# A GUIDE TO THE MODAL LOGICS OF KNOWLEDGE AND BELIEF: PRELIMINARY DRAFT

Joseph Y. Halpern
IBM Research Laboratory
San Jose, CA 95193

Yoram Moses*
Computer Science Department
Stanford University, Stanford, CA 94305
and
IBM Research Laboratory
San Jose, CA 95193

Abstract: We review and re-examine possible-worlds semantics for propositional logics of knowledge and belief with four particular points of emphasis: (1) we show how general techniques for finding decision procedures and complete axiomatizations apply to models for knowledge and belief, (2) we show how sensitive the difficulty of the decision procedure is to such issues as the choice of modal operators and the axiom system, (3) we discuss how notions of common knowledge and implicit knowledge among a group of agents fit into the possible-worlds framework, and (4) we consider to what extent the possible-worlds approach is a viable one for modelling knowledge and belief. As far as complexity is concerned, we show among other results that while the problem of deciding satisfiability of an S5 formula with one knower is NP-complete, the problem for many knowers is PSPACE-complete. Adding an implicit knowledge operator does not change the complexity substantially, but once a common knowledge operator is added to the language, the problem becomes complete for exponential time.

## 1. Introduction

Reasoning about knowledge and belief has long been an issue of concern in philosophy and artificial intelligence (cf. [Hil],[MH],[Mo]). Recently we have argued that reasoning about knowledge is also crucial in understanding and reasoning about protocols in distributed systems, since messages can be viewed as changing the state of knowledge of a system [HM]; knowledge also seems to be of vital importance in cryptography theory [Me] and database theory.

In order to formally reason about knowledge, we need a good semantic model. Part of the difficulty in providing such a model is that there is no agreement on exactly what the properties of knowledge are or should be. For example, is it the case that you know what facts you know? Do you know what you don't know? Do you know only true things, or can something you "know" actually be false?

Possible-worlds semantics provide a good formal tool for "customizing" a logic so that, by making minor changes in the semantics, we can capture different sets of axioms. The idea, first formalized by Hintikka [Hil], is that in each state of the world, an agent (or knower or player: we use all these words interchangeably) has other states or worlds that he considers possible. An agent knows p exactly if p is true in all the worlds that he considers possible. As Kripke pointed out [Kr], by imposing various conditions on this possibility relation, we can capture a number of interesting axioms. For example, if we require that the real world always be one of the possible worlds (which amounts to saying that the possibility relation is reflexive), then it follows that you can't know anything false. Similarly, we can show that if the relation is transitive, then you know what you know. If the relation is transitive and symmetric, then you also know what you don't know. (The one-knower models where the possibility relation is reflexive corresponds to the classical modal logic T, while the reflexive and transitive case corresponds to S4, and the reflexive, symmetric and transitive case corresponds to S5.)

Once we have a general framework for modelling knowledge, a reasonable question to ask is how hard it is to reason about knowledge. In particular, how hard is it to decide if a given formula is valid or satisfiable? The answer to this question depends crucially on the choice of axioms. For example, in the one-knower case, Ladner [La] has shown that for T and S4 the problem of deciding satisfiability is complete in polynomial space, while for S5 it is NP-complete,

* This author's work was supported in part by DARPA contract N00039-82-C-0250.

and thus no harder than the satisfiability problem for propositional logic.

Our aim in this paper is to reexamine the possible-worlds framework for knowledge and belief with four particular points of emphasis: (1) we show how general techniques for finding decision procedures and complete axiomatizat ions apply to models for knowledge and belief, (2) we show how sensitive the difficulty of the decision procedure is to such issues as the choice of modal operators and the axiom system, (3) we discuss how notions of common knowledge and implicit knowledge among a group of agents fit into the possible-worlds framework, and, finally, (4) we consider to what extent the possible-worlds approach is a viable one for modelling knowledge and belief.

We begin in Section 2 by reviewing possible-world semantics in detail, and proving that the many-knower versions of T, S4, and S5 do indeed capture some of the more common axiomatizations of knowledge. In Section 3 we turn to complexity-theoretic issues. We review some standard notions from complexity theory, and then reprove and extend Ladner's results to show that the decision procedures for the many-knower versions of T, S4, and S5 are all complete in polynomial space.* This suggests that for S5, reasoning about many agents' knowledge is qualitatively harder than just reasoning about one agent's knowledge of the real world and of his own knowledge.

In Section 4 we turn our attention to modifying the model so that it can deal with *belief* rather than knowledge, where one can believe something that is false. This turns out to be somewhat more complicated than dropping the assumption of reflexivity, but it can still be done in the possible-worlds framework. Results about decision procedures and complete axiomatiiations for belief parallel those for knowledge.

In Section 5 we consider what happens when operators for *common knowledge* and *implicit knowledge* are added to the language. A group has common knowledge of a fact p exactly when everyone knows that everyone knows that everyone knows ... that *p* is true. (Common knowledge is essentially what McCarthy's "fool" knows; cf. [MSHI].) A group has implicit knowledge of *p* if, roughly speaking, when the agents pool their knowledge together they can deduce p. (Note our usage of the notion of "implicit knowledge" here differs slightly from the way it is used in [Lev2] and [FH].) As shown in [HMI], common knowledge is an essential state for reaching agreements and coordinating action. For very similar reasons, common knowledge also seems to play an important role in human understanding of speech acts (cf. [CM]). The notion of implicit knowledge arises when reasoning about what states of knowledge a group can attain through communication, and thus is also crucial when reasoning about the efficacy of speech acts and about communication protocols in distributed systems.

It turns out that adding an implicit knowledge operator to the language does not substantially change the complexity of deciding the satisfiability of formulas in the language, but this is not the case for common knowledge. Using standard techniques from PDL (Propositional Dynamic Logic; cf. [FL],[Pr]), we can show that when we add common knowledge to the language, the satisfiability problem for the resulting logic (whether it is based on T, S4, or S5) is complete in deterministic exponential time, as long as there at least two knowers. Thus, adding a common knowledge operator renders the decision procedure qualitatively more complex. (Common knowledge does not seem to be of much interest in the in the case of one knower. In fact, in the case of S4 and S5, if there is only one knower, knowledge and common knowledge are identical.)

We conclude in Section 6 with some discussion of the appropriateness of the possible-worlds approach for capturing knowledge and belief, particularly in light of our results on computational complexity.

Detailed proofs of the theorems stated here, as well as further discussion of these results, can be found in the full paper ([HM2]).

## 2. Logics of knowledge

**2.1 Syntax: A logic of any kind needs a language. Although we consider a number of different logics here, the syntax for all of them is essentially the same. We wish to reason about a world consisting of a propositional reality ("nature") and $m$ agents, creatively named $1, \ldots, m$. Given a set of primitive propositions $\Phi = \{P, Q, R, \ldots\}$ and a set of $m$ agents, we define $\mathcal{L}_m(\Phi)$ to be the least set of formulas containing $\Phi$, closed under $\neg$, $\wedge$, and the modal operators $K_1, \ldots, K_m$. Thus, if $p$ and $q$ are formulas of $\mathcal{L}_m(\Phi)$, then so are $\neg p$, $p \wedge q$, and $K_i p$, for $i = 1, \ldots, m$ (where $K_i p$ is read "player $i$ knows $p$"). We use the standard abbreviations $p \vee q$ for $\neg(\neg p \wedge \neg q)$ and $p \supset q$ for $\neg(p \wedge \neg q)$. The *size* of a formula $p$ in $\mathcal{L}_m(\Phi)$, denoted $|p|$, is its length over the alphabet $\Phi \cup \{\neg, \wedge, (, ), K_1, \ldots, K_m\}$.**

* A problem is said to be *complete* with respect to a complexity class if, roughly speaking, it is the hardest problem in that class (see Section 3 for more details).

2.2 Possible-worlds semantics: Following Hintikka [Hil], Sato [Sa], Moore [Mo], and others, we use a *posaible-worlds* semantics to model knowledge. This provides us with a general framework for our semantical investigations of knowledge and belief. (Everything we say about "knowledge* in this subsection applies equally well to belief.) The essential idea behind possible-worlds semantics is that an agent's state of knowledge corresponds to the extent to which he can determine what world he is in. In a given world, we can associate with each agent the set of worlds that, according to the agent's knowledge, could possibly be the real world. An agent is then said to know a fact $p$ exactly if $p$ is true in all the worlds in this set; he does not know $p$ if there is at least one world that he considers possible where $p$ does not hold.

Kripke [Kr] introduced *Kripke structures* as a formal model for a possible-worlds semantics for the modal logic of necessity and possibility. A Kripke structure $M$ is a tuple $(S, \pi, P_1, ..., P_m)$, where $S$ is a set of states, $\pi(s)$ is a truth assignment to the primitive propositions of $\Phi$ for each state $s \in S$ (i.e., $\pi(s, P) \in \{\text{true}, \text{false}\}$ for each primitive proposition $P \in \Phi$ and state $s \in S$), and $P_i$ is a binary relation on the states of $S$, for $i = 1, \ldots, m$. A *Kripke world* (or just *world*) $w$ is a pair $(M, s)$, where $M = (S, \ldots)$ is a Kripke structure and $s \in S$. $P_i$ is intended to capture the possibility relation according to player $i$: $(s, t) \in P_i$ if, in world $(M, s)$, player $i$ considers $(M, t)$ a possible world.

One of the advantages of Kripke-style semantics is that we can view a Kripke structure as a labelled directed graph, where the nodes are the states of $S$, and there is an edge from $s$ to $t$ labelled $i$ exactly if $(s, t) \in P_i$. This graph-theoretic viewpoint will turn out to be particularly useful in our decision procedures (see Section 3).

We now formally define the notion of a formula being true at a world via the relation $\models$, a binary relation between worlds and formulas, where $w \models p$ is read "$p$ is true at $w$" or "$w$ satisfies $p$":

$(M, s) \models P$ (for $P \in \Phi$) iff $\pi(s, P) = \text{true}$

$(M, s) \models p \wedge q$ iff $(M, s) \models p$ and $(M, s) \models q$

$(M, s) \models \neg p$ iff $(M, s) \not\models p$

$(M, s) \models K_i p$ iff $(M, t) \models p$ for all $t$ s.t. $(s, t) \in P_i$.

The last clause in this definition captures the intuition that player $i$ knows $p$ in world $(M, s)$ exactly if $p$ is true in all worlds that $i$ considers possible.

A formula $p$ is said to be *valid* (resp. *satisfiable*) if $w \models p$ for all worlds $w$ (resp. some world $w$). We write $\models p$ if $p$ is valid. Note that $p$ is satisfiable iff $\neg p$ is not valid.

The following well-known theorem captures some of the formal properties of our $\models$:

**Theorem 1:**

(a) All instances of propositional tautologies are valid.

(b) For all formulas $p, q \in \mathcal{L}_m(\Phi)$ and $i = 1, \ldots, m$, the formula $[K_i p \wedge K_i(p \supset q)] \supset K_i q$ is valid.

(c) For all formulas $p, q \in \mathcal{L}_m(\Phi)$, if $\models p$ and $\models p \supset q$, then $\models q$.

(d) For all $p \in \mathcal{L}_m(\Phi)$ and $i = 1, \ldots, m$, if $\models p$ then $\models K_i p$. ⋈

**2.3 Axiom systems for knowledge:** Theorem 1 tells us that by subscribing to Kripke semantics we are forced to accept a number of constraints on the type of notions of knowledge that we can model.* We now show that in a precise sense these are the only constraints that we are forced to accept by using Kripke semantics. We do so by defining an axiom system $K_{(m)}$ corresponding to the above constraints, and then proving that these axioms characterise Kripke worlds.** Such results are well known (cf. [HC,Sa,Ch]). We reprove them here (using techniques originally due to Makinson [Ma]) in order to show the close correspondence between the axioms and a particular Kripke structure which we call the *canonical structure*.

$K_{(m)}$ consists of two axiom schemas:

A1. All tautologies of the propositional calculus

A2. $[K_i p \wedge K_i(p \supset q)] \supset K_i q, \quad i = 1, \ldots, m$

and two rules of inference:

R1. From $\vdash p$ and $\vdash p \supset q$ infer $\vdash q$ (Modus ponens)

R2. From $\vdash p$ infer $\vdash K_i p$ (Generalisation)

A formula $p$ is said to be $K_{(m)}$-*provable*, denoted $K_{(m)} \vdash p$, if $p$ is an instance of one of the axiom schemas, or if $p$ follows from provable formulas by one of the inference rules R1 and R2 (we omit the qualifier $K_{(m)}$ if it is clear from context). A formula $p$ is *consistent* if $\neg p$ is not provable. A finite set of formulas $\{p_1, \ldots, p_k\}$ is consistent exactly if $p_1 \wedge \ldots \wedge p_k$ is consistent, and an infinite set of formulas is consistent exactly if all of its finite subsets are consistent. Of course, a formula or set of formulas is said to be *inconsistent* exactly if it is not consistent. A set $F$ of formulas is a *maximal consistent set* if it is consistent and for all $p \in (\mathcal{L}_m(\Phi) \backslash F)$, the set $F \cup \{p\}$ is inconsistent.

* We discuss the ramifications of this point in Section 6.

** The name $K_{(m)}$ is inspired by the fact that for one knower, the system reduces to the well-known modal logic K.

Using standard techniques of propositional reasoning, we can show

**Lemma 2:** In any axiom system that includes A1 and R1:

(a) every consistent set F can be extended to a maximal consistent set,

(b) if F is a maximal consistent set, then for all formulas $p$, $q$:

(i) either $p \in F$ or $\neg p \in F$,

(ii) $p \wedge q \in F$ iff $p \in F$ and $q \in F$,

(iii) if $p \in F$ and $p \supset q \in F$, then $q \in F$,

(iv) if $p$ is a valid formula then $p \in F$. ⋈

An axiom system $S$ is *sound* with respect to a set of worlds $\mathcal{M}$ if every formula provable from $S$ is valid in $\mathcal{M}$ (i.e., is true in every world in $\mathcal{M}$). $S$ is *complete* with respect to $\mathcal{M}$ if every formula that is valid in $\mathcal{M}$ is provable from $S$. We think of an axiom system as characterising a set of worlds exactly if it provides a sound and complete axiomatisation of that set.

**Theorem 3:** $K_{(m)}$ is a sound and complete axiomatization for Kripke worlds.

**Proof:** Theorem 1 implies that $K_{(m)}$ is sound with respect to Kripke worlds. In order to prove completeness, it suffices to show that every consistent formula is satisfiable. We will do so by constructing a Kripke structure $M^c$, which we call the *canonical* Kripke structure for $K_{(m)}$, containing a state $s_V$ for each maximal $K_{(m)}$-consistent set $V$ in such a way that $(M^c, s_V) \models p$ for all $p \in V$. Since, by Lemma 2, every consistent set of formulas is contained in some maximal consistent set, this suffices. We proceed as follows. Given a set $V$ of formulas, define $V/K_i = \{p \mid K_i p \in V\}$. Let $M^c = (S, \pi, P_1, \ldots, P_m)$, where

$$S = \{s_V \mid V \text{ is a maximally consistent set}\}$$

$$\pi(s_V, P) = \begin{cases} \text{true} & \text{if } P \in V \\ \text{false} & \text{if } P \notin V \end{cases}$$

$$P_i = \{(s_V, s_W) \mid V/K_i \subseteq W\}.$$

We now show, by induction on the structure of $p$, that for all $V$ we have $(M^c, s_V) \models p$ iff $p \in V$. More precisely, assuming that the claim holds for all subformulas of $p$, we will also show that it holds for $p$. If $p$ is a primitive proposition $P$, this is immediate from the definition of $\pi(s_V, P)$ above. The cases where $p$ is a conjunction or a negation are simple and left to the reader. Assume that $p$ is of the form $K_i q$ and that $p \in V$. Then $q \in V/K_i$ and, by definition of $P_i$, if $(s_V, s_W) \in P_i$, then $q \in W$. By the induction hypothesis, $(M^c, s_W) \models q$ for all $W$ such that $(s_V, s_W) \in P_i$. By the definition of $\models$, it follows that $(M^c, s_V) \models K_i q$.

For the other direction, assume $(M^c, s_V) \models K_i q$. It follows that the set $(V/K_i) \cup \neg q$ is inconsistent. Suppose not. Then by Lemma 2 it would have a maximal consistent extension $W$, and, by construction, we would have $(s_V, s_W) \in P_i$. By the induction hypothesis we have $(M^c, s_W) \models \neg q$, and so $(M^c, s_V) \models \neg K_i q$, contradicting our original assumption. Since $(V/K_i) \cup \neg q$ is inconsistent, some finite subset, say $p_1, \ldots, p_k, \neg q$, must be inconsistent. Thus, by propositional reasoning, we have

$$\vdash p_1 \supset (p_2 \supset (\ldots (p_k \supset q) \ldots)).$$

By R2, we have

$$\vdash K_i(p_1 \supset (p_2 \supset (\ldots (p_k \supset q) \ldots))).$$

Since a maximal consistent set contains all tautologies, we must have

$$K_i(p_1 \supset (p_2 \supset \ldots (p_k \supset q) \ldots)) \in V.$$

And since $p_1, \ldots, p_k \in V/K_i$, we must also have

$$K_i p_1, \ldots, K_i p_k \in V.$$

Now, by repeated applications of axiom A2 and Lemma 2(b)(iii), we can easily show that $K_i q \in V$, as desired. ⋈

In the philosophical literature, one finds a great deal of discussion as to which axioms truly characterise knowledge (see [Len] for a discussion and review). Some of the ones more commonly considered include:

A3. $K_i p \supset p, \qquad i = 1, \ldots, m,$

the *knowledge axiom*, which states that only true facts can be known (this is usually taken as the essential property distinguishing knowledge from belief),

A4. $K_i p \supset K_i K_i p, \qquad i = 1, \ldots, m,$

the *positive introspection axiom*, which states that an agent knows what he knows, and

A5. $\neg K_i p \supset K_i \neg K_i p, \qquad i = 1, \ldots, m,$

the *negative introspection axiom*, which says that an agent knows what he does not know.

In the case of a single agent, the system K+A3 is classically known as T, T+A4 is known as S4, while S4+A5 is known as S5. In the case of $m$ agents, $m > 1$, we will call these systems $T_{(m)}$, $S4_{(m)}$, and $S5_{(m)}$ respectively.

Philosophers have spent years trying to determine which of these systems (if any) best captures knowledge (again, see [Len]). In view of Theorem 1, the best

that can be said is that we are modelling a rather idealised reaaoner, who knows all tautologies and all the logical consequences of his knowledge. If we take the classical interpretation of knowledge as true, justified belief, then an axiom such as A3 seems to be necessary. On the other hand, philosophers have shown that axiom A5 does *not* hold with respect to this interpretation ([Len]). However, the S5 axioms do capture an interesting interpretation of knowledge appropriate for reasoning about distributed systems (see [HM1] and Section 6). We continue here with our investigation of all these logics, deferring further comments on their appropriateness to Section 6.

Theorem 3 implies that the provable formulas of $K_{(m)}$ correspond precisely to the formulas that are valid for Kripke worlds. As Kripke showed [Kr], there are simple conditions that we can impose on the possibility relations $P_i$ so that the valid formulas of the resulting worlds are exactly the provable formulas of $T_{(m)}$, $S4_{(m)}$, and $S5_{(m)}$ respectively. We will try to motivate these conditions, but first we need a few definitions.

**We say that a binary relation $P$ on a set $S$ is *reflexive* if $(s, s) \in P$ for all $s \in S$; $P$ is *transitive* if, for all $s, t, u \in S$, if $(s, t) \in P$ and $(t, u) \in P$, then $(s, u) \in P$; $P$ is *symmetric* if, for all $s, t \in S$, whenever $(s, t) \in P$ then $(t, s) \in P$; $P$ is *Euclidean* if, for all $s, t, u \in S$, whenever $(s, t) \in P$ and $(s, u) \in P$, then $(t, u) \in P$; finally, $P$ is *serial* if, for all $s \in S$, there is some $t$ such that $(s, t) \in P$. A relation that is reflexive, symmetric, and transitive, is also commonly called an *equivalence relation*. Some of the relationships between these notions are described by the following lemma (cf. [Ch]):**

**Lemma 4:**

**(a) If $P$ is symmetric and transitive, then $P$ is Euclidean.**

**(b) $P$ is symmetric, transitive, and serial iff $P$ is reflexive and Euclidean iff $P$ is reflexive, symmetric, and transitive.** ⋈

**We say a world $(M, s)$ is *reflexive* (resp. *symmetric*, *transitive*, *Euclidean*, *rt*, *rst*, *serial*) exactly if all the possibility relations in $M$ are reflexive (resp. symmetric, transitive, Euclidean, reflexive and transitive, reflexive, symmetric, and transitive, serial).**

**To see the relationship between these notions and the axioms described above, consider the canonical Kripke structure $M^c$ defined in Theorem 2. Recall that $(s_V, s_W) \in P_i$ in $M^c$ exactly if $V/K_i \subseteq W$, where $V/K_i = \{p \mid K_i p \in V\}$. Now suppose that all instances of A3 are true at $s_V$. Then it is easy to see that $(s_V, s_V) \in P_i$, since $V/K_i \subseteq V$. This suggests that A3 corresponds to reflexivity. Indeed, it is easy to check that A3 is sound in all reflexive worlds. In terms of the possible worlds, if $P_i$ is reflexive, then in world $w$, player $i$ always considers $w$ to be one of his possible worlds. Thus, if in world $w$ player $i$ knows $p$, then $p$ must be true in $w$; i.e., $K_i p \supset p$.**

**A4 forces the possibility relations in the canonical structure to be transitive. To see this, suppose that $(s_V, s_W), (s_W, s_X) \in P_i$ and that all instances of A4 are true at $s_V$. Then if $K_i p \in V$, by A4 we have $K_i K_i p \in V$, and by the construction of $M^c$, we have $K_i p \in W$ and $p \in X$. Thus $V/K_i \subseteq X$ and $(s_V, s_X) \in P_i$ as desired. In terms of possible worlds, a transitive possibility relation says that if in world $w$, player $i$ considers $w'$ possible, and if in $w'$ player $i$ considers $w''$ possible, then in world $w$ player $i$ will already consider $w''$ possible.**

**Similar reasoning shows that axiom A5 forces the possibility relation in the canonical structure to be Euclidean. Note that the possibility relation in the canonical structure is forced to be symmetric by the axiom**

$$\neg p \supset K_i \neg K_i p,$$

**which can be shown to be a consequence of A3 and A5. This corresponds to the observation of Lemma 4(b) that a relation that is both reflexive and Euclidean is also symmetric.***

**Arguments essentially identical to those of Theorem 3 can now be used to show:**

**Theorem 5:**

**(a) $T_{(m)}$ is a sound and complete axiomatisation for reflexive worlds.**

**(b) $S4_{(m)}$ is a sound and complete axiomatisation for rt worlds.**

**(c) $S5_{(m)}$ is a sound and complete axiomatisation for rst worlds.** ⋈

**We close this section with some remarks on Kripke structures for S5 (i.e., the case of one agent). We define two worlds $(M, s)$ and $(M', s')$ to be *equivalent*, written $(M, s) \equiv (M', s')$, if, for all formulas $p$, we have $(M, s) \models p$ iff $(M', s') \models p$.**

* Since Lemma 4(b) says that a relation that is both reflexive and Euclidean must also be transitive, the reader may auspect that axiom A4 ia redundant in S5. Thia indeed ia the caae.

**Proposition 6:** Suppose $M = (S, \pi, P)$, where $P$ is an equivalence relation (so that $M$ is a model of S5), and $s \in S$. Then $(M, s) \equiv (M', s)$, where $M' = (S', \pi', P')$ and $S' = \{t \mid (s,t) \in P\}$ (so $S'$ is the equivalence class of $s$), $\pi'$ is $\pi$ restricted to $S'$, and $P' = \{(t,t') \mid t, t' \in S'\}$.

**Proof:** By a straightforward induction on the structure of formulas. ⋈

Proposition 6 intuitively says that in determining the truth of an S5 formula at a given state $s$, we can restrict our attention to states that are considered possible at $s$. It follows that we can assume without loss of generality that models of S5 have a particularly simple form: $M = (S, \pi, P)$, where for all $s, t \in S$, we have $(s,t) \in P$ (and in particular, S5 is a sound and complete axiomatisation for worlds $(M, s)$ where $M$ has this form). Note that for such models, we do not even have to mention the $P$ relation; we can simply assume that all states are related to each other via $P$. However, these remarks do not hold for $S5_{(m)}$ if $m > 1$.

## 3. Deciding the satisfiability of formulas

In this section we examine the inherent difficulty of determining whether a formula in a given logic is satisfiable. Of course, the problem of determining validity is a closely related one, since $p$ is valid iff $\neg p$ is not satisfiable. We consider this problem using tools from the computational complexity. We briefly review the necessary notions here; the reader should consult [HU] for further details.

The cost of solving a given problem is usually measured by the amount of time and or space (memory) required to compute the solution, as a function of the input size. Since the inputs we consider in this section are formulas, we will typically be interested in how difficult it is to determine if a formula $p$ is satisfiable or valid as a function of $|p|$. We are usually most interested in *deterministic* computations, where at any point in the computation, the next step of a computation is uniquely determined. However, thinking in terms *nondeterministic* computations — ones where the program may "guess" which of a finite number of steps to take — has been very helpful in classifying the intrinsic difficulty of a number of problems. The complexity classes we will be most concerned with here are $P$, $PSPACE$, $EXP$, and $NP$: the problems that are solvable in deterministic polynomial time, deterministic polynomial space, deterministic exponential time, and nondeterministic polynomial time, respectively. It is not hard to show that $P \subseteq NP \subseteq PSPACE \subseteq EXP$; it is also known that $P \neq EXP$. While it is conjectured that all the other subset relations are strict, proving this remains elusive. The $P = NP$ problem is currently considered the most important open problem in the field of computational complexity.

Roughly speaking, a problem $A$ is said to be *hard* with respect to a complexity class $C$ (eg. $NP$-hard, $PSPACE$-hard, etc.) if every problem in $C$ can be effectively *reduced* to $A$; i.e., for any problem $B$ in $C$, an algorithm for $B$ can be easily obtained from an algorithm for $A$. A problem is *complete* with respect to a complexity class $C$ if it is in $C$ and is $C$-hard. A well-known result due to Cook [Co] shows that the problem of determining whether a formula of propositional logic is satisfiable is $NP$-complete. In particular, this means that if we could find a polynomial-time algorithm for deciding satisfiability for propositional logic, we would also have polynomial-time algorithms for all other $NP$ problems. This is considered highly unlikely.

Since propositional logic is a sublanguage of all the logics we have considered, the satisfiability problem for all of them is $NP$-hard. Ladner showed that, at least for S5, it is no harder.

**Theorem 7** ([La]): The problem of deciding S5-satisfiability is $NP$-complete.

The key step in the proof of Theorem 7 lies in showing that satisfiable S5 formulas can be satisfied in very small worlds:

**Proposition 8** ([La]): An S5 formula $p$ is satisfiable iff it is satisfiable in a world $(M, s)$ where $M$ has less than $|p|$ states.

**Proof:** Assume that $p$ is satisfiable, and let $(M, s) \models p$. By Proposition 6 and the remarks following it, we can assume that $M = (S, \pi, P)$, where $(t,t') \in P$ for all $t, t' \in S$. Let $F$ be the set of subformulas of $p$ of the form $Kq$ for which $(M, s) \models \neg Kq$; i.e., $F$ is the set of subformulas of $p$ that have the form $Kq$ and are false at state $s$. For each formula $Kq \in F$, there must be a state $s_q \in S$ such that $(M, s_q) \models \neg q$. Let $M' = (S', \pi', P')$, where $S' = \{s\} \cup \{s_q \mid q \in F\}$, $\pi'$ is the restriction of $\pi$ to $S'$, and $P' = \{(t,t') \mid t, t' \in S'\}$. Note that $|S'| \leq |p|$. We now show that for all states $s' \in S'$ and for all subformulas $q$ of $p$ (including $p$ itself), $(M, s') \models q$ iff $(M', s') \models q$. As usual, we proceed by induction on the structure of $q$. The only nontrivial case is when $q$ is of the form $Kq'$. Suppose $s' \in S'$. If $(M, s') \models Kq'$, then $(M, t) \models q'$ for all $t \in S$, so, in particular, $(M, t) \models q'$ for all $t \in S'$. By induction hypothesis, $(M', t) \models q'$ for all $t \in S'$, so $(M', s') \models Kq'$. And if $(M, s') \not\models Kq'$, then $(M, s') \models \neg Kq'$. Since $M$ is a model of S5, we have $(M, s') \models K\neg Kq'$, so that $(M, s) \models \neg Kq'$ (since $(s', s) \in P$ by assumption). But then it follows that $Kq' \in F$, and $(M, s_{q'}) \models \neg q'$. By construc-

tion, $s_{q'} \in S'$, and by induction hypothesis, we also have $(M', s_{q'}) \models \neg q'$. Since $(s', s_{q'}) \in P'$, we have $(M', s') \models \neg Kq'$, and so $(M', s') \not\models Kq'$ as desired. Since $s \in S'$ and $(M, s) \models p$ by assumption, we also have $(M', s) \models p$. ⋈

**Proof of Theorem 7:** Because the propositional calculus is part of S5, Cook's Theorem implies that deciding S5 satisfiability is *NP*-hard. We now give an *NP* algorithm for deciding S5 satisfiability. Intuitively, given a formula $p$, we simply guess a world $(M, s)$ where $M$ has at most $|p|$ states, and verify that this world does satisfy $p$. More formally, we proceed as follows.

Given a formula $p$, where $|p| = n$, we nondeterministically guess a Kripke structure $M = (S, \pi, P)$, where $S$ is a set of $k \leq n$ states, $(s, t) \in P$ for all $s, t \in S$, and for all $s \in S$ and primitive propositions $P$ not appearing in $p$, $\pi(s, P) = \text{true}$. (Note that the only "guessing" that enters here is in the choice of $k$, and the truth value $\pi(s, P)$ for primitive propositions $P$ that appear in $p$ in the $k$ states in $S$.) Since at most $n$ primitive propositions appear in $p$, guessing such a Kripke structure can be done in nondeterministic time $O(n^2)$ (i.e., $\leq cn^2$ for some constant $c$). Next, we check whether $p$ is satisfied at some state $s \in S$. This can easily be done deterministically in time $O(n^2)$ by simply computing, by induction on structure, whether $(M, s) \models q$ for each state $s \in S$ and each subformula $q$ of $p$. We leave details to the reader. By Proposition 8, if $p$ is satisfiable, one of our guesses is bound to be right. (Of course, if $p$ is not satisfiable, no guess will be right.) Thus we have a nondeterministic $O(n^2)$ algorithm for deciding if $p$ is satisfiable. ⋈

As the following result shows, this technique will not work for deciding satisfiability for the other logics we have been considering.

**Proposition 9:** There is a constant $c > 0$ such that for all $n > 1$, there is a formula $p_n$ with $|p_n| = n$ that is $S5_{(3)}$- (resp. S4-, T-, K-) satisfiable such that $p_n$ is satisfiable only in worlds with $\geq 2^{\sqrt{cn}}$ (resp. $\geq 2^{cn}$, $\geq 2^{\sqrt{cn}}$, $\geq 2^{\sqrt{cn}}$) states. ⋈

In fact, as Ladner shows:

**Theorem 10** ([La]): The problem of deciding the satisfiability of formulas for the logics K, T, and S4 is *PSPACE*-complete. ⋈

It is easy to extend Ladner's techniques to show that for all $m$, $K_{(m)}$, $T_{(m)}$, and $S4_{(m)}$ are *PSPACE*-complete. Interestingly, although S5 is *NP*-complete, we can show that $S5_{(m)}$, for $m \geq 2$, is also *PSPACE*-complete. Thus, for S5, having more than one knower bumps the complexity of deciding satisfiability up from *NP*-complete to *PSPACE*-complete. The upper bound for $S5_{(m)}$ is proved using techniques similar to Ladner's. The lower bound follows from the following observation (which is also used in the proof of Proposition 9):

**Lemma 11:** Let $p \in \mathcal{L}_{(1)}(\Phi)$ (i.e., only one modal operator, $K$, appears in $p$). Let $\tau(p)$ be the $\mathcal{L}_{(2)}(\Phi)$ formula that results from replacing all occurrences of $K$ in $p$ by $K_1 K_2$. Then $p$ is T-satisfiable iff $\tau(p)$ is $S5_{(2)}$-satisfiable (iff $\tau(p)$ is $S4_{(2)}$-satisfiable).

As a consequence, we have:

**Theorem 12:** The problem of deciding the satisfiability of formulas in $K_{(m)}$, $T_{(m)}$, and $S4_{(m)}$ is *PSPACE*-complete. The problem of deciding the satisfiability of formulas in $S5_{(m)}$ for $m \geq 2$ is also *PSPACE*-complete. ⋈

The algorithms that prove that the logics are all in *PSPACE* are based on tableau techniques and depend crucially on the following observation. Define a structure $M = (S, \pi, P_1, \ldots, P_m)$ to be *treelike* if its graph forms a tree, with no backedges. The *r-closure* of $M$ is the structure that results when the possibility relations $P_1, \ldots, P_m$ are replaced by their reflexive closures. We can similarly define the *rt-closure* and *rst-closure* of a structure $M$.

**Proposition 13:** A $K_{(m)}$ (resp. $T_{(m)}$, $S4_{(m)}$, $S5_{(m)}$) formula $p$ is satisfiable iff it is satisfiable in a world $(M, s)$ where $M$ is a treelike structure (respectively, the r-closure, rt-closure, rst-closure of a treelike structure) of depth $\leq |p|$. ⋈

Using Proposition 13, we construct an algorithm that checks if $p$ is satisfiable by constructing the appropriate treelike structure for $p$ depth first, in a space-efficient manner. See [HM2] for details.

## 4. Belief

A number of recent papers (for example [Lev1]) have pointed out that the knowledge represented in a knowledge base is typically not required to be true. Thus the propositional attitude that philosophers have called *belief* seems more appropriate than knowledge for formalising the reasoning and deduction of a knowledge base. Since knowledge bases typically are assumed to have introspective powers, and so know what they know and do not know, this amounts to dropping A3 from the S5 axioms. However, since it is also assumed that knowledge bases do not have inconsistent beliefs, we must add:

A6. -K(false).

(Note that A6 follows from A3 by propositional reasoning, but is independent of the rest of the axioms if

we drop A3.) A6 is also called the axiom D, and the system consisting of A1, A2, A4, A5, A6, R1, and R2 is called **$KD45_{(m)}$ (cf. [Ch]) or *weak* $S5_{(m)}$.**

It now remains to find a model for $KD45_{(m)}$. In terms of possible worlds, the semantic impact of A6 is simply to say that the possibility relations must be serial. Since we have already argued that A3 corresponds to reflexivity, it would seem that we can get a model of $KD45_{(m)}$ simply by considering worlds where the possibility relation(s) are symmetric, transitive, and serial, although not necessarily reflexive. Unfortunately, this won't work; as Lemma 4 shows, any binary relation which is symmetric, transitive, and serial, must also be reflexive.

In the case of one knower, there are well-known ways to get around this problem: we consider a structure where one distinguished state describes what is true in the "real" world, and a set of states corresponds to the worlds that the agent thinks possible (cf. [Lev1]). This is analogous to the case for S5, where as observed in the remarks after Proposition 6, we can, without loss of generality, take a model to be a set of states (all related to each other by the possibility relation P), one of which will be the real world. Thus, in the case of one knower, the difference between knowledge and belief is that, in the case of belief, the real world is not necessarily one of the worlds the agent considers possible. But this approach does not extend to the many-knower case in any obvious way.

The solution to our problem is already implicit in our discussion in Section 2. Recall that axiom A5 corresponds to the possibility relation being Euclidean rather than symmetric. To understand the intuition behind Euclidean relations, observe that for a given state a, if $P$ is Euclidean then the restricton of $P$ to $\{t \mid (s,t) \in P\}$ is reflexive, symmetric, and transitive, i.e., an equivalence relation. Thus, for a Euclidean relation, the worlds that an agent thinks are possible form an equivalence relation, but do not necessarily include the real world. The fact that the relation is serial means that an agent always thinks *some* worlds are possible. Applying exactly the same techniques as those used in Theorem 5 we can now show (cf. [FV]):

Theorem 14: **$KD45_{(m)}$** is a sound and complete axiomatisation for Euclidean, transitive, and serial worlds.

Similarly, using the same techniques as in Theorems 7 and 12, we can show:

Theorem 15: The problem of deciding the satisfiability of KD45 formulas is NP-complete. For $m \geq 2$, the problem of deciding the satisfiability of $KD45_{(m)}$ formulas is PSPACEcomplete.

## 5. Incorporating Common Knowledge and Implicit Knowledge

In a number of situations it is useful to be able to reason about the state of knowledge of a group of agents, not just that of an individual agent. For example, we may want to reason about facts that are part of a group's "culture": not only does everyone know them, but everyone knows that everyone knows them, and everyone knows that everyone knows that everyone knows, and so on. These facts are said to be *common knowledge* . Put another way, these are essentially the facts that "any fool knows" (cf. [MSHI]).

**To capture these notions, we extend the language $\mathcal{L}_m(\Phi)$ by adding two new operators: $E$ and $C$. Thus, if $p$ is a formula, then so are $Ep$ ("everyone knows p") and $Cp$ ("$p$ is common knowledge"). We view $Ep$ as an abbreviation for $K_1p \wedge \ldots \wedge K_mp$, while $Cp$ is intended to represent the infinite conjunction $Ep \wedge EEp \wedge \ldots$. Note that if $m = 1$ then $Ep \equiv Kp$; thus common knowledge becomes interesting only if there are at least two agents.**

**We can capture the intended meaning of these constructs quite straightforwardly in our semantics. Given a structure $M = (S, \pi, P_1, \ldots, P_m)$, we define a state $t$ to be *reachable* from $s$ if there is some sequence $u_0, \ldots, u_n$ of states in $S$ such that $s = u_0$, $t = u_n$, and for all $i = 1, \ldots, n-1$, there is some $j$ such that $(u_i, u_{i+1}) \in P_j$. Then we have**

$$(M,s) \models Ep \text{ iff } (M,t) \models p \text{ for all } t \text{ such that } (s,t) \in P_1 \cup \ldots \cup P_m,$$

**and**

$$(M,s) \models Cp \text{ iff } (M,t) \models p \text{ for all } t \text{ reachable from } s.$$

**Somewhat surprisingly, even though $C$ is an "infinitary" operator, we can give a complete axiomatisation for it. Consider the following set of axioms (cf. [MSHI,Sa,Leh]):**

**A7.** $Ep \equiv K_1p \wedge \ldots \wedge K_mp$

**A8.** $Cp \supset p$

**A9.** $Cp \supset ECp$

**A10.** $[Cp \wedge C(p \supset q)] \supset Cq$

**A11.** $(p \supset Ep) \supset (p \supset Cp)$,

**and the rule of inference**

**R3.** From $\vdash p$ infer $\vdash Cp$.

Let $\mathbf{KC_{(m)}}$ **(resp.** $\mathbf{TC_{(m)}}$, $\mathbf{S4C_{(m)}}$, $\mathbf{S5C_{(m)}}$**)** be the system that results from adding A7-A11 and R3 to the axioms for $\mathbf{K_{(m)}}$ **(resp.** $\mathbf{T_{(m)}}$, $\mathbf{S4_{(m)}}$, $\mathbf{S5_{(m)}}$**).**

Theorem 16: For the language of common knowledge, $KC_{(m)}$ (resp. $TC_{(m)}$, $S4C_{(m)}$, $S50_{(m)}$) is a sound and complete axiomatisation for Kripke worlds (resp. reflexive worlds, rt worlds, rst worlds). ⋈

The common knowledge operator *C* adds a great deal of expressive power to the language. We can now make universal statements about what is true at all reachable worlds in the structure. One of the consequences of this is that the analogues to Theorem 12 and Proposition 13 no longer hold. In fact we have:

Proposition 17: There is a constant $\mathbf{c > 0}$ such that for all n > 2, there is a formula $p_n$ with $\mathbf{|p_n| = n}$ that is $\mathbf{KC_{(2)}}$**- (resp.** $\mathbf{TC_{(2)}}$**-,** $\mathbf{S4C_{(2)}}$**-,** $\mathbf{S5C_{(2)}}$**-)** satisfiable, but is not satisfiable in any world *(M, a)* where *M* is a treelike structure (resp. reflexive closure, rt closure, rst closure of a treelike structure) of depth $\mathbf{< 2^{\sqrt{cn}}}$. ⋈

Theorem 18: For $\mathbf{m \geq 2}$, the problem of deciding the satisfiability of $KC_{(m)}$ (resp. $TC_{(m)}$, $S4C_{(m)}$, $\mathbf{S5C_{(m)}}$) formulas is complete for exponential time. ⋈

The proof of the exponential-time lower bound follows from techniques similar to those used in [FL] to prove a similar bound for PDL. The upper bound can be obtained using techniques of [Pr] or [EH]. In fact, the techniques of [EH] allow us to combine the proof of the correctness of the algorithm with a proof of the completeness of the appropriate axiom system. Further details can be found in [HM2].

Besides the knowledge common to a group of agents, it is also often desirable to be able to reason about the knowledge that is *implicit* in the group, i.e., what someone who could combine the knowledge of all of the agents in the group would know.* Thus, for example, if Alice knows *p* and Bob knows $p \supset q$, then together they have implicit knowledge of q, even though it might be the case that neither of them individually knows *q*. Whereas common knowledge, in McCarthy's analogy, essentially corresponds to what *any fool* knows, implicit knowledge corresponds to what a (fictitious) "wise man" (one that knows exactly what each individual agent knows) would know. Implicit knowledge is a useful notion in describing the total knowledge available to a group of agents in a distributed environment (cf. [HM1]). Intuitively, a group has implicit knowledge of a fact if the knowledge of that fact is distributed among the members of a group. In a closed system, a group of cooperating agents cannot come to know a fact if it is not already implicit knowledge

In order to capture the notion of implicit knowledge in our language, we add a new modal operator $I$ that stands for "implicit knowledge". We can then capture implicit knowledge semantically as follows. Given a Kripke structure $M = (S, \pi, P_1, \ldots, P_m)$, we define

$$(M, s) \models Ip \text{ iff } (M, t) \models p \text{ for all } t \text{ such that } (s, t) \in P_1 \cap \ldots \cap P_m.$$

The intuition behind this definition is that if all the agents could "put their knowledge together", the only worlds they would consider possible are precisely those in the intersection of the sets of worlds that each one individually considers possible. Put another way, if some agent knows that a world *t* is not the real world, then the "wise man" should know this too. Thus the wise man would only consider possible the worlds that all agents consider possible. Note that in the case of a single agent (i.e., $m = 1$), we have $Ip \equiv Kp$; implicit knowledge just reduces to knowledge.

How can we be sure that this definition really does capture our intuitions regarding implicit knowledge? One way is to find a complete axiomatisation. If we view $Ip$ as saying "the wise man knows *p*", one axiom that suggests itself is

A12. $K_i p \supset Ip, \quad i = 1, \ldots, m;$

this axiom is easily seen to be sound with respect to the semantics given for $I$. We also expect the $I$ operator to act like a knowledge operator, and indeed it is easy to see that it satisfies the axiom schema A2:

$$[Ip \wedge I(p \supset q)] \supset Iq.$$

Moreover, if the $P_i$ relations are reflexive, so that knowledge satisfies A3, then so does implicit knowledge; similar remarks hold for A4 and A5. Let $\mathbf{KI_{(m)}}$ (resp. $\mathbf{TI_{(m)}}$, $\mathbf{S4I_{(m)}}$, $\mathbf{S5I_{(m)}}$) be the system that results from adding axiom A12 to the axiom system $\mathbf{K_{(m)}}$ (resp. $\mathbf{T_{(m)}}$, $\mathbf{S4_{(m)}}$, $\mathbf{S5_{(m)}}$) and assuming that $I$ also satisfies axiom schemas A2, A3, A4, and A5 (where applicable). Then we have

**Theorem 19:** For the language of implicit knowledge with $m \geq 2$ knowers, $\mathbf{KI_{(m)}}$ (resp. $\mathbf{TI_{(m)}}$, $\mathbf{S4I_{(m)}}$, $\mathbf{S5I_{(m)}}$) is a sound and complete axiomatisation for Kripke worlds (resp. reflexive worlds, rt worlds, rst worlds). ⋈

* Note that Levesque [Lev2] uses the term "implicit belief" in a somewhat different sense than we do here. In his case, an agent's "explicit" beliefs are not deductively closed, and the agent's "implicit* beliefs are roughly the deductive closure of his explicit beliefs.

We remark that if m = 1, we can get a complete axiomatization for implicit knowledge simply by adding the axiom schema I*p* = *Kp* to the axioms for knowledge.

In the discussion above, we also viewed implicit knowledge as the knowledge the agents would have by pooling their individual knowledge together. This suggests the following rule of inference:

**R4.** **From** $\vdash (q_1 \wedge \ldots \wedge q_m) \supset p$

**infer** $\vdash (K_1 q_1 \wedge \ldots \wedge K_m q_m) \supset Ip.$

Again, this inference rule is easily seen to be sound with respect to the semantics for I given above. Intuitively it says that if $q = q_1 \wedge \ldots \wedge q_m$ implies p, and if each of the agents knows a "part" of *q* (in particular, agent i knows $q_i$), then together they have implicit knowledge of *q,* and thus implicit knowledge of p.

It is easy to check that this inference rule is derivable from axiom A2, A12, and propositional reasoning. Conversely, A12 is derivable from R4 and the other axioms for knowledge. Thus, we can replace A12 by R4 and get another complete axiomatization for implicit knowledge. We omit details here. Finally, we observe that the addition of the I operator does not essentially affect the complexity of the language. We can extend the techniques of Theorem 12 to show:

Theorem 20: For $m \geq 2$, the problem of deciding the satisfiability of $\mathbf{KI}_{(m)}$ (resp. $\mathbf{TI}_{(m)}$, $\mathbf{S4I}_{(m)}$, $\mathbf{S5I}_{(m)}$) formulas is *PS PA* CIS-complete.

## 6. Conclusions

We have investigated various classical modal logics of knowledge and belief. It is reasonable at this point to consider to what extent these logics really do capture our intuitive notions. Our feeling in this regard is that there are several useful notions of knowledge and belief; some of them are captured by these logics, others are not. For example, consider a processor in a given distributed system that has received a certain set of messages (or a robot that has observed a certain set of facts). There are a number of global states of the system ("possible worlds*) that are consistent with the processor having received these messages (or the robot having made these observations). We can say that the processor knows *p* in this case if *p* is true in all these global states. Note that this is an "external" interpretation of knowledge, that does not require a processor to perform any reasoning to obtain knowledge, or even to be "aware" of this knowledge. This interpretation of knowledge precisely satisfies the $S5_{(m)}$ axioms, and turns out to be quite useful in practice (see [HMI] for further discussion).

When it comes to formalizing the reasoning of a knowledge base or of humans, computational complexity must be taken into account. We cannot expect a program to carry out exponential-time algorithms, much less a human! On the other hand, we must be careful in interpreting the lower bounds on complexity we have presented in the previous sections. These are *worst-case* results, and there is no reason to believe that most cases of interest should act like the worst case. Indeed, the evidence suggests that just the opposite is true. The complexity of deciding formulas that humans are interested in tends to be much better than the worst-case analysis would indicate. We have noted that for one-knower S5 and KD45, the decision procedure for satisfiability of formulas is NP-complete, just as it is for propositional logic. Resolution methods have proved to be quite efficient in practice for propositional logic, and it seems that similar techniques can also be applied successfully to S5 and KD45. And the fact that there are successful practical theorem-provers for linear-time temporal logic, a modal logic whose satisfiability problem is *PSPACE*complete, suggests that this is a feasible task even for the many-knower versions of the logics we have been considering.

These observations suggest that the logics we have been considering may provide reasonable approximations to the reasoning carried out by a knowledge base, but they still do not seem realistic models for human reasoning. Humans simply do not seem to be *logically omniscient* [Hi2], in the sense of Theorem 1: they do not know all tautologies, nor is their knowledge closed under deduction (i.e., it does not satisfy $[K_i p \wedge K_i(p \supset q)] \supset K_i q$). A number of attempts have been made to modify the possible-worlds framework to provide a more realistic semantic model of human reasoning. Most of these attempts have involved either allowing non-classical "impossible" worlds in addition to the regular possible worlds [Gr,Ra], using a non-classical truth assignment [Lev2,FH] or enriching the possible worlds with a syntactic "awareness* function [FH]. While none of these attempts appears as yet to provide the definitive solution, they do suggest that there is sufficient flexibility in the possible-worlds approach to make it worth pursuing.

Acknowledgements: We would like to thank Ron Fagin, Bob Moore, Nils Nilsson, and Moshe Vardi, for their helpful comments and criticisms. The first author would also like to thank the students of Stanford course CS400B for numerous interesting discussions on implicit knowledge.

## References


[Oh] B. F. Ohellas, *Modal Logic,* Cambridge University Press, 1980.

[CM] H. H. Clark and C. R. Marshall, Definite reference and mutual knowledge, in A. K. Joshi, B. L. Webber, and I. A. Sag (Eds.), *Elements of Discourse Understanding,* Cambridge University Press, 1981.

[Co] S. A. Cook, The complexity of theorem-proving procedures, in "Proceedings of the 3rd Annual ACM Symposium on the Theory of Computing', 1971, pp. 151-158.

[Or] M. J. Cresswell, *Logics and Languages,* Methuen and Co., 1973.

[EH] E. A. Emerson and J. Y. Halpern, Decision procedures and expressiveness in the temporal logic of branching time, J. C*omput. Systems Sci.* 80:1, 1985, pp. 1-24

[FH] R. Fagin and J. Y. Halpern, Belief, awareness, and limited reasoning, in "Proceedings of the Ninth International Joint Conference on AI (IJCAI-85)", 1985.

[FV] R. Fagin and M. Y. Vardi, An internal semantics for modal logic, in "Proceedings of the 18th Annual ACM Symposium on Theory of Computing", 1985, pp. 305-315.

[FL] M. J. Fischer and R. E. Ladner, Propositional dynamic logic of regular programs, *J. Comput. System Sci.* 18:2, 1979, pp. 194-211.

[HMI] J. Y. Halpern and Y. Moses, Knowledge and common knowledge in a distributed environment, in 'Proceedings of the 3rd ACM Conference on Principles of Distributed Computing", 1984, pp. 50-61; a revised version appears as IBM RJ 4421, 1984.

[HM2] J. Y. Halpern and Y. Moses, A guide to the modal logics of knowledge and belief, to appear as an IBM Research Report, 1985.

[Hil] J. Hintikka, *Knowledge and Belief,* Cornell University Press, 1962.

[Hi2] J. Hintikka, Impossible worlds vindicated, J. *Philosophy* 4, 1975, pp. 475-484.

[HU] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation,* Addison-Wesley, 1979.

[HC] G. E. Hughes and M. J. Cresswell, *An Introduction to Modal Logic,* Methuen, London, 1968.

[Kr] S. Kripke, Semantical considerations of modal logic, *Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik 9,*1963, pp. 67-96.

[La] R. E. Ladner, The computational complexity of provability in systems of modal propositional logic, *Siam Journal on Computing,* 6:3, 1977, pp. 467-480.

[Leh] D. Lehmann, Knowledge, common knowledge, and related pussies, in 'Proceedings of the 3rd Annual ACM conference on Principles of Distributed Computing", 1984, pp. 62-67.

[Len] W. Lensen, Recent work in epistemic logic, *Acta Philosophica Fennica 3*0, 1978, pp. 1-219.

[Levi] H. J. Levesque, A formal treatment of incomplete knowledge bases, Fairchild Technical Report No. 614, FLAIR Technical Report No. 3, 1982.

[Lev2] H. J. Levesque, A logic of implicit and explicit belief, in 'Proceedings of the 1984 National Conference on AI (AAAI-84)", pp. 198-202; a revised and expanded version appears as FLAIR Technical Report No. 32, 1984.

[Ma] D. Makinson, On some completeness theorems in modal logic, *Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik* 12,1966, pp. 379-384.

[MH] J. McCarthy and P. Hayes, Some philosophical problems from the standpoint of artificial intelligence, in *Machine Intelligence 4* (ed. D. Michie), American Elsevier, 1969, pp. 463-502.

[MSHI] J. McCarthy, M. Sato, T. Hayashi, S. Igarashi, On the model theory of knowledge, Computer Science Technical Report STAN-CS-78-657, Stanford University, April 1978.

[Me] M. J. Merritt, Cryptographic protocols, Ph.D. thesis, Georgia Institute of Technology, 1983.

[Mo] R. C. Moore, Reasoning about knowledge and action, Artificial Intelligence Center Technical Note 191, SRI International, 1980.

[Pr] V. R. Pratt, Models of program logics, in 'Proceedings of the 20th IEEE Symposium on the Foundations of Computer Science", 1979, pp. 115-122.

[Ra] V. Rantala, Impossible worlds semantics and logical omniscience, *Acta Philosophica Fennica,* 85, 1982, pp. 18-24.

[Sa] M. Sato, A study of Kripke-type models of some modal logics by Gentsen's sequential method, *Publications of the Research Institute for Mathematical Sciences,* Kyoto University 18:2, 1977.