RJ 5022 (52382) 1/29/86
Computer Science

Research Report

ON EPISTEMIC LOGIC AND LOGICAL OMNISCIENCE

Moshe Y. Vardi

IBM Almaden Research Center
650 Harry Road
San Jose, California 95120-6099

IBM Research Division
Yorktown Heights, New York • San Jose, California • Zurich, Switzerland

RJ 5022 (52382) 1/29/86
Computer Science

ON EPISTEMIC LOGIC AND LOGICAL OMNISCIENCE

Moshe Y. Vardi

IBM Almaden Research Center
650 Harry Road
San Jose, California 95120-6099

ABSTRACT: We consider the logical omniscience problem of epistemic logic. We argue that the problem is due to the way in which knowledge and belief are captured in Hintikka's possible worlds semantics. We describe an alternative approach in which propositions are sets of worlds, and knowledge and belief are simply a list of propositions for each agent. The problem of the circularity in the definition is solved by giving a constructive definition of belief and knowledge worlds. We show how to incorporate notions such as reasoning and context of use in our model. We also demonstrate the power of our approach by showing how we can emulate in it other epistemic models.

"Do I contradict myself?
Very well, then, I contradict myself.
(I am large, I contain multitudes.)"
Walt Whitman

1. Introduction

Epistemic logic, the logic of epistemic notions such as *knowledge* and *belief*, is a major area of research in artificial intelligence (cf. [FH85, Ko80, Le84, MH69, Mo85, XW83]). The major existing formal model of knowledge and belief, originated by Hintikka [Hi62], is based on the *possible worlds* approach. The basic notion of this approach is a set W of alternative worlds. Knowledge and belief of an agent, who is situated in the actual world w, consist of the agent distinguishing a subset W' of W as the set of worlds that are possible alternatives to w. Thus the agent knows or believes p in w if p is true in all the worlds in W'.

A fundamental problem with this model is the so called *logical omniscience* problem [Hi75]. Essentially, the problem is that an agent always knows all the logical consequences of her knowledge. That is, if agent a knows p and if p logically implies q, then a also knows q. In particular, a knows all valid sentences. (The problem with belief is analogous.) This situation is, of course, unintuitive and unrealistic.

One can accept this situation by endowing the epistemic notions with new pre-systematic interpretations. For example, one can restrict oneself to idealized agent with unbounded reasoning power [Mo85], or one can reinterpret knowledge and belief to be *implicit* rather then *explicit*, i.e., a believes p if p follows from a's explicit beliefs [HM84, Le84, RP85]. But this leaves us in want of a precise treatment of knowledge and belief in the customary senses.

As was pointed out by Cresswell [Cr70, Cr72, Cr73], the problem of logical omniscience can be dealt with by allowing *non-classical* worlds in our semantics, that is, worlds in which not all valid (in the standard sense of valid) formulas need be true, or in which inconsistent formulas may be true. (A weaker notion of a similar kind is contained in Kripke's notion of non-normal worlds [Kr65].) Such worlds are called *impossible* in [Hi75] and *nonstandard* in [RB79]. This approach was pursued further by Levesque [Le84], where the non-classical worlds are called, following [BP83], "situations". (Levesque claims that his approach is different from the possible worlds approach. A closer examination, however, shows that Levesque is actually pursuing the nonclassical possible worlds approach advocated by Cresswell.)

The non-classical worlds approach fails, however, to supply a satisfying solution to the logical omniscience problem for two reasons. First, the intuition behind the non-classical worlds, and in particular the inconsistent ones (called in [Le84] "incoherent situations"), is not clear at all. Thus it is far from obvious how to define the semantics of logical connectives in such worlds. Secondly, and worse, it turns out that the addition of non-classical worlds does not make the agents any less logically omniscient, but rather it just changes the logic in which the agents reason. That is, the agents still believe in all the "logical consequences" of their beliefs, where these are not the standard logical consequences but rather the logical consequences in some non-standard logical system. For example, the agents in Levesque's model [Le84] turn out to be perfect reasoners in Anderson's and Belnap's *relevance logic* [AB75]. Unfortunately, it does not seem that agents can reason perfectly in relevance logic any more than in classical logic.

We believe that the roots of the logical omniscience problem lie not only in the structure of the worlds (i.e., whether they are classical or non-classical) but also in the very fundamental way in which epistemic notions are captured in Hintikka's possible worlds approach [Hi62]. Let us have a closer look at the role played by the worlds.

At the most fundamental level, possible worlds theory is a theory that takes alternative possibilities as its basic primitive notion. While this theory is controversial in some circles (see [Ma73, St76, St85] for arguments pro and con), we are willing to accept it. The only assumption that this theory makes is that there are many *conceivable* states of affairs. Hintikka in [Hi62] went further to model knowledge and belief as a a relation between these conceivable states. According to this approach, at any state an agent has in mind a set of states that are *possible* relative to that state (the set of possible states is a subset of the set of all conceivable states). It is this set of possible states that captures the agent's knowledge or belief. Unfortunately, this way of capturing epistemic notions is far from being intuitive, and goes a long way beyond the basic assumption underlying the possible worlds theory. Thus by modelling knowledge and belief the way he did, Hintikka made a dubious metaphysical commitment, whose side-effect is the logical omniscience problem.

In the rest of this paper we study another approach that still has possible worlds as its basic notion, but does not make the metaphysical commitment made by Hintikka. The outline of the paper is as follows. In §2 we describe Montague's approach to modelling epistemic notions (since we wish to make as few metaphysical assumptions as possible we consider belief rather than knowledge), and we describe the shortcomings of that approach. In §3 we describe our approach, and we investigate its properties. In §4 we show how to model different modes of reasoning in our approach. §5 compares our approach with other approaches for modelling belief that have been suggested in the literature. We conclude with some comments in §6.

2. Belief in Propositions

The most simplistic way to model belief is by *belief sets*. Let L be the assertion language. Then a belief set for an agent a is a set of sentences of L. Intuitively, a believes that ϕ, where ϕ is a sentence of L, if and only if ϕ is in the belief set of a. This model makes no assumptions on the nature of belief. In fact not only can a have contradictory beliefs, but also a can even believe in inconsistent sentences such as $p\wedge\neg p$. Representing an agent's beliefs by a set of sentences in the basic idea underlying the models in [Eb74], [Ko83], [Ko85], and [MH79]. This approach, in which "a believes that" is viewed as a predicate on sentences rather than a sentential connective, is called in [Le84] the *syntactic* approach. Let us look now at the semantic analogue of this approach.

Let W be the set of all possible worlds, and suppose that the notion of *satisfaction* of a sentence ϕ in a world w, denoted $w\models\phi$, is defined. For the moment let us not worry about how worlds and satisfaction are defined. The *intension* of a sentence ϕ is the set of worlds in which it is satisfied, i.e., $I(\phi)=\{w : w\models\phi\}$. In this framework, the semantics of ϕ is fully determined by its intension. Thus if two sentences ϕ and ψ have the same intension, i.e., $I(\phi)=I(\psi)$, then they are *semantically equivalent*; in fact, they are *semantically synonymous*. Thus in a semantical model of belief, i.e., a model where agents believe in semantical objects and not in syntactical objects, agents cannot discern between synonymous sentences. So if a believes that ϕ, then she also believes that ψ.

Does such a model solve the logical omniscience problem? It does and it does not. It does because it is no longer the case that if a believes that ϕ and if ϕ semantically implies ψ, then a believes that ψ. In particular, an agent does not have to believe in all valid sentences, and it can believe in contradictory sentences. Nevertheless, if a believes that ϕ and if ϕ is semantically equivalent to ψ, then a believes that ψ. The latter phenomenon is, however, unavoidable as long as one wishes to keep the basic framework of the possible worlds approach and view "a believes that" as a sentential connective that applies to intensions. We refer the reader to [Cr85,Th80] where different approaches to propositional attitudes are taken, and we continue with the study of the possible worlds approach.

In trying to formalize the model, we encounter the difficulty that our informal definitions are in some sense circular. In order to define worlds, we have to define belief, in order to define belief, we have to define semantical equivalence, and in order to define semantical equivalence, we have to define worlds. To get around that difficulty we make the following observation. The relation of semantical equivalence partitions the set of all sentences of L into equivalence classes. Rather than saying that a believes that ϕ, we can say that a believes that $[\phi]$, where $[\phi]$ is the equivalence class of ϕ. Since all sentences in an equivalence class have the same intension, there is a natural correspondence between subsets of W and equivalence classes. Thus rather than list all the equivalence classes in which a believes, we can list the corresponding subsets of W (which are the intensions of the equivalence classes). That is, we can say that an agent's belief is a set of *propositions*, where a proposition is a set of worlds.

This motivates the following definition. We assume a set $\mathbf{P}$ of atomic propositions and a set $\mathbf{A}$ of agents. A *belief structure* is a triple $M=(W,N,\Pi)$, where W is a nonempty set, which we take to be the set of possible worlds, $\Pi:\mathbf{P}\rightarrow 2^W$ gives the intensions of the atomic propositions, and $N:\mathbf{A}\times W\rightarrow 2^{2^W}$ assigns to every agent the set of propositions in which the agent believes in any world.[1]

The language L is the smallest set that contains $\mathbf{P}$, is closed under Boolean connectives and contains $B_a\phi$ ("a believes that ϕ") if ϕ is in L and a is in $\mathbf{A}$. We now can define what it means for a world w in M to satisfy formulas.

- $M,w\models p$, where $p\in\mathbf{P}$, if $w\in\Pi(p)$.
- $M,w\models\neg\phi$ if $M,w\not\models\phi$.
- $M,w\models\phi\wedge\psi$ if $M,w\models\phi$ and $M,w\models\psi$.
- $M,w\models B_a\phi$ if $\{u : M,u\models\phi\}\in N(a,w)$ (that is, $B_a\phi$ is satisfied in w if the intension of ϕ is among the propositions that a believes in w).

What we have described so far is essentially Montague's *intensional logic* of belief [Mo70]. We find it, however, quite unsatisfying, for two reasons. First, this approach leaves the notion of a possible world as a primitive notion, and it does not gives us any intuition about the nature of these worlds. While this might be

[1] If there is only one agent then N can be taken as a function $N:W\rightarrow 2^{2^W}$. Such structures are called in the literature *neighborhood structures* [Se72] or *minimal structures* [Ch80]. They were introduced for technical reasons, which are unrelated to epistemic logic, by Montague [Mo68], who used them to interpret his *pragmatic* language, and Scott [Sc70].

seen as an advantage by the logician whose interest is in epistemic logic, it is a disadvantage for the "user" of epistemic logic whose interest is mostly in using the framework to model belief states. One might say that belief structure are models for epistemic logic and not for epistemic notions. For a further elaboration of this point see [FHV84,FV85]. Secondly, since the above approach does not elaborate on the issue of the nature of the possible worlds, it also leaves open the question of where one gets the set W of possible worlds in the first place. Note that the choice of W is quite significant, since it determines the relation of semantical equivalence, which we want to make as weak as possible. In the next section we describe another approach towards modelling belief in propositions.

3. Belief Worlds

In this section we define belief worlds constructively. This enables us to take W to be the set of *all* possible worlds. As a result, the relation of semantical equivalence will be as weak as it can be, i.e., it will be the relation of logical equivalence.

Basically, a world consists of a truth assignment to the atomic propositions and a collection of sets of worlds. This is, of course, a circular definition, and to make it meaningful we follow the methodology used in [FHV84, FV84, Va85], and define worlds inductively according to their *depth*.[2] A world of depth 0 is a truth assignment to the atomic propositions; a world of depth 1 is essentially a collection of sets of worlds of depth 0; a world of depth 2 is essentially a collection of sets of worlds of depth 1; etc. This process can be carried out to the limit, giving us worlds of infinite depth.

Formally, we define a *0th-order assignment*, f_0, to be a truth assignment $f_0{:}\mathbf{P}\rightarrow\{0,1\}$ to the atomic propositions. We call $<f_0>$ a *1-ary world* (since its "length" is 1). Assume inductively that k-ary worlds $<f_0, \ldots, f_{k-1}>$ have been defined. Let W_k be the set of all k-ary worlds. A *kth-order assignment* is a function $f_k{:}\mathbf{A}\rightarrow 2^{2^{W_k}}$. Intuitively, f_k associates with each agent a set of propositions, where each proposition is a set of k-ary worlds. There is a "compatibility" restriction on f_k's, which we shall discuss shortly. We call $<f_0, \ldots, f_k>$ a *$(k+1)$-ary world.* An infinite sequence $<f_0, f_1, f_2, \ldots>$, where each *prefix* $<f_0, \ldots, f_{k-1}>$ is a k-ary world, is called an *infinitary* world, to distinguish it from *finitary* worlds. W_ω is the set of infinitary worlds (there are uncountably many of them), and $\mathbf{W}$ is the set of all worlds.

The restriction that we mentioned earlier ensures that each level extends the preceding levels. That is, if we take the propositions in the $(k+1)$st level, and chop off their kth level, we should get the propositions of the kth level. To define this restriction formally, we need the following definition. Let $X \subseteq W_k$. Then $chop(X)$ is the set of worlds obtained by chopping off the last level in the worlds in X, that is,

$$chop(X)=\{<f_0, \ldots, f_{k-2}> : <f_0, \ldots, f_{k-2}, f_{k-1}> \in X\}.$$

We can now state the restriction as:

$$f_{k-1}(a)=\{chop(X) : X \in f_k(a)\}.$$

[2] In [FHV84, FV84, Va85] this methodology is used to model epistemic notions in the aforementioned Hintikka style.

The reader might be annoyed by the fact that we can carry our construction to the limit, assuming somehow an infinite amount of information. Nevertheless, in the next section we shall see that infinitary worlds can arise naturally.

We now define what it means for a finitary belief world to *satisfy* a formula of L.

- $<f_0, \ldots, f_r> \models p$, where p is a atomic proposition, if $f_0(p)=1$.
- $<f_0, \ldots, f_r> \models \neg\phi$ if $<f_0, \ldots, f_r> \not\models \phi$.
- $<f_0, \ldots, f_r> \models \phi \wedge \psi$ if $<f_0, \ldots, f_r> \models \phi$ and $<f_0, \ldots, f_r> \models \psi$.
- $<f_0, \ldots, f_r> \models B_a\phi$ if $r \geq 1$ and $\{w : w \in W_r \text{ and } w \models \phi\} \in f_r(a)$ (in particular, $<f_0> \not\models B_a\phi$.)

Since in our belief worlds higher levels always extend lower levels, to determine satisfaction it suffices to consider a long enough prefix. To formalize this statement we need to define *depth* of formulas, which is intuitively the depth of nesting of belief modalities.

- $depth(p)=0$, if p is atomic propositions.
- $depth(\neg\phi)=depth(\phi)$.
- $depth(\phi \wedge \psi)=max\{depth(\phi), depth(\psi)\}$.
- $depth(B_a(\phi))=1+depth(\phi)$.

Lemma 1: Assume that $depth(\phi)=k$ and $r \geq k$. Then $<f_0, \ldots, f_r> \models \phi$ iff $<f_0, \ldots, f_k> \models \phi$. []

Note, however, that the satisfaction relation is defined between all worlds and all formulas, regardless of the arity of the worlds and the depth of the formulas. The reader who is familiar with [FHV84, FV85, Va85] should note the difference in the methodology here and in those papers. Here we consider both finitary worlds and infinitary worlds to be full-fledged objects, while in the aforementioned papers finitary worlds are merely building blocks of infinitary worlds.

We can now define satisfaction for infinitary worlds. We say that the infinitary world $w=<f_0, f_1, \cdots>$ satisfies ϕ, written $w \models \phi$, if $<f_0, \ldots, f_k> \models \phi$, where $k=depth(\phi)$. This is a reasonable definition, since if $w' =<f_0, \ldots, f_r>$ is an arbitrary prefix of w such that $r \geq k$, it then follows from Lemma 1 that $w \models \phi$ iff $w' \models \phi$.

We now want to show that worlds can be extended *conservatively*, i.e., without changing the agents' beliefs. Let us say that two worlds w_1 and w_2 are *equivalent*, denoted $w_1 \equiv w_2$, if they satisfy exactly the same formulas, i.e., if $w_1 \models \phi$ if and only if $w_2 \models \phi$, for all formulas ϕ in L.

Theorem 2: For every k-ary world $w=<f_0, \ldots, f_{k-1}>$ there is a $(k+1)$-ary world $w' =<f_0, \ldots, f_{k-1}, f_k>$, such that $w \equiv w'$. Furthermore, for every k-ary world $w=<f_0, \ldots, f_{k-1}>$ there is an infinitary world $w' =<f_0, \ldots, f_{k-1}, f_k, \cdots>$, such that $w \equiv w'$. []

(This theorem does not have an analogue in [FHV84, FV85, Va85], since there the satisfaction relation between worlds and formulas is a partial relation.)

A formula ϕ is *valid* if it is satisfied by all belief worlds. Clearly if $depth(\phi)=k$, then it suffices to consider $(k+1)$-ary worlds. Since only finitely many worlds need be considered, we get:

Theorem 3: The validity problem for belief worlds is decidable. []

We can also axiomatize validity in the following way:

Theorem 4: The following formal system is sound and complete for validity in belief worlds:

(A1) All substitution instances of propositional tautologies.

(R1) From $\phi \equiv \psi$ infer $B_a\phi \equiv B_a\psi$. []

Thus validity is fully characterized by propositional reasoning plus substitutivity of equivalents. It follows that the logic of belief worlds is the generalization of the modal logic E [Ch80] to include multiple modalities.

We can now relate belief worlds, as defined in this section, to belief structures, as defined in the previous section. Intuitively, what we want to do is to construct a belief structure M so that the worlds in that structure will correspond to belief worlds. The natural choice for the set of possible worlds in M is, of course, the set **W** of all belief worlds. By Theorem 2, we can restrict ourselves, without loss of generality, to W_ω, the set of infinitary worlds. Also, it is easy to define the interpretation for atomic propositions. The nontrivial part is to define N. To do this we need some definitions.

Let $X \subseteq W_\omega$ be a set of infinitary worlds. Then $prefix_k(X)$ is the set of k-ary worlds that are prefixes of worlds in X. That is, $prefix_k(X)$ is the set

$$\{<f_0, \ldots, f_{k-1}> : <f_0, \ldots, f_{k-1}, f_k, \cdots> \in X\}$$

Let $w = <f_0, f_1, \ldots> \in W_\omega$, and $a \in \mathbf{A}$. We define $N_a(w)$ as the collection

$$\{X : X \subseteq W_\omega \text{ and } prefix_k(X) \in f_k(a), \text{for all } k \geq 0\}.$$

We can now define the desired belief structure.

Theorem 5: Let $M_b = (W_\omega, N, \Pi)$ be a belief structure, where W_ω is the set of all infinitary worlds, $\Pi(p) = \{w : w \models p\}$ for $p \in \mathbf{P}$, and $N(a,w) = N_a(w)$. Then $M, \mathfrak{f} \models \phi$ if and only if $\mathfrak{f} \models \phi$, for all formulas ϕ of L. []

Thus our approach has lead us to select a particular belief structure as the standard one, in the sense that for this structure semantical equivalence and logical equivalence are identical.

Theorem 5 claims that there is a belief structure that model the collection of all belief worlds. Thus belief structures are as expressive as belief worlds. One may suspect, however, that belief worlds are not as expressive as belief structures. The reason for this suspicion is as follows. The semantics of a possible world w in a belief structure $M = (W, N, \Pi)$ depends on three parameters: (a) the atomic propositions that are true in w, (b) $N(a,w)$ for each agent a, and (c) the set W. The set W is the *context* in which the agents operate [Cr73, Mo70, St80]; it is the set of possibilities that it is the point of the discourse to distinguish between; it is the set of possibilities that are compatible with the agents' *constraints* [BP83]. The presence of context is what makes belief structure adequate to interpret *pragmatic* languages [Mo68].[3] Belief worlds, on the other hand, seem not to have any notion of variable context. Put differently, belief worlds have a standard fixed context -

[3] According to Morris [Mo38], *pragmatics* is concerned with relations between linguistic expressions, the objects to which they refer, and the contexts of use of the expressions. The other two branches in the study of language are *syntax*, which is concerned with relations among linguistic expressions, and *semantics*, which is concerned with relations between expressions and the object to which they refer.

the set W. Thus it seems that in our effort to find a standard context we have lost the ability to model all other contexts. Surprisingly, this is not the case!

Theorem 6. Let $M=(W,N,\Pi)$ be a belief structure, and let $w\in W$. Then there is a belief world $\mathbf{f}_w$ such that $M,w\models\phi$ if and only if $\mathbf{f}_w\models\phi$, for all formulas ϕ of L. []

Theorem 6 is superficially similar to certain theorems in [FHV84, FV85, Va85], but its proof is quite different. The difference stems from the fact that the theorems in those papers deal with Kripke semantics that does not have the context component that Montague semantics (or rather pragmatics) has.

4. Reasoning

The agents in our model have very little reasoning power. This is clearly demonstrated by the fact that the formulas $B_a\phi\wedge B_a\psi$ and $B_a(\phi\wedge\psi)$ are incomparable, i.e., either of them can be satisfied in some belief world w, while the other is not satisfied in w. Suppose now that we do want to endow the agents with some reasoning power (which is probably the case if the agents we have in mind are supposed to have some intelligence). Here is a sample of formulas that we may want to be valid in our models:

(B1) $\neg B_a$**false**

(B2) B_a**true**.

(B3) $B_a(\phi\wedge\psi)\supset B_a\phi$.

(B4) $B_a\phi\wedge B_a\psi\supset B_a(\phi\wedge\psi)$.

Let us see now what conditions we have to put on belief structures to capture these modes of reasoning. It turns out that B1-B4 can be captured by imposing certain algebraic restrictions on belief worlds $<f_0,\ldots,f_k\cdots>$.

Consider first B1 and B2. The intension of **false** is the empty set, so to ensure that agents do not have false beliefs, we require that $\emptyset\notin f_k(a)$, for all $k\geq1$. The intension of **true** is the set of all worlds, so to ensure that agents believe in true things, we require that $W_{k-1}\in f_k(a)$, for all $k\geq1$. Note though that if a does not believe that **false**, then she does not believe any contradictory formula. Similarly, if a believes that **true**, then she believes all valid formulas.

Consider now B3 and B4. We know that the intension of $\phi\wedge\psi$ is the intersection of the intensions of ϕ and ψ. Thus to capture B3 we require that if $X\subseteq Y\subseteq W_k$ and $X\in f_k(a)$, then also $Y\in f_k(a)$, for all $k\geq1$, and to capture B4 we require that whenever $X,Y\in f_k(a)$ then also $(X\cap Y)\in f_k(a)$, for all $k\geq1$.

Note that B3 and B4 characterize completely different modes of reasoning. B3 characterizes drawing conclusion from individual facts one believes in, while B4 characterizes putting these facts together. If a believes that p when she thinks that the probability that p is true is very high, then a may employ B3 but not B4. The rationale for that is that if $\phi\wedge\psi$ is very probable, then so are ϕ and ψ, but if ϕ and ψ are very probable, it does not necessarily entail that $\phi\wedge\psi$ is very probable. (This explains the *lottery paradox* [Ky61]. One can believe for every particular lottery ticket holder x that x will not win the lottery, while still believing that someone will win the lottery.) We come back to this mode of reasoning in the next section.

Consider now reasoning that agents can do that involves introspection rather than deduction. Introspection about one's own beliefs can take two forms:

(B5) $B_a\phi \supset B_a B_a\phi$

(B6) $\neg B_a\phi \supset B_a \neg B_a\phi$.

In B5 agents are aware of their beliefs and in B6 they are aware of their doubts. For a discussion of epistemic introspection see [Le78]. The conditions needed to capture B5 and B6 are less intuitive than the restriction needed to capture B1-B4. To capture B5 we require that for all $k \geq 1$ if $X \in f_k(a)$, then $\{<g_0, \ldots, g_k> : X \in g_k(a)\} \in f_{k+1}(a)$. To capture B6 we require that for all $k \geq 1$ if $X \notin f_k(a)$, then $\{<g_0, \ldots, g_k> : X \notin g_k(a)\} \in f_{k+1}(a)$.

The restrictions that capture B5 and B6 have the significant consequence that they force the worlds to be infinitary. In other words, in finitary worlds agents do not have unlimited introspection, since introspection, by its reflexive nature, generates arbitrarily deep beliefs.

Finally, we mention how to modify belief worlds to obtain *knowledge* worlds. There is a general agreement that one property that distinguishes knowledge from belief is that the former is by definition correct. That is, if an agent knows that ϕ, then ϕ must be true; otherwise we would not have said that the agents know that ϕ, but rather we would have said that the agent believes that ϕ. It is not hard to see that in order to capture this property we have to require that $<f_0, \ldots, f_{k-1}> \in X$ for all $X \in f_k(a)$.

5. Models for Local Reasoning

It is well known that neighborhood semantics is more expressive than Kripke semantics. To further demonstrate the power of our approach we compare it here to models for belief with a particular mode of reasoning. We show that belief worlds can emulate two variants of Kripke structures.

In the previous section we saw different modes of reasoning that agents may employ. Let us consider again the reasoning characterized by B1, B2, and B3. The interpretation that we gave to that mode of reasoning was that a believe that ϕ if she thinks the the probability that ϕ is true is very high. Another possible interpretation is that agents have multiple "frame of minds" or "dispositions". Thus, a may believe that ϕ in one frame of mind, and she may believe that ψ in another frame of mind, but she never puts ϕ and ψ together to infer that $\phi \wedge \psi$. This mode of reasoning, which we call *local reasoning* following [FH85], has been considered by several authors [FH85, Le85, RB79, St85, Za85], who modelled it by different variants of Hintikka semantics. Our aim in this section is to show that our models and the models studied in the aforementioned papers are essentially equivalent.

The model that we have for local reasoning is belief worlds $<f_0, \ldots, f_k \cdots>$ that satisfy the following constraints, for all $a \in A$ and $k \geq 1$:

- $\emptyset \notin f_k(a)$.
- $W_{k-1} \in f_k(a)$.
- If $X \subseteq Y \subseteq W_{k-1}$ and $X \in f_k(a)$, then $Y \in f_k(a)$.

We call these worlds ***belief worlds for local reasoning*** or, for short, ***LR belief worlds***. We now describe two other approaches to modelling local reasoning that have been suggested in the literature.

Several authors have suggested modelling local reasoning by the following variant of Hintikka semantics. In Hintikka semantics each agents has a set of worlds that are thought as possible alternative. In the suggested variant, each agent has a collection of sets of worlds, where each set constitutes the alternatives in one frame of mind [FH85, Le85, St85, Za85]. An agent a believes that ϕ if ϕ is true in all alternatives in some frame of mind of A. Formally, an *FHLSZ belief structure* is a belief structure, i.e, it is a triple $M=(W,N,\Pi)$, where W is a nonempty set of worlds, $\Pi:\mathbf{P}\rightarrow 2^W$, and $N:\mathbf{A}\times W\rightarrow 2^{2^W}$. Here we interpret N an an assignment to agents of collection of sets of alternatives. Because we interpret N differently from before, the definition of satisfaction is also different.

- $M,w\models p$, where $p\in\mathbf{P}$, if $w\in\Pi(P)$.
- $M,w\models\neg\phi$ if $M,w\not\models\phi$.
- $M,w\models\phi\wedge\psi$ if $M,w\models\phi$ and $M,w\models\psi$.
- $M,w\models B_a\phi$ if for some $U\in N(a,w)$ we have that $M,u\models\phi$ for all $u\in U$.

Note that, even though FHLSZ belief structures look like belief structures, their semantics is closer to Kripke semantics than to Montague semantics.

Rescher and Brandon have suggested using *non-standard* possible worlds to model local reasoning [RB79]. Non-standard worlds are built from other worlds by two operations: *schematization* and *superposition.* The schematization operation combines worlds conjunctively. Thus, p holds in the schematization of u and v, denoted $u\Cap v$, if p holds both in u and v. The superposition operation combines worlds disjunctively. Thus, p holds in the superposition of u and v, denoted $u\Cup v$, if p holds in either u or v. Note that conjunctive worlds can be fuzzy, we can have that neither p nor $\neg p$ holds in $u\Cap v$, while disjunctive worlds can be over determined, we can have that both p and $\neg p$ hold in $u\Cup v$. We now proceed with formal definitions.

Let W be a set of worlds. We define the class of *non-standard possible worlds expressions* (expressions, for short) as the smallest class $\mathbf{E}$ that is closed under the following closure conditions:

- $W\subseteq\mathbf{E}$.
- If $w_i\in\mathbf{E}$ for all i in an index set I, then $\bigcap_{i\in I} w_i\in\mathbf{E}$.
- If $w_i\in\mathbf{E}$ for all i in an index set I, then $\biguplus_{i\in I} w_i\in\mathbf{E}$.

We now define *non-standard* (NS) belief structures. (Note: NS belief structures have not been defined in [RB79]. Our treatment, however, is based on ideas from [RB79].) An NS belief structure is a triple $M=(W,N,\Pi)$, where W is a nonempty set, $\Pi:\mathbf{P}\rightarrow 2^W$, and $N:\mathbf{A}\times W\rightarrow\mathbf{E}$. Here N assigns to each agent a non-standard world that the agent believes is the actual world. The divergence from Hintikka semantics is that there are no alternative worlds here. Rather, the agent believes that she exists in some particular world, albeit a non-standard one.

Satisfaction of formulas in worlds is defined as follows.

- $M,w\models p$, where $p\in\mathbf{P}$, if $w\in\Pi(P)$.
- $M,w\models\neg\phi$ if $M,w\not\models\phi$.
- $M,w\models\phi\wedge\psi$ if $M,w\models\phi$ and $M,w\models\psi$.
- If $N(a,w)=\bigcap_{i\in I} w_i$, then $M,w\models B_a\phi$ if $M,w_i\models\phi$ for all $i\in I$.

- If $N(a,w)= \bigcup_{i \in I} w_i$, then $M,w \models B_a\phi$ if $M,w_i \models \phi$ for some $i \in I$.

Our main result in this section is that the three approaches to local reasoning that we have described above, which seem on the surface to be rather different, are essentially equivalent.

Theorem 7.

(1) Let $\mathfrak{f}$ be an LR belief world. Then there is an FHLSZ belief structure M and a world w in M such that $\mathfrak{f} \models \phi$ if and only if $M,w \models \phi$ for all formulas ϕ of L.

(2) Let M be an FHLSZ belief structure, and let w be a world in M. Then there is an NS belief world N and a world u in N such that such that $M,w \models \phi$ if and only if $N,u \models \phi$ for all formulas ϕ of L.

(3) Let M be an NS belief structure, and let w be a world in M. Then there is an LR belief world $\mathfrak{f}$ such that such that $M,w \models \phi$ if and only if $\mathfrak{f} \models \phi$ for all formulas ϕ of L. []

6. Concluding Remarks

We have presented a framework in which epistemic notions are modelled by sets of propositions, where propositions are sets of possible worlds, and have shown how notions such as reasoning and context of use can be incorporated in the framework. We have also demonstrated the power of our approach by showing how we can emulate in it other epistemic models. Our framework alleviates the logical omniscience problem, but does not solve it completely - we still have substitutivity of equivalents in epistemic contexts. We now suggest two lines of possible attack on the problem.

First, one can try to incorporate the nonclassical worlds approach with our approach (cf. [Cr70, Cr72, RB79]). As we have argued before, this cannot solve the problem, but merely changes the standard logical system into a nonstandard one. An weakening of the logical system might, however, be desirable.

Alternatively, one can accept the fact the epistemic notions are not purely intensional. Rather than discard all semantics and embrace fully the aforementioned syntactic approach, one can try to inject small doses of syntax into the semantics (cf. [Cr75, FH85, Ra82]).

Acknowledgements. Part of this research was done while I was visiting the Center for Study of Language and Information at Stanford University with support from the System Development Foundation. I was inspired to investigate the model described in this paper by stimulating discussions with Ron Fagin and Joe Halpern about their model for limited reasoning [FH85]. I'd also like to thank Ron Fagin, Joe Halpern, Yoram Moses, and Ed Wimmers for their helpful comments on previous drafts of this paper.

REFERENCES

[BP83] Barwise, J., Perry, J.: *Situations and Attitudes*, Bradford Books, Cambridge, 1983.

[Ch80] Chellas, B.F.: *Modal logic*. Cambridge Univ. Press, 1980.

[Cr70] Cresswell, M.J.: Classical intensional logics. *Theoria* 36(1970), pp. 347-372.

[Cr72] Cresswell, M.J.: Intensional logics and logical truth. *J. Philosophical Logic* 1(1972), pp. 2-15.

[Cr73] Cresswell, M.J.: *Logic and Languages*, Methuen, London, 1873.

[Cr75] Cresswell, M.J.: Hyperintensional logic. *Studia Logica* 34(1975), pp. 25-38.

[Cr85] Cresswell, M.J.: *Structured Meanings.* MIT Press, 1985.

[Eb74] Eberle, R.A: The logic of believing, knowing, and inferring. *Synthese* 26(1974), pp. 356-382.

[FH85] Fagin, R., Halpern, J.Y.: Belief, awareness, and limited reasoning. *Proc. 9th Int'l Joint Conf. on Artificial Intelligence*, Los Angeles, 1985, pp. 491-501.

[FHV84] Fagin, R., Halpern, J.Y., Vardi, M.Y.: A model-theoretic analysis of knowledge. *Proc. 25th IEEE Symp. on Foundations of Computer Science*, West Palm Beach, 1984, pp. 268-278.

[FV84] Fagin, R., Vardi, M.Y.: An internal semantics for modal logic. *Proc. 17th ACM Symp. on Theory of Computing* Providence, 1985, pp. 305-315.

[Hi62] Hintikka, J.: *Knowledge and belief.* Cornell Univ. Press, 1962.

[Hi75] Hintikka, J.: Impossible possible worlds vindicated. *J. Philosophy* 4(1975), pp. 475-484.

[HM84] Halpern, J.Y., Moses, Y.: Knowledge and common knowledge in distributed environments. *Proc. 3rd ACM Symp. on Distributed Computing*, 1984, pp. 50-61.

[Ko80] Konolige, K.: A first-order formalization of knowledge and action for a multi-agent planning system. In *Machine Intelligence* 10 (J.E.Hayes, K.Michie, and Y-H. Pao, eds.), Ellis Howrood Limited, Chichester, England, 1982.

[Ko83] Konolige, K.: A deductive model of belief. *Proc. 8th Int'l Joint Conf. on AI*, 1983, pp. 377-381.

[Ko80] Konolige, K.: Belief and incompleteness. In *Formal Theories of the Commonsense World* (J.R. Hobbs and R.C. Moore, eds.), Ablex Publishing Company, 1985, pp. 359-404.

[Kr75] Kripke, S.A.: Semantical analysis of modal logic II : Non-normal propositional calculi. In *The theory of models* (L. Henkin and A. Tarski, eds.), North-Holland, 1965, pp. 206-220.

[Ky61] Kyburg, H.: *Probability and the logic of rational belief.* Wesleyan Univ. Press, 1961.

[Le78] Lenzen, W.: Recent work in epistemic logic. *Acta Phil. Fenn.* 30(1978), pp. 1-219.

[Le84] Levesque, H.J.: A logic of implicit and explicit belief. Proc. National Conf. on Artificial Intelligence, 1984, pp. 198-202.

[Le85] Levesque, H.J.: *Global and local consistency and completeness of beliefs.* Unpublished manuscript, 1985.

[Ma73] Mackie, J.L.: *Truth, probability, and paradox.* Clarendon Press, 1973.

[MH69] McCarthy, J., Hayes, P.J.: Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence* 4 (B. Meltzer and D. Michie, eds.), Edinburgh Univ. Press 1969, pp. 463-502.

[MH79] Moore, R.C., Hendrix, G.: Computational models of beliefs and the semantics of belief-sentences. Technical Note 187, SRI International, AI Center, 1979.

[Mo38] Morris, C.: *Foundations of the theory of signs.* Univ. of Chicago Press, 1938.

[Mo68] Montague, R.: Pragmatics. In *Contemporary Philosophy* (R. Kalibansky, ed.), La Nuova Italia Editrice, Florence, 1968, pp. 102-121.

[Mo70] Montague, R.: Universal grammar. *Theoria* 36(1970), pp. 373-398.

[Mo85] Moore, R.C.: A formal theory of knowledge and action. In *Formal Theories of the Commonsense World* (J.R. Hobbs and R.C. Moore, eds.), Ablex Publishing Company, 1985, pp. 319-358.

[Pu70] Purtill, R.L.: Believing the impossible. *Ajatus* 32(1970), pp. 18-24.

[Ra82] Rantala, V.: Impossible worlds semantics and logical omniscience. *Acta Phil. Fenn.* 35(1982), pp. 106-115.

[RB79] Rescher, N., Brandon, R.: *The logic of inconsistency.* Rowman and Littlefield, 1979.

[RP85] Rosenschein, S.J., Pereira, C.N.: *Knowledge and action in situated automata.* Forthcoming.

[Sa82] Saarinen, E.: Propositional attitudes are not attitudes towards propositions. *Acta Phil. Fenn.* 35(1982), pp. 131-162.

[Sc70] Scott, D.: Advice on modal logic. In *Philosophical problems in logic* (K. Lambert, ed.), Reidel, 1970, pp. 143-173.

[Se72] Segerberg, K.: *An essay on classical modal logic.* Uppsala, Philosophical Studies, 1972.

[St76] Stalnaker, R.C.: Possible worlds. *Nous* 10(1976), pp. 65-75.

[St80] Stalnaker, R.C.: Logical semiotics. In *Modern logic - A survey* (E. Agazzi, ed.), Reidel, 1980, pp. 439-456.

[St85] Stalnaker, R.C.: *Inquiry.* MIT Press, 1985.

[Th80] Thomason, R.H.: A model theory for propositional attitudes. *Linguistics and Philosophy* 4(1980), pp. 47-70.

[Va85] Vardi, M.Y.: A model-theoretic analysis of monotonic knowledge. *Proc. 9th Int'l Joint Conf. on Artificial Intelligence*, Los Angeles, 1985, pp. 509-512.

[XW83] Xiwen, M., Weide, G.: W-JS: A model logic of knowledge. *Proc. 8th Int'l Joint Conf. on AI*, 1983, pp. 398-401.

[Za85] Zadrozny, W.: *Explicit and implicit beliefs - a solution of a problem of H. Levesque.* Unpublished manuscript, 1985.