Reset reproduction of review published in *ACM SIGART Bulletin*, Vol. 2, No. 1 (January 1991), pp. 77–78. Reprinted July 1994.

Review of *Philosophical Logic and Artificial Intelligence*
Edited by Richmond H. Thomason
Dordrecht: Kluwer Academic Publishers, 1989

Jon Doyle DOYLE@LCS.MIT.EDU
Massachusetts Institute of Technology, Laboratory for Computer Science
Cambridge, Massachusetts 02139, U.S.A.

This book represents a welcome departure from the past. Many readers in AI avoid looking at books on philosophy and artificial intelligence, at least after seeing one or two of the genre. The reason for this disinterest is aptly captured by Thomason in his introduction to the present volume, which expands on a special issue of the *Journal of Philosophical Logic* published in 1988.

> There is an unfortunate reflex that is widespread in contemporary philosophy: when artificial intelligence is mentioned, philosophers are reminded of the problem of whether computers can think. Now, that is certainly an interesting problem, but reflecting on it doesn't require any detailed knowledge of artificial intelligence. (pp. 4–5)

The papers presented here offer great contrast to the usual philosophical discussions. They are rigorous treatments of philosophical problems that are fully informed about research in artificial intelligence, and are directly concerned with the specific logical and epistemological issues that arise in attempting to automate reasoning and represent knowledge. Each one speaks directly to the concerns of some subfield of AI, in language intelligible to both philosophy and AI.

The relevance of these philosophical articles to current AI thinking has a simple reason: each of the eight authors is a prominent researcher in AI or computer science. Only two (Thomason and Israel) received formal training as philosophers. Yet, as Thomason stresses, the work described in these papers represents research in philosophical logic.

Thomason's introductory article does not attempt to summarize the six principal papers of the collection, but instead examines something of the relationships between traditional work in logic and the new work in AI. He argues that while many of the problems addressed by the new work are entirely in line with the traditional concerns of logic as studied by philosophers, linguists, and mathematicians, it is the new methods offered by AI that provide the new work with a "competitive advantage." To quote again from the introduction:

> To me, at least, a major impediment to progress [in traditional work in philosophical logic] has been the difficulty of developing sound methodologies for motivating logics. You can point to examples of good motivation, but it is not easy to articulate principles for testing hypotheses about reasoning, or even about whether premisses and conclusions are true. (p. 3)
>
> But the technology [grounding work in AI] has played an important part in generating and evaluating the theories. The requirement that a logical theory

> should inspire the design of high performance AI algorithms adds a dimension to this work that has shaped it from beginning to end. The thought experiment in robotics that Carnap performed in the second edition of *Meaning and Necessity* has become a reality. I find this aspect of logical research in AI to be particularly exciting, because it promises not only to suggest new logical theories, but to provide new methods for motivating and evaluating the theoretical alternatives.

Thomason's introduction is mainly aimed at philosophers, explaining why they should consider looking at work in AI, but it also provides AI researchers with an encouraging perspective on the significance of their own field. At the same time, the perspective is humbling: while the AI authors make real contributions to philosophical logic, they also draw heavily on past work in logic and philosophy. The methods of AI may be wonderfully novel and powerful, but hard-won insights of past thinking on the same problems cannot be ignored if the field is to use its new methods wisely.

The first article, "I'm OK if you're OK: on the notion of trusting communication" by Ronald Fagin and Joseph Halpern, studies the problem of what someone needs to know in order to know that the statements he or she (or it) communicates to others are true. The theory presented here derives from work on distributed systems, and studies various notions of "honesty" in communication. An "honest" message is one that is known to be true when it is sent. Having every message be honest ensures that every message is true, but honesty in this sense is usually unattainable. Fagin and Halpern seek to broaden the theory, first by formalizing the notion of "conditional honesty," according to which a message is known to be true if all previous messages are true, and then by formalizing a notion of "trusting" messages, according to which a message is trusting if its sender knows that it is true if all previous messages are trusting. If every message is trusting, then every message is true. Fagin and Halpern show that trusting is, in general, not a well-defined notion, but that it can be precisely defined in terms of fixed points of a certain equation. These fixed points need not exist, and need not be unique if they do exist, but are shown to be unique in certain systems possessing global synchronized communication.

The formalism employed by Fagin and Halpern has close connections with traditional philosophical modal logics of knowledge. David Israel examines some more recent formal theories of knowledge and information in his paper "Concepts of information: comparative axiomatics." He presents a summary of a theory of knowledge and information developed by himself and John Perry. This theory, based on situation semantics, is used to argue that though some traditional modal logics may be suspect as theories of knowledge, they are reasonable as theories of information. Israel also compares his theory with the theory of situated automata developed by Stanley Rosenschein and the modal logics of knowledge developed by Halpern and Yoram Moses (on which the Fagin and Halpern article build). The main result is that while the theories of Rosenschein, Halpern, and Moses validate some traditional modal logics of knowledge (specifically, the logics S4 and S5), the Israel-Perry theory is incompatible with these logics. The situation-based theory explicitly acknowledges that information-carrying states may be subject to laws which constrain the content of information conveyed by the states, while the more traditional theories do not incorporate such constraints.

Hector Levesque uses his article, "Logic and the complexity of reasoning," to argue, in contrast to the judgment of much of modern psychology, that logic does have a significant

role to play in understanding thinking, at least when logic is used with awareness of its computational costs. His thesis is that some deviations from classical logic observed in human behavior are closely related to the deviations necessary to ensure the computational tractability of logical reasoning. Levesque first sets the stage by a general discussion of the significance of computability and efficiency in reasoning (especially notions like polynomial time computability), and then summarizes some approaches to making reasoning tractable, including his notion of "vivid" representations, the use of Horn and semi-Horn databases, relevance logic, nonmonotonic reasoning, and simulation of general logical inference.

In "Circumscriptive theories: a logic-based framework for knowledge representation," Vladimir Lifschitz introduces a fairly general form of circumscription. Circumscription draws conclusions by assuming that the extensions of certain predicates are as small as possible, given the axioms. While previous treatments of circumscription have involved a large variety of special rules for minimizing extensions depending on the inferential problem being addressed, in the present work Lifschitz shows how to express the rules and constraints governing predicate minimization as axioms within the knowledge base itself, in the same language as the axioms describing the objects of reasoning. The presentation assumes no prior familiarity with circumscription, and illustrates the new theory by recasting many familiar examples from earlier treatments.

John McCarthy, in "Artificial intelligence, logic and formalizing common sense," urges philosophers to contribute to AI by surveying a number of problems of interest to both fields. McCarthy shares Thomason's belief that examination of philosophical theories in the more detailed contexts offered by AI is necessary for philosophy to contribute much to AI, and is also likely to be a prerequisite for significant philosophical progress. He uses this paper to survey a number of philosophical problems from the perspective of AI, concentrating on the importance of common sense, the roles of logic, nonmonotonic reasoning, free will, knowledge and belief, reasoning about context, and some more general epistemological concerns. AI readers will find the article a useful, compact survey of the many topics McCarthy has examined over the years.

The final article, "Efficient reasoning about rich temporal domains" by Yoav Shoham, excerpts some of the discussion in Shoham's book *Reasoning About Change* (MIT Press, 1988), with remarks about some more recent work. Shoham first describes the "qualification" and "extended prediction" problems (the latter subsuming the famous "frame" problem) arising in mechanizing temporal reasoning. He then presents his approach to solving these problems by using nonmonotonic logics. He summarizes his semantical theory of nonmonotonic logic based on preferential entailment, and then summarizes his notions of chronological ignorance and causal theories. Causal theories enjoy the attractive property that any of their "chronologically maximally ignorant" models (models of knowledge and time in which things come to be known to be true as late as possible) satisfy the same set of atomic base sentences. Moreover, this unique set of atomic conclusions can be computed rapidly.

The articles in this book do a good job of conveying the importance of AI approaches in addressing philosophical problems. They also suggest many problems for further exploration. But perhaps their greatest contribution is service as a high standard against which AI and philosophy can measure their contributions and judge the directions of their investigations.