R. C. Moore, *Logic and Representation*, CSLI Lecture Notes, vol. 39, Standford, Cal., 1995, 196 + XIV pages.

Reviewed by

YVON GAUTHIER

Département de philosophie
Université de Montréal
C. P. 6128, Montréal
Québec, H3C 3J7, Canada.

This volume of essays is a collection of publications by the author over a period of ten years. Robert C. Moore is principal scientist of Artificial Intelligence Center of SRI International and his work can be seen as an attempt to bring formal logic and artificial intelligence on a par. The emphasis is on models for the representation of knowledge and the use of elementary formal logic as an analytic tool. It should be clear from this review that the author has not attempted to update his essays on account of the fast-growing literature on the subject and I shall limit my comments to the views expressed by the author.

Moore sets out in the first essays to define "The Role of Logic in Artificial Intelligence". Here the author goes over J. A. Robinson's resolution principle — which was indeed a breakthrough in the field — and logic programming (as in PROLOG), which was another major advance in the seventies. On Robinson's resolution principle, one should consult the lucid report of Francine F. Abeles, "Herbrand's Fundamental Theorem and the Beginning of Logic Programming", *Modern Logic*, vol. 4, no. 1 (January 1994), 63–73. One would look in vain for such historical insights in the work under review. The author advocates "the more general use of logic in automatic reasoning"; the second essay "A Cognitivist Reply to Behaviorism" is a brief comment on Skinner's "Behaviorism at Fifty" (1984).

Part II of the book deals with "Propositional Attitudes", a subject familiar to philosophers of language and philosophers of logic. It is no surprise if the author draws heavily from such people as Russell, Kaplan, Lewis, Kripke, since he bases his analysis on epistemic logic (Hintikka), possible-worlds semantics (Kripke) and offers an integrated theory of knowledge and action (chapter 3). This is a revised version of the author's 1980 doctoral dissertation. He goes on with "Computational Models of Belief and the Semantics of Belief Sentences" (chapter 4 with G. G. Hendrix) which is a philosophical rather than computational treatment of truth-conditional semantics of belief sentences. Chapter 5 is a general discussion of "Propositional Attitudes and Russellian Propositions" that centers around such philosophical (nonetheless important) themes as proper

names, Russell's theory of definite descriptions and Kripke's notion of rigid designators.

Part III contains the author's most original contributions to the field. "Autoepistemic Logic" is the logic of self-reflecting rational agents, that is the thinking subjects who take their own thoughts or beliefs as objects of thought. In "Semantical Considerations and Nonmonotonic Logic", (chapter 6), the author again locates the problem in the 1980 context of commonsense reasoning and the discussion of the McDermott and Doyle modal account. Chapter 7 "Possible-World Semantics for Autoepistemic Logic" is a brief report on the possible applications of possible-worlds semantics to the author's own brand of autoepistemic logic which is concisely revisited in chapter 8.

Part IV of this volume of essays is devoted to the "Semantics of Natural Language" and begins with "Events, Situations, and Adverbs" (chapter 9), a discussion of Davidson's and Perry's views on actions and the situations in which they evolve. Again, the overview is meant to be an introduction to the problem, rather than an alternative proposal. The final chapter is entitled "Unification-Based Semantic Interpretation" and is intended as a semantical variant to the more syntactical lambda-calculus approach.

In all, this collection of essays seems to me to be a faithful representative of one's itinerary at the crossroads of artificial intelligence and logic (and philosophy) as practised some ten years ago. To conclude, I would like to assess the situation from a more contemporary point of view. I limit myself to general critical remarks.

Cognitive science has not been taken into account in the author's panorama: Dennett's recent work, for example, has been ignored and the debates on intentionality (Putnam, Dreyfus, Searle among others) are not even mentioned. From a more logical point of view, if default reasoning deserves a brief discussion, the problem of negation as failure, the object of a voluminous literature nowadays, has not been tackled. Of course, the whole subject of dynamic or nonmonotonic logic invites an open attitude and the work of Clark, Reuter, McCarthy among others bears witness to the complexity of the problem. From a mathematical perspective, the revolution announced by Y. Gurevich in theoretical computer science is not echoed in the papers under review. Classical logic might not be suitable for logic programming after all, since the finite structures that are supposed to be modelized by logic programmes do not possess classical properties, like recursive enumerability as B. A. Trakhtenbrot had already shown in 1950. Completeness and compactness or Löwenheim-Skolem theorems are not available for finite structures: this imposes severe limitations on the usefulness of classical logic for knowledge representation and for natural or automated reasoning. Any temptation to override those questions condemns the theoretician to outmoded speculation or philosophical reserve. By all means, logic should not be confined to the traditional problem area where knowledge representation is at stake. Elementary logical methods might have exhausted their ressources and the author's ambivalent attitude towards possible-worlds semantics is symptomatic. A turning of the tables is taking place and we may speak of the problem of complexity, rather than the

complexity of the problem. Contemporary complexity theory forces us to consider new avenues, from algorithmic information theory to random algorithms, zero-knowledge interactive systems and probabilistic proof systems. This last topic has potential impact on the very foundations of the subject. A. Wigderson and his co-workers have introduced randomness as an essential ingredient in interactive proof systems where no preexisting knowledge is involved. Here, knowledge representation (beyond cryptology) is out of question, but logic is still very much alive. What would be the significance of a title like *Logic and Representation* in that context, if representation still means representation of knowledge?