Agency and Interaction What We Are and What We Do in Formal Epistemology

Jeffrey Helzner
Department of Philosophy
Columbia University

Vincent Hendricks
Department of Philosophy
University of Copenhagen

Abstract

Formal epistemology is the study of crucial concepts in general or mainstream epistemology including knowledge, belief (-change), certainty, rationality, reasoning, decision, justification, learning, agent interaction and information processing using a spread of different formal tools. These formal tools may be drawn from fields such as logic, probability theory, game theory, decision theory, formal learning theory, and distributed computing – such variety is typical in formal epistemology, a field in which interaction with topics outside of philosophy proper is the rule rather than the exception. Practitioners of formal epistemology include philosophers, computer scientists, social scientists, cognitive psychologists, theoretical economists, mathematicians, and theoretical linguists. The interdisciplinary nature of formal epistemology can make it difficult for those new to the field to have a sense of some of its basic agendas, actors, and issues. What follows is a breezy overview of formal epistemology as organized around notions of *agency* and *interaction*.

1 Introduction

Formal epistemology is a recent field of study in formal philosophy dating back only a decade or so (Hendricks 2006), (Helzner & Hendricks 2010), (Arlo-Costa et al. 2011), (Hansson & Hendricks 2001). This is not to say that formal epistemological studies had not been conducted prior to the late 1990's, but rather that the term introduced to cover the philosophical enterprise was coined around this time. Predecessors to the field may be found in the works of Carnap, Hintikka, Levi, Lewis, Putnam, Quine and other high-ranking officials in formal philosophy.

The point of departure of this essay is rooted in two philosophically fundamental and interrelated notions central to formal epistemology;

- *agency* – what agents are, and
- *interaction* – what agents do.

Agents may be individuals, or they may be groups of individuals working together. In each of the sections that follow, assumptions are made concerning the relevant features of the agents at issue. For example, such relevant features may include the agent's beliefs about its environment, its desires concerning various possibilities, the methods it employs in learning about its environment, and the strategies it adopts in its interactions with other agents in its environment. Fixing these features serves to bound investigations concerning interactions between the agent and its environment. The agent's beliefs and desires are assumed to inform its decisions. Methods employed by the agent for the purposes of learning are assumed to track or approximate or converge upon the facts of the agent's environment. Strategies adopted by the agent are assumed to be effective in some sense.

We believe that agency and interaction provide the basis of useful framework in which to understand much of what counts as formal epistemology. In what follows we will attempt to locate predominant paradigms – e.g., epistemic logic, interactive epistemology and game theory, formal learning theory, belief revision theory, probability theory, and decision theory – within such a framework.

2 Epistemic Logic

Epistemic logic started with the study of proper axiomatizations for knowledge, belief, certainty and other epistemic attitudes. Hintikka inaugurated the field with his seminal book (Hintikka 1962) which focuses on axiomatizing knowledge and belief in mainly mono-agent systems. Agents are syntactically represented as indices on epistemic operators in a formal logical language. From the semantic perspective, to be an agent is basically to be an index on an accessibility relation between possible worlds representing the epistemic alternatives over which the agent has to succeed in order to know some proposition (interesting alternative semantics to Kripke semantics have been developed by (Arlo-Costa & Pacuit 2006), (Baltag & Moss 2004) and others. Not much epistemic pertinence is derived from this construction in and by itself. Surely an epistemic logic may be both sound and complete but no epistemological role is really left for the agent to play on this conception, so agency is a rather thin concept early on. It is all about proper axiomatizations (Hendricks & Symons 2005), (Stalnaker 2006).

This neglect did not matter much to many practitioners of epistemic logic at the time, proofs of important (meta-)logical properties for these new logics sufficed. Thus the field was living an isolated life quite remote from the concerns of mainstream epistemology. Hintikka himself (and a few others like Lenzen (Lenzen 1978)) was a notable exception to this general attitude of epistemic logicians and insisted on telling a better story, not about what agents are in the logical language, but about what they do. Accordingly, Hintikka took axioms of epistemic logic to describe a certain kind of strong rationality much in sync with the autoepistemological tradition of G.E. Moore and especially Norman Malcolm. Axioms of epistemic logic are really prescriptions of rationality in mono-agent systems. Epistemic logic has since been used address a number of important philosophical problems including for instance the Fitch Paradox (Brogaard & Salerno 2009), the problem of logical omniscience (Duc 1997), (Parikh 2005), and various conceptual characterizations of knowledge and other epistemic attitudes (Kraus & Lehmann 1988).[1]

But rationality considerations are not only central to the singular agent acting in some environment, call it nature, but likewise, and perhaps especially, central to agents when in presence of other agents and interacting with these. Thus mono-agent systems had to be extended to multi-modal systems in order to get both agency and interaction off the epistemological ground for real.

3 Dynamic Epistemic Logic

A sea-change took place in epistemic logic in the late 1980's and the beginning of the 1990's especially due to the work of Joseph Halpern and his collaborators (Fagin 1995) and others (Meyer & Hoek 1995). Multiple agents were introduced into the logical language which, along with multiple epistemic accessibility relations on the semantic level, gave rise to a precise and adequate representation of the flow of information through an agent system, together with the nature of various protocols governing

[1]For solid overviews refer to (De Bruin 2008), (Gochet & Gribomont 2006)

such systems. In this setting, possible worlds are to be understood as the states of the system taken as a whole, or sometimes the possible histories or consecutive runs of the system as a whole, that are compatible with the state transition directives which rule the system. Stalnaker has recently summarized the consequences of this sea-change precisely:

> The general lesson I drew from this work was that it was useful for epistemology to think of communities of knowers, exchanging information and interacting with the world, as (analogous to) distributed computer systems. (Hendricks & Roy 2010): 78

Agent systems can now be thought of as encompassing everything from a group of robots on an assembly line to a group of poker players in Texas Hold 'Em. In turn, there is much more to what agents are nowadays, but also much more to what they do dynamically (as opposed to statically in terms of, say, epistemic axioms describing the rationality of single agents). Dynamic epistemic logic is a rich blend of studies ranging multi-agent axiomatizations of knowledge, belief, common knowledge and belief (Barwise 1988) certainty, uncertainty, doubt, ignorance and a host of other epistemic attitudes; models of the interplay between knowledge and games (Benthem 2001), (Benthem 2007), knowledge and justification in mainstream epistemology (Artemov & Nogina 2005), social software (Parikh 2002), knowledge and public announcement of information (Baltag et al. 2002), knowledge intertwined with preferences, actions and decisions (Ditmarsch et al. 2008); learning (Hendricks 2001), (Gierasimczuk 2009), belief revision (Baltag & Smets 2008), models of agent interaction in multi-agent systems; combined multi-agent and multi-modal systems in which for instance the development of knowledge over time may be scrutinized (Kraus & Lehmann 1988), relations between knowledge and deontic commitments investigated, divisions of cognitive labor modeled and so forth (for epistemic logic paired up with mainstream epistemological concerns, refer to (Hendricks 2006), (Williamson 2006), (Hendricks & Pritchard 2007), (Hendricks & Roy 2010)).

4 Interactive Epistemology

Theoretical economics is to a significant extent about understanding, anticipating and modeling phenomena like trading, stock speculation, real-estate dealing, hostile company take-overs, shareholding and so forth. Obviously, agency and interaction play a paramount role here. Independently, but informed by the developments in epistemic logic, economists have used game theory to scrutinize an extensive spread of the mentioned phenomena. By way of example, in 1976 the later Nobel Prize Laureate Robert Aumann published his famous Agreement Theorem in "Agreeing to Disagree" in which he describes conditions under which two "like minded" agents or players cannot "agree to disagree" in the sense that if the two players' posteriors of some event are common knowledge then they must coincide. In other words, in order to make trade possible, agents have to agree to disagree (Aumann 1976). That is agency in terms of players, interaction in terms of games.

On the way to this result Aumann made a host of assumptions about the nature knowledge much in tune with what is to be found in epistemic logic like the axiomatic strength of knowledge in order to infer the backwards induction equilibrium and assumptions about what is common knowledge among the players. In 1999, Aumann coined a term for these kinds of study in theoretical economics: "Interactive epistemology" (Aumann 1999). It denotes an epistemic program studying shared knowledge and belief given more than one agent or player in an environment and has, as already suggested, strong ties to game theoretic reasoning and questions of common knowledge and belief, backward induction, various forms of game equilibria and strategies in games, (im)perfect information games, (bounded) rationality etc (Stalnaker 2006), (Aumann & Brandenburger 1995).

Given its inauguration with Aumann, the program was in the beginning dominated by scholars drawn from theoretical economics and computer science rather than philosophy and logic, but recently philosophers and logicians have begun to pay close attention to what is going on in this striving program of formal epistemology. And for good reason too; the interactive epistemological approach to agency and interaction have close shaves with the major focal points in dynamic epistemic logic and much of the technical machinery is a common toolbox for both paradigms (Brandenburger 2007), (Hendricks & Hansen 2008).

5 Computational Epistemology

In epistemic logic, the knowledge of agents may be axiomatically characterized, in dynamic epistemic logic axiomatizations are extended to multi-agent, multi-modal systems and in interactive epistemology agents are players in a game against nature or other players. But agents may also be learners of knowledge and information as also modeled using the tools drawn from dynamic epistemic logic.

A different set of tools of the trade for learnability studies is provided by formal learning theory. Learning theory stems from computability theory rather than philosophical logic or game theory. Strictly speaking, formal learning theory, or "computational epistemology" as Kelly (Kelly 2000) recently dubbed the field, is not about knowledge but about learning, but learning is again about knowledge acquisition.

In computational epistemology agents are learning functions and the theory begins with the problem of finding true or empirically adequate, general theories from an ongoing stream of particular, empirical data (Schulte 2008). The interaction is given by the agents' responses to the empirical data in their ongoing attempts to formulate general empirical theories about some aspect of the world under scientific investigation. The basic idea is accordingly to seek epistemic justification in terms of truth-finding performance, rather than in terms of axiomatics. For example, one of the first publications in the area ((Putnam 1963)) involved a computational critique of the learning power of Carnap's confirmation theory. After that auspicious beginning, the field was developed mainly by mathematicians and computer scientists interested in the foundations of machine learning until the late 1980's, when Glymour (Glymour 1991), Kelly (Kelly 1996), Schulte (Schulte 1999), Osherson (Jain et al. 1999), Weinstein, and others began again to apply it to more traditional, epistemological concerns. Such applications include explications of empirical underdetermination and simplicity, critiques of Bayesianism, belief revision (Martin & Osherson 1998), (Kelly 1999) and internal realism (Kelly 2004), and the justification of inductive inference, Ockham's razor and causal discovery (Kelly 2008). Recently, given the work of (Hendricks 2001), (Hendricks 2006), Baltag, Smets (forthcoming) and others, computational epistemology has begun to interact more directly with the concerns of dynamic epistemic logic in terms of the learnability of epistemic axioms for different agents with different methodological recommendations of (hopefully) truth-tracking conduct and the division of epistemic labour.

6 Probability and Credence

Following [Levi 80], we assume that the agent is, at each point in time, committed to full belief in some set of propositions concerning its environment. Where the agent is not committed to full belief in a given proposition, the negation of that proposition is a serious possibility for the agent. The agent may judge some serious possibilities to be more probable than others. What can be said about these judgments? The received view, following a tradition that goes back to the work of Ramsey (Ramsey 1931), maintains that such credal judgments ought to be representable by a probability measure. This view has been criticized as being too weak and as being too strong. As for being too weak, the simple requirement that such judgments be representable by a probability measure says little about the extent to which these subjective probabilities should approximate objective probabilities, e.g., limiting frequencies in the sense of [von Mises 57] or perhaps even propensities in the sense of [Popper 59]. Various principles have been offered in order to require that the subjective probabilities of a rational agent are informed by that agent's knowledge of objective probabilities – readers may wish to consult [Kyburg 74], [Levi 78], and [Lewis 80] for important discussions along these lines. As for being too strong, requiring credal judgments to be representable by a probability measure implies, among other things, that such credal judgments are complete. That is, a consequence of such a requirement is that, for any given pair of serious possibilities, the agent either judges one of the possibilities to be more probable than the other or the agent regards the possibilities as being equally probable. Thus, the requirement bars situations in which the agent, because of a lack of information, is unable to supply such a judgment. Such considerations, which to some extent echo earlier, related concerns of [Keynes 21] and [Knight 21], have motivated some people – e.g., [Kyburg 68], [Levi 74], and [Walley 90] – to consider indeterminate probabilities, either in the form of interval-valued measures or sets of traditional measures, in representing rational credences.

7 Probabilism and the Status of Full Belief

What sorts of internal states are essential to the agent's representation of its environment? Notions of full belief – e.g., according to which the agent simply believes the proposition, in contrast to believing the proposition to some degree – are a traditional interest within mainstream epistemology. Some philosophers, e.g. [Jeffrey 92], have argued in favor of a doctrine known as *radical probabilism*. A central tenet of this doctrine is that various propositional attitudes of epistemic interest, especially full belief, are reducible to credal judgments. There are several ways that one might attempt such a reduction. Perhaps the most obvious is to identify full belief with maximal partial belief. For example, if we assume that the agent's credal state can be represented by a probability measure, then such a reduction would identify those propositions that are fully believed by the agent with those propositions that have maximal probability according to this representing measure. Following this proposal, it would seem that a proposition counts as a serious possibility for the agent just in case its negation is not assigned maximal probability according to the probability measure representing the agent's credal judgments. Hence, by the probability axioms, a proposition counts as seriously possible for the agent just in case it has nonzero probability under the representing measure. This leads to certain difficulties. For example, if the agent is concerned to estimate the height of an object that is sufficiently distant, then the agent might regard a continuum of values as possible – e.g., the height of the object is judged to be between three and four feet. According to the suggested reduction, such a continuum of possible values for the height of the object could not serve as a set of serious possibilities, since it is a mathematical fact that no probability measure can distribute positive probability to each element of such a continuum. The interested reader is urged to consult [van Fraassen 95] and [Arlo-Costa 01] for more sophisticated versions of probabilism.

8 Decision Theory

An agent interacts with its environment through the choices it makes. Choice presupposes alternatives, and a theory of rational choice should, at least, distinguish some of the available alternatives as admissible. As an example, consider those accounts of rational choice that are built on the concept of preference. One such account assumes that the agent has complete and transitive preferences over the set of available alternatives. Those alternatives that are optimal with respect to the given preference ranking are taken as admissible. This abstract preference-based account says nothing about the way in which preferences are informed by the agent's beliefs about its environment. Subjective expected utility theory [SEU], which is at the center of modern-day decision theory, provides significantly more detail than the abstract theory of preference optimization. SEU assumes that alternatives are acts, which, following Savage's classic formulation of SEU in [Savage 72], are functions from states to consequences. Drawing upon the earlier work of Ramsey (1931) on subjective probability and the work of [von Neumann and Morgenstern 47] on utility, Savage provides conditions on the agent's preferences over acts that guarantee the existence of a probability measure p and a utility function u such that the agent's preferences can be regarded as if they were the result of maximizing utility u with respect to probability p. According to the intended interpretation, the probability measure p represents the agent's degrees of belief concerning the possible states and the utility function u represents the extent to which the agent values the possible consequences.

The assumptions of SEU may be questioned in various ways. We focus on two ways that have generated significant interest among philosophers. First, why should it be that the rational agent's degrees of belief can be represented by a probability distribution p? As noted in Section 6, it is not clear why such an assumption should obtain in cases where the agent has very little information concerning the possible states. Second, in SEU it is assumed that the agent's subjective probability concerning the states is independent of the act that is chosen. Some question this assumption and offer examples in which a modification of SEU that provides for such dependencies, through the use of conditional probabilities, is supposed to give an irrational recommendation. The first line of questioning has led some – e.g., [Ellsberg 61], [Levi 74], and [Gardenfors and Sahlin 82] – to use indeterminate probabilities in their normative accounts of decision making under uncertainty. The second line of questioning has led some – e.g., [Gibbard and Harper 78], [Lewis 81], and [Joyce 99] – to investigate causal decision theory.

9 Belief Revision

An agent has beliefs about the environment with which it interacts. Sometimes these interactions are such that the agent, on pain of irrationality, must revise its beliefs. The classic example is that of a scientific agent who has beliefs about the world that might need to be revised in light of new data. The study of this sort of example has a long history in the philosophy of science, where it is often discussed at a relatively informal level in connection with topics such as underdetermination. In the context of formal epistemology, the study of belief revision has been generalized to include various sorts of epistemic agents. Questions such as the following suggest the range of theoretical options that are available in connection with such investigations:

How are the potential belief states to be interpreted? One might take the belief states to represent partial beliefs; e.g., the agent has a certain degree of belief in proposition P. Alternatively, one might be interested in states of full belief; e.g., the agent fully believes P. Further refinements have been considered. For example, one might consider those full beliefs with respect to which the agent manifests some level of awareness; e.g., I am aware of my belief that I am presently writing the words of this sentence. In contrast to a focus on conscious beliefs, one might consider those propositions that the agent is committed to fully believing; e.g., all of those propositions that are deducible from my conscious beliefs.

How are the potential belief states to be represented? The answers to this question depend, at least to some extent, on how the previous question is answered. For example, if partial beliefs are the issue, then probability distributions might be taken as the basis for the representation so that a potential belief state is represented as a probability measure over the possible states of nature. On the other hand, if full belief is the issue, then one might specify a suitably formalized language and represent each potential belief state as a subset of the language so that membership in the set indicates full belief.

How are revisions interpreted? If credal states are the concern, then modifications of the credal state might be understood in terms of something like conditionalization. The interested reader is urged to consult [Halpern 03] for a survey of various proposals concerning the representation and modification of credal states. What about revising states of full belief? When an instance of belief revision concerning full beliefs is the result of the agent selecting from a set of (full) belief states that the agent recognizes as potential alternatives, then such an instance may be regarded as the resolution of a decision problem. Isaac Levi has developed a decision-theoretic approach to belief change; important discussions of Levi's approach include [Levi 80], which considers belief change in the context Levi's general approach to epistemology, and [Arlo-Costa and Levi 06], which gives greater emphasis to the formal details concerning Levi's approach. Different connections between choice and belief revision are emphasized in [Rott 93]. Rott demonstrates an important correspondence between the "AGM" account of belief revision offered in [Alchourron et al. 85] and the economists' study of rational choice functions. Finally, it is worth noting that where both partial and full beliefs are considered, there may be significant dependencies between the modification of these two sorts of belief states. For example, if the credal judgments of rational agents are a function of their judgments of full belief, as some philosophers assume, then changes to the latter may result in changes to the former.

10 To Be and To Do

By organizing the discussion around the issues of *what agents are* (e.g., things that have beliefs, desires, methods, and strategies) and *what agents do* (e.g., make decisions, learn, and play games) an attempt has been made to survey the various topics that are studied in formal epistemology while also giving some sense of unity across an apparently diverse set of topics. Whether or not the reader finds the proposed way of dividing the subject illuminating, there are substantive questions that remain concerning the nature of formal epistemology. We now consider some of these questions.

Do agents of the sort considered above exist? If not, then in what sense is formal epistemology relevant? For the most part, it is difficult to maintain that the the accounts considered above are descriptive. Human beings are not logically omniscient, their credences typically fail to satisfy the probability axioms, and their decisions often violate standard norms of rational choice. Yet despite this descriptive inadequacy, such accounts can be applied to "human agents" in at least two ways. First, by providing a benchmark for interventions that are designed to improve the human agent. This same sort of virtue might extend to macro-level interventions that are intended to improve a system of interacting agents. Second, even if such accounts are descriptively inadequate at the level of individual agents, it still might be the case that macro phenomena concerning a system of such ideal agents provides a descriptively adequate account of the macro phenomena concerning a system of human agents. Of course, such "as if" interpretations have been very influential in economics. Note that the previous question could have just as well been asked of mainstream epistemology, and we suspect that the response on behalf of mainstream epistemology would not be substantially different from the normative one given above on behalf of formal epistemology; the distance between logic and psychology might not end up being as Frege had hoped, but a cursory examination of mainstream epistemology suggests that some distance remains.

So what is it that distinguishes formal epistemology from mainstream epistemology? In particular, what is formal about formal epistemology? It is worth noting that the term 'formal' had a more restricted meaning in the context of work on the foundations of mathematics, most clearly in connection with Hilbert's program. Many who now identify themselves as formal epistemologists either began their career by working in these areas or studied with those who did. Formalization in the context of Hilbert's program was understood in terms of ideas that are now taken as fundamental in the study of computation. Roughly, a concept is formal in this sense just in case there is an algorithm that can decide whether or not the concept applies. In this sense, the concept *is a proof in first-order logic* is formal. It seems clear that this formal-as-computable conception is not what distinguishes formal epistemology, since only a small proportion of the work is formal in the sense just described. Rather, the concepts of formal epistemology (e.g. *satisfies Savage's axioms*) are well-defined in the sense of ordinary mathematics. That is, the formalism serves to pick out a class of set-based structures, essentially the "models" of the formalism. By satisfying this standard there is, at least relative to the universe of sets, a definite 'yes' or 'no' answer to each question concerning whether or not a particular concept applies to a given thing. Fixing the boundaries of concepts in this way can be helpful in avoiding the sort of confusion that can result when several distinct concepts share a common label.

11 Acknowledgements

We are indebted to an anonymous referee for pertinent suggestions improving the presentation of formal epistemology and for pointing to additional references to be cited.

References

[Alchourron et al. 85] Alchourron, C., Gardenfors, P. & Makinson, D. (1985). "On the Logic of Theory Change: Partial Meet Functions for Contraction and Revision", *Journal of Symbolic Logic*, 50:510-530.

[Arlo-Costa 2001] Arlo-Costa, H. (2001). "Bayesian Epistemology and Epistemic Conditionals: On the Status of the Export-Import Laws", *The Journal of Philosophy*, 98(11): 555-593.

[Arlo-Costa & Levi 2006] Arlo-Costa, H. & Levi, I. (2006). "Contraction: On the Decision-Theoretical Origins of Minimal Change and Entrenchment", *Synthese*, 152,1:129-154.

[Arlo-Costa & Pacuit 2006] Arlo-Costa, H. & Pacuit, E. (2006). ”First Order Classical Modal Logic”, *Studia Logica*, 84, 2: 171-210.

[Arlo-Costa et al. 2011] Arlo-Costa, H., van Benthem, J.v. & Hendricks, V.F. (2011) (eds.). *Readings in Formal Epistemology*. Cambridge: Cambridge University Press.

[Artemov & Nogina 2005] S. Artemov and E. Nogina (2005). ”Introducing Justification into Epistemic Logic”, *Journal of Logic and Computation* 15: 1059-1073.

[Aumann 1976] Aumann, R. (1976). ”Agreeing to Disagree”, *The Annals of Statistics*, 4(6): 1236-1239.

[Aumann 1999] Aumann, R. (1999). ”Interactive Epistemology I”, *International Journal of Game Theory*, vol. 28(3): 263-300.

[Aumann & Brandenburger 1995] Aumann, R. & Brandenburger, A. (1995). ”Epistemic Conditions for Nash Equilibrium”, *Econometrica*, 63(5): 1161-1180.

[Baltag et al. 2002] Baltag, A., Moss, L.S., Solecki, S. (2002). ”The Logic of Public Annoucements, Common Knowledge, and Private Suspicion”. *Proceedings of TARK 1998*. Los Altos: Morgan Kaufmann Publishers: 43-56.

[Baltag & Moss 2004] Baltag, A. & Moss, L. (2004). ”Logics for Epistemic Programs”, *Synthese / Knowledge, Rationality and Action*, 139: 165–224.

[Baltag & Smets 2008] Baltag, A. & Smets, S. (2008). ”A Qualitative Theory of Dynamic Interactive Belief Revision”, in *Logic and the Foundation of Game and Decision Theory* LOFT7 (2008), Volume: 3. Amsterdam: Amsterdam University Press: 13-60.

[Barwise 1988] Barwise, J. (1988). ”Three Theories of Common Knowledge”, in *Proceedings of the 2nd conference on Theoretical aspects of Reasoning about Knowledge*. Pacific Grove: California: 365-379.

[Benthem 2001] Benthem, J.v. (2001). ”Games in Dynamic Epistemic Logic”, *Bulletin of Economic Research* 53:4, 219-24.

[Benthem 2007] Benthem, J.v. (2007). ”Logic Games, From Tools to Models of Interaction”, in A. Gupta, R. Parikh & J. van Benthem, (eds.), *Logic at the Crossroads*. Mumbai: Allied Publishers, 283-317.

[Brandenburger 2007] Brandenburger, A. (2007). ”The power of paradox: some recent developments in interactive epistemology”, *International Journal of Game Theory*, 35(4): 465-492.

[Brogaard & Salerno 2009] Brogaard, B. & Salerno, J. (2009). ”Fitch’s Paradox of Knowability”, *The Stanford Encyclopedia of Philosophy.* Principal editor Edward Zalta. Palo Alto: Stanford University.

[De Bruin 2008] De Bruin, B. (2008). ”Epistemic Logic and Epistemology”, in *New*

Waves in Epistemology, edited by Hendricks, V.F. & Prichard, D. London: Palgrave Macmillan.

[Ditmarsch et al. 2008] van Ditmarsch, H., Hoek, W.v.d. & Kooi, B. (2008). *Dynamic Epistemic Logic*. Dordrecht: Springer.

[Duc 1997] Duc, H.N. (1997). ”Reasoning about Rational, but not Logically Omniscient Agents”, *Journal of Logic and Computation* 7: 633-648.

[Ellsberg 1961] Ellsberg, D. (1961). ”Risk, ambiguity, and the savage axioms”, *The Quarterly Journal of Economics*, 75:643-669.

[Fagin 1995] Fagin, R., Halpern, J., Moses, Y. & Vardi (1995). *Reasoning about Knowledge*. Cambridge: MIT Press.

[Gardenfors & Sahlin 1982] Gardenfors, P. & Sahlin, N.E. (1982) ”Unreliable probabilities, risk taking, and decision making”, *Synthese*, 53: 361-386.

[Gibbard & Harper 1978] Gibbard, A. & Harper, W. (1978) “Counterfactuals and Two Kinds of Expected Utility”, in C.A. Hooker, J.J. Leach, & E.F. McClennen (eds.), *Foundations and Applications of Decision Theory*. Dordrecht: Reidel, 125-162.

[Glymour 1991] Glymour, C. (1991). ”The Hierarchies of Knowledge and the Mathematics of Discovery”, *Minds and Machines* 1: 75-95.

[Gochet & Gribomont 2006] Gochet, P. and Gribomont, P. (2006). ”Epistemic Logic” in *Logic* *and the Modalities in the Twentieth Century. Handbook of the History of Logic*. Amsterdam: Elsevier: 99-185.

[Gierasimczuk 2009] Gierasimczuk, N. (2009). "Learning by Erasing in Dynamic Epistemic Logic", *Lecture Notes in Computer Science*, vol. 5457: 362-373.

[Hansson & Hendricks 2001] Hansson, S.O. & Hendricks, V. F. (2011) (eds.). *The Handbook of Formal Philosophy*. Dordrecht: Springer.

[Halpern 2003] Halpern, J.Y. (2003). Reasoning about Uncertainty. Cambridge: MIT Press.

[Helzner & Hendricks 2010] Helzner, J & Hendricks, V. F. (2010). "Formal Epistemology," forthcoming in *Oxford Bibliography Online*.

[Hendricks 2001] Hendricks, V. F. (2001). *The Convergence of Scientific Knowledge – A View from the Limit*. Dordrecht: Springer.

[Hendricks 2001] Hendricks, V.F. (2003). "Active Agents", *Journal of Logic, Language and Information* 12: 469-495.

[Hendricks 2006] Hendricks, V.F. (2006). *Mainstream and Formal Epistemology*. New York: Cambridge University Press.

[Hendricks & Symons 2005] Hendricks, V.F. & Symons, J. (2005). "Epistemic Logic", in *Stanford Encyclopdia of Philosophy*.

Principal editor Edward Zalta. Palo Alto: Stanford University.

[Hendricks & Pritchard 2007] Hendricks, V. F. & Pritchard, D. (2007) (eds.). *Epistemology: 5 Questions*. New York / London: Automatic Press / VIP.

[Hendricks & Hansen 2008] Hendricks, V. F. & Hansen, P.G. (2008) (eds.). *Game Theory: 5 Questions*. New York / London: Automatic Press / VIP.

[Hendricks & Roy 2010] Hendricks, V. F. & Roy, O. (2010) (eds.). *Epistemic Logic: 5 Questions*. New York / London: Automatic Press / VIP.

[Hintikka 1962] Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Ithaca:

[Jeffrey 1992] Jeffrey, R. (1992) *Probability and the Art of Judgment*. New York: Cambridge University Press.

[Joyce 1999] Joyce, J. (1999) *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

[Kelly 1996] Kelly, K.T. (1996). *The Logic of Reliable Inquiry*. New York: Oxford University Press.

[Kelly 1999] Kelly, K. T. (1999). "Iterated Belief Revision, Reliability, and Inductive Amnesia", *Erkenntnis* 50: 11-58.

[Kelly 2000] Kelly, K. T. (2000). "The Logic of Success", *British Journal for the Philosophy of Science* 51,4: 639-660.

[Kelly 2004] Kelly, K. T. (2004). ”Learning Theory and Epistemology”, in *Handbook of Epistemology*, I. Niiniluoto, M. Sintonen, & J. Smolenski, eds. Dordrecht: Kluwer, 2004.

[Kelly 2008] Kelly, K.T. (2008) “Ockham’s Razor, Truth, and Information”, in *Handbook of the Philosophy of Information*, J. van Benthem and P. Adriaans, eds., Dordrecht: Elsevier.

[Keynes 1921] Keynes, J.M. (1921) *A Treatise on Probability*. MacMillan.

[Knight 1921] Knight,F.H. (1921) *Risk, Uncertainty and Profit*. Houghton-Mifflin.

[Kraus & Lehmann 1988] Kraus, S. & Lehmann, D. (1988). ”Knowledge, Belief and Time”, *Theoretical Computer Science* 58: 155-174.

[Kyburg 1968] Kyburg, H.E. (1968) ”Bets and Beliefs”, *American Philosophical Quarterly*. 5,1: 54-63.

[Kyburg 1974] Kyburg, H.E. (1974) *The Logical Foundations of Statistical Inference*. Dordrecht: Reidel.

[Lenzen 1978] Lenzen, W. (1978). Recent Work in Epistemic Logic, in *Acta Philosophica Fennica*, 30: 1-219.

[Levi 1974] Levi, I. (1974). ”On Indeterminate Probabilities”, *The Journal of Philosophy*, 71: 391-418.

[Levi 1977] Levi, I. (1977) ”Direct Inference”, *The Journal of Philosophy*, 74:5-29.

[Levi 1980] Levi, I. (1980) *The Enterprise of Knowledge*. Cambridge: MIT Press.

[Lewis 1980] Lewis, D. (1980). ”A Subjectivist’s Guide to Objective Chance”, in *Studies in Inductive Logic*, Vol II, ed. R. Jeffrey, Berkeley and Los Angeles: University of California Press.

[Lewis 1981] Lewis, D. (1981) ”Causal Decision Theory”, *Australasian Journal of Philosophy*, 59: 5-30.

[Martin & Osherson 1998] Martin, E. & Osherson, D. (1998). *Elements of Scientific Inquiry*. Cambridge: MIT Press.

[Meyer & Hoek 1995] Meyer, J.-J.Ch & Hoek, W. van der (1995). *Epistemic Logic for AI and Computer Science*. Cambridge Tracts in Theoretical Computer Science 41. Cambridge: Cambridge University Press.

[Jain et al. 1999] Jain, S., Osherson, D., Royer, J.S. & Sharma, A. (1999). *Systems That Learn - 2nd Edition: An Introduction to Learning Theory*. Cambridge: MIT Press.

[Parikh 2002] Parikh, R. (2002). “Social Software,” *Synthese*, 132: 187-211.

[Parikh 2005] Parikh, R. (2005). ”Logical omniscience and common knowledge: WHAT do we know and what do WE know?”, Proceedings of TARK05. Singapore: National University of Singapore: 62-77

[Popper 1959] Popper, K.R. (1959) ”The Propensity Interpretation of Probability”, *British Journal for the Philosophy of Science*, 10: 26-42.

[Putnam 1963] Putnam, H. (1963). ”Degree of Confirmation and Inductive Logic”, in *The Philosophy of Rudolf Carnap*, ed. P.A. Schilpp, La Salle, Ill: Open Court.

[Ramsey 1931] Ramsey, F.P. (1931). ”Truth and Probability”, in *The Foundations of Mathematics and other Logical Essays*, ed. R.B. Braithwaite, Routledge and Kegan Paul.

[Rott 1993] Rott, H. (1993). ”Belief Contraction in the Context of the General Theory of Rational Choice”, *The Journal of Symbolic Logic*. 58,4:1426-1450.

[Savage 1972] Savage, L.J. (1972) *The Foundations of Statistics*. New York: Dover. The Dover edition is a republication of the 1954 work.

[Schulte 1999] Schulte, O. (1999). ”Means-Ends Epistemology”, *The British Journal for the Philosophy of Science* 50: 1-31.

[Schulte 2008] Schulte, O. (2008). ”Formal Learning Theory” in *Stanford Encyclopedia of Philosophy*, principal editor Zalta, E. Palo Alto: Stanford University.

[Stalnaker 1996] Stalnaker, R. (1996). ”Knowledge, Belief and Counterfactual Reasoning in Games”, *Economics and Philosophy* 12: 133-163.

[Stalnaker 2006] Stalnaker, R. (2006). ”On Logics of Knowledge and Belief”, *Philosophical Studies*, 128, 1: 169-199.

[van Fraassen 1995] van Fraassen, B. (1995). ”Fine-Grained Opinion, Probability, and the Logic of Full Belief”, *Journal of Philosophical Logic*. 24,4:349-377.

[von Mises 1957] von Mises, R. (1957). *Probability, Statistics and Truth*, revised English edition, New York: Macmillan.

[von Neumann & Morgenstern 1947] von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

[Walley 1990] Walley, P. (1990) *Statistical Reasoning with Imprecise Probabilities*. New York: Chapman and Hall.

[Williamson 2006] Williamson, T. (2006). *Knowledge and Its Limits*. Oxford: Oxford University Press.