The formal semantics of logical L-KARL, procedural P-KARL and the three layers above are presented in separate sections, concentrating on a model-theoretical semantics for L-KARL. However the relationship between the languages and the layers is not made explicitly clear. A separate section or chapter on this aspect would have been helpful.

The conclusions include related work, providing a comparison with the knowledge specification language $(ML)^2$ developed as part of the KADS project, shortcomings of KARL, and future work. The reference list is usefully comprehensive and the index is adequate except for abbreviations (e.g., KADS, KARL, MIKE and $(ML)^2$) which are not included.

This work is based on a PhD thesis, and is thus fairly specialised. It will be of interest to researchers in the area of the formalisation of knowledge engineering, and of more indirect interest to those in the field of formal methods in general. Libraries with a comprehensive section on knowledge engineering or artificial intelligence will probably want to acquire this book but its readership will be limited by its very nature. The approach and presentation is very mathematically formal in general; although KARL is executable, there is no obvious indication as to what tool support is available or planned, if any. Without such support, the executable aspects to not seem very helpful.

In conclusion, I would prefer to see KARL considered as a formal modeling language rather than a specification language for reasons as outlined above. I am glad to see some unification and interconnections of approaches adopted by software engineering and knowledge engineering researchers; I hope that each community can learn something from the other.

Reviewed by Jonathan Bowen, The University of Reading, Department of Computer Science, Whiteknights, PO Box 225, Reading, Berks RG6 6AY, England

Reasoning about knowledge by Ronald Fagin, Joseph Y. Halpern, Yoram Moses and Moshe Y. Vardi, MIT Press, Cambridge, MA, 1995, pp 477, $45.00 cloth, ISBN 0-262-06162-7.

One way or another, academics spend much of their working lives reading. Sadly, most of us find that a large proportion of what we read is not terribly interesting. We review articles, correct student work and read books by the truck load simply to keep abreast of developments in our field. But little of this reading is done with real pleasure or interest; most of it results from a weary sense of duty. Fortunately, every now and then, a book comes along that reminds you what reading *should* be like; both enjoyable, and a real learning experience. I am happy to report that *Reasoning About Knowledge* is just such a book.

The tradition of using modal logics in the formal analysis of knowledge and belief is generally reckoned to have begun in earnest with the work of Jaakko Hintikka. It seems unlikely that Hintikka, publishing his classic 1962 book *Knowledge and Belief*, could have anticipated that his work would have any application outside the somewhat ivory tower tradition of formal philosophy within which he then worked. And yet, three decades later, modal logics of knowledge are a standard tool in both AI and theoretical computer science, and are the subject of much ongoing research.

Within AI, logics of knowledge are considered to be important because we are increasingly concerned with building *agents*: autonomous, self-contained computer systems, that are capable of communicating and cooperating with other agents in order to achieve their goals (Wooldridge & Jennings, 1995). It is commonly accepted that in order to do this, an agent must be able to manipulate representations of its peers. Logics of knowledge are intended to support exactly this kind of reasoning. To better understand what this reasoning looks like, consider the following puzzle, which is often used as a kind of benchmark for logics of knowledge:

> A certain king wishes to determine which of his three wise men is the wisest. He arranges them in a circle so that they can see and hear each other, and informs them that he will paint a white or black spot on each of

their foreheads. He also tells them that at least one spot will be white, and that he will offer his favour to the one who first correctly tells the king the colour of his spot. After a while, the wisest man announces that his spot is white. How does he know?

The solution goes as follows:

The wisest man starts by assuming that his spot was black. If this were the case, then the second may would see both a white spot and a black spot. In this case, the wisest man knows that the second man would start by assuming his spot was black. But in this case, the second man would know that the third man would have seen two black spots, and since at least one spot is white, the third man would know his spot was white. Since the third man made no such announcement, the second man would know that his assumption (that his spot was black) must be wrong, and hence that his spot was white. But the second man made no such announcement, and so the wisest man would know that his original assumption—that his spot is black—must be wrong. Hence he knows his spot is white.

Solving the puzzle thus involves reasoning about the knowledge and ignorance of the wise men. Clearly, the natural language version of the solution given here is cumbersome and hard to read, but an axiomatisation and proof of the solution in a modal logic of knowledge is fairly straightforward.

Logics of knowledge are also of great interest in theoretical computer science, and in particular, in the theory of distributed systems. To see why, simply observe that in specifying the behaviour of a communication protocol or network, one often makes use of informal reasoning like this:

IF
node 1 knows that node 2 has received message m_1
THEN (1)
node 1 should send node 2 message m_2

Logics of knowledge allow us to give a precise meaning to such statements.

Reasoning About Knowledge is the first modern textbook-style volume devoted solely to logics of knowledge, of the kind that may be used to solve the wisest man puzzle, and that form the basis for reasoning about distributed systems. The book could hardly have had authors that are more respected in the field: all of them, in particular Joseph Halpern, have made significant contributions to knowledge theory. The book can broadly be divided into three parts. The first part (Chapters 1 to 3) essentially provide the logical foundations upon which the remainder of the book stands. Chapter 1 introduces the basic idea of reasoning about knowledge by way of the *muddy children* puzzle (a variant of the wisest man puzzle presented above). Chapter 2 introduces the basic semantic model for knowledge, which is known as the *possible worlds* framework, or *Kripke semantics* (after their inventor, Saul Kripke). It is worth noting that many people find possible worlds semantics both abstract and obscure, but the description given here is extremely clear. The basic linguistic framework—normal modal logic—is introduced, and the notions of common knowledge (what everyone knows that everyone knows...) and distributed knowledge (intuitively, the knowledge that you would get by pooling together everyone's knowledge) are introduced. Chapter 3 discusses completeness and complexity of the decision problems for logics of knowledge. Unless you are particularly interested in complexity results, you could quite happily skip the complexity section, and most of the remainder of the book would remain accessible to you. But if you did this, you would miss out a very nice summary of complexity classes and NP-completeness, written by some of the best-known people working in the area. Chapter 3 closes with a discussion of first-order modal logics of knowledge. Introducing quantification into modal logics is a tricky subject, with some rather subtle issues, but once again the authors do a good job of making these issues plain. (Incidentally, the only definite mistakes I found in the book were in Chapter 3: some incorrect equation numbers on pp 84–85. Fortunately, it's easy to guess what the correct numbers are.)

The second part of the book (Chapters 4, 5, 7 and 10) is broadly concerned with the relationship between logics of knowledge and distributed systems. Chapter 4 introduces the by-now familiar interpretation of knowledge in such systems. This interpretation goes like this: a process in state s is said to know ϕ if in all possible situations where the process is in state s, ϕ is true. The authors show how a variety of (admittedly quite simple) systems can be modelled using this framework. This leads

into a discussion of various properties of distributed systems. For example, *synchrony* holds of a system if all components of the system change state simultaneously. Synchrony, like the other properties discussed (perfect recall, asynchronous message passing) turns out to be quite an intuitive property, with a simple semantic characterisation. Chapter 4 also presents one of the key results in knowledge theory, which intuitively states that in certain sorts of systems, common knowledge cannot be gained. That is, a group of processes can only have common knowledge of ϕ at time u if, and only if, they started with common knowledge of ϕ. This has implications for coordinating multi-process activity. The main subject of Chapter 4 is the idea of a *protocol*. Intuitively, a protocol defines the actions that may be performed by a process; we are often concerned with designing protocols to coordinate multi-process activity. Chapter 7 introduces the related idea of a knowledge-based program. Intuitively, a knowledge based program is one which takes the form of a set of rules rather like (1), above. The idea is that *executing* such a program would involve explicit *tests* for knowledge. The relationship between programs and protocols is discussed, and the concept of a protocol *implementing* a program is defined. (The idea of a protocol implementing a program may seem somewhat strange—usually we tend to think of *programs* implementing *protocols*. But in the knowledge-based case discussed here, programs are the more abstract representation.) Chapter 10 attempts to further bridge the gap between the fairly abstract representation of knowledge in possible worlds semantics, and real systems. It introduces the idea of algorithmic knowledge, i.e., knowledge that an agent can access via some computation.

The third part of the book, (Chapters 6, 8, 9 and 11) discusses what might be called related issues. Chapters 6 and 11 address the issue of common knowledge. It turns out that common knowledge is a particularly useful abstraction in considering certain problems (such as the *simultaneous Byzantine agreement* problem). However, common knowledge also turns is also quite a demanding property. It turns out that in systems that make even quite reasonable assumptions (such as the possibility of error in communication), common knowledge can be shown to be unattainable. Chapter 11 discusses some issues associated with common knowledge, including the idea of common knowledge as a fixed point. (If you don't know, or can't remember what a fixed point is, then don't worry about it: you will probably lead a happier life as a result.) Chapters 6 and 11 probably require the greatest mathematical sophistication of any sections in the book. Chapter 8 discusses the implications of introducing a temporal element into the basic logic of knowledge. Such an enriched language can be used to capture some simple properties of distributed systems (such as synchrony and perfect recall). One of the more intriguing issues in the book is the way that introducing even fairly simple semantic properties affects the complexity of the decision procedure for the resulting logic. (It is worth mentioning that the associated proofs, previously published in Halpern and Vardi (1989), are rather complex, and, presumably for this reason, have not been included in the book.) Finally, Chapter 9 discusses the issue of logical omniscience. Logical omniscience is a side effect of using possible worlds semantics to characterise knowledge. It is the undesirable prediction, made by these logics, that a knower has infinite, perfect reasoning abilities: it will know all the logical consequences of its knowledge. Logical omniscience is a problem because *real* knowers—you and I, and any system we care to build—will *not* have infinite, perfect reasoning abilities. A great many papers have been written on the subject of logical omniscience, and a similarly large number of formalisms have been developed as an alternative to possible worlds semantics, in an attempt to circumvent the problem. Chapter 9 discusses the nature of the logical omniscience problem in detail, and some of the formalisms developed to avoid it.

If I had to identify some shortcomings of the book, they would be as follows. First, no solutions are provided to any of the many exercises. This may seem like a trivial complaint, but some of the exercises are *hard*, and my guess is that many readers will not have access to experts in the field to answer their questions and help them with the exercises. Solutions to just some of the problems would be a real benefit here. I would also like to have seen proof theoretic aspects discussed. Proof methods for modal logics are a growth area, and given that the authors have previously done some relevant work (Halpern & Moses, 1992), it would presumably have been a fairly straightforward matter to include a summary. Another gripe is that I was not clear how close the theory is to

practice. I know of no real-world examples of systems that have been specified using logics of knowledge, and at one point (p 379), the authors actually point out that there is no methodology for going from knowledge-based programs to implementations. A larger discussion on systems that really *do* make use of logics of knowledge (such as situated automata (Rosenschein & Kaelbling, 1986) would have deflected this criticism. Finally, I was a little worried that the reader does not get a good feel for how the general area is developing, or what the major issues driving current research in knowledge theory are. For sure, there are excellent 'further reading' sections, and the exercises contain many research-level problems. But I would ideally have liked a single, coherent statement of what the current hot topics are, and the likely future development of the field. However, I am being pernickety. All things considered, this is an excellent book. It is a model of clarity and presentation, extraordinarily rigorous and thorough, but perhaps most importantly for a textbook that is essentially mathematical, *Reasoning About Knowledge* is eminently readable.

References

Halpern, JY and Moses, Y, 1992. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence* **54** 319–379.

Halpern, JY and Vardi, MY, 1989. The complexity of reasoning about knowledge and time. I. Lower bounds. *Journal of Computer and System Sciences* **38** 195–237.

Rosenschein, S and Kaelbling, LP, 1986. The synthesis of digital machines with provable epistemic properties. In Halpern, JY, ed, *Proceedings 1986 Conference on Theoretical Aspects of Reasoning About Knowledge* 83–98, Morgan Kaufmann.

Wooldridge, M and Jennings, NR, 1995. Intelligent agents: Theory and practice. *The Knowledge Engineering Review* **10**(2) 115–152.

Reviewed by Michael Wooldridge, Department of Computing, Manchester Metropolitan University, Manchester M1 5GD, United Kingdom