ALL WE BELIEVE FAILS IN IMPOSSIBLE WORLDS

A possible-world semantics for a "knowing at most" operator

by

S. Ben-David and Y. Gafni

Technical Report #593
November 1989
All We Believe Fails in Impossible Worlds
A possible-world semantics for a "knowing at most" operator

Shai Ben-David
Yael Gafni

Department of Computer Science
Technion
Haifa 32000
Israel

November 1989

Abstract

We extend the familiar possible-world semantics of modal logic by considering the "impossible" worlds of a Kripke structure. We obtain a simple semantics for Levesque's "All I know" logic. We provide a natural proof theory and prove the expected soundness and completeness theorems.

From a mathematical point of view we offer a natural generalization of modal logic that significantly strengthens its expressive power. Considered in the context of Knowledge Representation, such a logic is a standard monotonic logic that allows formal treatment of default reasoning.

*e-mail 'shai@techsel.bitnet', phone (972-4)294266.
1 Introduction

Autoepistemic logics, logics that enable reasoning about one's knowledge (or beliefs), should provide means to express properties of knowledge inside the formal language. An important property is being able to separate, (syntactically and semantically), true facts from facts believed to be true. It is possible that some true facts are not known, and it might also be the case that some false fact is believed to be true. Another property the language should account for is introspection. If an agent knows some fact, does he know that he knows it? Does he know that he doesn't know other facts? In Modal logic these properties are expressed quite naturally. Much work has been done using modal logic as a general framework for modeling knowledge. [1] [2]

Classical modal logic does not account for another important property of knowledge, namely it's being non monotonic. Sometimes acquiring knowledge of a new fact totally changes the state of knowledge: sentences formerly believed to be true are now believed to be untrue and vice versa. This is the case whenever default considerations are involved. Default assumptions are invoked by the lack of knowledge (rather then by some 'positive' piece of information). A logic that attempts to capture such properties should provide tools to express non-knowledge of sets of sentences.

Recently Levesque [6] has offered an elegant tool for this purpose. He augments the classical modal logic by a new modal operator $O\alpha$ with the intended meaning "$\alpha$ is all that I know". In order to interpret the $O\alpha$ operator Levesque reexamines the notion of Belief. The classical belief operator $B\alpha$ actually says that $\alpha$ is believed and maybe more, in other words at least $\alpha$ is believed. Under this interpretation if I believe at least $\alpha$ then I also believe anything that is a consequence of $\alpha$. But believing at least $\alpha$ gives us no information about sentences "stronger" or "not related" to $\alpha$, as we can see that both $\{B(\gamma \rightarrow \alpha), \alpha, B\gamma\}$ and $\{B(\gamma \rightarrow \alpha), \alpha, \neg B\gamma\}$ are satisfiable. Having defined a notion of 'believing at least' it is natural to define the dual notion of believing at most, which is denoted by a new operator $N$. Intuitively "I believe at most in $\alpha$" means that I believe $\alpha$ or less, in other words $\alpha$ is an upper bound on beliefs, the "strongest" sentence I possibly believe. Therefore anything "stronger" than $\alpha$ is not Believed. But by symmetry to the $B$ operator believing at most $\alpha$ gives us no information about sentences "weaker" then $\alpha$. Levesque defines Only $\alpha$ to be "believe at least $\alpha$ and believe at most $\alpha". Thus $O\alpha = (B\alpha \land N\alpha)$.

Levesque gives both a formal semantics and a proof theory for this extended modal language. Levesque's semantics is based on infinite structures of cardinality continuum that are defined somewhat implicitly, we feel it is difficult to gain intuition through them. Levesque's proof theory has the deficiency of not being expressible within the language. One of his axioms-

"$N\neg \alpha \rightarrow \neg B\alpha$ where $\alpha$ is any objective sentence that is falsifiable"

contains a metalogical notion "falsifiable".

This work began as a pursue of Levesque's approach to knowledge operators. It turned out that we have found ourselves defining new versions of modal logic and Kripke models. These new versions are of interest from a purely logical point of view. With a minimal addition to the language and to the definition of truth in a model, we have largely expanded the expressibility power of the logic. Properties of graphs that are not definable in modal logic, asymmetry and irreflexibility for example, are naturally definable in our extended logic. We postpone a thorough discussion of these issues to some other opportunity, here we shall concentrate on the autoepistemic implications of the logic.

The new logic provides an alternative semantics and proof system for Levesque's language. Quite suprisingly, our minor variation on the notion of a Kripke model allows a complete representation of the "all I know" operators. Consistent with our view of the $N$ op-
erator as a negative image of the modal \( B \) (or necessity) operator, we develop a proof theory for the new logic that is based on the classical axiom systems of modal logic. Our proof theory is purely inside the logic.

Similar to the situation in classical modal logic, our extended modal logic gives rise to a wide variety of axiom systems and their corresponding classes of models. Levesque's semantics can be viewed as a special case of such a class of models. We have soundness and completeness theorems to support our semantics and deductive calculus.

In section 2 we give a formal definition of the language and the semantics. In section 3 we show properties of the new operators and show that they capture their intended intuitive meaning. In section 4 we present the proof system. Section 5 is the technical heart of this paper - the soundness and completeness results. In section 6 we introduce special kinds of models that are significant for non-monotonic reasoning. In section 7 we compare our logic with Levesque's. In the last section we show how default reasoning can be done in the new logic.

2 The new language and it's semantics

Our logic will be based on two modal operators: \( B \) with the intended meaning \( \text{believe at least} \), and \( N \) with the intended meaning \( \text{believe at most} \). We would like to concentrate on the modality issues so we limit our discussion to the propositional part of the logic.

A formal definition of the language

Primitive symbols

1. \( p, q, r, \ldots \) [propositional variables]
2. \( \neg, \wedge, \vee, \rightarrow \) [propositional connectives]
3. \( B, N \) [monadic modal operators]
4. \( (,) \) [brackets]

A sentence is composed recursively by applying the propositional connectives and the two modal operators \( B, N \).

We use the following symbols as abbreviations:

1. \( O\alpha = B\alpha \wedge N\alpha \)
2. \( \Box^*\alpha = B\alpha \wedge N\neg\alpha \)
3. \( \Diamond^*\alpha = \neg\Box^*\neg\alpha \)

Definition of the semantics

We define a model \( M \) to be the triple \((W, R, V)\) where \( W \) is a nonempty set of "worlds", \( R \) is a binary relation over \( W \) that describes the accessibility graph, and \( V \) is an assignment of truth values to propositional variables in each world.

We define the notion of satisfiability in a model recursively:

1. for every variable \( p \), \((M, w) \models p \iff V(p, w) = 1 \)
2. Propositional connectives are interpreted by their usual truth tables.

3. \((M, w) \models B\alpha \text{ iff } (M, w') \models \alpha \text{ for every } w' \text{ s.t. } (w, w') \in R\)

4. \((M, w) \models N\alpha \text{ iff } (M, w') \not\models \alpha \text{ for every } w' \text{ s.t. } (w, w') \not\in R\)

Satisfiability and validity are defined as usual:
A sentence \(\alpha\) is satisfiable if there exists a model \(M\) and a world \(w\) in it s.t. \((M, w) \models \alpha\). A sentence \(\alpha\) is valid in a model \(M\) (denoted \(\models_M \alpha\)) iff \((M, w) \models \alpha\) for every \(w \in W\).
A sentence \(\alpha\) is valid if it is valid in every model \(M\).

3 Expressing the notion of believing at most

In this section we show that the formal semantics for the knowing at least and knowing at most operators captures the intuition behind them. We exhibit properties of these operators and investigate the relation between these properties and characteristics of models.

3.1 The meaning of the operators

The knowledge operator \(B\), views the accessibility relation as connecting a world \(w\) to the worlds that are 'possible' on the basis of \(w\)'s knowledge. One's knowledge can be inferred from the set of worlds considered possible.

You 'know at most' \(\alpha\) iff once a world satisfies \(\alpha\), you can't rule it out, i.e. you have to accept it as a possible world.

Definitions: Let \(M\) be a model and \(w\) a possible world.

1. The at least belief set of \(w\) is defined as the set of sentences \(w\) believes at least to be true: \(B(w) = \{\alpha : M, w \models B\alpha\}\)

2. The at most belief set of \(w\) is defined as the set of sentences \(w\) believes at most to be true: \(N(w) = \{\alpha : M, w \models N\alpha\}\)

3. \(\alpha\) is a logical consequence of \(\beta\) in a model \(M\), denoted \(\models_M \alpha\), iff for every world \(w \in W\) in the model \(M, M, w \models \alpha \rightarrow \beta\).

4. The set of logical consequences of \(\alpha\) in a model \(M\) is \(\text{Consequence}_M(\alpha) = \{\beta : \alpha \models_M \beta\}\)

Believing at least in \(\alpha\) means I believe at least in any sentence that is a consequence of \(\alpha\). The following proposition reflects this property:

Proposition 1: Let \(M\) be a model and \(w\) a world in it, then \(M, w \models B\alpha \iff \text{Consequence}_M(\alpha) \subseteq B(w)\)

Believing at most \(\alpha\) means \(\alpha\) is an upper bound on beliefs, so any sentence believed must be a consequence of \(\alpha\). The following proposition reflects this property:

Proposition 2: Let \(M\) be a model and \(w\) a world in it, then \(M, w \models N\alpha \Rightarrow B(w) \subseteq \text{Consequence}_M(\alpha)\)

The only if direction of proposition 2 does not always hold. Proposition 3 shows under what circumstances it is correct.
3.2 Extending the operators for sets of sentences

We have seen that our logic can express knowing at most single formulas. Sometimes we may wish to have a formal notion for expressing that an infinite set of formulas is an upper (or lower) bound on our knowledge.

Definitions: Let $X$ be a set of sentences, $M$ a model and $w$ a world in it, then
- $M, w \models BX$ iff $\forall \alpha \in X, M, w \models B\alpha$.  
- $M, w \models NX$ iff for every $w', s.t. (w, w') \notin R$ there is a sentence $\alpha \in X$ s.t. $M, w' \not\models \alpha$. 
We can also generalize believing exactly for a set of sentences based on the above definitions: 
- $M, w \models OX$ iff $M, w \models BX$ and $M, w \models NX$.

Note that $BX, NX$ and $OX$ are not formulas in the formal language, however, for finite sets $X$ we can regard $BX$ as an abbreviation to the formula $\bigwedge_{\alpha \in X} B\alpha$, and $NX$ as standing for the formula $N(\bigwedge_{\alpha \in X} \alpha)$. For such $X$, $OX$ is of course a formula too.

All the propositions in the previous section are true also for sets of sentences.

The intuition of a possible world is that the state described by this world agrees with ones knowledge, so if the state described by this world is possible, it is reasonable to require the world be an accessible world. One can define a class of models where every possible world is also an accessible one.

Definition: A model $M = (W, R, V)$ is called an $N$-standard model if it fulfills the following constraint: for every $w, w' \in W$, $(w, w') \in R$ iff $B(w) \subseteq w'$.

Proposition 3: A model $M$ is $N$-standard iff for every world $w$ in it and every set of formulas $X$, $B(w) \subseteq \text{Consequence}_M(X) \Rightarrow M, w \models NX$

As a corollary of the above propositions we have

Corollary 1: Let $M$ be an $N$ standard model, $w$ a world in it and $X$ a set of formulas, then $M, w \models OX \iff \text{Consequence}_M(X) = B(w)$

4 The Axiomatization

Classical modal logic provides us with a variety of modal systems each characterized by a class of models, and an axiomatization. The weakest system - $K$, is characterized by the class of all models. The system we introduce is analogous to $K$ in the sense that it is also characterized by the class of all models, consequently it is also the weakest system among various systems which can be defined over the new language. We name it $K_{BN}$.

The classical modal language can be embedded in our language as any of the following natural sublanguages. The language without $N$ operators which we denote $B$-basic, and the language without $B$ operators which we denote $N$-basic. The theorems of the $B$-basic language should correspond exactly to theorems of the standard modal language. So for the classical modal operator $B$ we have the usual axiom from the modal system $K$. For the modal operator $N$, note that $B$-basic sentences are symmetrical to $N$-basic sentences in the following manner: if $B\alpha$ is true in a world $w$ in model then $N\neg\alpha$ will be true in the same world in the complementary model (the complementary model of $M = (W, R, V)$ is defined as $(W, R', V)$, where $R'$ is the complement of the accessibility relation $R$). Following this idea we get the axiom characterizing $N$.
There is a resemblance between our logic, and the logic of two knowers. In the logic of two knowers there are two knowledge operators \( K_1 \) and \( K_2 \), one for each knower, and two accessibility relations: \( R_1 \) and \( R_2 \). Our models are similar but the relations \( R_1 \), \( R_2 \) are viewed as an accessibility and inaccessibility relations respectively, and thus must be complementary to each other. This constraint is not definable in modal logic, so we must settle for a weaker constraint - we enforce the union of the two relations to be an equivalence relation. The third and fourth axioms take care of this.

The System \( K_{BN} \)

Axioms

1. substitutions in propositional tautologies.

2a. \( B(\alpha \rightarrow \beta) \rightarrow (B\alpha \rightarrow B\beta) \)

2b. \( N\neg(\alpha \rightarrow \beta) \rightarrow (N\neg\alpha \rightarrow N\neg\beta) \)

3. \( \Box^*\alpha \rightarrow \alpha \)

4. \( \Diamond^*\Diamond^*\alpha \rightarrow \Box^*\alpha \)

Transformation rules

1. MP : if \( \vdash \alpha \) and \( \vdash \alpha \rightarrow \beta \) then \( \vdash \beta \)

2. Generalization : if \( \vdash \alpha \) then \( \vdash \Box^*\alpha \)

\( K_{BN} \), as we said, is the weakest system corresponding to the new language. We can obtain extensions of \( K_{BN} \) by adding extra axioms, such as the axiom for positive introspection \( B\alpha \rightarrow BB\alpha \), the axiom for negative introspection \( \neg B\alpha \rightarrow B\neg\alpha \), etc. By these additions we obtain systems resembling \( S4, S5 \) of classical modal logic. Symmetrically we can also add extra axioms corresponding to the \( N \) operator and obtain further extensions of the original system. In this way we can construct a new family of modal systems, each of them characterized by a class of models.

5 Soundness and Completeness of \( K_{BN} \)

**Theorem 1:** Every theorem of \( K_{BN} \) is valid in our semantics.

**Proof:** The proof is by induction on the length of the derivation. It is easy to see that all the axioms are valid, and the transformation rules preserve validity.

5.1 The completeness theorem

The proof of the completeness consists of building a model from a collection of maximal consistent sets. Towards this construction we need some preliminary notions.

**Definitions:** Let \( U \) be a collection of maximal consistent sets of sentences w.r.t. some fixed axiom system (when none is specified we mean \( K_{BN} \)).

1. For \( x \in U \) \( B(x) = \{ \alpha : B\alpha \in x \} \)
2. For $x \in U$ $\text{NF}(x) = \{ \alpha : \text{NF} \in x \}$

3. We define a relation $\sim$ on $U$: for $x, y \in U$, $x \sim y$ iff $B(x) \subseteq y$ or $\text{NF}(x) \cap y = \emptyset$.

**Lemma 1:** $\sim$ is an equivalence relation.

**Proof:** For any binary relation $R$, $R$ is an equivalence relation iff $R$ is reflexive and euclidian, so it suffices to show that $\sim$ is reflexive and euclidian. Using axiom 3 (\(\Box^* \alpha \rightarrow \alpha\)) one can show that $\sim$ is reflexive and using axiom 4 (\(\Diamond^* \alpha \rightarrow \Box^* \alpha\)) it can be shown that $\sim$ is euclidian.

**Definitions:** Let $U$ be a collection of consistent sets of sentences w.r.t. $K_{BN}$. A function $f : U \times U \rightarrow \{0,1\}$ is sensible if for every $x, y \in U$, if $f(x, y) = 1$ then $B(x) \subseteq y$, and if $f(x, y) = 0$ then $\text{NF}(x) \cap y = \emptyset$.

Intuition: consider a model who’s worlds are maximal consistent sets of sentences. Having $f(x, y) = 1$ is a commitment that $y$ will be accessible to $x$ in the model. $f(x, y) = 0$ is a commitment that $y$ will be inaccessible to $x$ in the model. If $f(x, y)$ is undefined we have not yet decided whether $y$ should be accessible or inaccessible to $x$. A function $f$ is sensible if the accessibility relation it reflects, agrees with the belief states of the maximal consistent sets. Given a partial function $f$ we would like to be able to extend it to a function from $U \times U$. The following lemma states when this can be done.

**Lemma 2:** Let $U$ be a collection of consistent sets of sentences w.r.t. $K_{BN}$. Every sensible function whose domain is a connected graph, can be extended to a sensible function on all of $U \times U$.

**Proof:** For every two sets $x, y \in U$ if $f$ is defined for them then $x \sim y$. As $\sim$ is an equivalence relation, and the domain of $f$ is a connected graph, it follows that $U$ is an equivalence class w.r.t. $\sim$. Thus $f$ can be extended to all of $U \times U$ and still conserve its sensibility property.

**Theorem 2:** (Completeness theorem) For every set of sentences $X$ and a sentence $\alpha$, if $X \vdash \alpha$ (in our semantics) then $X \vdash \alpha$ (in $K_{BN}$).

**Proof:** It suffices to show that for any $K_{BN}$-consistent set of sentences $X$ there is a model $M$ and a world $w$ in it s.t. $M, w \models X$. This holds by lemma 3.

**Lemma 3:** Let $X$ be a set of $K_{BN}$-consistent sentences, then there is a model $M$ and a world $w$ in it s.t. $M, w \models X$.

**Proof:** We construct a model for a given consistent set of sentences $X$ by defining a sequence of approximations $< M_i : i \in \omega >$, where each $M_i$ is a pair $(W_i, f_i)$. $W_i$ is a collection of maximal consistent sets of sentences, each consistent set representing a possible world. $f_i$ is a partial function from $W_i \times W_i$ to $\{0,1\}$.

The first approximation $(W_0, f_0)$ is determined by the consistent set $X$ for which the model is being constructed. $W_0 = \{w_0\}$, where $w_0$ is any extension of $X$ to a maximal consistent set. (The existence of such a set follows from Lindenbaum’s lemma). $f_0$ is the empty function.

Given the approximation $M_i$ we define $M_{i+1}$ as follows:

For each world $w$ that was added to the $i$-th approximation we add to the $i+1$-th approximation two sets of worlds:
A set of worlds that support w’s at least beliefs:
For each sentence α s.t. ¬Bα ∈ w define a world w', as an extension of B(w) ∪ {¬α} to a maximal consistent set. Such a set exists by lemma 5.
For each such world fi+1(w, w') = 1.

A set of worlds that support w’s at most beliefs:
For each sentence α s.t. ¬Nα ∈ w define a world w', as an extension of ¬N(w) ∪ {α} to a maximal consistent set. Such a set exists by lemma 6.
For each such world fi+1(w, w') = 0.

Definition of the model
We define M = (W, R, V):

1. The set of worlds of the model is: W = ∪i∈wWi
2. The accessibility relation R is determined according to any f which is a sensible extension of ∪i∈wfi. such an extension exists by lemma 2.
   For each pair of worlds w1, w2 ∈ W, (w1, w2) ∈ R iff f(w1, w2) = 1, and (w1, w2) ∉ R iff f(w1, w2) = 0.
3. The assignment of truth values to propositional variables in each world is determined by membership in the corresponding maximal consistent set: V(p, w) = 1 ↔ p ∈ w.

To end the proof, We still have to show that in a model, any sentence α is true in a world w iff α ∈ w, and this holds by lemma 4.

Lemma 4: Let M = (W, R, V) be a model defined as above, then for any sentence α and any w ∈ W, (M, w) |= α ↔ α ∈ w.

Proof: The proof is by induction on the construction of a sentence of KBN. We first note that if α is an atom, the lemma holds by the definition of V. We then prove that if the theorem holds for α, β then it holds for ¬α, α ∧ β, Bα and Nα.

To complete the construction of the model and the proof of lemma 4 we need the following lemmas:

Lemma 5: Let w be a maximaly consistent set of sentences, then if ¬Bα ∈ w then there exists a maximally consistent set w' s.t. B(w) ∪ {¬α} ⊆ w'.

Lemma 6: Let w be a maximaly consistent set of sentences, then if ¬Nα ∈ w then there exists a maximally consistent set w' ∈ W s.t. (N(w) ∪ {¬α}) ∩ w' = ∅.

Further Completeness results
It seems natural to define extensions of KBN, corresponding the classical T, S4, S5, etc. We can prove the expected completeness theorems by repeating the above procedure and checking that the resulting models belong to the appropriate classes (of graphs).

6 Special classes of models

In this section we present two modal systems that have special relevance to nonmonotonic reasoning.
6.1 canonical models

We have seen in section 3, that the relation \( \models_M \) reflects the intended meaning of the \( B \) and \( N \) operators. We would like to show similar results for the \( \models \) relation, and consequently, through a completeness theorem, for the provability relation in our logic.

**Definitions:** A model \( M = (W, R, V) \) is canonical iff for every satisfiable set of sentences \( X \) there is a world \( w \in W \) s.t. \( M, w \models X \).

**Proposition 4:** \( M \) is a canonical model iff for every set of sentences \( X \), \( \models_M X \) iff \( \models X \).

As it turns out finitely canonical models do not exist for our semantics. To see this suppose \( \{p, \neg q\} \subseteq X \) and \( \{Bq, Np\} \subseteq Y \). \( (X, Y) \) are obviously satisfiable. Suppose \( w_1 \) is a world in a model \( M \) that satisfies \( X \) and \( w_2 \) is a world in the same model that satisfies \( Y \). \( w_2 \) can not see \( w_1 \) because \( w_1 \) does not support all of \( w_2 \)'s at least beliefs - \( M, w_2 \models Bq \) but \( M, w_1 \models \neg q \). \( w_2 \) must not see \( w_1 \) but this is also not possible because \( w_1 \) does not support all of \( w_2 \)'s at most beliefs - \( M, w_2 \models NP \) but \( M, w_1 \models p \). Thus \( X \) and \( Y \) cannot share the same model.

As we cannot construct a canonical model we settle for a weaker notion. We define a model in which every set of propositional sentences is satisfied.

**Definitions:** A model \( M = (W, R, V) \) is propositionally canonical (PC for short) iff every satisfiable set of propositional sentences is satisfied in one of the worlds of the model.

Requiring every satisfiable set of sentences to be satisfied in the model is the same as requiring every truth assignment to be represented by some world in the model. Note that such models are uncountable.

**Definitions:** A model \( M = (W, R, V) \) is propositionally finitely canonical (PFC for short) iff every satisfiable propositional sentence is satisfied in one of the worlds of the model.

Requiring every satisfiable sentence to be satisfied in the model is the same as requiring every assignment to a finite subset of propositional variables be represented by some world in the model. The cardinality of such models can be countable.

Using a compactness argument we can get the following.

**Theorem 3:** A set of (modal) formulas is satisfiable in a PC model iff it is satisfiable in a countable PFC model.

**An axiom system for PFC models**

We introduce a set of axioms and show that it is sound and complete w.r.t. the class of PFC models.

Let \( \{p_1, p_2, \ldots\} \) be the propositional variables in the language. The set of axioms is:

\[
C = \{ \bigwedge^n (L_1 \land \cdots \land L_n) \mid n \in \omega, L_i \in \{p_i, \neg p_i\} \}
\]

**Theorem 4:** The axiom system \( K_BN \cup C \) is sound and complete w.r.t. The class of countable PFC models.

**Corollary 2:** The axiom system \( K_BN \cup C \) is sound and complete w.r.t. The class of PC models.
6.2 The class of \( \text{SiSo} \) models

We introduce a class of models where both the accessibility and inaccessibility relations are transitive and euclidian.

Definitions: A model \( M = (W, R, V) \) is a member of the class \( \text{SiSo} \), if both the accessibility relation \( R \) and the inaccessibility relation \( \overline{R} \) are transitive and euclidian.

These models have an interesting structure. The set of worlds divides into two subsets:
1. A set of worlds that are accessible to every world in the model (called sinks).
2. A set of worlds accessible to no world in the model (called sources).
Thus these are Sink-Source models, or \( \text{SiSo} \) models for short.

Another property of \( \text{SiSo} \) models is that all the worlds in the model assign the same truth value to subjective sentences, so actually we can speak of the belief set of a model rather than a belief set of a specific world in the model.

The weak \( S5_{BN} \) axiom system

We add the following axioms to the original axiom system presented in section 4:
1. \( B\alpha \rightarrow \Box^*B\alpha \)
2. \( \neg B\alpha \rightarrow \Box^*\neg B\alpha \)

Theorem 5: The axiom system weak \( S5_{BN} \) is sound and complete w.r.t. \( \text{SiSo} \) models.

7 Comparison to Levesque’s Logic

In this section we comment about the relation between our logic and that of Levesque. We refer the reader to [6] for a full description of Levesque’s logic.

Levesque’s emphasis is on the operator "All I know is ...", consequently his logic is tailored around properties of the knowledge operators. We offer a general framework that can be adapted to a wide variety of interpretations and applications. As we shall see below, both his semantics and proof theory can be viewed as a special case of our extension of modal logic. Moreover, as our semantics is a natural extension of Kripke’s semantics and our proof theory is, likewise, just a variation on the classical modal logic proof systems, we feel that they are easy to gain intuition to and to work with.

The Semantics

Levesque has offered two versions of semantics for his language [5] [6]. Here we discuss [6] as the semantics there seems to have more intuitive appeal and is closer to our semantics.

A Levesque (propositional) model can be presented as a \( \text{SiSo} \) model having two additional properties: (i) Its set of worlds is the set of all propositional truth assignments. (ii) The set of 'sinks' is a 'maximal set of assignments' in a sense described there. It follows that every Levesque model has cardinality continuum. The following theorems basically show that our class of countable PFC \( \text{SiSo} \) models is semantically equivalent to Levesque’s semantics.

Theorem 6: Let \( \Gamma \) be a set of sentences. If \( \Gamma \) is satisfiable in Levesque’s semantics then there exists a countable \( \text{SiSo} \) model \( M = (W, R, V) \) and a possible world \( w \) in it s.t. \( (M, w) \models \Gamma \).

Theorem 7: Let \( \Gamma \) be a set of sentences. If \( M \) is a PFC N-standard \( \text{SiSo} \) model and \( w \in W \) is a possible world such that \( (M, w) \models \Gamma \) then \( \Gamma \) is satisfiable in Levesque’s semantics.
The Proof Systems

Levesque uses as the following axiom scheme:
"N¬α ↝ ¬Bα where α is any objective sentence that is falsifiable".
As mentioned in [6] this axiom is problematic. It is a metalogical statement. For the predicate
logic it ruins the recursiveness of the proof system and for the propositional language it
distracts the tractability of the deductive calculus.

We replace it by axiom system C, which is in the language, and does not affect the
tractability of the proof system.

Corollary 3: A sentence α is provable in weak $S5_{BN} \cup C$ iff α is provable in Levesque's
proof system.

The Completeness Results

Levesque restricts his completeness theorem "X is consistent then X is satisfiable" to finite
sets X. Our completeness results apply to all X. (consequently we get for all X and α,
$X \models \alpha$ iff $X \vdash \alpha$). Corollary 3 shows that this result holds for his logic as well.

8 Default Reasoning

A default argument has the form: "If nothing better is known, then α", or more precisely,
"If on the basis of the available information you can't rule α out, then believe α". Let us
show how this can be modelled in our systems. The default rule is naturally expressed as
¬B¬α \rightarrow Bα. We shall show that adding such a formula to a data base that does not imply
¬α will make it imply Bα.

Lemma 7: If $D \cup \{\alpha\}$ is a consistent set of propositional formulas then $N(D) \models_{PC} \neg B\neg \alpha$.

Proof: Let $M$ be a PC model and let $w$ be a world in $M$ satisfying $N(D)$. It follows
that for every $w'$ in $M$ in which $D$ is satisfied $(w, w') \in R$. As $D \cup \{\alpha\}$ is a consistent set of
propositions and $M$ is canonical, there exists some $w^* \in M$ satisfying $D \cup \{\alpha\}$. It now follows
that $(w, w^*) \in R$ so $(M, w) \models \neg B\neg \alpha$.

Theorem 8: If $D \cup \{\alpha\}$ is a finite consistent set of prop. formulas then $N(D) \cup \{\neg B\neg \alpha \rightarrow
B\alpha\} \vdash_{KBNC} B\alpha$.

Proof: The theorem follows from the above lemma by the completeness theorem for PC
systems.

Remark: Note that in the lemma we did not demand that D is finite. This demand is
introduced in the Theorem because it is stated in syntactical terms and we do not have a
syntactical expression for $N(D)$ for infinite $D$'s.

Acknowledgements

We would like to thank Joe Halpern for stimulating discussions and for detecting several flaws
in a previous version, Hector Levesque, Gerhard Lakemeyer and Steve Bellantoni for drawing
our attention to the need for some 'canonicity' of the models, and Rina Dechter and Michal
Seidmann for their helpful comments.
References


