DEFAULT REASONING USING MONOTONIC LOGIC:
A Modest Proposal

Jane Terry Nutter
Department of Computer Science, Tulane University

ABSTRACT

This paper presents a simple extension of first order predicate logic to include a default operator. Rules of inference governing the operation of a logic operator are developed based on standard Tarskian semantics. The resulting system is trivially sound. It is argued that (a) this logic provides an adequate basis for default reasoning in A.I. systems, and (b) unlike most logician's proposals for this purpose, it retains the virtues of standard first order logic, including both monotonicity and simplicity.

I INTRODUCTORY COMMENTS

Reasoning from incomplete information and from default generalizations follows patterns which standard first-order predicate logic does not reflect. The most striking deviation involves making inferences whose conclusions may be counterindicated by further information which does not explicitly contradict anything previously known.

In standard logics, if a set of premises entails a conclusion, any larger set containing all those premises also entails that conclusion. Logics with this property are called monotonic. These departure from standard logic's reasoning patterns has led researchers to adopt and develop non-monotonic logics for use in A.I. systems (e.g. McDermott and Doyle, 1980; McDermott, 1982; Reiter, 1980; Aronson et al., 1980; Duda et al., 1978; and others).

Critics of non-monotonic logics have pointed out technical weaknesses (see e.g. Davis, 1973, and Israel, 1980) and persuasively that its supporters confuse logic with complex judgments of kinds logic cannot deal with. But unless some alternative appears for dealing with default reasoning, non-monotonic logic must continue to appeal.

I have argued elsewhere that (1) default reasoning only appears non-monotonic if we fail to distinguish warranted assertions from warranted assumptions (Nutter, 1982), and (2) any logic which adequately models default reasoning must be monotonic (Nutter, 1983).

If this is accepted, a conservative approach promises more than do radical ones, not because it is "safer", but because it simultaneously preserves what is of interest -- its simplicity, clarity, and adequacy in domains of complete information -- and allows distinguishing warranted assumptions from warranted assertions (generalizations from universals, default consequences from generalizations, etc.). This paper presents an almost trivial extension of first order predicate logic which distinguishes generalizations and the resulting presumptions from universal and other consequences and permits reasoning from default generalizations.

II DEFAULT AS PROPOSITIONAL OPERATORS

The logic we are aiming for must differ from standard logic in three ways. First, it must distinguish "guarded" from "unguarded" propositions (presumptions) from the outset. That is, we should make an explicit distinction between (1) default operators which may be included for default reasoning in A.I. systems, and (2) default operators which are currently not part of standard logic. The resulting system is trivially sound. It is argued that this system provides an adequate basis for default reasoning in A.I. systems, and that it retains the virtues of standard first order logic, including both monotonicity and simplicity.

If this is accepted, a conservative approach promises more than do radical ones, not because it is "safer", but because it simultaneously preserves what is of interest -- its simplicity, clarity, and adequacy in domains of complete information -- and allows distinguishing warranted assumptions from warranted assertions (generalizations from universals, default consequences from generalizations, etc.). This paper presents an almost trivial extension of first order predicate logic which distinguishes generalizations and the resulting presumptions from universal and other consequences and permits reasoning from default generalizations.

III GRAMMATICAL CONSTRUCTION

The logic we are aiming for must differ from standard logic in three ways. First, it must distinguish "guarded" from "unguarded" propositions (presumptions) from the outset. That is, we should make an explicit distinction between (1) default operators which may be included for default reasoning in A.I. systems, and (2) default operators which are currently not part of standard logic. The resulting system is trivially sound. It is argued that this system provides an adequate basis for default reasoning in A.I. systems, and that it retains the virtues of standard first order logic, including both monotonicity and simplicity.
at present occur unbound. In this section, we present the rules of grammar for the language of our logic.

The technical discussions in the sections below will presuppose familiarity with the usual ways that formal logics are described, and will presuppose a complete description of a first order predicate logic. The informal descriptions should convey an overview for readers who lack this familiarity and motivate some of the less obvious decisions embodied in the technical discussions.

A. Informal Description

The grammar holds relatively few surprises.

In simple terms, wherever in English one might reasonably prefix a clause with "There is reason to believe that," and quantify and ostrimize some such logic will allow its equivalent of the clause to be governed by our default operator $p$ (for "presumably").

The only real decision involves whether to let the operator govern formulas or only propositions. While its use governing formulas is probably eliminable, it seems to me both innocuous and well motivated. While "Birds generally fly," which becomes $\forall x (\text{Bird}(x) \rightarrow p \text{ Flies}(x))$ in the formal logic, probably comes to the same thing as "In general, birds fly," $p \forall x (\text{Bird}(x) \rightarrow \text{Flies}(x))$ in the long run, both are said. Furthermore, "If anything is a bird, then it has wings and presumably flies" must be considerably reformulated before it can be equivalently expressed without having $p$ bind a formula instead of a sentence. Hence, $p$ directly govern either a formula or a sentence.

B. Technical Presentation

Let $L$ be a standard first order predicate calculus without equality and without function symbols, with a Jaskowski-style natural deduction system (Jaskowski, 1934) containing explicit introduction and elimination rules for the standard connectives and quantifiers and a small set of "bookkeeping" rules. (The precise formal system is unimportant.) The language of $L$ is $L$, the logic with defaults which we will build from $L$ is $LD$, and its language is $LD$. The default operator is written $p$.

Throughout the discussion below, Roman capitals vary over formulas and propositions of $L$, while Greek letters $\alpha$ and $\beta$ vary over formulas and propositions of $LD$, and $\alpha$ varies over variables of $L$.

Def. The set of formulas of $LD$ is the smallest set satisfying the following clauses:

(a) All atomic formulas of $L$ are formulas of $LD$;

(b) For any formula $\phi$ of $LD$, $p\phi$ is a formula of $LD$;

(c) For any formulas $\phi, \psi$ or $\omega$, the following are all formulas of $LD$: $\neg \phi$; $\phi \lor \psi$; $\phi \land \psi$; $\phi \Rightarrow \psi$; $\psi \Rightarrow \phi$.

(d) For any formula $\phi$ of $LD$ in which $\alpha$ occurs open, the following are also formulas of $LD$: $\forall \alpha \phi$; $\exists \alpha \phi$.

For any formula $\phi$ of $LD$, $\alpha$ is a proposition of $LD$ if and only if $\alpha$ contains no open occurrences of variables.

IV DEDUCTIVE SYSTEM

As commented above, the deductive system is a standard natural deduction system, with the usual five connectives ("not", written "\neg", "and", written "\land", "or", written "\lor", "if...then...", written "\Rightarrow", and "if and only if", written "\iff") and both quantifiers ("for all", written "\forall", and "there is", written "\exists"). This will be the least altered portion of the logic.

A. Informal Description

The guarded status of default generalizations must be inherited through inferences, and from a guarded proposition, it must be possible to infer a guarded version of any conclusion which could be inferred from the unguarded version of the premise. That is, if you could infer from "Roger flies" that "If my dog sees Roger, he will find him fascinating," then from "Presumably Roger flies" you may infer "Presumably, if my dog sees Roger, he will find him fascinating," and so on. However, it generally takes more than one premise to warrant an inference. How many of the premises may be generalizations? There are several ways to go on this.

We could allow only one guarded premise per inference. But suppose we have "In general, if $A$ than $B$," and we have "There is reason to believe that $A$." We would normally feel free to infer: "There is reason to believe that $B$." But this inference has two guarded premises, and without further information, there is no way to avoid using both guarded premises in a single inference and still get the conclusion.

Alternatively, we could "string out" the guards, putting as many $p$ on the front of the conclusion as there were guarded premises contributing to it. This might seem to give not only a guard, but also a sort of certainty measure. Unfortunately, it's a very poor measure for anything interesting. Here's why.

Default generalizations are not statistical. They don't necessarily "accumulate". There are several reasons for this. First, they embody judgments of typicality. A priori, Roger flies" generalizations will frequently stand or fall together: if the instance is typical, they all hold, while if the instance is atypical, each is likely to fail. Certainly this phenomenon cannot be relied upon, but it is common enough to upset the measure. (For more on this and related points, see Nutter, 1982 and Hutter, 1992.)

Second, a single guarded proposition might be used five or six times in the course of a long derivation. Is the result any less certain for having that proposition enter more than once? By the time a "chain of inferences" has taken place, the number of $p$'s on the front may no longer represent even one of the distinct generalizations involved in reaching it.

Third, there is no reason to suppose that all the original generalizations are equally reliable. Five extremely strong generalizations probably provide better support than one. If we could assign weights reflecting reliability -- i.e. probabilities -- we would be dealing with statistical generalizations and not with defaults anywhay.

Furthermore, from "Generally, if $A$ than $B$" and "There is reason to believe that $A$," we don't conclude "There is reason to believe that $B$." Setting up a rule for inferring $B$ from $A$ and $pA$ aiming to argue from generalizations to a role in the kind of reasoning we are trying to model, it seems reasonable to suppose that language would reflect it. I have therefore decided to "inherit" the guarded proposition any or all premises involved it, without regard to how many. This is the force of the first rule of inference below.

Even with the rule in that form, it would be possible to generate long strings of $p$. The second rule says that if there are more than one consecutive qualifications, they may be collapsed into a single one.
There is a restriction on our interpretations which is not obvious from an intuitive standpoint. It arises from the fact that for any \( \psi \models \varphi \) Suppose we have \( \varphi \equiv \psi \), and also \( \varphi \models \). Then we can infer \( \psi \models \). For this to be sound, it must be the case that whenever \( \varphi \models \) and \( \psi \equiv \varphi \) are both true, so is \( \psi \).

But consider an interpretation on which \( \psi \) is false. Then \( \varphi \equiv \psi \) is true regardless of the truth value of \( \psi \) (or any other \( \psi \)'s). Hence it will be the case that whenever \( \varphi \models \) and \( \psi \equiv \varphi \) are both true, but \( \psi \) is already at least true. But \( \psi \) must also be at least true for false \( \psi \).

Hence for any interpretation, if even one false proposition has the property that there is reason to believe it on that interpretation, then all false propositions have that property on that interpretation (though the presumption forms of the true propositions may be only true, so this does not trivialize these interpretations.) This fact has its basis in the ordinary truth functional definition of "if... then...". Relevance logic can prevent the undesired deductions, but so far a deontological modal theory which distinguishes relevance logic from standard logic is lacking.

B. Interpretations

An interpretation \( I = \langle U, e, i \rangle \) of LD is defined from ordinary Tarskian interpretations of L by much the obvious sort of extension given the comments above, except that we alter the treatment of ordinary propositions slightly. In particular, where before we represented true and false, we will now use \( t \) and \( f \) respectively. That is, for all \( \varphi, \psi \in LD \), we would expect, for \( \varphi \in L, e(A) = (t) \) or \( e(A) = (f) \). It turns out that for \( \psi \in LD \) \( e(A) = (t) \) or \( e(A) = (t, f) \) (i.e. the presumption is either true, false, or both), but it will never happen that \( e(A) = t \).

In particular, having defined the evaluation function over atomic propositions of LD in the usual way (except for the substitution of \( t \) and \( f \) for \( t \) and \( t \) ), we use the following induction clauses:

1. For all \( \varphi \in LD \), if \( \varphi \) has the form \( \varphi \psi \) for \( A \in L \), then \( e(A) \subseteq e(\varphi) \).
2. For all \( \varphi \in LD \), if \( \varphi \) has the form \( \varphi \psi \) for some \( \psi \in LD \), then \( e(\varphi) = e(\psi) \).
3. For all \( \varphi, \psi \in LD \), we have
   
   \[ e(\varphi \psi) = (t) \text{ if } e(\varphi) = (t); \]
   
   \[ (t, f) \text{ if } e(\varphi) = (t, f); \]
   
   \[ (t, f) \text{ if } e(\varphi) \in e(\psi); \]
   
   \[ (t, f) \text{ otherwise.} \]
   
   \[ e(\varphi \psi) = (f) \text{ if } t \in e(\psi); \]
   
   \[ (t, f) \text{ if } t \in e(\varphi); \]
   
   \[ (t, f) \text{ if } e(\varphi) \in e(\psi); \]
   
   \[ (t, f) \text{ otherwise.} \]

   \[ e(\varphi \psi) = (t) \text{ if } e(\varphi) = (t) \text{ and } e(\psi) = (t); \]
   
   \[ (t, f) \text{ if } e(\varphi) = (t) \text{ and } e(\psi) = (f); \]
   
   \[ (t, f) \text{ if } e(\varphi) \in e(\psi); \]
   
   \[ (t, f) \text{ otherwise.} \]

   \[ e(\varphi \psi) = (f) \text{ if } e(\varphi) = (f) \text{ and } e(\psi) = (t); \]
   
   \[ (t, f) \text{ if } e(\varphi) = (f) \text{ and } e(\psi) = (f); \]
   
   \[ (t, f) \text{ if } e(\varphi) \in e(\psi); \]
   
   \[ (t, f) \text{ otherwise.} \]

   \[ e(\varphi \psi) = (t) \text{ if } e(\varphi) \in e(\psi); \]
   
   \[ (t, f) \text{ if } e(\varphi) \in e(\psi); \]
   
   \[ (t, f) \text{ if } e(\varphi) \in e(\psi); \]
   
   \[ (t, f) \text{ otherwise.} \]
e(∀a) = (f) if there is a ∈ U such that e(∀u/a) = (f);
(t) if for all u ∈ U, e(∀u/a) = (t);
(t; t) otherwise.

4. If there is a ∈ LD such that e(α) = (f) but t ∈ e(pφ), then for all ψ ∈ LD, φ ∈ e(ψ).
(Notational remark: e(∀u/a) means the result of evaluating e, treating a as if it were a constant
and e(a) = u.)

The standard equivalences follow trivially from these definitions. Now we define satisfaction as follows: for all (U, e) an interpretation of LD, if e(a) = u, then e(∀a) = (f) if and only if t ∈ e(α).

VI METALOGIC

As has been indicated from the outset, the characteristic we are most concerned with for logics in inference system implementations is soundness. For this logic, the following theorem follows trivially from the soundness of standard first order predicate logic:

Theorem The logic LD is sound, i.e. for all ψ ∈ LD, e(ψ) ∈ LD, if ψ → φ then ψ → φ.

VII CONCLUDING REMARKS

A. ISSUES OF RELEVANCE

As noted above, the apparently paradoxical features which result from adopting a truth-functional view of implication can be avoided by using a deductive system based on relevance logic: such a deductive system will permit all inferences which are a proper subset of those allowed under the traditional view presented here. (1983b) has developed a variant of relevance logic for belief revision, which has been implemented in SNePS (Shapiro, 1979). An inference system including default operators as described in this paper is being implemented on this belief revision system. Limitations of space prohibit fuller discussion here; for details, see Nutter

B. SIGNIFICANCE OF THE LOGIC

It is important not to overestimate what a logic for default reasoning can do once we have one. As Israel (1980) points out, arriving at useful generalizations and resolving conflicts among conclusions of default reasoning both exceed the scope of logic. To those who wonder what help the system will be once it has deduced "pα \& p เป", the answer is, none at all. If the non-logical portion of your system contains heuristics saying that conflicting evidence indicates a need for further investigation, and if your system further contains some subsystem for investigating, then the ability to deduce statements of the kind above can be used to trigger investigation; but this goes beyond the logic alone.

It does not follow that such a logic is useless. I have already argued that implementing a logic of the kind outlined here allows implementing useful salience rules in natural language generation for instance. Furthermore, this kind of inference will play an intimate role in any system which investigates failed expectations, conflicting expectations, and the like. Only it

VIll ACKNOWLEDGMENTS

Many thanks to Stuart Shapiro and to the SNePS Research Group for their many helpful comments and suggestions. This research was carried out in the Department of Computer Science, State University of New York at Buffalo, Amherst, New York.

REFERENCES


(11) Nutter, J. T., "Default reasoning in A.I. systems" (1983b) draft.
