4):291-310.

Verschure, P.F.M.J., Voegtlin, T. & Douglas, R.J. 2003. Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature*, 425: 620-624.

Wartofsky, M. W. 1979. *Models: representation and the scientific understanding*. Dordrecht: D.Reidel.

Webb, B. 1991. Do computer simulations really cognize? *Journal of Experimental and Theoretical Artificial Intelligence,* 3:247-254.

Webb, B. 1995. Using robots to model animals: a cricket test . *Robotics and Autonomous Systems*, 16, 117-134.

Webb, B. 2000. What does robotics offer animal behaviour? *Animal Behaviour* 60:545-558.

Webb, B. 2001. Can robots make good models of biological behaviour? *Behavioural and Brain Sciences* 24 (6):1033-1050.

Webb, B., Harrison, R. R., 2000. Integrating sensorimotor systems in a robot model of cricket behavior. *Sensor Fusion and Decentralised Control in Robotic Systems III* . Boston Nov 6-8: SPIE.

Webb, B., & Scutt, T. 2000. A simple latency dependent spiking neuron model of cricket phonotaxis. *Biological Cybernetics,* 82(3):247-269.

Wilson, S.W. 1985. Knowledge growth in an artificial animal. *Proceedings of the First International Conference on Genetic Algorithms and Their Applications* (pp. 16-23). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Wilson, S.W. 1991. The Animat Path to AI. In J. A. Meyer and S. W. Wilson (eds.), *From Animals to Animats: Proceedings of the First InternationalConference on the Simulation of Adaptive Behavior*, Cambridge, Massachusetts: The MIT Press/Bradford Books

Wimsatt, WC, 1987. False Models as means to Truer Theories. In: M. Nitecki and A. Hoffman, (eds.). London: Oxford University Press

 Winsberg, E. 1999. Sanctioning Models: The Epistemology. of. Simulation. *Science in Context*, 12(2):275-292.

Wohlers D.W., Huber F. 1982. Processing of sound signals by six types of neurons in the prothoracic ganglion of the cricket, *Gryllus campestris L. Journal of Comparative Physiology A,* 146: 161–173.