Logic and Formal Semantics for Epistemology

(For *The Routledge Companion to Epistemology*, Duncan Pritchard and Sven Bernecker eds.)

John Symons

Department of Philosophy

University of Texas, El Paso

Please note: This is a somewhat longer version of the chapter which will appear in print.

Introduction

This essay introduces some of the formal apparatus of epistemic logic and discusses its applicability to epistemological questions. The literature on knowability and belief revision are discussed in other chapters in this volume and so they will not be treated in detail here. (See Salerno and Fuhrmann #*in this volume*#) Since its origins in the early 1960s, the modern study of epistemic logic, and more specifically the study of the formal semantics of epistemic terms, has been a flourishing and diverse enterprise. Already in the late 1940s and into the 1950s Rudolf Carnap, Jerzy Łos, Arthur Prior, Nicholas Rescher, G.H. von Wright and other philosophers recognized that knowledge and belief exhibit systematic features that admit of an axiomatic-deductive treatment. Von Wright's insights were extended by Jaakko Hintikka in his book *Knowledge and Belief: An Introduction to the Logic of the Two Notions.* (1962) Hintikka provided a Kripke-style semantics for knowledge and belief. This formal framework was foundational for almost all subsequent work in the field.

For the most part, the semantics of epistemic logic is studied as a branch of modal logic. As described below, the modal systems to which philosophers and logicians were originally drawn (**S4** and **S5**) were generally regarded as unsatisfactory for philosophical reasons. Epistemological concerns, specifically the problem of logical omniscience, led to modifications of traditional possible world semantics. These modifications have generated their own questions and problems.

Informal philosophical concerns have always played a role in the development of epistemic logic. However, during the 1980s and 1990s many of the most important innovations in epistemic logic were the work of theoretical computer scientists rather than philosophers. In these decades, epistemic logicians took an increasingly pragmatic approach to their subject and epistemic logic was studied with an eye to its applicability to technological challenges.

To a certain extent, the philosophers who originally developed modern epistemic logic were also motivated by pragmatic concerns. Hintikka, in particular, has long emphasized the practical problem of how knowledge is acquired as opposed to the traditional emphasis on the justification and definition of knowledge in epistemology. However, by the 1980s the study of epistemic logic seemed to have drifted far away from its philosophical origins as theorists explored applications of epistemology in economics, artificial intelligence, robotics, database software design and network engineering.

The technological questions which occupied theoretical computer scientists involved the relationship between information, decision-making and action. So, for example, consider the problem of determining the point at which a computer should report that it does not have some item of information in its database? This problem has a very similar conceptual structure to traditional philosophical questions concerning self-knowledge. In this case, the programmer must determine when an agent can legitimately report that it knows that it does not know? As epistemic logic developed, topics which had previously been the purview of philosophers alone were being addressed in a technological context. (See Fagin et. al 1995)

In addition to the technological turn in epistemic logic, the most significant difference between work in the 1960s and what followed in the 1980s was a new emphasis on groups and dynamics rather than on the epistemic properties of individual agents. Individual agent based models favored by philosophers in the 1960s and 1970s were supplemented by multi-agent, multi-modal and dynamical approaches. In this context, the study of group knowledge came to the fore. Formal treatments of common knowledge and distributed knowledge have figured prominently in the recent literature. These developments touch on areas that are of great interest to epistemologists and merit more attention in the wider philosophical community.

For readers interested in a deeper understanding than this essay can provide, there are some excellent and comprehensive sources available. In their *Reasoning about Knowledge* Fagin, Halpern, Moses and Vardi (1995) provided an accessible and philosophically well-informed discussion of many of the most significant developments in theoretical computer science and artificial intelligence related to epistemic logic.

Meyer and van der Hoek's *Epistemic Logic for AI and Computer Science* was also published in 1995 and covered much of the same territory in greater technical detail. These two books served as canonical textbooks for modern epistemic logic and encapsulate the principal developments in the field from the 1980s and 90s. More recently, van Ditmarsch, van der Hoek and Kooi, published their *Dynamic Epistemic Logic* which shifts attention to treating knowledge and belief-change within a unified conceptual framework. (2007) Finally, Yoav Shoham and Kevin Leyton-Brown's *Multi-Agent Systems* (2009) provides an updated overview of the relevant literature on game theory and belief revision in a multi-agent setting. These four books present most of the necessary technical background for current work in epistemic logic in both theoretical computer science and philosophy.

There have been a number of recent attempts to bridge the gap between the epistemological community and epistemic logicians. Most valuable among these is Vincent Hendricks' *Mainstream and Formal Epistemology*. (2005) Hendricks provides a survey of the interplay between formal approaches and traditional epistemological questions and is a useful resource for interested philosophers.

This essay begins by considering the relationship between the post-Gettier project of modern epistemology and the formal project which emerged from the development of the semantics for epistemic logic. The remaining sections provide an introduction to the logic and formal semantics of epistemic notions and describe some of the achievements and open questions in the field.

Modern epistemology and its logic: Separated at birth

Most epistemologists favor informal methods. Common sense intuitions, thought experiments and counterexamples continue to be the principal tools of the mainstream epistemologist's trade. Matters of methodological taste are only partly responsible for the neglect of logic and semantics. More cogent reasons involve the concern that formal approaches are orthogonal to the project of providing a philosophically sound definition of knowledge. Since defining knowledge is what most contemporary epistemologists see as their principal task, their neglect of epistemic logic is understandable. Nevertheless, the early pioneers in epistemic logic saw the study of the logic and formal semantics of knowledge as a central part of the task of epistemology. These philosophers certainly did not regard themselves as neglecting epistemology in favor of some tangentially related formal enterprise. Thus, as described below, an important reason for the parting of ways, between so-called mainstream and formal epistemologists involved a meaningful disagreement over the nature of epistemology and more specifically over the role of epistemology in science and philosophy.

The most significant difference between the early figures in epistemic logic and the post-Gettier tradition concerns their views of the purpose of epistemology. For most epistemic logicians, epistemology is a matter of understanding and improving practices related to inquiry. By contrast, mainstream epistemologists in the post-Gettier tradition hoped to provide a definition of knowledge. The difference is quite simple: According to most epistemic logicians, the project of improving our ability to attain new knowledge is more important than the justification and definition of knowledge we have already attained.

Hintikka presents the most prominent and sustained defense of the view that epistemologists should be developing accurate and useful theories of inquiry rather than worrying about the nature of knowledge. (See Hilpinen 2006 and Hintikka 2007 11-37)While some philosophers, including Hendricks and Symons (2006) have argued that epistemic logic is, in fact quite relevant to the problem of defining knowledge, epistemic logicians are generally impatient with (or uninterested in) most of the ongoing debates in mainstream epistemology. This basic difference over the nature of epistemology has had consequences for the development of the field and has meant that mainstream epistemology has largely developed along a separate path from epistemic logic.

As Robert Stalnaker notes, contemporary epistemology and modern epistemic logic were separated soon after their births. Edmund Gettier's classic criticism of the Justified True Belief (JTB) analysis of knowledge (1963) was published shortly after *Knowledge and Belief* (1962). Hintikka provided the first formal semantics for epistemic concepts, thereby generating the modern research program in epistemic logic. Meanwhile, Gettier's paper sparked the familiar project wherein epistemologists attempt to refine or supplement the JTB model of knowledge.

Philosophers had long recognized that discourse about knowledge and belief exhibits systematic regularities that can be presented in an axiomatic-deductive system. However, prior to Hintikka's work, the lack of an appropriate semantics limited the philosophical usefulness of early reflections on epistemic logic. The development of Hintikka's semantics for epistemic logic was motivated in part by his view of the role of knowledge in inquiry. For Hintikka, epistemology is the theory of inquiry and the ultimate purpose of inquiry is to serve as a guide for decision making. The connection between decision making and the acquisition of new information is simple. On Hintikka's view, to know *p* means to be in the position to rule out possibilities in which it is not the case that *p.* Once we begin to think about inquiry and decision making in terms of ruling out possibilities, the modal character of epistemic terms is relatively obvious. Just as a necessary truth is one which is true in all possible worlds, an agent's knowledge can be understood as the set of truths which obtain in all of the agent's *epistemically* possible worlds. In other words, for an agent to know *p* means that in all worlds compatible with the agent's knowledge, it is the case that *p.*

> To take a simple example, let us suppose that I am getting ready to face a new day in the morning. How does it affect my actions if I know that it will not rain today? You will not be surprised if I say that what it means is that I am entitled to behave as if it will not rain – for instance to leave my umbrella home. However, you may be surprised if I claim that most of the important features of the logical behavior of knowledge can be teased out of such simple examples. (Hintikka 2007, 11-2)

Modern work in epistemic logic has built upon this view of the relationship between knowledge and possibility. So, for example, the connection between epistemic and pragmatic considerations continues into the 1980s and 1990s. Many researchers in Artificial Intelligence define belief, for instance, as the set of propositions which an agent would be willing to act upon. Nevertheless many mainstream epistemologists would take issue with the idea that we can simply develop a formal representation of knowledge without clarifying precisely what epistemic possibility amounts to. This is another way of framing the basic philosophical difference between the two camps. For the epistemic logician, as described below, the notion of epistemic possibility can be treated as an accessibility relation between possible worlds, one for each agent, and we are free to specify that relation in a variety of ways. As Stalnaker notes, "the idea was to give a precise representation of the structure of an epistemic state that was more or less neutral about more substantive questions about what constitutes knowledge, but that sharpened questions about the logic of knowledge." (2006, 171-2) Among theoretical computer scientists, the rejection of the philosophical goal of defining knowledge has been more explicitly stated and the following view expressed by Fagin, Halpern, Moses and Vardi is representative:

> We should emphasize here that we do not feel that the semantic model we represent in the next chapter is the unique "right" model of knowledge. We spend some time discussing the properties of knowledge in this model. A number of philosophers have presented cogent arguments showing that some of these properties are "wrong." Our concerns in this book are more pragmatic than those of the philosophers. We do not believe that there is a "right" model of knowledge. Different notions of knowledge are appropriate for different applications. (1995, 8)

In addition to being concerned about what they might regard as the excessively cavalier or pragmatic attitude towards the notion of epistemic possibility, many epistemologists were also concerned by what they saw as a very different kind of problem, namely the excessively idealized characterization of epistemic agency in epistemic logic. Early on, epistemologists were suspicious of assumptions which seemed built into the formal analysis, most importantly the assumption that epistemic agents are logically omniscient. Logically omniscient agents are committed to knowing all the logical consequences of their knowledge. For instance, a logically omniscient agent would know all logical truths and would recognize all the sentences which are logically equivalent statements of the sentences they know. Ordinary human believers do not fit this description. Epistemic logic has been challenged as setting unreasonably high standards for what counts as membership in the class of knowers (or believers!). If human beings never count as knowers, then from a strictly human perspective, epistemic logic seems to have made itself irrelevant to epistemology. The problem of logical omniscience and some of the solutions that epistemic logicians have proposed will be discussed in greater detail below.

It is important to avoid exaggerating the degree to which epistemology and epistemic logic have parted ways. To be sure, epistemic logic has encouraged epistemologists to explore questions concerning iterated knowledge, for example, whether knowing *p* implies knowing that one knows *p*. Furthermore, once we scratch the surface a little more deeply one finds a number of highly significant and philosophical results (and problems!) that are largely the result of attention to the formal character of the concepts of knowledge and belief, these include such familiar topics as Fitch's paradox of knowability, the problem of logical omniscience, the KK thesis and the like.

The systematic features of our epistemic discourse and the underlying logical structure of knowledge

Epistemic expressions like 'knows that' or 'believes that' have systematic properties that are amenable to formal study. Most obviously, statements containing epistemic expressions sometimes involve logical constants which behave in the usual way. So, for example, if you know *p* and *q* then you know *q*. The conceptual features of statements concerning knowledge and belief become more interesting and complicated when one begins to examine the characteristics of general principles governing the use of epistemic concepts. The behavior and interaction of these general principles has been the focus of epistemic logic. For example, as G.E. Moore pointed out, there seems to be something wrong with claiming

(1) "*p* and I do not believe *p*".

Assertions of this kind have a self-defeating quality because of the conceptual features of knowledge or belief and not simply by virtue of the syntactical features of the sentence or the character of the logical constants that are involved. As G. E. Moore noted "'I went to the pictures last Tuesday, but I don't believe that I did' is a perfectly absurd thing to say." (1952 543) The *perfect absurdity* here is due to a violation of a principle governing epistemic concepts. (1) is not obviously contradictory. Indeed, it is often a correct description of the state of affairs in question. For instance, since I recognize that I am fallible I am committed to the existence of cases where it is true that *p* and I do not believe *p*. Furthermore, I can assert, without paradox or contradiction, that there are some propositions *p* such that *p* and I do not believe *p.*

The paradox arises from the peculiarity of the agent in question attesting to particular instances of (1), where the variable *p* is replaced by an assertion concerning some state of affairs. Specifically, the paradox results from what John Austin called the illocutionary features of the claim. While I recognize that there may be cases where replacing *p* with some description of some state of affairs is true, I cannot sincerely attest to both parts of the conjunction contained in (1) at a particular moment for any specific instance of (1). In this sense, Moore's paradox sheds light on the properties of epistemic agents and the concept of belief. Reflecting on the *perfect absurdity* of Moore's examples, shows us that an agent's belief and its agency are related. However, we are not restricted to relying on epistemic intuitions in our consideration of Moore's paradox. Exploration of the principles or norms governing epistemic notions can take place in an axiomatic fashion. Hintikka, for example, provided a proof of the contradictory nature of the paradoxical form of the Moore statements, the case where the statement asserts that an agent believes *p* and *not-p*. (1962, 67) I can take as a rule for instance that it is prohibited or confused to say:

(2) "I know *p* but it is not the case that *p*"

If (2) is prohibited, it is due in part to the illocutionary considerations which applied in (1) but unlike assertions of belief, (2) is prohibited by virtue of another general principle, namely the veracity of knowledge.

(3) if one knows *p* then it is true that *p*,

If we accept that knowledge implies veracity then we can consider what the implications of taking it as an axiom might be and whether it is consistent with other general epistemic principles we might hold. As discussed below, considerations of this kind have been given an elegant formal framework by epistemic logicians. Wolfgang Lenzen (1978) provided an excellent overview of arguments from the 1960s and 1970s concerning the appropriate axioms for knowledge.

Observations of the kind emphasized by Moore, concerning the behavior of the term "knows that" served as the starting points for the development of modern epistemic logic. G.H. von Wright was the first to sketch an axiomatic treatment of the behavior of epistemic concepts. (1953 29-35) His discussion continues to inspire epistemologists to study the adequacy and implications of various epistemic axioms in a systematic manner. However, the project began to produce interesting results once Hintikka provided a semantic interpretation of epistemic and doxastic notions in the early 1960s.

Hintikka began by supplementing the language of propositional logic with two unary epistemic operators K_a and B_a such that K_ap reads 'Agent *a* knows *p*' and B_ap reads 'Agent *a* believes *p*' for some proposition *p*. In this way, candidate epistemic or doxastic axioms can be presented in formal terms. So, for instance, we have already seen that one intuitive axiom which we are likely to accept into our epistemic logic is:

(4) $K_cA \rightarrow A$

This is known as axiom T which we saw above as (3). With our modest addition to first-order logic in hand, we can begin to catalogue other plausible epistemic axioms.
A standard list of the axioms (following Lemmon (1977), Bull and Segerberg (1984)) that are relevant for epistemic logic run as follows:

K	$K_c(A \rightarrow A') \rightarrow (K_cA \rightarrow K_cA')$
D	$K_cA \rightarrow \neg K_c\neg A$
T	$K_cA \rightarrow A$
4	$K_cA \rightarrow K_cK_cA$
5	$\neg K_cA \rightarrow K_c\neg K_cA$
.2	$\neg K_c\neg K_cA \rightarrow K_c\neg K_c\neg A$
.3	$K_c(K_cA \rightarrow K_cA') \vee K_c(K_cA' \rightarrow K_cA)$
.4	$A \rightarrow (\neg K_c\neg K_cA \rightarrow K_cA)$

We can consider the philosophical merits of each axiom to a certain extent without the introduction of additional formalism. However, Hintikka's approach to the semantics of epistemic notions offers an important supplement to our intuitive reflections. In order to begin thinking about the relative merits of these axioms one can begin by considering the familiar interpretation of the *K* and *B* operators using possible world semantics along the lines discussed above:

K_cA: *In all possible worlds compatible with what c knows, it is the case that A*

B_cA: *In all possible worlds compatible with what c believes, it is the case that A*

The basic assumption is that any ascription of propositional attitudes like knowledge and belief, involves dividing the set of possible worlds in two: Those worlds compatible with the attitude in question and those that are incompatible with it.

The central idea in possible worlds semantics is the notion of accessibility. Accessibility is a relation which is defined on the set of possible worlds. In standard modal logic we say that some world *w* is accessible from some world *w′* just in case *w* is *possible* relative to *w′*. Specifically, the relation can be characterized as a subset of the Cartesian product of the set of possible worlds. As described below, determining the accessibility relation is the most basic step in determining the properties of our semantical framework. So, for example, whether one assumes that the accessibility relation is symmetric, transitive, reflexive or some combination of the three, will make a significant difference in how one thinks about the modal or epistemic properties of the system in question. In the epistemic context, the set of worlds accessible to an agent (its set of epistemic alternatives) depends on its informational resources at an instant. This dependency is captured via the specification of the accessibility relation, *R*, on the set of possible worlds. To express the idea that for agent c, the world w' is compatible with his information state, or accessible from the possible world w which c is currently in, it is required that R holds between w and w'. This relation is written Rww' and reads "world w' is accessible from w". The world w' is said to be an *epistemic* or *doxastic alternative* to world w for agent c, depending on whether knowledge or belief is under consideration. We can give this a semantic interpretation, by saying that if a proposition A is true in all worlds which agent c considers possible then c knows A.

A possible world semantics for a propositional epistemic logic with a single agent c then consists of a *frame* $\mathcal{F}$ which in turn is a pair $<W,R_c>$ such that W is a non-empty set of possible worlds and R_c is a binary accessibility relation (relative to agent c) over W. A *model* $\mathcal{M}$ for an epistemic system consists of a frame and a denotation function φ assigning sets of worlds to atomic propositional formulas. Propositions are taken to be sets of possible worlds; namely the set of possible worlds in which they are true. Let *atom* be the set of atomic propositional formulae, then $\varphi : atom \mapsto P(W)$, where P denotes the powerset operation.

The model $\mathcal{M} = <W,R_c,\varphi>$ is called a Kripke-model and the resulting semantics Kripke-semantics (Kripke 1963): An atomic propositional formula, **a**, is said to be true in a world w in $\mathcal{M}$ (written $\mathcal{M},w \vDash \mathbf{a}$) iff w is in the set of possible worlds assigned to **a**, i.e., $\mathcal{M},w \vDash \mathbf{a}$ iff $w \in \varphi(\mathbf{a})$ for all $\mathbf{a} \in atom$. The formula K_cA is true in an world w (i.e., $\mathcal{M},w \vDash K_cA$) iff $\forall w' \in W$, if R_cww', then $\mathcal{M},w' \vDash A$. The semantics for the Boolean connectives follow the usual recursive recipe. Similar semantics may be formulated for the belief operator. Since a belief is not necessarily true but rather probably true, possibly true, or likely to be true, we must modify our approach to the semantics of belief appropriately. For instance, belief may be modeled by assigning a sufficiently high degree of probability to the proposition in question and determining the doxastic alternatives accordingly. The truth-conditions for the doxastic operator are defined in a way similar to that of the knowledge operator and the model may also be expanded to accommodate the two operators simultaneously.

A modal formula is said to be *valid* in a frame iff the formula is true for all possible assignments in all worlds in the frame.

An important feature of possible world semantics is that the epistemic axioms listed above, correspond to algebraic properties of the frame in the following sense: A modal axiom is valid in a frame if and only if the accessibility relation satisfies some algebraic condition. (See Hendricks and Symons (2006)) For example, the axiom expressing the veridicality property that if a proposition is known by *c*, then *A* is true,

(5) $K_c A \rightarrow A$,

is valid in all frames in which the accessibility relation is reflexive in the sense that $\forall w \in W$: Rww. Given reflexive accessibility, every possible world is accessible from itself. Similarly if the accessibility relation satisfies the condition that

(6) $\forall w, w', w'' \in W$: $Rww' \wedge Rw'w'' \rightarrow Rww''$

which is also known as transitivity, then the axiom (7) is valid.

(7) $K_cA \rightarrow K_cK_cA$,

(7) is called axiom 4 and is also known as the axiom of self-awareness, positive introspection or the KK-thesis. In this case, the axiom captures the idea that if the agent knows *p* then it has knowledge of its knowledge that *p*. Other axioms require yet other relational properties to be met in order to be valid in all frames: If the accessibility relation is reflexive, symmetric and transitive, then

(8) $\neg K_cA \rightarrow K_c\neg K_cA$

is valid. (8) is called axiom 5 also better known as the axiom of wisdom. This is the much stronger thesis that an agent has knowledge of its own ignorance: If *a* does not know *p*, it knows that it doesn't know *p*. The axiom is also known as the axiom of negative introspection.

One contentious axiom which is valid in all possible frames

(9) $K_c (A \rightarrow B) \rightarrow (K_cA \rightarrow K_cB)$

is the closure condition for knowledge, also known as axiom K, or the axiom of deductive cogency: If the agent *a* knows $p \rightarrow q$, then if *a* knows *p*, a also knows *q*. As discussed below, this axiom leads to the most difficult philosophical problem for epistemic logicians, namely the apparent commitment to logical omniscience. It seems, that if one accepts this axiom, that an epistemic agent must know everything that follows logically from its knowledge.

Other axioms of epistemic import require yet other relational properties to be met in order to be valid in all frames. When combined in various ways, these axioms make up epistemic modal systems of varying strength. Their strengths vary according to the modal formulas valid in the respective systems and given the algebraic properties assumed for the accessibility relation.

Returning to the axioms listed above, we can begin to see how we might compare their relative strengths. The weakest system of epistemic interest is usually considered to be system **T**. The reader should take care to distinguish the epistemic operator *K*, the modal axiom K and the system of axioms **K** in what follows. Similarly, we distinguish the axiom T from the system **T**. **T** is a system of modal logic which is characterized by reflexive frame with the axioms T and K as valid axioms.

Additional modal strength may be obtained by extending **T** with other axioms drawn from the above pool altering the frame semantics to validate the additional axioms. By way of example, while

(10) $K_c A \rightarrow A$

is valid in system **T**,

(11) $K_c A \rightarrow A$, $K_c A \rightarrow K_c K_c A$ and $\neg K_c A \rightarrow K_c \neg K_c A$

are all valid in **S5** but not in **T**.

System **T** has a reflexive accessibility relation, **S5** has reflexive, transitive and symmetrical accessibility relations. The arrows in the table below indicate that the system to which the arrow is pointing is included in the system from which the arrow originates and hence reflect relative strength. Then **S5** is the strongest and **S4** the weakest of the ones listed.

Epistemic Systems			
KT4	=	**S4**	
KT4 + .2	=	**S4.2**	↑
KT4 + .3	=	**S4.3**	↑
KT4 + .4	=	**S4.4**	↑
KT5	=	**S5**	↑

Table 2: Relative Strength of Epistemic Systems Between S4 and S5

One of the important tasks of epistemic logic is to catalogue all sound and complete systems of such logics in order to allow us to pick the most ‘appropriate’ ones. The logics range from **S4** over the intermediate systems **S4.2–S4.4** to **S5**. By way of example, Hintikka settled for **S4** (1962), Kutschera argued for **S4.4** (1976), Lenzen suggested **S4.2** (1978), van der Hoek has proposed to strengthen knowledge according to system **S4.3**

(1996). van Ditmarsch, van der Hoek and Kooi together with Fagin, Halpern, Moses and Vardi (Fagin et al. 1995) and others assume knowledge to be **S5** valid.

In the doxastic context, we can also catalogue the completeness properties of the alternative systems in a similar fashion. Of course in doxastic logic we drop axiom T, which is usually replaced by D. This avoids committing doxastic logic to the truth of beliefs while retaining the condition that beliefs be consistent. Replacing T with D generates systems like **KD4–KD45**. This approach permits the combination of epistemic and doxastic systems and for studying the interplay between knowledge and belief (see Voorbraak 1993). There are some important philosophical concerns with such combined doxastic and epistemic systems. Lenzen (1978) and Stalnaker (1996) point out that such combined systems risk conflating knowledge and belief.

How does semantic formalization relate to epistemology? By way of example, it is worth returning briefly to our discussion of Moore's problem to see what kind of light Hintikka's formalization shed on that case. In *Knowledge and Belief*, he was able to prove that statements of the sort "*p* and I do not believe *p*" are *perfect absurdities* not because they run afoul of some kind of epistemic intuition, but because, when properly analyzed, they generate a contradiction. More importantly, the analysis allows us to recognize which epistemic commitments are involved in generating the contradiction. These commitments are formulated as rules for epistemic alternatives in model systems. So, for example, the proof of the absurdity of "*p* and I do not believe *p*" (1962, 68) relies on the conditions governing the semantics of sentences concerning belief. The difference between the kind of reasoning we find in Moore and Hintikka with respect to "*p* and I do not believe *p*" boils down to difference with respect to the degree of explicitness and control that the philosophers aspire to in their arguments. For Hintikka, unlike Moore, the point is to achieve the same level of explicitness in epistemology as is found in logic:

> The word “logic” which occurs in the subtitle of this work is to be taken seriously. My first aim is to formulate and to defend explicit criteria of consistency for certain sets of statements - criteria which, it is hoped, will be comparable with the criteria of consistency studied in the established branches of logic. (1962, 3)

Logical omniscience and idealized epistemic agents

Epistemic logic inevitably traffics in idealizations. As discussed below, the problem of logical omniscience (a product of accepting the axiom of deductive cogency or axiom K and standard possible world semantics) encouraged theorists to craft formal systems which more adequately reflected the actual properties of epistemic agents. Since real epistemic agents modify their beliefs and engage in inquiry, there was some philosophical interest in attempting to formally capture the dynamical features of inquiry. Developments since *Knowledge and Belief*, principally those since Kutchera’s (1976) and Lenzen’s (1978) attempted to integrate broader insights from modal logic with epistemic logic and have made it possible to formally model some prominent features of the dynamical nature of epistemic agency. Gärdenfors’ (1988) account of belief revision was particularly important in setting the stage for a slew of dynamical models of knowledge.

Logical omniscience is related to closure properties. Axiom K can, under certain circumstances, be generalized to a closure property for an agent's knowledge which is implausibly strong: Whenever an agent *c* knows all of the formulas in a set Γ and *A* follows logically from Γ, then *c* also knows *A*. In particular, *c* knows all theorems (letting Γ = ø), and he knows all logical consequences of any formula which he knows (letting Γ consist of a single formula).

In response to the threat of logical omniscience, some epistemologists raised the question of whether the very idea of a logic of knowledge makes any epistemological sense. For instance, Hocutt challenged the applicability of logic to any realistic account of knowledge. (1972) Because there is no guarantee that a knower will recognize that it is committed to some proposition that is logically equivalent to some proposition to which it readily assents, the very idea of an epistemic logic is on slippery ground.

Some of the first proposals for solving the problem of logical omniscience introduce semantical entities which explain why the agent appears to be, but in fact is not really logically omniscient. These entities were called 'impossible possible worlds' by Hintikka (1975). The basic idea is that an agent may mistakenly count among the worlds consistent with his or her knowledge, some worlds containing logical contradictions. The mistake is simply a product of limited resources; the agent may not be in a position to detect the contradiction and may erroneously count them as genuine possibilities.

'Seemingly possible' worlds are introduced by Veikko Rantala (1975) in his urn-model analysis of logical omniscience. Rantala devised a way of alleviating the mismatch between our model theoretic reasoning about knowledge and our proof theoretic commitments: He asks us to conceive of our epistemic relationship with the world by analogy with an urn from which we may draw balls (individual units of information) one by one over time. With each new piece of information drawn from the urn, we can modify our models. The idea is, simply, that inquiry is a dynamical process in which our model of the world changes with new information. Rantala has provided a formalism which incorporates an intuitively reasonable notion of change in a model. Such change can be understood as a change in the properties of individuals of the model or a change in its universe of discourse.

Representing how the agent's model might dynamically update is one way of thinking about epistemic agency in a more realistic manner. However, on any realistic account of epistemic agency, the agent is likely to consider (albeit inadvertantly) worlds in which the classical laws of logic do not hold. In this context, the general problem of establishing a set of epistemic principles for a realistic agent is unavoidable. Rantala's approach provides a way of making the appearance of logical omniscience less threatening, but at the cost of introducing a degree of arbitrariness along with impossible or seemingly possible worlds. (See Rantala 1982) In Rantala's discussion of the semantics for impossible worlds (1982) the truth condition is completely free, insofar as any contradiction among an agent's beliefs can be represented by a model containing an impossible world. While logical omniscience is avoided, the price we pay is high, since no real epistemic principles hold broadly enough to encompass impossible and seemingly possible worlds. (See Meyer and van der Hoek 1995, 87-88)

Some conditions must be applied to epistemic models such that they cohere with epistemic principles. Computer scientists have proposed that what is being modeled in epistemic logic is not knowledge simpliciter but a related concept which is immune to logical omniscience. The epistemic operator $K_c A$ should be read as 'agent c knows implicitly *A*', '*A* follows from *c*'s knowledge', '*A* is agent *c*'s possible knowledge', etc.

Propositional attitudes like these should replace the usual ‘agent *c* knows *A*’. While there exists some variation, the locutions all suggest modeling implicit knowledge or what is implicitly represented in an agent's information state rather than explicit knowledge (Fagin et al. 1995, and others). The agents neither have to compute knowledge nor can they be held responsible for answering queries based on their knowledge under the implicit understanding of knowledge. Logical omniscience is an epistemological condition for implicit knowledge, but the agent may actually fail to realize this condition.

There are a variety of ways of responding to these kinds of challenges. (See Vardi 1986 for an excellent review of the issue) One rather unpromising approach is to deny that epistemic logic is under any obligation to connect with more general epistemological concerns. As Lenzen argues:

> The search for the correct analysis of knowledge, while certainly of extreme importance and interest to epistemology, seems not significantly to affect the object of epistemic logic, the question of validity of certain epistemic-logical principles. (Lenzen, 1978, p. 34)

Rather than treating epistemic logic as a purely formal exercise, a preferable response involves maintaining that epistemic logic does carry epistemological significance but in an inevitably idealized sort of way. One restricts attention to a class of rational agents where rationality is defined by certain postulates. Thus, agents have to satisfy at least some minimal conditions to simply qualify as rational. This is, for example, what Lemmon originally suggests. (Lemmon, 1959) One such condition would involve assuming that rational agents should acknowledge the laws of logic. For instance, if the agent knows p and $p \rightarrow q$, it should be able to recognize that q follows validly.

These 'rationality postulates' for knowledge exhibit a striking similarity to the laws of modal and epistemic logic. One may in turn legitimately attempt to interpret the necessity operator in alethic axioms as a knowledge operator and then justify the modal axioms as axioms of knowledge. While Lemmon constructs the rational epistemic agent directly from the axiomatization of the logic, another way of justifying the epistemic axioms involves reference to their semantical features. This is the line of thought that Hintikka pursued in *Knowledge and Belief*. Hintikka stipulated that the axioms or principles of epistemic logic are conditions descriptive of a special kind of general (strong) rationality. The statements which may be proved false by application of the epistemic axioms are not inconsistent meaning that their truth is logically impossible. They are rather rationally 'indefensible'. Indefensibility is fleshed out as the agent's epistemic laziness, sloppiness or perhaps cognitive incapacity whenever to realize the implications of what he in fact knows. Defensibility then means not falling victim of 'epistemic negligence' as Chisholm calls it (Chisholm, 1963, 1977). The notion of indefensibility gives away the status of the epistemic axioms and logics. Some epistemic statement for which its negation is indefensible is called 'self-sustaining'. The notion of self-sustenance actually corresponds to the meta-logical concept of validity. Corresponding to a self-sustaining statement is a logically valid statement. But this will again be a statement which is rationally indefensible to deny. So, in conclusion, epistemic axioms can be understood to be descriptions of rationality. This argument is spelled out in detail by Hilpinen (2002).

Common Knowledge and Distributed Knowledge

So far, this essay has discussed the epistemic properties of individual agents. However, many recent developments in epistemic logic concern the study of the formal properties of systems of interacting agents. Work in these domains has been relatively underexploited by epistemologists and is likely to offer considerable philosophical insight. This section introduces two of the most prominent notions in the study of multi-agent systems: common and distributed knowledge.

When we consider the epistemic properties of human beings, it is obvious that we are involved in a dynamical social context in which one's knowledge of what one's neighbors know or fail to know is of great importance. It is also obvious that human knowledge and inquiry is related to the kinds of communication structures or networks to which we belong. The structure and behavior of one's social and communications networks is a significant factor in determining one's epistemic resources.

So, for example, when we consider agents who are connected via some network we can study the effect of new information, presented to part of, or made public to the whole, group. Formal grasp of the role of announcements in a complex network of agents has important practical consequences for our understanding of cooperation and competition. The manner in which new information moves through a multi-agent system and how it causes individual agents to modify their beliefs is, in part, dependent on the character of the networks connecting those agents. However, analysis of the cooperative and competitive behavior of agents immediately brings into focus the dependency of these behaviors on a prior shared epistemic medium known as common knowledge.

Common knowledge, as distinguished from shared or mutual knowledge begins with the knowledge that one's fellow agents know that *p*. However, it is more than that, because not only do all the agents in a group know that *p*, they also know that *all other* agents know *p* and that furthermore that they all know that they all know *p*, and so on. I know that my fellow drivers know that they ought to stop at the red light and they know that I know that they know that they should stop, etc. Thus, in one sense common knowledge is a very powerful kind of phenomenon. The social role played by common knowledge in a broad range of human activities has long been recognized by philosophers from David Hume (1740) to David Lewis. (1969)

Common knowledge became a concern for theoretical computer scientists given the difficulties faced by projects in artificial intelligence which focus too narrowly on the epistemic condition of single agents. These difficulties encouraged theoretical computer scientists to focus on the social and conventional features of knowledge and belief. Common knowledge is the basic background knowledge which supports the kinds of social entanglements that are crucial for sophisticated forms of intelligence; it is what any fool knows (to echo John McCarthy (1979)). Philosophical accounts of common knowledge see it as carrying a great burden; supporting the very possibility of the kind of collaborative activity that defines human intelligence. In order to approach anything resembling human epistemic agency, a significant level of social scaffolding needs to be in place. Clearly, for example, we depend on shared epistemic starting points for most basic social interactions, including, prominently, membership in a linguistic community.

As Lewis noted (1969) a convention requires common knowledge among the agents that observe it. Robert Aumann also emphasized the centrality of common knowledge with respect to norms, social and linguistic practices, agent interactions and games. (1994).

A detailed treatment of the various formal techniques for tackling common knowledge is beyond the scope of this essay. (see Vanderschaaf and Sillari 2007 for an excellent overview) However, given our account of the logical landscape for single-agent systems above, it is possible to introduce some of the main features of multi-agent systems with just a little oversimplification. The primary difference between the semantics of single- and multi-agent semantics is that more than one accessibility relation is introduced. A modal system for *n* agents results from combining *n* modal logics in cases where it may be assumed that the agents are homogenous in the sense that they may all be described by the same logical system. Thus, in the simplest case, an epistemic logic for *n* agents consists of *n* copies of a certain modal logic. In such an extended epistemic logic it is possible to express that some agent in the group knows a certain fact, that an agent knows that another agent knows a certain fact etc.

It is possible to develop the logic even further: Not only may an agent know that another agent knows a fact, but they may all know this fact simultaneously. From here it is possible to express that everyone knows that everyone knows that everyone knows that everyone knows that … some fact holds. This is what is meant by common knowledge.

One way of defining common knowledge involves defining common knowledge for the entire group of agents rather than partitioning the group of agents into subsets with different common 'knowledges'. Once multiple agents have been added to the syntax, the language is augmented with an additional operator *C*. *CA* is then interpreted as ‘It is common knowledge among the agents that *A*’. Well-formed formulas follow the standard recursive recipe with modifications which account for the the multiple agents. So, for instance the operator *E* is introduced such that *EA* means ‘Everyone knows that *A*’. *EA* is defined as the conjunction $K_1A \wedge K_2A \wedge \ldots \wedge K_nA$.

To semantically interpret *n* knowledge operators, binary accessibility relations R_n are defined over the set of possible worlds *W*. A special accessibility relation, $R°$, is introduced to interpret the operator of common knowledge. The relation must be flexible enough to express the relationship between individual and common knowledge. The idea is to let the accessibility relation for *C* be the transitive closure of the union of the accessibility relations corresponding to the knowledge operators for the individual agents.

The model $\mathcal{M}$ for an epistemic system with *n* agents where the agents have common knowledge is a structure $\mathcal{M} = \langle W, R_1, R_2, \ldots, R_n, R°, \varphi \rangle$, where *W* is a non-empty space of possible worlds, $R_1, R_2, \ldots, R_n, R°$ are accessibility relations over *W* for which $R° = (R_1 \cup R_2 \cup \ldots \cup R_n)$ and φ again is the function assigning worlds to atomic propositional formula $\varphi : atom \mapsto P(W)$. The semantics for the Boolean connectives remain intact. The formula K_iA is true in a world *w*, i.e., $\mathcal{M}, w \vDash K_iA$ iff $\forall w' \in W$: if R_iww', then $\mathcal{M}, w' \vDash A$. So, *A* is common knowledge in a world *w*, when $\mathcal{M}, w \vDash CA$ iff $R°ww'$ implies $\mathcal{M}, w' \vDash A$.

Varying the properties of the accessibility relations $R_1, R_2, \ldots, R_n$, results in different epistemic logics. For instance system **K** with common knowledge is determined by all frames, while system **S4** with common knowledge is determined by all reflexive and transitive frames. Similar results can be obtained for the remaining epistemic logics (Fagin *et al*. 1995).

Informally speaking, the claim that some proposition is common knowledge is extremely strong. Therefore, those propositions for which we can claim common knowledge tend to be very weak. To claim that some proposition is common knowledge, that everyone knows that everyone knows *A*, implies that everybody knows *A*, which implies individual knowledge of *A*.

If we think of common knowledge as involving very strong claims about the epistemic state of a group of agents, at the opposite end of the spectrum is the notion of distributed knowledge. Distributed knowledge is an epistemic property which is captures the idea that there is an aggregated store of knowledge in a group, some of which may not necessarily be possessed by any individual member of the group. If even one agent in a group knows *A* then *A* is part of the distributed knowledge of the group. Where common knowledge is very strong (and its argument is rather weak), Distributed knowledge is weaker, but can be obtained for much stronger facts, as we shall see.

A good way to think about distributed knowledge is to recognize its relationship to a communication network. Something is distributed knowledge in a group, if it could be known by the individuals, were they able to talk to each other. For instance, in a crowd of 100 people, when 2 people have the same birthday, this might not be individually known, but could be known if the members were able to talk to each other. So, to take a simple case, if A knows that B is older than C or D and E knows that B is not older than D, then while no individual agent knows that A is older than C, that knowledge is distributed throughout the group and could be elicited given the right kinds of communication. This is a kind of knowledge which can be ascribed to some collection of agents (given certain conditions) but which need not necessarily be ascribed to any agent in isolation. Of course the knowledge that an individual agent has is also part of the aggregated store of distributed knowledge.

When we consider the epistemic properties of groups such as corporations or scientific communities, distributed knowledge that the group exhibits is likely to be one of the properties of interest. For instance, if we are entitled to say that the electric company as a whole knows how to maintain the power supply, we do so by reference to the distributed knowledge which exists in the group. Similarly, in debates concerning the nature of cognition which relies on resources beyond the confines of the individual brain and body; those problems related to the possibility of extended cognition, the formal study of distributed knowledge may prove useful. The distribution of knowledge in a community may seem like a rather nebulous or metaphysically extravagant notion until we begin to examine it in a formal setting.

The most famous example of the formal study of features of group knowledge involves scenarios like the muddy children problem. The scenario involves n children and their father. k children have mud on their foreheads. The children can see each other but they cannot see whether they have mud on their own foreheads. The children trust their father, do not cheat, are rational and do not communicate with one another. The scenario involves the effect of the father's announcement on the behavior of the group and on the epistemic states of the members of the group. The father announces: `There is at least one child with mud on its forehead. Will all the children who know they have mud on their foreheads please step forward?' If k is greater than one, no child steps forward. When their father makes his announcement the *k*-th time all muddy children step forward.

In order to explain this scenario, we first consider the simplest case. Where k =1 then the child with mud on its head knows that it must be the muddy one since it sees no other child with mud on its head. So, when k = 1 the explanation is clear. In the case where k = 2 we can imagine the following scenario with two muddy children *a* and *b*; k=2.

$a^* \rightarrow \quad b^*$

$\downarrow$

c

Muddy child *a* can see that *b* is muddy, but it does not know whether it is muddy itself. After the father's first announcement, when *b* does not step forward, *a* knows that *b* does not know whether *b* has a muddy forehead. This means that *b* sees at least one muddy forehead on either *a* or *c*. Since *a* can see that *c* has a clean forehead, *a* reasons that the muddy forehead that *b* saw, was *a*'s. *b* reasons in the same way based on the failure of muddy headed *a* to step forward after the father's first announcement. So, after the second announcement, both children step forward. By induction from cases with one and two muddy children, we can easily see how cases with greater numbers of muddy children would proceed. A full treatment of the muddy children problem can be found in

Meyer & van der Hoek, `Epistemic Logic for AI and CS', (56) or in Fagin et al., `Reasoning About Knowledge'. (3)

There is a range of cases like these in which we must account for the interaction of multi-agent systems and in which certain collective features of group behavior must be explained. The kinds of epistemic systems under consideration include cases, like the muddy children, which are sensitive to the introduction of new information via public announcement and in which the interactions of the agents contributes another dynamical component which has an effect on the unfolding states of the system. The muddy children problem is a simple example of the kinds of dynamical features that epistemic logic tackled in the 1980s and 1990s.

Game Theory, Belief Revision and the Properties of Agents

In multi-agent settings, it is natural to consider the role of competition and cooperation. Thus, as epistemic logic began to attend to the dynamics of groups, game theory began to play a more prominent role in reflections on epistemic agency. Aumann, van Benthem, Brandenburger, Fagin, Halpern, Keisler, Moses, Stalnaker, Vardi and others have contributed to uncovering important features of agent rationality showing how game theory adds to the general understanding of notions like knowledge, belief and belief revision. By the end of the 1990s Baltag, Moss, Solecki had combined epistemic logic with belief revision theory to study actions and belief updates in games (Baltag *et al*. 1999).

In the 1980s Alchourrón, Gärdenfors and Makinson developed a theory of belief revision theory (AGM) which provides an account of rational change of belief in light of novel evidence (Alchourrón 1985, Gärdenfors 1988). Expansions, contractions and revisions in an agents' set of belief are characterized formally. 'Revision' here means additions of beliefs to the agent's belief set which maintain consistency. Revision is distinguished from simple expansion, which takes place without regard for consistency, and contraction, where beliefs are removed from the set. In order to be considered rational, an agent who revises his beliefs must obey the AGM postulates. Taking K to be the agent's initial set of beliefs and * to be the revision operation and letting A be the additional information that the agent encounters, the basic postulates are presented by Robert Koons as follows (2009):

(1) $K*A$ is closed under logical consequence.

(2) A belongs to $K*A$.

(3) $K*A$ is a subset of the logical closure of $K \cup \{A\}$.

(4) If $\neg A$ does not belong to K, then the closure of $K \cup \{A\}$ is a subset of $K*A$.

(5) If $K*A$ is logically inconsistent, then either K is inconsistent, or A is.

(6) If A and B are logically equivalent, then $K*A = K*B$.

(7) $K*(A \& B)$ is a subset of the logical closure of $K*A \cup \{B\}$.

(8) If $\neg B$ does not belong to $K*A$, then the logical closure of $K*A \cup B$ is a subset of $K*(A \& B)$.

Following Andre Fuhrmann's development of the idea of translating AGM into dynamical modal logic, (1988, 1991) de Rijke also showed that the AGM-postulates governing expansion and revision may be translated into the object language of dynamic modal logic (de Rijke 1994). At about the same time, Segerberg demonstrated how the theory of belief revision could be formulated in modal logic. Segerberg merged the static first generation doxastic logic with the dynamics of belief change into 'dynamic doxastic logic' (Segerberg 1995). Doxastic operators in the logic of belief like B_cA may be captured by AGM in the sense that '*A* is in *c*'s belief-set *T*', or $\neg B_c\neg A$ becomes '$\neg A$ is not in *c*'s belief-set *T*'. An immediate difference between the two perspectives is that while AGM can express dynamic operations on belief-sets like expansions ('*A* is in *c*'s belief-set *T* expanded by *D*', i.e., $A \in T+D$), revisions ('*A* is in *c*'s belief-set *T* revised by *D*', i.e., $A \in T*D$), and contractions ('*A* is in *c*'s belief-set *T* contracted by *D*', i.e. $A \in T-D$), no such dynamics are immediately expressible in the standard language of doxastic logic. On the other hand, action languages include operators like $[\nu]$ and $<\nu>$ which are prefixed to a well-formed formula *A*. On Segerberg's interpretation, $[\nu]A$ ($<\nu>A$) mean that 'after every (some) way of performing action ν it is the case that *A*'. By introducing three new operators [+], [*], and [−] into the doxastic language, the three dynamic operations on belief-sets may be rendered as $[+D]B_cA$, $[*D]B_cA$ and $[-D]B_cA$.

After revising the original belief revision theory such that changes of beliefs happen in 'hypertheories' or concentric spheres enumerated according to entrencement Segerberg (1999a, 1999b) provided several axiomatizations of the dynamic doxastic logic together with soundness and completeness results. The dynamic doxastic logic paradigm may also be extended to iterated belief revision as studied by Lindstrøm and Rabinowicz

(1997) and accommodate various forms of agent introspection. A related approach drawn up by van Ditmarsch, van der Hoek and Kooi's new ‘dynamic epistemic logic’ studies how information changes and how actions with epistemic impact on agents may be modeled (Hoek *et al*. 2003, Ditmarsch *et al*. 2007). For a more detailed discussion of belief revision theory, see André Fuhrmann ‘The Logic of Belief Revision’ in this volume.

One may also choose to endow the agents with *epistemic capacities* facilitating special epistemic behaviors. Fagin, Halpern, Moses and Vardi have for instance considered ‘perfect recall’ (Fagin *et al*. 1995): interacting agents' knowledge in the dynamic system may increase as time goes by but the agents may still store old information. The agent's current local state is an encoding of all events that have happened so far in the run. Perfect recall is in turn an epistemic recommendation telling the agent to remember his earlier epistemic states.

There are other structural properties of agents being studied in the literature of dynamic epistemic logics. In an epistemic logic suited for modeling various games of imperfect information van Benthem refers to such properties as ‘styles of playing’ (van Benthem 2000). Properties like ‘bounded memory’, various ‘mechanisms for information updates’ and ‘uniform strategies’, infallibility, consistency etc. have been investigated. Yoav Shoham and Kevin Leyton-Brown’s *Multi-Agent Systems* (2009) provides an updated overview of the relevant literature on game theory and belief revision in a multi-agent setting. Agents as explcitly learning mechanims are also integral parts of Kelly's (1996) computational epistemology and a related approach called modal operator epistemology (Hendricks 2001, 2003). Researchers in artificial intelligence have additionally been trying to describe and specify the behavior of intelligent/rational agents by extensions of epistemic of logic by augmenting logics of time, action and belief with modalities for desires and intentions (see Meyer 2003), in particular, his discussion of the BDI-framework of Rao and Georgeff in Section 5.2).

Conclusion

Recent work in epistemology often involves informal reflection on the semantics of the terms 'know' and 'believe'; understanding knowledge is widely thought to be a matter of having the right semantical story to tell about our use of the word 'know'. So, for example, whether one adopts invariantism, relativism, contextualism or some alternative, depends, at least to some extent, on how one understands the truth conditions for sentences containing the word 'know'. Since epistemic logic is largely the study of precisely this kind of semantics and since epistemic logicians have been concerned with attempting to characterize reasoning under varying conditions and constraints, it seems likely that epistemic logic should play an important role in epistemology.

Furthermore, the problem of defining knowledge is inextricably linked to the task of defining the class of worlds, scenarios, or propositions compatible with what an agent knows. In this respect, the formal semantics of epistemic logic is arguably the most appropriate domain for undertaking this kind of definitional task. This is not to say that epistemologists will all their problems solved or their questions answered by formal techniques. Instead, the point of examining epistemological questions in a formal setting is that the implications of (and problems with) one's views are sometimes more obvious.

Obviously, by itself, tweaking the formal apparatus of epistemic logic does not solve traditional epistemological problems. Epistemic logic can help us to navigate through problems in a systematic fashion by unpacking the logic of the problematic concepts, it can also lead us to recognize problems that we had not anticipated. This is basically analogous to the role that modal logic has played in contemporary metaphysics. At the beginning of this essay I noted the divergence between the goal of defining knowledge and the goal of improving our capacity to improve our ability to gain new knowledge. As we have seen, pursuit of the latter goal has led to advances in our capacity to model the dynamics of epistemic and doxastic states. Elsewhere, (2006) together with Vincent Hendricks, I have argued that the apparent divergence of both enterprises can be reconciled once one recognizes that both goals bear on a third problem, namely the problem of understanding the rationality of inquiry. This problem, of course, is of equal importance to both mainstream and formal epistemologists. Dynamical treatments of epistemic logic and insights from epistemic logicians into the logic of inquiry speak directly to the goal of understanding the rationality of inquiry.

Acknowledgments

I owe a special debt of gratitude to Wiebe van der Hoek, whose detailed criticism of this chapter has improved it considerably. I am very grateful to Vincent Hendricks for his comments and for graciously permitting me to use some of our previously co-authored material in this chapter. Specifically, pages ##MS16-18## are largely taken from our co-authored 2006. Thanks also to Duncan Pritchard, Clifford Hill and Emmanuel Genot for helpful criticism of an earlier draft of this essay.

References

Alchourrón, C.E., Gärdenfors, P. and Makinson, D. (1985). 'On the Logic of Theory Change, *Journal of Symbolic Logic*, 50: 510-530.

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

van Benthem, J.F.A.K. (2000). 'Logic and Game Theory--close encounters of the third kind', in *Prooceedings of LLC99*, Loon, I.v., Mints, G. and Muskens, R. (eds). Stanford: CSLI Publications.

Bull, R. and Segerberg, K. (1984). 'Basic Modal Logic', in Gabbay 1984, 1-88.

Chisholm, R. (1963) The Logic of Knowing, *Journal of Philosophy* 60, 773–795.

van Ditmarsch, H., van der Hoek, W. and Kooi, B. (2007) *Dynamic Epistemic Logic* . Dordrecht: Springer

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

Fuhrmann, A., (1988) *Relevant Logic, Modal Logic and Theory Change*, Ph.D Thesis Department of Philosophy and Automated Reasoning Project, Institute of Advanced Studies, Australian National University, Canberra.

Fuhrmann, A., (1991) On the modal logic of theory change, in Fuhrmann, A. and M. Morreau *The Modal Logic of Theory Change.* Lecture Notes in Artificial Intelligence 465. Springer Heidelberg-Berlin-New

Gabbay, D. and Guenthner F. (1984). *Handbook of Philosophical Logic*, vol. II: *Extensions of Classical Logic*. Dordrecht: D. Reidel Publishing Company.

Gärdenfors, P. (1988). *Knowledge in Flux — Modelling the Dynamics of Epistemic States*. Cambridge: MIT Press.

Gettier ,E (1963)Is Justified True Belief Knowledge? Analysis 23 121-123.

Hendricks, V. F. (2001). *The Convergence of Scientific Knowledge--a view from the limit*. Trends in Logic: Studia Logica Library Series, Dordrecht: Kluwer Academic Publishers.

Hendricks, V. F. (2003). ‘Active Agents’, *Journal of Logic, Language and Information*, van Benthem, J. and van Rooy, R. (eds.). 12, 4: 469-495

Hendricks, V. F. (2005). *Mainstream and Formal Epistemology*. New York: Cambridge University Press.

Hendricks, V. F. and Symons J. (2006). Where's the Bridge? Epistemic Logic and Epistemology, *Philosophical Studies*, 128, 1: 2-26.

Hendricks, V.F and Symons, J. (2007) "Epistemic Logic", *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/logic-epistemic/>.

Hilpinen, R. (2002) ‘Δεοντιχ, Τεμποραλ, ανδ Επιστεμιχ Μοδαλ Λογιχσ□ in Dale Jacquette (ed) *A Companion to Philosophical Logic*. Malden, MA: Blackwell. 491-509

Hilpinen, R. (2006) Jaakko Hintikka on Epistemic Logic and Epistemology in *The Philosophy of Jaakko Hintikka*, ed. by R. E. Auxier, Chicago: Open Court Publishing Company, 783-818.

Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell: Cornell University Press.

Hintikka, J. (1975). 'Impossible Possible Worlds Vindicated', *Journal of Philosophical Logic*, 4: 475-484.

Hintikka, J (2007) Epistemology without Knowledge and Without Belief in J. Hintikka, *Socratic Epistemology: Explorations of Knowledge Seeking by Questioning.* Cambridge: Cambridge University Press. 11-37

Hocutt, M.O. (1972) Is Epistemic Logic Possible? *Notre Dame Journal of Formal Logic* 13, 433–453.

Hoek, W.v.d., Ditmarsch, H., Kooi, B. (2003). 'Concurrent Dynamic Epistemic Logic,' in Hendricks, V.F., Jørgensen, K.F, and Pedersen, S.A. (eds.) *Knowledge Contributors*. Synthese Library, vol. 322. Dordrecht: Kluwer Academic Publishers. 105-173

Hume, D. (1740, 1888) 1976, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge. rev. 2nd. ed., ed. P. H. Nidditch. Oxford: Clarendon Press

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

Kutschera, F. von. (1976). *Einführung in die intensional Semantik*. Berlin: W. de Gruyter

Koons, R., Defeasible Reasoning, *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/reasoning-defeasible/>.

Lemmon, E.J. (1959): 'Is There Only One Correct System of Modal Logic?', *Aristotelian Society Supplementary* Volume XXXIII, 23–40.

Lemmon, E. J. (1977). *An Introduction to Modal Logic*, in collaboration with D. Scott. Oxford: Basil Blackwell Publishers

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

Lewis, D. (1969), *Convention: A Philosophical Study*, Cambridge: Harvard University Press

Lindström, S. and Rabinowicz, W. (1997). 'Extending Dynamic Logic: Accommodating Iterated Beliefs and Ramsey Conditionals within DLL', in Lindahl, L., Needham, P. and Sliwinski, R. (eds.)*For Good Measure*,. *Uppsala Philosophical Studies*, 46: 123-153.

McCarthy, J (1979) Ascribing Mental Qualities to Machines. Technical Report STAN-CS-79-725 Stanford University.

Meyer, J.-J. Ch. (2003). 'Modal, Epistemic and Doxastic Logic,' in: *Handbook of Philosophical Logic* (2nd edition) (D. Gabbay & F. Guenthner, eds.) Vol. 10, Dordrecht: Kluwer, 1-38.

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

Moore, G. E. (1952) *The Philosophy of G. E. Moore*: Second Edition, edited by Paul Arthur Schilpp, New York: Tudor Publishing Company.

Rantala, V. (1975). 'Urn Models: A New Kind of Non-Standard Model for First-Order Logic', *Journal of Symbolic Logic*, 4: 455-474.

Rantala, V. (1982). Quantified Modal Logic: Non-Normal Worlds and Propositional Attitudes. Studia Logica 41 (1)41-66 de Rijke, M. (1994). ‘Meeting Some Neighbours: A Dynamic Modal Logic Meets Theories of Change and Knowledge Representation’, in *Logic and Information Flow*, Eijck, J.v and Visser, A. (eds.). Cambridge: MIT Press, 170–196.

Segerberg, K. (1995). ‘Belief Revision from the Point of View of Doxastic Logic’, *Bulletin of the IGPL*, 3: 535-553.

Segerberg, K. (1999a). ‘The Basic Dynamic Doxastic Logic of AGM’, *Uppsala Prints and Preprints in Philosophy*, 1.

Segerberg, K. (1999b). ‘A Completeness Proof in Full DDL’, in *Philosophical Crumbs: Essays Dedicated to Ann-Mari Henschen-Dahlqvist on the Occasion of her Seventy fifth Birthday*, Sliwinski, R. (ed.). *Uppsala Philosophical Studies*, 49: 195-207.

Shoham, Y. and Leyton-Brown, K. (2009) *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations* Cambridge: Cambridge University Press.

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

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

Vanderschraaf, P. and G. Sillari, (2007) Common Knowledge, *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/common-knowledge/>.

Voorbraak, F. (1993). *As Far as I Know*. Epistemic Logic and Uncertainty, Dissertation, Utrecht University, Utrecht. Published as volume 7 in *Queastiones Infinitae*. Department of Philosophy, Utrecht University.

von Wright, G., H. (1951). *An Essay on Modal Logic*. Amsterdam: North-Holland Publishing Company.

BIO:

John Symons is Associate Professor of Philosophy at the University of Texas at El Paso. His work spans logic, philosophy of mind, epistemology, and metaphysics but focuses on the problem of emergence. In addition to numerous articles and chapters he is the author or editor of eight books including *On Dennett*, (2002) *The Routledge Companion to Philosophy of Psychology*, (2009) *Logic Epistemology and the Unity of Science* (2004). Since 2002 he has been editor of *Synthese.*