BEHAVIORAL AND BRAIN SCIENCES (2001) **24,** 1033–1050
Printed in the United States of America

Can robots make good models of biological behaviour?

Barbara Webb
Department of Psychology, Centre for Computational and Cognitive Neuroscience, University of Stirling, Stirling FK9 4LA, Scotland, United Kingdom
b.h.webb@stir.ac.uk **www.stir.ac.uk/psychology/Staff/bhw1/**

Abstract: How should biological behaviour be modelled? A relatively new approach is to investigate problems in neuroethology by building physical robot models of biological sensorimotor systems. The explication and justification of this approach are here placed within a framework for describing and comparing models in the behavioural and biological sciences. First, simulation models – the representation of a hypothesis about a target system – are distinguished from several other relationships also termed "modelling" in discussions of scientific explanation. Seven dimensions on which simulation models can differ are defined and distinctions between them discussed:

1. Relevance: whether the model tests and generates hypotheses applicable to biology.
2. Level: the elemental units of the model in the hierarchy from atoms to societies.
3. Generality: the range of biological systems the model can represent.
4. Abstraction: the complexity, relative to the target, or amount of detail included in the model.
5. Structural accuracy: how well the model represents the actual mechanisms underlying the behaviour.
6. Performance match: to what extent the model behaviour matches the target behaviour.
7. Medium: the physical basis by which the model is implemented.

No specific position in the space of models thus defined is the *only* correct one, but a good modelling methodology should be explicit about its position and the justification for that position. It is argued that in building robot models biological relevance is more effective than loose biological inspiration; multiple levels can be integrated; that generality cannot be assumed but might emerge from studying specific instances; abstraction is better done by simplification than idealisation; accuracy can be approached through iterations of complete systems; that the model should be able to match and predict target behaviour; and that a physical medium can have significant advantages. These arguments reflect the view that biological behaviour needs to be studied and modelled in context, that is, in terms of the real problems faced by real animals in real environments.

Keywords: animal behaviour; levels; models; neuroethology; realism; robotics; simulation

1. Introduction

"Biorobotics" can be defined as the intersection of biology and robotics. The common ground is that robots and animals are both moving, behaving systems; both have sensors and actuators and require an autonomous control system that enables them to successfully carry out various tasks in a complex, dynamic world. In other words "it was realised that the study of autonomous robots was analogous to the study of animal behaviour" (Dean 1998, p. 60), hence robots could be used as models of animals. As summarised by Lambrinos et al. (1997), "the goal of this approach is to develop an understanding of natural systems by building a robot that mimics some aspects of their sensory and nervous system and their behaviour" (p. 185).

Dean (1998) reviews some of this work, as do Meyer (1997), Beer et al. (1998), Bekey (1996), and Sharkey and Ziemke (1998), although the rapid growth and interdisciplinary nature of the work make it difficult to comprehensively review. Biorobotics will here be considered as a new methodology in biological modelling, rather than as a new "field" per se. It can then be discussed directly in relation to other forms of modelling. Rather than vague justification in terms of intuitive similarities between robots and animals, the tenets of the methodology can be more clearly stated and a basis for comparison to other approaches established. However, a difficulty that immediately arises is that a "wide divergence of opinion . . . exists concerning the proper role of models" (Reeke & Sporns 1993, p. 597) in biological research.

For example, the level of mechanism that should be represented in the model is often disputed. Cognitivists criticise connectionism for being too low level (Fodor & Pylyshyn 1988), while neurobiologists complain that connectionism abstracts too far from real neural processes (Crick 1989). Other debates address the most appropriate means for implementing models. Purely computer-based simulations are criticised by advocates of sub-threshold transistor technology (Mead 1989) and by supporters of "real-world" robotic implementations (Brooks 1986). Some

Barbara Webb obtained her B.Sc. in Psychology from the University of Sydney in 1988, and her Ph.D. in Artificial Intelligence from the University of Edinburgh in 1993. She is currently a lecturer in the Psychology Department at Stirling University, having previously lectured at the University of Nottingham and the University of Edinburgh.

worry about oversimplification (Segev 1992), while others deplore overcomplexity (Koch 1999; Maynard Smith 1974). Some set out minimum criteria for "good" models in their area (e.g., Pfeifer 1996; Selverston 1993); others suggest there are fundamental trade-offs between desirable model qualities (Levins 1966).

The use of models at all is sometimes disputed, on the grounds that detailed models are premature and more basic research is needed. Croon and van de Vijver (1994) argue that "developing formalised models for phenomena which are not even understood on an elementary level is a risky venture: what can be gained by casting some quite gratuitous assumptions about particular phenomena in a mathematical form?" (pp. 4–5). Others argue that "the complexity of animal behaviour demands the application of powerful theoretical frameworks" (Barto 1991, p. 94) and "nervous systems are simply too complex to be understood without the quantitative approach that modelling provides" (Bower 1992, p. 411). More generally, the formalization involved in modelling is argued to be an invaluable aid in theorising – "important because biology is full of verbal assertions that some mechanism will generate some result, when in fact it won't" (Maynard Smith 1988, p. 231).

Beyond the methodological debates, there are also "meta"-arguments regarding the role and status of models in both pure and applied sciences of behaviour. Are models essential to gaining knowledge or just convenient tools? Can we ever really validate a model (Oreskes et al. 1994)? Is reification of models mistaken, that is, can a model of a process ever be a replica of that process (Pattee 1989; Webb 1991)? Do models really tell us anything we didn't already know?

In what follows, a framework for the description and comparison of models will be set out in an attempt to answer some of these points, and the position of biorobotics with regard to this framework will be made clear. Section 2 will explicate the function of models, in particular to clarify some of the current terminological confusion, and define "biorobotic" modelling. Section 3 will describe different dimensions that can be used to characterise biological models, and discuss the relationships between them. Section 4 will lay out the position of robot models in relation to these dimensions, and discuss how this position reflects a particular perspective on the problems of explaining biological behaviour.

2. The process of modelling

2.1. *The "model muddle"* (Wartofsky 1979)

Many discussions of the meaning and process of modelling can be found: in the philosophy of science, for example, Hesse (1966), Harre (1970b), Leatherdale (1974), Bunge (1973), Wartofsky (1979), Black (1962), and further references throughout this article; in cybernetic or systems theory, particularly Zeigler (1976); and in textbooks on methodology – recent examples include Haefner (1996), Giere (1997), and Doucet and Sloep (1992). They also arise as part of some specific debates about approaches in biology and cognition: in ecological modelling, for example, Levins (1966) and Orzack and Sober (1993); in cognitive simulation, for example, Fodor (1968), Colby (1981), Fodor and Pylyshyn (1988), Harnad (1989); in neural networks, for example, Sejnowski et al. (1988), Crick (1989); and in Artificial Life, for example, Pattee (1989), Chan and Tidwell (1993). However the situation is accurately summed up by Leatherdale (1974): "the literature on 'models' displays a bewildering lack of agreement about what exactly is meant by the word 'model' in relation to science" (p. 41). Not only "model" but most of the associated terms – such as "simulation," "representation," "realism," "accuracy," "validation" – have come to be used in a variety of ways by different authors. Several distinct frameworks for describing models can be found, some explicit and some implicit, most of which seem difficult to apply to real examples of model building. Moreover, many authors seem to present their usage as the obvious or correct one and thus fail to spell out how it relates to previous or alternative approaches. Chao (1960) noted 30 different, sometimes contradictory, definitions of "model" and the situation has not improved.

There does seem to be general agreement that modelling involves the relationship of representation or correspondence between a (real) target system and something else.[1] Thus "A model is a representation of reality" (Lamb 1987, p. 91) or "all [models] provide representations of the world" (Hughes 1997, p. 325). What might be thought uncontroversial examples are: a scale model of a building which corresponds in various respects to an actual building; and the "billiard-ball model" of gases, suggesting a correspondence of behaviour in microscopic particle collisions to macroscopic object collisions. Already, however, we find some authors ready to dispute the use of the term "model" for one or other of these examples. Thus, Kaplan (1964) argues that purely "sentential" descriptions like the billiard-ball example should *not* be called "models"; whereas Kacser (1960) maintains that *only* sentential descriptions should be called "models" and physical constructions like scale buildings should be called "analogues"; and Achinstein (1968) denies that scale buildings *are* analogies while using "model" for both verbal descriptions and some physical objects.

A large proportion of the discussion of models in the philosophy of science concerns the problem that reasoning by analogy is not logically valid. If A and A* correspond in factors $x_1, \ldots, x_n$, it is not possible to deduce that they will therefore correspond in factor x_{n+1}. "Underdetermination" is another aspect of essentially the same problem – if two systems behave the same, it is not logically valid to conclude that the cause or mechanism of the behaviour is the same; so, a model that behaves like its target is not necessarily an explanation of the target's behaviour. These problems are sometimes raised in arguments about the practical application of models, for example, Oreskes et al. (1994) use underdetermination to argue that validation of models is impossible. Weitzenfeld (1984) suggests that a defence against this problem can be made by arguing that if there is a true *isomorphism* between A and A*, the deduction is valid, and the problem is only to demonstrate the isomorphism. Similar reasoning perhaps explains the frequently encountered claim that a model *is* "what mathematicians call an 'isomorphism'" (Black 1962, p. 222) – a one to one mapping – of "relevant aspects" (Schultz & Sullivan 1972), or "essential structure" (Kaplan 1964). Within cybernetic theory one can find formal definitions of models (e.g., Klir & Valach 1965) that require there to be a *complete* isomorphic or homomorphic mapping of all elements of a system, preserving all relationships.

However, this is not helpful when considering most actual examples of models (unless one allows there "to be as many definitions possible to isomorphism as to model," Conant & Ashby 1991, p. 516). In the vast majority of cases, models are not (mathematical) isomorphisms, nor are they intended to be. Klir and Valach (1965) go on to include as examples of models "photos, sculptures, paintings, films . . . even literary works" (p. 115). It would be interesting to know how they intend to demonstrate a strict homomorphism between *Anna Karenina* and "social, economic, ethical and other relations" in nineteenth century Russia. In fact, it is just as frequently (and often by the same authors) emphasised that a model *necessarily* fails to represent everything about a system. For example, Black (1962) goes on to warn of "risks of fallacies of inference from inevitable irrelevancies or distortions in the model" (p. 223) – but if there is a true isomorphism, how can there be such a risk? A "partial isomorphism" is an oxymoron; and more to the point, cannot suffice for models to be used in valid deduction. Moreover, this approach to modelling obscures the fact that the purpose in modelling is often to *discover* what are the "relevant features" or "essential structures," so model usage cannot depend on prior knowledge of what they are to establish the modelling relationship.

2.2. What use are models?

> *There are things and models of things, the latter being also things, but used in a special way*
> *(Chao 1960, p. 564)*

Models are intended to help us deal in various ways with a system of interest. How do they fulfill this role? It is common to discuss how they offer a convenient/cost-effective/manageable/safe substitute for working on or building the real thing. But this does not explain *why* working on the model has any relevance to the real system, or provide some basis by which relevance can be judged, that is, what makes a model a *useful* substitute? It is easier to approach this by casting the role of modelling as part of the process of explanation and prediction described in Figure 1.

Figure 1 can be regarded as an elaboration of standard textbook illustrations of either the "hypothetico-deductive" approach or the "semantic" approach to science (see below). To make each part of the diagram clear, consider an example. Our target – selected from the world – might be the human cochlea and the human behaviour of pitch perception. Our **hypothesis** might be that particular physical properties of the basilar membrane enable differently positioned hair cells to respond to different sound frequencies. One *source* of this idea may be the Fourier transform, and associated notion of a bank of frequency filters as a way of processing sound. To see what is predicted by the physical properties of the basal membrane we might build a symbolic **simulation** of the physical properties we think perform the function, and run it using computer *technology*, with different simulated sounds to see if it produces the same output frequencies as the cochlea (in fact Bekesy 1960 first investigated this problem using rubber as the technology to represent the basilar membrane). We could interpret the dominant output frequency value as a "pitch percept" and compare it to human pitch perception for the same waveforms: insofar as it fails to match we might conclude our hypothesis is not sufficient to explain human pitch perception. Or, as Chan and Tidwell (1993) concisely summarise this process, we theorise that a system is of type T, and construct an analogous system to T, to see if it behaves like the target system.

I have purposely not used the term "model" in the above description because it can be applied to different parts of this diagram. Generally, in this paper, I take "modelling" to correspond to the function labelled "simulation": models are something added to the "hypothesis-prediction-observation" cycle merely as "prostheses for our brains" (Milinski 1991). That is, modelling aims to make the process of producing predictions from hypotheses more effective by enlisting the aid of an analogical mechanism. A mathematical model such as the Hodgkin-Huxley equations sets up a correspondence between the processes in theorised mechanism – the ionic conductances involved in neural firing, – and processes defined on numbers – such as integration. We can more easily manipulate the numbers than the chemicals so the results of a particular configuration can be more easily predicted. However, limitations in the accuracy of the correspondence might compromise the validity of conclusions drawn.

However, under the "semantic" approach to scientific explanation (Giere 1997), the hypothesis itself is regarded as a "model," that is, it specifies a hypothetical system of which

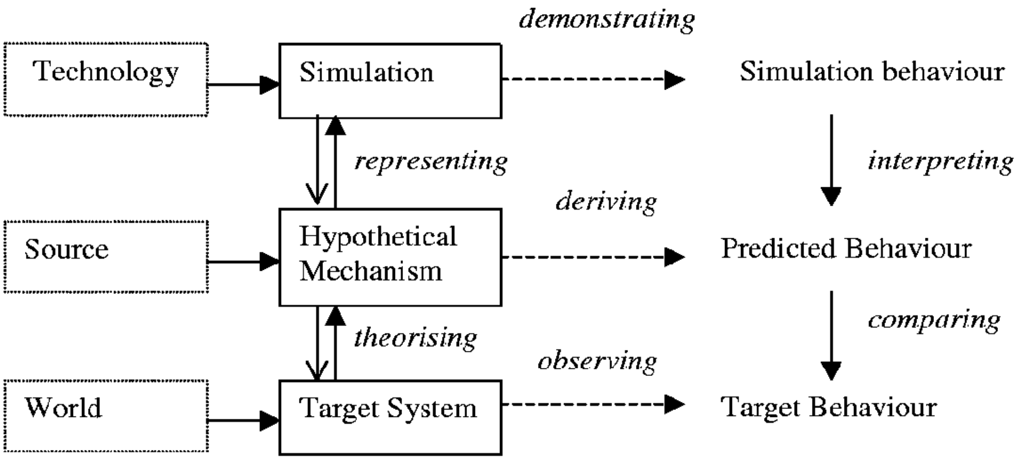

Figure 1. Models and the process of explanation

the target is supposed to be a type. The process of prediction is then described as "demonstration" (Hughes 1997) of how this hypothetical system should behave like the target. Demonstration of the consequences of the hypothesis may involve "another level" of representation in which the hypothesis is represented by some other system, also called a model. This system may be something already "found" – an analogical or source model – or something built – a simulation model (Morgan 1997). Moreover, the target itself can also be considered a "model," in so far as it involves abstraction or simplification in selecting a system from the world (Cartwright 1983). This idea perhaps underlies Gordon's (1969) definition of model: "we define a model as the body of information about a system gathered for the purpose of studying the system" (p. 5).

2.3. Theories, models, simulations, and sources

While the usage of "model" to mean the target is relatively rare, it is common to find "model" used interchangeably with "hypothesis" and "theory"[2]: even claims that "A model is a description of a system" (Haefner 1996, p. 4); or "A scientific model is, in effect, one or a set of statements about reality" (Ackoff 1962, p. 109). This usage of "model" is often qualified, most commonly as the "theoretical model," but also as the "conceptual model" (Brooks & Tobias 1996; Rykiel 1996; Ulinski 1999), "sentential model" (Harre 1970a), "abstract model" (Spriet & Vansteenkiste 1982), or, confusingly, the "real model" (Maki & Thompson 1973), or "base model" (Zeigler 1976). The tendency to call the hypothesis a "model" seems to be linked to how formal or *precise* is the specification it provides (Braithwaite 1960), as hypotheses can range from vague qualitative predictions to Zeigler's (1976) notion of a "well-described" base model, which involves defining all input, output, and state variables and their transfer and output functions, as a necessary prior step to simulation. The common concept of the theoretical model is that of a hypothesis that *describes* the components and interactions thought sufficient to produce the behaviour: "the actual building of the model is a separate step" (Brooks & Tobias 1996, p. 2).

This separate step is *implementation*[3] as a simulation, which involves representing the hypothesis in some physical instantiation – taken here in its widest sense, that is, including carrying out mathematical calculations or running a computer program, as well as more obviously "physical" models. But as Maki and Thompson (1973) note: "in many cases it is very difficult to decide where the real model [the hypothesis] ends and the mathematical model [the simulation] begins" (p. 4). Producing a precise formulation may have already introduced a number of "technological" factors that are not really part of the hypothesis, in the sense that they are there only to make the solution possible, not because they are really considered to be potential components or processes in the target system. Grice (cited in Cartwright 1983) called these "properties of convenience" and Colby (1981) makes this a basis for distinguishing models from theories: all statements of a theory are intended to be taken as true, whereas some statements in a model are not.

Simulation[4] is intended to augment our ability to deduce consequences from the assumptions expressed in the hypothesis: "a simulation program is ultimately only a high speed generator of the consequences that some theory assigns to various antecedent conditions" (Dennett 1979, p. 192); "models . . . help . . . by making predictions of unobvious consequences from given assumptions" (Reeke & Sporns 1993 p. 599). Ideally, a simulation should clearly and accurately represent the whole of the hypothesis and nothing but the hypothesis, so conclusions based on the simulation are, in fact, correct conclusions about the hypothesis. However, a simulation must also necessarily be precise in the sense used above, that is, all components and processes must be fully specified for it to run. The "formalization" imposed by implementation usually involves elaborations or simplifications of the hypothesis to make it tractable, which may have no theoretical justification. In other words, as is generally recognised, any actual simulation contains a number of factors that are not part of the "positive analogy" between the target and the model.

In the philosophy of science, discussion of "simulation" models has been relatively neglected. Rather, as Redhead (1980) points out, the extensive literature on models in science is mostly about modelling in the sense of using a "source" analogy. A source[5] is a pre-existing system used in devising the hypothesis. For example, Amit (1989) describes how concepts like "energy" from physics can be used in an analogical sense to provide powerful analysis tools for neural networks, without any implication that a "physics level" explanation of the brain is being attempted. Though traditionally the "source" has been thought of as another physical system (e.g., a pump as the source of hypotheses for the functioning of the heart), it is plausible to consider mathematics to be a "source." That is, mathematical knowledge provides a pre-existing set of components and operations we can put in correspondence to the hypothesised components and operations of our target. Mathematics just happens to be a particularly widely applicable analogy (Leatherdale 1974).

It is worth explicitly noting that the *source* is not in the same relation to the hypothesis as the *technology,* that is, what is used to implement the hypothesis in a simulation. Confusion arises because the same system can sometimes be used both as a source and as a technology. Mathematics is one example, and another of particular current relevance is the computer. The computer can be used explicitly as a source to suggest structures and functions that are part of the hypothesis (such as the information processing metaphor in cognition), or merely as a convenient way of representing and manipulating the structures and functions that have been independently hypothesised. It would be better if terms like "computational neuroscience" that are sometimes used strongly in the "source" sense – computation as an explanatory notion for neuroscience – were not so often used more loosely in the "technology" sense: "not every research application that models neural data with the help of a computer should be called computational neuroscience" (Schwartz 1990, p. x). Clarity is not served by having (self-labelled) "computational neuroethologists," for example, Beer (1990) and Cliff (1991) who apparently reject "computation" as an explanation of neuroethology.

2.4. Biorobotic models

Figure 1 suggests several different ways in which robots and animals might be related through modelling. First, there is a long tradition in which robots have been used as the *source* in explaining animal behaviour. Since at least Descartes (1662), regarding animals as merely complex machines, and explaining their capabilities by analogy with man-made systems has been a common strategy. It was most explicitly articulated in the cybernetic approach, which, in Wiener's subtitle to "Cybernetics" (Wiener 1948), concerned "control and communication in the animal and the machine." It also pertains to the information processing approaches common today, in which computation is the source for explaining brains. Much work in biomechanics involves directly applying robot-derived analyses to animal capacities, for example, Walker (1995) attempts "to analyse the strengths and weaknesses of the ancient design of racoon hands from the point of view of robotics" (p. 187).

Second, animals can be regarded as the *source* for hypotheses in robot construction. This is one widely accepted usage of the term "biorobotics" – sometimes called "biomimetic" or "biologically-inspired" robotics. For example, Ayers et al. (1998) suggest "the set of behavioural acts that a lobster or lamprey utilises in searching for and identifying prey is exactly what an autonomous underwater robot needs to perform to find mines." Pratt and Pratt (1998b) in their construction of walking machines "exploit three different natural mechanisms," the knee, ankle, and swing of animal legs to simplify control. The connection to biology can range from fairly exact copies of mechanisms, for example, Franceschini et al.'s (1992) electronic copy of the elementary motion detection circuitry of the fly, to adopting some high level principles, for example, using the ethological concept of "releasing stimuli" to control a robot via simple environmental cues (Connell 1990), or the approach described in Mataric (1998).

For the following discussion, however, I wish to focus on a third relationship: robots used as simulations of animals, or how "robots can be used as physical models of animals to address specific biological questions" (Beer et al. 1998, p. 777). The potential for building such models has increased enormously in recent years due to advances in both robot technology and neuroethological understanding, allowing "biologists/ethologists/neuroscientists to use robots instead of purely computational models in the modelling of living systems" (Sharkey & Ziemke 1998, p. 164).

The following criteria have been adopted for the inclusion of work in what follows as "biorobotic modelling," to avoid the necessity of discussing an unmanageably large body of work in robotics and biological modelling:

It must be robotic: The system should be physically instantiated and have unmediated contact with the external environment; the transduction is thus constrained by physics. The intention is to rule out purely computer-based models (i.e., where the environment as well as the animal is represented in the computer); and also computer sensing systems that terminate in descriptions rather than actions. This somewhat arbitrarily discounts verbal behaviour (e.g., visual classification) as sufficient; but to do so is consistent with most people's understanding of "robotic."

It must be biological: One aim in building the system should be to address a biological hypothesis or demonstrate understanding of a biological system. The intention is to rule out systems that might use some biological mechanisms but have no concern about altering them in ways that make it a worse representation, for example, industrial robot arms, most computer vision, most neural net controllers. It also rules out much of the "behaviour-based" approach in robotics which uses "algorithms specifying robot behaviours that have analogy to behaviours of life-form[s]" (Yamaguchi 1998, p. 3204) but makes no serious attempt to compare the results to natural systems. Probably the largest set of borderline cases thus excluded is the use of various learning mechanisms for robot behaviour, except those specifically linked to animal behavioural or physiological studies.

There is already a surprisingly substantial amount of work done even applying these criteria. The earliest examples come from mid-century, where theories of equilibrium (Ashby 1952), learning (Shannon 1951), and sensorimotor control (Walter 1961) were tested by building "animal" machines of various kinds – a number of other early examples are discussed in Young (1969). Current work tends to be more focused on specific biological systems, and ranges across the animal kingdom, from nematodes to humans. Table 1 lists a selection of recent studies, and to illustrate the approach I will describe three examples here in more detail.

1. *A robot model of rat hippocampus:* Burgess et al. (1997; 1998; 2000) have presented a model of the rat hippocampus implemented on a robot. "The use of a robot ensures the realism of the assumed sensory inputs and enables true evaluation of the navigational capability" (Burgess et al. 1997, p. 1535). The robot uses edge-filtering on a camera image to sense the distance of walls in its environment, and a combination of visual and odometric information to link the distance to the allocentric direction of the walls, rotating in place to cover a sufficient field of view. These researchers argue that these mechanisms "provide realistic simulation . . . since the rat's visual and odometric system appear to be relatively unsophisticated" (Burgess et al. 2000, p. 306). This sensory information is encoded computationally by sensory "cells" that effectively have "receptive fields" for different directions and distances of walls. These feed to an array of "entorhinal cells" which combine connections from sensory cells. These connect to the layer of "place cells" with the connection pattern modifiable by competitive learning: thus representing the learnt place dependent activity of cells observed in rat hippocampus. These cells further connect to a small number of goal cells, which also receive input from "head direction" cells. By Hebbian learning of these connections when a goal is encountered, the network forms a representation which can be used to guide the robot's movement back to a goal position from novel locations.

"[To] maintain close contact with the experimental situations in which the place cell data constraining the model was collected, the robot was tested in simple rectangular environments" (Burgess et al. 2000, p. 306). The results show the robot is capable of good self localisation while wandering in the environment and can reliably return to the goal position from novel locations. The effects of changing the environment (e.g., the proportions of the rectangle, or adding a new barrier) on the place cell representation and the search behaviour can be compared to the results in rats; some predictions from the model have been supported (Burgess et al. 2000). They further predict that cells with "receptive fields" for direction and distance of barriers will be found within or upstream of the entorhinal cortex, but this is yet to be confirmed.

2. *A robot model of desert ant navigation:* The impressive homing capabilities of the desert ant *Catyglyphis* have long been the subject of study (Wehner 1994). Several aspects of this behaviour have been investigated in robot

Table 1. *Examples of biorobot research. This is intended to be a representative sampling, not a fully comprehensive listing.*

Subject area	Examples	References
Simple sensorimotor control		
Chemical	Moth pheromone tracking	Kuwana et al. 1995; Ishida et al. 1999; Kanzaki 1996
	Ant trail following	Sharpe & Webb 1998; Russell 1998
	Lobster plume following	Grasso et al. 1996; Ayers et al. 1998
	C. elegans gradient climb	Morse et al. 1998
Auditory	Cricket phonotaxis	Webb 1995; Lund et al. 1998; Webb & Scutt 2000
	Owl sound localisation	Rucci et al. 1999
	Human localisation	Horiuchi 1997; Huang et al. 1995
	Bat sonar	Kuc 1997; Peremans et al. 1998
Visual	Locust looming detection	Blanchard et al. 1999; Indiveri 1998
	Frog snapping	Arbib & Liaw 1995
	Fly motion detection to control movement	Franceschini et al. 1992; Hoshino et al. 1998; Huber & Bulthoff 1998; 1997; Harrison & Koch 1999; Srinivasan & Venkatash 1997
	Praying mantis peering	Lewis & Nelson 1998
	Human oculomotor relflex	Horiuchi & Koch 1999; Shibata & Schaal 1999
	Saccade control	Clark 1998; Schall & Hanes 1998
Other	Ant polarized light compass	Lambrinos et al. 1997
	Lobster anemotaxis	Ayers et al. 1998
	Cricket wind escape	Chapman & Webb 1999
	Trace fossils	Prescott & Ibbotson 1997
Complex motor control		
Walking	Stick insect	Cruse et al. 1998; Pfeiffer et al. 1995
	Cockroach	Espenschied et al. 1996; Nelson & Quinn 1998; Binnard 1995
	Four-legged mammal	Ilg et al. 1998; Berkemeier & Desai 1996
Swimming	Tail propulsion	Triantafyllou & Triantafyllou 1995; Kumph 1998
	Pectoral fin	Kato & Inaba 1998
	Undulation	Patel et al. 1998
	Flagellar motion	Mojarrad & Shahinpoor 1997
Flying	Insect wings	Miki & Shimoyami 1998; Fearing 1999
	Bat	Pornsin-Siriak & Tai 1999
Arms/hands	Spinal circuits	Hannaford et al. 1995
	Cerebellar control	Fagg et al. 1997
	Grasping	Leoni et al. 1998
	Rhythmic movement	Schaal & Sternad 2001
	Haptic exploration	Erkman et al. 1999
Humanoid		Special issue *Advanced Robotics* 11(6): 1997
		Brooks & Stein 1993
		Hirai et al. 1998
Other	Running & hopping	Raibert 1986
	Brachiation	Saito & Fukuda 1996
	Mastication	Takanobu et al. 1998
	Snakes	Hirose 1993, Review in Worst 1998
	Paper wasp nest construct	Honma 1996
Navigation		
Landmarks	Ant/bee landmark homing	Moller 2000; Möller et al. 1998
Maps	Rat hippocampus	Burgess et al. 1997
		Gaussier et al. 1997
Search	*review*	Gelenbe et al. 1997
Collective behaviours		Beckers et al. 1996
		Melhuish et al. 1998
Learning		Edelman et al. 1992; Sporns 2001
		Scutt & Damper 1997
		Saksida et al. 1997
		Voegtlin & Verschure 1999
		Chang & Gaudiano 1998

models that operate in the same Sahara environment (Lambrinos et al. 1997; 2000; Moller et al. 1998). Insects can use the polarisation pattern of the sky as a compass, with three "POL" neurons in the brain integrating the response from crossed-pairs of filters at three different orientations. This sensor-neural morphology has been duplicated in the robot. Two different models for extracting compass direction were considered: a "scanning" mechanism that rotates to find a peak response which indicates the solar meridian (as had been previously proposed for the ant); and a novel "simultaneous" mechanism that calculates the current direction from the pattern of neural output. The "simultaneous" mechanism was substantially more efficient as the robot (or ant) does not need to rotate 360 degrees each time it wants to refer to the compass. This compass was successfully used in a path integration algorithm, reducing the error in the robot's return to its starting location.

A further development of the robot allowed the testing of hypotheses about landmark navigation. A conical mirror placed above a camera enabled the robot to get a 360 degree view of the horizon comparable to that of the ant. The "snapshot" model proposed by Cartwright and Collett 1983 was implemented first: this matches the landmarks in a current view with a stored view, to create a set of vectors whose average is a vector pointing approximately in the home direction. The ability of this model to return the robot to a location was demonstrated in experiments with the same black cylinders as landmarks as were used for the ant experiments. Further, a simplification of the model was proposed, in which the robot (or animal) only stores an "average landmark vector" rather than a full snapshot, and it was shown that the same homing behaviour could be reproduced. Möller (2000) recently implemented this in analog electronic hardware to provide "insights as to how the visual homing might be implemented in insect brains" (p. 243), and successfully tested this implementation on a robot in reproductions of experiments performed on bees in which landmarks are moved or removed.

3. *A robot model of human motor control:* Schaal and Sternad (2001) present a comparison of human and robot behaviour to analyse the control of motor trajectories. This is used to addressed a critical question – does the apparent "2/3 power law" relating endpoint velocity to path curvature in human movement represent an explicit parameter implemented directly in the nervous system, or is it merely the by-product of other control mechanisms? The study measured humans making cyclic drawing motions, and modelled the behaviour using a seven degree-of-freedom anthropomorphic robot arm, with PID control of joint movements based on simple sinusoidal target trajectories. The frequency, amplitude, and phase of the sinusoids were estimated from measurements on the human subjects. Schaal and Sternad found that "As in the human data, for small perimeter values [the 2/3 law] was produced quite accurately, but, as in the human subjects, the same deterioration of power law fits were apparent for increasing pattern size" (p. 67). Moreover, they could explain these deviations as a consequence of nonlinearities in the kinematic transform from joint control to end-effector trajectories, and explain the power law as emergent from mechanisms for ensuring smooth movement in joint space.

It can thus be seen that useful results for biology have been already been gained from robotic modelling. But it is still pertinent to ask: Why use robots to simulate animals? How does this methodology differ from alternative approaches to modelling in biology? To answer these questions it is necessary to understand the different ways in which models can vary, which will now be examined.

3. Dimensions for describing models

Figure 2 presents a seven-dimensional view of the "space" of possible biological models. If the "origin" is taken to be using the system itself as its own model (to cover the view expressed by Rosenblueth & Wiener 1945) as "the best material model of a cat is another, or preferably the same, cat" p. 316), then a model may be distanced from its target in terms of abstraction, approximation, generality or relevance. It may copy only higher levels of organisation, or represent the target using a very different material basis, or only roughly reproduce the target's behaviour. Exactly what is meant here by each of the listed dimensions, and in what ways they are (or are not) related will be discussed in detail

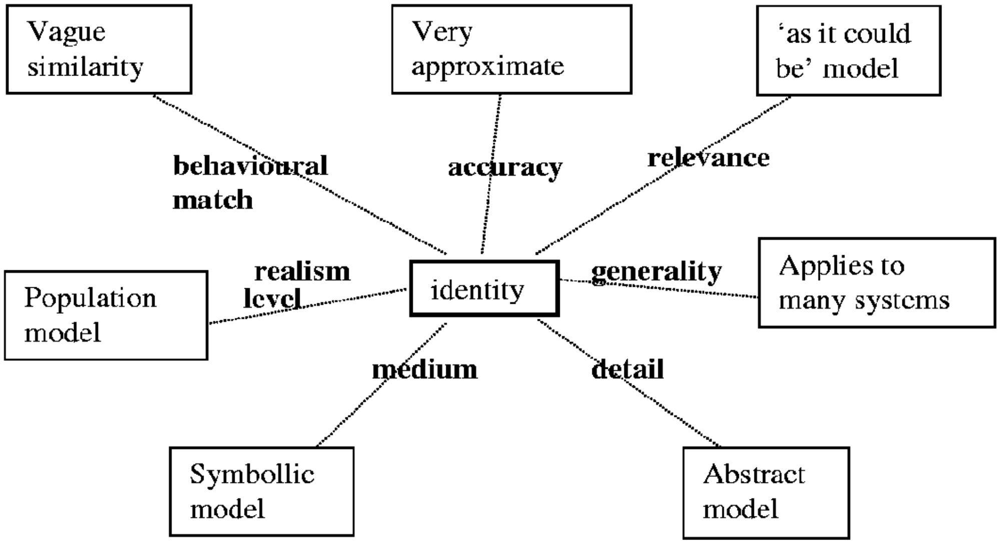

Figure 2. Dimensions for describing models

in what follows. They are presented as an attempt to capture, with a manageable number of terms, as much as possible of the variation described and discussed in the literature on modelling, and to separate various issues that are often conflated.

Though it is generally more popular in the literature to classify models into types (see e.g., the rather different taxonomies provided by Achinstein 1968; Black 1962; Haefner 1996; and Harre 1970b), there are precedents for this kind of dimensional description of models. Some authors attempt to use a single dimension. For example, Shannon (1975) presents a diagram of models situated on a single axis that goes from "exact physical replicas" at one end to "highly abstracted symbolic systems" at the other. By contrast, Schultz and Sullivan (1972) present a list of some 50-odd different dimensions by which a model may be described. One set of dimensions widely discussed in ecological modelling was proposed by Levins in 1966. He suggested that models could vary in realism, precision, and generality (in his 1993 reply to Orzack and Sober's 1993 critique, he notes that this was not intended to be a "formal or exhaustive" description). Within the "systems" approach to modelling the most commonly discussed dimensions are "complexity," "detail," and "validity" as well as more practical or pragmatic considerations such as cost (e.g., Rothenberg 1989 includes "cost-effectiveness" as part of his *definition* of simulation). Brooks and Tobias (1996) discuss some proposed methods for measuring these factors, and also point out how some of the connections between these factors are not as simple as seems to be generally thought.

Many of the debates about "appropriate" biological simulation assume that there are strict relations between certain aspects of modelling. Neural nets are said to be more accurate than symbol processing models because they are lower level; Artificial Life models are said to be general because they are abstract; neuromorphic models are said to be more realistic because they use a physical implementation. However, none of these connections follow simply from the nature of modelling but depend on background assumptions about biology. Is inclusion of a certain level essential to explaining behaviour? Can general laws of life be found? Are physical factors more important than information processing in understanding perception? The arguments for using robot models in biology, as for any other approach, reflect particular views about biological explanation. This will be further discussed in Section 4 which applies the defined dimensions to describe the biorobotic approach.

3.1. Biological relevance

Is the biological target system clearly identified? Does the model generate hypotheses for biology?

Models can differ in the extent to which they are intended to represent, and to address questions about, some real biological system. Work in biorobotics varies in biological relevance. For example, Huber and Bulthoff (1998) use a robot to test the hypothesis that a single motion-sensitive circuit can control stabilisation, fixation, and approach in the fly. This work is more directly applicable to biology than the robot work described by Srinivasan et al. (1999) utilising bee-inspired methods of motor control from visual flow-fields, which does not principally aim to answer questions about the bee. Similarly, the "robotuna" (Triantafyllou & Triantafyllou 1995) and "robopike" were specifically built to test hypotheses for fish swimming – "The aim of these robots is to help us learn more about the complex fluid mechanics that fish use to propel themselves" (Kumph 1998) – whereas the pectoral fin movements implemented on a robot by Kato and Inaba (1998), though based on close study of black bass, are not tested primarily for how well they explain fish swimming capability.

Another expression of this dimension is to distinguish between investigation of "the model as a mathematical statement and the model as empirical claim about some part of the physical world" (Orzack & Sober 1993, p. 535). Investigating a model for its own sake is often regarded critically. Hoos (1981) describes as "modelitis . . . being more interested in the model than the real world and studying only the portions of questions that are amenable to quantitative treatment" (p. 42). Bullock (1997) criticises Artificial Life where "simulations are sometimes presented as 'artificial worlds' worthy of investigation for their own sake . . . However this practice is theoretically bankrupt, and such [result] statements have no scientific currency" (p. 457). But Caswell (1988), for example, defends the need to investigate "theoretical problems" raised by models independently of their fit to reality. Langton's (1989) advocacy of investigating "life as it could be" is an example. As in "pure" maths, the results may subsequently prove to have key applications, but of course there is no guarantee that the "model-creating cycle" will not end up "spiralling slowly but surely away from reality" (Grimm 1994, p. 645) without any reconnection occurring.

It is worth explicitly mentioning in this context that a model that is "irrelevant" for biology might have utility in other respects. Models may serve for purposes of communication or education; or be employed for prediction and control. Moreover, there may be some value in investigating the technological aspects of a model: the mechanisms may have utility independent of their adequacy in explaining their origin. Arkin (1998) describes robots that abstract and use "underlying details" from biological sciences "unconcerned with any impact on the original discipline" (p. 32). Such "models" should then be evaluated with respect to engineering criteria,[6] rather than how well they represent some natural system.

Biologically "irrelevant" models, then, are those too far removed from biology to connect their outcomes back to understanding the systems that inspired them. For a non-robotic example, doubts are expressed about the relevance of artificial neural networks by, for example, Miall (1989): "it is not clear to what extent artificial networks will help in the analysis of biological networks" (p. 11). The main criteria for relevance could be taken to be the ability of the model to *generate testable hypotheses* about the biological system it is drawn from. For example the robot studies of Triantafyllou and Triantafyllou (1995) mentioned above suggest that fish use the creation of vortexes as a means of efficient tail-fin propulsion.

Arbib and Liaw (1995) provide as their definition of a "biological model": "a schema-based model . . . becomes a biological model when explicit hypotheses are offered as to how the constituent schemas are played over particular regions of the brain" (p. 56) (in their case, this involves the use of simulated and robot models of the visual guidance of behaviour in the frog). Generalised, this seems an appropriate test for relevance: are the mechanisms in the model explicitly mapped back to processes in the animal, as hypotheses about its function? In biorobotics this may sometimes concern neural circuitry, for example, in a model of auditory localisation of the owl (Rucci et al. 1999). But it can also occur at a relatively high level, such as using "shaping" methods in learning (Saksida et al. 1997) or involve testing a simple algorithm such as the sufficiency of a small set of local rules to explain collecting and sorting behaviour in ants (Holland & Melhuish 1999; Melhuish et al. 1998). The point is to use the robot model to make a serious attempt at addressing biological questions, at whatever level these may exist.

This notion of "relevance" appears to be what at least some authors mean by the term "realism" in describing models. Churchland and Sejnowski (1988) appear to define "realistic" in this way: "realistic models, which are genuinely and strongly predictive of some aspect of nervous system dynamics or anatomy" versus "simplifying models, which though not so predictive, demonstrate that the nervous system could be governed by specific principles" (p. 744). But this is rather different to their definition in Sejnowski et al. (1988) of a realistic model as a "large scale simulation that tries to incorporate as much of the cellular detail as is available" made "realistic by adding more variables and more parameters" (p. 1300). It seems unlikely that they believe only models "realistic" in the latter sense can be "realistic" in the former sense – indeed, they argue in Churchland et al. (1990) that "a genuine perfect model, faithful in every detail, is as likely to be incomprehensible as the system itself" (p. 54). However, "realistic" *is* often used to mean "detailed," or "not abstract." For example: Beer et al. (1998) specify "realistic" in relation to robot models as those which "literally try to emulate in every detail a particular species of insect" (p. 32); Manna and Pnueli (1991) define realism as "degree of detail"; Palsson and Lee (1993) directly equate "realistic" to "complex" – a decision on realism is a decision on how many factors to include; and Orzack and Sober (1993) redefine Levins' (1966) concept of realism as "tak[ing] into account more independent variables known to have an effect" (p. 534).

However, it is clear that Levins (1966) was concerned to argue *against* the assumption that a model can only be made "realistic" by being more detailed. His discussion of "real and general" models includes a number of quite simple and abstract examples: the issue of realism is the extent to which they improve understanding of the biological system, that is, what I have here called relevance. Schultz and Sullivan (1972) make a useful distinction between modelling that tries to build a complete "picture of reality" versus building a device for *learning* about reality: that is, it may be possible for a model to be too detailed (or "realistic" in one sense) to actually be productive of hypotheses (or "realistic" in the other sense). Collin and Woodburn (1998) similarly refer to the possibility of "a model in which the incorporated detail is too complex . . . for it to contribute anything to the understanding of the system" (pp. 15–16). The relevance of a model to biology, and the detail it includes, are separable issues which should not be conflated under the single term "realism."

3.2. Level

What are the base units of the model?

This dimension concerns the hierarchy of physical/processing levels that a given biological model could attempt to represent. Any hypothesis will usually have "elemental units" whose "internal structure does not exist or can be ignored" (Haefner 1996, p. 4). In biology these can range from the lowest known mechanisms such as the physics of chemical interactions through molecular and channel properties, membrane dynamics, compartmental properties, synaptic and neural properties, networks and maps, systems, brains and bodies, perceptual and cognitive processes, up to social and population processes (Shepherd 1990). The level modelled in biorobotics usually includes mechanisms of sensory transduction, for example the sonar sensors of bats (Kuc 1997) including the pinnae movements (Peremans et al. 1998), or of motor control, such as the six legs of the stick insect (Pfeiffer et al. 1995) or the multijointed body of the snake (Hirose 1993). The central processing can vary from a rule-based level through high level models of brain function such as the control of eye movements (Schall & Hanes 1998), to models of specific neuron connectivity hypothesised to underlie the behaviour, such as identified neural circuitry in the cricket (Webb & Scutt 2000), and even the level of dendritic tree structure that explains the output of particular neurons such as the "looming" detector found in the locust and modelled on a robot by Blanchard et al. (1999). The data for the model may come from psychophysics (e.g., Clark's 1998 model of saccades), developmental psychology (Scassellati 1998), or evolutionary studies (Kortmann & Hallam 1999), but most commonly comes from neuroethological investigations.

This notion of level corresponds to what Churchland and Sejnowski (1988) call "levels of organisation" and, as they note, this does not map onto Marr's well-known discussion of "levels of analysis" (Marr 1982). Marr's discussion of levels (computational, algorithmic, and implementational) applies rather to any explanation across several levels of organisation, and describes how one level (be that network, neuron, or channel) considered as an algorithm relates to the levels above (computation) and below (implementation). In fact, this point was made clearly by Feibleman (1954): "For any organisation, at any given level, its mechanism lies at the level below and its purpose at the level above" (p. 61).

One source of the conflict over the "correct level" for biological modelling may be that levels in biology are relatively close in spatio-temporal scale, as contrasted with macro and micro levels in physics by Spriet and Vansteenkiste (1982). They point out that "determination of an appropriate level is consequently less evident" (p. 46) in biological sciences. Thus, it is always easy to suggest to a modeller that they should move down a level; whereas it is obviously impractical to pursue the strategy of always working at the lowest level. Koch (1990) makes the interesting point that low-level details may be unimportant in analysing some forms of collective neural computation, but may be critical for others – the "correct level" may be problem specific, and "which really are the levels relevant to explanation in the nervous system is an empirical, not an a priori, question" (Churchland et al. 1990, p. 52).

Another problem related to levels is the misconception that the level of a model determines its biological relevance. A model is not made to say more about biology *just* by including lower-level mechanisms. For example, using a mechanism at the "neural" level does not in itself make a model realistic: most "neural network" controlled robots have little to do with understanding biology (Zalzala & Morris 1996). Moreover, including lower levels will generally make the model more complex, which may result in its being intractable and/or incomprehensible. Levins (1993) provides a useful example from ecological models: it is realistic to include a variable for the influence of "nutrients"; less realistic to include specific variables for "nitrogen" and "oxygen" if, thereby, other nutrient effects are left out. It is also important to distinguish level from accuracy (see below) as it is quite possible to inaccurately represent any level. Shimoyama et al. (1996) suggest that to "replicate functionality and behaviour . . . not necessarily duplicate their anatomy" in building robot models of animals is to be "not completely faithful" (p. 8): but a model can "faithfully" replicate function at different levels.

3.3. Generality

How many systems does the model target?

A more general model is defined as one that "applies to more real-world [target] systems" (Orzack & Sober 1993, p. 534). Some researchers in biorobotics appear sanguine about the possibility of generality, for example, Ayers et al. (1998) claim "locomotory and taxis behaviours of animals are controlled by mechanisms that are conserved throughout the animal kingdom" and thus their model of central pattern generators is taken to be of high generality. Others are less optimistic about general models. Hannaford et al. (1995), regarding models of motor control with "broad" focus, opines "because of their broad scope, it is even more difficult for these models to be tested against the uncontroversial facts or for them to predict the results of new reductionist experiments." This suggests that increasing generality decreases relevance, so it should be noted that, strictly speaking, a model *must* be relevant to be general – if it does not apply to any specific system, then how can it apply to many systems (Onstad 1988)? But a model does not have to be general to be relevant.

The obvious way to test if a model is general is to show how well it succeeds in representing a variety of different specific systems. For many models labelled "general" this does not happen. When it is attempted, it usually requires a large number of extra situation or task specific assumptions to actually get data from the model to compare to the observed target. This is a common criticism of optimal foraging studies (Pierce & Ollanson 1987): that, provided enough task specific assumptions are made, any specific data can be made to fit the general model of optimality. A similar critique can be made of "general" neural nets (Verschure 1996) – a variety of tasks can be learned by a common architecture, but only if the input vectors are carefully encoded in a task specific manner. Raaijmakers (1994) makes a similar point for memory models in psychology and pertinently asks – is this any better than building specific models in the first place?

The most common confusion regarding generality is that what is abstract will thereby be general. This can often be found in writings about artificial life simulations, and Estes (1975), for example, makes this claim for psychological models. Shannon (1975) talks about "the most abstract and hence the most general models" (p. 10) and Haefner (1996) suggests more detail necessarily results in less generality. Sejnowski et al. (1988) describe "simplifying models" as abstracting from individual neurons and connectivity to potentially provide "general findings" of significance for the brain. Sometimes this argument is further conflated with "levels," for example, Wilson (1999) discusses how "component neurons may be described at various levels of generality" (p. 446) contrasting the "abstraction" of spike rates to the "detail" of ionic currents – but an ionic current description is actually more general as it applies to both spiking and nonspiking neurons. The membrane potential properties of neurons are very general across biology but not particularly abstract; whereas logical reasoning is quite abstract but not very general across biological systems. Obviously some concepts are both abstract and general – such as feedback – and many concepts are neither. Moreover, precisely the opposite claim, that is, that *more* detail makes models more general, is made by some authors, for example, Black (1962), Orzack and Sober (1993). The reasoning is that adding variables to a model will increase its scope, because it now includes systems where those variables have an influence, whereas before it was limited to systems where they do not.

Grimm (1994) points out that insofar as generality appears to be lost when increasing detail, it may simply be because the systems being modelled are in fact unique, rather than because of an inherent trade-off between these factors. This raises the important issue that "generality has to be found, it cannot simply be declared" (Weiner 1995, p. 155). That is to say, the generality of a model depends on the true nature of the target(s). If different animals function in different ways then trying to generalise over them won't work – you are left studying an empty set. Robertson (1989) makes the same point with regard to neural networks "[neural] circuits that are unique in their organisation and operation demand unique models if such models are to be useful" (p. 262); Taylor (1989) similarly argues for ecology that simple models are "not shortcuts to ecological generality." Consequently, one strategy is to work instead on understanding specific systems, from which general mechanisms, if they exist, will emerge (Arbib & Liaw 1995). Biology has often found that the discovery and elucidation of general mechanisms tends to come most effectively from close exploration of well-chosen specific instantiations (Miklos 1993), such as the fruitfly genome or squid giant axon.

3.4. Abstraction

How many elements and processes from the target are included in the model?

Abstraction concerns the number and complexity of mechanisms included in the model; a more detailed model is less abstract. The "brachiator" robot models studied by Saito and Fukuda (1996) illustrate different points on this spectrum: an early model was a simple two-link device, but in more recent work they produce a nine-link, twelve degree-of-freedom robot body with its dimensions based on exact measurements from a 7–8 year-old female simiang skeleton. "Abstraction" is not just a measure of the simplicity/complexity of the model however (Brooks & Tobias 1996) but is relative to the complexity of the target. Thus a simple target might be represented by a simple, but not abstract, model, and a complex model still be an abstraction of a very complex target.

Some degree of abstraction is bound to occur in most model building. Indeed, it is sometimes taken as a defining characteristic of modeling – "A model is something simple made by the scientist to help them understand something complicated" (Segev 1992, p. 414). It is important to note that abstraction is not directly related to the level of modelling: a model of a cognitive process is not, of its nature, more or less abstract than a model of channel properties. The amount of abstraction depends on how the modeller chooses to describe and represent the processes, not what kind of processes they represent. Furthermore, the fact that some models – such as biorobots – have a hardware "medium" (see below) does not make them necessarily less abstract than computer simulations. A simple pendulum might be used as an abstract physical model for a leg, whereas a symbolic model of the leg may include any amount of anatomical detail. As Etienne (1998) notes, "Robots tend to simulate behaviour and the underlying neural events on the basis of a simplified architecture and therefore less precisely than computers" (p. 286).

How much abstraction is considered appropriate seems to largely reflect the "tastes" of the modeller: should biology aim for simple, elegant models or closely detailed system descriptions? Advocates of abstraction include Maynard Smith (1974): "Should we not therefore put into the model everything that we think might be important? . . . construction of such complex models is a waste of time" (p. 116); and Molenaar (1994): "[it is] precisely by simplification and abstraction that models are most useful" (p. 101). The latter gives as reasons for preferring more abstract models that complex models are harder to implement, understand, replicate or communicate. An important point is that they thereby become hard for reviewers to critique or check (e.g., Rexstad & Innis 1985 report a surprising number of basic errors in published models they were attempting to reimplement to test simplification techniques). Simpler models are easier to falsify, and reduce the risk of mere data-fitting, by having fewer free parameters. Their assumptions are more likely to be transparent. Another common argument for building a more abstract model is to make the possibility of an analytical solution more likely (e.g., the abstraction of neural "sprouting" proposed by Elliot et al. 1996).

However, abstraction carries risks. The existence of an attractive formalism might end up imposing its structure on the problem so that alternative, possibly better, interpretations are missed. Segev (1992) argues that in modelling neurons, we need to build complex detailed models to discover what are appropriate simplifications. Details abstracted away might turn out to actually be critical to understanding the system. As Kaplan (1964) notes, the issue is often not just "over-simplification" per se, but whether we have "simplified in the wrong way" or, that "what was neglected is something important for the purposes of that very model" (p. 281). For explaining biological behaviour, abstracting away from the real problems of sensorimotor interaction with the world is argued, within biorobotics, to be an example of the latter kind: in this case, abstraction reduces relevance because the real biological problem is not being addressed.

3.5. Structural accuracy

Is the model a true representation of the target?

Accuracy is here intended to mean how well the mechanisms in the model reflect the real mechanisms in the target. This is what Zeigler calls structural validity: "if it not only produces the observed real system behaviour but truly reflects the way in which the real system operates to produce this behaviour" (1976, p. 5) as distinct from replicative and predictive validity, that is, how well the input/output behaviour of the system matches the target.[7] This notion has also been dubbed "strong equivalence" (Fodor 1968). Brooks and Tobias (1996) call this the "credibility" of the model, and Frijda (1967) suggests "[input/output] performance as such is not as important as convincing the reader that the reasons for this performance are plausible" (p. 62). Thus, Hannaford et al. (1995) lay out their aims in building a robot replica of the human arm as follows: "Although it is impossible to achieve complete accuracy, we attempt to base every specification of the system's function and performance on uncontroversial physiological data."

One major issue concerning the accuracy of a model is "how can we know?" (this is also yet another meaning of "realism"). The anti-realist interpretation of science says that we cannot know. The fact that certain theories appear to work as explanations is not evidence to accept that they represent reality, because the history of science has shown us to be wrong before (the "pessimistic meta-induction," Laudan 1981). On the other hand, if they do not approximately represent reality then how can we build complex devices that actually work based on those theoretical assumptions (the "no miracle argument," Putnam 1975)? Not wishing to enter this thorny territory, it will suffice for current purposes to argue for no more than an instrumentalist position. If we can't justifiably *believe* our models, we can justifiably *use* them (Van Fraassen 1980). Accuracy in a model means there is "acceptable justification for scientific content of the model" (Rykiel 1996, p. 234) relative to the contemporary scientific context in which it is built; and that it is rational (Cartwright 1983) to attempt "experimental verification of internal mechanisms" (Reeke & Sporns 1993, p. 599) suggested by the model.

Inaccuracies in models should affect our confidence in using the model to make inferences about the workings of the real system (Rykiel 1996), but do not rule out all inference, provided "assumptions . . . [are] made explicit so that the researcher can determine in what direction they falsify the problem situation and by how much" (Ackoff 1962, p. 109). Horiuchi and Koch (1999) make this point for neuromorphic electronics: "By understanding the similarities and differences . . . and by utilising them carefully, it is possible to maintain the relevance of these circuits for biological modelling" (p. 243). Thus, accuracy can be distinguished from relevance. It is possible for a model to address "real" biological questions without utilising accurate mechanisms. Many mathematical models in evolutionary theory fit this description. Dror and Gallogly (1999) describe how "computational investigations that are completely divorced, in practice and theory, from any aspect of the nervous system . . . can still be relevant and contribute to understanding the biological system" (p. 174), for example, as described by Dennett (1984) to "clarify, sharpen [and] systematise the purely semantic level characterisation" (p. 203) of the problem to be solved.

Accuracy is not synonymous with "amount of detail" included in the model. This is well described by Schenk (1996) in the context of "tree" modelling. He notes that researchers often assume that a model with lots of complex detail is accurate, without actually checking that the details are correct. Or, a particular simplification may be widely used, and justified as a necessary abstraction, without looking at alternatives that may be just as abstract but more accurate. Similarly, it has already been noted that accuracy does not relate directly to the level of the representation – a high-level model might be an accurate representation of a cognitive process where a low-level model may turn out to be not accurate to brain biology.

A widely used term that overlaps with both "relevance" and "accuracy" is "biological plausibility." This can be taken simply to mean the model is applicable to some real biological system; or used to describe whether the assumptions on which the model are based are biologically accurate. More weakly, it is sometimes used to mean that the model "does not require biologically unrealistic computations" (Rucci et al. 1999, p. 96). In fact, this latter usage is probably a better interpretation of "plausible," that is, it describes models where the mechanism merely *could be* like the biology, rather than those where there are stronger reasons to say the mechanism *is* like the biology – the latter is "biological accuracy," and neither is a pre-requisite for "biological relevance" in a model.

3.6. Match

To what extent does the model behave like the target?

This dimension describes how the model's performance is assessed. In one sense it concerns testability: can we potentially falsify the model by comparing its behaviour to the target? For example, the possibility that the lobster uses instantaneous differences in concentration gradients between its two antennules to do chemotaxis was ruled out by testing a robot implementation of this algorithm in the real lobster's flow-tank (Grasso et al. 2000). However, assessment of a biorobot may be simply in terms of its capabilities rather than directly relate back to biology. While a significant role for robot models is the opportunity to compare different control schemes for their success (e.g., Ferrell 1995 looks at three different controllers, two derived from biology, for six-legged walking) simply reporting what will work best on a (possibly inaccurate) robot model does not necessarily allow us to draw conclusions about the target animal behaviour.

When a direct comparison with biology is attempted, there is still much variability on this dimension regarding the nature of the possible match between the behaviours. Should the behaviours be indistinguishable or merely similar? Are informal, expert, or systematic statistical investigations to be used as criteria for assessing similarity? Is a qualitative or quantitative match expected? Can the model both reproduce past data and predict future data? Some modelling studies provide little more than statements that, for example, "the overall behaviour looked quite similar to that of a real moth" (Kuwana et al. 1995, p. 375). Others make more direct assessment, for example, Harrison and Koch (1999) have tested their analog VLSI optomotor system in the real fly's flight simulator and "repeated an experiment often performed on flies," showing, for example, that the transient oscillations observed in the fly emerge naturally from inherent time-delays in the processing on the chip. Even where biological understanding is not the main aim of the study, it is possible that "animals provide a benchmark" for evaluating the robot system, such as Berkemeier and Desai's (1996) comparison of their "biologically-styled" leg design to the hind limb of a cat at the level of force and stiffness parameters.

There are inevitable difficulties in drawing strong conclusions about biological systems from the results of robot models. As with any model, the performance of similar behaviour is never sufficient to prove the similarity of mechanisms – this is the problem of underdetermination. Some authors are concerned to stress that behavioural match is never sufficient evidence for drawing conclusions about the accuracy or relevance of a model (e.g., Deakin 1990; Oreskes et al. 1994). Uttal (1990) goes so far as to say that "no formal model is verifiable, validatable or even testable with regard to internal mechanisms" and claims this is "generally accepted throughout other areas of science." But the widespread use of models in exactly the way so deplored suggests that most modelers think a reasonable defence for the practice can be made in terms of falsification or coincidence. If the model does not match the target then we can reject the hypothesis that led to the model or at least know we need to improve our model. If it does match the target, better than any alternatives, then the hypothesis is supported to the extent that we think it unlikely that such similar behaviour could result from completely different causes. This is sometimes more formally justified by reference to Bayes' theorem (Salmon 1996).

However, there are some limitations to this defence. Carrying out the comparison of model and target behaviours can be a sufficiently complex process so that neither of the above points apply. First, how can we be sure that the measurements on the real system are correct? If the model does not match we may reject the measurements rather than the model. Second, an interpretation process is required to convert the behaviour of the model and target into comparable form. This interpretation process may be wrong, or more worryingly, may be adjusted until the "match" comes out right – "interpretive steps may inadvertently contain key elements of the mechanism" (Reeke & Sporns 1993, p. 598). Third, it is not uncommon for models to have their parameters "tuned" to improve the match. As Hopkins and Leipold 1(996) demonstrate, this practice can in fact conceal substantial errors in the model equations or in the data. Finally, Bower and Koch (1992) provide a sobering view of the likelihood of a model being rejected on the basis of failure to match experiments:

> experiments needed to prove or disprove a model require a multi-year dedicated effort on the part of the experimentalist . . . falsification of any one such model through an *experimentum crucis* can be easily countered by the introduction of an additional *ad hoc* hypothesis or by a slight modification of the original model. Thus the benefit, that is, the increase in knowledge, derived from carrying out such time- and labour-intensive experiments is slight. (p. 459)

3.7. Medium

What is the simulation built from?

Hypotheses can be instantiated as models in various different forms, and hardware implementation is one of the most striking features of biorobotics compared to other biological models. Doucet and Sloep (1992) list "mechanical," "electric," "hydraulic," "scale," "map," "animal," "game," and "program" as different forms a model might take. A popular taxonomy is "iconic," "analog," and "symbolic" models (e.g., Black 1962; Chan & Tidwell 1993; Kroes 1989; Schultz & Sullivan 1972), but the definitions of these terms do not stand up to close scrutiny. "Iconic" originally derives from "representation," meaning something used to stand in for something else, and is used that way by some authors (Harre 1970b; Suppe 1977) to mean any kind of analogy-based model. However, it is now often defined specifically as using "another instance of the target type" (Chan & Tidwell 1993), or "represent the properties by the same properties with a change of scale" (Schultz & Sullivan 1972, p. 6). One might assume this meant identity of materials for the model and the target, as discussed below, but the most cited example is Watson and Crick's scale model of DNA, which was built of metal, not deoxyribonucleic acid. Yet "analog" models are then distinguished from "iconic" as models that introduce a "change of medium" (Black 1962) to stand in for the properties. A popular example of an analog model is the use of electrical circuit models of mechanical systems. Some authors include computer models as analogs, for example, Achinstein (1968) whereas others insist they are symbolic, for example, Lambert and Brittan (1992). But whether the properties are shared or analogous or "merely" symbolically represented depends entirely on how the properties are defined: whether the "essence" of a brain is its chemical constitution, its connectivity pattern or its ability to process symbols depends on what you are trying to explain. All models are "iconic," or share properties, precisely from the point of view that makes the model usefully stand in for the target for a particular purpose (Durbin 1989 calls this "the analogy level"). Hence, I will abandon this distinction and consider the medium more literally as what the model is actually built from.

A model can be constructed from the same materials as its target. Bulloch and Syed (1992) describe "culture models," that is, the reconstruction of simplified networks of real neurons in vitro as models of networks in vivo; and Miklos (1993) argues for the use of transgenic techniques to "build novel biological machines to test our hypotheses" (p. 843). Kuwana et al. (1995) use actual biological sensors – the antennae of moths – on their robot model and note these are 10,000 times more sensitive than available gas sensors. In these cases the representation of the target properties is by identity in the model properties.

However, most models are not constructed from the same materials. They may share some physical properties with their targets, for example, a vision chip and an eye both process real photons. Gas sensing is substituted for pheromone sensing in Ishida et al.'s (1999) robot model of the moth, but they replicate other physical features of the sensor, for example, the way that the moths wings act as a fan to draw air over the sensors. Models may use similar physical properties. This may mean that the properties can be described by the same mathematics, for example, the subthreshold transistor physics used in neuromorphic design are said to be equivalent to neuron membrane channel physics (Etienne-Cummings et al. 1998). Or it may be a "looser" mapping. The robot model of chemotaxis in *Caenorhabolitis Elegans* (Morse et al. 1998) uses a light source as an analog for a chemical gradient in a petri dish, while preserving a similar sensor layout and sensitivity. Models may also use quite different properties to stand in for the properties specified in the target, for example, a number in the computer processor labelled "activity" to represent the firing rate of a neuron, or the use of different coloured blocks in a robot arena to represent "food" and "mates."

In fact, nearly all models use all these modes of representation to various extents in creating correspondences to the hypothesised target variables. Thus "symbolic" computer simulations frequently use time to represent time (Schultz & Sullivan 1972); "iconic" scale models tend to use materials of analogous rigidity rather than the same materials; mathematical models can be treated as a short-hand for building a physically "analogous" system. Rather than sharply contrasting "kinds" of models, what is of relevance are the constraints the medium imposes on the operation of the model. What makes a representation more "symbolic" is that the medium is more arbitrary or neutral with respect to representing the target properties. Symbols rest on "arbitrary conventions – no likeness or unlikeness it may bear to a its subject matter counts as a reason why it is a symbol for, or of, α" (Harre 1970, p. 38). More "physical" models are chosen *because* the medium has some pre-existing resemblance to the properties we wish to represent, such as the use of analog VLSI to implement neural processing (Mead 1989). The medium may contribute directly to the accuracy and relevance of the model, or simply make it easier to implement, run or evaluate as described by Quinn and Espenscheid (1993):

> Even in the most exhaustive [computer] simulations some potentially important effects may be neglected, overlooked or improperly modelled. It is often not reasonable to attempt to account for the complexity and unpredictability of the real world. Hence implementation in hardware is often a more straightforward and accurate approach for rigorously testing models of nervous systems. (p. 380)

Doucet and Sloep (1992) point out "the way physical models operate is, as it were, ruled by nature itself – rules for functioning of conceptual [symbolic] models – we make ourselves" (p. 281). Symbolic models may implicitly rely on levels of precision in processing that are unlikely to be possible to real systems. Computer programs can represent a wider range of possible situations than we can physically model, but physical models cannot break the laws of physics.

4. The position of biorobotics

In section 2.4, I discussed in what sense biorobots can be considered biological models – in particular, how robots can be used as physical simulations of organisms, to test hypotheses about the control of behaviour. How, then, does biorobotics differ from other modelling approaches in biology? If it is suggested that "the use of a robot ensures the realism" (Burgess et al. 1997, p. 1535) of a model, does this mean making the model more relevant for biology, making it more detailed, making it more accurate, making it more specific (or general?), making it a "low-level" model, making the performance more lifelike, or just that the model is operating with "real" input and output?

In this section, I will use the framework developed above to clarify how biorobotics differs, on various dimensions, from other kinds of biological models. I will also advance arguments for why the resulting position of biorobots in modelling "space" is a good one for addressing some fundamental questions in explaining biological behaviour. I do not intend to suggest that it is the only correct approach – "there is no unique or correct model" (Fowler 1997, p. 8) of a system. However "there are good and bad models" relative to the purposes of the model builder. Thus, this discussion will also illustrate for what purposes in understanding biology, biorobotics appears to have particular strengths.

4.1. Relevance to biology

A notable feature that distinguishes recent biorobotic research from earlier biologically-inspired approaches in robotics (such as the "behaviour-based" approach articulated by Brooks 1986 and Arkin 1998) is the increased concern as to whether and how the robot actually resembles some specified biological target. Thus, most of the robot studies listed in Table 1 cite the relevant biological literature that has guided decisions on what to build, how to build it, and how to assess it; frequently a biological investigator has initiated or collaborated directly in the research. The likelihood of being able to apply the results back to biology is thus increased, even where this was not the primary aim in the initial robot construction. Biorobotics has been able to confirm, develop, and refute theories in several areas of biology, as already described in a number of examples above.

A distinction was drawn in previous sections between using biorobots as biological models and using them for engineering, and it is sometimes argued that these are incompatible, or at least orthogonal, concerns (e.g., Hallam 1998; Pfeifer 1996). Nevertheless, many of those working in biorobotics claim to be doing both. For example, Hirose et al. (1996) include as "biorobotics" both "build[ing] robots that emulate biological creatures" *and* "us[ing] development of robots to understand biological and ethological concepts" (p. 95). Espenschied et al. (1996), in describing their work on robot models of insect walking, claim that "results that demonstrate the value of basing robot control on principles derived from biology . . . also . . . provide insight into the mechanisms of locomotion in animals and humans" (p. 64). Lambrinos et al. (1997), regarding their robot model of desert ant navigation, suggest: "On the one hand, the results obtained . . . provide support for the underlying biological models. On the other hand . . . computationally cheap navigation methods for mobile robots are derived." Raibert (1986), in discussing methods for legged locomotion, points out: "In solving problems for the machine, we generate a set of plausible algorithms for the biological system. In observing the biological behaviour, we explore plausible behaviours for the machine" (p. 195).

Indeed, even where the explicit aim in building the robot model is said to be just "engineering" or just "biology," the process is very likely to involve both. It is the engineering requirement of making something that actually works that creates much of the hypothesis testing power of robotic models of biological systems. This is well described by Raibert (1986):

> To the extent that an animal and a machine perform similar locomotion tasks, their control systems and mechanical structure must solve similar problems. By building machines we can get new insights into these problem, and we can learn about possible solutions. Of particular value is the rigor required to build physical machines that actually work. (p. 3)

In the other direction, building a robot "inspired" by an animal source presupposes a certain degree of knowledge about that source. If, as Ayers et al. (1998) claim, "biologically-based reverse engineering is the most effective procedure" to design robots, then we need to understand the biology to build the robots – in Ayers' case this has involved exhaustive analysis of the underlying "units" of action in the locomotion behaviour of the lobster. That is, our goal is, as defined by Shimoyama et al. (1996), "to understand activation and sensing of biological systems so that we can build equivalents" (p. 8) or, as Leoni et al. (1998) put it, "a proper comprehension of human biological structures and cognitive behaviour . . . is fundamental to design and develop a [humanoid] robot system" (p. 2274). The robot designer's motives thus overlap substantially with those of the biologist.

4.2. Level

It is sometimes argued in biorobotics that this methodology should focus on lower levels, or working from "bottom-up." In fact, Taddeucci and Dario (1998) describe explicitly, in the context of models of eye-hand control, what most biorobotics researchers do implicitly, that is, work both top-down and bottom-up on the problems. The possible influence of lower level factors is kept in mind, and the exploration of the interaction of levels is engaged in. While this is perhaps true of many other modelling approaches, robotic implementation specifically supports the consideration and integration of different levels of explanation because of its emphasis on requiring a complete, behaving system as the outcome of the model. For example, Hannaford et al. (1995) primarily consider their robot arm as a "mechanism or platform with which to integrate information," particularly the interaction of morphology and neural control. Thus, the context of the behaviour of the organism is always included in a robot model, counteracting the tendency in biological studies to lose sight of this context in close study of small parts of the underlying mechanisms.

The level of mechanism modelled by the robot will reflect the level of information currently available from biology. Interest in a particular level of explanation (such as single neuron properties) may bias the choice of target system, for example, towards invertebrate systems in which identified neurons have been mapped (Franceschini 1996). On the other hand, interest in a particular target may determine the level at which an accurate model can be attempted. For example, Etienne (1998) reviews the behavioural and physiological data on mammalian navigation and concludes that lack of information about the actual interactions of the neural systems "leaves the field wide open to speculative modelling" (p. 283) at the level of networks.

In addition, biorobotic systems emphasise the "physical" level in the performance of sensing and action. That is, the dynamics of the physical interaction of the robot/animal and its environment are seen to be as critical in explaining its behaviour as the processing or neural connectivity (Chiel & Beer 1997). It is often found that engaging closely in modelling the periphery simplifies central or higher level processing. For example, Mura and Shimoyama (1998) note that copying the circuitry of insect visual sensors "closely integrates sensing and early stage processing" to "ease off decision making at a higher processing level" (p. 1859), and Kubow and Full (1999) discuss the extent to which running control is actually done by the mechanical characteristics of the cockroach's legs. Some of the most interesting results from biorobotic modelling demonstrate that surprisingly simple control hypotheses can suffice to explain apparently complex behaviours when placed in appropriate interaction with the environment. Examples include the use of particular optical motion cues to achieve obstacle avoidance that slows down the robot in cluttered environments without explicit distance cues being calculated (Franceschini et al. 1992), the "choice" between sound sources with different temporal patterns resulting from a simple four-neuron circuit in the robot cricket (Webb & Scutt 2000), and the use of limb linkage through real world task constraints to synchronise arm control (Williamson 1998).

4.3. Generality

In engineering, robots built for specific tasks have to date been more successful than "general purpose" ones. Similarly, in biorobotics the most successful results to date have been in the context of modelling specific systems – particular competencies of particular animals. There is some doubt whether modelling "general" animal competencies (e.g., by simulating "hypothetical" animals such as Pfeifer's [1996] "fungus eater" or Bertin and van de Grind's [1996] "paddler") will tell us much about any real biological system. Without regrounding the generalisations by demonstrating the applicability of the results to some specific real examples, the problem modelled may end up being "biological" only in the terminology used to describe it.

An example of the tendency to more specificity is the shift in research described by Nelson and Quinn (1998) from generic six-legged walkers (Espenscheid et al. 1996) to a robot that closely copies the anatomy and mechanics of the cockroach. As they explain, the desired movement capabilities for the robot – fast running and climbing abilities – depend on quite specific properties such as the different functions of the rear, middle, and front pair of legs. Hence, the specific morphology has to be built into the robot if it is to be able to exploit features such as the propulsive power of the rear legs and the additional degrees of freedom in the front legs that enable the cockroach to climb.

If important factors in understanding behaviour lie in the specific sensorimotor interface, then it is necessary to model specific systems in sufficient detail to encompass this. "Generalising" a sensorimotor problem can result in changing the nature of the problem to be solved. What is lost are the properties described by Wehner (1987) as "matched filters," the specific fit of sensor (or motor) mechanisms to the task. The sound localisation of crickets is a good illustration. Crickets have a unique auditory system in which the two ears are connected by a tracheal tube to form a pressure difference receiver. This gives them good directionality but only for a specific frequency – that of the calling song they need to localise. Copying this mechanism in a robot model, it was possible to demonstrate that this factor alone can suffice to reproduce the cricket's ability to respond only to songs with the carrier frequency of conspecifics (Lund et al. 1997).

Note that while "matched filters" are by their nature specific to particular animals, the concept is a general one. Similarly, while the neural circuitry modelled in the cricket robot is highly specific to the task (and hence very efficient), the idea it uses of exploiting timing properties of neural firing is a general one. Thus we can see general principles emerging from the modelling of specific systems. Moreover, the "engineering" aspect of biorobotics enhances the likelihood of discovering such generalities as it attempts to transfer or apply mechanisms from biology to another field, the control of man-made devices.

4.4. Abstraction

It might be assumed that the aims discussed so far of increasing relevance by having a clearly identified target system, and increasing specificity rather than trying to invent general models, require that biorobotic models become more detailed. Beer et al. (1997) suggest as a principal for this approach "[generally to] err by including more biology than appears necessary" (p. 33). However, others believe that abstraction does not limit relevance, for example, according to McGeer (1990) "it seems reasonable to suppose that our relatively simple knee jointed model has much to say about walking in nature" (p. 1643). Indeed, it has been suggested that a key advantage of biorobotics is the discovery of "simpler" solutions to problems in biology because it takes an abstract rather than analytic approach (Meyer 1997). It is clear that some quite abstract robot representations have usefully tested some quite specific biological hypotheses. For example, there is minimal representation of biological details in the physical architecture of Beckers et al.'s (1996) robot "ants," Burgess et al.'s (1997) "rat" or indeed the motor control of the robot "cricket" mentioned above, but nevertheless, it was possible to demonstrate interesting resemblance in the patterns of behaviour of the robots and the animals, in a manner appropriate to testing the hypotheses in question.

Rather than being less abstract, it might better be said that biorobotics has adopted different abstractions from simulations (or from standard robot control methods; Bekey 1996; Pratt & Pratt 1998a). Robots are not less abstract models just because they are physically implemented – a two-wheeled robot is a simpler model of motor control than a six-legged simulation. What does distinguish abstraction in biorobotics from simulations is that it usually occurs by leaving out details, by substitution, or by simplifying the representation, rather than by *idealising* the objects or functions to be performed. Thus, even two-wheeled motor control has to cope with friction, bumps, gravity, and so on; whereas a six-legged computer simulation may restrict itself to representing only the kinematic problems of limb control and ignore the dynamics entirely.

Different aspects of the systems are often abstracted to different degrees in biorobotics. Thus, models involving quite complex sensors often use very simple two-wheeled motor control rather than equally complex actuators. Edelman et al. (1992) describe relatively complex neural models but test them in rather abstract tasks. Though some robots are tested in quite complex environments, the majority have a simplified environment constructed for them (though in some cases this is not much different from the controlled environment used to test the animals). Pfeifer (1996) and Cruse (2001) have made the point that this *imbalance* in abstraction may itself lead to a loss of biological relevance. What is needed is to ensure that the assumptions involved in the abstraction are clear, and justified. A good example is the description by Morse et al. (1998) of the simplifications they adopted in their robot model of chemotaxis in *C. elegans*, such as the biological evidence for abstracting the motor control as a constant propulsive force plus a steering mechanism provided by contraction of opposing muscles.

4.5. Accuracy

If "highest possible accuracy" was considered to be the aim in biorobotics, then there are many ways in which existing systems can be criticised. Most robot sensors and actuators are not directly comparable to biological ones: they differ in basic capability, precision, range, response times, and so on. Binnard (1995) in the context of building a robot based on some aspects of cockroach mechanics, suggests that the "tools and materials . . . are fundamentally different" (p. 44), particularly in the realm of actuators. Ayers (1995) more optimistically opines that "Sensors, controlling circuits and actuators can readily be designed which operate on the same principles as their living analogs." The truth is probably somewhere in between these extremes. Often the necessary data from biology is absent or not in a form that can easily be translated into implementation (Delcomyn et al. 1996). The process of making hypotheses sufficiently precise for implementation often requires a number of assumptions that go well beyond what is accurately known of the biology. As for abstraction, there is also a potential problem in having a mismatch in the relative accuracy of different parts of the system. For example, it is not clear how much is learnt by using an arbitrary control system for a highly accurate anatomical replica of an animal; or conversely, by applying a detailed neural model to control a robot carrying out a fundamentally different sensorimotor task.

Biorobotics researchers are generally more concerned with building a complete, but possibly rough or inaccurate model, than with strict accuracy per se. That is, the aim is to build a complete system that connects action and sensing to achieve a task in an environment, even if this limits the individual accuracy of particular parts of the model because of necessary substitutions, interpolations, and so on. While greater accuracy is considered worth striving for, a degree of approximation is considered a price worth paying for the benefits of gaining a more integrated understanding of the system and its context, in particular the "tight interdependency between sensory and motor processing" (Pichon et al. 1989, p. 52). This is exemplified in their robot "fly" by the use of self movement to generate the visual input required for navigation.

Projects that set out to build "fully accurate" models tend not to get completed, and we can learn more from several somewhat inaccurate models than from one incomplete one. In several cases the accuracy has then been increased iteratively, for example, the successive moves from a slower, larger robot implementation of the cricket robot (Webb 1994; 1995), to a robot capable of processing sound at cricket speed (Lund et al. 1998), and then to a controller that more closely represents the cricket's neural processing (Webb & Scutt 2000). Indeed, Levins (1966) argues that building multiple models is a useful strategy to compensate for inevitable inaccuracies because results common to all the models are "robust" with respect to the individual inaccuracies of each.

4.6. Match

It should be admitted that the assessment of the behaviour relative to the target is still weak in most studies in biorobotics. It is more common to find only relatively unsupported statements that a robot "exhibited properties which are consistent with experimental results relating to biological control systems" (Leoni et al. 1998, p. 2279). One encouraging trend in the direction of more carefully assessing the match is the attempt to repeat experiments with the same stimuli for the robot and the animal. For example, Touretsky and Saksida (1997) describe how they "apply our model to a task commonly used to study working memory in rats and monkeys – the delayed match to sample task" (p. 219). Sharpe and Webb (1998) draw on data in ant chemical trail-following behaviour for methods and critical experiments to assess a robot model's ability to follow similar trails under similar condition variations, such as changes in chemical concentration. Some behaviours lend themselves more easily than others to making comparisons – for example, the fossilised worm trails reproduced in a robot model by Prescott and Ibbotson (1997) provide a clear behavioural "record" to attempt to copy with the robot.

The accuracy of the robot model may impose its own limits on the match. Lambrinos et al. (1997) note, when testing their polarisation compass and landmark navigation robot in the Sahara environment, that despite the same experimental conditions "it is difficult to compare the homing precision of these agents, since both their size and their method of propulsion are completely different" (p. 200). There is also the inherent problem in any modelling, that reproducing the same behaviour is not proof that the same underlying mechanism is being used by the robot and the animal. There are some of ways in which the biorobotics approach can attempt to redress these limitations. By having a specific target, usually chosen because there is substantial existing data, more extensive comparisons can be made. Using a physical medium and more accurately representing environmental constraints reduces the possibility that the "world model" is being tuned to make the animal model work, rather than the reverse. The interpretation of the behaviour is more direct. Voegtlin and Verschure (1999) argue, in their robot implementation of models of classical conditioning, that by combining levels, and thus satisfying constraints from "anatomy, physiology and behaviour" the argument from match is strengthened.

Finally, biorobotic modelling has been instrumental in driving the collection of further data from the animal. Quinn and Ritzmann (1998) describe how building a cockroach-inspired robot has "required us to make detailed neurobiological and kinematic observations of cockroaches" (p. 239). Correctly matching the behaviour is perhaps less important then revealing what it is we need to know about the animal to select between possible mechanisms demonstrated in the robot.

4.7. Medium

The most distinctive feature of the biorobotics approach is the use of hardware to model biological mechanisms. It is also perhaps the most often questioned – what is learnt that could not be as effectively examined by computer simulation? One justification relates to the issue of building "complete" models discussed above – the necessity imposed by physical implementation that all parts of the system function together and produce a real output. Hannaford et al. (1995) argue that "Physical modelling as opposed to computer simulation is used to enforce self consistency among co-ordinate systems, units and kinematic constraints" in their robot arm.

Another important consideration is that using identity in parts of a model can sometimes increase accuracy at relatively little cost. Using real water or air-borne plumes, or real antennae sensors, saves effort in modelling and makes validation more straightforward. Dean (1998) proposes that by capturing the body and environmental constraints, robots provide a stronger "proof in principle" that a certain algorithm will produce the right behaviour. In engineering, demonstration of a real device is usually a more convincing argument than simulated results. Thus one direction of current efforts in biorobotics is the attempt to find materials and processes that will support better models. Dario et al. (1998) review sensors and actuators available for humanoid robots. Kolacinski and Quinn (1998) discuss elastic storage and compliance mechanisms for more muscle-like actuators. Mojarrad and Shahinpoor (1997) describe polymeric artificial muscles that replicate undulatory motions in water, which they use to test theoretical models of animal swimming. On a similar basis some researchers use dedicated hardware for the entire control system (i.e., not a programmed microcontroller). Franceschini et al.'s (1992) models of the fly motion detection system used to control obstacle avoidance are developed as fully parallel, analog electronic devices. Maris and Mahowald (1998) describe a complete robot controller (including contrast sensitive retina and motor spiking neurons) implemented in analog VLSI.

Cited advantages of hardware implementations include the ability to exploit true parallelism, and increased emphasis on the "pre-processing" done by physical factors such as sensor layout. It is important to note, however, that simply using a more "physical" medium does not reduce the need for "ensuring that the relevant physical properties of the robot sufficiently match those of the animal relative to the biological question of interest" (Beer et al. 1998, p. 777). Electronic hardware is not the same medium as that used biology, and may lend itself to different implementations – a particular problem is that neural connectivity is three dimensional where electronic circuits are essentially two-dimensional.

However, a more fundamental argument for using physical models is that an essential part of the problem of understanding behaviour is understanding the environmental conditions under which it must be performed – the opportunities and constraints that it offers. If we simulate these conditions, then we include only what we already assume to be relevant, and moreover represent it in a way that is inevitably shaped by our assumptions about how the biological mechanism works. Thus, our testing of that mechanism is limited in a way that it is not if we use a real environment, and the potential for further discovery of the actual nature of the environment is lost. Thus, Beckers et al. (1996) suggest "systems for the real world must be developed in the real world, because the complexity of interactions available for exploitation in the real world cannot be matched by any practical simulation environment" (p. 183). Flynn and Brooks (1989) argue that "unless you design, build, experiment and test in the real world in a tight loop, you can spend a lot of time on the wrong problems" (p. 15).

5. Conclusions

> *It was by learning the inner workings of nature*
> *that man became a builder of machines.*
> *(Hoffer, cited by Arkin 1998, p. 31)*

> *We've only rarely recognised any mechanical*
> *device in an organism with which we weren't*
> *already familiar from engineering.*
> *(Vogel 1999, p. 311)*

Biorobotics, as the intersection of biology and robotics, spans both views represented by the quotes above – understanding biology to build robots, and building robots to understand biology. It has been argued that robots can be "biological models" in several different senses. They can be modelled *on* animals – the biology as a source of ideas when attempting to build a robot of some target capability. They can be models *for* animals – robotic technology or theory as a source of explanatory mechanisms in biology. Or they can be models *of* animals – robots as a simulation technology to test hypotheses in biology. Work on this last kind of "biorobot," and the potential contribution it can make to biology, has been the main focus of discussion in this target article.

To assess biorobotics in relation to other kinds of simulations in biology, a multi-dimensional description of approaches to modelling has been proposed. Models can be compared with respect to their relevance, the level of organisation represented, generality, abstraction, structural accuracy, behavioural match, and the physical medium used to build them. Though interrelated, these dimensions are separable: models can be relevant without being accurate, general without being abstract, match behaviour at different levels, and so on. Thus, a decision with respect to one dimension does not necessarily constrain a modeller with respect to another.

I agree with Levins (1993) that a dimensional description should not be primarily considered as a means of ranking models as "better" or "worse" but rather as an elucidation of potential strategies. The strategy of biorobots has here been characterised as: increasing relevance and commitment to really testing biological hypotheses; combining levels; studying specific systems that might illustrate general factors; abstracting by simplification rather than idealisation; aspiring to accuracy but concerned with building complete systems; looking for a closer behavioural match; and using real physical interaction as part of the medium. The motivations for this strategy have been discussed in detail above, but can be compactly summarised as the view that biological behaviour needs to be studied in context, that is, in terms of the real problems faced by real animals in real environments.

Thus, the justification of the biorobotic approach is grounded in a particular perspective on the issues that need to be addressed. Different approaches to modelling will reflect differing views about the processes being modelled, and the nature of the explanations required. One aim of this paper is to encourage other modellers to clarify their strategies and the justification for them – even if it is only by disagreement over the included dimensions. Indeed, different views of "models" reflect different views of the "nature of explanation," as has been long discussed in the philosophy of science. It has not been possible to pursue all these meta-issues, some of which seem in any case to have little relevance to everyday scientific use of simulation models. What is critical is that the conclusions that can be drawn from a model are only as good as the representation provided by that model. In this respect, by working on real problems in real environments, robots can make good models of real animals.

NOTES

1. Suppe (1977) distinguishes this "representational" use of "model" from "model" used in the mathematical sense of a semantic interpretation of a set of axioms such that they are true. There is not space in this article to discuss this model theoretic approach in the philosophy of science (Carnap 1966; Nagel 1961; Suppe 1977), or the formal systems theoretic approach to models developed by Zeigler (1976) and adopted in many subsequent works (e.g., variants in Halfon 1983; Maki & Thompson 1973; Spriet & Vansteenkiste 1982; Widman & Loparo 1989). These formal/logical definitions are in any case not easy to apply to real examples of models in science where "Modelling is certainly an art, involving a number of logical gaps" (Redhead 1980, p. 162).

2. Black (1962) suggests this generic usage of "model" is "a pretentious substitute for theory" whereas Stogdill (1970) calls it "an unpretentious name for a theory."

3. Implementation is sometimes taken to mean actually reproducing a real copy of the system (Harnad 1989), that is, replication; this is not intended here.

4. It should be acknowledged that there are several, fairly widespread, definitions of simulation more restricted than the usage I have adopted here. First, there is the usage that contrasts iterative solutions for mathematical systems to analytical solutions (Forrester 1972). Second, there is the emphasis on simulations being processes, that is, "dynamic" versus "static" models or an "operating model . . . one that is itself a process" (Schultz & Sullivan 1972, p. 2). These distinctions have some validity, but I am going to ignore them for convenience, as analytical and static models stand in the same relationship to targets and hypotheses as iterative or temporal ones. Third, is the usage of "simulation" to refer to relatively detailed models of specific systems (e.g., of a particular species in certain niche) as opposed to more "general" models (e.g., of species propagation) which may also be implemented on computers for iterative solutions (e.g., Levins 1993; Maynard Smith 1974). Fourth, is the distinction of simulations as models that only attempt to match input-output behaviour (e.g., Dreyfus 1979; Ringle 1979) as opposed to models that are supposed to have the same internal mechanisms as their target. These latter distinctions often carry the implication that "simulations" are used for applications and "models" for science, that is, these distinctions tend to be polemic rather than principled (Palladino 1991), and they are certainly not clear-cut.

5. The term "source" is taken from Harre (1970b) who discusses this notion extensively. Unfortunately, the term "source" is also occasionally used for what I have called the "target," by some authors.

6. "Biologically-inspired" robots can be criticised at times for using "biological" as an excuse for not evaluating the mechanism against other engineered solutions, while using "inspired" as a disclaimer for being required to show it applies to biology.

7. Some authors do use "accuracy" in the sense of "replicative validity," for example, Bhalla et al. (1992): "accuracy is defined as the average normalized mean square difference between the simulator output and the reference curve" p. 453). The term "match" is used instead in this article (see sect. 3.7).

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

Some robotic imitations of biological movements can be counterproductive

Ramesh Balasubramaniam and Anatol G. Feldman

Research Center, Rehabilitation Institute of Montreal, University of Montreal, Montreal, Quebec H3S 2J4, Canada. **mahavishnu@sympatico.ca**
Feldman@ere.umontreal.ca

Abstract: It is proposed here that Webb's ideas about robots as possible models of animals need some rethinking. In our view, even though widely used biorobotics strategies are fairly successful at reproducing the macroscopic behavior of biological systems, there are still several problems unresolved on the side of robotics as well as biology. Both mathematical and hardware-like robotics models should be feasible physiologically. Control principles elaborated in robotics are not necessarily applied to biological control systems. Although observations of flying birds inspired aerodynamics and thus modern airplanes, little knowledge has been added to the neurophysiological principles underlying flight in birds. Chess playing computers might outperform most chess players, but they cannot be considered as physiologically feasible models of human thinking.

Barbara Webb presents a comprehensive report of issues in biorobotics with respect to relevance, level, generality, abstraction, structural accuracy, performance match, and medium. We contend that even if a robotic model of biology satisfies all of the abovementioned constraints, there are still issues that can make mechanical instantiations of biological behavior different from the actual behavior of biological systems.

General purpose vs. special purpose systems. The author draws inspiration from the idea that if an animal and a machine perform the same behavior, their control systems and mechanical structure must solve similar problems (Raibert, 1986). We believe that this idea can be quite misleading. Consider the example of a human arm. The human arm can be assembled to throw, to catch, to reach, to grasp, move rhythmically, and so on. Each of these behaviors requires a different functional assembly of the arm, with different activation of muscles and joints in a task specific way (Feldman & Levin 1995). A robotic arm can be engineered to perform all these behaviors but all the contexts of constraint have to be built into the system. Most conventional robots are general purpose devices, in the sense that they use the same dynamical regime for all types of tasks (e.g., spatial trajectory planning for the terminal device and velocity servoing use the same device). A biological system, on the one hand, carries its own contexts of constraint (Turvey 1990). The generality of a biological system comes from the fact that task-specific solutions are orchestrated by capitalizing on peculiarities of the task such as invariants, lawful regularities, and so on. Special purpose devices, on the other hand, are softly-assembled devices that are put together functionally to meet the demands of specific tasks (Greene 1975).

How much biology can we understand from machines? Webb borrows from the ideas of Beckers et al. (1996) and Flynn and Brooks (1989) that physical model systems for the real world must be developed not in any practical simulation environment but by exploiting the complexity of interactions available in the real world itself. While we agree with this general contention, there are some issues here that need to be resolved. The modern airplane was inspired by centuries of observations of bird flight. The physical principles of flight in an airplane and a bird are quite similar. Both the airplane and the bird take advantage of the same real world physical principles of aerodynamics, such as Bernoulli's principle, flow fields, and turbulence. Even if the structural accuracy of the airplane to model target behavior (that of the bird) was developed further – in the sense that if the airplane could flap its wings, land with soft collisions, and so on, – it is unlikely that much more would be identified about the neurophysiology or the behavioral organizational principles of bird flight.

Even though the biorobotic control systems mentioned by Webb are a significant improvement over the classic cybernetic models, they still make assumptions similar to the cybernetic approach. To illustrate this, let us consider the example of the simplest control system, the thermostat. Temperature regulation is achieved in the thermostat by minimizing the difference between the actual temperature of the room and that of the set point. The set point is prescribed by an external source and is not specified by the system itself. In contrast, biological control systems have the capacity to internally modify the set points of different subsystems and thus intentionally initiate the transition from one steady state to another, a dynamical process underlying voluntary actions (Feldman & Levin 1995). For example, the human arm as controlled by the nervous system can reach desired positions by prescribing a set point, which is achieved by active engagement with the environment, gravity, inertial, and reactive forces. Muscle activation patterns, forces, torques, and trajectories are not programmed or computed but are largely emergent in biological systems in the process of interaction with the environment. Whereas, even in the most sophisticated examples used by Webb, such as modern robotic approaches called force control models (Wolpert et al. 1998), these variables are directly programmed and computed. The dynamical mechanism of state resetting is largely ignored, simply because the computational principles underlying imitations of biological movements in robotics conflict with the natural, dynamical nature of the resetting mechanism underlying intentional movements.

To emphasize this point, consider the posture-movement problem in biological control of movement as formulated by Von Holst and Mittelstaedt (1950/1973). He noticed that there are powerful neuromuscular mechanisms ("postural reflexes") that generate electromyographic (EMG) activity and forces in order to resist perturbations that deflect the body from an initial posture. At the same time, it is clear that the organism can intentionally adopt different postures. Each new posture adopted by the system might be considered as a deflection from the initial one. The deflection would result in resistance tending to return the system to its initial position. How then is an intentional movement from the initial posture and the achievement of a new posture of the body possible without resistance? It has been established that the nervous system can reset the postural state by changing length-dimensional parameters – muscle activation thresholds (Asatryan & Feldman 1965; Feldman & Orlovksy 1972; Matthews 1959). By resetting these thresholds, the system shifts the spatial coordinates at which an equilibrium posture can be reached and maintained. Thereby, the initial posture appears to be a deflection from the newly specified posture. Therefore, the same neuromuscular mechanisms that produce EMG signals and forces in response to deflections from the initial position produce, without any programming, EMG signals and forces tending to eliminate the deflection from the new posture and thus move the system to it (Feldman & Levin 1995; St-Onge et al. 1997). This postural resetting mechanism was also confirmed by the finding that many systems, including cortico-spinal descending ones, have the capacity to regulate the activation thresholds. Further support stems from recent studies showing dramatic movement problems following deficits in the regulation of activation thresholds in neurological patients (Levin & Dimov 1997).

Stemming from robotics, force control models fail to answer the basic question posed by Von Holst and Mittelstaedt (1950) on how the system can actively move from an initial posture without triggering resistance. By disregarding the empirical mechanism of postural resetting (shifts in muscle activation thresholds), force control models produce movements by overcoming such resistance. For example, Schweighofer et al. (1998) simulated planar point-to-point arm movements using a force control strategy. Their equations show that after the movement offset, muscles generate tonic activity in proportion to the distance between the initial and the final muscle lengths. This implies that the final position is reached by overcoming the resistance to the deflection of the arm from the initial position. Thereby, at the final position, the muscle activity cannot be minimized without driving the limb back to the initial position. This prediction of the force control strategy obviously conflicts with the common observation that after transition of the arm to a new position, muscle activation can be minimized without arm motion. Control strategies that tolerate the resistance to deflections from the initial posture each time when an active movement is produced are highly inefficient in terms of energy costs. Incorporating the empirically established mechanism of postural resetting in a motor control theory comes with a price: the resetting mechanism implies that output, mechanical variables do not need to be directly programmed or computed to make adequate actions. This implication conflicts with the basic, computational principles underlying force control models and thus questions their physiological feasibility, despite their efficiency in robotics.

We conclude that while great strides have been made in biorobotics, there is still a long way to go before robotics can make pertinent contributions to biology. The contributions of biorobotics to biology can be greatly accelerated if engineering approaches take into account the context in which biological systems generate solutions to real world problems.

From reflex to planning: Multimodal versatile complex systems in biorobotics

Jean-Paul Banquet,[a] Philippe Gaussier,[b] Mathias Quoy,[b] and Arnaud Revel[b]
[a]*Neuroscience et Modélisation, INSERM 483, Université Pierre et Marie Curie, 75252 Paris Cedex 5, France;* [b]*Neurocybernetics Group, ETIS Lab, Université de Cergy-Pontoise, Cergy-Pontoise, 95014, France.*
banquet@ccr.jussieu.fr **{gaussier;quoy;revel}@ensea.fr**
http://www-etis.ensea.fr/~neurocyber

Abstract: As models of living beings acting in a real world biorobots undergo an accelerated "philogenic" complexification. The first efficient robots performed simple animal behaviours (e.g., those of ants, crickets) and later on isolated elementary behaviours of complex beings. The increasing complexity of the tasks robots are dedicated to is matched by an increasing complexity and versatility of the architectures now supporting conditioning or even elementary planning.

The edge of biorobotics over plain mathematical modelling. Robotic models result from a back-and-forth interaction between mathematical models' simulation in "gedanken" experiments, and robotic models' experiments in the real world. As such, they do not constitute a different type of, but actually a step further to, classical modelling.

From our experience of complex systems, a few points are emphasized. First, a complete behaving system, as in a robotic model, requires a necessary "horizontal" mechanistic integration, at the basic level of network interactions between different components of the global architecture (sensory, motor, associative, timing, planning, and so on), that enforces self-consistency among coordinated systems (at variance with the "vertical" integration between different levels of explanation alluded to in the target article). This "horizontal" integration is rarely achieved in network modelling characterized by a tendency to build dedicated architectures for specific tasks without caring about their functional integration in a system. Less accuracy or some degree of approximation is the price to pay. The specific import of biorobotics in the "vertical" integration has been to stress the interdependence between agent morphology and control system, and to provide a test of congruence between the model of the agent and the model of its environment thanks to dynamic interactions with the external world during robotic experiments.

Second, model-grounding in the real world by compliance with physical laws (such as spatio-temporal continuity, gravity, and so on) is a lever to the system's simplification rather than a constraint. Learning a real environment (during an individual lifespan or a species' evolutionary period) is paradoxically a source of simplification in this type of solution and solution-space dimensionality. Combinatorial explosion of the solution domain is prevented by probabilistic sorting out of the situations actually encountered in the real world. As illustrations from our model: (1) Assuming spatio-temporal continuity expressed by the formalism of neural fields (Schoener et al. 1995) allows stable target selection and smooth, robust control. (2) Control architectures can take advantage of inherent perceptual ambiguity to perform complex tasks in a simple way, with disambiguation resulting from the dynamics of the behavior. (3) Sensory and motor modules designed in compatible or similar coordinate systems achieve a great simplification of the information flows (Gaussier et al. 1999).

Third, two dimensions of neural organisation, relevant for humans as well as robots, account respectively for the nature of the functions performed (e.g., spatio-temporal processing and navigation, linguistic processing, episodic learning), and the level of performance (reflex, conditioning, planning). The first horizontal dimension specifies, according to combined perceptuo-motor modalities, the nature of the parallel processes performed, whatever their level of performance. It depends on the combination of the multimodal associations (e.g., visual [exteroceptive] and movement related [interoceptive] inputs in the computation of space during navigation; auditory but also visual inputs in speech and language processing; and so on). Each peculiar combination of associations specifies a function. The second vertical dimension specifies the level of performance (stimulus-response reflex behaviour, conditioned automatic behaviour, planned controlled behaviour), whatever the nature of the functions involved. This second sequential, or rather, iterative aspect unfolds in cortico-subcortical loops characterized by the dual process of convergence-contraction and divergence-expansion of information, and gives rise to different levels of pattern-recognition (uni- or multimodal events, transitions, chunks, sequences, plans).

Illustration by a generic spatio-temporal control system. Depending on these two dimensions, dedicated architectures can become generic and, as such, used for multiple implementations according to the input-output nature and the level of processing. We assumed common mechanisms for spatio-temporal processing during navigation and declarative-episodic memory. Both depend on spatio-temporal sequence learning based on a cascade of associations and pattern recognition performed in cortico-hippocampal loops.

The computational model implemented as a robot control system features three levels of organization (Fig. 1a) linked by intrahippocampal and cortico-hippocampal loops. First, a basic hippocampal level learns events (whatever their nature, e.g., places), transitions between events, and chunks. Second, the intermediate level links sensory information to motor responses in relation with drives and reinforcement. Finally, the cortical level links and stores sequences of chunks to form graphs and maps that can be used for planning in relation with goals and motivations. According to the nature of the input-output modalities, the system can be used to learn timing and temporal sequences, motor sequences for

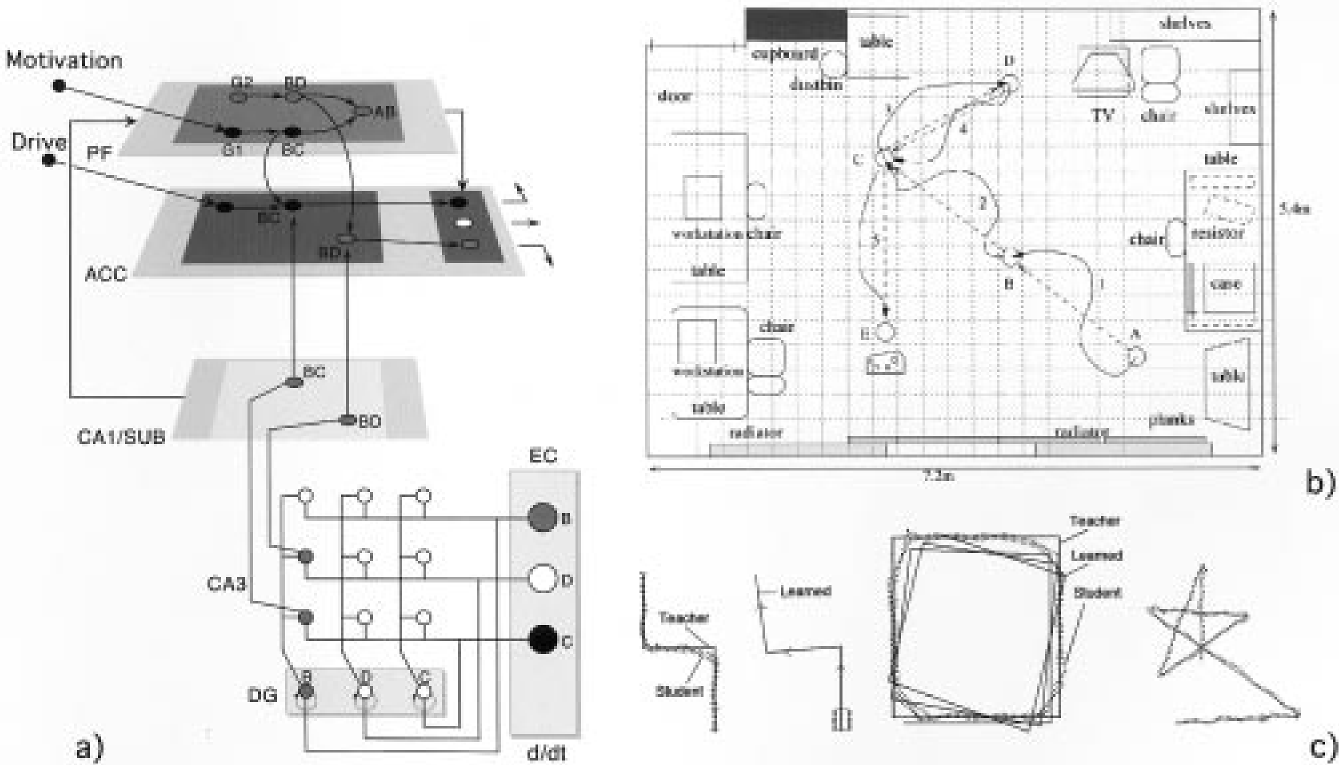

Figure 1 (Banquet et al.). (a) Three-level generic architecture featuring hippocampus (EC, entorhinal cortex; DG, dentate gyrus; CA3-CA1; SUB, subiculum), intermediate level of basal ganglia (ACC, accumbens), and cortex (PF, prefrontal). (b) Learning of places (A,B,C,D,E), transitions, and associated trajectories while exploring an indoor environment. (c) Imitation and learning of the teacher trajectories by a student.

imitation, or temporo-spatial sequences for navigation (Banquet et al. 2001).

In our model of timing and sequence learning, cell populations with different time constants extend an event-related phasic signal for different delays (so learning the timing between two events). The transition between two successive events, thus made co-occurrent, is learned by a hetero-associative network (Banquet et al. 1997; 1998). In one trial, event-transitions and temporal sequences are learned, as for a melody.

The same basic architecture, complemented with a sensorimotor module that feedbacks proprioceptive inputs to the hippocampal system performs protoimitations (Gaussier et al. 1998) and learns arbitrary trajectories (Fig. 1c). The system, based on a homeostasis principle, minimizes the difference between teacher-movement perception and student movement. Structuring the movements of the teacher in movement-transitions allows online imitations and learning of the imitated sequences. A playful version of this robot model (imitating vs. being imitated) is implemented as an aid to understanding autistic behaviour whose main handicap results from a deficit in social interactions (Andry et al. 2001).

The same architecture (Fig. 1a), when receiving visuo-spatial inputs, learns not only transitions between places during the exploration of an environment, but also (thanks to path-integration) the ideal trajectory between two places, and associates this trajectory with the transition (Fig. 1b). Learning can either be latent (Hebbian without reinforcement) or reinforced (higher order conditioning) by a reward. Yet, transition-learning, combined with the propagation along the graph of the activation initiated by goal representation and motivation (Fig. 1a), allows discovering transitions never experienced before, and managing several simultaneously active goals (Gaussier et al. 2001), according to their salience.

The biorobotic models certainly do not constitute a proof of the existence of similar neurobiological mechanisms, but rather, a guarantee of functional realism and plausibility, and a questioning tool capable of suggesting unexpected hypotheses on biological systems.

Models of complexity: The example of emotions

Catherine Belzung[a] and Catherine Chevalley[b]
[a]*EA 3248 Psychobiologie des émotions, UFR Sciences et Techniques, Université F. Rabelais, Parc Grandmont, Tours F-37200, France;*
[b]*Département de Philosophie, Centre d'Etudes Supérieures de la Renaissance, Université F. Rabelais, Tours F-37000, France.*
Belzung@univ-tours.fr **Chevalley@univ-tours.fr**

Abstract: Using the example of the difficulties which emerge when trying to model complex behaviors – such as emotional expression – that result from stochastic interactions between different components, we argue that biorobotics may well describe one possible evolution of certain features of a biological system, but cannot pretend to be a simulation of the whole behavior of the system.

Robots are believed to mimic the behavior of biological systems, but do they model complex behaviors, such as emotional expression? Several robots have been built that include the so-called "emotional model." For example, the AIBOT, which is a home entertainment robot simulating a dog's behavior, seems able to express "emotional behavior" (Pransky 2001), while human head-like robots can communicate with humans by changing facial color expression (Miwa et al. 2001a). The major reason for including such abilities may be to advance the sociability of the robots; emotional models facilitate communication between robots and humans (Miwa et al. 2001a) and machines carrying them are so much more attractive (Ogata & Shigeki 2000). Other reasons for including "emotional models" could be to facilitate adaptation to natural and unpredictable environments (autonomy of the system), and to improve cognitive processes. Indeed, it has been suggested that emotions may be a process crucial for cognition (Chevalley & Belzung 2001). In fact, there is substantial evidence indicating that emotions may be mediators between low-level reactive behaviors and high-level rational behaviors. What kind of model do these "emotional systems" use? The human head-like robots use the so-called "equation of emotion" (Miwa et al. 2001b) which consists of the following process: (1) the robot senses the stimulus (Miwa et al. 2001a); (2) the robot appraises the stimulus according to three dimensions (pleasantness, activation level, certainty); (3) the robot generates a "mental state" using an equation (called "the equation of emotions") integrating the three dimensions; (4) it loads the response, which can consist of a modification of its facial color by using red EL (electroluminescent) sheets. Other models are based upon the imitation of the human endocrine system to adjust various internal conditions such as motor output or sensor gain (Ogata & Shigeki 2000).

So, the models used are based upon the production of a serial linear process, occurring at the psychological or the physiological levels. This is quite a simplistic modeling, because in biological systems emotional expression may in fact result from complex interactions between different causes, including ones related to the natural history of the species (genetic factors) or to the events the subject has been faced with (epigenetic factors), to psychological states, to brain circuitry, neurotransmitter systems, and so on. Each of these factors participates in the generation of emotions by activating a complex set of parallel distributed processes (Chevalley & Belzung 2001), which permanently interact with all the others so as to yield an unpredictable response. Therefore, there may be some elements of the behavior of biological models which cannot be simulated in robots. Furthermore, it is rather probable that even though robots may imitate some aspects of the emotional expression (such as changes in face color), they cannot have an emotional feeling. This further emphasizes how impossible it is to mimic emotional behavior.

The case of emotions well exhibits the difficulties we face when attempting to model human behavior. But it also exhibits the ambiguities of our conceptions of what modeling is. The methodological approach that was associated with modeling at the time of a crisis in the fundamental concepts of Mechanics in the 1880s was a very sophisticated one. Heinrich Hertz (1894/1956) was the first to use the word "model" in connection with a new conception of the "theory of knowledge" (see, notably, his definition of what is a "dynamical model" in Bk. II, sect. 418), and this inspired people as different as Boltzmann, Wittgenstein, Cassirer, Bohr, and Heisenberg. Assuming that a model is a representation that is a construct of the mind and may have no resemblance whatsoever to the thing it represents, Hertz based his epistemology of the *Scheinbilder* on the idea that the agreement between Mind and Nature can be compared to the agreement between two systems one of which is a model of the other. There must exist between the two something like what Helmholtz (1878/1921) had called "parallelism in law-likeness": namely, there must be a strict correlation, not between the system modeled and its model, but between the law of *evolution* of the system modeled and the law of *evolution* of the model. Among other motivations, such as introducing "hidden masses" in the science of Mechanics, this was meant at the time as an argument against all naïve "pre-Kantian" conceptions of knowledge based on the notion of a resemblance between things and ideas, or between systems and their symbolic expression. To Hertz, the benefit of modeling was to allow us to focus not on objects, but on law-likeness (*Gesetzlichkeit*).

From the beginning, then, a model was different from a simulation. This is the background of Bohr's claim that a representation of the atom could not in any way "look like" the planetary system, while it should account for the discontinuities observed in experimenting on radiation phenomena. Building a model is not simulating a process, rather, it is building one possible interpretation of the evolution of a set of data or any kind of observed behavior.

If that is the case, it follows that biorobotics may well describe one possible evolution of such and such feature of a biological system, but that it cannot pretend to be a simulation of the whole or entire behavior of the system. To be slightly provocative, we might note that, indeed, there exist excellent models of the behavior of living beings, and notably of emotional behavior (love, avarice, jealousy, and so on): these are the many theatre plays, poems, and great novels that our literature has produced. In the same year, Racine and Corneille both wrote the story of Berenice, thus proposing two different modelings of the same predicament. Indeed, in the sciences too, our understanding seems bound to bounce from one representation to another, while renouncing the fallacious hope to ever find the one true theory.

Biorobotics researcher: To be or not to be?

Carolina Chang
Departmento de Computación y Tecnología de la Información, Universidad Simón Bolívar, Caracas, 1080, Venezuela. **cchang@ldc.usb.ve**
http://www.ldc.usb.ve/~cchang

Abstract: Much confusion exists within the robotics and the biology communities regarding the definition of biorobotics and the aims and strategies that characterize this approach. Not even the basic criteria for identifying biorobotic research are being applied consistently. Barbara Webb has taken a crucial step towards setting a common ground from which biorobotic systems can be described, analyzed, and compared.

In having organized conference sessions on biorobotics, co-edited a journal issue on biomimetic robotics (Chang & Gaudiano 2000), and also in everyday teaching and research, I have come across many diverse views about the biorobotics approach. On the one hand, the approach is often misunderstood or attacked. On the other hand, it is drawing an increasing interest from the robotics community. At conferences and talks I have sensed a special interest and enthusiasm for this field, to the extent that some people seem to believe that a piece of research is "better" or "more serious" if its author can claim that the work belongs to biorobotics.

Many researchers who focus on engineering and classical robotics simply don't understand what the goal of biorobotics is. This may be because their robots achieve more striking results for the time being. So, what would be the point of having, for example, a robot with a model of the rat hippocampus, if such a robot can navigate only in simple rectangular environments? By constructing symbolic models of the environment and planning usage based on these, AI researchers have built robots that, to some extent, can navigate more complex indoors and outdoors environments. This and other kinds of models have proven successful for robot navigation tasks. This approach is fine as long as one is interested mainly in getting the robot moving around. However, it will not be useful if one is interested in generating testable hypotheses for biology. Of course, biological relevance isn't and shouldn't be the aim of an engineered solution. When two pieces of research have different goals, results are not directly comparable. Therefore, we cannot say a priori whether a robot with a model of rat hippocampus is better or worse than a robot with a symbolic model of the environment. It all depends on what we are pursuing with the model, which in turn determines what dimensions best describe it. Clearly, people will disagree on what the goals and the important dimensions should be.

I have no problem with this type of disagreement. What really puzzles me is finding work that claims to be based on a biorobotic approach, but that, in my opinion and according to Webb's article, does not belong to this field.

Work that is not based on robots. I agree with Webb that simulations and sensing systems that do not end in actions are not robotic systems. I don't think, however, that most people share this understanding of robotics. The need for physical robots has been recognized for years in the behavior-based robotics approach (Brooks 1991), and this principle has been the source of much debate. Part of the problem here is due to the sometimes vague definition of what constitutes a robot. I used to think that lay people, especially kids, were doing better than some researchers at classifying robotic from non-robotic systems. However, today the terms "robot" and "bot" are being used widely, especially on the Internet, to name software agents. Even ordinary people are beginning to regard them as robots. In robotics, it is common to find researchers who believe that simulated robots are robots. Biorobotics should keep letting people know that it is devoted to physical robotic systems.

Work that has no grounding in biology. Just because a given robot has six legs, it is not a biorobotic insect. Similarly, even though some robots learn by means of a neural network controller or some sort of reinforcement learning mechanism, they are not biorobots if they do not "address a biological hypothesis or demonstrate understanding of a biological system" (target article, sect. 2.4). Unfortunately, the terms biorobotics and biologically-inspired robotics have been used frequently to describe systems that are only loosely related to biological systems. I tend to prefer the term "biomimetic robotics," because mimicry in a biological sense implies a close resemblance, sometimes of one organism to another of a different species (Wickler 1968). In our case, we mean a close resemblance of a robot to a biological organism, at the anatomical, physiological, functional or behavioral level. But the term "biomimetic robotics" has had the same fate as the terms "biorobotics" and "biologically-inspired robotics." It doesn't matter how we name this research area, the wrong usage of its name will continue until its definition is stated clearly and broadly.

Finally, if the biorobotics approach is not well understood within the robotics community, it is even more obscure for a large part of the biology community. Some biologists are disappointed by the minimal relevance of many so-called biologically-inspired robots, others are concerned about the lack of scientific methodology in robotics research, and yet others are simply not aware or not optimistic about the potential results of the intersection of biology and robotics.

We have been lacking a common ground from which the biorobotic approach could be described, analyzed, and discussed. In this regard, Webb's article stands as a very valuable contribution to the field. I strongly believe that it will help us achieve a better understanding of the biorobotics approach. I won't argue whether or not the seven dimensions that she has proposed are the best choice for framing the approach. As I said before, I anticipate some disagreement on the selection and description of these dimensions. What is really important is that Webb has plotted the biorobotics approach very thoughtfully and very carefully along the chosen dimensions. We can now locate our own research within the context of a well-characterized approach.

Webb's efforts to provide and promote a standard definition of this field, and a metric by which to classify work, should be commended. The continued careless, gratuitous use of terms such as biorobotics or biomimetic robots only hurts this fascinating and promising field of research by undermining its credibility.

Programs, models, theories, and reality

Robert I. Damper
Department of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom.
rid@ecs.soton.ac.uk **www.ecs.soton.ac.uk/~rid**

Abstract: The question "Are 'biorobots' good models of biological behaviour?" can be seen as a specific instance of a more general question about the relation between computer programs and models, between models and theories, and between theories and reality. This commentary develops a personal view of these relations, from an antirealism perspective. Programs, models, theories and reality are separate and distinct entities which may converge in particular cases but should never be confused.

Webb asks: "Are 'biorobots' good models of biological behaviour?" It seems her preferred answer – like that of other biorobotics practitioners – is that they can be. (Why else go to all that bother?) The difficulty is in justifying this answer in the face of proper scientific scepticism. Webb's review of the philosophy of science literature on the relations between simulations, models, and theories in general uncovers a degree of disagreement (if not downright confusion) which barely helps the quest for enlightenment. Her solution is essentially pragmatic: She lists seven dimensions on which robotic models can vary and uses these to prompt a set of "acid tests." Any particular model should, in some degree, pass at least a subset of these if claims of usefulness are to be substantiated. But robots are artificial computational devices irrespective of whether there was a biological inspiration for their construction or not. Hence, Webb's question can be seen as a specific instance of the more general question: "Can computer programs serve as good models of real-world phenomena?" It is useful to reframe things thus, since the latter version has been extensively argued.

It could be objected that a robot is more than a computer, and a behaving robot is more than a computer program: One might not normally think of a computer as mobile and autonomous. But with increasing microminiaturisation and concomitant advances in software engineering, computers are becoming ever more pervasive in a variety of mobile applications, often displaying a considerable degree of autonomy of function. In any event, what are the wheels and associated driving motors of a mobile robot but peripheral output devices, and what are its sensors but peripheral input devices? Webb hints at this when she writes that the hardware implementation of biorobots "does not necessarily make them less abstract than computer simulations." According to the cartoon below, an amusing role-reversal play on the Turing test, robots and computers are certainly synonymous!

With these preliminaries out of the way, then, the two questions which I would like to discuss further are:

1. What is the relation between computer program and model, and between model and theory?
2. What is the relation between model/theory and physical reality?

As Webb convincingly demonstrates with reference to the literature, a range of opinions exists on question 1, but I hold to the view succinctly expressed by Moor (1978, p. 220) some years ago. As he pointed out, "computer scientists often talk as if there is no distinction among programs, model, and theories; and discussions slide easily from programs to models and from models to theories." Webb tacitly assumes that we already have some computer program (running, say, on a mobile robot) which is to be viewed as a model – the task is to decide if this model is "good" (i.e., has some useful or interesting relation to theory and/or reality) or not. This assumption is common in the field. For instance, Dean (1998, p. 64) writes: "For a biologist, an animat is a version of a familiar tool, a model." So let us agree to ignore any distinction between "robot" (or "animat," or "program") and "model" and move on. Now, Moor's view is that "a computer model does not automatically embody a theory in the important sense that one knows what the theory is . . . . The theory must be statable independent of the program" (p. 221). In other words, sometimes a computer model can embody a theory but sometimes it does not.

Note that we have followed Moor's careful wording here, when he writes of a model which can "embody a theory." The concept is similar to Nagel's (1961) "model of a theory" which can be either a "substantive" or a "formal" analogy (p. 110). The former takes a more or less familiar physical system, with known properties or relationships between its elements, and uses it to suggest corresponding properties or relationships in the theory which is being developed so as to understand a less familiar system. This sort of correspondence is no doubt what Webb has in mind in mentioning "substitution" as a mechanism of abstraction in biorobotics (sect. 3.1). Often the terms used to describe the familiar, model system are simply transplanted into the less familiar situation. That is, the model is "semantic." The "formal" analogy is typified

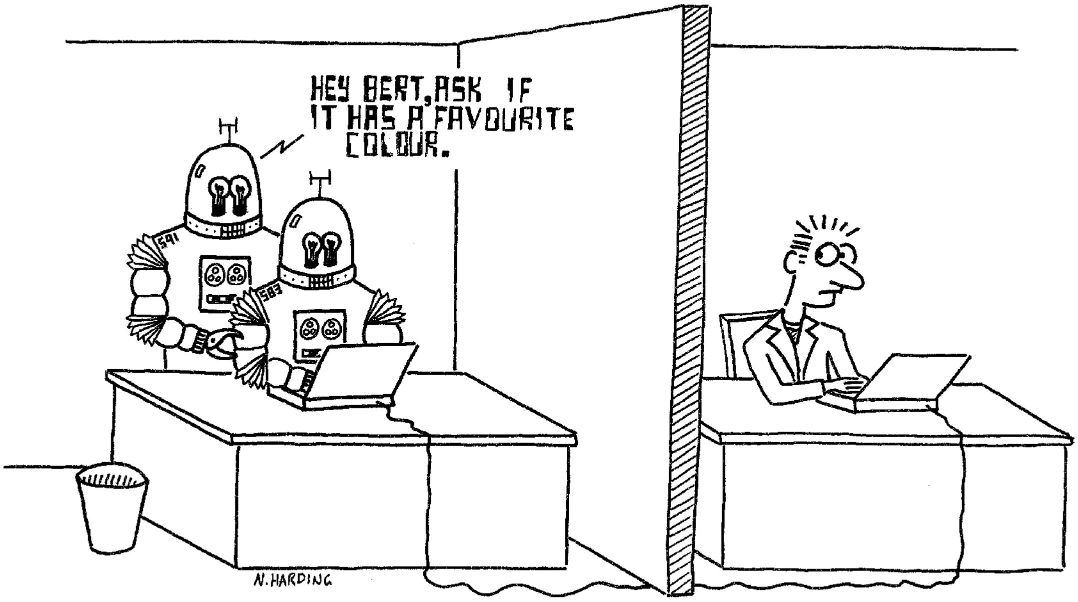

Figure 1 (Damper). (Unpublished cartoon. Copyright Nick Harding. Permission to reproduce granted by Nick Harding.)

by a mathematical model in which the relation between abstract objects, devoid of content, and operations upon these objects, are exploited in defining the model. The difficulty of devising unequivocal definitions, which Webb takes up in detail, accords with Nagel's warning: "The only point that can be affirmed with confidence is that a model for a theory is not the theory itself" (p. 116).

One can easily think of models which embody a theory in Moor's sense, statable independent of any particular program implementation. One obvious example (of the "mathematical" kind) is the Hodgkin-Huxley model of the nerve cell action potential (Hodgkin & Huxley 1952). As an example which does not pass the "independently statable" test, Moor gives the celebrated AI program ELIZA (Weizenbaum 1966) which embodies not a theory of "intelligence" but a collection of ad hoc programming tricks. A useful and interesting argument might then be had about which side of the divide to place currently-popular connectionist models, but space does not allow us to develop this topical issue here. Instead, the reader is referred to the relevant literature; particularly Zipser (1992), O'Reilly (1998), and Dror and Gallogly (1999).

Turning now to question 2, Webb (perhaps wisely) declines "to enter this thorny territory" but briefly argues "for no more than an instrumentalist position." It is commonplace in the philosophy of science to set instrumentalism ("antirealism") against realism as irreconcilable alternatives that one chooses between according to taste – see Kim (1998, p. 90) for a succinct statement of the issue and Nagel (1961, Ch. 6) for a more lengthy development. As a practically-minded engineer rather than a philosopher, and not being given to adherence to any particular school, I believe both have something to offer. However, instrumentalism definitely takes precedence. For, until a theory has proved itself as an instrument of prediction, it is not even worth considering its claims to reality. Take the standard model of physics – a highly-successful theory from the instrumentalist (or any other!) viewpoint. As it has developed, as its predictive power has improved, so it has tended to become a better and better "fit" to reality. For this is how good science works. Still, it remains a model/theory, logically distinct from reality, and not to be confused with it.

Biorobotic models can contribute to neurobiology

Fred Delcomyn

Department of Entomology, University of Illinois, Urbana, IL 61801.
delcomyn@life.uiuc.edu http://www.life.uiuc.edu/delcomyn

Abstract: The idea that biorobots can be used as a testbed for the evaluation of hypotheses about how an animal functions is supported. Generation of realistic feedback is a major advantage of biorobotic models. Nevertheless, skeptics can only be convinced that this approach is valid if significant biological insights are generated from its application.

This is a wonderful paper. In the first section, Webb analyzes the concept of a "model" and describes the various types. Although arguments over what a model is and whether a model can be useful in understanding biological phenomena are certainly not new (e.g., Vowles 1964), Webb does an unusually thorough job of considering the various points of view. In the second part of the paper, Webb makes a good case for the view that a "biorobot" represents a particular kind of model, one that offers considerable advantages for the study of biological systems. In particular, she suggests that biorobots can be designed and built as "models *of* animals" (her emphasis). By this she means that robots can be used as a kind of testbed for the evaluation of hypotheses about how an animal functions, or how the behavior of that animal is controlled or coordinated.

I wholeheartedly support this position. A significant difficulty with most other forms of modeling, be they paper and pencil circuit diagrams or computer simulations of biological processes, is that they require the modeler to make many assumptions about factors that may influence the performance of the model but that are not directly a part of it. For example, a variety of formal models have been developed of the way in which insect locomotion (walking) is generated and coordinated (Dean 1991; Graham 1977). The main objective of these models is to show how the coordinated patterns of motor output to leg muscles that underlie walking can be generated. However, although sensory feedback from the moving legs is an important component of these models, the sensory signals that are used are arbitrary, in the sense that the model is designed so that it receives the types and strengths of sensory signals that the researchers believe will be generated at various times during the stepping cycle. Although the signals used in the models are derived as much as possible from experimental data, they are nevertheless unlikely to be as variable as sensory signals in living animals are (Ridgel et al. 2000). Therefore, in so far as values of the feedback used in the models are at variance from the values of the sensory feedback actually present during walking in an insect, the models will yield an output that does not accurately reflect what would happen in the living animal.

By building a real mechanical device, researchers can avoid this difficulty. If electromechanical analogs of important sensory structures on the legs are built into the robot, sensor feedback similar to that actually produced by sense organs in a walking insect will be generated as the robot walks. Hence, a computer model of the control system for walking can be evaluated by making it the control system for the robot and analyzing how well it performs. Not only will sensor feedback then depend on actual movements of and loads on the legs, the necessity for the control system to control a physical body will allow an evaluation of the control architecture (which incorporates ideas of how the central nervous system controls walking) that will be considerably more useful than any test that can be run on a computer simulation model (Ritzmann et al. 2000). This is the basis of the trend toward building robots that are modeled as closely as possible after particular animals (e.g., Delcomyn & Nelson 2000; Pfeiffer et al. 1995; Quinn & Ritzmann 1998).

Nevertheless, Webb's paper will undoubtedly have its detractors, either those who do not consider modeling a useful activity in neurobiology, or those who would emphasize the very real difficulties that modelers must operate under to assure the biological relevance of what they do. For some, the reaction will undoubtedly be more emotional than rational. They like to work with living tissue and may have little regard for formal models. For others, the question of how much like the animal the model must be may seem overwhelming. The fact that neither computer simulations or robots can really be "like" the animals they attempt to emulate renders virtually any model biologically irrelevant in their minds.

It is certainly possible to build a strong case that, given the variety of detailed computer or robotic "models" of biological systems being designed and investigated now – as discussed in part by Webb – such modeling has entered the mainstream (e.g., Beer et al. 1998; Churchland & Sejnowski 1992). However, an objective analysis of papers on the topic relative to more traditional physiological papers might suggest otherwise.

The fact of the matter is that modeling is an approach that does not come naturally to most neurobiologists, who are attracted to the field for reasons other than an interest in computers or robotics. It is up to those of us who embrace these methods to demonstrate the biological relevance of the models with which we work if we expect our research to be recognized by our physiologically-minded peers as relevant to what they do.

For this recognition to occur, we must do two things. First, we must be relentless in our attempts to link our models to biological reality. Sentiments such as "our control system generates behavior similar to that of an animal and therefore it reflects what happens inside the animal's nervous system" are seductive but not sufficient. Modelers must be their own harshest critics when it comes to accepting simulation or biorobotic results as relevant to biology.

Second, we must make every effort to investigate problems of central biological interest and to tie our modeling results back to these problems. Shedding light on important questions will always catch people's attention. Anything less will doom this promising approach to the backwaters of research, a kind of fringe that may be interesting in a "gee whiz" kind of way, but one without relevance to the main problems in which researchers have an interest.

ACKNOWLEDGMENTS
Research and preparation for this commentary have been supported by grant NAGS-8781 from the National Aeronautics and Space Administration, and grant BCS 92-16562 from the National Science Foundation.

Robotic search: What's in it for comparative cognition?

Carlo De Lillo

Department of Psychology, University of Leicester, Leicester LE1 7RH, United Kingdom. **cd12@le.ac.uk**
http://www.le.ac.uk/psychology/cd12/

Abstract: Although the advantage of biorobotics over traditional modelling tools is not always evident from the studies on animal search addressed in the target article, this commentary argues that testing different robotic architectures and specific biological organisms in structured search spaces, where environmental constraints matter, might prove one of the most promising research strategies in comparative cognition.

The strongest rationale for biorobotics, permeating most of the dimensions along which models are evaluated in the target article, resides in its potential for providing complete (from perception to action) physically instantiated models which are capable of capturing the interrelationships between agent and environmental constraints (the very convincing "medium" argument).

As such, robotic simulations are particularly appropriate where the structure of the environment plays a primary role in goal directed activities such as spatial search and foraging. Simple robotic architectures aimed at implementing foraging behaviour are provided by behaviour-based approaches in robotics (Arkin 1998). However, these are detached from their biological counterparts. "Wander – acquire – retrieve" characterisations of foraging tasks do not do justice to the fact that animals never, not even micro-organisms, ramble randomly; that their search strategies can become more sophisticated with experience; and that animals often keep track of locations explored to avoid wasteful re-visits of exploited resources (Bell 1991).

Since biological relevance is the first dimension used by Webb to evaluate robotics research, the lack of biological plausibility of some of these models justifies their exclusion from the studies listed in Table 1. The target article is, however, too parsimonious in providing alternative examples of successful bio-robotic modelling of animal search.

The only article explicitly focusing on search in animals and robots included in the list (Gelenbe et al. 1997) reviews studies on search in artificial agents on the one hand, and biological ones, on the other, but does not offer specific examples of robotic simulations which have lead to insights concerning the mechanisms that control search strategies in animals. The examples considered there are of formal models of decision-making based on data obtained from operant "analogues" of foraging tasks where, for example, pecking behaviour in birds might show how rates of reinforcement control "patch" (where a patch is one of the pecking keys) choice (e.g., Stephens & Patton 1986).

Such studies have had an impact on theories of animal behaviour, and it is possible to see their relevance for engineers, for example, who – with specific applied aims in mind – seek to empower their systems with useful decision-making rules. Nevertheless, the added value of a robotic implementation of these formal models for the understanding of how animals control their behaviour in a foraging task is not transparent.

If the spatial distribution of patchy resources and the structure of the environment is not relevant to the task, and the animals are faced with binary choices between adjacent keys, mathematical models might be more powerful than robotic simulations for testing specific theoretical hypotheses. The benefits of biorobotic modelling are also obscured in other paradigms where the interrelationship between organisms and environmental constraints is not central to the task, such as in delayed Matching-to-Sample, even if the experimental apparatus allows the presentation of the same stimuli to rats, monkeys, and robots (Touretsky & Saksida 1997).

The scope of robotics could be much greater in the context of tasks, such as those featuring multiple loci to search for hidden baits, where organisms have the opportunity to exploit the spatial structure of the search space and organise adaptively their patterns of movements on the basis of the constraints it affords.

Systematic manipulations of the structure of the search space reveal that monkeys' search efficiency is affected by the spatial constraints afforded by the environment (De Lillo et al. 1997), that they restructure their search patterns in ways that minimise the implicit costs (in terms of memory load or distance travelled) of unprincipled search trajectories (De Lillo et al. 1998), and that they may plan up to two steps ahead in their sequence of movements (Cramer & Gallistel 1997).

The assessment of the behavioural match between different robotic architectures and different organisms faced with the above tasks, which can be easily set up for use with animals and robots, might reveal the extent to which particular forms of control suffice to account for the search behaviour of organisms characterised by different levels of cognitive complexity. For example, the extent to which search organisation can be explained on the basis of a limited set of simple hard-wired rules, repeated at each step of the behavioural sequence; or, what sort of costs have to be monitored for dynamic adaptation to environmental changes to emerge; or, whether or not planning (a controversial issue among proponents of different approaches to robotics) is required to match the behaviour of the target system.

As mentioned by Webb, often robotic research has been able to show that when constraints are taken into account, simple control procedures suffice in explaining apparently complex behaviours. In problems where the spatial structure of the environment is unimportant, even minimalist approaches to non-robotic simulations can be useful for disambiguating the sufficient conditions for successful task solution and highlight the problems faced by simple systems when the structure of the problem changes slightly (De Lillo et al. 2001). However, in spatial search non-robotic modelling would be particularly exposed to the risk, as rightly pointed out by Webb, of overlooking some of the numerous potentially relevant variables, such as the complexity of the spatial structure of the environment, or, in this specific case, the fact that the organism continuously changes vantage point during search.

A final point concerns the fact that, apart from a few examples, to date biorobotic models seem to have largely targeted invertebrate species (as reflected in the range of biological organisms surveyed in Webb's article). This obviously undermines the scope of biorobotics in animal cognition research.

However, if over-reliance on available data is avoided, and an effort is undertaken to develop new research programs specifically aimed at assessing the behavioural match of particular artificial and biological agents – in the appropriate tasks and under systematic experimental manipulations – the (otherwise extremely valid) argument for targeting simple systems, of which we have detailed behavioural and neurophysiological knowledge, does not have to be taken too far. Under these conditions robotics has the potential to become one of the best heuristic tools in animal and comparative cognition.

An intentional dynamics approach to comparing robots with their biological targets

Judith A. Effken[a] and Robert E. Shaw[b]

[a]*College of Nursing, University of Arizona, Tucson, AZ 85721;* [b]*Department of Psychology, University of Connecticut, Storrs, CT 06269.*
jeffken@nursing.arizona.edu **reshaw@uconnvm.edu**
http://www.nursing.arizona.edu/scripts/facpageV2.asp?txtLoginName=jeffken

Abstract: After identifying similarities in the paradigmatic problems of biorobotics and ecological psychology, we suggest a way to compare the performance of robots with that of their biological targets. The crucial comparison is between the intentional dynamics of the robot and those of the targeted animal, a measure that depends critically on recognizing and describing the underlying affordance-effectivity match of the target system.

As ecological psychologists, our response to Webb's target article was largely positive. In this commentary, we first identify areas of similarity in the fundamental problems and emerging solutions of biorobotics and ecological psychology, then offer a potential answer from ecological psychology to what Webb correctly characterizes as an unsolved problem in robotics: that is, how best to compare the performance of the model with the target biological system.

Ecological psychology and biorobotics have much in common, both in their paradigmatic problems and in their proposed solutions (Effken & Shaw 1992). In each domain, the paradigmatic problem has been solving how animals/robots can navigate successfully through a cluttered environment. Increasingly, the solutions proposed by bioroboticists for robot movement and vision resemble those of ecological psychologists. Bioroboticists have found that coordinating a robot's movement requires reducing the number of degrees of freedom to be controlled. In ecological psychology, the problem of coordinating human movement has culminated in the discovery of low-dimensional dynamical laws that govern the complex coordinative structures that subserve any physical action (e.g., Turvey 1990). Gibson (1958) intuited that animal navigation would turn out to be governed by low dimensional rules for the perceptual control of action. For example, if you want to effect a soft contact with an object – or another individual – move in such a way as to cancel the accelerative centrifugal flow of the optic array, which, since the flow is a function of visual angle tangent, will bring the approach into soft contact with the object's surface precisely at the time when the visible contour of the image is maximal. Ecological psychologists have gone on to uncover the informational basis for this kind of solution; specifically, the complex visual invariants that animals use to guide their navigation (e.g., Lee 1976; and see Warren 1998, for an excellent review). Gibson (1986) made it clear that his proposed "rules" were not commands issued by a brain, but emerge from the animal-environment system. As Webb notes, in biorobotics, similar low-dimensional rules have begun to replace complex computational algorithms for robot vision (e.g., Franceschini et al. 1992; Webb & Scutt 2000; Williamson 1998).

Although she advocates the inclusion of context and testing in the real world, Webb defines medium rather narrowly as "what the model is actually built from." From that point of view, the crux of the problem is whether the medium supports or interferes with the evaluation of performance. However, from an ecological psychology point of view, including environmental constraints is crucial for understanding the problem and making the actor (human or robot) contribution simpler (Effken et al. 1997; 2001). For example, the ecological approach to interface design begins by identifying environmental constraints because they are more stable than social or user constraints (Vicente 2000). Indeed, ecological psychology is founded on a series of dual (mutual and reciprocal) relationships, of which the most general is the relationship of animal and environment. The environment comprises sources of information (affordances) for an animal with particular real-time control capabilities (effectivities). The same environment (e.g., a chair) affords different action opportunities for an adult human and a mouse because of their different effectivities. Similarly, the same adult human may respond quite differently in different environments (e.g., at home and at work). Consequently, it may be more useful to model specific target systems and derive more general rules, as Webb suggests.

Webb notes that comparing the robot's behavior to the target remains a weakness in most biorobotics studies. If the minimal unit of analysis is the affordance-effectivity match in a particular animal-environment system, then a very different approach to measurement may be required. Measuring the fit of the animal to the environment can be achieved by using dimensionless invariants (e.g., Shaw et al. 1995; Warren 1984; Warren & Whang 1987) or measures derived from conjugate field theory (Kadar & Shaw 2000).

However, the crucial comparison to be made is between the intentional (goal-directed) dynamics of the robot and the intentional dynamics of the targeted animal. This match can be measured by the strength of a coupling term between the two models (the observed robot path space versus the intended animal path space). Selecting an appropriate measure of coupling strength depends critically on recognizing and describing the underlying affordance-effectivity match of the target system. The environment, under Gibson's affordance description (what the environment offers an animal, for good or ill), sets up opportunities for action by the target actor. Consequently, any description of the environment must include those affordances that the target animal requires to remain viable (i.e., able to meet its needs and satisfy the goals that support the satisfaction of those needs). Similarly, any description of the target animal must include the effectivities (real-time control capabilities) needed to realize the essential affordance goals. This affordance-effectivity match provides the criterion against which the robot's performance must be compared.

For measurement purposes, the paths the robot can potentially follow can be partitioned into those that are functionally useful, those that are irrelevant, and those that are actually counterproductive for achieving the goal. This partitioning can be achieved by using the inner product (the metric tensor of the path space), defined as the difference in the direction cosines of the two paths (target's path versus model's path) multiplied by their projected lengths. From this an overlap integral (index of agreement) can be computed. The overlap measure is based on the integral of this inner product difference measure and provides a kind of correlation ratio between the path set of the robot and the path set of the target animal. This correlation ratio provides an "index of agreement" that is not sensitive to nonlinearity, as ordinary correlations are, but is sensitive to asymmetry between the systems coupled (e.g., the model system is led by the target system but not vice-versa). By using this approach, researchers can quantify the affordance-effectivity match as a function of how well the behavior of the model can be predicted by the behavior of the target.

Biorobotic simulations might offer some advantages over purely computational ones

Donald R. Franceschetti

Department of Physics and Institute for Intelligent Systems, The University of Memphis, Memphis, TN 38152. **dfrncsch@memphis.edu**
www.people.memphis.edu/~physics/drf/FRANCESC.HTML

Abstract: A slight modification of Webb's diagrammatic representation of the dimensions for describing models is proposed which extends it to cover a range of theoretical models as well as material models. It is also argued that beyond a certain level robotic simulations could offer a number of real advantages over computer simulations of organisms interacting with their environment.

The target article provides a fairly compact framework within which physical robotic models of biological systems can be classified. The framework is summarized in the article's Figure 2 in which the system being modeled is itself the center, with its models then characterized along seven dimensions depending on the abstractions or generalizations made of the system. The notion that the best possible material model of a system is the system itself, drawn from Rosenblueth and Wiener (1945) has a certain logical appeal, but raises a serious question. If, as these authors note, the main use of models is as an aid to understanding phenomena that cannot be understood as directly observed – and most biological processes are likely to persist in this category for some time, – then the system itself is at the same time inherently unusable as a model of itself. On the other hand, in the study of organism-environment interactions, the system being modeled includes both the environment and the organism, and the biorobotics approach allows these to be separately modeled, often with the environment being taken as the most accurate model of itself, and, within the robot, some of the component processes by themselves as well.

A case can be made for visualizing the target article's Figure 2, not as laid out in a plane, but as transcribed to a globe, with the label "identity" at one pole and the seven boxes placed about the equator. The outgoing lines in the figure then become longitudes which would again converge at the opposite pole, perhaps labeled "formal" which would be reserved for the highly abstract, general and symbolic models which have provided insight into some of the basic biological processes common to all, or at least, a great many living organisms. While this geometrical representation should not be taken too literally, it is reasonable to view the set of very simple, abstract and symbolic models as representing a convergence of the outgoing lines as one moves away from the material and towards the theoretical. The axis of the globe would then correspond to the single line continuum suggested by Shannon (1975). The globe thus envisioned has a certain Aristotelian character, but that need not militate against its usefulness.

Near the formal pole of this structure one would find such highly abstract but informative models as Von Neumann's (1966) cellular automata, which try to capture the essence of reproduction; Turing's (1952) description of morphogenesis, which still stimulates work in developmental biology; and Fontana's (1992) algorithmic chemistry, which has been used to model the origin of metabolism in prebiotic days. More sophisticated models of the dynamical system (van Gelder 1998) or dynamical field type (Thelen et. al. 2001) would appear in the lower hemisphere but closer to the equator.

It is clear that the lower hemisphere belongs to computer simulation, except perhaps the area near the pole, where closed form mathematical results might be obtainable. Near the "identity" or material pole, it is tempting to think that robotic implementations or animats might have the upper hand. There are several reasons why this might be so. As noted above, when the focus is on sensory and motor behavior, one must model both the organism and its environment, and in biorobotics these models need not be at the same level. If the robot is to scale, and the robot sensors detect the same features of the environment as the organism, then the environment can serve as its own model. There is then no need to worry over how best to model the physical properties of, for example, desert sand or the internal structure of a chemical plume. Even if the robot is not to scale or senses a different stimulus, it still may be relatively easy to build a credible model of the environment. Although keeping track of the organism boundary is certainly possible in a computer simulation, it can be computationally expensive. It is done automatically with a robot. Some physical phenomena within the organism, say, diffusion within an irregularly shaped space, are likewise somewhat expensive to model computationally, but solve themselves in a material model.

The best models are like the best students – one gets more out of them than one puts in. This was certainly true historically for the simple theoretical models assigned to the formal pole, and Webb cites instances in which it is becoming true for biorobotics, still a very youthful endeavor. Despite the inevitable introduction of the abstract in the definition of any model, modeling, or at least the drawing of conclusions from them, remains an empirical affair. Robotics is still a very young technology, and it is far too soon to gauge what the limits of its contributions to the understanding of behavior will be.

Models as implementations of a theory, rather than simulations: Dancing to a different drummer

Stan Franklin

Institute for Intelligent Systems, University of Memphis, Memphis, TN 38152.
franklin@memphis.edu **http://www.msci.memphis.edu/~franklin**

Abstract: Robots, as well as software agents, can be of use in biology as implementations of a theory rather than as simulations of specific real world target systems. Such implementations generate hypotheses rather than representing them. Their behavior is not predicted, but rather observed, and is not expected to duplicate that of a target system. Scientific knowledge is gained through the testing of generated hypotheses.

Robots, or software agents, can be of value to biology as generators of hypotheses as well as in the more traditional ways so well described by Webb. Let me illustrate with a software agent example. Essentially the same architecture could control a robot. Intelligent Distribution Agent (IDA) (Franklin 2000b; Franklin et al. 1998) implements global workspace theory, a psychological and neuropsychologicial theory of consciousness and cognition (Baars 1988; 1997; Franklin & Graesser 1999), as well as parts of other theories (Barsalou 1999; Glenberg 1997; Kintsch 1998; Sloman 1999). Supported by the US Navy, IDA is intended to replace a human personnel agent in the real world task of job distribution (Franklin 2001).

Models such as IDA have the potential to play a synergistic role in modeling a scientific theory. The theory constrains the design of the agent or robot that implements (models) that theory. While a theory, including many from cognitive neuroscience, may be abstract and only broadly sketch an architecture, an implemented computational design provides a fully articulated architecture and a complete set of at least computational mechanisms. This architecture and set of mechanisms provide a richer, more concrete and more decisive theory, as well as both a conceptual and a computational model. Moreover, every design decision taken during an implementation translates into a hypothesis that serves to flesh out the theory. These hypotheses may motivate experiments to test them, thus providing direction for biological research. Conversely, the results of such experiments motivate corresponding modifications of the architecture and mechanisms of the agent or robot and the cycle starts again. These ideas have been discussed more fully elsewhere (Franklin 1997).

IDA also implements William James' theory of voluntary action (Franklin 2000a; James 1890; Kondadadi & Franklin 2001). The experimental work of neuroscientist Benjamin Libet lends support to this implementation of voluntary action as mirroring what happens in humans (Libet 1999; Libet et al. 1983). The IDA implementation offers a new interpretation of Libet's work and suggests possible experimental tests (Franklin & Graesser 2001).

It is also likely that observing the behavior of such software agents or robotic implementations will lead to additional hypotheses that can then be tested. The full IDA prototype is now complete. We've just begun running her continuously so that such observations can be made.

The nature and function of models

Ronald N. Giere

Center for Philosophy of Science, University of Minnesota, Minneapolis, MN 55455. **giere@umn.edu** **http://www.tc.umn.edu/~giere/**

Abstract: There is no *best* scientific model of anything; there are only models more or less good for different purposes. Thus, there is no general answer to the question of whether one should model biological behavior using computer simulations or robots. It all depends on what one wants to learn. This is not a question about models, but about scientific goals.

1. Introduction. Scientists usually do not engage in methodology for its own sake, but because there is a particular issue or problem that requires methodological reflection. Webb's issue is the legitimacy of biorobotics, and, in particular, the use of robots to model biological systems. She focuses on the notion of modeling itself. As is typical in methodological debates, she appeals to authorities well beyond her own scientific specialty, including some in my own specialty, the philosophy of science. My purpose here is not so much to criticize as to sharpen and emphasize some of the many points she has made.

2. Questions about models. There are many questions one can ask about scientific models. There are, for example, *ontological* questions (What *kinds* of things are models?); *functional* questions (What can one *do* with models?); *evaluative* questions (What makes a model a *good* model?); and *epistemic* questions (How can we *know* that a given model is a good model?). Webb considers all of these sorts of questions. I think the ontological and functional questions are primary. Answers here provide a framework for answering other sorts of questions. But one should not expect to find the single best answer for any of these questions. There are different good answers depending on one's purposes. I will assume Webb's purposes.

3. Models as representational objects. I suggest we begin with a universally acknowledged example of a good scientific model, Watson's original physical model of DNA constructed out of tin and cardboard. This model was a physical object. It was not a set of equations, a body of knowledge, or any other sort of purely linguistic or symbolic object, the sorts of things Webb reports others as claiming models to be. One could say that this model *embodies* knowledge, for example, of the chemical composition of DNA, even knowledge expressible in formulas. But the model itself was not composed of such things.

The primary function of this model seems clearly to have been to *represent* the physical structure of a DNA molecule. Discovering the physical structure of DNA was Watson's primary scientific goal. Building such models was central to his method for achieving this goal, and he achieved success with construction of the final model. So here we have one physical object that is being used to represent, or stand in for, another physical object.

Watson's model exhibits several fundamental features of representational models. The latter represent only selected features of their target. Watson's model had neither the size nor chemical composition of DNA. Moreover, his model represented the structure of DNA with less than perfect accuracy. So another fundamental characteristic of representational models is that they represent features only to some approximate degree of accuracy. These two features of models reflect the idea, quoted by Webb in several different forms, that the only perfect model of anything is the thing itself. Thus, paradoxically, merely possessing a perfect model of something provides no increase in our scientific understanding of that thing.

These two features of representational models have the very important consequence that there is no such thing as *the* model of anything. There are many different models representing different features to different degrees of accuracy. Which of these many possible models one wishes to consider depends on the *purposes* for which the model is being constructed. Thus, one cannot eliminate the purposes of scientists from the evaluation of any model.

4. Abstract models. Many of the representational models used in the sciences are not physical objects in the way Watson's original model of DNA was a physical object. They are abstract objects in the way that mathematical objects are abstract. Consider the simple harmonic oscillator treated in all textbooks of classical mechanics. Its motion is characterized by the force function, $F = -k\,x$. Now we know that no actual oscillating system is a simple harmonic oscillator. All real systems exhibit dissipating forces, such as friction. So the simple harmonic oscillator must be an abstract object. One may also say it is an idealized object, an unrealized ideal. Nevertheless, we can use the abstract object, the model, to represent real oscillators even though the fit will never be perfect. Moreover, we cannot identify this abstract object, the model, with the mathematical equations used to describe it simply because the same abstract structure, thus the same model, can be described using many different equations.

5. Computer simulations and robots. A computer simulation is just a fancy way of investigating the mathematical features of an abstract model characterized by a set of equations. The physical computer puts constraints on what sorts of functions can be used in the characterization of the model and how fast they can be solved. But this is no different in principle from the limitations of a person solving differential equations by hand. The main limitation to computer simulation is that literally everything has to be characterized symbolically. If one wants to model an organism in an environment, one has to model the environment as well. On the other hand, computer simulations have the desirable feature that everything about the model is explicitly represented.

Here is an advantage of using robots as models of biological systems. If locomotion is the object of study, for example, no calculations of the physical forces on a limb are required. The relevant forces just operate as they will. And gravity, of course, is the same for a robot as for a real animal. On the other hand, robots have the disadvantage that one does not get an explicit representation of the physical forces.

6. Conclusion. The question whether to use computer simulations or robotic models is really a question about what one wants to learn about particular types of living systems and for what purposes. There is no basis for claiming that one form of modeling is inherently superior to any other.

Can robots without Hebbian plasticity make good models of adaptive behaviour?

Jørn Hokland[a] and Beatrix Vereijken[b]

[a]*Department of Computer and Information Sciences, Faculty of Physics, Informatics and Mathematics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway;* [b]*Human Movement Science Section, Faculty of Social Sciences and Technology Management, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.* **Jorn.Hokland@idi.ntnu.no** **Beatrix.Vereijken@svt.ntnu.no** **www.idi.ntnu.no/~hokland/** **www.svt.ntnu.no/psy/Beatrix.Vereijken/**

Abstract: No. Animals' primary problem is the shaping of movements, guided by and adapting to sensory signals. This requires a narrower class of biorobotic models than that spanned by Webb's dimensions and examples. We claim that all model variables and mechanisms must have real counterparts, input vectors must model known sensor fields, internal state vectors and transformations must model neurophysiological processes, and output vectors must model coordinated muscle signals.

> *The ultimate function of the nervous system is control of behavior. The brain is the organ that moves the muscles. It does many other things, but all of them are secondary to making our bodies move.*
> (Carlson 2001, p. 243)

> *Behaviour in turn, is shaped by learning.*
> (Kandel 2000, p. 1247)

Webb argues that "biological behaviour needs to be studied and modelled in context, that is, in terms of the real problems faced by real animals in real environments." The real problem for nervous systems is to produce sequences of motor output vectors by transforming internal state variables and sensory input vectors, and to adapt this transformation based on experience. The real problem in biorobotic modelling is to describe this transformation and its change mechanisms, and to test this description. Clearly, some models are better suited to solve this problem than others. Within Webb's dimensions describing the variability of models, biorobotic models should therefore occupy a quite narrow 7-dimensional subspace.

Performance match. The primary aim of biorobotics is to seek a match between simulated behaviour and target behaviour. But "should the behaviours be indistinguishable or merely similar?" (target article, sect. 3.6). In our view, indistinguishable is clearly better, and Webb does not seem to give counterarguments. Thus, the benchmark is quantitative matching of movements. Unfortunately, this makes the testing of sensor models less straightforward. Although such targeting may guarantee realistic inputs, that is, close "sensor matches," the prevailing output of low-dimensional action commands, often to wheels, is less suited for behaviour matching than are high-dimensional movement patterns. This implies that biorobotics best validates models describing animal coordination, targeting not *what* action to make, but *how* behaviour is created.

Abstraction, structural accuracy, relevance, and level. In order for biorobotics to help us learn how movements are shaped by vision, hearing, taste, and so on, the simulation should depend on input vectors representing real sensor signals only. A testable target system is the circuit from retina via superior colliculus to eye muscles, since it produces target movements known as saccades. However, when targeting more complex behavioural patterns, the fifth sense comes first: "Sensory information from muscles, joints, and skin . . . is essential for regulating movement. Without this somatosensory input, gross movements tend to be imprecise, while tasks requiring fine coordination in the hands, such as fastening buttons, are impossible" (Kandel et al. 2000, p. 713). Several of Webb's examples applying her criteria are not based on realistic inputs, for instance, the PID control of robot arm joint movements, which is based on simple sinusoidal target trajectories (sect. 2.4).

Webb states that "the performance of similar behaviour is never sufficient to prove the similarity of mechanisms – this is the problem of underdetermination" (sect. 3.6). We trust she refers to *neural network* mechanisms, that is, the model should at least carry the details of a connectionist network with internode weights and intranode activations. Most state variables and mechanisms of a simulation should have uncontroversial physiological counterparts or, exceptionally, one hypothetical mechanism. Should the simulation match the real behaviour, the hypothetical mechanism must then be experimentally tested. In Webb's words, "The main criteria for relevance could be taken to be the ability of the model to *generate testable hypotheses* about the biological system it is drawn from" (sect. 3.1, her emphasis).

The connectionist level being the highest, what is the lowest? Must the Hodgkin-Huxley model of ionic currents be included, or will a simple threshold of weighted inputs do? Is scalar-valued synaptic efficacy sufficient to obtain performance match, leaving irrelevant the details of presynaptic transferase and release, of transmitter esterase and reuptake, and of short- and long-term postsynaptic potentiation? Such questions represent the very essence of neuroscientific modelling, and researchers in biorobotics must make risky choices, "the purpose in modelling is often to *discover* what are the 'relevant features' or 'essential structures'" (Webb, sect. 2.1).

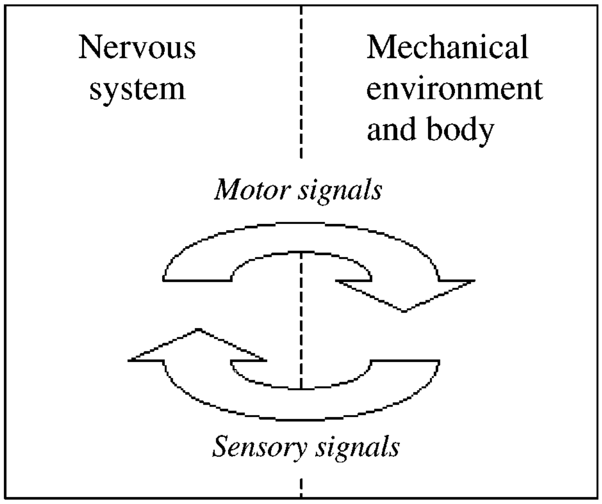

Figure 1 (Hokland & Vereijken). Some necessary variables and mechanisms of adaptive biorobotic models. Left: Hebbian plasticity and integrate-and-fire action. Right: Newtonian mechanics, including muscled limbs. Arrows: sequences of coordinated motor output and sensor (including kinaesthetic) input vectors.

Medium. Although "two-wheeled motor control has to cope with friction, bumps, gravity, and so on" (sect. 4.4), we regard this more as the challenge of wheelchair engineering than of behavioural science. In her definition of biorobotic modelling, Webb rules out purely computer-based models, since they can break or neglect physical laws given that "we include only what we already assume to be relevant" (sect. 4.7). This is true, but equally so for wheeled devices, as the ultimate function of the brain is to move the muscles (Carlson 2001). Thus, a simulated body with legs may be a better biorobotic model than some physical device, assuming that the simulated robot and environment are Newtonian. When modelling creatures on land, space must be three-dimensional, inertia must exist, the body must have limbs with mass and joints with muscle forces, and the environment must have gravity and some friction surface to walk on. Such robots may succeed where physical robots fail, since most neural net simulations will fall behind in transforming real-time somatosensory inputs into proactive muscle commands (see Fig. 1).

Generality. Webb's pessimistic premise, "if different animals function in different ways, then trying to generalise over them won't work" (sect. 3.3), is only half true. Some functions are species-specific, others are not. Locomotor patterns, for example, are remarkably similar across different vertebrates and even insects – normally walking using a tripod gait – spontaneously adopt a typical tetrapod gait when one leg from each side is removed, regardless of what combination of four legs remains (von Buddenbrock 1921). Therefore, a biorobotic model of locomotion could be general.

Related to the concept of generality, but not mentioned by Webb, is the multi-functionality *within* an animal. A real animal is inherently a multi-purpose device that continuously constrains itself to act as a special-purpose device, in interaction with the environment and learning from experience. Synaptic plasticity is suggested as the primary mechanism of development and learning, and "once synaptic contact is established . . . continued development depends on the coordination of neural activity of pre- and postsynaptic neurons" (Kandel et al. 2000, p. 1128). Thus we coin biorobotic models containing synaptic state variables and some Hebbian change mechanism *adaptive.*

Adaptivity. Answering our title question, only such adaptive models, fitting the narrower 7-dimensional subspace we have outlined here, can meet the challenge of Sanes et al. (2000, p. *xiii*): "neuroscientists have come to realize that they are studying a moving target: growth and change are integral to brain function and form the very basis by which we can learn anything about it."

The usefulness property of biorobotic sensorimotor models: A natural source of prosthetic designs

Kristen N. Jaax
School of Medicine, University of Washington, Seattle, WA 98115-2500.
jaax@u.washington.edu
http://rcs.ee.washington.edu/brl/people/kjaax

Abstract: This commentary addresses an additional feature of biorobotic modeling: usefulness in prosthetic design. By implementing structurally and behaviorally accurate models as prosthetics, biologically accurate restoration can be approximated. This has potential to reestablish important peripheral elements of the sensorimotor control system, including limb biomechanics, proprioception, and vision. Examples given include musculoskeletal prosthetics and retinal implants.

The target article does an excellent job proposing a well-defined vocabulary and structure by which the modeling community can analyze the choices made by a given modeler. Hopefully this will also resolve some of the controversies arising from disparate definitions of what a "good" model entails. Webb's discussion of the merits of biorobotic modeling is a powerful treatment of a question often asked of this young field: What is the purpose of building a model in hardware rather than simply using a computer simulation? She clearly demonstrates the unique perspective offered by this new modeling medium and its resultant ability to generate relevant biological hypotheses. In this commentary, I wish to extend this analysis to consider an additional category describing biological models: usefulness. The class of biologically-relevant biorobotic hardware described by Webb, hereafter referred to as "biorobotic hardware," distinguishes itself in this category by offering an ideal source for biomorphic prosthetic designs. For this subset of engineering, it is proposed that the design goals are closely congruous to those of biological modeling. Several cases are presented in which biorobotic hardware is being applied to biomorphic prosthetic design.

Why develop biomorphic prosthetics? The biological faithfulness of biorobotic hardware offers distinct advantages in integration of a prosthesis into the wearer's sensorimotor system. As Webb suggests, "engaging closely in modeling the periphery simplifies central or higher level processing" (target article, sect. 4.2). The behavior of peripheral sensorimotor organs is laden with organism specific functionality. As noted by Kubow and Full (1999), the biomechanics of the limb and its actuators is important in the control of motor tasks such as running. Sensory organs play an important role as well, filtering incoming stimuli and delivering refined signals to the central nervous system (CNS). Since prosthetic communication to the CNS is severely bandwidth limited, accurately restoring biological filtering in prosthetic hardware has the potential to dramatically improve the wearer's perception of sensory stimulus. Finally, using biorobotic hardware which uses structurally accurate mechanisms to generate behavior, minimizes the risk of omitting essential behaviors due to our limited appreciation of the mechanisms of the sensorimotor system.

The goals of biorobotic modeling match well with the goals of biomorphic prosthetic design. They are of course similar in their use of engineering hardware as a modeling medium. Biological relevance is important to ensure that the inevitable abstractions preserve the essential functional characteristics. The need for completeness results in integration of multiple levels with a primary focus on the physical level of sensing and actuation in a real world environment. Biomorphic prosthetics specifically model the sensorimotor organ they are designed to replace, although data availability may require model development based on another species. The goal of "matched filters" (Wehner 1987), however, necessitates eventual transformation of scale to that of the human body. Abstraction occurs through simplification, not idealization, on account of the need to function in the real world. Functional completeness necessarily supersedes full structural accuracy in importance, with the goal of a device which closely approximates the function of the true biological mechanisms. Finally, behavioral match is critical to attain the benefit of biomorphic prosthetic design. One could argue that structural accuracy is irrelevant so long as behavioral match is assured. Given our incomplete understanding of the mechanisms involved, though, striving for structural accuracy minimizes the likelihood of inadvertently omitting a key behavior.

Klute et al. (2000) have applied their biorobotic models to the design of "a biorobotic prosthetic limb engineered to mimic the performance of an amputated limb." The design of their biorobotic below-knee prosthetic incorporates artificial muscles, called Gaylord-McKibbon actuators, and artificial tendons with the intent of capturing some of the benefits of normal leg biomechanics including energy conversion, storage, and return. The group originally used the Gaylord-McKibbon actuators as artificial muscles as part of a biorobotic model to test biological hypotheses regarding human arm movement (Hannaford et al. 1995). The artificial tendons were designed specifically for the prosthesis project to replicate the properties of biological tendons "to provide appropriate connective tissue between an artificial muscle and the skeleton of a biorobotic system" (Klute et al 2000). In developing these models, both engineering specifications and behavioral match were assessed based on performance data from the biological literature. By partially restoring normal limb biomechanics, this prosthesis has the potential to greatly improve amputee mobility.

Prosthetics technology is at the cusp of enabling direct interface to the CNS (Rizzo et al. 2001), which will enable the application of biorobotic sensor models as well to prosthetics applications. A robotic muscle spindle (Jaax & Hannaford 2002) has been developed whose core technologies could be used to provide proprioceptive feedback to a powered prosthesis driven by the amputee's own CNS. The biorobotic model was developed based on performance data from the biological literature with the intent of testing biological hypotheses (Jaax & Hannaford 2002). Such a device could be used to restore the complex mechanical filtering, transduction and encoding transfer functions characteristic of the native organ, communicating muscle length and velocity to the CNS in its native language.

Finally, Boahen (2000) is currently developing a biorobotic VLSI model of the retina with the intent of using it as an ocular implant to restore vision. The retina, with its many layers of neural circuitry, performs extensive manipulations on incoming image prior to passing the data to the visual cortex. Given the inevitable pixel limitations of a retinal prosthesis, it is advantageous to maximize bandwidth usage by supplying preprocessed data to the neural interface. Accordingly, Boahen hopes to "match the retina's coding efficiency by morphing anatomically identified retinal microcircuits into micropower VLSI CMOS circuits." In doing so, he is developing not only a biorobotic model of the retina, but a "retinomorphic" prosthetic which could restore vision to patients with devastating ocular diseases such as *retinitis pigmentosa* (Boahen 2000). These examples demonstrate a vital characteristic of the biorobotic modeling technique: landmark medical applications with the potential to transform the lives of disabled individuals.

Doing versus knowing

Peter R. Killeen

Department of Psychology, Arizona State University, Tempe, AZ 85287-1104.
killeen@asu.edu
www.asu.edu/clas/psych/dresearch/blab.html

Abstract: Aristotle's four causes frame Webb's question. Comprehension requires specification of trigger, function, mechanism, and representation. Robots are real models of *function.* Physical, biological, and epigenetic constraints delimit the hypothesis space for candidate *mechanisms.* Robots constitute a simplified system more susceptible to formal *representation* than the target system. They thus constitute an important tool in a constructivist development of scientific knowledge.

Like science in general, modeling is an enterprise that we do better than we understand. To say this implies that doing is not the same as understanding. It follows that fabricating robots that behave is not the same as understanding behavior.

Modeling modeling. Webb's overview of modeling reveals its protean nature. A leitmotiv in the chorus of characterizations is:

> *A model is a structure put into correspondence with a target structure.*

Then what? Aristotle's four [be]causes frame the picture provided by real models such as robots (Fig. 1), and say what. *Efficient* causes are triggers of subsequent effects. *Formal* causes are syntactic models, or descriptions. *Material* causes are substrates and mechanisms. *Final* causes are functions and purposes (Killeen 2001). Real models provide a tool for learning about each cause of a target system.

Parts 'n parts. Models may be put into correspondence with systems, or with their components (*mechanisms*), or with *their* components, *ad* almost *infinitum.* A clock represented as a pacemaker and register supports a corresponding recurrent stochastic model; the pacemaker may be periodic or stochastic, the register may be analog or digital, each supporting different formal models – the digital register may be a binary counter; the counter may incur errors in incrementing stages, and that may be captured by its own stochastic process, and so on. Formal models direct the search for potential mechanisms.

Depth of representation is not necessary for models. Newton's model of the world mapped the orbits of the planets but gave no clue to machinery that could bridge the distances. This defeat occasioned his resolve to "frame no hypotheses" about gratuitous machinery. Words to the wise. Requiring biological plausibility may tell us something about potential machinery at the next level, and it can improve our sense of what is plausible. But plausibility becomes less plausible the deeper it is pushed. The elephants supporting the universe may rest on other elephants, but as we descend it quickly becomes academic whether the next substrate is elephant or turtle. In many cases the constraint of plausibility merely serves, like the chalk marks on a tennis court, to keep the game interesting.

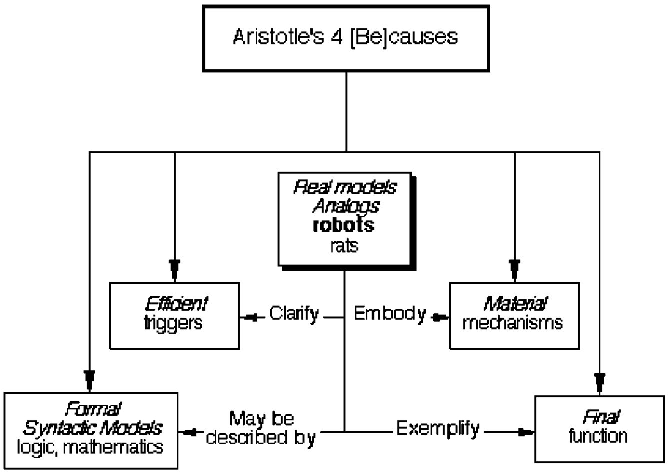

Figure 1. (Killeen). A frame for robots. The success of a robot model depends on correspondence with the function of the target system. It embodies mechanisms that can serve as hypotheses concerning those in the target system. By identifying the inputs/ triggers necessary for the function, it clarifies a central question of psychology – What is the stimulus? As a semblance of the target system it is more susceptible to syntactic models than is the target.

Syntactic models. Models can be real or syntactic. Formal models are stated in a language that has a well-defined syntax: English, logic, geometry. Often they are subsets of mathematics, employing some features and not others (e.g., real but not complex numbers, ordinal but not interval properties). They are proper subsets of the language, else a prediction and its negation would both be part of the model. It follows that computers per se are not models of cognition, even though programs implemented on them may be. Because scientists are familiar with the elements and relations in the modeling language, they can readily manipulate it to test and predict. Einstein held that understanding a phenomenon involves reducing it to a system we already have intuitions about. For most scientists, mathematics is that system.

Real models. Some real models provide only similitude, not function. Cartoons, advertising images, and toy models need only look their part. Such topographical correspondence is of limited utility to science, which prefers robots to androids, function to form. Often an unknown or incompletely understood real model is put into correspondence with the target system. Advances in conquering a disease usually require the discovery of a "preparation" – a model biological system in which the disease can be studied. The most common model in the biological sciences is a rodent. Rats and mice are used as models of humans to predict the effects of drugs or other procedures; as models of another rodent whenever a control group is run. A rat is used as a model of itself in within-subject designs. Real models are often associated with binary outcomes and null hypothesis statistical tests. The investigator merely wants to know whether a procedure is an efficient cause for a target response. The critical question about a real model is: *How well does it scale?* Threshold effects, square-cube laws, and change in the model with uncontrolled variables (e.g., simple passage of time, placebo effects, demand characteristics) are threats to their validity.

Remodeling. Some formal models are well-defined but interact in complex ways as a function of parameter and input. Newton's attempts to predict the orbit of the moon was frustrated by its participation in a three-body problem. His model was valid but not so useful until it could be instantiated in a computer. Models of models, such as computers that sedulously crank the maths, constitute the same formal models in a different medium.

The medium can, however, change the message. No real model can represent most real numbers. Computers truncate, transforming reals and transcendentals into rational approximations. In complex dynamic systems truncation error grows with iteration, eroding the solutions to noise. Successful physical implementations can carry more conviction than syntactic ones because they demonstrate function despite such threats. "Yes, . . . in theory" is a reproach of formal models. This is not the case for real models; but because they set unknowns in correspondence with unknowns, "Yes, . . . but so what?" is their reproach.

What is it like to be a robot? Robots that assemble parts of a car are not faulted for being poor drivers. Robots provide a candidate system that *functions* like a target system in *delimited* ways. This is often nontrivial, but never more than logically sufficient: Theirs is not the solution (mechanism) that other robots, or even the target system, necessarily use. If they provide candidate mechanisms, they are a contribution; otherwise a curiosity. Often robots are a hodge-podge of gizmos not easily described formally. Then the sense of understanding that (often falsely) accompanies a formal model is compromised. That a robot passes a Turing test is to claim no more than that $x = y$, without assigning values or confidence intervals to the variables.

Model husbands. Different people need different types of models: One woman's model husband is another's model bore. Robots underscore the physical constraints and candidate workarounds of target systems. Robots cannot replace syntactic models; but they can instantiate them, as a slide rule does the logarithm, lending credibility. Construction under constraint clarifies the constraints. Nature's prominent coping strategy is *repurposing*, which generates a path-dependency in mechanisms that is seldom on the agenda of a bioroboticist. Why swim so differently than fishes (unless the targets are marine mammals that found their way back), build routes of innervation so circuitous (unless the organ drifted over generations), egg retrieval so awkward (unless gulls can't infer "my wing could do it better"), males so often profligate? Biological plausibility depends not only on the machinery of wetware, but also on its epigenetic provenance. All laboratories should have the query "What is the question?" posted on their wall, and all modelers should answer it before being given the keys to the computer, or to the foundry. The answer is negotiable; but must be stated and addressed if tinkering is to amount to science.

What robots are good for. Comprehension of a target system requires understanding each of the four causes: What makes it go, what it does, how it does it, and how to talk about it. The goal of science is to reduce the hypothesis spaces for those causes. Successful robots are a triumph of function; their doing improves one aspect of our knowing. Constraints – physical, biological, or epigenetic, – reduce the solution spaces for mechanisms. Subsumption architectures provide a complementary demonstration that complex machinery may be unnecessary for the function. Robots enforce a parsimony on all other attempts to model a target system. What robots are specially good for is sharpening Ockham's razor.

Models are better than their theory

Rolf Kötter

C. & O. Vogt Brain Research Institute and Institute of Anatomy II, Heinrich Heine University, D-40225 Düsseldorf, Germany.
rk@hirn.uni-duesseldorf.de http://www.hirn.uni-duesseldorf.de/rk

Abstract: As modelling becomes a popular approach in the study of biological systems it is necessary to clarify its concepts and dimensions. This helps to characterize and to distinguish models but cannot establish their quality. The virtue of a model depends on the insight gained in respect to a specific scientific question, and it is hard to measure this with a theory.

Modelling plays an important role in our conceptualisation of the world and has a long tradition in physical science. In biology, by contrast, the reputation of modelling and models is poor: working hard to unravel highly complex systems, we do not want the facts to be confounded with meta-data, simplistic approximations, and speculations. Thus, if models are necessary – and there is evidence that we rely on them not only for understanding complex systems, – then we need criteria that allow us to objectively assess the quality of models and to progress more surely towards better ones.

In trying to assess the "model muddle," it is useful to clarify the terminology and the general procedure: The process of modelling involves that "we theorise that a system is of type T, and construct an analogous system to T, to see if it behaves like the target system" (Chan & Tidwell 1993). It has to be realised, however, that the term model has been ascribed to different steps in this process (preferably to the theory). On this basis, models can be classified, and there have been many suggestions in the past as to the number and names of their classificatory dimensions. Barbara Webb describes seven dimensions, which admittedly are not completely orthogonal, but are very useful as a reference frame for comparing models more easily and putting them in perspective. Following the dimensions of "biological relevance," "level," and "generality," it is the dimension of "abstraction" which seems most prone to confusion. Being referred to as "detail" in Figure 2, most of Webb's discussion points to "simplification" as the appropriate term, if it were not for the unclear distinction in this dimension between "simplification" and "idealisation" when comparing robots to simulations later in section 4.4. Furthermore, this section lacks reference to the notion of a "minimal model" as the simplest model that still adequately represents a certain hypothetical mechanism. A well-known example is the reduction of Traub's model (Traub et al. 1991) of hippocampal cells by Pinsky and Rinzel (1994). Adding "structural accuracy," "behavioural match," and implementation "medium," the author makes an important contribution by disentangling much of the confusion and by laying out a fairly sophisticated repertoire of distinctions for classifying models in biology, be they robots or computer simulations. Another important dimension appears to be the degree of integration among conceptual levels: While it is commonplace to emphasise the multi-level organisation of biological organisms, most models address processes only at a single level, for example, biochemical, neuronal, population, or functional module. It is easy to see how this limitation arises during the modelling process from favourite ways of thinking and the application of particular mathematical tools and their implementations. Whether these levels are similarly independent in real biological systems is an interesting question – if not, this might lead to radically different model architectures both for computer programmes and robots.

None of this discussion, however, answers the question whether robots can make good models of biological behaviour. The literature is full of unproven claims about what is a good feature – and equally blunt statements to the opposite. For example, the amount of detail to be included in a model ranges from the austerity of the "minimal model" to as much detail as possible, if not the precise copy of the real system. Even from the strongest reiterated argument, namely, that robots can make good models of animals "by working on real problems in real environments," the virtue of biorobotics does not follow. None of the seven model dimensions distinguishes good from bad. What they do is to collectively provide a multi-dimensional classification that helps to identify the characteristics of bio-robotics and to distinguish them from other modelling approaches. In this sense, the title of the target article is misleading: it begs the question. However, this article raises a far more general and important point: Can the quality of models in biology be objectively assessed and thereby be systematically improved? The answer is also indicated: A model is good if it provides insight and helps to gain knowledge, ideally (but certainly not exclusively) by behaving like the target system. Knowledge may be gained independently of whether it is a model *of* (emphasizing simulation technology), *for* (taking theory as a source of explanation), or *on* (using biology as a source of ideas) vis-à-vis animals; at the intersection of biology and robotics all directions are possible, and no single view is exclusively good. Of course, this is also true for the field of computational neuroscience at the intersection of neuroscience, computer science, and other disciplines: the technology sense, the source sense, and the target sense are all valid views but none of them is exclusively valid or a sufficient definition (Kötter 1999). Insights are also observer-dependent: designers, vendors, and users have very different requirements for models of the same system.

The virtue of a model has to be judged empirically in the context of the question that the model is designed to answer for the investigator. Thus, we need good models that address good scientific questions. Developing good models seems more advanced than the development of a general theory of what makes a model a good model.

ACKNOWLEDGMENT
The author is a member of the EU Thematic Network "Computational Neuroscience and Neuroinformatics."

There is more to biological behavior than causation and control

Mark A. Krause
Department of Psychology, University of Texas at Austin, Austin, TX 78712.
krause@psy.utexas

Abstract: Biorobots may model the causation of relatively simple behaviors, but many animal behaviorists are concerned with complex cognitive traits and their evolution. Biorobotics seems limited in its ability to model cognition and to provide evolutionary explanations. Also, if robots could model complex traits, such as theory of mind, underdetermination could be problematic. Underdetermination is also a challenge for comparative psychologists.

The causes and evolution of behavior are of concern to most animal behaviorists. Based on Webb's target article it appears that many animal behaviorists use biorobots to model the behavior of animals, and to generate testable hypotheses. The issues I raise about biorobotics apply mainly to the field of comparative psychology and the evolution of complex cognitive traits. In comparative psychology, if two or more related species perform the same or very similar behavior, an inference about the evolution of that behavior can often be made. If a biorobot were made to model that same behavior, then perhaps the biorobot could be used to make predictions of the behavior of several species. In addition to studying (and comparing) the behavior of animals, comparative psychologists are interested in determining the underlying causes of behavior. However, under certain circumstances it can be erroneous to assume an identical cause for a behavior exhibited by two or more species. Furthermore, identifying the cause of behavior influences the evolutionary interpretation of the behavior. Webb acknowledges the controversy surrounding reasoning by analogy (underdetermination). However, she does not elaborate much on this important concern.

Can robots make good models *of* animals? This is a focal question of Webb's article. A complete explanation *of* animal behavior requires an evolutionary account. Cladistics is a common method used to delineate the evolutionary history of behavior. In this case, real animals make the best "models" of other real animals. However, even such comparisons among living species sometimes fall short of providing adequate explanations of the causes of behavior. For example, there is ongoing debate over whether nonhuman species possess higher order cognitive abilities such as a "theory of mind." Nonhuman species may behave as if they reason about the mental states of others, but arguments by analogy based on comparisons among living species (or biorobots) may result in a faulty evolutionary interpretation.

If a robot were designed to follow the eye gaze direction of another species, and appeared to understand the relationship between seeing and knowing, then it could be inferred that the robot models how a living species reasons about the mental states of others. Chimpanzees are sensitive to the eye gaze direction of others, and in captivity they are known to use pointing in conjunction with eye gaze alternation between a human and a referent (Krause 1997; Leavens et al. 1996). Thus, at a behavioral level, it appears that chimpanzees infer something about the mental states of others, as pointing and the establishment and manipulation of joint visual attention indicate the presence of a "theory of mind." However, the evidence that chimpanzees have a theory of mind is incomplete and highly ambiguous (Heyes 1998). It may be that chimpanzees in fact do not possess a theory of mind, despite the fact that their pointing and eye gaze behavior often appear so similar to that of humans. Based on some critical experiments testing how chimpanzees use (or do not use) visual attention during communication, it appears that they in fact do not possess a theory of mind (Povinelli et al. 2000). Now consider what would happen if a robot was designed to follow the eye gaze direction of another species, and if, based on the robot's behavior, it appeared to understand the relationship between seeing and knowing. Among other conclusions, the robot's behavior would seem to indicate an understanding of the mental states of others. In this instance, using a robot to model the pointing and joint attention behavior of chimpanzees or humans would not resolve the controversy if the robot simply mimicked the behavior of either species. In many cases, it seems unlikely that a robot could tell us anything we do not already know or could not answer with experiments on live animals.

As another example comparing remotely related species, human children dress up their dolls and talk to things they know will not answer; they "animate" their own worlds with creative imagination. In a similar fashion, domestic cats are known for their propensity to treat inanimate objects as though they were prey. Common sense suggests that pretense underlies both examples, but pretense is probably an over-interpretation of the cat's behavior. More than likely, the human child and the cat have come upon similar behavioral solutions to different problems. A robot constructed to "animate" its world probably would not explain the evolutionary differences that underlie human and nonhuman behavior.

In comparative psychology and ethology, misinterpreting a behavior as homologous is avoided by examining phylogeny, carefully testing adaptation, and by closely exploring the underlying mechanisms controlling a behavior. When these steps are taken in the comparison of species, the interpretation of the causal mechanisms of behavior may differ. Arguments by analogy have the allure of common sense, but they can lead to the misinterpretation of data. This is an issue for scientists studying animal behavior, and therefore one for people working in biorobotics. The field of biorobotics has considerable promise and I hope that it will continue to flourish and generate hypotheses. However, Webb should further address this "argument by analogy" problem with regard to comparisons made between the behavior of robots and organisms. Would a biorobot programmed to "animate" its world be more like a cat or a child? Would it have a theory of mind or not? Can biorobots really answer questions about behavioral evolution? Furthermore, of what use are biorobots to comparative psychologists studying complex cognitive traits?

Like the perfect animal, there's no such thing as the perfect institution

Susanne Lohmann
Department of Political Science, University of California, Los Angeles, Los Angeles, CA 90077. **lohmann@ucla.edu**
www.polisci.ucla.edu/faculty/lohmann/

Abstract: Models of biological and political systems share in common an irreconcilable tension. They must extract the essential features of the system to make its workings comprehensible to the human observer, and yet the omitted underbrush is essential to the workings of the system. A good model accommodates both the workings of the system and the cognitive makeup of the observer.

Why build a robot to model an animal? At first blush, the idea of developing a physical robot model to represent a biological sensorimotor system seems bizarre. How can a clunky mechanical robot possibly tell us something about a living breathing animal? And if we need a simplified account of what makes an animal tick, why not a mathematical model on paper, or a computer simulation – why the concreteness and messiness of an engineering contraption?

To explore the "why a robot" question, I draw an analogy between the study of biological and political systems.[1] Examples of biological systems are ant colonies and bat sonar. Examples of political systems are the United States Congress and the first-past-the-post system used for elections in the United Kingdom.[2]

Political and biological systems are alike in several respects.

First, there is a biological component to the workings of political institutions because they are "inhabited" by human beings. Thus, the functioning of a political institution depends at least in part on how it interacts with the human body and brain. Trivially, an institution that does not respect the human need for sleep is not going to last for long. Less trivially, well-functioning institutions respect the quirks of human cognition, emotion, and moral reasoning. They rely on formal structures (simple majority rule, federalism, written constitutions) and informal modes of interaction (culture, ideology, morale, and leadership). They take into account limitations on the human ability to pay attention (what do people naturally attend to? what is the human attention span?). They pander to the human desire for narrative (people tend to work hard for institutions that have a vision and give them a sense of purpose). Great institutions make productive use of the human propensity for factionalism (people tend to form clusters, and they cooperate within and compete across clusters; people also seem to have a close-to-infinite capacity to form clusters-within-clusters).

Second, political institutions are complex systems with emergent properties. Just like the brain possesses consciousness and yet one cannot find the seat of consciousness in it, a political institution can contribute to economic stability and growth even though its internal world consists of a Kafkaesque bureaucracy obsessed with partisan squabbles and prone to petty corruption. We have theories connecting the internal world of an institution to its aggregate performance. Even so, there is something magical about well-performing institutions, which is to say that our theories are thin simplifications.

Third, political institutions evolve in conflict with their environment. Some institutions emerge bottom up as a result of the decentralized actions of individuals. In a process comparable to annealing, people's interactions become increasingly structured until a complex institution crystallizes out of the mess.[3] Other institutions are designed from scratch by political actors, but they evolve over time as they accommodate the political pressures of disaffected individuals and groups or change their structure to better cope with environmental pressures – or they break down and are replaced by other institutions that are better able to stand the political heat.

So what characterizes a good model of political institutions? For starters, we need to understand why we need a model in the first place. Why can't we just "see and understand" an institution the way it really works? After all, the human brain is capable of seeing and understanding very complex phenomena. Indeed, in many cases (recognizing faces, making sense of gossip) we only realized how complex these phenomena are after our man-made computers choked over analyzing them.

This is where evolutionary psychology comes in. The social nature of the human brain was shaped in a hunter-gatherer environment, roughly 30,000 to 300,000 years ago. People lived together in small groups and spent a lot of time gossiping with and about each other. It is not all that surprising that human beings are good at extracting subtle cues from facial expressions or reading other minds ("if I say this she will think he left unsaid that . . . "). Fortunately or not, there was nothing even remotely resembling the United States Congress floating around in the hunter-gatherer environment. As a result, our brain is poorly equipped to reason about the workings of the United States Congress (or bat sonar for that matter).

The purpose of a model is to take a complex system that our brain cannot comprehend, boil it down to its essential features, and thereby make it transparent. This is where we run into a snag. The underbrush we omit as non-essential is in fact essential to the workings of the system. The thin simplification that is our model would drop down dead if it were forced to confront the environment in which the true complex system survives and thrives. What a good model does is to reconcile as best as it can an irreconcilable tension – it accommodates both the workings of the system (it provides a decent approximation as measured by some yardstick of usefulness) and the cognitive makeup of the human observer (it is illuminating to us).

When it comes to the cognitive makeup of the human observer, there are human universals and then there are cultural and individual-specific differences. Cultural differences include differences in "seeing and believing" across scientific disciplines and subfields – indeed, the process of becoming an expert through graduate training and mentorship can be thought of as a process of getting brainwashed into slicing up the world in a certain way and being blind (and, interestingly, hostile) to other ways of slicing up the world. As a result, different disciplines and subfields come up with different models of complex phenomena.

Indeed, the lack of disagreement carries over to meta-level debate about the purpose and workings of the scientific enterprise. When scientists self-reflect, they come up with competing and partially contradictory models of "what is a model" and "what makes for a good model." And this is good so. If scientific progress occurs, it is because the scientific enterprise is in sync with the cognitive makeup of its human inhabitants – their diversity and their propensity for forming factions and factions-within-factions.

NOTES

1. The ideas developed in this commentary are drawn from the author's published work (e.g., Lohmann 2000).

2. In a given district, the candidate who gets the most votes becomes Member of Parliament regardless of whether he or she is supported by more than 50% of the voters. In the House of Commons, the party with the most seats forms the government regardless of whether it has won a majority of seats.

3. The concept of annealing comes out of the thermodynamics of how liquids freeze and crystallize or metals cool and anneal. At high temperatures, the molecules of a liquid metal move around freely. If the liquid is cooled slowly, the atoms line up to form a pure crystal that is completely regular and represents the state of minimum energy for the system. If the liquid is cooled quickly, it does not reach the minimum energy state but instead gets stuck in a higher energy state. A process that allows the molecules to move around and then gradually quenches them has certain optimality properties.

How building physical models can reduce and guide the abstraction of nature

Malcolm A. MacIver

Division of Engineering and Applied Science & Computation and Neural Systems Program, 104–44, California Institute of Technology, Pasadena, CA 91125. **maciver@robotics.caltech.edu**
http://robotics.caltech.edu/~maciver

Abstract: Animals detect and acquire resources through a sequence of shape changes. This process is tightly coupled to the sensory and mechanical ecology of the animal. Building physical models allow us to prescind from modeling these aspects of the environment, which may not yet be described or suitably abstracted. The significance of this hybrid of physical modeling and experimentation to the acquisition of scientific knowledge is discussed.

The sparsity of resources within a mobile animal's domain compels a certain logic, one that all energy-consuming autonomous agents must follow. Resources must be detected, and behavioral programs engaged that terminate in the acquisition of those resources. In carrying out this imperative of their continued existence, animals exert changes in their geometric configuration in space (in brief, shape) to several ends. These include: (1) for sensing: shape changing to enhance the quality of information from its sensor arrays (e.g., bats manipulating their pinnae position during echolocation; dogs bringing their snouts to the substrate to follow a trail); (2) for locomotion: shape changing to undergo net movement, typically toward the detected resource (e.g., a fish bending its body to swim forward; a biped extending a leg to walk forward);

3) for physical coupling: shape changing to eat or grasp the resource (e.g., depression of the lower jaw of fish to create negative buccal pressure for prey capture; bats flipping the tail membrane up to bring an insect to their mouth).

Animals exhibit an astonishing sophistication in their manipulation of the mechanical properties of their world to achieve these ends. For example, shape changing for sensing in electric fish can be seen in rolling behavior following prey detection (MacIver et al. 2001). This behavior centers the fish's top edge – a region of high sensor density – under the prey; allows the fish to approach the prey by slicing its narrowest cross-section through the water, thereby minimizing added-mass effects; and may provide a simple control strategy for reaching the prey by balancing the stimulus between the two sides of its body and ascending the gradient of sensory signal strength (MacIver et al. 2001; Nelson & MacIver 1999). As described further below, investigations of shape change for locomotion in insects and fish are demonstrating that these animals utilize phenomena within fluids quite beyond those that we utilize in our flying and underwater machines, phenomena that we are still discovering, to say nothing of having a good analytical approach toward.

Shape changing for resource detection and acquisition is clearly fundamental to the sensorimotor intelligence of animals that we so desire to understand. As the examples indicate, these shape changes are tightly coupled to the sensory and mechanical ecology of the animal. Yet, modeling the environment, which animals have demonstrated an unerring capacity to exploit in ways we are hardly aware of, let alone capable of simulating accurately, presents a high obstacle to the integrative computer simulations that are currently our best shot at understanding these coupled sensorimotor processes.

As Webb and others have pointed out (target article, sects. 3.7 & 4.7; Beckers et al. 1996; Flynn & Brooks 1989; Quinn & Espenschied 1993), a great advantage of building physical models is that this allows us to prescind from modeling the undiscovered or unabstracted aspects of the environment on which the target behavior depends. Although Webb's article is very helpful in clarifying the maze of issues surrounding the building of physical models, I believe that this key point is one which merits further elaboration. In what follows, I place the building of physical models in the broader context of the acquisition of scientific knowledge, inquire into the nature of their contribution to this process, and briefly describe some recent examples.

Understanding involves abstraction. These abstractions are expressed in some language for communication and verification. Mathematics provides one such language, but what follows applies to abstractions expressed in any language. Suppose we express our abstractions of some biological phenomenon using the language of mathematics. The next logical step is to calculate predictions from these expressions in order to test their fidelity to the phenomenon (in the case of a spoken language, we would use practices of informal logic to derive verbal predictions). The expressions may need to be approximated to make them computable in finite memory machines in finite time. The calculated predictions are compared to empirically obtained observations, and an interwoven process of theory adjustment, algorithm development, and experimental work ensues. Where can building physical models contribute?

The tragedy of abstraction is that it requires the loss of information. Otherwise, we haven't abstracted. In the process of generating predictions from abstractions, there will be some predictions that will therefore not be computed; namely, those that rely upon the information excluded from our abstractions, or lost in the approximations of those abstractions required by computational expediency. I will use the phrase "abstraction load" (in analogy to "cognitive load") to refer to the work needed to obtain the abstractions and computational methods that will generate the observations we have failed to compute.

Building physical models has the advantage of reducing the potentially insurmountable abstraction load associated with computing all the aspects of the environment on which the target phenomenon depends (where "environment" refers to any aspect external to the phenomenon we are trying to abstract). To simulate the phenomenon adequately, this work would have to be done; but building the object and letting reality supply the physics obviates the need to do some of this work. The crucial point is, we haven't thereby given up the game completely – we are neither pinned into the muck and goo of pure experimentation, nor caged in the assumption-permeated world of pure simulation, but find ourselves at some interesting halfway point.

For example, following similar work by McGeer, Ruina and colleagues developed computational models of a "passive walker" – a walker that has a human-like bipedal gait down inclined planes without actuation or control. The simulations predicted that the walker would not be stable, but it was built in order to test some other issues. To their surprise, the model did walk (Coleman & Ruina 1998). The functioning of the physical model directed the development of a simple quantitative model to explain its stability (Coleman et al. 2001). Similarly, in recent work on fish swimming and insect flying, a number of fluid phenomena have either been discovered or made more observable as a result of the use of robotic devices that approximate the movement of these animals (Ahlborn et al. 1997; Bandyopadhyay et al. 2000; Barrett et al. 1999; Birch & Dickinson 2001; Dickinson et al. 1999).

In allowing the full complexity of the environment to work on what could be called "reduced robotic preparations," this research is cracking open the black box of complex deformable-body and fluid dynamics phenomena to new theoretical advances. The epistemic accessibility afforded by building these robotic devices is analogous to that obtained by traditional instruments such as the microscope and telescope. The building of physical models not only reduces abstraction load, but in illuminating that part of nature we most urgently need to abstract in order to account for a phenomenon, it provides a saliency filter for the immense richness of opportunities for abstraction effort that arise at every turn in the course of experimental work.

When robots fail: The complex processes of learning and development

Ludovic Marin[a] and Olivier Oullier[b]

[a]*Sport, Performance and Health Laboratory, University of Montpellier 1, F-34000 Montpellier, France;* [b]*Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL 33431.*
l.marin@staps.univ-montp1.fr **oullier@walt.ccs.fau.edu**
http://oullier.free.fr

Abstract: Although robots can contribute to the understanding of biological behavior, they fail to model the processes by which humans cope with their environment. Both development and learning are characterized by complex relationships that require constant modification. Given present technology, robots can only model behaviors in specific situations and during discrete stages. Robots cannot master the complex relationships that are the hallmark of human behavior.

In her article, Webb offers a convincing argument for instances in which robots can be good models for understanding biological behavior. In her account, mechanical models, like those used in research on other animals, can similarly help researchers gain insight into human behavior. Robots offer experimental advantages in certain situations because robots can be programmed, and even demonstrate very simple learning strategies within a given environment and context. For example, robots can be used as "stand-ins" for humans in experimental situations that are too dangerous for live subjects (e.g., removing the primer of a rocket). Using robots eliminates emotions like fear or anxiety that affect experimental outcome. This said, however, we are not convinced that present technology allows mechanical models of humans to produce one of the most fundamental hallmarks of behavior – adaptation to changes and variations in environmental constraints.

Behavior is affected by both developmental evolution (biological and physiological changes) and learning, or discovering new ways to cope with novel situations in the environment. These variables must be perceived and processed together in order for an appropriate behavior/response to result. As we will attempt to illustrate, the on-line nature of these interactions and the vast amount of variability and complexity within changes of a physiological or an environmental nature are impossible to capture with a mathematical model.

There is no better illustration of variability in behavior than that of babies learning to perform new skills for the first time. Reaching, sitting, crawling, grasping, walking, and throwing are just a few of the plethora of skills babies master in their first few years of life. One major characteristic of these early developmental milestones is that they are manifested in a nonlinear process. We've all seen new parents boast that their baby has taken steps and can walk, even though baby still prefers crawling as a means of getting from point A to B. After the very first steps that a baby takes, it is often a few weeks or even months before the baby is actually described to others as a "walker." During this time, the baby may take ten steps on one day, zero steps on the following two days, and five steps a week later. Motivation for locomotion, as well as the baby's physical ability to put one foot in front of the other in a given moment both factor into whether the new walker will actually decide to walk rather than crawl, scoot, roll, or slide. Given all this variability, we are not convinced that a robot can take into account the process by which physiological change and contraints in the environment relate to and directly affect each other.

Learning a new skill requires that physiological and biological properties are modified with respect to the constraints of the environment and to the task. Although it is possible, as Webb herself points out, for robots to learn through conditioning, we argue that human learning is fortunately much more complex than simple conditioning. Newell (1986) showed that learning all new skills is based on the interactions among the intrinsic properties of the learner him/herself (morphology, muscle), the environment, and the task constraints. For example, when a gymnast learns a new and complex tumbling skill, he/she must first resolve the relationship between individual and environmental constraints. Furthermore, gymnasts must learn how to constantly modify this relationship with respect to psychological demands like fear, fatigue, and motivation. The interaction of these multiple constraints is very difficult to reproduce in a model because all interactions must first be *identified.* If interaction is not the sum of its parts but, in fact, an original entity (Kelso 1995; Kofka 1935), how is it possible to program and model an interaction without defining, or at the very least identifying each constraint? We maintain that the various elements that contribute to the learning process cannot be gathered in one general model that captures every variation. It is only possible to model the behavior of one given interaction in a specific situation at a given moment with given constraints. As is the case in development, mechanical models are more likely the reflection of one very specific instance (a "snapshot" of a special stage in learning) rather than a model of the learning *processes.*

We maintain that existing experimental methods remain the best way to truly address the questions of human development in the contexts of learning and development. We think that it will remain so, as long as robots are created as computer analogy. Any model for biological behavior must take into account the various interactions that are continually present in learning and development. Some interesting studies have already moved in this direction (Schöner et al. 1995). The new generation of robots, the so-called "animats," are autonomous systems exhibiting self-organizing properties (Kodjabachian & Meyer 1995; Meyer & Wilson 1991). Based on neural network modeling, these robots are nonlinear systems that can learn from their environment and lead to "the emergence of interesting and ecologically valid behaviors" (Damper et al. 2000). We expect that future generations of robots will make further progress in accurately representing the processes involved in nonlinear long term change.

ACKNOWLEDGMENTS
We are very grateful to Idell Weise for many constructive criticisms on this commentary, and thank Jean-Philippe Herbien and Alf Capastaga for helpful discussions.

Embodiment and complex systems

Giorgio Metta and Giulio Sandini
LIRA-Lab - DIST, University of Genova, 16145 Genova, Italy.
pasa@lira.dist.unige.it **giulio@dist.unige.it**
http://pasa.lira.dist.unige.it/ **http://giulio.lira.dist.unige.it/**

Abstract: In agreement with the target article, we would like to point out a few aspects related to embodiment which further support the position of biorobotics. We argue that, especially when complex systems are considered, modeling through a physical implementation can provide hints to comprehend the whole picture behind the specific set of experimental data.

Beside the many examples described in the target article, we argue that one of the possible uses of biorobotics is for the study of complex systems with the aim of elucidating general principles rather than perfectly simulating them in every detail. We agree that the most effective research to date has been devoted to the analysis and simulation of specific subsystems (e.g., chemotaxis in *C. Elegans,* locomotion in insects, simple visual motion detectors, etc.), and consequently it addresses rather specific biological questions. On the other hand, the real potentiality lies mostly in the possibility of further analyzing complexity in general terms by, for instance, devising experiments that could not be carried out on real biological systems for various reasons. One possibility we pursued is that of investigating *development* of sensori-motor coordination and cognition: in particular, the first year of life (Metta 2000; Sandini et al. 1997).

The main hypothesis we put forward is that development can be regarded as controlling the complexity of the learner (Metta 2000). We proposed development in contrast to the classical *modular approach,* not only as a source of inspiration, but rather, as a possible design alternative. The criticism we have against the modular approach, especially in engineering, is that, very often, to make the problems tractable, complex systems are divided in small parts, which are then analyzed in isolation. Complexity is addressed by breaking the system into components. This has been successful so far but it has also hit its own limits (Brooks 2001). Most of the time large-scale system integration has either failed or has been successful only at the expense of generality and adaptation.

A different approach is taken by biological systems. Newborns, for example, are already, at birth, an *integrated* system. Many "modules" are still non-functional or they function differently from their "adult" counterpart: neural growth is not completed (Leary 1992), motor control limited (Konczak et al. 1995), but the sensorial, motor, and cognitive abilities are nicely matched. Submodules develop harmonically, resulting in a system whose components always fit one to another during growth. Adaptation is inherent in the very fabric of the system: we can observe the general tendency of a smooth shift from *simpler to more complicated.* Limitations, such as poor sensory resolution, are thought to be an advantage rather than a drawback (Turkewitz & Kenny 1982). Newborns are maximally efficient in collecting data (making new experiences) and their behaviors are optimized for learning rather than simply for efficiency in absolute terms; even "negative" aspects such as noise contribute instead to the long-term performance by supporting exploratory behaviors. Examples of such simple early behaviors are motor reflexes that, although clearly not that much adaptive, provide the simplest controller yet are able to generate useful sensori-motor synergies.

A more fundamental difference – and concurrently, a powerful constraint posed on the developing agent – is that the training data does not come for free: gathering information always has a cost. Getting representative data of the complete learning problem in a vast state space (e.g., imagine a body with many degrees of freedom and a diverse sensory system) can take a very long time – the learner has to visit and try out too many different configurations in this space. The agent cannot devote all its efforts to exploring the environment because otherwise no task would be ever achieved. This issue has been called the *exploration-exploitation dilemma* (Sutton & Barto 1998). Any *real* learning system has to face this problem and adopt strategies to cope with it.

It is clear that casting biorobotics in this "ecological" context shifts the learning problem towards that of collecting the training data. This is to say that we have to deal with action in the first place, and the very capacity to learn involves finding representative data without incurring severe penalties in terms of basic drives (e.g., feeding, mating, or artificial versions of them). Further, the way exploration is performed – the quality of training – depends strongly on how the system acts. This is also why perception is doomed to be derived from actions. The capacity of categorization emerges out of the sensori-motor coordination patterns (Pfeifer & Scheier 1998).

On a more practical basis, this suggests that action is at the foundation of more cognitive functions, such as categorization. Developmentally, if action has to be a prerequisite of perception, we should observe a different developmental progression of the motor control versus the perceptual abilities. This question has been investigated, for example, by Kovacs (2000) who provided supporting evidence in this direction. Along the same line, and, in our view, supporting this stricter integration of motor and perceptual abilities are the studies of Milner and Goodale (1995) and Jeannerod (1997). The more recent discovery of mirror neurons is (Fadiga et al. 2000; Rizzolatti 1996) believed to provide another piece of the link between action and perception. In short, it seems plausible that motor information is needed to perform visual recognition of observed actions. In other words, our ability to recognize an action is based on the activation of the same neurons used to actually perform it. For this reason a certain similarity in terms of motor abilities and kinematics between the observer and the experimenter is required.

In robotics, theories where action had a sort of primacy have been already proposed in the past; for example, the active vision or purposive vision paradigms (Aloimonos 1990; Rajcsy 1985) but, although we generally agree with them, action was eventually exploited a little. It was never a fundamental component in the sense we are proposing here.

All these considerations point toward embodiment and adaptation as fundamental characteristics of intelligent living systems. Firstly, action allows the agent (either biological or artificial) to gather "learning information" efficiently. Secondly, this very capacity of acting can indeed facilitate more fundamental cognitive tasks such as categorization in cases where an external teacher (as in supervised learning) is not necessarily available. Finally, as interpretation of mirror neurons seems to suggest, a body – and possibly a body with a certain kinematics – is needed to recognize actions and communicate properly with others.

In summary, we suggest that if the aim is to investigate complex systems (as the brain certainly is), biorobotics might provide a complementary view on some aspects and, in particular, realistically take into account the agent-environment interaction. It can be regarded as another tool, similar to what tools like Matlab and systems theory are for the simulation of biological control systems.

ACKNOWLEDGMENTS
The research described in this commentary has been supported by the EU Project COGVIS (IST-2000-29375) and MIRROR (IST-2000-28159), and by the Italian Space Agency (ASI). Georgio Metta, a postdoctoral associate at MIT Artificial Intelligence Lab, has been supported with a grant by DARPA as part of the "Natural Tasking of Robots Based on Human Interaction Cues" project under contract number DABT 63-00-C-10102.

Robots aren't the only physical models

Peter E. Midford

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721. **pmidford@u.arizona.edu**
http://www.mesquiteproject.org/midford/

Abstract: Webb compares robots to other behavior models in a novel space, but neglects significant portions of this space. The article's analysis of models would have been strengthened by including a broader coverage of physically embodied models and of models of social behavior. Nevertheless, Webb is correct to claim that robot construction will contribute to understanding behavior.

The title of the target article suggested that the article would discuss and ultimately defend robots as models of biological behavior. This enterprise ought to be relatively uncontroversial. The abundance of examples, both cited in the text and listed in the table, speak for themselves. The article also presents a seven-dimensional space for describing models. However, several types of physical models, relevant to the study of behavior, are not discussed in connection with this space. The omission of these models is significant, as they point to difficulties in Webb's definitions of robots and of models.

There is much to agree with in this article. Webb usefully distinguishes what she calls biorobots from other types of biologically motivated robot research. She has elsewhere (Webb 2000) made a more general case for the use of robots in the study of behavior, so I read this article expecting a comparison of robots to other approaches to modeling behavior. Although the article presented many examples of robotic models and comparisons with other information processing models, there was very little discussion of other physical models.

The absence of any mention of Michelsen et al.'s (1992) mechanical dancing bee was particularly surprising as it stands as one of the best known examples of using robots to study social behavior. If Webb felt that this device was something less than a robot, it would have made an excellent counterexample to buttress the informal notion of "robotic" used in section 2.4. Webb never makes clear what makes a mechanical device robotic.

In addition to the bee, two other models serve to illustrate the point. The first is a die-cast replica of a lizard such as has been used in studies of heat absorption and loss in physiological ecology. As sunlight and the surrounding air interact with the surface of the model, its internal temperature does change in response to its environment. Although these changes are not "behavior" in and of themselves, they represent events that elicit thermoregulatory behaviors. The second example is a mount of a predator, such as the hawk used by Lima and Bednekoff (1999) in their studies of antipredatory vigilance in juncos. The hawk does not directly respond to the behavior of the juncos, it simply slides down a monofilament cable towards an individual focal junco.

These three models (bee, lizard, and hawk) share a physical medium. By virtue of their physical implementation, they all tell us something about the interaction of environment and behavior, exactly as the best robot models do. Each is relevant to at least one particular system, though the lizard and hawk models also generalize to other terrestrial animals and aerial predators respectively. In the case of the bee and hawk, the subject animals, rather than the experimenter, decide the quality of match. Together, these examples suggest that robotics represents an evolutionary advance in the study of behavior as usual, rather than a fundamentally new approach. The greater dissimilarity seems to be between these models and the computational/mathematical models Webb frequently mentions.

Webb rightly observed that robots, like other physical models, can tell us something new about their environment. In doing so, they cease to be pure models of their "source" and also serve as instruments for detecting and changing properties of that environment. Although Webb hints at this dual role for robots, its consequences in relation to her modeling framework were never examined.

The continued evolution of robot hardware promises a continuing stream of interesting robot models across the range of animal taxa. Having both studied behavior and coded robot software in the past, articles like this motivate me to find out more about recent advances in appropriate hardware. Although the introduction of robots represents merely an evolutionary advance in existing methods for studying behavior, contact with and interest from the robotics community should provide a long overdue rejuvenation for classical ethology.

Finally, robots as models in specific systems suggest a role for robots in comparative studies of mechanism. If two target species are phylogenetically close, a robot implementing the behavior of one ought to be easy to convert to a model of the other. Violations of this expectation suggest that either something is wrong with the model or something special happened in the evolutionary history of one of the species. We currently lack good models for evolutionary change in behavior. Comparing model implementations may provide insight in cases where comparing descriptions or measurements of behavior do not do so.

ACKNOWLEDGMENT
This material is based on work supported by the National Science Foundation under a grant awarded in 2000; Grant No. DBI-0074524.

Differentiating robotic behavior and artificial intelligence from animal behavior and biological intelligence: Testing structural accuracy

Ralph R. Miller and Francisco Arcediano
Department of Psychology, State University of New York at Binghamton, Binghamton, NY 13902-6000. **rmiller@binghamton.edu**
arcedian@binghamton.edu

Abstract: We emphasize the feature of Webb's presentation that bears most directly on contemporary research with real animals. Many neuroscience modelers erroneously conclude that a model that performs like an animal must have achieved this goal through processes analogous with those used by the animal. A simulation failure justifies rejecting a model, but success does not justify acceptance. However, an important benefit of models, successful or otherwise, is to stimulate new research.

Webb provides an illuminating and commendably comprehensive analysis of the nature and role of models in the study of behavior, and for this we are grateful. However, unlike Webb, we are not meta-modelers; we are part of the infantry that uses models to do experimental science with live animals. Here we focus on what Webb calls "structural accuracy" (Zeigler's [1976] "structural validity"), that is, how well a model (including one implemented in a robot) represents the actual mechanisms underlying the target behavior. Toward assessing structural validity, we note that comparison of behavioral data to predictions of a behavioral model can provide a basis for rejection of the model. However, a relatively successful prediction of the model does not constitute grounds for accepting the model. Real world behavior (data) can be described by any of an infinite set of equations that predict behavior (dependent variable) on the basis of assorted environmental and genetic factors (independent variables). However good the fit to the data, these equations are no more than descriptive unless the various terms have psychological meaning. For example, the terms of long Fourier series or exponential series ordinarily lack psychological meaning. Such series constitute no more than curve fitting, which is not to say that all curve fitting is undesirable. Initial curve fitting sometimes leads to psychologically meaningful processes being attached to the terms and parameters of the equation fit to the data. Moreover, curve fitting can sometimes have useful predictive value even if it lacks explanatory value. But curve fitting, intentional or serendipitous, rarely illuminates underlying psychological or neurophysiological processes. Restated as a simple example, consider a crate with four rows of three bottles. The number of bottles can be ascertained by simple counting, multiplying three bottles by four rows, or rotating the crate 90 degrees and multiplying four bottles by three rows. We could even add the total numbers of bottle tops and bottoms and then divide by two. All of these approaches correctly yield twelve bottles, but provide no insight concerning how a specific evolved, experienced, real person would actually determine the number of bottles.

We must be careful with the once widely held assumption that "the study of autonomous robots was analogous to the study of animal behaviour" (Dean 1998, as cited by Webb). At the level of investigation, they are in some ways analogous. But an analogy between them does not mean a correspondence between them in underlying process. Autonomous robots are relevant for researchers of natural sciences if they implement a theory based on the study of natural-living organisms. Engineers can build an infinite array of autonomous robots with very few analogies with the processes within natural organisms. Webb is aware of and briefly speaks to this issue. Although a functional robot is hardly evidence of a successful model, the problems in implementing a model through a robot surely provide an excellent means of detecting incompleteness within a model.

The ways in which our minds currently work are constrained not only by our experiential histories, but by our evolutionary histories. Any model of behavior that fails to incorporate these constraints will never be successful in modeling real world behavior, which often is far from optimal as we ordinarily define "optimal." Yet, we are at a very primitive stage in our understanding of evolution, and its exploration challenges the traditional laboratory approach. However, newly developed computer models of evolution do allow us to examine evolutionary hypotheses. Hence, a particularly fruitful avenue for modeling at this time is simulation of evolutionary scenarios.

Our knowledge concerning the behavior of biological organisms is currently so incomplete that one primary goal must be to improve our description of behavior. There are a limitless number of experiments that we could in principle perform and observations that might be made toward fleshing out our knowledge of behavior. Thus, in our view, a primary goal of a behavioral model at this time is to guide further research, that is, incite us to perform an experiment with a living subject that tests a novel hypothesis of the model. The present models are all incorrect to greater or lesser degree; the goal is to gain more illuminating data; and these data will constrain future models which, in turn, will stimulate additional research. This is not to deny the many other uses of models (e.g., for engineering purposes) that Webb describes. But for researchers studying live animals, the primary function of models is to assist the investigator in deciding which experiments to perform. It is from these experiments – experiments that often refute models – that new models are born. This evolution of models is the stuff of scientific progress. Models that are demonstrably invalid can still be relevant (in Webb's sense of the word) if researchers continue to use them to inspire research, which they do until a better model comes along. The Rescorla and Wagner (1972) model is a good example of this.

Modeling forces scientists to use clearer definitions of variables and of the algorithms to manipulate those variables in order to model a system (a living organism, part of it, or a function of it [e.g., learning]). Thus, one of the advantages of the modeling approach is that implementation of theories using robots (or computers) could provide us with an explicit and common language (computational algorithms) to discuss biological models, as well as, for example, psychological models (e.g., models of learning, memory, perception, motivation, etc.) Once people learn a common language, discussion about underlying processes as well as specific models (and their implementations) is often more fruitful than prior to the development of a common language. However, there is a down side to researchers using a common language. A common language often constrains one's thinking, thereby inhibiting the development of radically new models.

Additionally, robotic implementations and simulations facilitate top-down and bottom-up connections with other levels of explanations (e.g., neurophysiology [neuroscience], associations analysis, normative analysis, and higher-order cognitive processes including decision-making and language). Moreover, such implementations also facilitate integration within the same level of explanation (e.g., between Pavlovian conditioning and instrumental conditioning, between acquisition and retrieval processes, and between learning, memory, perception, attention, motivation, etc.). Webb briefly discusses some of these relationships between different levels of explanations, but does not consider the potential of biorobotics to integrate knowledge in the same level of explanation, which is one of the most interesting aspects of this approach (although perhaps more futuristic).

Webb argues for examination of models in "real" (meaning "complex") environments. Surely such testing is necessary to determine ecological validity. But to provide the principles to be tested in the real world, initial studies in the sterile (ecologically invalid) laboratory are often essential. Just as models are attempts at simplifications of animal information processing, so too, laboratory settings can be simplifications of ecological niches.

ACKNOWLEDGMENT
Support for the preparation of this commentary was provided by National Institute of Mental Health Grant 33881.

Artificial systems as models in biological cybernetics

Titus R. Neumann,[a] Susanne Huber[b]
and Heinrich H. Bülthoff[a]

[a]*Max Planck Institute for Biological Cybernetics, 72076 Tübingen, Germany;*
[b]*Friedrich Miescher Laboratory, 72076 Tübingen, Germany.*
{titus.neumann;susanne.huber;heinrich.buelthoff}@tuebingen.mpg.de
http://www.kyb.tuebingen.mpg.de/~titus
http://www.fml.tuebingen.mpg.de/schwarzer/
http://www.kyb.tuebingen.mpg.de/~hhb

Abstract: From the perspective of biological cybernetics, "real world" robots have no fundamental advantage over computer simulations when used as models for biological behavior. They can even weaken biological relevance. From an engineering point of view, however, robots can benefit from solutions found in biological systems. We emphasize the importance of this distinction and give examples for artificial systems based on insect biology.

One of the primary goals of biological cybernetics is to find the control processes and invariances of biological behavior which can often be represented as formal or mathematical models. A possible way of testing and evaluating these models is to implement them in artificial systems such as computer simulations or robots which reproduce certain aspects of the target system's behavior. From this perspective, artificial systems can be a useful contribution to the understanding of biological findings. From an engineering point of view, however, biological relevance is not important. Instead, particular sensor, effector, and control problems have to be solved. This can be facilitated by applying findings from biological studies that show how natural organisms accomplish similar tasks. The following examples address both the modeling and the engineering aspects of building biologically motivated artificial systems.

Franz et al. (1998) present a purely vision-based scheme for learning a topological representation of an open environment, a so-called view graph. In computer simulations and robot experiments they demonstrate that complex visual exploration and navigation tasks can be performed without using explicit metric representations or "mental 3D models" of the environment. Huber et al. (1999) apply models of the visual processing system of flies to artificial agents in order to generate appropriate motor signals for course stabilization, obstacle avoidance, and fixation of stationary targets. The control algorithms were tested on both a simulated agent and a mobile robot using visual stimuli comparable to those in the original experiments with real flies. In the above examples, computer simulations and robot experiments give comparable results and are both used to test and support the plausibility of biological models. At the same time, the robot implementations demonstrate that the control algorithms derived from these models can be applied to "real world" engineering problems.

In our view, robots have no fundamental advantage over computer simulations in terms of biological relevance, whereas computer simulations can provide a more accurate representation of an organism and its environment. We agree with Webb that "computer programs can represent a wider range of possible situations than we can physically model, but physical models cannot break the laws of physics." However, robot implementations are restricted by the available sensor and actuator hardware, which in most cases differ fundamentally from the biological target system. As an example, walking or flying insects do not use CCD cameras as visual sensors or wheels as effectors for locomotion. Building a robot usually involves considerable engineering efforts in order to make the system work under these hardware constraints. This might require adaptations of the sensorimotor processing and simplifications of the robot's environment in order to match the capabilities and noise characteristics of the given hardware. Although both the robot and the biological target system move and behave in the "real world," that is, both are subject to the laws of physics and have to cope with effects like noise, friction, and other external influences, these effects may be totally different in both systems on account of different hardware for sensors and motors. Therefore, the mere presence of such effects cannot be used as an argument for the use of robot models instead of computer simulations.

Idealization, another point in Webb's criticism of computer simulations, is in our view possible in both robots and computer simulations, for example, by using global or external knowledge about the world. These so-called oracles provide information that is not directly available to the biological target system in its environment, such as the position or distance of objects, or idealized optic flow fields as input stimuli. Idealization can be avoided in both robots and computer simulations by using only information that is also accessible for the biological target system.

In computer simulations, not only does the organism itself need to be modeled, but the environment in which it moves and behaves also has to be represented in an appropriate way. The rapid development in computer graphics and virtual reality technology provides the means to generate photorealistic virtual environments. The stimuli produced with these techniques have reached a quality and realism that is even sufficient for studies in human psychophysics (Bülthoff & van Veen 2001). It is also possible to make highly realistic simulation models of physical properties of an agent and its environment, such as motion dynamics, gravity, friction, or noise in the sensor and motor systems. Terzopoulos et al. (1994) demonstrate physics-based computer simulations of artificial fishes showing a variety of behaviors such as locomotion, obstacle avoidance, and tracking. Faloutsos et al. (2001) show a highly realistic physics-based simulation of complex human body motions such as falling and standing up using a detailed anatomical model of the human skeleton. It would be a major challenge to achieve this degree of anatomical accuracy in a physical robot model.

Even if both implementations are feasible, computer simulations could be more useful tools in modeling biological behavior than robots, since they provide full control over the entire action-perception cycle. Furthermore, simulations are not restricted to real time or real size, so they can represent biological processes that are too slow, too fast, too large or too small for a real-world robot implementation. Neumann and Bülthoff (2001) use computer simulations to demonstrate that three-dimensional flight with all six degrees of freedom can be visually stabilized using models of spatial orientation strategies found in insects. These strategies exploit the distribution of local light intensities and local image motion in an omnidirectional field of view, and include mechanisms for attitude control, course stabilization, obstacle avoidance, and altitude control. The motor system and flight dynamics of the artificial agent is a simplified model of the fruitfly *Drosophila* and includes effects of drag due to air viscosity. With computer simulations it is possible to represent such effects, which would be extremely difficult to achieve in a robot implementation.

Research, robots, and reality: A statement on current trends in biorobotics

Ernst Niebur,[a] Mounya Elhilali,[b] Iyad Obeid,[c] Justin Werfel,[d] Mark Blanchard,[e] Mattia Frasca,[f] Kaushik Ghose,[g] Constanze Hofstoetter,[e] Giovanni Indiveri[h] and Mark W. Tilden[i]

[a]*Zanvyl Krieger Mind/Brain Institute and Department of Neuroscience, Johns Hopkins University, Baltimore, MD 21218;* [b]*Institute for Systems Research, Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742;* [c]*Department of Biomedical Engineering, Duke University, Durham, NC 27708;* [d]*Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139;* [e]*Institute of Neuroinformatics, University of Zurich/ETH Zurich, CH-8057 Zurich, Switzerland;* [f]*Department of Electrical, Electronical and Systems Engineering, University of Catania, 95125 Catania, Italy;* [g]*Department of Psychology, University of Maryland at College Park, College Park, MD 20742;* [h]*FhG-AiS, Fraunhofer Institute for Autonomous Intelligent Systems Technology, 53754 Sankt Augustin, Germany;* [i]*Institute for Physical Sciences, Los Alamos, NM 87544.* **niebur@jhu.edu http://cnslab.mb.jhu.edu mounya@eng.umd.edu http://www.ece.umd.edu io@duke.edu http://www.duke.edu/~io jkwerfel@mit.edu http:/hebb.mit.edu/people/jkwerfel jmb@ini.phys.ethz.ch http://www.ini.unizh.ch mfrasca@dees.unict.it kghose@wam.umd.edu http://www.wam.umd.edu/~kghose/research.html connie@ini.phys.ethz.ch http://www.ini.unizh.ch giovanni.indiveri@ais.fraunhofer.de http://www.ais.fraunhofer.de mwtilden@qwest.net**

Abstract: While robotics has benefited from inspiration gained from biology, the opposite is not the case: there are few if any cases in which robotic models have lead to genuine insight into biology. We analyze the reasons why biorobotics has been essentially a one-way street. We argue that the development of better tools is essential for progress in this field.

We will here use the term robot to describe a hardware model of a biological system whose interaction with the physical environment, both in terms of sensors and of actuators, forms an essential part of the model. The question asked in the title of Webb's article is whether such robots can be useful for understanding biology. Our perspective as a group working at the interface between biology and robotics is that robot models have the potential to make considerable contributions, with significant advantages over other styles of analysis, but that this potential is not being fully exploited at this time.

Robots vs. other models. We compare robot models with computational models (i.e., numerical simulations) and theoretical models (i.e., mathematical abstractions) on the one hand, and biological models, on the other. Computational and theoretical models are devoid of any physical substrate. Although they are comparatively easy to implement, compared to a robotic model of the same biological system, oversimplification due to abstraction gone too far is a significant risk for them. For instance, the modeller must decide what (if any) external noise is to be included and what form to give it; this decision may influence the outcome strongly. A robot model will by its nature be subject to all the actual constraints and conditions of the real world, which cannot be ignored or finessed away. Another disadvantage of computational models is that some properties of the system or its environment may actually be more difficult or costly to simulate in software than in hardware (e.g., nonlinear friction, requirement for real-time response, etc.).

Biological models – those using organisms, cultured cells, brain slices, and so on as their substrate – have other limitations. First, they are vastly more complicated than hardware models, involving complex biological tissue or even whole organisms. Gaining a deep understanding of the system may therefore be difficult. Not only is a robot model simpler than an animal model; since we construct it ourselves, its components and their interactions are known down to the lowest level.

In principle, using robot models rather than animal models may also be preferable because of ethical concerns. At this time, we feel that this is of limited importance because the current level of robotics does not allow detailed modelling of behaviors that are only found in animals of higher phyla, for which strong ethical considerations come into play.

A one-way street – so far. Despite these benefits, however, the flow of information between biology and robotics is at present almost entirely one-directional. While machine builders receive inspiration from biology, examples of significant discoveries in biological systems that were inspired by building robots are, at best, rare. Webb lists some examples in her target article, but they are few and far between. It is not clear whether there are yet any cases in which robot models lead to nontrivial, successful predictions that have been actually confirmed in animals. This is in marked contrast to other modeling techniques, notable especially in computational neuroscience, where computer modelling has become a respected technique among biologists; the surest sign of this being that many experimental groups routinely develop computational simulations themselves.

Why has robotics not been similarly successful? One reason is because the field is still relatively new and small. Biorobotics, in the sense of robots being used to provide insight into biology, arguably started about fifteen years ago with a paper by Brooks (1985), several decades after computational models were first introduced. Furthermore, the number of active researchers in the field is still very small. This is not counting the large number of those who may have completed one robotics project and then reverted to more classical methods. We have encountered a large number of cases of model recidivism, in which a computational model was implemented in hardware but, in the further course of the project, the hardware implementation was abandoned in favor of future development of the computational model. Presumably, it was found that pursuing the hardware implementation is more difficult, expensive, and time-consuming than the implementation of a simulation. This brings us to the second issue, the difficulty of the approach.

Constructing robots is a difficult, expensive process that takes a long time from original design to finished prototype. Moreover, materials are non-standard, and at present, essentially every model has to be developed from scratch. Without doubt, the field would make much faster progress if a robotic equivalent of the PC existed – a low-cost, universally available, and standardized platform allowing rapid prototyping and seamless collaboration of large groups of researchers. A few candidate systems exist, both in hardware (e.g., Lego Mindstorms™ [http:www.legomindstorms.com], the K-team robots [http:www.k-team.com], Tilden's bugs [Haslacher & Tilden 1995]), and in software (e.g., IQR [Verschure & Voeglin 1998]) but at this time it is not clear whether any of these, or any others, will be able to play the role for biorobotics that the PC played for computational modelling.

The way out. While all of this may sound pessimistic, we remain hopeful about the future. First, we believe that systems-level approaches will increase in importance. Second, although we are still far away from the situation in computation where a nearly universal hardware infrastructure is cheaply and readily available, prices of robotic equipment have come down by several orders of magnitude in less than a decade, and the trend continues. Developing hardware models of biological systems may never become a method for everyone, but it will play a larger role once tools become available that will make robotic modeling accessible to a larger part of the scientific world.

ACKNOWLEDGMENT
We thank the National Science Foundation and the Gatsby Foundation for the support of the Telluride Workshop on Neuromorphic Engineering where this work was performed.

The conundrum of correlation and causation

Irene M. Pepperberg

The Media Lab, Massachusetts Instititute of Technology, Cambridge, MA, 02139. **impepper@media.mit.edu** **http://www.alexfoundation.org**

Abstract: Biology can inspire robotic simulations of behavior and thus advance robotics, but the validity of drawing conclusions about real behavior from robotic models is questionable. Robotic models, particularly of learning, do not account, for example, for (a) exaptation: co-opting of previously evolved functions for new behavior, (b) learning through observation, (c) complex biological reality, or (d) limits on computational capacity.

Although Webb presents an important review of robotic models, including excellent guidelines for their biological relevance, she admits that "a model that behaves like its target is not necessarily an explanation of the target's behavior," that is, that correlation is not necessarily causation. Such oft-repeated statements, however, do not dampen her enthusiasm for robotics as a means to understand biological form and function. What her account lacks is additional appreciation of potential problems inherent in using robotics to answer biological questions. She fails to acknowledge exaptation, observational learning, and complexity as biological reality; she underestimates limits on computational capacity.

Exaptation, or the co-opting of previously evolved functions to do new things, can seriously compromise robotic simulation. Evolutionary forces work on existent biology, and thus real-life biological solutions may involve mechanisms less efficient than those used robotically. Hewes (1973), for example, argues that spoken language was derived from gestural forms without major neural restructuring. Data supporting Hewes' hypothesis – and the notion that exaptation of gestural neural substrates for communication may be extremely widespread – are that parallel development of physical and communicative combinatorial acts exists in humans, nonhuman primates, and even Grey parrots (Greenfield, 1991; Pepperberg & Shive 2001). Mechanisms used by a robotic system to model acquisition of spoken language *de novo* might reproduce data, but are unlikely to use circuits derived initially for stacking cups in order to combine labels. Thus, its mechanisms would be removed from, and say little about, those of biological systems.

Observational learning is also widespread in animals (Heyes & Galef 1996). Animals would die before they could reproduce if they had to learn skills such as predator avoidance or what to eat via the trial-and-error mechanisms that are currently the basis for computer modeling (Pepperberg 2001). Even in the most elegant attempts at imitation simulation, which involve some form of programming by example, the extent to which the computer learns is limited (e.g., Lieberman 2001; Weng et al. 2001). Thus, the current relevance of robotics to forms of learning beyond simple associationist principles, and to real-life systems, is limited.

In a related vein, biology is complex. Advanced learning involves the ability to choose the set of rules, among many learned possibilities, from which the appropriate response can be made, and the creativity to build upon learned information to devise novel solutions to a problem. In contrast, conditioned learning is limited in scope in that it does not allow a robot even the ability to alter behavior quickly based on the immediate past, much less allow immediate flexibility to respond to changing conditions. True, brute force systems such as Big Blue win chess games with stunning success (e.g., Campbell 1996), but such systems cannot learn in a broad manner, that is, cannot integrate new and existing knowledge to solve novel problems, take knowledge acquired in one domain to solve problems in another, or form and manipulate representations to attain concrete goals. The point is not that associative/conditioned learning is irrelevant: It exists, is a basis for learning, can be seen as basic to the programming language of learning . . . ; but it is not the appropriate overall program for learning, because it does not engender generalization, transfer, or insightful behavior. The simple initial association of stimulus and response may be what is first linked in memory in humans, but for humans repeated interactions in the real world both sharpen and broaden the connections (Bloom 2000); what results is a representation. Robots can indeed be programmed so that repeated interactions improve their decision-making ability, and one might even argue that statistically-based similarity coding might constitute a representation. Advanced learning, however, derives from manipulation of representations. What is needed to devise an intelligent learning machine, therefore, is not a more efficient program that takes a stimulus as input and uses various rules to produce an expected response, but one that takes that stimulus and uses creativity, reasoning, and decisions based on context to produce an appropriate, adapted, adaptive behavior. So far, robotics cannot simulate such behavior.

Finally, the computational or robotic capacity used to produce a model might be less than the computational capacity of the living system; we cannot discount real-life mechanisms because simulations cannot reproduce the data. Webb cites Kuwana et al. (1995), who must use the actual antenna of moths on their robot model because available gas sensors are ten thousand times less sensitive than the biological system. Later she comments on the rejection of lobsters' use of instantaneous differences in concentration gradients between their two antennules to do chemotaxis, simply because robotic implementation of this algorithm in the real lobsters' flow-tank failed (Grasso et al. 2000) – that is, she implies that failure could be merely a consequence of the quality of the robotic sensor. I applaud Webb's inferences, but suggest that these problems are more serious than she surmises.

In sum, robotic design can advance from attempts to simulate animal behavior without worrying about simulating exact mechanisms. But using current robotic simulations (which for learning are predominantly based on associationist principles) to answer questions about real-life systems can lead into a trap identical to that of Skinnerian behaviorism, which found many situations these same laws could not explain. Anomalous activities of animals whose natural responses to stimuli could not be reshaped by behavioristic training (e.g., Breland & Breland 1961; see review by Roitblat 1987) required a new paradigm in which animals were seen as multi-level processors of information (Kamil 1984; 1988; Pepperberg 1990). This need was made even clearer by behavioral ecologists, whose data could be explained only by positing mechanisms such as selective attention and long-term memory (e.g., Kamil & Sargent 1981; Pyke et al. 1977; see also Roitblat 1987), which were not part of the behaviorist tool-box. An existence proof via robotic simulation that shows that an animal does not, for example, need a cognitive map does not mean that such a map is not used. "Simple" mechanisms are not necessarily how biology works, and trying to create complex mechanisms by using brute processing force to link enormous numbers of simple mechanisms can lead to serious misunderstandings of real cognitive and neural processes.

ACKNOWLEDGMENT
Preparation of this commentary was supported by the MIT Media Lab.

Living and learning

John Pickering
Psychology Department, Warwick University, Coventry, CV4 7AL United Kingdom. **j.pickering@warwick.ac.uk**

Abstract: To be plausible, biorobots will need to build themselves. Such autopoietic systems will be autonomous, active learners whose functional architecture is a joint product of factors supplied by the designer and factors learned from encountering an environment. Creating such biorobots will require appropriate theories of cognition, learning, and evolution, all of which are available.

Models of biological behaviour have been with us for centuries, if not millennia. From Daedalus to the comic mechanical animals of the eighteenth century, we see a progression of attempts to model biological behaviour in artefacts of one sort or another. What Barbara Webb's comprehensive review shows is that this project continues with the biorobotics research programme which is growing vigorously. This comment addresses something that may be missing, or at least under-represented in the programme: the role of learning.

Living cognitive systems actively explore the meaning of their surroundings and learn to act more effectively from what they encounter. They create themselves through action, something which Maturana and Varela call "autopoiesis" (Maturana & Varela 1987). In Piagetian terms this is the outcome of a dynamic balance between assimilation and accommodation. Organic systems assimilate information about the world mediated by their internal structure. This structure then accommodates, adaptively, in turn allowing a broader range of environmental inputs to be assimilated. This growth of intelligence is driven by spontaneous living behaviour. Behaviour is the driver of change, not just over the lifetime of a single organism but also over a longer evolutionary timescale. As Piaget puts it: "Behaviour is the motor of evolution" (Piaget 1979, Ch. 9).

Webb shows that biorobots are becoming capable of fairly autonomous behaviour. If they could also learn, they might develop and even evolve in biologically plausible ways. Something like this has been in the air for some time, roughly since the advent of connectionism, which offers new ways to describe the interaction of innate and learned factors (e.g., Bates et al. 1998; Maynard-Smith 1987).

Now, some models of biological behaviour may be so closely tied to an actual biological substrate that learning would not figure to any significant degree. Models of cell membrane dynamics might be an example here. However, for neural networks and whole-organism models such as the behaviour-based robotics of recent years, learning through activity is more important. Not just because learning is itself a biological behaviour but also because it can help to create the robot. Then the plausibility of a biorobot ceases to be simply a function of design and becomes a more complex function of design, learning, and a history of growth and change. This history is dependant on what the robot encounters as it behaves in a given environment.

While many of the current generation of robots can learn, the learning is sometimes of a rather passive Skinnerian sort. Skinner concentrated on how animals learned the consequences of their actions. He was less concerned with the internal changes that accompanied learning and with what it was that made animals active in the first place. As such, his is a poor lead to follow when trying to model the real biological role of learning.

Instead, it would be more helpful to regard learning as the coupling the internal structure of an active cognitive system with the environment in which it acts. This "structural coupling" according to Maturana and Varela arises from the mutual evolution of autopoietic organisms and their environments (Maturana & Varela 1987). Indeed, Maturana and Varela have proposed a theory of learning, action, and cognition that might serve well as the basis for plausible biorobots.

This is not to reject current efforts but to suggest how to make the functional architecture of biorobots more interesting and plausible. If this architecture is simply specified in advance by the designer, then a biorobot runs the risk of being a mere simulacrum, a distant, if more sophisticated, descendant of the eighteenth century automata. By contrast, if a robot's functional architecture is the result of autonomous learning, then its behaviour becomes the more interesting and the less liable to the usual critique of robots that "you only get out what you put in."

Moreover, it aligns biorobotics with an important shift in cognitive science. There is a move away from simple computational metaphors and towards the language of dynamic systems theory in which cognition is treated as embodied action distributed within specific situations (e.g., Clark 1997; 1998; 2001). Learning, action, and change are central to this shift. As Maturana and Varela put it: "All knowing is doing and all doing is knowing" (Maturana & Varela 1987, p. 27). Cognition is biological action and not mere computation. It depends on having a body of a certain sort that has, through active learning, become adapted to particular situations.

Thus, any biorobotic model has to take into account its own physical properties as well as the situation in which its cognitive skills are expressed. Moreover, what actually comprises the system is not merely that which is internal to the robot but also the structure of the environment to which it is coupled. If this coupling emerges as an active biorobot learns from its encounter with its environment, then it will be a more plausible model of biological behaviour.

Modelling criteria: Not just for robots

George N. Reeke
Laboratory of Biological Modelling, The Rockefeller University, New York, NY 10021. **reeke@lobimo.rockefeller.edu** **www.rockefeller.edu**

Abstract: Webb's scheme for classifying behavioral models is applicable to a wide range of theories and simulations, nonrobotic as well as robotic. It is suggested that a meta-analysis of existing models, characterized according to the proposed scheme, could identify regions of the seven-dimensional modelling space that are particularly likely to lead to new insights in understanding behavior.

Barbara Webb has given us a very rich and yet very practical set of dimensions that can be used to characterize and categorize models of cognitive and neural systems, along with a quite comprehensive survey of the biorobotics literature to show how various approaches fit into her scheme. While she has confined her discussion to robotic models, professedly to keep the work to a manageable size, it is important to note that the proposed classification (with the possible exception of dimension 7, medium of implementation) is in fact applicable to essentially the entire range of models proposed in the brain and behavioral sciences. This is clear when one considers that biorobots are, after all, instantiations of more or less abstract theories.

Further, Webb has quite rightly refrained in the target article from attaching value judgements to different locations in her modelling space; nonetheless, it is tempting to think that some locations in this space may be more favorably situated than others to provide fruitful theories of behavior. For this reason, it can be hoped that other authors will take up the challenge and undertake a meta-analysis of existing models with the aim of identifying positions along Webb's dimensions where the most productive models can be found.

In this spirit, I would like to amplify a bit on several of the points that Webb has made in order perhaps to identify a few regions where robotic models may have particular advantages:

Simulating the whole hypothesis and nothing but the hypothesis. As G. M. Edelman and I have pointed out some time ago (Reeke & Edelman 1988), models that address only a circumscribed part of a cognitive system must be very carefully constructed and interpreted in order to avoid the danger of inadvertently introducing homunculi at the interface between the model and the world – unspecified components (usually the mind of the modeller) that are not part of the model but that carry out functions critical to its successful performance. Biorobots, or at least the autonomous variety of biorobots, because they interface directly with the world, avoid this hazard. Of course, it is also possible, with suitable care, to avoid homunculi in nonrobotic model systems (Reeke et al. 1990).

Dealing with the real world. Dealing with the real world has another big advantage: It forces the modeller to address problems that a real organism has to face, but which are easily avoided by felicitous choice of input and output representations in more abstract simulations. This is similar to the homunculus problem mentioned above, except now the homuncular mechanisms are explicit rather than inadvertent. Examples of such problems would include: recognizing visual patterns in the face of scale and rotational perturbations caused by the relative motion of the subject and the object; manipulating objects that respond in unexpected ways to touch, for example, by falling over or running away; compensating for unreliable effector mechanisms. It is easy to dismiss these problems as secondary when developing an abstract model; building a robot forces one to see that they are fundamental.

Practicality: Operation in real time. Robots need to respond quickly enough to changes in the environment that their "plans" for solving a problem are not rendered irrelevant by new circumstances. While some allowance can be made for the relative speeds of available microprocessors versus brains, ultimately, cognitive models, whether instantiated as robots or not, must meet this test.

Isomorphism: The balance between realism and simplicity. Webb makes it quite clear that realism and simplicity are not mutually exclusive. She reminds us of the warning by Churchland et al. (1990) that an overly faithful model is as likely to be incomprehensible as the system itself. The other side of this coin is that an overly simplified model is likely to miss key aspects of the problem that the biological system has to solve. A widely cited example of this is the "Wickelfeature" representation of words used by Rumelhart and McClelland (1986) in their early model of English past tense learning. Prince and Pinker (1988) have discussed in detail why this representation was too simple. Webb's discussion of abstraction provides some helpful guidelines for arriving at an appropriate balance. Perhaps the trick is to abstract the solution without abstracting too much of the problem. Instantiation of a model as a robot can be a valuable aid in evaluating the costs versus paybacks of adding complications to that model.

Dimensions of modelling: Generality and integrativeness

Jeffrey C. Schank

Department of Psychology, University of California at Davis, Davis, CA 95616. **jcschank@ucdavis.edu**
http://psychology.ucdavis.edu/Schank/

Abstract: Webb has articulated a clear, multi-dimensional framework for discussing simulation models and modelling strategies. This framework will likely co-evolve with modelling. As such, it will be important to continue to clarify these dimensions and perhaps add to them. I discuss the dimension of generality and suggest that a dimension of integrativeness may also be needed.

Webb's multi-dimensional framework provides a good starting point for discussing modelling strategies. As a framework, it should not be construed as a formal philosophical analysis of models or modelling. Instead, it is likely to co-evolve with the development of modelling strategies. I will focus on two issues. One concerns the dimension of generality. The other concerns the long-term need for integrative modelling strategies.

Generality is a key dimension of modelling articulated by Webb, and by Levins (1966), with the aim of discussing model generality in the *practice of modelling.* While there are many philosophical views of generality, most are not applicable to the practice of modelling. However, in Webb's view, by studying specific systems we may discover core ideas that are used over and over again in models applied to a broad range of systems. This suggests that generality emerges from applications of modelling, and that general models share core ideas in common.

To further clarify this important idea, it is useful to contrast it with the opposite idea that general models are more detailed models (Orzack & Sober 1993). To evaluate these contrasting views and their relevance to modelling practice, it is best to examine them in actual practice. The biological control of insect pests provides a simple yet representative area of biological modelling for comparing alternative views of generality.

Modelling biological control has had an extensive history of application to host-parasitoid systems in part because modelers could construct relatively simple yet biologically plausible models (Mills & Getz 1996). For example, the specificity of many parasitoids makes it biologically plausible to treat host-parasitoid systems as closed; the generation times of hosts and parasitoids are often nearly the same; and because adult parasitoid females attack hosts, age structure can be simplified or safely ignored.

For discrete generations, Nicholson and Bailey (1935) developed a basic model with two core ideas. First, the number N_e of encounters of hosts with P_t parasitoids is proportional to the number of hosts N_t. This core idea is expressed as $N_e = a\,N_t P_t$ (Hassell 1978). Second, the N_e encounters are randomly distributed among the available hosts. Thus, the probability that a host is not attacked is the zero term of the Poisson distribution: $p_0 = \exp(-N_e/N_t)$, where N_e/N_t is the mean number of encounters per host. The actual number of hosts parasitized is $N_a = N_t[1 - \exp(-N_e/N_t)]$. Now, assuming that (i) $N_e/N_t = aP_t$, (ii) each host parasitized produces one adult parasitoid (e.g., phorid flies lay a single egg in each fire ant attacked), and (iii) the rate of increase of hosts is λ, then the basic discrete-generation model is illustrated in Figure 1**a**.

Many variations of this basic model have been proposed over the years to address problems such as functional responses and non-random searches of parasitoids, but the fundamental problem facing these models from the beginning was population stability (Hassell 1978; Mills & Getz 1996). Indeed, three models were proposed that directly addressed this problem (Fig. 1**b, c, d**; Mills & Getz 1996). All three produce qualitatively similar behavior but not necessarily for the same range of parameter values, and biologically unrealistic behavior can occur for some parameter values of m and n in **b** and **c** respectively (Hassell 1978).

$$\begin{tikzcd}
 & \mathbf{a}\quad \begin{array}{l} N_{t+1} = \lambda N_t e^{-aP_t} \\ P_{t+1} = N_t(1 - e^{-aP_t}) \end{array} \arrow[dl] \arrow[d] \arrow[dr] & \\
\mathbf{b}\quad \begin{array}{l} N_{t+1} = \lambda N_t^{(1-m)} e^{-aP_t} \\ P_{t+1} = N_t^{(1-m)}(1 - e^{-aP_t}) \end{array} & \mathbf{c}\quad \begin{array}{l} N_{t+1} = \lambda N_t e^{-aP_t^{(1-n)}} \\ P_{t+1} = N_t(1 - e^{-aP_t^{(1-n)}}) \end{array} & \mathbf{d}\quad \begin{array}{l} N_{t+1} = N_t e^{r(1-N_t/K)-aP_t} \\ P_{t+1} = N_t(1 - e^{-aP_t}) \end{array} \\
 & \mathbf{e}\quad \begin{array}{l} N_{t+1} = \lambda N_t^{(1-m)} e^{r(1-N_t/K)-aP_t^{(1-n)}} \\ P_{t+1} = N_t^{(1-m)}(1 - e^{-aP_t^{(1-n)}}) \end{array} \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}$$

Figure 1 (Schank). Two views of generality: relationships between the basic Nicholson-Bailey model (**a**), three stability models (middle), and a model, **e**, from the middle models are derivable.

In Orzack and Sober's (1993) view of generality, model **e** is both the most *general* and realistic form of these stability models. This is because models **b**, **c**, and **d** are mathematically derivable from **e,** given specific assumptions (**a** is also derivable from **b**, **c**, and **d,** given specific assumptions). For example, **b** is derivable from **e** on the assumptions that $n = 0$, K approaches infinity, and $e^r = \lambda$. However, most biological modelers would consider **e** neither general nor realistic but rather biologically absurd (Levins 1993).

The process of modelling does not start from the most complex set of equations that can be conceived and then subsequently pared down to specifically applicable models. Instead, modelers search for core ideas that can be used to model a variety of systems as Webb suggests. Thinking about generality in the context of modelling leads to a radically different view of the products of modelling: A view in which the products of modelling are more like phylogenies of models rather than extraordinarily elaborate and detailed models.

Another issue that is likely to become increasingly important in future research concerns how to integrate different types of models and modelling strategies to achieve the aims of increasingly sophisticated research programs. Organisms are paradigms of complex systems in which causal influences on behavior circulate up and down levels of organization (Schank 2001). A long-term challenge will be to develop strategies for integrating the different types of models required to model these complex bio-behavioral systems. These strategies must allow us to efficiently and rationally build models that facilitate deeper theoretical and empirical understanding of organisms and their behavior. This may be especially challenging because, broadly construed, simulation models include not only mathematical and robotic models but also experimental designs (Levins 1993).

Such integrative modelling strategies will require more than the coordinated use of different types of models. We will need to articulate strategies for developing multi-level models that integrate different types of models at multiple levels of organization (Schank 2001). We will need strategies for how one type of modelling (e.g., robotic) may inform and guide the development of other types of modelling (e.g., designing experiments). We will also need to articulate how information about model failures can be used to build better models (Wimsatt 1987), and to find ways to use model failures of a given type (e.g., robotic model failures) to build better models of other types (e.g., experimental designs or computer simulations). These considerations strongly suggest that future discussions of modelling will require an additional integrative dimension for discussing how different types of models can facilitate integrative strategies.

ACKNOWLEDGMENT
I thank Ted Morris for helpful comments.

Is there more to "model" than "muddle"?

Matthias Scheutz
Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556. **mscheutz@nd.edu**
http://www.nd.edu/~mscheutz/

Abstract: Any discussion comparing different models with respect to their quality *qua* models must presuppose a notion of model, that is, what it is to be a model. While Webb provides seven criteria to assess the quality of various proposed biorobotic models, she does not clarify the very notion of "model of animal behavior" itself.

The term "model," like any widely used term of significant theoretical importance, connotes a variety of concepts. Barbara Webb demonstrates this convincingly, quoting from many research articles in the various subfields of the behavioral sciences. This "model muddle" is unfortunate, especially because researchers involved in "modelling" animal behavior have clear intuitions about the aims and significance of their endeavors, even if it is not obvious how to formulate the underlying concept of model precisely. If the notion of "model," despite its numerous, partly incompatible and even opposing construals, is of enough theoretical and practical interest to be part of our explanatory framework, then significant effort should be devoted to making this notion precise, for it underwrites the very practice of "modelling."

Unfortunately, "modelling" *qua* practice is often covertly conflated with the "modelling relationship" between two systems. Webb's article is no exception (e.g., see sect. 2.2). If the purpose of "modelling" is "often to discover what are the 'relevant feature' or 'essential structures'," then there have to be such things as *features* or *structures* of target systems in the first place, to make sense out of what modellers do when they try to "model" them. In other words, the practice of modelling is conceptually dependent upon a notion of model. Webb, however, claims that "model usage cannot depend on prior knowledge of what [the relevant features and essential structures are] to establish the modelling relationship," conflating the epistemological process of establishing "the modelling relationship" between two systems with the independent ontological question of whether, in fact, any such relationship exists between them. Even if we focus on the epistemological question, it seems unlikely that modellers approach their task without any theoretical prejudice as to why a particular system (simulation, robotic, etc.) may count as a model of the target system. That certain features of the model may turn out to be of higher relevance to the investigated target than others in the course of the modelling processes, does not take away from the fact that there has to be a modelling relationship in the first place that enables one to assess the relevance of various features of the model.

While Webb takes "modelling" to involve "the relationship of representation or correspondence between a (real) target system and something else," she seems to dismiss the possibility of a general notion of model, based on a notion of correspondence. In section 2.1, for example, she infers from the fact that "in the vast majority of cases, models are not (mathematical) isomorphisms, nor are they intended to be," that any attempt to define the notion of model in terms of structural relationships (such as homomorphisms and isomorphisms) is doomed to fail. In particular, she confers the label "oxymoron" upon the notion of *partial isomorphism* and claims that it "cannot suffice to be used in valid deduction." Not only is this conclusion contrary to standard practice in (logical) "model theory," where partial isomorphisms are routinely employed to study the relationship of structures that are not themselves isomorphic, yet share isomorphic substructures (e.g., Ebbinghaus et al. 1980), but it is also not true. To see why, one needs to distinguish between *descriptions* of systems and the systems themselves: whereas the modelling relationship is a structural relationship between two systems, the notion of deduction pertains to formal proof systems, which in turn presuppose a (formal) language (in order to deduce properties of the target system from properties of the model system, a proof system is required in a language, in which both the target system and the model can be described).

The logical notion of partial isomorphism is defined for two (formal) structures in a common language as a bijective mapping from a subset of the domain from one structure to a subset of the domain of the other, which respects all functions and relations of the language. This can be extended to real-world systems by taking the domain of a structure to consist of the parts of a system (at some level of description). Consequently, the mapping needs to be established between parts of the target and parts of the model such that the structural properties (as determined by the common language) of either system are preserved.

Once the notion of partial isomorphism is introduced for two systems, it can very well be used for deductions as long as they are restricted to isomorphic subsystems under partial isomorphisms. In fact, it is exactly this notion of partial isomorphism that seems to lie at the heart of a general notion of model, as it captures our intuitions about what it means for parts of the model system to *correspond* to parts of the target system (in a systematic way). And it is the existence of such a mapping between substructures that warrants calling one system *a model of* the another.

Webb focuses on "robots used as simulations of animals", that is, on robotic systems that "address a biological hypothesis or demonstrate understanding of a biological system," but what this phrase means remains unclear. It seems that she, after having overtly acknowledged the difficulties buried in the notion of model, covertly uses phrases like the one above or like "the robot model of X," where X is a function or capacity of an animal, as if they were entirely unproblematic. Yet, it is left open when, and to what extent, a robotic model can be said to exhibit or have X; that is, the crucial question of what exactly warrants the claim that a robot system is a model of an animal with respect to X is not answered. While the suggested seven dimensions, along which models in biology can vary, seem helpful in assessing the quality of various proposed biorobotic models, they tacitly presume an agreement on what it is to be a model (of an animal or animal behavior). Webb's article does not make this explicit, despite its intrinsic merit as a detailed overview and comprehensive summary (including a long, very useful bibliography) of current modelling endeavors in biorobotics.

Biomimetic robots and biology

Allen I. Selverston
INLS-0402, University of California, San Diego, La Jolla, CA 92093.
aselverston@ucsd.edu

Abstract: Using robots that operate in the real world as opposed to computer simulations of animal behavior is a form of modeling that may provide some biological insights. However, since engineering principles and materials differ significantly from those used in biology, one should be extremely cautious in interpreting robot biomimicry as providing an explanation of biological mechanisms.

Like the term neural networks, the term "biomimetic robots" has a certain seductive quality that promises more than it can deliver. Just as artificial neural networks do not have anything to do with real neural networks, biomimetic robots cannot duplicate living processes with nonliving materials. For example, there is nothing even remotely similar to muscle tissue available to roboticists. Currently, linear actuators and the computer programs which act as controllers for them are not analogous to muscles and nervous systems. This is true not only at the material level, but in terms of our understanding of the biological principles for controlling movements. Perhaps the greatest deficiency for biomimetic robots, however, is their inability to mimic the biochemical mechanisms that are at the heart of physiological processes. This applies not only to phenomena like energy balance (batteries are a poor substitute for ATP), but to hormonal control of global processes like motivation and drive, as well as the biochemical pathways underlying cellular mechanisms like synaptic transmission. Because biology has been spectacularly successful at the cell and molecular level, neurobiologists today are awash with reductionistic data. Yet, it is a truism that all of this data will not in and of itself be adequate to tell us how nervous systems work. Modeling, therefore, is essential as one way to organize this data into larger structures and eventually into whole organisms. This will enable testing hypotheses and making predictions that can be experimentally verified. Typically, such modeling has been accomplished with computer simulations. Webb suggests that robots usually built by engineers which can be "inspired" by biology, can in fact serve as models for understanding biological mechanisms. In fact, she postulates, they may even be better than computer simulations because they might uncover problems that exist only in real world situations.

There is no universal response to the question of whether or not a physical model might be more advantageous than a computer simulation in answering biological questions. The nervous system has to process physical stimuli by transduction and neural coding using organic structures that are individually and collectively different from man-made physical devices. Nevertheless, some neural systems can work in a way that can be mimicked with engineering principles; and for these processes, we might indeed learn something of biological relevance. The three examples Webb cites illustrate this, and there are many others. However, I think one should be extremely cautious in assuming we can understand all neural mechanisms with physical models. The neurophysiological basis for most behaviors is very incomplete in terms of neural circuits and neuronal properties, so validating behavioral mechanisms with robots that are controlled by computer programs is not going to suggest anything about the circuit mechanisms that underlie animal behavior.

But as Webb has shown in her own work, some of the underlying principles of cricket acoustic behavior can be mimicked without building an entire robotic cricket and, instead, letting the hardware of interest control a mobile platform. Her results can aid in understanding the neural mechanisms used by real crickets even though the circuitry involved in cricket phonotaxis is not completely understood. Where circuits are known in detail, they may be helpful in building robots that perform tasks different from the task the circuit normally controls. We know enough general principles about the circuits operating invertebrate CPGs to use them as the basis for designing controllers for specific sensorimotor tasks other than the ones the circuit may have been specifically designed for. One will probably learn more with known circuits that can be modeled and modified, than by modeling ungrounded arbitrary circuits, especially large scale models in which almost everything is guessed at.

Finally, it seems to me the greatest gap lies in the area of what Webb calls the "performance match" dimension. A simple underdetermined robot that can perform one or two sensorimotor tasks similar to some organism, does not demonstrate biological mimicry.

Most modelers do not test their models sufficiently, and, to the extent that roboticists work with biologists and vice versa, the match between robotic and biological behavior will improve. But at the moment engineers can still get more from biology than the other way around.

The methodology of the artificial

Luc Steels
Sony Computer Science Laboratory – Paris, and VUB AI Laboratory, Brussels, 1050, Belgium. **steels@arti.vub.ac.be** **http://arti.vub.ac.be**

Abstract: Biorobotics research should not only target "realistic" models of living systems and be judged exclusively from that perspective. It should pay just as much attention to formal models and artificial systems. They allow the examination of assumptions which do not necessarily hold for living systems, but precisely therein lies their value. They generate insight by enabling a comparison between the artificial and the real.

Barbara Webb is to be applauded for her courageous effort to examine the methodological assumptions of those researchers in robotics and artificial intelligence who do not exclusively focus on building practical applications, but try to advance scientific knowledge on adaptive behaviour and cognition. Webb focuses on the question of how far robots can be good models of biological behaviour, where "modelling involves the representation or correspondence between a (real) target system and something else" (target article, sect. 2.1). This type of realistic modeling is very common and several robotics researchers reviewed in her paper have stated that their robotic experiments should be interpreted as such. But it is not the only way.

There is a second type of scientific activity which consists in making formal models, as illustrated by work in theoretical biology or theoretical economics. Formal models do not necessarily describe a natural system. But they examine the implications of certain assumptions that can then be used to understand natural systems. The assumptions need not be realistic; on the contrary, unreality of assumptions is often seen as a virtue: "to be important . . . a hypothesis must be descriptively false in its assumptions," says economist Milton Friedman (1953, p. 14). Unrealistic assumptions make it possible to investigate boundary conditions, isolate factors, highlight implications which would otherwise go unnoticed, perform demonstrations by reductio ad absurdum, and so on. Thus, Nowak et.al. (2000) make a number of obviously false assumptions in their formal models of the evolution of language, for example, that the lexicon and grammar of a language are transmitted genetically. But this does not diminish the strength of their claims about why compositionality may be selectively advantageous. A lot of mathematically oriented research in neural networks, and many papers found in conferences on "simulation of adaptive behavior" or "artificial life," similarly explore formal models rather than realistic ones.

Building artificial systems is a third, and in my view, truly alternative type of scientific activity – one that many biorobotics researchers implicitly practice. It consists in building a machine that has a functionality similar to the one performed by a natural system, for example, an airplane capable of artificial flight or an artificial walker capable of walking. An artificial system is, on the one hand, more "realistic" than a formal model because it involves building physical systems that undergo the constraints of nature or of the cultural and social environment in which they are put. But the researcher is not restricted to mimicking natural systems. Building artificial systems is therefore similar to formal modeling: it can explore alternative solution paths, use other boundary conditions, adopt pragmatic solutions that are very different from biological implementations. An airplane is not a realistic model of a flying bird. Its wings do not have feathers, it does not flap its wings, it does not run on two legs. Early attempts to build airplanes by modeling biological systems failed miserably. Similarly, an artificial walker, or an artificial face recognition system, does not have to mimic the solutions adopted by human beings, which are hardly understood anyway. What matters is good performance and this can be measured accurately and objectively. Artificial systems require a level of detail which is not necessary in formal modeling. Thus, Nowak et al. (2000) can simply assume that agents learn a grammar with a certain probability without having to specify the learning algorithm itself – whereas an artificial system researching the same questions would need an operational learning scheme that can work on realistic linguistic input and use self-generated meaning anchored in the world through a sensory-motor apparatus (Steels 1998). Artificial systems are more constrained, however, because the solution must work in reality. This gives less freedom in exploration.

Why would we want to use the methodology of the artificial? I see at least three reasons: We may want to replicate the functionality in a cheaper, more reliable, or more robust way. For example, we may want to build airplanes to carry people and cargo reliably through the air. We may want to understand how a particular functionality can be achieved at all, and thus comprehend a mystery that nature somehow solved. Or, 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 in the sense discussed by Webb (target article). On the contrary, it is a way to gain insight by confronting it with something that is different but still sufficiently similar to make the comparison interesting.

By pretending that the main goal of present-day biorobotics research is to make faithful models of biological systems, we miss out on the opportunities for insight that formal models and artificial systems give us, and we only increase the existing misunderstanding about the methodology of the artificial. Hence, we make it even more difficult for researchers following other approaches to see the point or to learn from our results. There is no doubt that this misunderstanding exists. My current research focuses on language communication, particularly on questions concerning the origins and acquisition of language and meaning (Steels 1998). I try to get to a point where robots construct a shared communication system about the real world perceived through their sensors. I am often asked whether it is really necessary to build these physical robots, given the effort involved, and whether the same insights cannot be gained by computer simulations. Our papers are occasionally rejected by conferences in linguistics or natural language processing on grounds that the artificial languages constructed by these robotic agents are not natural, – as if a sentence constructed for the occasion by a generative linguist or computational language processing without semantics and pragmatics are more natural! In any case, such reactions miss the main points of the methodology, namely: (1) artificial systems are developed in the first place to examine the consequences of certain assumptions, just as formal models are; (2) they require much more realism than formal models and hence provide much deeper insight; but (3) the goal is not to build realistic replicas of natural systems. Their value for understanding nature lies in providing points of comparison with natural systems. Such an activity is as valuable as making faithful realistic models.

ACKNOWLEDGMENTS

Funding from a governmental GOA project to the VUB AI Lab and CNRS-OHLL project to Sony CSL is gratefully acknowledged. Thanks to Marleen Wynants for editorial comments.

Robotic modeling of mobile ball-catching as a tool for understanding biological interceptive behavior

Thomas Sugar[a] and Michael McBeath[b]

[a]*Mechanical and Aerospace Engineering, Arizona State University, Tempe, AZ 85287;* [b]*Department of Psychology, Arizona State University, Tempe, AZ 85287.* **thomas.sugar@asu.edu** **http://www.eas.asu.edu/~tsugar**
michael.mcbeath@asu.edu

Abstract: We support Webb's insights into the potential benefits of using robotic modeling to better understand biological behavior. We defend the major points put forward by Webb by presenting a specific case study in which robotic modeling of mobile ball catching has helped refine and clarify aspects of our understanding of biological interceptive behavior.

In this commentary, we support Webb's insights into the potential benefits and pitfalls of using robotic modeling to better understand biological behavior. We defend the major points put forward by Webb by presenting a specific case study in which robotic modeling of mobile ball catching has helped refine and clarify aspects of our understanding of biological interceptive behavior. We show

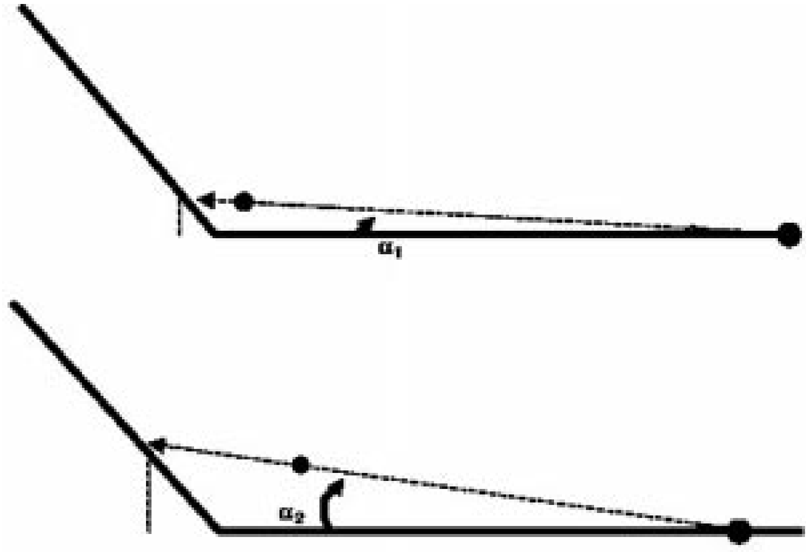

Figure 1 (Sugar & McBeath). The OAC Model. Shown is a side view of temporally successive images of a fielder (large dot) running up to catch an approaching fly ball (small dot). The fielder runs along a path that maintains a constant rate of increase in the vertical tangent angle of the ball, $\frac{d}{dt}\tan(\alpha)$. α is defined as the vertical optical angle of the ball from the perspective moving fielder. The optical change is equivalent to that produced by an imaginary elevator that rises at a constant speed along the tilted back-plane shown to the left.

how a robotic model designed to only use information that is perceptually available allows us to examine whether proposed perceptual heuristics can accurately account for biological interceptive behavior. A tunable robot also allows us to systematically control parameters such as thresholds, reaction lag time, and resolution to see if these accurately account for the patterns of variance observed in biological interception.

Recent perceptual models of how humans catch fly balls have been based largely on ecological principles that presume an ideal observer (and in much of the earlier work, on an ideal environment with no air-resistance or mechanical lag). Examples include the Optical Acceleration Cancellation (OAC) model introduced by Chapman (1968) and later refined by others (e.g., McLeod & Dienes 1996; Michaels & Oudejans 1992) and the Linear Optical Trajectory Model introduced by us (McBeath et al. 1995; 1996; Shaffer & McBeath 2002). Both of these approaches utilize control principles to guide the fielder to the correct destination by locomoting along a path that maintains constant movement of the image of the ball (Marken 1997). In the case of OAC, the vertical optical speed is maintained to be constant, while in the case of LOT, the optical trajectory is maintained to monotonically increase along a straight line (see Figs. 1 and 2). These models also treat the fielder as a point receptor, and assumed that ongoing calculations of optical angles from the perspective of the running fielder are recalibrated independent from the direction he is facing.

When we began trying to simulate biological interceptive behavior with a robotic model, we realized that important behavioral aspects were not clearly elucidated and defined in the previous perceptual models. In particular, we noted that there are two ways that the optical constancy of the trajectory of the ball can be encoded. One approach, which we refer to as "passive," is to keep the eye (or camera) stationary relative to the environmental reference frame, and control the position of the fielder to maintain constant optical ball movement across the stationary retina. A second approach, which we refer to as "active," is to move the eye relative to the environment. Here, the fielder or robot tracks the ball by moving the optics to keep them directed toward it, and then

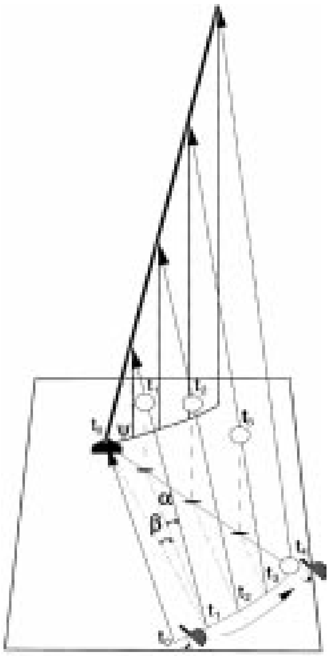

Figure 2 (Sugar & McBeath). The LOT Model. Shown is a bird's eye view of a fielder (hat) running along a path to intercept the ball (white dot) in successive temporal intervals (t_0-t_4). The fielder runs along a path that keeps the image of the ball moving along a monotonically rising Linear Optical Trajectory (LOT). Mathematically, a constant optical projection angle, Ψ, is maintained. The optical change is equivalent to that produced by an imaginary elevator that rises at a constant rate along the tilted line.

the position of the fielder is controlled to maintain constant movement of the eye (or camera) relative to the environment (McBeath et al. 2001; Sugar & McBeath 2001; Suluh et al. 2001).

From the perspective of the previous perceptual models and the assumption of an ideal observer, these two approaches are equivalent. But using the robotic simulation, we were able to demonstrate that the active approach is more robust against noise, and allows for a higher gain, so it is superior in a real world setting, with real mechanical and inertial constraints. This prompted us to confirm that biological fielders also appear to follow the active approach, and gave us insight as to why. Next we plan to systematically alter perceptual thresholds, lag, and gain parameters of the robotic model and verify that the variance in robotic fielding behavior matches that found in biological domains. The robotic platform has thus helped us to articulate and refine our biological model of interceptive behavior.

Our robotic simulation is an example that meets Webb's conditions for a useful scientific endeavor. It is relevant because it applies to real interceptive behavior in humans and animals (Collett & Land 1975; Jablonski 1998; Masters et al. 1985). It has the appropriate level, in that it models at the control parameter level just as do the perceptual interceptive models. It is generalizable in that it applies to all biological navigation that utilizes systematic control of optical variables. It has an appropriate level of abstraction, in that the complexity of the robotic and human perceptual models match well and account well for the variability of the running behavior during the interception task of catching a fly ball. It is structurally accurate, in that the design of the robotic model is a straightforward mapping from proposed biological control heuristics. It results in a good performance match with the types of running paths found with biological interception. And finally, the medium is the same as with biological interception, both are real-world, ball-catching tasks that utilize only information that is realistically available from the perspective of a moving organism.

In short, we designed and tested a robotic model of interceptive behavior and found that it meets Webb's validity test demands, and that it has added insight to our understanding of biological interceptive behavior by demanding a more refined and complete biological model to control an autonomous robot.

Soul searching and heart throbbing for biological modeling

Daniel L. Young[a] and Chi-Sang Poon[b]

[a]*Division of Health Sciences and Technology, Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139;* [b]*Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139.* **dlyoung@mit.edu** **cpoon@mit.edu**
http://cybernet.mit.edu/

Abstract: Biological models are useful not only because they can simulate biological behaviors, but because they may shed light on the inner workings of complex biological structures and functions as deduced by top-down and/or bottom-up reasoning. Beyond the stylistic appeal of specific implementation methods, a model should be appraised according to its ability to bring out the underlying organizing and operating principles – which are truly the model's heart and soul.

In the target article, Webb proposes a set of useful criteria (or "dimensions") for assessing biological models. These criteria present a definitive lexis for "bottom-up" models that aim to simulate overall behavior based on the observed mechanisms at elemental levels. Another approach to modeling, which is alluded to in passing by Webb, is a "top-down" strategy that aims to infer the elemental processes from observed overall behavior. Although less well appreciated, top-down modeling is routinely practiced in hypothesis-driven investigations in the life sciences as well as the physical (Lightman 1992) and social sciences (Bradley & Schaefer 1998; Poon 1994).

Specifically, top-down models (Fig. 1) translate integrated phenomena into hypotheses about sub-level components and their interrelationships, often through the formulation of some overarching law or "theory." Because of their reliance on deductive reasoning, top-down models are necessarily controversial and could be met with skepticism, misunderstanding or outright indifference by others, particularly when the system is complex, the theory is abstract, and hard evidence is lacking. Nonetheless, such controversies may create an impetus for scientific inquiry into potentially revolutionary ideas which, if proven correct, could have far-reaching scientific implications in comparison to research approaches that target a specific reduced structure.

A masterpiece of top-down modeling comes from the legendary Mendelian theory of heredity, which laid the foundation of the genetics discipline. Like any avant-garde, however, the significance of this historic breakthrough was disregarded altogether by Mendel's peers, only to be resurrected with utmost reverence decades later when the cellular and molecular basis of the genetic code began to unfold.

By contrast, bottom-up models (Fig. 1) are grounded in experimental evidence of elemental structures and mechanisms, and their usefulness is determined by how well they match overall behaviors. This process of model building is well served by the conventional reductionist approach, which is an efficient strategy of systematically amassing discrete data. The feasibility of such a strategy is ultimately dictated by the advent of modern technology, namely, more powerful experimental tools afford further miniaturization, modularization, and proliferation of observations thereby laying the groundwork for model building at more elemental levels.

The bottom-up approach mistakenly presumes that the discovery of prime data, rather than modeling per se, is key to the understanding of biological mechanisms and behaviors. This oversimplification is not always true, however. Take, for example, the landmark discovery of the structure of DNA. Here, the reductionist's approach produced an impressive set of elemental data, and yet it was the ingenious modeling effort of Watson and Crick that ultimately fit all the pieces of the puzzle together, making sense of them. In most instances, both top-down and bottom-up approaches may be needed in order to solve a complex problem, and a model is fully validated only when bottom-up meets top-down (Lisberger & Nusbaum 2000; Poon 1992).

At another extreme, the recent mapping of the human genome has proved to mark just the beginning – rather than the end – of an odyssey to explain biological behavior from bottom up. This challenging task is hampered by its intrinsic combinatorial complexity which, in the absence of any unifying theories or models as guiding principles, may prove to be intractable in practice (XIII Oxford Conference, 2001). Could there be the equivalent of such grand deductive theories as law of gravity, relativity and evolution in the bioinformatics of genes and neurons?

That said, then how good are biological models – in particular bio-robotic models? In our view, a model should embody the key (observed or hypothesized) organizing and operating principles that relate top-level to bottom-level mechanisms – and vice versa – based on (observed or predicted) processes and behaviors at either or both ends as well as the intermediate levels. As such, the specific modalities and media for model implementation, such as bio-robotic or computer models, are secondary for the purpose of explaining biological behavior. The crux of a model is whether it captures (definitively or hypothetically) the fundamental working principles behind intricate biological structures and functions, which are truly the model's heart and soul. A model that only simulates the behavior without illuminating the underlying principles is nothing but a lifeless body.

This emphasis on principles over forms is consistent with that

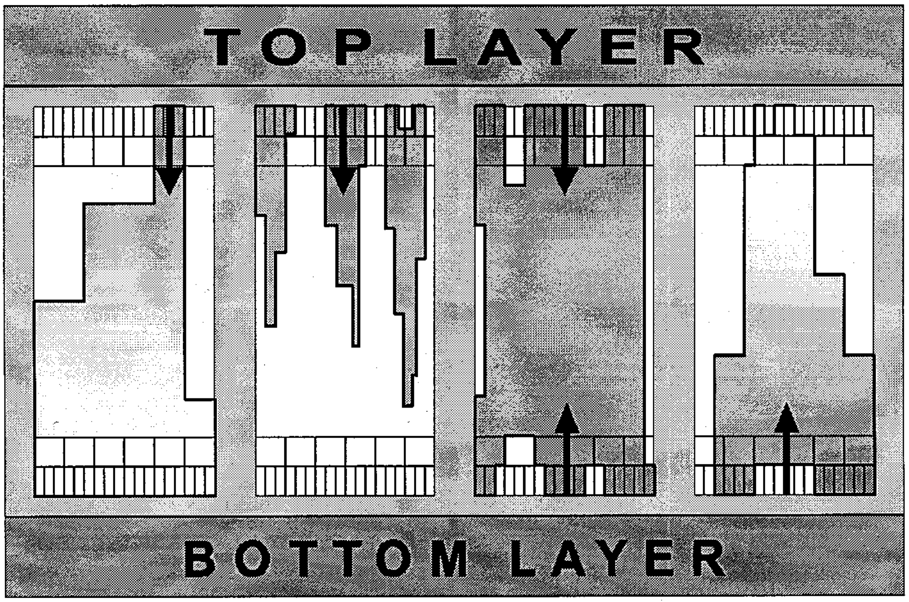

Figure 1 (Young & Poon). Illustration of top-down and bottom-up scientific approaches in four scenarios. Upper layers (in relative scale) represent more integrated phenomena and lower layers the more elemental, with top-down modeling being shown in blue and bottom-up in red as indicated by directions of arrows. The width of a colored region indicates the generality of model predictions and/or breadth of direct evidence at corresponding layer, while fading indicates the lack of direct evidence in support of model predictions at that layer. *Leftmost:* Top-down inference of large-scale fundamental phenomena by deductive generalization of specific experimental observations at top. *Left-of-center:* Early stage of reductionist approach in which a large-scale problem is divided into a set of reduced problems, each being tackled incrementally and recursively by successive applications of top-down data-driven hypothesis formulation and bottom-up hypothesis-driven experimentation on a small scale. *Right-of-center:* Final stage of reductionist approach, when a large-scale problem is solved with accumulation of knowledge through the process of divide-and-conquer. *Rightmost:* Bottom-up solution of a large-scale problem beginning from prime data at bottom and then working upward – a strategy that is made possible by modern technology but may be limited by the problem's complexity.

envisioned by the noted mathematician N. Wiener in his coining of the term *cybernetics* to describe the "entire field of control and communication theory, whether in the machine or in the animal" (Wiener 1948). In particular, an artificial machine, such as a biorobot, should make a good model of biological behavior provided it brings to life the underlying working principles, whether isomorphically or not.

Such widespread muddle about biological modeling in the face of its longstanding acceptance as a basic scientific tool, calls for soul searching. It is perhaps worthwhile to recall that, as with any genre of modeling, the makings of a heartthrob lie not only in the mere physique, but in the inner beauty of the heart and the soul.

ACKNOWLEDGMENTS
Supported by a NIH individual predoctoral fellowship award 1F31MH12697 (DLY) and by NIH grants HL52925 and HL60064 (CSP).

Author's Response

Robots can be (good) models

Barbara Webb
Department of Psychology, Centre for Computational and Cognitive Neuroscience, University of Stirling, Stirling FK9 4LA, Scotland, United Kingdom. **b.b.webb@stir.ac.uk**
www.stir.ac.uk/psychology/Staff/bhw1/

Abstract: It appears there is general support amongst the commentaries for the potential usefulness of biorobots as models, with some caveats. These include the issue that not all areas of biology have been addressed by this methodology (and perhaps some cannot be?); and that other methodologies may sometimes be more useful. Which dimensions of biorobotic (or other models) are considered important varies with the goals of the investigator. These goals are also an essential part of the "modelling relationship."

R1. Introduction

To date, the interaction of biology and robotics has been principally regarded as an approach to robotics, that is, the use of biological inspiration to address engineering problems. In the target article I hoped to establish that there was now a substantial body of "biorobotic" work that could be regarded as an approach to biology, that is, the use of engineering technology to address biological problems. It seemed evident that this work should be set in the context of other models in biology, but in many ways, setting out the grounds for comparison between models proved to be the larger endeavour. This resulted in a review that tried to tackle the notion of modelling in general, as well as robot models in particular. In doing so it perhaps tried to do too much, leaving some confusion over the target article's intentions, or the claims being made. In response to the commentaries, I will here try to clarify several such points. In some cases, this may remove disagreements; in other cases it may at least sharpen the focus on the grounds of disagreement. I will start with the more specific issues regarding robots and move on to the more general considerations of modelling.

R2. Some robots have proved to be good models of some biology

It would seem that the majority of commentators agree that biorobots have something to offer as models of biology. Some are relatively enthusiastic in support (**Banquet et al.; Chang; Delcomyn; De Lillo; Effken & Shaw; Franceschetti; Jaax; MacIver; Metta & Sandini; Midford; Reeke; Sugar & McBeath**), some agree, but with caveats (**Hokland & Verijken; Miller & Arcediano; Pickering; Selverston; Steels; Marin & Oullier**). A few cast doubt on whether the approach is really worthwhile (**Neumann et al.**), or whether it has yet demonstrated its usefulness in biology (**Niebur et al.**), or whether robot models can sufficiently represent certain complex biological issues (**Balasubramaniam & Feldman; Belzung & Chevalley; Krause; Pepperberg**). The remaining commentaries are essentially noncommittal with regard to biorobots, pursuing other issues raised in the article, such as the nature of models and modelling (**Damper; Franklin; Giere; Killeen; Kötter; Lohmann; Schank; Scheutz; Young & Poon**).

Among the supporters of the general approach, some limitations and weaknesses of current work in biorobotics are raised. **Delcomyn** takes up the point that models need to be aimed at critical biological questions, and properly assessed, if they are to be considered relevant and recognised by biologists. **Effken & Shaw** suggest that the problems in evaluating the match of robot and animal behaviour might be helped by adopting some measurement techniques from ecological psychology. **Neumann et al.**, in a critique that will be discussed further below, note that simply being a real world device does not mean that the robot's physical interaction with the world is comparable to the animal: it is still important to consider the accuracy of the representation of noise, physics, and so on. **Selverston** also points out the substantial differences between organic and man-made devices, such as muscles and standard robot actuators. Selverston also argues that more is needed to demonstrate the behavioural match and that working with biologists is a good route towards this. I agree with all these points.

Nevertheless, the "abundance of examples" (**Midford**) discussed in the article make it hard to accept **Niebur et al.**'s suggestion that "the flow of information between biology and robotics is at present almost entirely one-directional" and that there are few if any cases in which robotic models have led to genuine insight into biology. Within the other commentaries are further examples, such as **Sugar & McBeath**'s description of how, using a robot to model ball-catching, they discovered issues not covered in previous mathematical models, and found one approach advantageous under real world constraints, prompting experiments to confirm if this approach was used by biological fielders. I hope this goes some way towards satisfying **Niebur et al.**'s demand for "non-trivial successful predictions" from work in biorobotics.

R2.1. *Not all biology is currently well-modelled by robotics*

A number of the commentaries express reservations regarding biorobotics on the basis that the robot work described in the article fails to include aspects of behaviour that they consider important. Thus, **Pepperberg** mentions "exaptation, observational learning, and complexity"; **Marin & Oullier:** problems of human development and learning; **Belzung & Chevalley:** the range of interacting systems underlying emotional expression; **Krause** wonders, "of what use are biorobots to comparative psychologists studying complex cognitive traits"; **Pickering** suggests that "to be plausible, biorobots will need to build themselves" that is, be autopoietic systems; and **Hokland & Vereijken** feel that adaptivity (Hebbian plasticity) is essential.

Not all these issues are applicable to all biological systems. It is hardly a failure for robots aimed at explaining cricket phonotaxis or ant polarised-light navigation if they do not demonstrate intelligent observational learning, complex cognition, or believable human emotion. Nor was it intended to claim that biorobots will ultimately explain all biology: of course we "should be extremely cautious in assuming we can understand all neural mechanisms with physical models" (**Selverston**). Moreover, it seems obvious that biorobots (no more than other kinds of models) can make little headway in areas where we really as yet have no idea of the biological mechanisms underlying the behaviour. As discussed in the article, biorobots should be deployed at a level which is relevant to current biological understanding. With these considerations in mind, what can be said about the potential for a contribution from biorobotics in the areas raised by the commentators?

R2.1.1. Higher cognition? **Effken & Shaw** rightly note that the biorobotic approach has a certain affinity with ecological psychology, with its stress on the interaction of the organism and environment as an essential determinant of the behaviour, and on the organism as an active system. Such views are used by some as arguments that the biorobotic approach will have an important role in explaining more complex, cognitive behaviours. **Metta & Sandini** argue the necessity for understanding cognition in the context of action; **Pickering** cites Piaget on the importance of active behaviour, and concludes that "cognition is biological action"; and **Banquet et al.** present a robot system that uses grounding in the real world as a "lever" to simplification in complex tasks (including a form of imitation learning). However, it is fair to question whether such approaches will "scale-up," as not all subscribe to this view of how cognition should be understood. The future success or failure of biorobots based on such principles may help answer this question.

R2.1.2. Adaptivity? The issue of "adaptivity" also reflects a difference of views within biology, this time regarding the importance of adaptive processes for understanding neural systems. For example, can we look at an adult mammalian retina, and try to "reverse engineer" it as a more or less hardwired system; or must the fact that it evolved, developed, and may in future be shaped "adaptively" take precedence in any and every analysis of the function? Much of what we currently understand best about neural circuits has come from focus on the "hardwired" aspects, though there is no doubt that in the future issues of growth and change need to be tackled. In any case, as many examples of biorobots do include mechanisms of neural plasticity, this is clearly not a "in principle" problem with the approach. This is not to say that "simple conditioning" will suffice (as **Marin & Oullier** and **Pepperberg** seem to assume is claimed in the article) – biorobots will need to model more sophisticated learning mechanisms to emulate biology. But there seems no a priori reason why they cannot do so.

R2.1.3. Evolution? The problem of exaptation, or the evolutionary history of animal mechanisms is mentioned by several commentators (**Pepperberg; Krause; Killeen; Miller & Arcediano; Belzung & Chevalley**). It is true that being built by evolution is an extremely important constraint on animal mechanisms, and also that a robot built by design to perform the same task, not having to take the same development route, may well have a different mechanism. But does exaptation prevent us from understanding the current functioning of a biological mechanism? If not, then we can attempt to build a robot model that is accurate to the biology even if that is not the approach that "pure" engineering would suggest. There is also an entire subfield of robotics that tries to use evolutionary principles as part of the design process (Nolfi & Floreano 2000). Brooks (1986) has argued for an "exaptation" approach to "behaviour based" robotics, by using a "subsumption" architecture in which layers of increasingly complex behavioural controllers are built on top of simpler ones. **Midford**'s commentary makes another interesting suggestion: that claims about evolutionary homology in animals might be tested by trying to convert a successful robot model of the behaviour of one animal into a model of the other.

R2.1.4. Life itself? Finally, I would note that the endeavour to model living systems is not necessarily an attempt to create living systems. If biorobots fall short of true autopoietic autonomy (**Pickering**), this does not necessarily make them failures as models (no more than the inability to "duplicate living processes with nonliving materials" [**Selverston**] or the worry that "robots . . . cannot have an emotional feeling" [**Belzung & Chevalley**]). Indeed, a self-creating robot – fascinating as such a thing would be – might be less useful as a model because we may understand no more about how it works than the original biological system. On the other hand, Maturana and Varela's (1987) views on living systems are certainly a motivation for undertaking the environment-oriented modelling that biorobots provide, although I have found (contra **Pickering**'s suggestion) that their theory falls somewhat short of the specificity needed to produce working robot systems.

R2.2. Not all robots are good biological models

It should be apparent from my discussion of criteria for including only some systems as "biorobots" (sect. 1.4) that I do not intend to suggest that *all* robots make good biological models, or that all of engineering has application to biology. So, to discuss, as do **Balasubramaniam & Feldman**, how conventional robots, airplanes, chess-playing computers, and thermostats make bad biological models is simply beside the point. If, to take their main example, force-control models for motor control are structurally inaccurate and conflict with "common observations," and better models are available, then all I am suggesting is that it might be informative to test the better models in robot implementations. I am not trying to defend as good biology, engineering solutions that contradict basic biological data.

Two issues have to be kept in mind here. One is discussed particularly by **Chang**: not all robots currently presented as "biorobots," "bio-mimicks," or "bio-inspired" are seriously intended as biological models. However, the labels make it understandably difficult for the biologist to judge what is being claimed, and they may well feel disappointed in the field when they encounter "bio"-robots that really have little connection to biology. I can only plead that this is not sufficient grounds to condemn the whole approach.

The second issue is that a "good" biorobot model, in the sense of one that *is* productive for further understanding in biology, may nevertheless fall a long way short of "simulation of the whole behaviour of the system" (**Belzung & Chevalley**). **Miller & Arcediano** argue that the primary purpose of models is to incite investigators to perform experiments. Inadequate models can often serve this purpose by prompting detailed critique, counter-modelling, and further investigations aimed at demonstrating their inadequacy. In this sense they can still be "good" models even if clearly not sufficient to explain the behaviour of interest. This is one reason why, in cases where we have no better model, an engineering solution to a task that closely resembles a biological problem may be useful to consider. Moreover, it is not always true that "from an engineering point of view . . . biological relevance is not important" (**Neumann et al.**), as illustrated by **Jaax**'s example of biorobotics in prosthesis design, where "design goals are closely congruous to those of biological modelling" and "striving for structural accuracy minimises the likelihood of inadvertently omitting a key behaviour" required for a match in capability.

R2.3. Undetermination

A more general criticism raised in **Pepperberg; Krause; Balasubramaniam & Feldman;** and **Miller & Arcediano** is that the article doesn't take seriously enough the problem of underdetermination. That is, having a robot behave like an animal is no guarantee that the animal works the same way. **Krause** raises the example of a hypothetical robot built to imitate the kinds of eye-gaze following and pointing behaviours taken as evidence for "theory of mind" in humans and primates.[1] He suggests the robot would not help resolve the controversy over whether such behaviour in chimpanzees really means they have a "theory of mind."

On the contrary, it would be an excellent way to resolve this controversy. The current problem is that we cannot know, independently of observing its behaviour, if the chimp really has a "theory of mind." If we knew the robot had no "theory of mind" module or mechanism, but could still behave like the chimp, then it would be clearly a mistake to attribute a "theory of mind" to the chimp purely on the basis of this behaviour. On the other hand, if we found it impossible to generate such behaviour without coming up with a "theory of mind" mechanism for the robot, then we would feel more convinced that the chimp's behaviour was good evidence for such capability. In other words, a robot can be a powerful way to explore the connection between behaviour and causal mechanisms, and hence evaluate arguments that assume causal similarity from analogical behaviour.

De Lillo offers a similar example, in the context of animal search behaviour, of how, by using a robot, we can test whether relatively simple controllers can actually explain some apparently complex behaviours (particularly when in interaction with complex spatial structures in the environment) before assuming more complex cognitive capabilities are indicated by the behaviour. **Pepperberg** correctly asserts that "an existence proof by robotic simulation that shows that an animal does not, for example, need a cognitive map does not mean that such a map is not used." But it does suggest that biologists should provide better behavioural evidence before building complex theories based on, or trying to discover the neural correlates of, such a map. As **Killeen** concludes, robots can help sharpen Ockham's razor.

R3. Biorobotic modelling is not the only valid approach

To argue that biorobots make good models is not to argue that other kinds of models are not also valid, nor would it be surprising if "hardware models of biological systems will never become a method for everyone" (**Niebur et al.**). A variety of techniques and approaches is always healthy in any science. Biorobots are not going to replace computer simulations, as they are complementary tools. Nor is the fact (alluded to by Neibur et al.) that some researchers return to simulation after building a robot an indication that they learnt nothing by building the robot, any more than a return to direct experimentation indicates that time spent modelling was wasted.

R3.1. Simulations may sometimes be better models

Neumann et al. make a more concerted argument that biorobots "have no fundamental advantage" over computer simulations for biology, and that computer simulations can be more accurate, more tractable, and more useful tools. This conclusion is to some extent a result of basing their discussion on the particular systems they are investigating. Virtual reality may offer "photorealistic" environments for visual stimuli, but modelling turbulent chemical plumes or even sound propagation in an environment of varied surfaces is still extremely difficult. Although some simulations may include all the environmental details of the real world, the simple fact is that the majority of simulations do not. Rather, they include what the modeller thinks to be important, that is, they tend to be biased towards the hypotheses to be tested. For example, **Delcomyn** describes how insect walking simulations generally have "the types and strengths of sensory signals that the researchers believe will be generated at various times during the stepping cycle" rather than the "sensor feedback [depending] on actual movements of and loads of the legs" derived from legged-robot implementations. Despite excellent efforts at simulating motion dynamics, gravity, friction, and noise, engineers still find physical models useful to confirm their conclusions. **MacIver** provides an example where a "passive walker" turned out to be stable when simulations predicted it would not, and also mentions important insights in fluid phenomena for animal propulsion discovered through a physical modelling.

The issue is not one of restricting ourselves only to robotics or computer simulations, but using either or both (or going back to experimentation) depending on the particular questions, state of knowledge, and state of hardware and software technology we can bring to bear. In this respect, it is important to note that computer software, mathematics, and language are also "substrates" that can limit or influence the models we build (**Killeen; Miller & Arcediano**).

R3.2. Robots may be used in ways other than as models

Steels suggests a division in scientific activity between constructing the kind of biorobots focused on in my article, formal modelling, and "building artificial systems." My initial reaction is to see these as different approaches within the broad continuum of "model space," rather than to agree that we have a "truly alternative type of scientific activity" in the latter case. Thus, in the "methodology of the artificial," he describes models that are good in match, physical in medium, require more detail than formal models, but are not necessarily accurate in structure.

One problem here is the word "realistic." **Steels** seems to use it at least four different ways within his commentary. Is the robot a "faithful realistic" replica (detail)? Does it have realistic assumptions or mimic biological solutions (accuracy)? Does it have to operate under realistic constraints (medium)? Does it represent a real target system (relevance)? Insofar as my article attempts to restrict what counts as "biorobotics," it is only relevance (the biology) and medium (the robotics) that are required. As I tried to point out in the article, I do not believe "faithful" models are always essential (e.g., in sect. 3.5: "It is possible for a model to address 'real' biological questions without utilising accurate mechanisms").

However, I think the key disagreement underlying **Steels**' discussion is my grouping together of the two questions I list under "relevance": "Is the biological target system clearly identified? Does the model generate hypotheses for biology?" And this relates to his contention that "comparing is not the same as mimicking or modelling." A model that is "irrelevant" in the sense of having no clear target is arguably not a model at all, because it is not intended to represent any (real) thing. However, as **Steels** notes, such a system might still potentially generate hypotheses for biology, through a process of comparison. It seems the current "misunderstanding of the methodology of the artificial" may be another symptom of the "model muddle": because the methodology can resemble modelling but is not (always) intended *as* modelling. In fact, it does not seem to me that the "artificiality" is key here; we might also generate hypotheses by comparing two biological systems without wanting to call one a model of the other.[2]

My defence, in this case, is that the scope of the article was only to assess those robots that *are* meant to be models of biology, rather than assess all possible uses of robots in advancing biological understanding. This defence is also pertinent to **Midford**'s commentary. He suggests several other "physical models" that could have been included in my discussion: a mechanical bee, a die-cast lizard replica used to study thermoregulation, and a mount of a hawk used to study prey reactions to predators. Presumably, he is only suggesting the first is actually a robot. It is debatable whether even this (preprogrammed) device could be said to "have sensors and actuators and require an autonomous control system that enables [it] to successfully carry out various tasks in a complex, dynamic world" (target article, sect. 1; this definition in the introduction was assumed when later I narrowed down "robotic" as meaning *real* sensors, actuators, and worlds). However, even it if it is a robot, it is not so obviously a model. The robot was being used to *stimulate,* not simulate, the mechanisms of dance recognition in bees. The same applies to the hawk – the target system to be explained is the behaviour of the juncos, and the hawk is an experimental tool to test hypotheses about this behaviour, rather than a direct representation of the hypothetical mechanisms.[3]

The lizard, on the other hand, is a good example of a (non-robot) physical model, and indeed illustrates some similar advantages of physical modelling as those discussed in the article. I am willing to accept the suggestion that "biorobots" are "merely an evolutionary advance in existing methods for studying behaviour" (**Midford**) – it is then a question of definition how much "evolutionary advance" suffices to classify something as a 'new species" of methodology.

R4. Determining dimensions

R4.1. The suggested dimensions are not definitive

Several commentators seem to have taken my intention in describing the different dimensions as an attempt to specify "acid tests . . . any particular model should . . . pass at least a subset of these if claims of usefulness are to be substantiated" (**Damper**) or "constraints" that must be "satisfied" (**Balasubramaniam & Feldman**) or "criteria for assessment" (**Young & Poon**). **Sugar & McBeath** seem to be implying the same when they discuss meeting "Webb's criteria for a useful scientific endeavour" or "validity test demands," but in practice they use the dimensions rather more as I intended, that is, to spell out and justify the nature of their model. The dimensions are meant to be a way for researchers to describe their modelling approach more clearly, rather than a set of tests that define a "good" model. Indeed, **Kötter** suggests I have "begged the question" in the paper's title by failing to specify what position on the dimensions makes a model good; **Reeke** feels that I have "quite rightly refrained from value judgements" with regard to the dimensions.

I agree with **Giere** that what makes a model "good" is dependent on scientific goals, or as I said in my target article conclusion, "Different approaches to modelling will reflect differing views about the processes being modelled, and the nature of the explanations required." What I hoped to make clear in the article was how "the justification of the biorobotic approach is grounded in a particular perspective on the issues that need to be addressed," that is, to discuss some of the goals with respect to which robots make good models. Along the way I hoped to illustrate that many apparent disagreements over the "correct" position on the dimensions are in fact disagreements over goals.

This diversity can be clearly seen in the different opinions expressed within the commentaries as to which are the most important dimensions. Thus, **Delcomyn** and **Kötter** emphasise the relevance – answering real and critical questions in biology – as the key issue for biorobotics. **Selverston** and **Hokland & Vereijken** are concerned about the level – for Hokland & Vereijken this must be no higher than a neural network, and Selverston is concerned by the "inability [of biomimetic robots] to mimic biochemical processes that are at the heart of physiological processes." By contrast, **Chang** accepts as "biomimicry" the "close resemblance of a robot to a biological organism at the anatomical, physiological, functional or behavioural level." **Schank** focuses on generality and the importance of finding general mechanisms, or "core ideas" through modelling different systems. **Reeke**, **Lohmann** and **Kötter** discuss the problem of getting the abstraction level right: **Kötter** raises the notion of the "minimal model," which might seem to be the ideal point in this dimension, that is, "the simplest model that still adequately represents a certain hypothetical mechanism." **Young & Poon** consider accuracy to be the "heart and soul" of modelling, while medium is relatively unimportant, and **Neumann et al.** and **Balasubramaniam & Feldman** also seem to use accuracy as their main criteria for assessing models. **Hokland & Vereijken** argue that the match should be "indistinguishable" at the level of limb movements; while **Miller & Arcediano**, **Pepperberg**, and **Krause** stress that "merely" matching behaviour proves nothing. Finally, the medium is seen as the really defining feature of this modelling approach by **Steels** (who does not consider accuracy as necessary), by **Midford** (who sees a natural grouping with other physical models), by **Metta & Sandini**, and by **De Lillo**, who endorses the "medium argument" as providing the right kind of "complete" structural accuracy constraints.

R4.1. The description of dimensions could evolve

Schank is quite correct to characterise my framework as *not* a formal analysis, and as likely to co-evolve with modelling. Indeed, the framework went through several modifications in different drafts of the paper, and I am pleased to see several further modifications suggested by the commentaries (though surprised not to see more). **Franceschetti** somewhat misunderstands my Figure 2 as representing a plane as opposed to a seven-dimensional "hyperspace." He then suggests an initially appealing picture in which the opposite ends of the dimensions from "identity" converge to describe a formal or "abstract, general, and symbolic" model. However, this is effectively to reintroduce some of the assumptions I intended to challenge – for example, that being symbolic was necessarily related to being abstract; that abstraction would automatically confer greater generality and so on. **Franceschetti** also suggests the dimensions need "extending" to cover non-material models, and **Young & Poon** suggest that I provide a lexis for bottom-up models but not top-down models. However, I think that the general, abstract, symbolic models they discuss do fall within the space I have delineated, albeit at the opposite end of the dimensions from most of the models I discuss.

R4.1.1. Integrativeness. Several commentators suggest that "integrativeness" would be a useful additional dimension (**Schank, Miller & Arcediano, Kötter, Belzung & Chevalley, Banquet et al.**). This seems plausible to me. It is not directly correlated with any of the other dimensions. It can refer both to between and within level attempts to combine information. It may in certain situations be preferred to detail or to part-by-part accuracy. It may still be specific to a particular system, described from different points of view, or be a way to aim at generalising across different systems. It can be attempted in any medium, and can aim for different degrees of relevance or match. It is often a consideration in building or assessing models. It is a good way to characterise the idea of "complete" models that I discuss in section R4.5.

R4.1.2. Usefulness? Jaax also suggests "usefulness" as an additional category. In fact, in earlier drafts of the target article I included a dimension of "utility," meant to encompass various kinds of potential usefulness for models: for advancing biological understanding (although this overlaps with relevance); for applications and technology as Jaax is suggesting (which perhaps overlaps with generality, that is, other systems the model can be used for); or for other purposes such as biological control, communication of ideas or education. It was not included in the final scheme because it seems obvious that a "good" model should be useful, and that few scientists set out to build useless models. A possible distinction could be drawn, however, between the intended purpose of the model and how useful it actually turns out to be.

R5. Do we know what models are?

Rather to my surprise, only two commentaries offered alternative accounts of the notion of "models" and "modelling." **Belzung & Chevalley** refer to the idea of a model presented by Hertz. It is important first to note that, insofar as they argue that "models" are different from "simulations," they are using "simulation" in the sense of "resemblance." This is not the same as the idea of "correspondence" that is use to characterise "simulation models" in the article. In fact, Hertz (1894/1956) in his Introduction describes how we endeavour to draw inferences by forming "images" ("innere Scheinbilder") or symbolic descriptions of external objects, such that the consequences of the images match the consequences observed in nature. This seems to be quite consistent with my discussion of "hypothetical models." Hertz also draws a distinction between such theoretical images and the actual representation of them in some notation, which parallels my discussion of representing hypotheses in simulations, using a particular technology. However, it is true that in his later discussion of dynamical models (Hertz 1894/1956, Book II, p. 418), he suggests the relationship of mental images to things is "precisely the same" as the relationship of bijective mapping of co-ordinates between two physical systems (referring in this case specifically to the kinetic behaviour of connected systems of masses) – that is, suggesting it is an isomorphic relationship. Although it is not obvious that **Belzung & Chevalley** would agree with this (they mention "*strict* correlation" in laws of evolution between system and model [my emphasis], but they also suggest that novels are "excellent models of the behaviour of living things"), it leads us conveniently to the issues raised by **Scheutz.**

Scheutz points out that I have incorrectly used a term from "model theory" in mathematical logic, where "partial isomorphism" is defined as a relationship between systems that have isomorphic substructures. I meant to refer the illogical notion of "incomplete" isomorphisms, for example, cases like "Anna Karenina" and "nineteenth century Russia" where though some correspondences can be found, a complete bijective mapping preserving all structural properties is not (even in principal) possible to demonstrate for the structure of (or any well-defined substructure within) the two systems.

More importantly, he uses the notion of "partial isomorphism" to argue against my contention that a general and explicit notion of "model," that can explain the practice of modelling in science, is not available. He starts by making a distinction between systems and the description of these systems in a (formal) language that can support deduction. So, to establish a modelling relationship between two systems, Scheutz suggests that we perform the following steps:

1. Describe the parts of the target system at some level.
2. Describe the parts of the model system in a common language.
3. Establish a bijective mapping that preserves structure between subsets of the two descriptions.

This allows us to perform deduction on the (description of the) model system to derive conclusions we can apply to the (description of the) target system, provided we restrict ourselves to the relevant subsets.

It seems to me that (1) is equivalent to the step of stating a theory or hypothesis about a system, and thus the mapping described in (3) is the representation of a hypothesis in a model. That is, it is the relationship described by the arrows between "hypothetical mechanism" and "simulation" (or "source") in my Figure 1.[4] In the target article I described this process as "modelling aims to make the process of producing predictions from hypotheses more effective by enlisting the aid of an *analogical mechanism*" (sect. 2.2). What **Scheutz** is proposing then (if I have followed correctly) is that we can replace "*analogical mechanism*" in this phrase with "*partially isomorphic system,*" and thereby use logical model theory to validate our reasoning. Note that what we thus "deduce" with the model are not conclusions about the target system itself but rather conclusions about our description of the target system. Unless we already have reason to believe our hypothesis is a correct description, we still need to test these conclusions against the actual behaviour of the system to validate the theory.

So far so good, and I will even agree that in this sense the formal notion of isomorphism "captures our intuitions" about models. For example, I might say my cricket robot is a model of the cricket because the neural simulation in the robot's control program can be mapped to my theory of the neural connections underlying the relevant behaviour in the animal. The problem is that I would be hard put to *prove* this was a "bijective mapping" that fully preserves structural properties. Not having done so, I cannot claim to have established a "partial isomorphism" in the model-theoretic sense. Is it therefore the case that my "cricket" robot (and all the other robots described in the article) are not actually models? If not, they would at least be in good company, for few other so-called "models" in biology fit this criterion. To take just one famous example – Crick and Watson's physical model of DNA used shaped bits of metal that (as **Giere** points out) were rather approximate in their representation of hypothesised molecular structures. In Watson's (1968) description of their work with this model he notes "It was all too easy to fudge a successful series of atomic contacts so that, while each looked almost acceptable, the whole collection was energetically impossible." In other words, their model was an imperfect representation that could result in faulty deductions.

But if we cannot demonstrate a partial isomorphism, what, asks **Scheutz**, "warrants the claim that a robot system is a model of an animal with respect to X"? I had no intention of being covert in what I meant by "a robot system is a model of an animal." I meant the robot *is being used* as a model of the animal. That is, I *intentionally* conflated the "modelling relationship" with "modelling" *qua* practice. To be explicit, I consider there to be no "independent ontological" question, any more than, when a poet uses one thing as a metaphor for another, there is an "independent ontological" question as to whether "such a metaphorical relationship" actually exists between the two things.[5] That is, going beyond **Giere**, not only is it the case that you "cannot eliminate the purposes of scientists from the evaluation of any model," it is my belief that you cannot eliminate their purposes from the question of whether the system is a model at all.

NOTES

1. In fact, there is an actual robot built to investigate this issue, discussed in Scassellati 2001.

2. Some might argue that the operation of comparison, requiring that two systems be put in some kind of correspondence so that we can observe the similarities and differences, is a sort of *de facto* modelling (e.g., **Krause** describes cladistics as the use of real animals as "models" of other animals). However, I do not think this is generally appropriate.

3. There is a sense in which an experiment can (also) be described as a kind of model, in that we are setting up a system to generate the behaviour that follows from our hypotheses (e.g., using a particular group of juncos to represent all prey, and the artificial hawk behaviour to represent the hypothesised causes for the prey's response). **Miller** mentions lab experiments as representations of real ecological niches. **Killeen** discusses how an experimental rat may not only be a model of a human, but also a model of "itself in a within-subjects' design." **Schank** notes that Levin's discussion of modelling suggested an extension of the notion of models to cover experiments. Nevertheless, (as for comparison in Note 2) it seems reasonable to draw a general distinction between experiments on a target system and simulation of a target system.

4. Or, as **Damper** quotes Moor (1978), it is the idea that a model can "embody a theory"; – note that he also recognises there can be a problem in determining exactly what theory a particular model embodies.

5. This is not to deny that the choice of a metaphor or model, and its subsequent productivity or value, will depend in some way on the richness of the structural mappings we can make between their domains and their targets. The point is we can construct and use untidy and incomplete mappings without ever being close to proving the existance of an isomorphism (partial or otherwise). There has been much interesting work recently on the theory of metaphor (Lakoff 1993) and the nature of analogical reasoning (Gentner 1983) that seems more applicable than logical model theory in understanding how we use models in scientific practice.

References

Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively.

Achinstein, P. (1968) *Concepts of science: A philosophical analysis.* Johns Hopkins Press. [aBW]

Ackoff, R. L. (1962) *Scientific method.* Wiley. [aBW]

Ahlborn, B., Chapman, S., Stafford, R., Blake, R. W. & Harper, D. G. (1997) Experimental simulation of the thrust phases of fast-start swimming of fish. *Journal of Experimental Biology* 200(17):2301–12. [MAM]

Aloimonos, J. (1990) Purposive and qualitative active vision. Paper presented at the Proceedings of the Tenth Annual International Conference on Pattern Recognition, Atlantic City, June 1990. [GM]

Amit, D. J. (1989) *Modeling brain function: The world of attractor neural networks.* Cambridge University Press. [aBW]

Andry, P., Gaussier, P., Moga, S., Banquet, J.-P. & Nadel, J. (2001) Learning and communication via imitation: An autonomous robot perspective. *IEEE Journal: Man, Systems and Cybernetics.* (in press). [RB]

Arbib, M. & Liaw, J. (1995) Sensorimotor transformations in the worlds of frogs and robots. *Artificial Intelligence* 72:53–79. [aBW]

Arkin, R. C. (1998) *Behaviour-based robotics.* MIT Press. [CD, aBW]

Asatryan, D. G. & Feldman, A. G. (1965) Functional tuning of the nervous system with control of movement or maintenance of a steady posture: I. Mechanographic analysis of the work of the limb on execution of a postural task. *Biophysics* 10:925–35. [RB]

Ashby, W. R. (1952) *Design for a brain.* Chapman and Hall. [aBW]

Ayers, J. (1995) A reactive ambulatory robot architecture for operation in current and surge. In: *Proceedings of the Autonomous Vehicles in Mine Countermeasures Symposium, pp. 15–31.* Naval Postgraduate School. http://www.dac.neu.edu/msc/nps95mcm%20manuscript.html [aBW]

Ayers, J., Zavracky, P., Mcgruer, N., Massa, D., Vorus, V., Mukherjee, R. & Currie, S. (1998) A modular behavioural-based architecture for biomimetic autonomous underwater robots. *Autonomous Vehicles in Mine Countermeasures Symposium* .http://www.dac.neu.edu/msc/biomimeticrobots98.html [aBW]

Baars, B. J. (1988) *A cognitive theory of consciousness.* Cambridge University Press. [SF]

(1997) *In the theater of consciousness.* Oxford University Press. [SF]

Bajcsy, R. K. (1985) Active perception vs. passive perception. Paper presented at the Third IEEE Workshop on Computer Vision: Representation and Control, Bellaire, Michigan. [GM]

Bandyopadhyay, P. R., Castano, J. M., Nedderman, W. H. & Donnelly, M. J. (2000) Experimental simulation of fish-inspired unsteady vortex dynamics on a rigid cylinder. *Journal of Fluids Engineering – Transactions of the ASME* 122(2):219–38. [MAM]

Banquet, J.-P., Gaussier, P., Contreras-Vidal, J., Gissler, A., Burnod, Y. & Long, D. L. (1998) A neural network model of memory, amnesia, and cortico-hippocampal interactions. In: *Fundamentals of neural networking,* ed. R. W. Parks, D. S. Levine & D. L. Long. MIT Press. [J-PB]

Banquet, J.-P., Gaussier, P., Dreher, J. C., Joulain, C., Revel, A. & Günther, W. (1997) Space-time, order and hierarchy in fronto-hippocampal system: A neural basis of personality. In: *Cognitive science perspectives on personality and emotion,* ed. G. Matthews. Elsevier Science. [J-PB]

Banquet, J.-P., Gaussier, P., Revel, A. & Moga, S. (2001) Sequence learning and timing in hippocampus, prefrontal cortex, and accumbens. In: *International Joint Conference on Neural Networks,* pp. 1053–58, ed. K. Marko & P. Werbos. Omnipress. [J-PB]

Barrett, D. S., Triantafyllou, M. S., Yue, D. K. P., Grosenbaugh, M. A. & Wolfgang, M. J. (1999) Drag reduction in fish-like locomotion. *Journal of Fluid Mechanics* 392:183–212. [MAM]

Barsalou, L. W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22:577–609. [SF]

Barto, A. G. (1991) Learning and incremental dynamic programming. *Behavioral and Brain Sciences* 14:94–95. [aBW]

Bates, E., Elman, J. L., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1998) Innateness and emergentism. In: *Companion to cognitive science,* ed. W. Bechtel & G. Graham. Blackwell. [JP]

Beckers, R., Holland, O. E. & Deneubourg, J. L. (1996) From local actions to global tasks: Stigmergy and collective robotics. In: *Proceedings of the Fourth International Conference on Artificial Life, Boston, July 1994,* ed. R. Brooks & P. Maes. MIT Press. *Artificial Life IV:*181–89. [MAM, aBW]

Beer, R. D. (1990) *Intelligence as adaptive behaviour: An experiment in computational neuroethology.* Academic Press. [aBW]

Beer, R. D., Chiel, H. J., Quinn, R. D. & Ritzmann, R. E. (1998) Biorobotic approaches to the study of motor systems. *Current Opinion in Neurobiology* 8:777–82. [FD, aBW]

Beer, R., Quinn, R., Chiel, H. & Ritzmann, R. (1997) Biologically inspired approaches to robotics. *Communications of the ACM* 40:31–38. [aBW]

Bekey, G. (1996) Biologically inspired control of autonomous robots. *Robotics and Autonomous Systems* 18:21–31. [aBW]

Bell, J. (1991) *Searching behaviour.* Chapman and Hall. [CD]

Berkemeier, M., & Desai, K. (1996) Design of a robot leg with elastic energy storage, comparison to biology, and preliminary experimental results. In: *Proceedings of IEEE International Conference on Robotics and Automation, Minneapolis, MN, 22–28 April 1996,* pp. 213–18. IEEE. [aBW]

Bertin, R. J. V. & van de Grind, W. A. (1996) The influence of light-dark adaptation and lateral inhibition on phototaxic foraging: A hypothetical animal study. *Adaptive Behaviour* 5:141–67. [aBW]

Bhalla, U. S., Bilitch, D. H. & Bower, J. M. (1992) Rallpacks: A set of benchmarks for neural simulators. *Trends in Neurosciences* 15:453–58. [aBW]

Binnard, M. B. (1995) Design of a small pneumatic walking robot. ftp://ftp.ai.mit.edu/pub/users/binn/SMthesis.zip. [aBW]

Birch, J. M. & Dickinson, M. H. (2001) Spanwise flow and the attachment of the leading-edge vortex on insect wings. *Nature* 412(6848):729–33. [MAM]

Black, M. (1962) *Models and metaphors.* Cornell University Press. [aBW]

Blanchard, M., Verschure, P. F. M. J. & Rind, F. C. (1999) Using a mobile robot to study locust collision avoidance responses. *International Journal of Neural Systems* 9:405–10. [aBW]

Bloom, L. (2000) The intentionality model: How to learn a word, any word. In: *Becoming a word learner: A debate on lexical acquisition,* ed. R. Michnick Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. L. Woodward, N. Akhtar, M. Tomasello & G. Hollich. Oxford University Press. [IMP]

Boahen, K. A. (2000) Retinomorphic vision chips for ocular implants. *Annals of Biomedical Engineering* 228(S1):S – 115. [KNJ]

Bower, J. M. (1992) Modeling the nervous system. *Trends in Neurosciences* 15:411–12. [aBW]

Bower, J. M. & Koch, C. (1992) Experimentalists and modelers: Can we all just get along? *Trends in Neurosciences* 15:458–61. [aBW]

Bradley, W. J. & Schaefer, K. C. (1998) *The uses and misuses of data and models. The mathematization of the human sciences.* Sage. [DLY]

Braithwaite, R. B. (1960) Models in the empirical sciences. In: *Logic methodology and philosophy of science,* ed. E. Nagel, P. Suppes & A. Tarski. Stanford University Press. [aBW]

Breland, K. & Breland, M. (1961) The misbehavior of organisms. *American Psychologist* 16:661–64. [IMP]

Brooks, R. A. (1985) A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation* 2(1):14–23. [EN]

(1986) Achieving artificial intelligence through building robots. *Artificial Intelligence Memo 89.* MIT Artificial Intelligence Lab. [arBW]

(1991) *Intelligence without reason.* A. I. Memo No. 1293. MIT Artificial Intelligence Laboratory. [CC]

(2001) The relationship between matter and life. *Nature* 409(1):409–11. [GM]

Brooks, R. A. & Stein, L. A. (1993) Building brains for bodies. *Artificial Intelligence. Memo No. 1439.* MIT Artificial Intelligence Lab. [aBW]

Brooks, R. J. & Tobias, A. M. (1996) Choosing the best model: Level of detail, complexity and model performance. *Mathematical Computer Modelling* 24:1–14. [aBW]

Bulloch, A. G. M. & Syed, N. I. (1992) Reconstruction of neuronal networks in culture. *Trends in Neurosciences* 15:422–27. [aBW]

Bullock, S. (1997) An exploration of signalling behaviour by both analytic and simulation means for both discrete and continuous models. *Fourth European Conference on Artificial Life.* MIT Press. [aBW]

Bülthoff, H. H. & van Veen, H. A. H. C. (2001) Vision and action in virtual environments: Modern psychophysics in spatial cognition research. In: *Vision and attention,* ed. M. Jenkin & L. Harris. Springer-Verlag. [TRN]

Bunge, M. (1973) *Method, model and matter.* D. Reidal. [aBW]

Burgess, N., Donnett, J. G., Jeffery, K. J. & O'Keefe, J. (1997) Robotic and neuronal simulation of the hippocampus and rat navigation. *Philosophical Transactions of the Royal Society of London B* 352:1535–43. [aBW]

Burgess, N., Donnett, J. G. & O'Keefe, J. (1998) Using a mobile robot to test a model of the rat hippocampus. *Connection Science* 10:291–300. [aBW]

Burgess, N., Jackson, A., Hartley, T. & O'Keefe, J. (2000) Predictions derived from modelling the hippocampal role in navigation. *Biological Cybernetics* 83:301–12. [aBW]

Campbell, M. S. (1996) "An enjoyable game": How HAL plays chess. In: *HAL's legacy: 2001's computer as dream and reality,* ed. D. G. Stork. MIT Press. [IMP]

Carnap, R. (1966) *Philosophical foundations of physics.* Basic Books. [aBW]

Carson, N. R. (2001) *Physiology of behavior.* Allyn and Bacon. [JH]

Cartwright, B. & Collett, T. (1983) Landmark learning in bees. *Journal of Comparative Physiology A* 151:521–43. [aBW]

Cartwright, N. (1983) *How the laws of physics lie.* Clarendon Press. [aBW]

Caswell, H. (1988) Theories and models in ecology – a different perspective. *Ecological Modelling* 43:33–44. [aBW]

Chan, K. H. & Tidwell, P. M. (1993) The reality of Artificial Life: Can computer simulations become realizations? In: *Third International Conference on Artificial Life.* (submitted). [RK, aBW]

Chang, C. & Gaudiano, P. (1998) Application of biological learning theories to mobile robot avoidance and approach behaviours. *Journal of Complex Systems* 1:79–114. [aBW]

(2000) Editorial: Biomimetic robotics. *Robotics and Autonomous Systems* 30(1–2): Elsevier Science. [CC]

Chao, Y. R. (1960) Models in linguistics and models in general. In: *Logic, methodology and philosophy of science,* ed. E. Nagel, P. Suppes & A. Tarski. Stanford University Press. [aBW]

Chapman, S. (1968) Catching a baseball. *American Journal of Physics* 36:836–70. [TS]

Chapman, T. & Webb, B. (1999) A neuromorphic hair sensor model of wind-mediated escape in the cricket. *International Journal of Neural Systems* 9:397–403. [aBW]

Chevalley, C. & Belzung, C. (2002) Emotional expression as the result of stochastic interactions: A process crucial for cognition. *Behavioral Procedures.* (in press). [CB]

Chiel, H. & Beer, R. (1997) The brain has a body: Adaptive behaviour emerges from interactions of nervous system, body and environment. *Trends in Neurosciences* 20:553–57. [aBW]

Churchland, P. S., Koch, C. & Sejnowski, T. J. (1990) What is computational neuroscience? In: *Computational neuroscience,* ed. E. L. Schwartz. MIT Press. [GNR, aBW]

Churchland, P. S. & Sejnowski, T. J. (1988) Perspectives on cognitive neuroscience. *Science* 242:741–45. [aBW]

(1992) *The computational brain.* MIT Press. [FD]

Clark, A. (1997) *Being there: Putting brain, body and world back together again.* MIT Press. [JP]

(1998) Embodied, situated and distributed cognition. In: *Companion to cognitive science,* ed. W. Bechtel & G. Graham. Blackwell. [JP]

(2001) *Mindware: An introduction to the philosophy of cognitive science.* Oxford University Press. [JP]

Clark, J. J. (1998) Spatial attention and saccadic camera motion. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, 16–20 May 1998,* IEEE. [aBW]

Cliff, D. (1991) Computational neuroethology: A provisional manifesto. In: *From animals to animats,* ed. J.-A. Meyer & S. W. Wilson. MIT Press. [aBW]

Colby, K. M. (1981) Modeling a paranoid mind. *Behavioural and Brain Sciences* 4:515–60. [aBW]

Coleman, M. J., Garcia, M., Mombaur, K. & Ruina, A. (2001) Prediction of stable walking for a toy that cannot stand. *Physical Review E* 64:022901. [MAM]

Coleman, M. J. & Ruina, A. (1998) An uncontrolled walking toy that cannot stand still. *Physical Review Letters* 80(16):3658–61. [MAM]

Collett, T. S. & Land, M. F. (1975) Visual control of flight information in the hoverfly, *Syritta Pipiens L. Journal of Comparative Physiology* 99:1–66. [TS]

Collin, C. & Woodburn, R. (1998) Neuromorphism or pragmatism? A formal approach. In: *Neuromorphic systems: Engineering silicon from neurobiology,* ed. L. Smith & A. Hamilton. World Scientific. [aBW]

Conant, R. C. & Ashby, W. R. (1991) Every good regulator of a system must be a model of that system. In: *Facets of system science,* ed. G. J. Klir. (Reprint). Plenum Press. [aBW]

Connell, J. H. (1990) *Minimalist mobile robotics: A colony style architecture for an artificial creature.* Academic Press. [aBW]

Cramer, A. E. & Gallistel, C. R. (1997) Vervet monkeys as travelling salesman. *Nature* 38:464. [CD]

Crick, F. (1989) The recent excitement about neural networks. *Nature* 337:129–32.

Croon, M. A. & van de Vijver, F. J. R. (1994) Introduction. In: *Viability of mathematical models in the social and behavioural science,* ed. M. A. Croon & F. J. R. van de Vijver. Swets and Zeitlinger. [aBW]

Cruse, H. (2001) Building robots with a complex motor system to understand cognition. In: *Biorobotics: Methods and applications,* ed. B. Webb & T. Consi. AAAI Press/MIT Press. [aBW]

Cruse, H., Kindermann, T., Schumm, M., Dean, J. & Schmitz, J. (1998) Walknet – a biologically inspired network to control six-legged walking. *Neural Networks* 11:1435–47. [aBW]

Damper, R. I., French, R. L. B. & Scutt, T. W. (2000) ARBIB: An autonomous robot based on inspirations from biology. *Robotics and Autonomous Systems* 31:247–74. [LM]

Dario, P., Laschi, C. & Guglielmelli, E. (1998) Sensors and actuators for humanoid robots. *Advanced Robotics* 11:567–84. [aBW]

Deakin, M. A. B. (1990) Modelling biological systems. In:*Dynamics of complex interconnected biological systems,* ed. T. L. Vincent, A. I. Mees & L. S. Jennings. Birkhauser. [aBW]

Dean, J. (1991) A model of leg coordination in the stick insect, *Carausius morosus.* II. Description of the kinematic model and simulation of normal step patterns. *Biological Cybernetics* 64:403–11. [FD]

(1998) Animats and what they can tell us. *Trends in Cognitive Sciences* 2(2):60–67. [RID, aBW]

Delcomyn, F. & Nelson, M. E. (2000) Architectures for a biomimetic hexapod robot. *Robotics and Autonomous Systems* 30:5–15. [FD]

Delcomyn, F., Nelson, M. & Cocatre Zilgien, J. (1996) Sense organs of insect legs and the selection of sensors for agile walking robots. *International Journal of Robotics Research* 15:113–27. [aBW]

De Lillo, C., Aversano, M., Tuci, E. & Visalberghi, E. (1998) Spatial constraints and regulatory functions in monkey's (*Cebus apella*) search. *Journal of Comparative Psychology* 112:353–62. [CD]

De Lillo, C., Floreano, D. & Antinucci, F. (2001) Transitive choices by a simple, fully connected, backpropagation neural network: Implications for the comparative study of transitive inference. *Animal Cognition* 4:61–68. [CD]

De Lillo, C., Visalberghi, E. & Aversano, M. (1997) The organisation of exhaustive searches in a "patchy" space by capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology* 111:82–90. [CD]

Dennett, D. C. (1979) *Brainstorms: Philosophical essays on mind and psychology.* Harvester Press. [aBW]

Descartes, R. (1662/1911) *Trait de l'homme,* trans. E. S. Haldane & G. R. T. Ross. Cambridge University Press. (Cambridge English edition, 1911). [aBW]

Dickinson, M. H., Lehmann, F. O. & Sane, S. P. (1999) Wing rotation and the aerodynamic basis of insect flight. *Science* 284(5422):1954–60. [MAM]

Doucet, P. & Sloep, P. B. (1992) *Mathematical modelling in the life sciences.* Ellis Horwood. [aBW]

Dreyfus, H. L. (1979) A framework for misrepresenting knowledge. In: *Philosophical perspectives in Artificial Intelligence,* ed. M. Ringle. Harvester Press. [aBW]

Dror, I. E. & Gallogly, D. P. (1999) Computational analyses in cognitive neuroscience: In defense of biological implausibility. *Psychonomic Bulletin and Review* 6(2):173–82. [RID, aBW]

Durbin, R. (1989) On the correspondence between network models and the nervous system. In: *The computing neuron,* ed. R. Durbin, C. Miall & G. Mitchison. Addison-Wesley. [aBW]

Ebbinghaus, H. D., Flum, J. & Thomas, W. (1980) *Mathematical logic. Undergraduates Texts in Mathematics.* Springer-Verlag. [MS]

Edelman, G. M., Reeke, G. N., Gall, W. E., Tononi, G., Williams, D. & Sporns, O. (1992) Synthetic neural modeling applied to a real world artifact. *Proceedings of the National Academy of Sciences USA* 89:7267–71. [aBW]

Effken, J. A., Kim, N.-G. & Shaw, R. E. (1997) Making the constraints visible: Testing the ecological approach to interface design. *Ergonomics* 40(1):1–27. [JAE]

Effken, J. A., Loeb, R., Johnson, K., Johnson, S. & Reyna, V. (2001) Using cognitive work analysis to design clinical displays. In: *Proceedings of MedInfo – 2001* (pp. 27–31), ed. V. L. Patel, R. Rogers & R. Haux. IOS Press. [JAE]

Effken, J. A. & Shaw, R. E. (1992) Ecological perspectives on the new artificial intelligence. *Ecological Psychology* 4:247–70. [JAE]

Elliott, T., Howarth, C. I. & Shadbolt, N. R. (1996) Neural competition and statistical mechanics. *Proceedings of the Royal Society of London B* 263:601–606. [aBW]

Erkman, I., Erkman, A. M., Takkaya, A. E. & Pasinlioglu, T. (1999) Haptic perception of shape and hollowness of deformable objects using the Anthrobot-III robot hand. *Journal of Robotic Systems* 16:9–24. [aBW]

Espenschied, K., Quinn, R., Beer, R. & Chiel, H. (1996) Biologically based distributed control and local reflexes improve rough terrain locomotion in a hexapod robot. *Robotics and Autonomous Systems* 18:59–64. [aBW]

Estes, W. K. (1975) Some targets for mathematical psychology. *Journal of Mathematical Psychology* 12:263–82. [aBW]

Etienne, A. S. (1998) Mammalian navigation, neural models and robotics. *Connection Science* 10:271–89. [aBW]

Etienne-Cummings, R., Van der Spiegel, P. & Mueller, P. (1998) Neuromorphic and digital hybrid systems. In: *Neuromorphic systems: Engineering silicon from neurobiology,* ed. L. Smith & A. Hamilton. World Scientific. [aBW]

Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (2000) Visuomotor neurons: Ambiguity of the discharge of "motor" perception? *International Journal of Psychophysiology* 35(2–3):165–77. [GM]

Fagg, A., Sitkoff, N., Barto, A. & Houk, J. (1997) Cerebellar learning for control of a two-link arm in muscle space. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Alberquerque, NM, April 1997,* pp. 2638–44. IEEE Press. [aBW]

Faloutsos, P., van de Panne, M. & Terzopoulos, D. (2001) Composable controllers for physics-based character animation. In: *Proceedings of the Association for Computing Machinery, ACM SIGGRAPH 2001, pp. 251–60.* ACM Press. [TRN]

Fearing, R. S. (1999) Micromechanical flying insect [Web Page]. http://robotics.eecs.berkeley.edu/~ronf/mfi.html. [aBW]

Feibleman, J. K. (1954) Theory of integrative levels. *British Journal of the Philosophy of Science* 5:59–66. [aBW]

Feldman, A. G. & Levin, M. F. (1995) The origin and use of positional frames of reference in motor control. *Behavioral and Brain Sciences* 18:723–806. [RB]

Feldman, A. G. & Orlovsky, G. N. (1972) The influence of different descending systems on the tonic stretch reflex in the cat. *Experimental Neurology* 37:481–94. [RB]

Ferrell, C. (1995) Comparison of three insect-inspired locomotion controllers. *Robotics and Autonomous Systems* 16:135–59. [aBW]

Flynn, A. M. & Brooks, R. A. (1989) Battling reality. *Artificial Intelligence Memo 1148.* MIT Artificial Intelligence Lab. [MAM, aBW]

Fodor, J. A. (1968) *Psychological explanation.* Random House. [aBW]

Fodor, J. A. & Pylyshyn, Z. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* 28:3–71. [aBW]

Fontana, W. (1992) Algorithmic chemistry. In: *Artificial Life II,* ed. C. G. Langton, C. Taylor, J. D. Farmer & S. Rasmussen. Addison Wesley. [DRF]

Forrester, J. W. (1972) *Principles of systems.* Wright-Allen Press. [aBW]

Fowler, A. C. (1997) *Mathematical models in the applied sciences.* Cambridge University Press. [aBW]

Franceschini, N. (1996) Engineering applications of small brains. *FED Journal* 7:38–52. [aBW]

Franceschini, N., Pichon, J. M. & Blanes, C. (1992) From insect vision to robot vision. *Philosophical Transactions of the Royal Society of London B* 337:283–94. [JAE, aBW]

Franklin, S. (1997) Autonomous agents as embodied AI. *Cybernetics and Systems* 28:499–520. [SF]

(2000a) Deliberation and voluntary action in "conscious" software agents. *Neural Network World* 10:505–21. [SF]

(2000b) Modeling consciousness and cognition in software agents. In: *Proceedings of the Third International Conference on Cognitive Modeling, Groeningen, The Netherlands, March 2000,* pp. 100–109, ed. N. Taatgen. Universal Press. [SF]

(2001) Automating human information agents. In: *Practical applications of intelligent agents,* ed. Z. Chen & L. C. Jain. Springer-Verlag. [SF]

Franklin, S. & Graesser, A. (1999) A software agent model of consciousness. *Consciousness and Cognition* 8:285–305. [SF]

(2001) Modeling cognition with software agents. In: *CogSci2001. Proceedings of the 23rd Annual Conference of the Cognitive Science Society (August 1–4, 2001, Series Eds. 301–306),* ed. J. D. Moore & K. Stenning. Erlbaum. [SF]

Franklin, S., Keleman, A. & McCauley, L. (1998) IDA: A cognitive agent architecture. In: *IEEE Conference on Systems, Man and Cybernetics (Series Eds. 2646–51).* IEEE Press. [SF]

Franz, M. O., Schölkopf, B., Mallot, H. A. & Bülthoff, H. H. (1998) Learning view graphs for robot navigation. *Autonomous Robots* 5:111–25. [TRN]

Friedman, M. (1953) *Essays in positive economics.* University of Chicago Press. [LS]

Frijda, N. J. (1967) Problems of computer simulation. *Behavioural Science* 12:59–67. [aBW]

Gaussier, P., Banquet, J. P., Joulain, C., A. Revel, A. & Zrehen, S. (1997) Validation of a hippocampal model on a mobile robot. In: *Proceedings of the International Conference on vision, recognition, action: Neural models of mind and machine, Boston University, May 28–31, 1997.* Boston University Press. [aBW]

Gaussier, P., Joulain, C., Banquet, J. P., Leprêtre, S. & Revel, A. (1999) The visual homing problem: An example of robotics/biology cross fertilization. *Robotics and Autonomous Systems* 30:155–80. [J-PB]

Gaussier, P., Moga, S., Banquet, J.-P. & Quoy, M. (1998) From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. *Applied Artificial Intelligence* 12:701–27. [J-PB]

Gaussier, P., Revel, A., Banquet, J.-P. & Babeau, V. (2001) From view cells and place cells to cognitive maps: Processing stages of the hippocampus system. *Biological Cybernetics.* (in press). [J-PB]

Gelenbe, E., Schmajuk, N., Staddon, J. & Reif, J. (1997) Autonomous search by robots and animals: A survey. *Robotics and Autonomous Systems* 22:23–34. [CD, aBW]

Gentner, D. (1983) Structure mapping: A theoretical framework for analogy. *Cognitive Science* 7:155–70. [rBW]

Gibson, J. J. (1958) Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology* 49:182–94. [JAE]

(1986) *The ecological approach to visual perception.* Erlbaum. (Original work published 1979). [JAE]

Giere, R. N. (1997) *Understanding scientific reasoning.* Harcourt Brace. [aBW]

Glenberg, A. M. (1997) What memory is for. *Behavioral and Brain Sciences* 20:1–19. [SF]

Gordon, G. (1969) *System simulation.* Prentice Hall. [aBW]

Graham, D. (1977) Simulation of a model for the coordination of leg movement in free walking insects. *Biological Cybernetics* 26:187–98. [FD]

Grasso, F., Consi, T., Mountain, D. & Atema, J. (1996) Locating odor sources in turbulence with a lobster inspired robot. In: *Sixth International Conference on Simulation of Adaptive Behaviour: From animals to animats 4,* ed. P. Maes, M. J. Mataric, J. A. Meyer, J. Pollack & S. W. Wilson. MIT Press. [aBW]

(2000) Biomimetic robot lobster performs chemo-orientation in turbulence using a pair of spatially separated sensors: Progress and challenges . *Robotics and Autonomous Systems* 30:115–31. [IMP, aBW]

Greene, P. (1975) Problems of organization of motor systems. In: *Progress in theoretical biology, vol. 2,* ed. R. Rosen & F. Snell. Academic Press. [RB]

Greenfield, P. M. (1991) Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences* 14:531–95. [IMP]

Grimm, V. (1994) Mathematical models and understanding in ecology. *Ecological Modelling* 75/76:641–51. [aBW]

Haefner, J. W. (1996) *Modeling biological systems.* Chapman and Hall. [aBW]

Halfon, E. (1983) Is there a best model structure? I. Modeling the fate of a toxic substance in a lake. *Ecological Modelling* 20:135–52. [aBW]

Hallam, J. (1998) Can we mix robotics and biology? *Workshop on Biomorphic Robots.* Victoria BC, Canada. [aBW]

Hannaford, B., Winters, J., Chou, C.-P. & Marbot, P.-H. (1995) The anthroform biorobotic arm: A system for the study of spinal circuits. *Annals of Biomedical Engineering* 23(4):399–408. [KNJ, aBW]

Harnad, S. (1989) Minds, machines and Searle. *Journal of Experimental and Theoretical Artificial Intelligence* 1:5–25. [aBW]

Harre, R. (1970a) *The method of science.* Wykeham. [aBW]

(1970b) *The principles of scientific thinking.* Chicago University Press. [aBW]

Harrison, R. R. & Koch, C. (1999) A robust analog VLSI motion sensor based on the visual system of the fly. *Autonomous Robotics* 7:211–24. [aBW]

Hassell, M. P. (1978) *Arthropod predator-prey systems.* Princeton University Press. [JCS]

Hasslacher, B. & Tilden, M. W. (1995) Living machines. In: *Robotics and autonomous systems: The biology and technology of intelligent autonomous agents,* ed. L. Steels. Elsevier. [EN]

Hertz, H. (1894/1956) *Der Prinzipien der Mechanik.* English translation: *The principles of mechanics, Books I and II,* trans. D. E. Jones & J. T. Walley. Prefaced by H. von Helmholts. Dover Books. (Original publication 1894, Dover reprint 1956). [CB, rBW]

Hesse, M. B. (1966) *Models and analogies in science.* Notre Dame University Press. [aBW]

Hewes, G. W. (1973) Primate communication and the gestural origin of language. *Current Anthropology* 33:65–84. [IMP]

Heyes, C. M. (1998) Theory of mind in nonhuman primates. *Behavioral and Brain Sciences* 21:101–48. [MAK]

Heyes, C. M. & Galef, B. G., Jr., eds. (1996) *Social learning in animals: The roots of culture.* Academic Press. [IMP]

Hirai, K., Hirose, M., Haikawa, Y. & Takenaka, T. (1998) The development of the Honda humanoid robot. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, May 1998,* pp. 1321–26. IEEE Press. [aBW]

Hirose, S. (1993) *Biologically inspired robotics.* Oxford University Press. [aBW]

Hirose, S., Raibert, M. & Pack, R. T. (1996) General systems group report (International Workshop on Biorobotics). *Robotics and Autonomous Systems* 18:95–99. [aBW]

Hodgkin, A. L. & Huxley, A. F. (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology* 117:500–44. [RID]

Holland, O. & Melhuish, C. (1999) Stigmergy, self-organization and sorting in collective robotics. *Artificial Life* 5:173–202. [aBW]

Honma, A. (1996) Construction robot for three-dimensional shapes based on the nesting behavior of paper wasps. *Seimitsu Kogaku Kaishi/Journal of the Japan Society for Precision Engineering* 62: 805–09. [aBW]

Hoos, I. R. (1981) Methodology, methods and models. In: *Systems thinking: 2,* ed. F. E. Emery. Penguin. [aBW]

Hopkins, J. C. & Leipold, R. J. (1996) On the dangers of adjusting parameter values of mechanism-based mathematical models. *Journal of Theoretical Biology* 183:417–27. [aBW]

Horiuchi, T. (1997) An auditory localization and co-ordinate transform chip. In: *Advances in neural information processing systems 8.* MIT Press. [aBW]

Horiuchi, T. K. & Koch, C. (1999) Analog VLSI-based modeling of the primate oculomotor system. *Neural Computation* 11:243–64. [aBW]

Hoshino, H., Mura, F., Morii, H., Suematsu, K. & Shimoyama, I. (1998) A small sized panoramic scanning visual sensor inspired by the fly's compound eye. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, May 1998,* pp. 1641–46. IEEE Press. [aBW]

Huang, J., Ohnishi, N. & Sugie, N. (1995) A biomimetic system for localization and separation of multiple sound sources. *IEEE Transactions on Instrumentation and Measurement* 44:733–38. [aBW]

Huber, S. A. & Bulthoff, H. H. (1998) Simulation and robot implementation of visual orientation behaviour of flies. In: *From animals to animats 5,* ed. R. Pfeifer, B. Blumberg, J. A. Meyer & S. W. Wilson. MIT Press. [aBW]

Huber, S. A., Franz, M. O. & Bülthoff, H. H. (1999) On robots and flies: Modeling the visual orientation behavior of flies. *Robotics and Autonomous Systems* 29:227–42. [TRN]

Hughes, R. I. G. (1997) Models and representation. *Philosophy of Science* 64:S325–36. [aBW]

Ilg, W., Berns, K., Jedele, H., Albiez. J., Dillmann, R., Fischer, M., Witte, H., Biltzinger, J., Lehmann, R. & Schilling, N. (1998a) BISAM: From small mammals to a four-legged walking machine. In: *From animals to animats 5,* ed. R. Pfeifer, B. Blumberg, J. A. Meyer & S. W. Wilson. MIT Press. [aBW]

Indiveri, G. (1998b) Analog VLSI model of locust DCMD neuron response for computation of object approach. In: *Neuromorphic systems: Engineering silicon from neurobiology,* ed. L. Smith & A. Hamilton. World Scientific. [aBW]

Ishida, H., Kobayashi, A., Nakamoto, T. & Moriisumi, T. (1999) Three dimensional odor compass. *IEEE Transactions on Robotics and Automation* 15:251–57. [aBW]

Jaax, K. & Hannaford, B. (2002) A biorobotic structural model of the mammalian muscle spindle primary afferent response. *Annals of Biomedical Engineering* 30(1):84–96. [KNJ]

Jablonski, P. (1998) A rare predator exploits prey escape behavior: The role of tail fanning and plumage in foraging of the painted redstart (*Myioborus pictus*). *Behavioral Ecology* 10:7–14. [TS]

James, W. (1890) *The principles of psychology.* Harvard University Press. [SF]

Jeannerod, M. (1997) *The cognitive neuroscience of action.* Blackwell. [GM]

Kacser, H. (1960) Kinetic models of development and heredity. In: *Models and analogues in biology: Symposia of the Society for Experimental Biology.* Cambridge University Press. [aBW]

Kadar, E. & Shaw, R. (2000) Toward an ecological field theory of perceptual control of information. *Ecological Psychology* 12(2):141–80. [JAE]

Kamil, A. C. (1984) Adaptation and cognition: Knowing what comes naturally. In: *Animal cognition,* ed. H. L. Roitblat, T. G. Bever, & H. S. Terrace. Erlbaum. [IMP]

(1988) A synthetic approach to the study of animal intelligence. In: *Nebraska symposium on motivation: Comparative perspectives in modern psychology, vol. 7,* ed. D. Leger. University of Nebraska Press. [IMP]

Kamil, A. C. & Sargent, T. D. (1981) *Foraging behavior: Ecological, ethological, and psychological approaches.* Garland Press. [IMP]

Kandel, E. R., Schwartz, J. H. & Jessell, T. M. (2000) *Principles of neural science.* McGraw-Hill. [JH]

Kanzaki, R. (1996) Behavioral and neural basis of instinctive behavior in insects: Odor-source searching strategies without memory and learning. *Robotics and Autonomous Systems* 18:33–43. [aBW]

Kaplan, A. (1964) *The conduct of enquiry.* Chandler. [aBW]

Kato, M. & Inaba, T. (1998) Guidance and control of fish robot with apparatus of pectoral fin motion. *Proceedings – IEEE International Conference on Robotics and Automation* 1:446–51. [aBW]

Kelso, J. A. S. (1995) *Dynamic patterns.* MIT Press. [LM]

Killeen, P. R. (2000) Modeling games from the 20th century. *Behavioural Processes* 54:33–52. (Available at author's URL: www.asu.edu/clas/psych/dresearch/blab.html). [PRK]

Kim, J. (1998) *Philosophy of mind.* Westview/Perseus. [RID]

Kintsch, W. (1998) *Comprehension.* Cambridge University Press. [SF]

Klir, J. & Valach, M. (1965) *Cybernetic modelling.* ILIFFE Books. [aBW]

Klute, G. K., Czerniecki, J. M. & Hannaford, B. (2000) Artificial tendons: Biomechanical design properties for prosthetic lower limbs. Paper presented at the World Congress on Medical Physics and Biomedical Engineering, Chicago, IL. [KNJ]

Koch, C. (1999) *Biophysics of computation.* Oxford University Press. [aBW]

Kodjabachian, J. & Meyer, J.-A. (1995) Evolution and development of control architectures in animats. *Robotics and Autonomous Systems* 16:161–82. [LM]

Kofka, K. (1935) *Principles of Gestalt psychology.* Harcourt Brace Jovanovich. [LM]

Kolacinski, M. & Quinn, R. D. (1998) A novel biomimetic actuator system. *Robotics and Autonomous Systems* 25:1–18. [aBW]

Konczak, J., Borutta, M., Topka, H. & Dichgans, J. (1995) Development of goal-directed reaching in infants: Hand trajectory formation and joint force control. *Experimental Brain Research* 106:156–68. [GM]

Kondadadi, R. & Franklin, S. (2001) A framework of deliberative decision making in "conscious" software agents. In: *Proceedings of the Sixth International Symposium on Artificial Life and Robotics (AROB-01), Tokyo, Japan, January 15–17, 2001,* pp. 280–83). Gotanda. [SF]

Kortmann, R. & Hallam, J. (1999) Studying animals through artificial evolution: The cricket case. In: *Proceedings of the Fifth European Conference on Artificial Life,* ed. D. Floreano, J.-D. Nicond & F. Mondada. Springer-Verlag. [aBW]

Kötter, R. (1999) Trends in European computational neuroscience. *Reviews in Neuroscience* 10:175–79. [RK]

Kovacs, I. (2000) Human development of perceptual organization. Vision Research 40:1301–10. [GM]

Krause, M. A. (1997) Comparative perspectives on pointing and joint attention in children and apes. *International Journal of Comparative Psychology* 10:137–57. [MAK]

Kroes, P. (1989) Structural analogies between physical systems. *British Journal for the Philosophy of Science* 40:145–54. [aBW]

Kubow, T. M. & Full, R. J. (1999) The role of the mechanical system in control: A hypothesis of self-stabilisation in the cockroach. *Philosophical Transactions of the Royal Society of London B* 354:849–61. [KNJ, aBW]

Kuc, R. (1997) Biomimetic sonar recognizes objects using binaural information. *Journal of the Acoustical Society of America* 102:689–96. [aBW]

Kumph, J. M. (1998) MIT Robot Pike Project [Web Page]. URL http://www.mit.edu/afs/athena/org/t/towtank/OldFiles/www/pike/index.html [aBW]

Kuwana, Y., Shimoyama, I. & Miura, H. (1995) Steering control of a mobile robot using insect antennae. *IEEE International Conference on Intelligent Robots and Systems* 2:530–35. [IMP, aBW]

Lakoff, G. (1993) The contemporary theory of metaphor. In: *Metaphor and thought, 2nd edition,* ed. A. Ortony. Cambridge University Press. [rBW]

Lamb, J. R. (1987) Computer simulation of biological systems. *Molecular and Cellular Biochemistry* 73:91–98. [aBW]

Lambert, K. & Brittan, G. G. (1992) *An introduction to the philosophy of science, 4th edition.* Ridgview. [aBW]

Lambrinos, D., Maris, M., Kobayashi, H., Labhart, T., Pfeifer, R. & Wehner, R. (1997) An autonomous agent navigating with a polarized light compass. *Adaptive Behaviour* 6:175–206. [aBW]

Lambrinos, D., Moller, R., Labhart, T., Pfeifer, R. & Wehner, R. (2000) A mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems* 30:39–64. [aBW]

Langton, C. G. (1989) *Artificial Life.* Addison-Wesley. [aBW]

Laudan, L. (1981) A confutation of convergent realism. *Philosophy of Science* 48:218–49. [aBW]

Leatherdale, W. H. (1974) *The role of analogy, model and metaphor in science.* North Holland. [aBW]

Leavens, D. A., Hopkins, W. D. & Bard, K. A. (1996) Indexical and referential pointing in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 110:346–53. [MAK]

Lee, D. (1976) A theory of visual control of braking based on information about time-to-collision. *Perception* 5:437–59. [JAE]

Leoni, F., Guerrini, M., Laschi, C., Taddeucci, D., Dario, P. & Staritao, A. (1998) Implementing robotic grasping tasks using a biological approach. *Proceedings of the IEEE International Conference on Robotics and Automation* 3:2274–80. [aBW]

Levin, M. F. & Dimov, M. (1997) Spatial zones for muscle coactivation and the control of postural stability. *Brain Research* 757:43–59. [RB]

Levins, R. (1966) The strategy of model building in population biology. *American Scientist* 54:421–31. [JCS, aBW]

(1993) A response to Orzack and Sober: Formal analysis and the fluidity of science. *Quarterly Review of Biology* 68:547–55. [JCS, aBW]

Lewis, M. A. & Nelson, M. E. (1998) Look before you leap: Peering behaviour for depth perception. In: *From animals to animats 5,* ed. R. Pfeifer, B. Blumberg, J. A. Meyer & S. W. Wilson. MIT Press. [aBW]

Libet, B. (1999) Do we have free will? *Journal of Consciousness Studies* 6:47–57. [SF]

Libet, B., Gleason, C. A., Wright, E. W. & Pearl, D. K. (1983) Time of conscious intention to act in relation to onset of cerebral activity (Readiness-Potential): The unconscious initiation of a freely voluntary act. *Brain* 106:623–42. [SF]

Lieberman, H., ed. (2001) *Programming by example.* Morgan Kaufmann/Academic Press. [IMP]

Lightman, A. (1992) *Great ideas in physics.* McGraw-Hill. [DLY]

Lima, S. L. & Bednekoff, P. A. (1999) Back to the basics of antipredatory vigilance: Can nonvigilant animals detect attack? *Animal Behaviour* 58:537–43. [PEM]

Lisberger, S. G. & Nusbaum, M. P. (2000) Understanding motor circuits: Where bottom-up meets top-down. *Current Opinions in Neurobiology* 10:673–75. [DLY]

Lohmann, S. (2000) Sollbruchstelle: Deep uncertainty and the design of monetary institutions. *International Finance* 3:391–411. [SL]

Lund, H. H., Webb, B. & Hallam, J. (1997) A robot attracted to the cricket species *Gryllus Bimaculatus.* In: *Fourth European Conference on Artificial Life,* ed. P. Husbands & I. Harvey. MIT Press. [aBW]

(1998) Physical and temporal scaling considerations in a robot model of cricket calling song preference. *Artificial Life* 4:95–107. [aBW]

MacIver, M. A., Sharabash, N. M. & Nelson, M. E. (2001) Prey-capture behavior in gymnotid electric fish: Motion analysis and effects of water conductivity. *Journal of Experimental Biology* 204(3):543–57. [MAM]

Maki, D. P. & Thompson, M. (1973 *Mathematical models and applications.* Prentice-Hall. [aBW]

Manna, Z. & Pnueli, A. (1991) On the faithfulness of formal models. *Lecture Notes in Computer Science* 520:28–42. [aBW]

Maris, M. & Mahowald, M. (1998) Neuromorphic sensory motor mobile robot controller with pre-attention mechanism. In: *Neuromorphic systems: Engineering silicon from neurobiology,* ed. L. Smith & A. Hamilton. World Scientific. [aBW]

Marken, R. S. (1997) The dancer and the dance: Methods in the study of living control systems. *Psychological Methods* 2:436–46. [TS]

Marr, D. (1982) *Vision.* W. H. Freeman. [aBW]

Masters, W. M., Moffat, A. J. M. & Simmons, J. A. (1985) Sonar tracking of horizontally moving targets by the big brown bat, *Eptesicus fuscus. Science* 228:1331–33. [TS]

Mataric, M. J. (1998) Behavior-based robotics as a tool for synthesis of artificial behavior and analysis of natural behavior. *Trends in Cognitive Sciences* 2:82–87. [aBW]

Matthew, P. B. C. (1959) The dependence of tension upon extension in the stretch reflex of the soleus muscle in the decerebrated cat. *Journal of Physiology* 147:521–46. [RB]

Maturana, U. & Varela, F. (1987) *The tree of knowledge: The biological roots of human understanding.* Shambala. [JP, rBW]

Maynard Smith, J. (1974) *Models in ecology.* Cambridge University Press. [aBW]

(1987) When learning guides evolution. *Nature* 329:761–62. [JP]

(1988) *Did Darwin get it right?* Penguin. [aBW]

McBeath, M. K., Shaffer, D. M. & Kaiser, M. K. (1995) How baseball outfielders determine where to run to catch fly balls. *Science* 268(5210):569–73. [TS]

(1996) On catching fly balls. *Science* 273(5272):256–59. [TS]

McBeath, M. K., Sugar, T. & Shaffer, D. (2001) Comparison of active versus passive ball catching control algorithms using robotic simulations. *Journal of Vision* 1(3):193.< http://www.journalof vision.org/1/3/193> [TS]

McGeer, T. (1990) Passive walking with knees. In: *Proceedings of the IEEE Conference on Robotics and Automation, Cincinnati, Ohio, May 1990,* pp. 1640–45. IEEE Press. [aBW]

McLeod, P. & Dienes, Z. (1996) Do fielders know where to go to catch the ball or only how to get there? *Journal of Experimental Psychology: Human Perception and Performance* 22:531–43. [TS]

Mead, C. (1989) *Analog VLSI and Neural Systems.* Addison-Wesley. [aBW]

Melhuish, C., Holland, O. & Hoddell, S. (1998) Collective sorting and segregation in robots with minimal sensing. In: *From animals to animats 5,* ed. R. Pfeifer, B. Blumberg, J. A. Meyer, & S. W. Wilson. MIT Press. [aBW]

Metta, G. (2000) Babybot: A study on sensori-motor development. Unpublished doctoral dissertation, University of Genova, Genova, Italy. [GM]

Meyer, J. (1997) From natural to artificial life: Biomimetic mechanisms in animat designs. *Robotics and Autonomous Systems* 22:3–21. [aBW]

Meyer, J.-A. & Wilson, S. W. (1991) *From animals to animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior.* MIT Press. [LM]

Miall, C. (1989) The diversity of neuronal properties. In: *The computing neuron,* ed. R. Durbin, C. Miall & G. Mitchison. Addison-Wesley. [aBW]

Michaels, C. F. & Oudejans, R. D. (1992) The optics and actions of catching fly balls. *Ecological Psychology* 4:199–222. [TS]

Michelson, A., Andersen, B. B., Storm, J., Kirchner, W. H. & Lindauer, M. (1992) How honeybees perceive communication dances, studied by means of a mechanical model. *Behavioral Ecology and Sociobiology* 30:143–50. [PEM]

Miklos, G. L. G. (1993) Molecules and cognition: The latter day lessons of levels, language and lac. *Journal of Neurobiology* 24:842–90. [aBW]

Milinski, M. (1991) Models are just protheses for our brains. *Behavioral and Brain Sciences* 14:101. [aBW]

Mills, N. J. & Getz, W. M. (1996) Modelling the biological control of insect pests: A review of host-parasitoid models. *Ecological Modelling* 92:121–43. [JCS]

Milner, D. A. & Goodale, M. A. (1995) *The visual brain in action, vol. 27.* Oxford University Press. [GM]

Miwa, H., Umetsu, T., Takanishi, A. & Takanobou, H. (2001a) Human-like robot head that has olfactory sensation and facial color expression. *Proceedings of the IEEE International Conference on Robotics and Automation* 1:459–64. [CB]

Miwa, H., Umetsu, T., Takamishi, A. & Takanobou, H. (2001b) Robot personality based on the equations of emotion defined in the 3D mental space. *Proceedings of the IEEE International Conference on Robotics and Automation* 3:2602–607. [CB]

Mojarrad, M. & Shahinpoor, M. (1997) Biomimetic robotic propulsion using polymeric artificial muscles. *Proceedings -IEEE International Conference on Robotics and Automation* 3:2152–57. [aBW]

Molenaar, I. M. (1994) Why do we need statistical models in the social and behavioural sciences. In: *Viability of mathematical models in the social and behavioural science,* ed. M. A. Croon & F. J. R. van de Vijver. Swets and Zeitlinger. [aBW]

Möller, R. (2000) Insect visual homing strategies in a robot with analog processing. *Biological Cybernetics* 83:231–43. [aBW]

Möller, R., Lambrinos, D., Pfeifer, R., Labhart, T. & Wehner, R. (1998) Modeling ant navigation with an autonomous agent. In: *From animals to animats 5,* ed. R. Pfeifer, B. Blumberg, J. A. Meyer & S. W. Wilson. MIT Press. [aBW]

Moor, J. H. (1978) Three myths of computer science. *British Journal for the Philosophy of Science* 29:213–22. [RID]

Morgan, M. S. (1997) The technology of analogical models: Irving Fisher's monetary worlds. *Philosophy of Science* 64:S304–14. [aBW]

Morse, T. M., Ferree, T. C. & Lockery, S. R. (1998) Robust spatial navigation in a robot inspired by chemotaxis in *Caenorrhabditis elegans. Adaptive Behaviour* 6:393–410. [aBW]

Mura, F. & Shimoyama, I. (1998) Visual guidance of a small mobile robot using active, biologically-inspired, eye movements. *Proceedings of the IEEE International Conference on Robotics and Automation* 3:1859–64. [aBW]

Nagel, E. (1961) *The structure of science: Problems in the logic of scientific explanation.* Routledge & Kegan Paul. [aBW] Harcourt, Brace and World/Hackett. (1979 Hackett edition cited by Damper). [RID]

Nelson, G. M. & Quinn, R. D. (1998) Posture control of a cockroach-like robot. *Proceedings of the IEEE International Conference on Robotics and Automation* 1:157–62. [aBW]

Nelson, M. E. & MacIver, M. A. (1999) Prey capture in the weakly electric fish *Apteronotus albifrons:* Sensory acquisition strategies and electrosensory consequences. *Journal of Experimental Biology* 202(10):1195–203. [MAM]

Neumann, T. R. & Bülthoff, H. H. (2001) Insect inspired visual control of translatory flight. In: *ECAL 2001. Lecture Notes in Computer Science 2159: Lecture Notes in Artificial Intelligence,* 627–36, ed. J. Keleman & P. Sosik. Springer-Verlag. [TRN]

Newell, K. M. (1986) Constraints on the development of coordination. In: *Motor development in children: Aspects of coordination and control,* ed. M. G. Wade & H. T. A. Whiting. Martinus Nijhoff. [LM]

Nicholson, A. J. & Bailey, V. A. (1935) The balance of animal populations. Part I. *Proceedings of the Zoological Society of London* 3:551–98. [JCS]

Nolfi, S. & Floreano, D. (2000) *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines.* MIT Press. [rBW]

Nowak, M. A., Plotkin, J. B. & Jansen, V. A. (2000) The evolution of syntactic communication. *Nature* 404:495–98. [LS]

Ogata, T. Shigeki, S. (2000) Emotional communication between humans and the autonomous robot W AMOBEA-2 (Waseda amoeba) which has the emotional model. *JSME International Journal, Series C: Mechanical Systems, Machine Elements and Manufacturing* 43:568–74. [CB]

O'Leary, D. D. M. (1992) Development of connectional diversity and specificity in the mammalian brain by the pruning of collateral projections. *Current Opinion in Neurobiology* 2:70–77. [GM]

Onstad, D. W. (1988) Population-dynamics theory – the roles of analytical, simulation, and supercomputer models. *Ecological Modelling* 43:111–24. [aBW]

O'Reilly, R. C. (1998) Six principles for biologically based computational models of cognition. *Trends in Cognitive Science* 2(11):455–62. [RID]

Oreskes, N., Shrader-Frechette, K. & Belitz, K. (1994) Verification, validation and confirmation of numerical models in the earth sciences. *Science* 263:641–46. [aBW]

Orzack, S. H. & Sober, E. (1993) A critical assessment of Levin's "The strategy of model building in the social sciences" (1966). *Quarterly Review of Biology* 68:533–46. [JCS, aBW]

Palladino, P. (1991) Defining ecology – ecological theories, mathematical-models, and applied biology in the 1960s and 1970s. *Journal of the History of Biology* 24:223–43. [aBW]

Palsson, B. O. & Lee, I. (1993) Model complexity has a significant effect on the numerical value and interpretation of metabolic sensitivity coefficients. *Journal of Theoretical Biology* 161:299–315. [aBW]

Patel, G. N., Holleman, J. H. & DeWeerth, S. P. (1998) Analog VLSI model of intersegmental coordination with nearest neighbour coupling. In: *Advances in neural information processing systems 10,* ed. M. I. Jordan, M. J. Kearns & S. A. Solla. MIT Press. [aBW]

Pattee, H. H. (1989) Simulations, realizations and theories of life. In: *Artificial life,* ed. C. Langton. Addison-Wesley. [aBW]

Peirce, G. J. & Ollanson, J. G. (1987) Eight reasons why optimal foraging theory is a complete waste of time. *Oikos* 49:111–25. [aBW]

Pepperberg, I. M. (1990) Cognition in an African Grey Parrot (*Psittacus erithacus*): Further evidence for comprehension of categories and labels. *Journal of Comparative Psychology* 104:41–52. [IMP]

(2001) Evolution of avian intelligence. In: *The evolution of intelligence,* ed. R. Sternberg & J. Kaufman. Erlbaum. [IMP]

Pepperberg, I. M. & Shive, H. R. (2001) Simultaneous development of vocal and physical object combinations by a Grey Parrot (*Psittacus erithacus*): Bottle caps, lids, and labels. *Journal of Comparative Psychology* 115:376–84. [IMP]

Peremans, H., Walker, A. & Hallam, J. C. T. (1998) 3D object localization with a binaural sonar head – inspirations from biology. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, May 1998,* pp. 2795–2800. IEEE Press. [aBW]

Pfeifer, R. (1996) Building "fungus eaters": Design principles of autonomous agents. In: *From animals to animats 4: Proceedings of the 4th International Conference on Simulation of Adaptive Behaviour* 3–11. MIT Press. [aBW]

Pfeifer, R. & Scheier, C. (1998) Representation in natural and artificial agents: An embodied cognitive science perspective. Paper presented at the Workshop on Natural Organisms, Artificial Organisms, and their Brains, Bielefeld, Germany, March 8–12, 1998. [GM]

Pfeiffer, F., Etze, J. & Weidemann, H.-J. (1995) Six-legged technical walking machine considering biological principles. *Robotics and Autonomous Systems* 14:223–32. [FD, aBW]

Piaget, J. (1979) Behaviour and evolution, trans. D. Nicholson-Smith. Routledge & Kegan Paul. [JP]

Pichon, J.-M., Blanes, C. & Franceschini, N. (1989) Visual guidance of a mobile robot equipped with a network of self-motion sensors. In: *Mobile robots IV,* ed. W. J. Wolfe & W. H. Chun. Society of Photo-optical Instrumentation Engineers. [aBW]

Pinsky, P. F. & Rinzel, J. (1994) Intrinsic and network rhythmogenesis in a reduced Traub model for CA3 neurons. *Journal of Computational Neuroscience* 1:39–60. [RK]

Poon, C.-S. (1992) Introduction: Optimization hypothesis in the control of breathing. In: *Control of breathing and its modeling perspective,* ed. Y. Honda, Y. Miyamoto, K. Konno & J. Widdicombe. Plenum Press. [DLY]

(1994) Quantitative social sciences. *Nature* 368:297–98. [DLY]

Pornsin-Sirirak, N. & Tai, Y.-C. (1999) Microbat [Web Page]. http://secretary.erc.caltech.edu/~tai/research/nick/. [aBW]

Povinelli, D. J., Bering, J. M. & Giambrone, S. (2000) Toward a science of other minds: Escaping the argument by analogy. *Cognitive Science* 24:509–41. [MAK]

Pransky, J. (2001) AIBOT – the no. 1 selling service robot. *Industrial Robot* 28:24–26. [CB]

Pratt, J. & Pratt, G. A. (1998a) Intuitive control of a planar biped walking robot. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, May 1998,* pp. 2014–21. IEEE Press. [aBW]

(1998b) Exploiting natural dynamics in the control of a planar biped walking robot. In: *Proceedings of the Thirty-sixth Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, September 1998.* University of Illinois Press. [aBW]

Prescott, T. J. & Ibbotson, C. (1997) A robot trace-maker: Modelling the fossil evidence of early invertebrate behaviour. *Artificial Life* 3:289–306. [aBW]

Prince, A. & Pinker, S. (1988) Rules and connections in human language. *Trends in Neurosciences* 11:195–202. [GNR]

Putnam, H. (1975) *Philosophical papers, vol. 1.* Cambridge University Press. [aBW]

Pyke, G. H., Pulliam, H. R. & Charnov, E. L. (1977) Optimal foraging: A selective review of theory and tests. *Quarterly Review of Biology* 52:137–54. [IMP]

Quinn, R. D. & Espenscheid, K. S. (1993) Control of a hexapod robot using a biologically inspired neural network. In: *Biological neural networks in invertebrate neuroethology and robotics,* ed. R. D. Beer, R. E. Ritzmann & T. McKenna. Academic Press. [MAM, aBW]

Quinn, R. D. & Ritzmann, R. E. (1998) Construction of a hexapod robot with cockroach kinematics benefits both robotics and biology. *Connection Science* 10:239–54. [FD, aBW]

Raaijmakers, J. G. W. (1994) Mathematical models in memory research. In: *Viability of mathematical models in the social and behavioural science,* ed. M. A. Croon & F. J. R. van de Vijver. Swets and Zeitlinger. [aBW]

Raibert, M. H. (1986) *Legged robots that balance.* MIT Press. [aBW]

Redhead, M. (1980) Models in physics. *British Journal for the Philosophy of Science* 31:145–63. [aBW]

Reeke, G. N., Jr. & Edelman, G. M. (1988) Real brains and artificial intelligence.

Daedalus. Proceedings of the American Academy of Arts and Sciences 117:143–73. [GNR]

Reeke, G. N., Jr., Finkel, L. H., Sporns, O. & Edelman, G. M. (1990) Synthetic neural modeling: A multilevel approach to the analysis of brain complexity. In: *Signal and sense: Local and global order in perceptual maps,* ed. G. M. Edelman, W. E. Gall & W. M. Cowan. Wiley. [GNR]

Reeke, G. N. & Sporns, O. (1993). Behaviourally based modeling and computational approaches to neuroscience. *Annual Review of Neuroscience* 16:597–623. [aBW]

Rescorla, R. A. & Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement. In: *Classical conditioning II: Current research and theory,* ed. A. H. Black & W. F. Prokasy. Appleton, Century, Crofts. [RRM]

Rexstad, E. & Innis, G. S. (1985). Model simplification: Three applications. *Ecological Modelling* 27:1–13. [aBW]

Ridgel, A. L., Frazier, S. F., DiCaprio, R. A. & Zill, S. N. (2000) Encoding of forces by cockroach tibial campaniform sensilla: Implications in dynamic control of posture and locomotion. *Journal of Comparative Physiology A* 186:359–74. [FD]

Ringle, M. (1979). Philosophy and A.I. In: *Philosophical perspectives in artificial intelligence,* ed. M. Ringle. Harvester Press. [aBW]

Ritzmann, R. E., Quinn, R. D., Watson, J. T. & Zill, S. N. (2000) Insect walking and biorobotics: A relationship with mutual benefits. *Bioscience* 50:23–33. [FD]

Rizzo, J. F., IIIrd, Wyatt, J., Humayun, M., de Juan, E., Liu, W., Chow, A., Eckmiller, R., Zrenner, E., Yagi, T. & Abrahms, G. (2001) Retinal prosthesis: An encouraging first decade with major challenges ahead. *Ophthalmology* 108(1):13–14. [KNJ]

Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. (1996) Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3(2):131–41. [GM]

Robertson, M. (1989) Idiosyncratic motor units generating innate motor patterns: Neurones and circuits in the locust flight system. In: *The computing neuron,* ed. R. Durbin, C. Miall & G. Mitchison. Addison-Wesley. [aBW]

Roitblat, H. L. (1987) *Introduction to comparative cognition.* W. H. Freeman. [IMP]

Rosenblueth, A. & Wiener, N. (1945) The role of models in science. *Philosophy of Science* 12:316–21. [DRF, aBW]

Rothenberg, J. (1989) The nature of modeling. In: *Artificial intelligence, simulation and modeling,* ed. L. E. Widman, K. A. Loparo & N. R. Nielson. Wiley. [aBW]

Rucci, M., Edelman, G. & Wray, J. (1999) Adaptation of orienting behavior: From the barn owl to a robotic system. *IEEE Transactions on Robotics and Automation* 15:960–110. [aBW]

Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tenses of English verbs. In: *Parallel distributed processing: Explorations in the microstructure of cognition, vol. 2,* ed. J. L. McClelland, D. E. Rumelhart & The PDP Research Group. MIT Press. [GNR]

Russell, R. (1998) Odour sensing robot draws inspiration from the insect world. In: *Proceedings of the 2nd International Conference on Bioelectromagnetics, Melbourne, Australia (pp. 49–50),* ed. B. Lithgow & I. Cosic. IEEE. [aBW]

Rykiel, E. (1996) Testing ecological models: The meaning of validation. *Ecological Modelling* 90:229–44. [aBW]

Saito, F. & Fukuda, T. (1996) A first result of the brachiator III – a new brachiation robot modeled on a siamang. In: *Proceedings of a life V,* ed. C. Langton & K. Shimohara. MIT Press. [aBW]

Saksida, L. M., Raymond, S. M. & Touretzky, D. S. (1997) Shaping robot behavior using principles from instrumental conditioning. *Robotics and Autonomous Systems* 22:231–49. [aBW]

Salmon, W. C. (1996) Rationality and objectivity in science or Tom Kuhn meets Tom Bayes. In: *The philosophy of science,* ed. D. Papineau. Oxford University Press. [aBW]

Sandini, G., Metta, G. & Konczak, J. (1997) Human sensori-motor development and artificial systems. Paper presented at the International Symposium on Artificial Intelligence, Robotics, and Intellectual Human Activity Support for Nuclear Applications (AIR&IHAS '97), Japan, November 1997. [GM]

Sanes, D. H., Reh, T. A. & Harris, W. A. (2000) *Development of the nervous system.* Academic Press. [JH]

Scassellati, B. (1998) Imitation and mechanisms of joint attention: a developmental structure for building social skills on a humanoid robot. In: *Computation for metaphors, analogy and agents: Springer Lecture Notes in Artificial Intelligence, vol. 1562,* ed. C. Nehaniv. Springer-Verlag. [aBW]

(2001) Investigating models of social development using a humanoid robot. In: *Biorobotics: Methods and applications,* ed. B. Webb & T. Consi. AAAI Press. [rBW]

Schaal, S. & Sternad, D. (2001) Origins and violations of the 2/3 power law in rhythmic three-dimensional arm movements. *Experimental Brain Research* 136:60–72. [aBW]

Schall, J. & Hanes, D. (1998) Neural mechanisms of selection and control of visually guided eye movements. *Neural Networks* 11:1241–51. [aBW]

Schank, J. C. (2001) Beyond reductionism: Refocusing on the individual with individual-based modelling. *Complexity* 6:33–40. [JCS]

Schenk, H. (1996) Modeling the effects of temperature on growth and persistence of tree species: A critical review of tree population models. *Ecological Modelling,* 92:1–32. [aBW]

Schöner, G., Dose, M. & Engels, C. (1995) Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems* 16(2–4):213–45. [J-PB, LM]

Schultz, R. L. & Sullivan, E. M. (1972) Developments in simulation in social and administrative science. In: *Simulation in social and administrative science,* ed. H. Guetzkow, P. Kotter & R. L. Schultz. Prentice Hall. [aBW]

Schwartz, E. L. (1990) Introduction. In: *Computational neuroscience,* ed. E. L. Schwartz. MIT Press. [aBW]

Schweighofer, N., Arbib, M. A. & Kawato, M. (1998) Role of cerebellum in reaching movements in humans. I. Distributed inverse dynamic control. *European Journal of Neuroscience* 10:86–94. [RB]

Scutt, T. & Damper, R. (1997) Biologically-motivated learning in adaptive mobile robots. In: *SMC '97 IEEE International Conference on Systems, Man and Cybernetics, Orlando, FL, October 1997,* pp. 475–80. IEEE Press. [aBW]

Segev, I. (1992) Single neuron models: Oversimple, complex and reduced. *Trends in Neurosciences* 15:414–21. [aBW]

Sejnowski, T. J., Koch, C. & Churchland, P. S. (1988) Computational neuroscience. *Science* 241:1299–305. [aBW]

Selverston, A. I. (1993) Modeling of neuronal circuits: What have we learned? *Annual Review of Neuroscience* 16:531–46. [aBW]

Shaffer, D. M. & McBeath, M. K. (2002) Baseball outfielders maintain a linear optical trajectory when tracking uncatchable fly balls. *Journal of Experimental Psychology: Human Perception and Performance* 28(2):335–48. [TS]

Shannon, C. E. (1951) Presentation of a maze-solving machine. In: *Transactions of the 8th Conference of the Josiah Macy Foundation,* ed. V. Foerster. [aBW]

Shannon, R. E. (1975) *Systems simulation: The art and science.* Prentice Hall. [DRF, aBW]

Sharkey, N. E. & Ziemke, T. (1998). Editorial: Biorobotics. *Connection Science* 10:163–66. [aBW]

Sharpe, T. & Webb, B. (1998) Simulated and situated models of chemical trail following in ants. In: *From animals to animats 5: Proceedings of the Fifth International Conference on the Simulation of Adaptive Behaviour, Zurich,* ed. R. Pfeifer, B. Blumberg, J. A. Meyer & S. W. Wilson. MIT Press. [aBW]

Shaw, R. E., Flascher, O. M. & Kadar, E. E. (1995) Dimensionless invariants for intentional systems: Measuring the fit of vehicular activities to environmental layout. In: *Global perspectives on the ecology of human-machine systems,* ed. J. Flach, P. Hancock, J. Caird & K. Vicente. Erlbaum. [JAE]

Shepherd, G. M. (1990) The significance of real neuron architectures for neural network simulations. In: *Computational neuroscience,* ed. E. L. Schwartz. MIT Press. [aBW]

Shibata, T. & Schaal, S. (1999) Robot gaze stabilisation based on mimesis of oculomotor dynamics and vestibulocerebellar learning. *Advanced Robotics* 13:351–52. [aBW]

Shimoyama, I., Bekey, G. & Asaad, S. (1996) Introduction – Special Issue from International Workshop on Biorobotics. *Robotics and Autonomous Systems* 18:7–11. [aBW]

Sloman, A. (1999) What sort of architecture is required for a human-like agent? In: *Foundations of rational agency,* ed. M. Wooldridge & A. Rao. Kluwer Academic. [SF]

Sporns, O. (2001) The emergence of complex neuronal properties in an embodied model of the visual system. In: *Biorobotics,* ed. B. Webb & T. Consi. AAAI Press/MIT Press. [aBW]

Spriet, J. A. & Vansteenkiste, G. C. (1982) *Computer-aided modelling and simulation.* Academic Press. [aBW]

Srinivasan, M. V., Chahl, J. S., Weber, K. & Venkatesh, S. (1999) Robot navigation inspired by principles of insect vision. *Robotics and Autonomous Systems* 26:203–16. [aBW]

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

Steels, L. (1998) The origins of syntax in visually grounded robotic agents. *Artificial Intelligence* 103(1,2):133–56. [LS]

Stephens, D. W. & Patton, S. R. (1986) How constant is the constant of risk-aversion? *Animal Behaviour* 34:1659–67. [CD]

Stogdill, R. M. (1970) Introduction: The student and model building. In: *The process of model building in the behavioral sciences,* ed. R. M. Stogdill. Ohio State University Press. [aBW]

St.-Onge, N., Adamovich, S. V. & Feldman, A. G. (1997) Control processes underlying elbow flexion movements may be independent of kinematics and electromyographic patterns: Experimental study and modeling. *Neuroscience* 79:295–316. [RB]

Sugar, T. & McBeath, M. (2001) Spatial navigational algorithms: Applications to mobile robotics. Paper presented at the Vision Interface Annual Conference, Ottawa, Canada, July 2001. [TS]

Suluh, A., Sugar, T. & McBeath, M. (2001) Spatial navigational principles: Applications to mobile robotics. In: *Proceedings in the International Conference on Robotics and Automation, Seoul, Korea, May 2001, pp. 1689–94.* [TS]

Suppe, F. (1977) *The structure of scientific theories.* University of Illinois Press. [aBW]

Sutton, R. S. & Barto, A. (1998) *Reinforcement learning: An introduction.* MIT Press. [GM]

Taddeucci, D. & Dario, P. (1998) Experiments in synthetic psychology for tactile perception in robots: Steps toward implementing humanoid robots. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, May 1998,* pp. 2262–67. IEEE Press. [aBW]

Takanobu, H., Yajima, T., Nakazawa, M., Takamishi, A., Ohtsuki, K. & Onishi, M. (1998) Quantification of masticatory efficiency with a masticatory robot. In: *Proceedings of the IEEE International Conference on Robotics and Automation, Leuven, Belgium, May 1998,* pp. 1635–40. IEEE Press. [aBW]

Taylor, P. J. (1989) Revising models and generating theory. *Oikos* 54:121–26. [aBW]

Terzopoulos, D., Tu, X. & Grzeszczuk, R. (1994) Artificial fishes: Autonomous locomotion, perception, behavior, and learning in a simulated physical world. *Artificial Life* 1(4):327–51. [TRN]

Thelen, E., Schoner, G., Schier, C. & Smith, L. B. (2001) The dynamics of embodiment: A field theory of infant perserverative reaching. *Behavioral and Brain Sciences* 24:1–86. [DRF]

Touretsky, D. S. & Saksida, L. M. (1997) Operant conditioning in Skinnerbots. *Adaptive Behaviour* 5:219–47. [CD, aBW]

Traub, R. D., Wong, R. K., Miles, R. & Michelson, H. (1991) A model of a CA3 hippocampal pyramidal neuron incorporating voltage-clamp data on intrinsic conductances. *Journal of Neurophysiology* 66:635–50. [RK]

Triantafyllou, M. S. & Triantafyllou, G. S. (1995) An efficient swimming machine. *Scientific American* 272:40–48. [aBW]

Turing, A. M. (1952) The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London, Series B* 237:37–72. [DRF]

Turkewitz, G. & Kenny, P. A. (1982) Limitation on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. *Developmental Psychology* 15:357–68. [GM]

Turvey, M. T. (1990) Coordination. *American Psychologist* 45:938–53. [RB, JAE]

Ulinski, P. S. (1999) Modeling cortical circuitry. In: *Cerebral cortex, vol. 13,* ed. P. S. Ulinski, E. G. Jones & A. Peters. Kluwer Academic. [aBW]

Uttal, W. R. (1990) On some two-way barriers between models and mechanisms. *Perception and Psychophysics* 48:188–203. [aBW]

Van Fraassen, B. C. (1980) *The scientific image.* Clarendon Press. [aBW]

van Gelder, T. (1998) The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* 21:615–65. [DRF]

Verschure, P. F. M. J. (1996) Connectionist explanation: Taking positions in the mind-brain dilemma. In: *Neural networks and a new AI,* ed. G. Dorfner. Thompson Press. [aBW]

Verschure, P. F. M. J. & Voegtlin, T. (1998) A bottom up approach towards the acquisition and expression of sequential representations applied to a behaving real-world device: Distributed Adaptive Control III. *Neural Networks* 11:1531–49. [EN]

Vicente, K. J. (1999) *Cognitive work analysis: Toward safe, productive, and healthy computer-based work.* Erlbaum. [JAE]

Voegtlin, T. & 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* 10:291–310. [aBW]

Vogel, S. (1999) *Cat's paws and catapults.* Penguin Books. [aBW]

von Bekesy, G. (1960) *Experiments in hearing.* McGraw-Hill.

von Buddenbrock, W. (1921) Der Rhythmus der Schreitbewegungen der Stabheuschrecke *Dyxippus. Biologisches Zentralblatt* 41:41–48. [JH]

von Helmholtz, H. (1878/1921) Die Tatsachen in der wahrnehmung. In: H. von Helmholtz, *Schriften zur Erkenntnis theorie,* ed. P. Hertz & M. Schlick. Springer. (Conference from 1878; Springer edition published 1921). [CB]

Von Holst, E. & Mittelstaedt, H. (1950) Das reafferenzprincip. Wechselwirkungen zwischen Zentralnerven-system und Peripherie. *Naturwissenschaft* 37:467–76. English translation: The reafference principle. In: *The behavioral physiology of animals and man. The collected papers of Erich von Holst, vol. 1,* trans. R. Martin. University of Miami Press. [RB]

von Neumann, J. (1966) *Theory of self-reproducing automata.* University of Illinois Press. [DRF]

Vowles, D. M. (1964) Models and the insect brain. In: *Neural theory and modeling,* ed. R. F. Reiss. Stanford University Press. [FD]

Walker, I. D. (1995) A successful multifingered hand design – the case of the racoon. *IEEE/RSJ International Conference on Intelligent Robots and Systems,* 186–93. IEEE Computer Society Press. [aBW]

Walter, W. G. (1961) *The living brain.* Penguin Books. [aBW]

Warren, W. (1984) Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance* 10:683–703. [JAE]

(1998) Visually controlled locomotion: 40 years later. *Ecological Psychology* 10(3–4):177–219. [JAE]

Warren, W. & Whang, S. (1987) Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance* 13:371–83. [JAE]

Wartofsky, M. W. (1979) *Models: Representation and the scientific understanding.* D.Reidel. [aBW]

Watson, J. D. (1968*) The double helix.* Penguin Books. [rBW]

Webb, B. (1991) Do computer simulations really cognize? *Journal of Experimental and Theoretical Artificial Intelligence* 3:247–54. [aBW]

(1994) Robotic experiments in cricket phonotaxis. In: *From animals to animats 3: Proceedings of the Third International Conference on the Simulation of Adaptive Behaviour, Brighton,* ed. D. Cliff, P. Husbands, J. A. Meyer & S. W. Wilson. MIT Press. [aBW]

(1995) Using robots to model animals: A cricket test. *Robotics and Autonomous Systems* 16: 117–34. [aBW]

(2000) What does robotics offer animal behaviour? *Animal Behaviour* 60:545–58. [PEM]

Webb, B. & Scutt, T. (2000) A simple latency dependent spiking neuron model of cricket phonotaxis. *Biological Cybernetics* 82:247–69. [JAE, aBW]

Wehner, R. (1987) Matched filters – neural models of the external world. *Journal of Comparative Physiology A* 161:511–31. [KNJ, aBW]

(1994) The polarization-vision project: Championing organismic biology. In: *Neural basis of behavioural adaptations,* ed. K. Schildberger & N. Elsner. Gustav Fischer Verlag. [aBW]

Weiner, J. (1995) On the practice of ecology. *Journal of Ecology* 83:153–58. [aBW]

Weitzenfeld, J. S. (1984) Valid reasoning by analogy. *Philosophy of Science* 51:137–49. [aBW]

Weizenbaum, J. (1966) ELIZA – a computer program for the study of natural language communication between man and machine. *Communications of the Association for Computing Machinery* 9(1):36–45. [RID]

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M. & Thelen, E. (2001) Autonomous mental development by robots and animals. *Science* 291:599–600. [IMP]

Wickler, W. (1968) *Mimicry in plants and animals.* World University Library. McGraw-Hill. [CC]

Widman, L. E. & Loparo, K. A. (1989) A critical survey. In: *Artificial intelligence, simulation and modeling,* ed. L. E. Widman, K. A. Loparo & N. R. Nielson. Wiley. [aBW]

Wiener, N. (1948) *Cybernetics. Or, control and communication in the animal and the machine.* MIT Press. [aBW, DLY]

Williamson, M. W. (1998) Rhythmic robot arm control using oscillators. In: *1998 IEEE/RSJ International Conference on Intelligent Robots and Systems,* pp. 77–83. IEEE Computer Society Press. [JAE, aBW]

Wilson, H. R. (1999) Non-fourier cortical processes in texture, form and motion perception. In: *Cerebral cortex,* ed. P. S. Ulinski, E. G. Jones & A. Peters. Kluwer Academic. [aBW]

Wimsatt, W. C. (1987) False models as means to truer theories. In: *Neutral models in biology,* ed. M. Nitecki & A. Hoffman. Oxford University Press. [JCS]

Worst, R. (1998) Robotic snakes. In: *Proceedings of the Third German workshop on Artificial Life.* C. Wilke, S. Allmeyer & T. Martinetz. Harri Deutsch. [aBW]

XIII Oxford Conference Panel on Biomedical Modeling (2001) Modeling in the 21st century: An executive summary. In: *Frontiers in modeling and control of breathing: Integration at molecular, cellular and systems levels,* ed. C.-S. Poon & H. Kazemi. *Advances in Experimental Medicine and Biology, vol. 499, pp. 9–14.* Kluwer Academic/Plenum. [DLY]

Yamaguchi, H. (1998) A cooperative hunting behaviour by mobile robot troops. In: *Proceedings of the IEEE Conference on Robotics and Automation, Leuven, Belgium, May 1998,* pp. 3204–209. IEEE Press. [aBW]

Young, J. F. (1969) *Cybernetics.* Iliffe Books. [aBW]

Zalzala, A. M. S. & Morris, A. S., eds. (1996) *Neural networks for robotic control: Theory and Applications.* Ellis Horwood. [aBW]

Zeigler, B. P. (1976) *Theory of modelling and simulation.* Wiley. [RRM, aBW]

Zipser, D. (1992) Identification models of the nervous system. *Neuroscience* 47(4):853–62. [RID]