Epistemic Logic: : Annotated Bibliography

Ernest Davis*
New York University
New York, New York

Leora Morgenstern
IBM Watson Research Center
Hawthorn, New York

Disclaimer: To produce a comprehensive bibliography of the literature of epistemic logic and its applications in computer science and artificial intelligence would be a very substantial undertaking, and we have not attempted it. The works cited here are those that we personally have encountered and found important or helpful. We have also included a few philosophical works of epistemology that we have found particularly inspiring or thought-provoking.

For most of these references, we have given a brief annotation describing the contents. In a few cases we consider the title sufficient summary and have therefore omitted the annotation.

Beth Allen (1990). "Costly Acquisition of (Differentiated) Information." In (Parikh, 90), pp. 169-184. Analysis of the economics of knowledge.

Nicholas Asher and Hans Kamp (1986). "The Knower's Paradox and Representational Theories of Attitudes." In (Halpern, 86).

R.J. Aumann. (1976). "Agreeing to disagree." *Annals of Statistics.* pp. 1236-1239. Knowledge attained by communicating agents.

Jon Barwise and John Etchemendy (1987). *The Liar: An Essay intro Truth and Circularity.* Oxford University Press. Situation theory applied to self-referential sentences.

Jon Barwise and John Perry (1983). *Situations and Attitudes.* MIT Press, Cambridge, MA. A new approach to developing logic, with applications to the logic of knowledge.

T. Burge (1977). "Belief *De Re.*" *Journal of Philosophy,* June, 1977, pp. 338-362. Discussion of the *de re / de dicto* distinction.

H.H. Clark and C.R. Marshall (1981). "Definite Reference and Mutual Knowledge." In A.K. Joshi, B.L. Webber, and I.A. Sag (eds.), *Elements of Discourse Understanding.* Cambridge University Press. Common knowledge in interpreting referential phrases in discourse.

Philip Cohen and Hector Levesque (1986). "Persistence, Intention, and Commitment." In M. Georgeff and A. Lansky, *Reasoning about Actions and Plans: Proc. 1986 Workshop, Timberline, Oregon.* Morgan Kaufmann, San Mateo, CA. Formal development of the BDI (belief, desires, intention) model of cognition, and its application to plan analysis.

Ernest Davis (1988). "Inferring Ignorance from the Locality of Visual Perception." *Proc. AAAI-88,* pp. 786-790. Formal theory for describing and reasoning about perception and its connection to knowledge.

Ernest Davis (1989). "Solutions to a Paradox of Perception with Limited Acuity." *Proc. First Intl. Conf. on Principles of Knowledge Representation and Reasoning.* Morgan Kaufmann, San Mateo,

*The writing of this paper was suppported in part by NSF grant #IRI-9001447

CA. Extension of [Davis, 88] for imperfect perception.

Ernest Davis (1990). *Representations of Commonsense Reasoning.* Morgan Kaufmann, San Mateo, CA. Textbook on knowledge representation. Chapters 2,8,9,10 address epistemic representations and their applications.

Ernest Davis (1993). "Knowledge Preconditions for Plans." NYU Tech. Rep. #637. Generalization of (Moore, 1985) to indeterminate plans.

Daniel Dennett. (1978) *Brainstorms.* MIT Press, 1978. Highly readable philosophical essays on the nature of cognition.

Daniel Dennett (1981). "Beyond Belief." In Andrew Woodfield, ed., *Thought and Object.* Oxford University Press, 1981. Critique of the *de re / de dicto* distinction.

Jennifer Elgot-Drapkin and Don Perlis (1986). "Step Logics: An Alternative Approach to Limited Reasoning." *Proc. ECAI-86.* Model of cognition where an agent may one inference per time unit.

O. Etzioni, S. Hanks, D. Weld, D. Draper, N. Lesh, and M. Williamson (1992). "An Approach to Planning with Incomplete Information." *Third Intl. Conf. on Princ. of Knowledge Representation and Reasoning,* Morgan Kaufmann, San Mateo, CA. Non-linear planner that constructs plans with steps that gain knowledge.

Ronald Fagin and Joseph Halpern (1985). "Belief, Awareness, and Limited Reasoning." *Proc. IJCAI-85* pp. 491-501. Modal logic of explicit belief that avoids the consequential closure assumption.

Ronald Fagin, Joseph Halpern, and Moshe Vardi (1986). "What can machines know? On the epistemic properites of machines." *Proc. AAAI.* pp. 428-434.

Ronald Fagin and Moshe Vardi (1986). "Knowledge and Implicit Knowledge in a Distributed Environment." In (Halpern, 86).

Alan M. Frisch and Richard B. Scherl. (1991). "A General Framework for Modal Deduction." *Second Intl. Conf. on Princ. of Knowledge Representation and Reasoning,* pp. 196-207. Automated reasoning for modal logics.

Haim Gaifman (1988). "Operational Pointer Semantics: Solutions to Self-Referential Sentences." in (Vardi, 88). Avoiding paradoxes of self-reference. Meaning is considered as applying to utterances rather than to sentences.

John Geanakoplos. (1990) "Common Knowledge in Economics." In (Parikh, 90). Brief tutorial abstract.

Christophe Geissler and Kurt Konolige (1986). "A Resolution Method for Quantified Modal Logics of Knowledge and Belief." In (Halpern, 86).

Michael Genesereth and Nils Nilsson (1987). *Logical Foundations of Artificial Intelligence.* Morgan Kaufmann, San Mateo, CA. Graduate AI textbook. Chapter 9 deals with epistemic logics.

E. Gettier (1967). "Is Justified True Belief Knowledge?" In A.P. Griffiths, ed. *Knowledge and Belief.* Oxford University Press. pp. 144-146. Examples where an agent may have a justified true belief of P but still would not be said to know P.

S. Goldwasser, S. Micali, and C. Rackoff. (1985). "The Knowledge Complexity of Interactive Proof-Systems." *17th Symposium on the Theory of Computing.* pp. 291-304. Zero-knowledge interactive proofs.

H. Paul Grice (1957). "Meaning." *Philosophical Review.* Vol. 66 pp. 377-388. The necessity of common knowledge of convention and meaning for communication in natural language.

H. Paul Grice (1975) "Logic and Conversation", Syntax and Semantics, Vol. 3 Speech Acts, Peter Cole and Jerry L. Morgan, eds. Academic Press. Maxims of conversation.

Adam Grove (1993). "Semantics for Knowledge and Communication." *Third Intl. Conf. on Princ. of Knowledge Representation and Reasoning,* pp. 213-224.

Adam Grove and Joseph Halpern. (1991). "Naming and Identity in a Multi-Agent Epistemic Logic" *Second Intl. Conf. on Princ. of Knowledge Representation and Reasoning,* pp. 301-312. Epistemic theory when the identity of agents is only partially specified.

Anil Gupta (1982). "Truth and Paradox." *Journal of Philosophical Logic,* vol. 11, no. 1, pp. 1-60. Solution to paradox of self-reference.

Andrew Haas (1983). "The Syntactic Theory of Belief and Knowledge." Bolt, Baranek, and Newman, Report No. 5368. Syntactic theory applied to belief and knowledge to avoid the consequential closure problem.

Andrew Haas. (1986). "A Syntactic Theory of Belief and Action." *Artificial Intelligence,* vol. 28, no. 3, pp. 245-292. Syntactic theory applied to the problem of characterizing knowledge preconditions for actions.

Joseph Halpern, (ed.) (1986). *Theoretical Aspects of Reasoning About Knowledge.* Morgan Kaufmann, San Mateo, CA. TARK is the major biennial conference on epistemic logics and their applications. See (Vardi, 88), (Parikh, 90), (Moses, 92).

Joseph Halpern (1986a). "Reasoning about Knowledge: An Overview." In (Halpern, 86). Survey paper.

Joseph Halpern and Yoram Moses (1984). "Knowledge and common knowledge in a distributed environment." *Proc. 3rd ACM Conf. on Principles of Distributed Computing.* pp. 50-61.

Joseph Halpern and Yoram Moses (1985). "A Guide to the Modal Logics of Knowledge and Belief." *Proc. IJCAI-85,* pp. 480-490. Survey of propositional modal logics, their connections to possible worlds theories, and basic results in computational complexity.

Joseph Halpern and Moshe Vardi (1991). "Model Checking vs. Theorem Proving: A Manifesto". *Second Intl. Conf. on Princ. of Knowledge Representation and Reasoning,* pp. 325-334. Vivid reasoning for epistemic logics, by explicitly constructing Kripke models.

Joseph Halpern and L.D. Zuck (1987). "A little knowledge goes a long way: Simple knowledge-based derivations and correctness proofs for a family of protocols." *Proc. 6th ACM Symposium on Principles of Distributed Computing.* pp. 269-280.

Hans Herzberger (1982). "Notes on Naive Semantics." *Journal of Philosophical Logic,* vol. 11, no. 1, pp. 61-102. Solution to paradox of self-reference.

Jaako Hintikka. (1962). *Knowledge and Belief.* Cornell University Press, Ithaca, NY. The first extensive study of modal logics for knowledge and belief.

Jaako Hintikka (1969). "Semantics for Propositional Attitudes." In L. Linsky, ed., *Reference and Modality,* Oxford University Press, pp. 145-167. Application of Kripke's possible worlds semantics to epistemic logic.

Jaako Hintikka (1986). "Reasoning about Knowledge in Philosphy: The Paradigm of Epistemic Logic." In (Halpern, 86).

G.E. Hughes and M.J. Cresswell (1968). *An Introduction to Modal Logic.* Methuen and Co., London. Basic textbook for modal logics of all kinds.

Leslie P. Kaelbling (1986). "An Architecture for Intelligent Reactive Systems." In M. Georgeff and A. Lansky, *Reasoning about Actions and Plans: Proc. 1986 Workshop, Timberline, Oregon.* Morgan Kaufmann, San Mateo, CA. Application of epistemic logic to the creation of a hardware design for an intelligent reactive system.

David Kaplan and Richard Montague. (1960). "A Paradox Regained." *Notre Dame Journal of Formal Logic,* vol. 1, no. 3, pp. 79-90. Discussion of the "Knower" paradox in syntactic theories of knowledge.

David Kaplan. (1968). "Quantifying In." *Synthese.* Discussion of quantifying into modal contexts.

Kurt Konolige (1985). "Belief and Incompleteness." In J. Hobbs and R. Moore (eds). *Formal Theories of the Commonsense World.* Ablex Pubs., Norwood, NJ. An epistemic theory that avoids the assumption of consequential closure, using a characterization of an incomplete inference engine and a syntactic representation of knowledge.

Kurt Konolige (1986). "What Awareness Isn't: A Sentential View of Implicit and Explicit Belief." In (Halpern, 86).

Kurt Konolige (1988). "On the Relation between Default and Autoepistemic Logic." *Artificial Intelligence.* vol. 35, pp. 343-382.

Saul Kripke. (1963). "Semantical considerations on modal logic." *Acta Philosophica Fennica, Modal and Many-Valued Logics* 1963, pp. 83-94; Reprinted in L. Linsky, ed., *Reference and Modality.* Oxford University Press, London, 1971, pp. 63-72. Introduction of the possible worlds semantics for modal logics.

Saul Kripke. (1972). "Naming and Necessity." In D. Davidson and G. Harmon, eds., *Semantics of Natural Language.* D. Reidel Publishing Co., Dordrecht, Holland, 1972, pp. 253-355; Published as a book by Harvard University Press, 1980. Discussion of rigid designators and necessary truths.

Saul Kripke. (1975). "Outline of a Theory of Truth." *Journal of Philosophy,* Vol. 72, pp. 690-716. Solution to the Liar's Paradox.

R.E. Ladner (1977). "The computational complexity of provability in systems of modal propositional logic." *SIAM Journal on Computing.* vol. 6, no. 3, pp. 467-480.

W. Lenzen (1978). "Recent Work in Epistemic Logic." *Acta Philosophica Fennica.* vol. 30, pp. 1-219.

Hector Levesque (1984) "A Logic of Explicit and Implicit Belief." *Proc. AAAI-84* pp. 198-202.

Hector Levesque (1987). "All I know: An abridged report." *Proc. AAAI-87* pp. 426-431. Autoepistemic inference.

Fangzhen Lin and Yoav Shoham (1992). "A Logic of Knowledge and Justified Assumptions." *Artificial Intelligence.* vol 57, pp. 271-289. Unification of default logic with auto-epistemic logic.

Benson Mates. (1972). *Elementary Logic.* Oxford University Press, 1972. Introductory logic textbook. Readable and clear.

John McCarthy and Patrick Hayes. (1969). "Some Philosophical Problems from the Standpoint of Artificial Intelligence." In B. Meltzer and D. Michie, eds., *Machine Intelligence 4,* Edinburgh University Press, Edinburgh, 1969, pp. 463-502. Reprinted in Matthew L. Ginsberg (ed.) *Readings in Nonmonotonic Reasoning.* Morgan Kaufmann, San Mateo, 1987. Seminal paper in knowledge representation, with some discussion of epistemic representations and the knowledge precondition problem for actions.

P. Milgrom. (1981) "An axiomatic characterization of common knowledge." *Econometrica* vol. 49, no. 1 pp. 219-222.

P. Milgrom and N. Stokey (1982). "Information, trade, and common knowledge." *Journal of Economic Theory,* vol. 26, no. 1, pp. 17-26.

Robert Moore. (1980). "Reasoning about Knowledge and Action." Technical Note 191, SRI International, Menlo Park, CA. Reasoning about knowledge by translating modal language into a first-order possible worlds representation. Analysis of the knowledge preconditions problem.

Robert Moore and Gary Hendrix. (1980). "Computational Models of Belief and the Semantics of Belief Sentences." Technical Note 187, SRI International, Menlo Park, CA. Modelling belief by membership in a knowledge base.

Robert Moore. (1985). "A Formal Theory of Knowledge and Action." In J. Hobbs and R. Moore (eds). *Formal Theories of the Commonsense World.* Ablex Pubs., Norwood, NJ. Summary and extension of (Moore, 1980).

Robert Moore (1985). "Semantical Considerations on Nonmonotonic Logic." *Artificial Intelligence,* Vol. 25, pp. 75-94. Auto-epistemic logic.

Leora Morgenstern (1987). "Knowledge Preconditions for Actions and Plans." *Proc. IJCAI-87,* pp. 867-874. Knowledge preconditions problem for multiagent plans using a syntactic theory.

Leora Morgenstern (1988). "Foundations of a Logic of Knowledge, Action, and Communication." NYU Ph.D. Thesis. Syntactic theory of knowledge and its applications to reasoning about plans and communication.

Leora Morgenstern (1991). "Knowledge and the Frame Problem." In K. Ford and P. Hayes, (eds), *Reasoning Agents in a Dynamic World: The Frame Problem.* JAI Press, Greenwich. Dealing with the additional complexities of the frame problem in environments where agents may be ignorant of the events that have taken place.

Yoram Moses (ed.) (1992). *Prof. Fourth Conference on Theoretical Aspects of Reasoning About Knowledge.* Morgan Kaufmann, San Mateo, CA. See (Halpern, 86).

Yoram Moses and Cynthia Dwork (1986). "Knowledge and Commonse Knowledge in Byzantine Environments I: Crash Failures." In (Halpern, 86). Application of epistemic theory to distributed systems.

Yoram Moses and Yoav Shoham (1992). "Belief as Defeasible Knowledge." *Artificial Intelligence,* ??. Belief defined as uncertain knowledge.

Madhura Nirkhe, Sarit Kraus, and Donald Perlis (1993). "Situated Reasoning within Tight Deadlines and Realistic Space and Computation Bounds." *2nd symposium on Logical Formalizations of Commonsense Reasoning.*

Rohit Parikh (1984). "Logics of Knowledge, Games, and Dynamic Logic." *FST-TCS,* Springer Lecture Notes on Computer Science #181, pp. 202-222.

Rohit Parikh (1987). "Knowledge and the Problem of Logical Omniscience." *ISMIS-87*, North Holland, pp. 432-439.

Rohit Parikh (ed.) (1990). *Prof. Third Conference on Theoretical Aspects of Reasoning About Knowledge.* Morgan Kaufmann, San Mateo, CA. See (Halpern, 86).

Rohit Parikh (1990a). "Recent Issues in Reasoning about Knowledge." In (Parikh, 1990) pp. 3-10. Tutorial.

Rohit Parikh and R. Ramanujam (1985). "Distributed processing and the logic of knowedge." *Proceedings of the Brooklyn College Workshop on Logics of Programs.* pp. 256-268.

Donald Perlis. (1985). "Language with Self-Reference I: Foundations." *Artificial Intelligence,* vol. 25, pp. 301-322. Avoiding paradoxes of self-reference in syntactic theories.

Plato. *Theaetetus.* In *The Dialogues of Plato,* tr. Benjamin Jowett, Macmillan, 1892. The classic philosophical discussion of the definition of knowledge.

Stanley Rosenschein and Leslie Kaelbling. (1986). "The Synthesis of Digital Machines with Provable Epistemic Properties." In (Halpern, 86). Application of epistemic logic to the creation of a hardware design for an intelligent reactive system.

Bertrand Russell (1948). *Human Knowledge: Its Scope and Limits.* Simon and Schuster. Study of the foundations of knowledge.

Gilbert Ryle. (1949). *The Concept of Mind.* Barnes and Noble. An alternative approach to epistemology. Ryle arguets "knowing how to do A" is more fundamental than "knowing that P."

R. Scherl and H. Levesque (1993). "The Frame Problem and Knowledge Producing Actions." *Proc. AAAI-93*, to appear. Formulating suitable frame axioms governing agents' knowledge.

Yoav Shoham (1986). "Chronological Ignorance: Time, Non-monotonicity, Necessity, and Causal Theories." *Proc. AAAI-86.* Application of epistemic logic to temporal reasoning.

Raymond Smullyan. (1978). *What is the name of this book?* Prentice Hall. One of a series of popular books by this author exploring paradoxes of self-reference and related issues.

Alfred Tarski (1956). "The Concept of Truth in Formalized Languages." In *Logic, Science, and Metamathematics.* Oxford University Press. Syntactic theories and a solution to the Liar paradox.

Richmond Thomason (1980). "A Note on Syntactic Treatments of Modality." *Synthese.* vol. 44 pp. 391-395.

Moshe Vardi (ed.) (1988). *Prof. Second Conference on Theoretical Aspects of Reasoning About Knowledge.* Morgan Kaufmann, San Mateo, CA. See (Halpern, 86).