EPISTEMIC SEMANTICS FOR FIXED-POINTS NON-MONOTONIC LOGICS

Fangzhen Lin
Yoav Shoham
Department of Computer Science
Stanford University

Abstract

Default Logic and Autoepistemic Logic are the two best-known fixed-points non-monotonic logics. Despite the fact that they are known to be closely related and that the epistemic nature of Autoepistemic Logic is obvious, the only semantics that have been offered for Default Logic to date are complex and have little to do with epistemic notions [Etherington 1987]. In this paper we provide simple uniform epistemic semantics for the two logics. We do so by translating them both into a new logic, called GK, of Grounded Knowledge, which embodies a modification of preference semantics [Shoham 1987]. Beside their simplicity and uniformity, the semantics have two other advantages: They allow easy proofs of the connections between Default Logic and Autoepistemic Logic, and suggest a general class of logics of which the two logics are special cases.

1 INTRODUCTION

Existing nonmonotonic logics can be divided roughly into two kinds. One is based on fixed points, like default logic [Reiter 1980] or autoepistemic logic ([Moore 1983], [Konolige 1988], and [Levesque 1989]). The other is based on minimal models, like circumscription ([McCarthy 1980, 1984] and [Lifschitz 1987]) or minimal knowledge ([Halpern and Moses 1984], [Shoham 1987], and [Lin 1988]). While all those that are based on minimal models have a clear and intuitive semantics, some of those that are based on fixed-points do not have one yet. In spite of the fact that autoepistemic logic has a clear possible worlds semantics, and a close correspondence has been established between default and autoepistemic logics [Konolige 1988], similar semantics for default logic are still missing.[1]

[1]Etherington (1987) provides the first semantics for default logic. The semantics seem complicated and have little to do with autoepistemic logic, in spite of the close relationship between the two logics.

In this paper we provide a possible-world epistemic semantics for default logic as well as for autoepistemic logic. Among other things, the semantics clearly show the relationship, both similarities and differences, between the two logics. The semantics are largely in the spirit of Shoham's preference semantics [Shoham 1987]. In fact, they are largely motivated by the failure of Shoham's early attempt to provide preference semantics for default logic. Shoham's key idea is to translate a closed default

$$p : q/r$$

into something like

$$Kp \wedge \neg K \neg q \supset Kr$$

with knowledge (K) minimized. Although the translation works for normal defaults, i.e. when q and r are identical, it does not work for the general case. Intuitively speaking, the translation fails due to its wrong interpretation of the above default as "if p is known and $\neg q$ is not known (q is consistent) to be true, then r is also known to be true." This wrong interpretation does not meet the requirement that knowledge be not only minimal but also *grounded.* Our new semantics can be thought of accounting for groundedness of knowledge by the following informal reading of the same default rule: "if p is known to be true and $\neg q$ is not *assumed* to be true (q is *assumed* to be consistent), then r is known to be true – provided the assumption of the consistency of q can eventually be justified."

This suggests employing *two* epistemic modalities, K (for knowledge) and A (for assumption), which will be related to one another. Indeed, we will translate the above default rule into the sentence

$$Kp \wedge \neg A \neg q \supset Kr$$

and the process of justifying the assumption that q is consistent will correspond to first minimizing the knowledge K with A fixed, and then comparing the resulting knowledge and assumptions to see whether they agree.

Interestingly, the treatment of autoepistemic logic will be almost identical. It was shown in [Konolige 1988] that any autoepistemic theory can be transformed into an equivalent one in which every sentence has the form

$$Lp \wedge \neg L \neg q_1 \wedge \ldots \wedge \neg L \neg q_n \supset r$$

We will translate each such sentence into

$$Ap \wedge \neg A \neg q_1 \wedge \ldots \wedge \neg A \neg q_n \supset Kr$$

and keep the rest as in the default logic translation. Thus the only change required when moving from default logic to autoepistemic logic is the replacement of the first K by an A!

We now proceed to make this precise. We first define our basic logical language, which is simply a propositional one augmented by two epistemic modalities. Then in section 3 we define the logic GK of Grounded Knowledge, which is the result of imposing an augmented version of preference semantics on our basic language. In section 4 we offer translations of both default logic and autoepistemic logic into GK, translations that are similar but subtly different. We prove the correctness of the translations, and show how the common framework explicates the relationhsip between the two logics. Finally in section 5, we outline some of our ongoing further work.

2 THE BASIC LOGICAL LANGUAGE

Our language is a propositional one, augmented with two modalities K and A, which are at this point mutually unconstrained. Well-formed formulas are defined as usual. Intuitively $K\varphi$ means that φ is known or believed (the distinctions between the two are not important in this paper) to be true, while $A\varphi$ means that φ is assumed to be true.

A *Kripke structure* is a tuple (W, π, R_K, R_A), where W is a nonempty set, $\pi(w)$ a truth assignment to the primitive propositions for each $w \in W$, and R_K, R_A are binary relations over W (the accessibility relations for K and A, respectively). A *Kripke interpretation* M is a pair $(w, (W, \pi, R_K, R_A))$, where $w \in W$ and (W, π, R_K, R_A) is a Kripke structure.

We have not placed restrictions on the two accessibility relations so far. Indeed, our results are surprisingly insensitive to the properties of these relations. In particular, the reader may assume any of the major systems that have been used to capture epistemic notions – $S5$, $K45$, $KD45$ or $S4$.

The satisfaction relation "$\models$" between Kripke interpretations and formulas is defined as follows:

1. $(w, (W, \pi, R, T)) \models p$ iff $\pi(w)(p) = 1$, where p is a primitive proposition.
2. $M \models \varphi_1 \vee \varphi_2$ iff $M \models \varphi_1$ or $M \models \varphi_2$.
3. $M \models \neg\varphi$ iff it is not the case that $M \models \varphi$.
4. $(w, (W, \pi, R_K, R_A)) \models K\varphi$ iff $(w', (W, \pi, R_K, R_A)) \models \varphi$ for any $w' \in W$ such that $(w, w') \in R_K$.
5. $(w, (W, \pi, R_K, R_A)) \models A\varphi$ iff $(w', (W, \pi, R_K, R_A)) \models \varphi$ for any $w' \in W$ such that $(w, w') \in R_A$.

We say that a Kripke interpretation M is a *model* of a set of formulas S if M satisfies every member of S.

We conclude this section with two definitions that will be used throughout the remainder of the paper:

$$K(M) = \{\varphi \mid M \models K\varphi,\ \varphi \text{ is a base formula}\}$$

$$A(M) = \{\varphi \mid M \models A\varphi, \varphi \text{ is a base formula}\}$$

where a *base formula* is one that does not contain modal operators.

3 PREFERENCE SEMANTICS AND THE LOGIC GK

In this section we modify the semantics of our basic language, defining the logic GK. We will augment the semantical framework for nonmonotonic logics developed by Shoham (1987). Shoham's basic idea is that a nonmonotonic logic (called a *preference logic* by Shoham) can be viewed semantically as the result of imposing a partial order (called a *preference relation*) "$\sqsubset$" on interpretations of a monotonic logic. The syntax of the resulting logic is unchanged, but the definition of semantic entailment changes: Φ now entails φ if φ is true in all models of Φ that are minimal according to that partial order (but not necessarily in all models of Φ). This new notion of entailment is called *preferential entailment*, and is denoted by $\models_{\sqsubset}$. We start by endowing our basic language with such preference semantics.

To do so we define a preference relation on the Kripke interepretations defined in the previous section. The following relation "$\sqsubset$" has the effect of minimizing knowledge with assumptions fixed.

Definition 3.1 *Let M_i, $i = 1, 2$, be two Kripke interpretations. We say that M_1 is preferred over M_2, written $M_1 \sqsubset M_2$, if*

1. $A(M_1) = A(M_2)$
2. $K(M_1) \subset K(M_2)$.

M is a *minimal* model of S if M is a model of S and there is no other model M' of S such that $M' \sqsubset M$.

These semantics, however, are not sufficient for our purposes, as they do not capture the property of "groundedness" of knowledge. As was said in the introduction, the assumptions made by the agent should be justified. We capture this justification by an added requirement of the minimal models – that the assumptions coincide with the knowledge. This naturally leads to the following definition:

Definition 3.2 *Let S be a set of formulas, and M a Kripke interpretation. We say that M is a* preferred model *of S if*

1. *M is a minimal model of S, and*
2. $K(M) = A(M)$

Definition 3.3 *The logic GK is defined as follows:*

Syntax: The syntax of our basic language.

Semantic entailment: $\Phi \models_{GK} \varphi$ *iff* φ *holds in all preferred models of* Φ.

Example 3.1 In the following, p, q, r are primitive propositions.

$S_1 = \{Kp\}$ has a unique preferred model in the sense that M is a preferred model of $\{Kp\}$ iff $K(M)$ is the tautological closure of p, and $A(M) = K(M)$.

$S_2 = \{Kp \vee Kq, Kp \wedge \neg A\neg r \supset Kr, Kq \wedge \neg A\neg r \supset Kr\}$. We can distinguish two kinds of preferred models of S_2: One in which $Kp \wedge \neg Kq$ and Kr are true; and the other in which $Kq \wedge \neg Kp$ and Kr are true. Notice that Kr is true in all preferred models.

$S_3 = \{\neg Ap \supset Kp\}$ has no preferred models. For let M be a preferred model of S_3. If Ap is true in M, then Kp must be false because of the minimality of M. But then $K(M) \neq A(M)$ and M can not be a preferred model of S_3. Now if Ap is false in M, then Kp must be true in M because it is a model of S_3, this again will make $K(M) \neq A(M)$. ■

As we will see in the next section, S_1 corresponds to the default theory $(p, \emptyset)$, and S_3 to $(\emptyset, \{: \neg p/p\})$. There is no default theory that corresponds to S_2.

We conclude this section by proving a generalization of Reiter's Theorem 2.1 in [Reiter, 1980].

Theorem 1 *Let* $S = S_1 \cup S_2$ *be a set of formulas such that every member of* S_1 *has the form* Kp *and every member of* S_2 *the form*

$$Kp \wedge Aq \wedge \neg Ar_1 \wedge \ldots \wedge \neg Ar_n \supset Ks$$

where p, q, r_1, ..., r_n, s *are base formulas,* $n \geq 0$, *and both* Kp *and* Aq *may be absent (either* S_1 *or* S_2 *may be empty). Then a Kripke interpretation* M *is a preferred model of* S *iff*

1. $K(M) = A(M)$;
2. $K(M) = E_1 \cup E_2 \cup \ldots$, *where* E_i, $i = 1, 2, \ldots$, *are defined inductively as follows:*
   (a) $E_1 = \{p \mid Kp \in S_1\}$.
   (b) $E_{i+1} = Th(E_i) \cup \{s \mid Kp \wedge Aq \wedge \neg Ar_1 \wedge \ldots \wedge \neg Ar_n \supset Ks \in S_2$, *where* $p \in E_i, q \in A(M)$ *and* $r_1, r_2, \ldots, r_n \notin A(M)\}$, *where* $i > 0$ *and* $Th(E_i)$ *is the tautological closure of* E_i.

4 FIXED-POINTS NONMONOTONIC LOGICS

The two major fixed-points nonmonotonic logics in the AI literature are Reiter's default logic [Reiter, 1980] and Moore's autoepistemic logic [Moore, 1985]. By the above Theorem 1 and Theorem 2.1 in [Reiter, 1980], it is straightforward to translate default theories into sets of formulas in our logic GK. As we shall show in this section, a closely related but subtly different translation also exists for autoepistemic theories.

4.1 Default logic

Default logic was proposed by Reiter (1980) as a formalism for default reasoning. Default logic was originally defined with respect to a first-order language. For the sake of consistency with autoepistemic logic, the default logic adopted here is with respect to a propositional language. It is straightforward to extend the results to the first-order case (this is true because open defaults are defined as a shorthand for sets of closed defaults. See the concluding section for a discussion of quantifications over default rules). In this subsection, we define sentences as base formulas.

According to Reiter, a default theory, adopted to a propositional language, is a pair (W, D), where W is a set of sentences and D a set of defaults, which are expressions of the form

$$p : q_1, \ldots, q_n / r$$

where $p, q_1, \ldots, q_n, r$ are sentences. A set of sentences E is an extension of a default theory $\Delta = (W, D)$ if $\Gamma(E) = E$, where Γ is the operator defined as follows: for any set of sentences S, $\Gamma(S)$ is the smallest set of sentences satisfying the following conditins:

D1. $W \subseteq \Gamma(S)$

D2. $\Gamma(S)$ is closed under the tautological deduction.

D3. If $(p : q_1, \ldots, q_n/r) \in D$ and $p \in \Gamma(S)$, and $\neg q_1, \ldots, \neg q_n \notin S$ then $r \in \Gamma(S)$.

A default theory $\Delta = (W, D)$ is translated into the following set of formulas Δ_{GK} in GK:

1. If $p \in W$ then $Kp \in \Delta_{GK}$
2. If $(p : q_1, \ldots, q_n/r) \in D$ then

$$Kp \wedge \neg A \neg q_1 \wedge \ldots \wedge \neg A \neg q_n \supset Kr \in \Delta_{GK}$$

Theorem 2 *A consistent set of sentences E is a default extension of Δ iff there is a preferred model M of Δ_{GK} such that $E = K(M)$.*

4.2 Autoepistemic logic

Autoepistemic logic was proposed by Moore (1985) as a reformulation of McDermott and Doyle's nonmonotonic logic (1980). The language Moore used is a propositional one augmented by a modal operator L for belief. In the following, by a L-sentence we mean one that may contain the modal operator L but not K nor A.

Let S be a set of L-sentences. According to Moore, a set of L-sentences E is a *stable expansion (or AE extension)* of S if

$$E = Th(S \cup \{L\varphi \mid \varphi \in E\} \cup \{\neg L\varphi \mid \varphi \notin E\}) \tag{1}$$

where Th is the tautological closure operator. Konolige (1988) proves that for any set of L-sentences S, there is a set of L-sentences S' such that (a) a set E is a stable expansion of S iff it is a stable expansion of S'; (b) every member of S' is in *normal* form, that is, of the form:

$$Lp \wedge \neg Lq_1 \wedge \ldots \wedge \neg Lq_n \supset r$$

where $p, q_1, ..., q_n, r$ are base formulas, $n \geq 0$, and Lp may be absent. Thus without the lose of generality, in the following we assume that S is always a set of L-sentences of the normal form.[2]

For any such S, we define S_{GK} by the following equation:

$$S_{GK} = \{Ap \wedge \neg Aq_1 \wedge \ldots \wedge \neg Aq_n \supset Kr \mid Lp \wedge \neg Lq_1 \wedge \ldots \wedge \neg Lq_n \supset r \in S\}$$

Parallel to Theorem 2, we have the following result:

Theorem 3 *A consistent stable set of L-sentences E is a stable expansion of S iff there is a preferred model M of S_{GK} such that $K(M) = Base(E)$, where $Base(E)$ is the set of base formulas in E.*

By Theorem 1, the theorem is equivalent to the following propostion which is announced in [Lin and Shoham, 1989] and independently proved by Marek and Truszczynski (1989):

Proposition 4.1 *A stable set E is a stable expansion of S iff $Base(E) = E_2$ where*

1. $E_1 = \{p \mid p \in S$ *is a base formula*$\}$
2. $E_2 = Th(E_1 \cup \{r \mid Lp \wedge \neg Lq_1 \wedge \ldots \wedge \neg Lq_n \supset r \in S$, *where* $p \in Base(E)$ *and* $q_1, ..., q_n \notin Base(E)\})$.

[2]Theoretically, it is not necessary to use the normal form. Under the assumption that A satisfies S45 system, an arbitray L-sentence can be translated into a sentence in our language by inserting K in front of the L-sentence and replacing every L by A. The following Theorem 3 will also be true for this transformatioin.

Therefore by Theorem 2 and Theorem 3, under Konolige's transformation [Konolige 1988]:

$$p : q_1, \ldots, q_n/r \Rightarrow Lp \wedge \neg L\neg q_1 \wedge \ldots \wedge \neg L\neg q_n \supset r$$

the difference between default and autoepistemic logics lies in their different interpretations of the premise p. While default logic treats the premise p and the conclusion r in the same way (as knowledge in our terminology), autoepistemic logic treats the premise p and the consistency assumptions q's in the same way (as assumptions in our terminology). The effect of this difference is that the notion of default extensions is stronger than that of stable expansions, that is, under Konolige's transformation, if E is a default extension of a default theory and T is a stable set such that $Base(T) = E$, then T is also a stable expansions of the corresponding autoepistemic theory, but the converse is not true in general [Konolige 1988]. There is, however, a special case. If Δ is a default theory without premises, that is, if every default in Δ has the form:

$$: q_1, \ldots, q_n/r$$

then Konolige's transformation is exact, that is, a set of base sentence E is a default extension of Δ iff there is a stable set T such that $Base(T) = E$ and T is a stable expansion of the corresponding autoepistemic theory (this result is also proved in [Lin and Shoham 1989], and is now trivial according to our results).

5 CONCLUSIONS

We have defined the logic GK, of Grounded Knowledge, and provided two very similar transformations from default and autoepistemic logics into GK. The transformations provide for the first time a uniform epistemic semantics for both default and autoepistemic logics, and thus a common semantic background against which the two logics can be clearly compared.

As for our future work, the most important one is the extension of GK to the first-order case. As one might expect, there are several different ways of doing this. Ideally, we would like the extension to have the following properties:

1. It should provide a first-order extension of autoepistemic logic.

2. It should provide a truly first-order extension of default logic, i.e., we should have something like $\forall x(P(x) : Q(x)/R(x))$, instead of having only the open default $P(x) : Q(x)/R(x)$, which is considered a shorthand for a set of closed defaults.

3. It should be able to capture circumscription (since both GK and circumcription are minimal-model based). Particularly, like circumscription, it should have the ability to infer universal statements.

It turns out however that although it is easy to satisfy any single one of the above properties separately, it is quite difficult, if not impossible, to have a first-order GK to satisfy the three properties at the same time. We hope that we shall have a separate paper about first-order GK in the near future.

References

[1] Etherington, D.W. (1987), A semantics for default logic, *Proceedings of IJCAI-87*, Milan, Italy.

[2] Halpern, J. Y. and Y. O. Moses (1984), Towards a theory of knowledge and ignorance: preliminary report, IBM Technical Report RJ 4448 (48136), 1984.

[3] Konolige, K. (1988), On the relation between default logic and autoepistemic logic, *Artificial Intelligence*, 35 (1988) 343–382.

[4] Levesque, H. (1989), All I know: a study in autoepistemic logic, Technical Report KRR-TR-89-3, University of Toronto, 1989.

[5] Lifschitz, V. (1987a), Pointwise circumscription, in *Readings in Nonmonotonic Reasoning*, Morgan Kaufmann, 1987, ed. by M. Ginsberg.

[6] Lin, F. (1988), Circumscription in a modal logic, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, Asilomar, California, 1988.

[7] Lin, F. and Y. Shoham (1988), Argument systems: a uniform basis for nonmonotonic reasoning, in the *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, 1989.

[8] McCarthy, J. (1980), Circumscription — A form of non-monotonic reasoning, *Artificial Intelligence* 13(1980) 27–39.

[9] McCarthy, J. (1986), Applications of circumscription to formalizing commonsense knowledge, *Artificial Intelligence* 28 (1986), 89–118.

[10] McDermott, D. and J. Doyle (1980), Nonmonotonic logic I, *Artificial Intelligence* 13(1980), 41–72.

[11] Marek, W. and M. Truszczynski (1989), Relating autoepistemic and default logics, in the *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, Toronto, 1989.

[12] Moore, R. (1983), Semantical considerations on nonmonotonic logic, *Conference Proceedings of IJCAI-83*.

[13] Reiter, R. (1980), A logic for default reasoning, *Artificial Intelligence* 13(1980) 81–132.

[14] Shoham, Y. (1987), Nonmonotonic logics: meaning and utility, in *Proceedings of IJCAI-1987*, Milan, Italy.