Hierarchic Autoepistemic Theories for Nonmonotonic Reasoning*

Kurt Konolige
Artificial Intelligence Center
Center for the Study of Language and Information
SRI International
333 Ravenswood, Menlo Park, Ca. 94025

Abstract

Nonmonotonic logics are meant to be a formalization of nonmonotonic reasoning. However, for the most part they fail to capture in a perspicuous fashion two of the most important aspects of such reasoning: the explicit computational nature of nonmonotonic inference, and the assignment of preferences among competing inferences. We propose a method of nonmonotonic reasoning in which the notion of inference from specific bodies of evidence plays a fundamental role. The formalization is based on autoepistemic logic, but introduces additional structure, a hierarchy of evidential subtheories. The method offers a natural formalization of many different applications of nonmonotonic reasoning, including reasoning about action, speech acts, belief revision, and various situations involving competing defaults.

1 Introduction

The nonmonotonic character of commonsense reasoning in various domains of concern to AI is well established. Recent evidence, especially the work connected with the Yale Shooting Problem (see [Hanks and McDermott, 1987]) has illuminated the often profound mismatch between nonmonotonic reasoning in the abstract, and the logical systems proposed to formalize it. This is not to say that we should abandon the use of formal nonmonotonic systems; rather, it argues that we should seek ways to make them model our intuitive conception of nonmonotonic reasoning more closely. Generally speaking, current formal nonmonotonic systems suffer from two shortcomings:

1. They have no computationally realizable implementation.
2. They have only limited means for adjudicating among competing nonmonotonic inferences.

To briefly review just the current major formalisms in this regard: Circumscription [McCarthy, 1980] and related model-preference systems [Shoham, 1987], default logic [Reiter, 1980], and autoepistemic (AE) logics [Moore, 1985; Levesque, 1982] are computationally intractable; and various proposals based on the notion of *defeasible rules* (see, for example, [Poole, 1985]) have yet to be given an implementation. The standard means of arriving at an implementation is to restrict the language, but of course this restricts the expressivity of the resulting system, often to a rather severe extent.

The importance of having a flexible means for deciding among competing nonmonotonic inferences has become clear in the recent debate over the Yale Shooting Problem. It also arises in other contexts, such as taxonomic hierarchies [Etherington and Reiter, 1983] or speech act theory [Appelt and Konolige, 1988]. Prioritized circumscription [Lifschitz, 1984] gives circumscription the capability of assigning priorities to various default assumptions. To some extent, preferences among default inferences can be encoded in AE and default logics by introducing auxiliary information into the statements of defaults; but this method does not always give a satisfactory correspondence with our intuitions. The most natural statement of preferences is with respect to the multiple *extensions* of a particular theory, that is, we prefer certain extensions because the default rules used in them have a higher priority over ones used in alternative extensions.

Hierarchic autoepistemic logic (HAEL) is a modification of autoepistemic logic [Moore, 1985] that addresses these two considerations. In HAEL, the primary structure is not a single uniform theory, but rather, a collection of subtheories linked in a hierarchy. Subtheories represent different sources of information available to an agent, and the hierarchy expresses the way in which this information is combined. For example, in representing taxonomic defaults, more specific information would take precedence over more general attributes. HAEL thus permits a natural expression of preferences among defaults, and, in general, a more natural translation of our informal conception of nonmonotonic reasoning into a formal system. Further, given the hierarchic nature of the subtheory relation, there is a well-founded constructive semantics for the autoepistemic operator, in contrast to the usual self-referential fixedpoints. We can then easily arrive at computational realizations that make use of resource-bounded inference methods.

HAEL has been implemented and integrated with KADS, a resolution theorem-proving system for commonsense reasoning [Geissler and Konolige, 1986]. We have developed axiomatizations for reasoning about action, a preliminary form of belief revision, and speech-act theory. Currently, the resolution system and speech-act axiomatization are being employed in a natural-language generation system [Appelt and Konolige, 1988].

The rest of this paper is divided into two sections. In the first, we present an informal overview of HAEL, its re-

*This research was supported by the Office of Naval Research under Contract No. N00014-85-C-0251, by subcontract from Stanford University under the Defense Advanced Research Projects Administration under Contract No. N00039-84-C-0211, and by a gift from the System Development Foundation.

lation to AE logic, and its applicability for nonmonotonic reasoning. The second section contains the formal characterization of HAEL, including its semantics and characterization in terms of *stable sets*. Because of the shortness of this abstract, we do not include proofs of propositions, and cannot present an extended example of the application of the logic.

2 Hierarchic autoepistemic theories

Hierarchic AE logic is derived from AE logic by splitting a uniform belief set into components, called *subtheories*. In AE logic, an agent is assumed to have an initial set of premise sentences A. The language of A contains an operator L for talking about self-belief: $L\phi$ is intended to mean that ϕ is one of the agent's beliefs. A belief set T that is derivable from the premises A by an ideal agent will be *stable* with respect to self-belief, that is, a sentence ϕ is in T if and only if $L\phi$ is. This interpretation of L is clearly self-referential, since it refers to the theory in which L itself is embedded.

In the hierarchic modification of AE logic, the dependence of L on T is broken by dividing T into a hierarchy of subtheories, and indexing L so that it refers to subtheories beneath it in the hierarchy. For example, we might divide T into two subtheories T_0 and T_1, with T_1 succeeding T_0 in the hierarchy. Subtheory T_1 may contain atoms of the form $L_0\phi$, which refer to the presence of ϕ in the subtheory T_0. The interpretation of L is constructive as long as the hierarchy is well-founded (no infinite descending chains) and every subtheory contains only modal operators referring to lower subtheories.

HAEL is still an *auto*epistemic logic, because the subtheories together comprise the agent's belief set. In fact, HAEL could be considered a more natural formalization of autoepistemic reasoning than AE logic, because of its hierarchic structure. In AE logic, we found it necessary to characterize extensions in terms of the groundedness of inferences used in their construction (see [Konolige, 1987]), in order to exclude those containing circular reasoning. No such device is necessary for HAEL; circularity in the derivation of beliefs is impossible by the very nature of the logic.

Breaking the circularity of AE logic has other advantages. Given a fairly natural class of closure conditions, every HAEL structure has exactly one "extension," or associated theory. So HAEL, although a nonmonotonic logic, preserves many of the desirable properties of first-order logic, including a well-defined notion of *proof* and *theorem*, and a well-founded, compositional semantics. Computationally, HAEL is still not even semi-decidable in the general case; but unlike AE logic, it lends itself readily to proof-theoretic approximation.

The subtheories of HAEL are meant to serve as bodies of evidence, as discussed in the previous section. Those subtheories lower in the hierarchy are considered to be stronger evidence, and conclusions derived in them take precedence over defaults in subtheories higher in the hierarchy. Priorities among defaults and bodies of evidence are readily expressed in the structure of the hierarchy. Many different domains for nonmonotonic reasoning can be fruitfully conceptualized in this fashion. The most natural case is taxonomic hierarchies with exceptions, since the structure of the subtheories mimics the taxonomy (we give a very simple taxonomic example in the next section). Speech act theory is a very complex and interesting application domain, since the sources of information (agents' mental states, the content and linguistic force of the utterance) interact in complicated ways to induce belief revision after the utterance. In this case, we model the structure of the belief revision process with subtheories that reflect the relative force of the new information on old beliefs (see [Appelt and Konolige, 1988]).[1]

3 HAEL structures and their semantics

We now present the formal definition of HAEL structures, and two independent semantics for these structures. The first is based on the notion of a stable set, an idea introduced by Stalnaker [Stalnaker, 1980] and used extensively in the development of AE logic [Moore, 1985; Konolige, 1987]. Stable sets are defined using closure conditions that reflect the end result of introspection of an ideal agent on his own beliefs. The second semantics is a classical approach: first-order valuations modified to account for the intended interpretation of the L_i-operators. This semantics is taken directly from AE logic, and shows many of the same properties. However, there are some significant differences, due to the hierarchical nature of HAEL structures. In AE logic, a belief set that follows from a given assumption set A via the semantics is called an *extension* of A. There may be none, one, or many mutually conflicting extensions of A. HAEL structures always have exactly one extension, and thus a well-defined notion of theorem.

There is also a mismatch in AE logic between stable set semantics and autoepistemic valuations. A stable set for A which is minimal (in an appropriate sense) is a good candidate for a belief set; yet there exist such minimal stable sets that are not extensions of A. In HAEL, we show that the two semantics coincide: the unique minimal stable set of an HAEL structure is the extension of that structure given by its autoepistemic valuations.

3.1 HAEL structures

In AE logic, one starts with a set of premise sentences A, representing the initial beliefs or knowledge base of an agent. The corresponding object in HAEL is an *HAEL structure*. A structure τ consists of an indexed set of subtheories τ_i, together with a well-founded, irreflexive partial order on the set. We write $\tau_i \prec \tau_j$ if τ_i precedes τ_j in the order. The partial order of subtheories reflects the relative strength of the conclusions reached in them, with preceding subtheories having being stronger. The condition of well-foundedness means that there is no infinite descending chain in the partial order; the hierarchy always "bottoms out."

Each subtheory τ_i contains an initial premise set A_i, and also an associated first-order deduction procedure I_i. The

[1] It should be noted that this is the first formalization of speech act theory in a nonmonotonic system that attempts to deal with a nontrivial belief revision process.

deduction procedures are sound (with respect to first-order logic) but need not be complete. The idea behind parameterizing HAEL structures by inference procedures in the subtheories is that ideal reasoning can be represented by using complete procedures, while resource-bounded approximations can be represented by incomplete but efficient procedures. In the rest of this paper, we shall assume complete first-order deduction in each subtheory; HAEL structures of this form are called *complete.*

The language $\mathcal{L}$ of HAEL consists of a standard first-order language, augmented by a indexed set of unary modal operators L_i. If ϕ is any sentence (no free variables) of the first-order language, then $L_i\phi$ is a sentence of $\mathcal{L}$. Note that neither nesting of modal operators nor quantifying into a modal context is allowed. Sentences without modal operators are called *ordinary.* The intended meaning of $L_i\phi$ is that the sentence ϕ is an element of subtheory τ_i.

Within each subtheory, inferences are made from the assumption set, together with information derived from subtheories lower in the hierarchy. Because subtheories are meant to be downward-looking, the language $\mathcal{L}_i \subseteq \mathcal{L}$ of a subtheory τ_i need contain only modal operators referring to subtheories lower in the hierarchy. We formalize this restriction with the following statement:[2]

(1) *The operator L_j occurs in $\mathcal{L}_i$ if and only if $\tau_j \prec \tau_i$.*

Here is a simple example of an HAEL structure, which can be interpreted in terms of the taxonomic example presented in the preceding section, by letting the intended meaning of $F(x)$ be "x flies," $B(x)$ be "x is a bat," and $M(x)$ be "x is a mammal."

$$
\begin{aligned}
&\tau_0 \prec \tau_1 \prec \tau_2\\
&A_0 = \{B(a)\}\\
(2)\quad &A_1 = \{\forall x.Bx \supset Mx,\\
&\qquad\qquad L_0B(a) \wedge \neg L_0\neg F(a) \supset F(a)\}\\
&A_2 = \{L_1M(a) \wedge \neg L_1F(a) \supset \neg F(a)\}
\end{aligned}
$$

There are three subtheories, with a strict order (heritable) between them. Subtheory τ_0 is the lowest, and contains the most specific information (based on the taxonomy). In the assumption set A_1, there is a default rule about bats flying: if it is known in τ_1 that a is a bat, and unknown in τ_0 that a does not fly, then it will be inferred that a flies. The assumption set A_2 is similar to A_1; it also permits the deduction that bats are mammals. The information that a is a bat and a mammal will be passed up to τ_2, along with any inferences about its ability to fly.

The partial order of an HAEL structure is well-founded, and so it is possible to perform inductive proofs using it. At times we will need to refer to unions of sets derived from the subtheories preceding some subtheory τ_n; to do this, we use $\bigcup_{j \prec n} X_j$, where j ranges over all indices for which $\tau_j \prec \tau_n$.

[2]We can relax this restriction so that L_i can occur in $\mathcal{L}_i$ under certain circumstances. The semantics of HAEL structures is simpler to present without this complication, however, so we will not deal with it here.

3.2 Complex stable sets

Stalnaker considered a belief set Γ that satisfied the following three conditions:

1. Γ is closed under first-order consequence.[3]
2. If $\phi \in \Gamma$, then $L\phi \in \Gamma$.
3. If $\phi \notin \Gamma$, then $\neg L\phi \in \Gamma$.

He called such a set *stable*, because an agent holding such a belief set could not justifiably deduce any further consequences of his beliefs. In HAEL, these conditions must be modified to reflect the nature of the L_i-operators, as well as the inheritance of sentences among subtheories.

DEFINITION 3.1 *A* complex stable set *for a structure τ is a sequence of sets of sentences $\Gamma_0, \Gamma_1, \ldots$, corresponding to the subtheories of τ, that satisfies the following five conditions:*

1. *Every Γ_i contains the assumption set A_i.*
2. *Every Γ_i is closed under the inference rules of τ_i. In the case of an ideal agent, the closure is first-order logical consequence.*
3. *If ϕ is an ordinary sentence of Γ_j, and $\tau_j \prec \tau_i$, then ϕ is in Γ_i.*
4. *If $\phi \in \Gamma_j$, and $\tau_j \prec \tau_i$, then $L_j\phi \in \Gamma_i$.*
5. *If $\phi \notin \Gamma_j$, and $\tau_j \prec \tau_i$, then $\neg L_j\phi \in \Gamma_i$.*

To illustrate complex stable sets, consider the previous example of flying bats. Let $\mathrm{Cn}_i(X)$ stand for the first-order closure of X using language $\mathcal{L}_i$, and define the set $S = S_0, S_1, \cdots$ by

$$
\begin{aligned}
&S_0 = \mathrm{Cn}_0(B(a))\\
&S_1 = \mathrm{Cn}_1(B(a), L_0B(a), \neg L_0\neg B(a), \neg L_0\neg F(a),\\
(3)\quad &\qquad\qquad \forall x.Bx \supset Mx,\ M(a),\ F(a), \cdots)\\
&S_2 = \mathrm{Cn}_2(B(a), M(a), F(a), L_0B(a), L_1B(a),\\
&\qquad\qquad L_1F(a), \cdots)
\end{aligned}
$$

The set S is a complex stable set for the HAEL structure τ as defined in Equations (2). The lowest set S_0 contains just the first-order consequences of $B(a)$. S_1 inherits this sentence, and has the additional information $M(a)$ from its assumption set. Modal atoms of the form $L_0\phi$ and $\neg L_0\phi$ are also present, reflecting the presence or absence of sentences in S_0; the sentence $F(a)$ is derived from these and the assumption set. Finally, S_2 inherits all ordinary sentences from S_1, as well as $L_1F(a)$.

The subsets S_i of S are *minimal* in the sense that we included no more than we were forced to by the conditions on complex stable sets. For example, another stable set S' might have $S_0' = \mathrm{Cn}_0(B(a), \neg F(a))$, with the other subtheories defined accordingly. The sentence $\neg F(a)$ in S_0' is not justified by the original assumption set A_0, but there is nothing in the definition of complex stable sets that forbids it from being there. So, a complex stable set is a candidate for the extension of an HAEL structure only if it is minimal. But what is the appropriate notion of minimality here? For simple stable sets, minimality can be defined in terms of set inclusion of the ordinary sentences of the stable sets. Complex stable sets have multiple subtheories, and the definition of minimality must take into account the relative strength of information in these subtheories.

[3]Stalnaker considered propositional languages and so used tautological consequence.

DEFINITION 3.2 *A stable set S for the HAEL structure τ is* minimal *if for each subset S_i of S, there is no stable set S' for τ that agrees with S on all $S_j \prec S_i$, and for which $S'_i \subset S_i$.*

A complex stable set for τ is minimal if each of its subsets is minimal, given that the preceding subsets (those of higher priority) are considered fixed.

There is exactly one minimal complex stable set for an HAEL structure. We now prove this fact, and give an inductive definition of the set.

PROPOSITION 3.1 *Every HAEL structure τ has a unique minimal complex stable set, which can be determined by the following inductive procedure.*

Define:

$$\begin{array}{lcl} \mathrm{Cn}_i(X) & = & \text{the first-order closure of } X \text{ in } \mathcal{L}_i \\ \mathrm{Ord}(X) & = & \text{the ordinary sentences of } X \\ \mathrm{L}_i(X) & = & \{L_i\phi \mid \phi \in X \text{ and } \phi \text{ ordinary}\} \\ \mathrm{M}_i(X) & = & \{\neg L_i\phi \mid \phi \notin X \text{ and } \phi \text{ ordinary}\} \end{array}$$

For minimal τ_i (that is, there is no τ_j such that $\tau_j \prec \tau_i$), let

$$S_i = \mathrm{Cn}_i(A_i)$$

For nonminimal τ_n, define

$$S_n = \mathrm{Cn}_n(A_n \cup \bigcup_{j \prec n} \mathrm{Ord}(S_j) \cup \mathrm{L}_j(S_j) \cup \mathrm{M}_j(S_j)) \, .$$

S is the unique minimal complex stable set for τ.

The existence of a unique minimal complex stable set for every HAEL structure gives us a means of defining the theorems of a structure. Let S be the complex stable set for τ. We say that a sentence ϕ is *derivable in the subtheory* τ_i if and only if it is an element of S_i, and write $\tau \vdash_i \phi$ if this holds. For the bat example, the following derivations exist (where τ is the HAEL structure (2)):

$$(4)\quad \begin{array}{l} \tau \vdash_0 B(a) \\ \tau \not\vdash_0 \neg F(a) \\ \tau \vdash_1 B(a) \wedge M(a) \wedge \neg L_0 \neg F(a) \wedge F(a) \\ \tau \vdash_2 B(a) \wedge M(a) \wedge L_1 F(a) \wedge F(a) \end{array}$$

3.3 HAEL semantics

We have used complex stable sets to give a proof-theoretic notion of theorem to HAEL structures. An alternative approach is to develop a semantics for these structures, and define a notion of validity with respect to the semantics. As with autoepistemic logic, the semantic picture is complicated by the presence of self-referential elements, and validity must be determined by use of a fixedpoint equation. Happily, for HAEL structures validity turns out to be equivalent to derivability, so that the sentences which are valid logical consequences of a structure are exactly those given by its minimal complex stable set.

We start with the notion of a *valuation* of an HAEL structure τ. In classical logic, a valuation assigns true or false to each sentence of the language, and a valuation is said to *satisfy* a theory if all the sentences of the theory are assigned true. If the valuation v assigns true to the sentence ϕ, we write $v \models \phi$. Restrictions on valuations single out the intended semantics of the theory, e.g., first-order valuations must respect the intended meaning of the quantifiers and boolean operators.

In autoepistemic logic, the interpretation of the modal operator L adds an additional complication to valuations. Since the intended interpretation of $L\phi$ is that ϕ be in the belief set of the agent, an AE valuation consists of a first-order valuation v and a set of sentences (the belief set) Γ (see [Moore, 1985]). We call Γ the *modal index* of the valuation. The interpretation rules for AE valuations are as follows (we let ϕ stand for an arbitrary *ordinary* sentence).

$$(5)\quad \begin{array}{lll} \langle v, \Gamma \rangle \models \phi & \text{iff} & v \models \phi \\ \langle v, \Gamma \rangle \models L\phi & \text{iff} & \phi \in \Gamma \end{array}$$

The interpretation of the L-operator is completely decoupled from the first-order valuation.

The *autoepistemic extension* of an assumption set A is a set of sentences T that are the logical consequences of A under AE valuations. Because the intended interpretation of L is *self*-belief, only those AE valuations that respect this interpretation can be used. Let $A \models_\Gamma \phi$ mean that every AE valuation with modal index Γ that satisfies the set A also satisfies ϕ. An extension T of A is defined by the following equation (see [Konolige, 1987]):

$$(6)\quad T = \{\phi \mid A \models_T \phi\}$$

By fixing the modal index as T, we are assured that the interpretation of L is with respect to the belief set T itself. Of course, the equation defining extensions is self-referential, and as we have pointed out, self-reference creates problems from a computational point of view.

The semantics of HAEL structures is similar to AE assumption sets, but is complicated by the presence of multiple subtheories. The interpretation of the indexed operators L_i must be with respect to a sequence of belief subsets, instead of a single belief set Γ. So an HAEL valuation $\langle v, \Gamma_1, \cdots, \Gamma_n, \cdots \rangle$ consists of a first-order valuation v, together with the indexed belief subsets Γ_i, which we call a *complex belief set*. The interpretation rules for HAEL valuations are similar to that for AE valuations (again, ϕ stands for an arbitrary ordinary sentence).

$$(7)\quad \begin{array}{lll} \langle v, \Gamma_1, \cdots, \Gamma_n, \cdots \rangle \models \phi & \text{iff} & v \models \phi \\ \langle v, \Gamma_1, \cdots, \Gamma_n, \cdots \rangle \models L_i\phi & \text{iff} & \phi \in \Gamma_i \end{array}$$

The interpretation of each L_i is with respect to the appropriate belief subset. Note that there is no necessary relation in valuations among the interpretations of the modal operators, or between the modal operators and the first-order valuation.

An autoepistemic extension of an HAEL structure τ consists of a sequence of a complex belief set, $T = T_1, \cdots, T_n, \cdots$, corresponding to the subtheories of the structure. Again, we require that extensions be defined using only those valuations that respect the nature of the L_i-operators as *self*-belief. Also, because each subtheory inherits the ordinary sentences of preceding subsets, the assumption set must be augmented appropriately.

DEFINITION 3.3 *The complex belief set T is an* extension *of τ if it satisfies the equations*

$$T_i = \{\phi \in \mathcal{L}_i \mid A_i \cup \bigcup_{j \prec i} \mathrm{Ord}(T_j) \models_T \phi\} \, .$$

As with AE logic, the definition of extensions for HAEL appears to be self-referential, since T_i appears on both sides of the equation. However, this self-reference is illusory from the point of view of the individual subtheories, because they contain L_i-operators referring only to subtheories lower in the hierarchy. In fact, every HAEL structure has a unique extension, and that extension is the minimal complex stable set.

PROPOSITION 3.2 *Every HAEL structure τ has a unique extension T, which is the complex stable set for τ.*

Having a single extension is a nice feature of HAEL structures, because there is a single notion of *theorem*, and the problem of choosing among competing multiple extensions (as in AE logic) does not exist. However, there is a price to pay. In AE logic, multiple extensions arise because there are conflicting defaults: the classic Nixon diamond is a well-known example, where the default that Republicans are not pacifists conflicts with the default that Quakers are. In HAEL, if both these defaults are placed in the same subtheory, an inconsistency will occur (there will still be a single extension, but the subtheory will consist of all sentences because of closure under logical consequence). Thus the HAEL structure must be constructed so that conflicts of this sort within the same subtheory are avoided.

3.4 Proof theory

Proposition 3.1 is important in that it makes the notion of "theorem" well-founded for HAEL structures. It also is the basis for proof methods on HAEL structures. Consider the previous example of the bat taxonomy (Equation 2). We want to know whether a flies, that is, whether $F(a)$ or $\neg F(a)$ is provable in T_2. Suppose we set $\neg F(a)$ as a goal in T_2. There is only one axiom which applies, and this gives the subgoal $L_1 M(a) \wedge \neg L_1 F(a)$. To establish the first conjunct, we set up $M(a)$ as a goal in T_1. Using the universal implication, we arrive at the subgoal $B(a)$, which matches with $B(a)$ in T_0. Hence we have shown that $L_1 M(a)$ holds in T_2.

In a similar manner, we set up $F(a)$ as a subgoal in T_1. Using the second axiom of A_1, we have the conjunctive goal $L_0 B(a) \wedge \neg L_0 \neg F(a)$. The first subgoal is easily proven, since $B(a)$ is in T_0. Now we try to prove $\neg F(a)$ in T_0. This is not possible, so $\neg L_0 \neg F(a)$ is proven in T_1. We have just shown $F(a)$ to be provable in T_1, so $\neg L_1 F(a)$ is not provable in T_2. Our attempt to prove $\neg F(a)$ in T_2 fails. On the other hand, along the way we have shown $F(a)$ to be provable in T_1; hence by inheritance it is also in T_2.

In this example, we used backward-chaining exclusively as a proof method. Other methods are also possible, e.g., intermixtures of forward and backward chaining, resolution, etc. Whenever there is a question as to the provability of a modal atom, an appropriate subgoal is set up in a preceding subtheory, and the proof process continues.

It should be noted that no proof process can be complete when the nonmodal language is undecidable, because the inference of $\neg L_i \phi$ requires that we establish ϕ to be not provable in T_i. However, a proof method can readily approximate the construction of Proposition 3.2, by assuming that ϕ cannot be proven after expending a finite amount of effort in attempting to prove it. Given enough resources, a proof procedure of this sort will converge on the right answer.

We have implemented HAEL on a resolution theorem-proving system, modified to accept a belief logic of the sort described in [Geissler and Konolige, 1986]. The implementation was straightforward, and involved adding a simple negation-as-failure component to the prover. The implementation has been successfully applied to reason about speech acts in a natural-language understanding project [Appelt and Konolige, 1988].

References

[Appelt and Konolige, 1988] Douglas E. Appelt and Kurt Konolige. A nonmonotonic logic for reasoning about speech acts and belief revision. *submitted to the Workshop on Non-Monotonic Reasoning*, 1988.

[Etherington and Reiter, 1983] D. W. Etherington and R. Reiter. On inheritance hierarchies with exceptions. In *Proceedings of the American Association of Artificial Intelligence*, 1983.

[Geissler and Konolige, 1986] Christophe Geissler and Kurt Konolige. A resolution method for quantified modal logics of knowledge and belief. In Joseph Y. Halpern, editor, *Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 309–324, 1986.

[Hanks and McDermott, 1987] S. Hanks and D. McDermott. Nonmonotonic logic and temporal projection. *Artificial Intelligence*, 33(3), November 1987.

[Konolige, 1987] Kurt Konolige. *On the Relation between Default Logic and Autoepistemic Theories.* Technical Note 407, SRI Artificial Intelligence Center, Menlo Park, California, 1987.

[Levesque, 1982] Hector J. Levesque. *A Formal Treatment of Incomplete Knowledge Bases.* Technical Report 614, Fairchild Artificial Intelligence Laboratory, Palo Alto, California, 1982.

[Lifschitz, 1984] Vladimir Lifschitz. Some results on circumscription. In *AAAI Workshop on Non-Monotonic Reasoning*, 1984.

[McCarthy, 1980] John McCarthy. Circumscription — a form of nonmonotonic reasoning. *Artificial Intelligence*, 13(1–2), 1980.

[Moore, 1985] Robert C. Moore. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25(1), 1985.

[Poole, 1985] D. Poole. On the comparison of theories: preferring the most specific explanation. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 144–147, Los Angeles, 1985.

[Reiter, 1980] Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13(1–2), 1980.

[Shoham, 1987] Yoav Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence.* MIT Press, Cambridge, Massachusetss, 1987.

[Stalnaker, 1980] R. C. Stalnaker. A note on nonmonotonic modal logic. 1980. Department of Philosophy, Cornell University.