that allow us to interpret the words used in posing the problem in an unintended fashion; and *scenario* effects that incline us (mistakenly) to assimilate the task to quite a different project than any standard interpretation of the words would permit. He describes modifications of the original experiments that would be expected to generate much higher success rates (modifications that block cueing of the wrong patterns) and reports some of the results.

The last six chapters deal with a different type of evidence. Margolis looks to the history of science, in particular to the great revolutionary changes in science, to see if the formulation, defence and discussion of radically new ideas can be illuminated by his account of cognition. Here, success depends on showing that those who originate new concepts (Copernicus and Darwin are the exemplars) were in an especially good position to do so, on showing how their arguments induced the right kinds of patterns in those who supported the new cause, and on explaining the resistance of the opposition.

Margolis devotes most attention to the copernican revolution. He argues that Copernicus's route to discovery could indeed have begun with his professed dissatisfaction with the use of equants in Ptolemaic astronomy, even though the equant is logically irrelevant. Copernicus was particularly well-placed to articulate a new pattern (a view of the Solar System in which the ordering of the planets was both determinate and consonant with the periods of rotation) because of his freedom from teaching, because of his own early research on the anomalies of the motions of the Moon and because of the impact of new discoveries in geography. Together, Margolis says, these factors would have diminished Copernicus's commitment to the ptolemaic pattern of viewing the Solar System as a group of nested spheres. Once freed from that commitment, he was able to develop an alternative perspective until its intrinsic harmony became apparent. Moving the Earth is a small by-product of the new vision.

Margolis continues with some thought-provoking analyses of the reception of copernicanism and the trial of Galileo. There are places in which his history seems strained or lop-sided: in his neglect of the standard explanation (Kepler's predilection for bizarre numerology) for Galileo's failure to refer to Kepler's laws (p. 286), and especially in his downplaying of the role of aristotelianism in the copernican debate. But these are minor blemishes on a provocative reinterpretation.

More fundamental is a difficulty which besets not only the historical analysis but much of the book. It is often hard to know whether the claims that are made about

subjects' performances, in twentieth-century academia or Renaissance *palazzi*, really involve the cueing and miscueing of patterns. At various key stages in the book Margolis describes behaviour that could more easily be analysed in everyday terms or in the idioms of other cognitive theories. Does the notion of pattern-recognition appear so powerful because it is so plastic? Could a clever exponent deploy it to cover anything?

Naive falsificationists, posing such questions, might conclude that this book is

worthless. They would be wrong. *Patterns, Thinking, and Cognition* is complex and compressed, but the scope and suggestiveness of Margolis's proposals indicate that the effort involved in making them more precise and definite may be very worthwhile indeed. We are offered a broad and challenging vision of human cognition. Formal logic and its offspring may never look the same again. □

Philip Kücher is a Professor in the Department of Philosophy, University of California at San Diego, La Jolla, California 92093, USA.

Ways of knowledge

Drew McDermott

Logical Foundations of Artificial Intelligence. By Michael R. Genesereth and Nils J. Nilsson. *Morgan Kaufmann: 1987. Pp.405. Distributed by W.H. Freeman, \$36.95, £29.95.*

ANYONE interested in artificial intelligence (AI) should be interested in the topic of this book, that is, computer manipulation of knowledge expressed in logical formalism. Unfortunately, the book is at best a partial success.

Knowledge is important to AI because an intelligent agent must know something about the world it lives in, and about how to solve problems that arise in that world. Many practitioners of AI, including the authors of this book, are driven by the idea that knowledge should be represented explicitly — 'declaratively'. And many who espouse that position further believe that mathematical logic is the proper framework for studying the syntax, semantics and inference mechanisms of declarative knowledge representation systems. This is the position known as 'logicism'. Genesereth and Nilsson place themselves in that camp.

It would be good to have a thorough exegesis of logicism, which enjoys a position of prestige as one of the few theoretically weighty branches of AI. But the book doesn't fill that role. It fails to present an adequate case for why AI should concentrate so strongly on deduction as the paradigm for inference. Worse, it is rather parochial and fails to do justice to a lot of logicist research. Many people outside Stanford will be surprised to find that their work is either not mentioned at all or is relegated to bibliographical notes.

The book does have some high points. Chapter 1 is a concise presentation of the logicist position, and in Chapter 2 Genesereth and Nilsson do a good job of explaining denotational semantics for AI purposes. The notion of *conceptualization*, which the authors depend on, is a nice adaptation of the idea of *intended model* to AI. But they are not sure what to

say about inference. They classify inference processes into two groups, sound and unsound, and acknowledge that most interesting inference algorithms are unsound (liable to draw invalid conclusions), but their attention is focused on the sound ones such as resolution theorem proving. So, after the presentation of deductive techniques in Chapters 3 to 5, the discussion bifurcates into treatments of various nondeductive forms of inference, and unrelated applications of deduction to particular domains. As an example, after the illuminating discussion of 'nonmonotonic' logic in Chapter 6 (that is, logic that allows defeasible conclusions to be drawn), no use is made of it again. Instead, in Chapters 11 and 12, on temporal reasoning and planning, problems are defined in terms of classical monotonic deduction where possible. For instance, the planning problem is defined as the problem of proving that a series of actions will bring about a desired state of the world. This is not a definition that an active researcher in the field would accept.

There are a few places where the presentation is lucid and thorough. The account of nonmonotonic logic in Chapter 6 is instructive, especially the parts about John McCarthy's 'circumscription' technique. This chapter is as up to the minute as a textbook can hope to be. Chapter 9, on the logic of knowledge and belief, is another winner. Here, two approaches are described: Kurt Konolige's syntactic approach, in which belief is a predicate on formulae in an agent's database; and Jaakko Hintikka's possible-worlds approach, adapted to AI by Robert Moore, in which belief in a proposition is attributed to an agent if the proposition is true in all possible worlds compatible with what he believes. Both 'conceptualizations' are interesting and powerful.

Sadly, however, many of the other chapters are either out of date or aimless. The reader is not given much internal evidence of the value of the logicist approach, and may find his doubts about it unassuaged. □

Drew McDermott is a Professor in the Department of Computer Science, Yale University, PO Box 2158, Yale Station, New Haven, Connecticut 06520, USA.