*PINS, 2004, 30, 77-79*

# PSYCHOLOGY AS MODELLING

**Book review**
Dawson, M R W (2004) **Minds and machines: Connectionism and psychological modelling.** Oxford: Blackwell. ISBN 1-4051-1349-9.

*David Spurrett*
*Philosophy*
*University of KwaZulu-Natal*
*Howard College Campus*
*Durban 4041*
*Email: spurrett@ukzn.ac.za*

Loose talk about “models” is, like loose talk about “deconstruction”, common and unhelpful. People who use “deconstruct” as no more than a stylistic variant of “interpret” or “analyse” abuse and devalue the distinctive theoretical claims of post-structuralism. Similarly, people who use “model” as a stylistic variant of “theory” or “mechanism” fail to show understanding of the distinctive demands that must be met before something counts as a model in science, or of the different sorts of model that there can be.

Dawson’s lively, engagingly written and wide-ranging book on psychological modelling makes no such mistake. For a start, he is clear about some of the major distinctions between types of model. Models of data, in which empirical measurements are subject to statistical analysis to find patterns aren’t the same thing as mathematical models that attempt to find equations that can produce the patterns observed in the data. As Dawson notes, mathematical models are, inter alia, less dependent on existing data, less likely to be linear, and more likely to be able to surprise us. Neither of these sorts of model, each of which is treated in a brief and judicious chapter, is Dawson’s main interest, though, to a significant extent because they do not themselves *behave*. That is, although a mathematical model of change in performance with practice might be useful, the model itself neither practises nor performs.

His main interest is in models that do behave, and so the bulk of the book is concerned with computer, and to a lesser extent robotic, simulations. More specifically, as the book’s sub-title makes clear, he is interested in connectionist models of psychological processes. This means that non-connectionist modelling approaches, such as “production systems” and other mainstays of classical artificial intelligence and cognitive psychology are handled rather briskly. Connectionist models use networks of artificial (and almost always virtual) neurons, with features abstracted away from biological

neurons.[1] These networks can be "trained" to perform a variety of tasks by various methods, and Dawson discusses the most common approaches to network training. Such networks have produced impressive results, especially in perceptual and control tasks at which conventional symbolic computing has been less successful.

Dawson's main project is to defend the prospects for a "synthetic psychology" (inspired by Braitenberg, 1984) paradigmatically based on connectionist models, although including discussions of some examples from robotics and the study of simple living systems. Dawson is amply qualified to produce such a survey, having himself done important work on building connectionist models, and contributed to our understanding of the technical capabilities of networks, and ways of building and training them.

Dawson suggests a schema for the progress of synthetic psychology, starting with the *synthesis* of a working system, perhaps in quite opportunistic and/or weakly premeditated ways, followed by studying the system in operation, at which stage *emergence* may be observed. Although there is considerably more loose talk about emergence than there is about modelling and deconstruction combined, Dawson invokes the concept with a careful eye to some of the more responsible exemplars, including Holland (1998), to label functionality that arises from (non-linear and typically recursive) interaction of multiple sub-systems and the environment. Finally, we have *analysis* in which an attempt is made to understand how it is that the working system is effective, and how the emergent features arise. Unlike a number in the field, including Braitenberg himself in some moods, Dawson does think analysis is an important stage in its own right, i.e. that it isn't enough to build a "working" model and stop, since successful replication leaves un-answered questions concerning what, if anything, the model tells us about the real system that might have inspired it, and how we might seek to further develop the system by adding functionality.

In the bulk of the book this schema is tested against a succession of examples, including some famous connectionist systems such as NETtalk (Sejnowski & Rosenberg, 1988), and some frequently discussed examples of problem solutions in living systems, such as phonotaxis in the cricket, and in robotics, including some of Rodney Brooks' better-known early mobots, as well as simple robots built by Dawson's students. The later chapters in particular focus on the challenges of analysis, and provide a clear introduction to a selection of the main tools for analysing connectionist networks (such as cluster analysis, that attempts to determine the representational or computational division of labour within a trained network that successfully performs some task) as well as developing Dawson's argument that analysis is important, and that it carries its own distinctive demands and challenges.

An important virtue of this book is that the content and order of presentation has clearly been tested at length in the classroom of a dedicated and creative teacher. The book has many illustrations from teaching practice, and would be an excellent basis for a senior undergraduate or introductory graduate course on cognitive modelling, and I'd be delighted to use it for that purpose myself. The book also provides a well-organised

[1] Growing understanding of the details of the working of biological neurons make clear that some of the abstractions are very severe, and more recent work that Dawson does not discuss seeks to try to close the gap.

platform for explorations into the wider and more recent literature, and could readily form the backbone of a more advanced course. It covers many of the same examples and problems as Clark's well known (1997) but in a way generally better suited to the classroom, and more clearly focussed on the problems of working cognitive scientists. This is a fine book, and I suspect it would be a valuable resource for those who don't know much about synthetic psychology but would like to get a clear sense of the lie of the land.

**REFERENCES.**

Braitenberg, V (1984) **Vehicles: Explorations in synthetic psychology.** Cambridge, MA: MIT Press.

Clark, A (1997) **Being there.** Cambridge, MA: MIT Press.

Holland, J (1998) ***Emergence*.** Reading, MA: Perseus Books.

Sejnowski, T J & Rosenberg, C R (1988) NETtalk: A parallel network that learns to read aloud, in Anderson, J A & Rosenfeld, E (eds) **Neurocomputing: Foundations of research.** Cambridge, MA: MIT Press.