On Inheritance Hierarchies With Exceptions

David W. Etherington[1]
University of British Columbia

Raymond Reiter[2]
University of British Columbia
Rutgers University

Abstract

Using default logic, we formalize NETL-like inheritance hierarchies with exceptions. This provides a number of benefits:

(1) A precise semantics for such hierarchies.

(2) A provably correct (with respect to the proof theory of default logic) inference algorithm for acyclic networks.

(3) A guarantee that acyclic networks have extensions.

(4) A provably correct quasi-parallel inference algorithm for such networks.

1. Introduction

Semantic network formalisms have been widely adopted as a representational notation by researchers in AI. Schubert [1976] and Hayes [1977] have argued that such structures correspond quite naturally to certain theories of first-order logic. Such a correspondence can be viewed as providing the semantics which "semantic" networks had previously lacked [Woods 1975].

More recent work has considered the effects of allowing exceptions to inheritance within networks [Brachman 1982, Fahlman 1979, Fahlman et al 1981, Touretzky 1982, Winograd 1980]. Such exceptions represent either implicit or explicit cancellation of the normal property inheritance which IS-A hierarchies enjoy.

In this paper, we establish a correspondence between such hierarchies and suitable theories in Default Logic [Reiter 1980]. This correspondence provides a formal semantics for networks with exceptions in the same spirit as the work of Schubert and Hayes for networks without exceptions. Having established this correspondence, we identify the notion of correct inference in such hierarchies with that of derivability in the corresponding default theory, and give a provably correct algorithm for drawing these inferences. As a corollary of the correctness of this algorithm, the default theories which formalize inheritance hierarchies with exceptions can be seen to be coherent, in a sense which we will define.

We conclude, unfortunately, on a pessimistic note. Our results suggest the unfeasibility of completely general massively parallel architectures for dealing with inheritance structures with cancellation (c.f. NETL [Fahlman 1979]). We do observe, however, that limited parallelism may have some applications, but that these appear to be severely restricted in general.

2. Motivation

In the absence of exceptions, an inheritance hierarchy is a taxonomy organized by the usual IS-A relation, as in Figure 1.

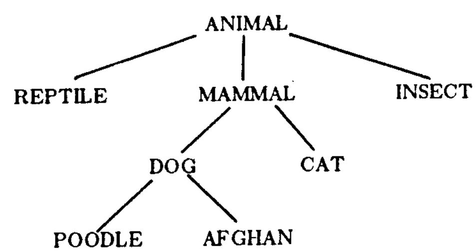


Figure 1 — Fragment of a Taxonomy

The *semantics* of such diagrams can be specified by a collection of first order formulae, such as:

$(x).POODLE(x) \supset DOG(x)$
$(x).DOG(x) \supset MAMMAL(x)$
$(x).MAMMAL(x) \supset ANIMAL(x)$
etc.

If, as is usually the case, the convention is that the immediate subclasses of a node are mutually disjoint, then this too can be specified by first order formulae:

$(x).MAMMAL(x) \supset \neg REPTILE(x)$
$(x).MAMMAL(x) \supset \neg INSECT(x)$
etc.

The significant features of such hierarchies are these:

(1) Inheritance is a logical property of the representation. Given that $POODLE(Fido)$, $MAMMAL(Fido)$ is provable from the given formulae. Inheritance is simply the repeated application of modus ponens.

(2) Formally, the node labels of such a hierarchy are unary predicates: e.g. $DOG(*)$, $ANIMAL(*)$.

(3) No exceptions to inheritance are possible. Given that Fido is a poodle, Fido must be an animal, regardless of what other properties he enjoys.

The logical properties of such hierarchies change dramatically when exceptions are permitted; non-monotonicity can arise. For example, consider the following facts about

[1] This research was supported in part by I.W. Killam Predoctoral and NSERC Postgraduate Scholarships.

[2] This research was supported in part by the Natural Science and Engineering Research Council of Canada grant A7642, and in part by the National Science Foundation grant MCS-8203954.

elephants:

(1) Elephants are gray, except for albino elephants.

(2) All albino elephants are elephants.

It is a feature of our common sense reasoning about prototypes like "elephant" that, when given an individual elephant, say Fred, not known to be an albino, we can infer that he is gray. If we subsequently discover — perhaps by observation — that Fred is an albino elephant, we must retract our conclusion about his grayness. Thus, common sense reasoning about exceptions is non-monotonic, in the sense that new information can invalidate previously derived facts. It is this feature which precludes first order representations, like those used for taxonomies, from formalizing exceptions.

In recent years, there have been several proposed formalisms for such non-monotonic reasoning (See e.g. [AI 1980]). For the purpose of formalizing inheritance hierarchies with exceptions, we shall focus on one such proposal — Default Logic [Reiter 1980]. A default theory consists of a set, W, of ordinary first order formulae, together with a set, D, of rules of inference called *defaults*. In general, defaults have the form:

$$\frac{\alpha(x_1,\ldots,x_n) : \beta(x_1,\ldots,x_n)}{\gamma(x_1,\ldots,x_n)}$$

where α, β, and γ are first order formulae whose free variables are among $x_1,\ldots,x_n$. Informally, such a default can be understood to say: For any individuals $x_1,\ldots,x_n$, if $\alpha(x_1,\ldots,x_n)$ is inferrable and $\beta(x_1,\ldots,x_n)$ can be consistently assumed, then infer $\gamma(x_1,\ldots,x_n)$. For our elephant example, the first statement would be represented by a default:

$$\frac{ELEPHANT(x) : GRAY(x) \,\&\, \neg ALBINO\text{-}ELEPHANT(x)}{GRAY(x)}$$

From the informal reading of this default, one can see that when given only ELEPHANT(Fred), GRAY(Fred) & ¬ALBINO-ELEPHANT(Fred) is consistent with this; hence GRAY(Fred) may be inferred. On the other hand, given ALBINO-ELEPHANT(Fred) one can conclude ELEPHANT(Fred) using the first order fact (x).ALBINO-ELEPHANT(x) ⊃ ELEPHANT(x), but ALBINO-ELEPHANT(Fred) "blocks" the default, thereby preventing the derivation of GRAY(Fred), as required.

The formal details of Default Logic are beyond the scope of this paper. Roughly speaking, however, for a default theory, (D,W), we think of the defaults of D as extending the first order theory given by W. Such an *extension* contains W and is closed under the defaults of D as well as first order theoremhood. It is then natural to think of an extension as defining the "theorems" of a default theory; these are the conclusions sanctioned by the theory. However, these extensions need not be unique [Reiter 1980]. For a default theory with more than one extension, any one of its extensions is interpreted as an acceptable set of beliefs that one may entertain about the world represented by that theory.

In the next section, we show how inheritance hierarchies with exceptions can be formalized as default theories. Default Logic will then be seen to provide a formal semantics for such hierarchies, just as first order logic does for IS-A hierarchies. As was the case for IS-A hierarchies, inheritance will emerge as a logical feature of the representation. Those properties, $P_1,\ldots,P_n$, which an individual, b, inherits will be precisely those for which $P_1(b)$, ..., $P_n(b)$ all belong to a common extension of the corresponding default theory. Should the theory have multiple extensions — an undesirable feature, as we shall see — then b may inherit different sets of properties depending on which extension is chosen.

3. A Semantics for Inheritance Hierarchies With Exceptions

We now show that Default Logic can provide a formal semantics for inheritance structures with exceptions. We adopt a network representation with five link types. Although other approaches to inheritance may omit one or more of these, our formalism subsumes these as special cases. The five link types,[3] with their translations to default logic, are:

(1) Strict IS-A: A.——▶.B: A's are always B's.
Since this is universally true, we identify it with the first order formula: (x).A(x) ⊃ B(x).

(2) Strict ISN'T-A: A.++++++▶.B: A's are never B's.
Again, this is a universal statement, identified with: (x).A(x) ⊃ ¬B(x).

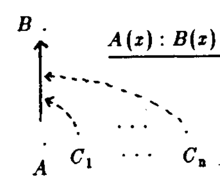
(3) Default IS-A: A.———>.B: Normally A's are B's, but there may be exceptions.
To provide for exceptions, we identify this with a default:

$$\frac{A(x) : B(x)}{B(x)}$$

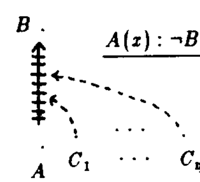
(4) Default ISN'T-A: A.++++++>.B: Normally A's are not B's, but exceptions are allowed.
Identified with:

$$\frac{A(x) : \neg B(x)}{\neg B(x)}$$

(5) Exception: A.------>
The exception link has no independent semantics; rather, it serves only to make explicit the exceptions, if any, to the above default links. There must always be a default link at the head of an exception link; the exception then alters the semantics of that default link. There are two types of default links with exceptions; their graphical structures and translations are:



$$\frac{A(x) : B(x) \,\&\, \neg C_1(x) \,\&\ldots\&\, \neg C_n(x)}{B(x)}$$



$$\frac{A(x) : \neg B(x) \,\&\, \neg C_1(x) \,\&\ldots\&\, \neg C_n(x)}{\neg B(x)}$$

We illustrate with an example from [Fahlman et al 1981].

Molluscs are normally shell-bearers.
Cephalopods must be Molluscs but normally are *not* shell-bearers.
Nautili must be Cephalopods and must be shell-bearers.

---

[3] Note that strict and default links are distinguished by solid and open arrowheads, respectively.

Our network representation of these facts is given in Figure 2.

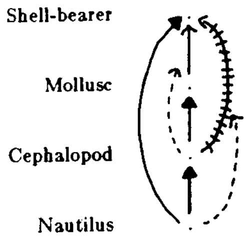


Figure 2 — Network representation of our knowledge about Molluscs.

The corresponding default theory is:

$$\{\frac{M(x) : Sb(x) \,\&\, \neg C(x)}{Sb(x)},\ (x).C(x) \supset M(x),\ (x).N(x) \supset C(x),$$

$$\frac{C(x) : \neg Sb(x) \,\&\, \neg N(x)}{\neg Sb(x)},\ (x).N(x) \supset Sb(x)\}.$$

Given a particular Nautilus, this theory has a unique extension in which it is also a Cephalopod, a Mollusc, and a Shell-bearer. A Cephalopod not known to be a Nautilus will turn out to be a Mollusc with no shell.

It is instructive to compare our network representations with those of NETL [Fahlman et al 1981]. A basic difference is that in NETL there are no strict links; all IS-A and ISN'T-A links are potentially cancellable and hence are defaults. Moreover, NETL allows exception (*UNCANCEL) links only for ISN'T-A (*CANCEL) links. If we restrict the graph of Figure 2 to NETL-like links, we get Figure 3,

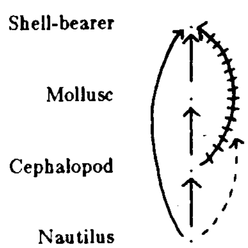


Figure 3 — NETL-like network representation of our knowledge about Molluscs.

which is essentially the graph given by Fahlman. This network corresponds to the theory:

$$\{\frac{M(x) : Sb(x)}{Sb(x)},\ \frac{C(x) : M(x)}{M(x)},\ \frac{N(x) : C(x)}{C(x)},$$

$$\frac{C(x) : \neg Sb(x) \,\&\, \neg N(x)}{\neg Sb(x)},\ \frac{N(x) : Sb(x)}{Sb(x)}\}.$$

As before, a given Nautilus will also be a Cephalopod, a Mollusc, and a Shell-bearer. A Cephalopod not known to be a Nautilus, however, gives rise to *two* extensions, corresponding to an ambivalence about whether or not it has a shell. While counter-intuitive, this merely indicates that an exception to shell-bearing, namely being a Cephalopod, has not been explicitly represented in the network. Default Logic resolves the ambiguity by making the exception explicit, as in Figure 2. NETL, on the other hand, cannot make this exception explicit in the graphical representation, since it does not permit exception links to point to IS-A links.

How then does NETL conclude that a Cephalopod is not a Shell-bearer, without also concluding that it is a Shell-bearer? NETL resolves such ambiguities by means of an inference procedure which prefers shortest paths. Interpreted in terms of default logic, this "shortest path heuristic" is intended to favour one extension of the default theory. Thus, in the example above, the path from Cephalopod to ¬Shell-bearer is shorter than that to Shell-bearer so that, for NETL, the former wins. Unfortunately, this heuristic is not sufficient to replace the excluded exception type in all cases. Reiter and Criscuolo [1983] and Etherington [1982] show that it can lead to conclusions which are unintuitive or even invalid — i.e. not in any extension. Fahlman et al [1981] and Touretzky [1981, 1982] have also observed that such shortest path algorithms can lead to anomalous conclusions and they describe attempts to restrict the form of networks to exclude structures which admit such problems. From the perspective of default logic, these restrictions are intended to yield default theories with unique extensions.

An inference algorithm for network structures is correct only if it can be shown to derive conclusions all of which lie within a single extension of the underlying default theory. This criterion rules out shortest path inference for unrestricted networks. In the next section, we present a correct inference algorithm.

4. Correct Inference

The correspondence between networks and default theories requires defaults all of which have the form:

$$\frac{\alpha(x_1, \cdots, x_n) : \beta(x_1, \cdots, x_n) \,\&\, \gamma(x_1, \cdots, x_n)}{\beta(x_1, \cdots, x_n)}.[4]$$

Such defaults are called ***semi-normal*** , and can be contrasted with *normal* defaults, in which $\gamma(x_1, \ldots, x_n)$ is a tautology. Our criterion for the correctness of a network inference algorithm requires that it derive conclusions all of which lie within a single extension of the underlying default theory. Until recently, the only known methods for determining extensions were restricted to theories involving only normal defaults [Reiter 1980]. Etherington [1982] has developed a more general procedure, which involves a relaxation style constraint propagation technique. This procedure takes as input a default theory, (D,W), where D is a finite set of closed defaults,[5] and W is a finite set of first order formulae. In the presentation of this procedure, below, the following notation is used:

S $\vdash \omega$ means formula ω is first order provable from premises S.
S $\nvdash \omega$ means that ω is not first order provable from S.
$CONSEQUENT(\frac{\alpha : \beta}{\gamma})$ is defined to be γ.

---

[4] $\alpha(x_1, \cdots, x_n)$ and $(\beta(x_1, \cdots, x_n) \,\&\, \gamma(x_1, \cdots, x_n))$ are called the *prerequisite* and *justification* of the default, respectively.

[5] A default, $\frac{\alpha : \beta}{\gamma}$, is *closed* iff α, β, and γ contain no free variables.

$$
\begin{array}{l}
H_0 \leftarrow W;\ j \leftarrow 0; \\
\textbf{repeat} \\
\quad j \leftarrow j + 1;\ h_0 \leftarrow W;\ GD_0 \leftarrow \{\ \};\ i \leftarrow 0; \\
\quad \textbf{repeat} \\
\qquad D_i \leftarrow \{ \frac{\alpha : \beta}{\gamma} \in D \mid (h_i \vdash \alpha), (h_i \not\vdash \neg\beta), (H_{j-1} \not\vdash \neg\beta) \}; \\
\qquad \text{if } \neg null(D_i - GD_i) \textbf{ then} \\
\qquad\quad \textbf{choose } \delta \textbf{ from } (D_i - GD_i); \\
\qquad\quad GD_{i+1} \leftarrow GD_i \bigcup \{\delta\}; \\
\qquad\quad h_{i+1} \leftarrow h_i \bigcup \{CONSEQUENT(\delta)\};\ \textbf{endif}; \\
\qquad i \leftarrow i + 1; \\
\quad \textbf{until } null(D_{i-1} - GD_{i-1}); \\
\quad H_j = h_{i-1} \\
\textbf{until } H_j = H_{j-1}
\end{array}
$$

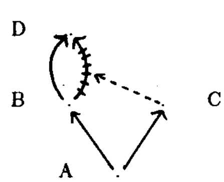
Extensions are constructed by a series of successive approximations. Each approximation, H_j, is built up from any first-order components by applying defaults, one at a time. At each step, the default to be applied is chosen from those, not yet applied, whose prerequisites are "known" and whose justifications are consistent with both the previous approximation and the current approximation. When no more defaults are applicable, the procedure proceeds to the next approximation. If two successive approximations are the same, the procedure is said to *converge*.

The choice of which default to apply at each step of the inner loop may introduce a degree of non-determinism. Generality requires this non-determinism, however, since extensions are not necessarily unique. Deterministic procedures can be constructed for theories which have unique extensions, or if full generality is not required.

Notice that there are appeals to ***non-provability*** in this procedure. In general, such tests are not computable, since arbitrary first order formulae are involved. Fortunately, such difficulties disappear for default theories corresponding to inheritance hierarchies. For these theories, all predicates are unary. Moreover, for such theories, we are concerned with the following problem: Given an individual, b, which is an instance of a predicate, P, determine all other predicates which b inherits — i.e. given $P(b)$ determine all predicates, $P_1, \ldots, P_n$, such that $P(b)$, $P_1(b)$, ..., $P_n(b)$, belong to a common extension. For this problem it is clear that predicate arguments can be ignored; the appropriate default theory becomes purely propositional. For propositional logic, non-provability is computable.

Example

Consider the network of Figure 4. Given an instance of A, the corresponding default theory has a unique extension in which A's instance is also an instance of B, C, and D. When the procedure is applied to this theory, it generates the approximations shown. (The formulae in each approximation are listed in the order in which they are derived.) ¬D occurs in H_1 since it can be inferred *before* C.



$H_0 = \{ A \}$

$H_1 = \{ A, B, \neg D, C \}$

$H_2 = \{ A, B, C \}$

$H_3 = H_4 = \{ A, B, D, C \}$

Figure 4 — Example of Procedure Behaviour

The following result is proved in [Etherington 1983]:

> *For default theories corresponding to acyclic inheritance networks with exceptions, the procedure always converges on an extension.*

As a simple corollary we have:

> *The default theory corresponding to an acyclic inheritance network with exceptions has at least one extension.*

The latter result is comforting. It says that such networks are always coherent, in the sense that they define at least one acceptable set of beliefs about the world represented by the network.

5. Parallel Inference Algorithms

The computational complexity of inheritance problems, combined with some encouraging examples, has sparked interest in the possibility of performing inferences in parallel. Fahlman [1979] has proposed a massively parallel machine architecture, NETL. NETL assigns one processor to each predicate in the knowledge base. "Inferencing" is performed by nodes passing "markers" to adjacent nodes in response to both their own states and those of their immediate neighbours. Fahlman suggests that such architectures could achieve logarithmic speed improvements over traditional serial machines.

The formalization of such networks as default theories suggests, however, that there might be severe limitations to this approach. For example, correct inference requires that all conclusions share a common extension. For networks with more than one extension, inter-extension interference effects must be prevented. This seems impossible for a one pass parallel algorithm under purely local control, especially in view of the inadequacies of the shortest path heuristic.

Even in knowledge bases with unique extensions, structures requiring an arbitrarily large radius of communication can be created. For example, both the default theories corresponding to the networks in Figure 5 have unique extensions. A network inference algorithm must reach F before propagating through B in the first network and conversely in the second. The salient distinctions between the two networks are not local; hence they cannot be utilized to guide a purely local inference mechanism to the correct choices. Similar networks can be constructed which defeat marker passing algorithms with any fixed radius.

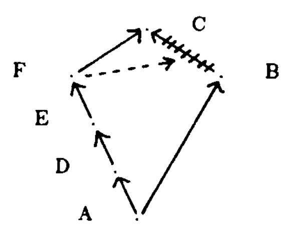


Figure 5a

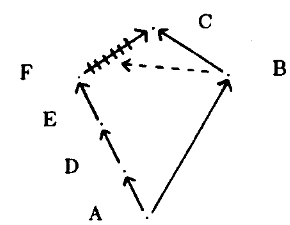


Figure 5b

Touretzky [1981] has observed such behaviour and has developed restrictions on network structures which admit parallel inferencing algorithms. In part, such restrictions appear to have the effect of limiting the corresponding default theory to one extension. Unfortunately, it is unclear how these restrictions affect the expressive power of the resulting networks. Moreover, Touretzky has observed that it is not possible to determine in parallel whether a network satisfies these restrictions.

A form of limited parallelism can be achieved by partitioning a network into a hierarchy of subnetworks, using an algorithm given in [Etherington 1983]. A parallel algorithm can then be applied to the individual subnets, in turn. The number of subnets which must be processed is bounded by the number of exception links in the network. Unfortunately, it can be shown that this technique may exclude some extensions of the theory. We have not yet characterized the biases which this induces in a reasoner.

6. Conclusions

By formalizing inheritance hierarchies with exceptions using default logic we have provided them with a precise semantics. This in turn allowed us to identify the notion of correct inference in such a hierarchy with that of derivability within a single extension of the corresponding default theory. We then provided an inference algorithm for acyclic inheritance hierarchies with exceptions which is provably correct with respect to this concept of derivability.

Our formalization suggests that *for unrestricted hierarchies,* it may not be possible to realize massively parallel marker passing hardware of the kind envisaged by NETL. Fortunately, this pessimistic observation does not preclude parallel architectures for suitably restricted hierarchies. This raises several open problems:

1a) Determine a natural class of inheritance hierarchies with exceptions which admits a parallel inference algorithm yet does not preclude the representation of our commonsense knowledge about taxonomies.

1b) Define such a parallel algorithm and prove its correctness with respect to the derivability relation of default logic.

2) In connection with (1a), notice that it is natural to restrict attention to those hierarchies whose corresponding default theories have unique extensions. Characterize such hierarchies.

7. Acknowledgments

We would like to thank Robert Mercer and David Touretzky for their valuable comments on earlier drafts of this paper.

8. References

AI (1980), Special issue on non-monotonic logic, *Artificial Intelligence 13 (1,2)*, April.

Brachman, R. (1982), "What 'IS-A' Is and Isn't", *Proc. Canadian Soc. for Computational Studies of Intelligence-82*, Saskatoon, Sask., May 17-19, pp 212-220.

Etherington, D.W. (1982), *Finite Default Theories*, M.Sc. Thesis, Dept. Computer Science, University of British Columbia.

Etherington, D.W. (1983), *Formalizing Non-monotonic Reasoning Systems*, TR-83-1, Dept. Computer Science, University of B.C. (To appear).

Fahlman, S.E. (1979), *NETL: A System For Representing and Using Real-World Knowledge*, MIT Press, Cambridge, Mass.

Fahlman, S.E., Touretzky, D.S., and van Roggen, W. (1981), "Cancellation in a Parallel Semantic Network", *Proc. IJCAI-81*, Vancouver, B.C., Aug. 24-28, pp 257-263.

Hayes, P.J. (1977), "In Defense of Logic", *Proc. IJCAI-77*, Cambridge, Mass., pp 559-565.

Reiter, R. (1980), "A Logic for Default Reasoning", *Artificial Intelligence 13*, (April) pp 81-132.

Reiter, R., and Criscuolo, G. (1983), "Some Representational Issues in Default Reasoning", *Int. J. Computers and Mathematics*, (Special Issue on Computational Linguistics), to appear.

Schubert, L.K. (1976), "Extending the Expressive Power of Semantic Networks", *Artificial Intelligence 7(2)*, pp 163-198.

Touretzky, D.S. (1981), Personal Communication.

Touretzky, D.S. (1982), "Exceptions in an Inheritance Hierarchy", Unpublished Manuscript.

Winograd, T. (1980), "Extended Inference Modes in Reasoning", *Artificial Intelligence 13*, (April), pp 5-26.

Woods, W.A. (1975), "What's In A Link?", *Representation and Understanding*, Academic Press, pp 35-82.

PROCEEDINGS

of

The National Conference on Artificial Intelligence

Sponsored by the American Association for Artificial Intelligence

August 22-26, 1983

Washington, D.C.

AAAI-83

in One Volume

Distributed by
William Kaufmann, Inc.
95 First Street
Los Altos, California 94022

ISBN 0-86576-065-9
Printed in the United States of America