(2)

A Heterogeneous Inheritance System Based on Probabilities

Fahiem Bacchus*
Department of Philosophy &
Department of Computer Science
University of Rochester
Rochester, New York
14627

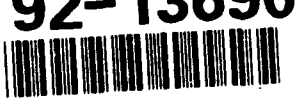
February 1, 1988

Abstract

A heterogeneous inheritance system is presented. Heterogeneous systems, in contrast to the homogeneous systems which have been the focus of recent research, make an explicit differentiation between strict IS-A links which do not allow exceptions and defeasible IS-A links which do. Some problems which arise in homogeneous systems, due to their lack of this distinction, are discussed. The inheritance system possesses a natural formal semantics which is based on a statistical interpretation of probabilities. The semantics gives meaning to all of the pieces of the system and resolves certain ambiguities present in inheritance systems. The system is not limited to acyclic inheritance nets, and can be further generalized through its foundation in a more general formal logic.

Length: ≈ 4500 words Topic: Commonsense Reasoning—Default Reasoning

*This research was supported by the University of Alberta through their Dissertation Fellowship, and through a Post-Doctoral fellowship funded by the U.S. Army Signals Warfare Laboratory; thanks also to the Department of Computer Science, University of Rochester, which provided document preparation facilities.



REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1988	Unknown

4. TITLE AND SUBTITLE

A Heterogeneous Inheritance System Based on Probabilities

5. FUNDING NUMBERS

DAAB10-86-C-0567

6. AUTHOR(S)

Fahiem Bacchus

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Rochester
Department of Philosophy
Rochester, NY 14627

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army CECOM Signals Warfare Directorate
Vint Hill Farms Station
Warrenton, VA 22186-5100

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

92-TRF-0018

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Statement A; Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

A heterogeneous inheritance system is presented. Heterogeneous systems, in contrast to the homogeneous systems which have been the focus of recent research, make an explicit differentiation between strict IS-A links which do not allow exceptions and defeasible IS-A links which do. Some problems which arise in homogeneous sytems, due to their lack of this distinction, are discussed. The inheritance system possesses a natural formal semantics which is based on a statistical interpretation of probabilities. The semantics gives meaning to all the pieces of the system and resolves certain ambiguities present in inheritance systems. The system is not limited to acyclic inheritance nets, and can be further generalized through its foundation in a more general formal logic.

14. SUBJECT TERMS

Artificial Intelligence, Data Fusion, Heterogeneous Inheritance Probabilities

15. NUMBER OF PAGES

19

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to **stay within the lines to meet optical scanning requirements.**

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	**PR** - Project
G - Grant	**TA** - Task
PE - Program Element	**WU** - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - DOD - Leave blank.
DOE - DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - NASA - Leave blank.
NTIS - NTIS - Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

1 Introduction

This paper presents an exception allowing multiple inheritance system. The system is heterogeneous; i.e., there are two types of links—strict and defeasible. This is in contrast with most of the recent work in this area (e.g., Touretzky [1], Horty et al. [2], Pearl [3]) which has explored homogeneous systems in which all links are identical.[1] Heterogeneous systems have been proposed before, e.g., in Etherington [5] and more recently in Geffner and Pearl [6]. However, to quote Touretzky et al., "Heterogeneous systems are not yet well understood." One of the purposes of this paper is an attempt to make such systems more understandable and to point out the advantages they possess.

The second feature of this system is that it has a natural formal semantics based on sets and probabilities. Probabilistic formal semantics for inheritance systems is the main focus of work by Pearl [3]; however, the semantics presented here are quite different from Pearl's ϵ-semantics and in my opinion are more natural. Understandable formal semantics have been lacking in previous systems (e.g., Touretzky [1], Horty et al. [2], Sandewall [7]).

The formal semantics provides a clear justification for all of the inferences made, resolving the ambiguities present in inheritance systems of this nature (see Touretzky et al. [4]).[2] Some may not agree with the interpretation given by these semantics; however, since the interpretation is explicit, at least a developer can decide if it is appropriate for his domain. Also, some types of inferences which seem intuitive are not supported by these semantics; however, the semantic model identifies counter examples to these inferences, showing that they are not uniformly valid. Furthermore, the additional assumptions such inferences depend on can be identified.

The following sections examine the two main features of the system, heterogeneity and probabilistic semantics. Next, the system is presented. After this, the behavior of the system is examined through the use of some common examples which have appeared in previous literature. The paper closes with some brief comments on the ambiguities of inheritance systems and on the

[1] This terminology comes from Touretzky et al. [4]

[2] The semantics does not solve these ambiguities; rather, it gives the knowledge expressed in the system a particular interpretation, and under that interpretation there is no ambiguity.

relationship between the system and a more general formal logic (Bacchus [8]). Unfortunately space prohibits a discussion of the design space taxonomy presented in Touretzky et al. [4] which is useful for comparing the system's behaviour with other systems. This task is undertaken in the longer version of this paper (Bacchus [9]) which also contains more detail on all of the points mentioned here.

2 Homogeneity

Homogeneous inheritance systems suffer from at least two major problems. The first was discussed by Brachman in [10]. The main focus of his sound criticisms is that since there is no differentiation between defeasible and necessary properties, all conclusions drawn by such systems must be defeasible. That is, it is always possible that the system may be lying. This makes it impossible for such systems to accurately represent compositional classes, e.g., three legged elephants. One could never conclude with certainty that three legged elephants possess three legs, even through, they do simply by definition. See Brachman's article for the detailed arguments behind this claim.

Another difficulty with homogeneous systems is that from a formal standpoint there is no a priori limit on the depth of an inheritance net, and thus no a priori limit on the length of paths down which properties can be inherited. Inheritance down an arbitrarily lengthy path of strict IS-A links presents no problems. For example, given the strict IS-A path

$$Tweety \Rightarrow Bird \Rightarrow Animal \Rightarrow Physical\ Object \Rightarrow Occupies\ Space,$$

the inference that Tweety occupies space is perfectly valid and intuitive. However, property inheritance down a path of defeasible links can quickly lead to counter-intuitive conclusions. For example, the path of defeasible links

$$Helicopter \rightarrow Flying\ Object \rightarrow Has\ Wings$$

leads to the counter-intuitive conclusion that helicopters have wings after only two perfectly reasonable defeasible inferences. It might be argued that in any inheritance net the node *Helicopter* should have an exception link to the node *Has Wings*. However, it is easy to see that in general avoiding all

such counter-intuitive conclusions would necessitate examining all allowable (non-preempted) paths in the inheritance net and then adding the required exception links. Besides its impracticality, I would argue that this vitiates the very purpose of inheritance systems. Inheritance systems, like any reasoning systems, are intended to generate plausible conclusions which go beyond the explicit knowledge actually contained in them. To require the addition of such intuition preserving exception links is, in a sense, requiring that the conclusions be known prior to any reasoning being performed.

In summary, through their inability to differentiate between necessary and defeasible links homogeneous systems cannot know when a lengthy chain of inheritance leads to a valid conclusion and when it leads to a counter-intuitive conclusion.

3 Probabilistic Semantics

The semantic model presented here is based on set theory and probability theory. In particular, P's are strict IS-A Q's means that the set of P's is a subset of the set of Q's, and P's are probably IS-A Q's means that greater than 100c% of all P's are Q's, where $0.5 \leq \mathbf{c} < 1$.[3]

This interpretation of defeasible links can be contrasted with the interpretation given by Pearl [3]. Pearl interprets all links (his is a homogeneous system) as holding with probability $1 - \epsilon$, where ϵ is arbitrarily close to zero. The interpretation given here is much more conservative in its claims; all that is claimed is that more that the majority of instances of P possess property Q. I feel that this gives a more natural rendering of defeasible properties like "Most birds can fly." The heterogeneous system of Geffner and Pearl [6] is also subject to this criticism. In their system defeasible links are interpreted as holding with probability almost one. Furthermore, this means that their heterogeneous system still sanctions property inheritance down arbitrarily lengthy chains of defeasible links.

As mentioned before, some intuitive inferences are not supported by this system. In particular, with this statistical interpretation of defeasibility only one defeasible link can be traversed in any path. It is easy to see that if more than 100c% of all P's are Q's and more than 100c% of all Q's are S's there is

[3]Thus it is left to the developer to decide how high c should be for his domain, or if such knowledge is not available c can be left with this loose interpretation.

no reason to suppose that more than 100c% of all P's are S's.[4] This means that property inheritance down more than one defeasible link can never be *uniformly* valid within these semantics, and indeed, the helicopter example shows that it is *not* uniformly valid.

With an additional independence assumption property inheritance can occur down more that one defeasible link. If the conditional probability of S given Q is independent of P, i.e., $[S(x)|Q(x) \wedge P(x)]_x = [S(x)|Q(x)]_x$, then it is deducible that $> (100)\mathbf{c}^2\%$ of all P's are S's. This has the natural result that each time a defeasible property is used the final conclusion becomes less certain. This possibility will not be explored further as the validity of the independence assumption is domain dependent.

Lengthy chains of inheritance, however, are not excluded—contrary to the criticism of probabilistic semantics used by Sandewall [7]. Property inheritance can occur down arbitrary lengthy chains, but all of the links in the chain will be strict IS-A links except, possibly, for one defeasible link.

It has been claimed that probabilities are inappropriate to encode notions of typicality like "birds lay eggs" (Carlson [11], Nutter [12]). It is clear that laying eggs is a typical property of birds which is not possessed by a majority of birds. This example demonstrates the need for a careful distinction between different notions of typicality. Brachman [10, page 98] has pointed out that there is a difference between prototypical properties, which are characteristic of a kind, and descriptions which specify the properties which typically apply to instances of a kind. It is certainly true that prototypical properties may not be probable properties and as such not representable with the semantics used here.[5] However, as Brachman argues, inheritance nets are more concerned with the the expression of properties which typically apply to instances of a kind, rather than prototypical properties. Indeed, given an instance of a kind the inheritance net is used to reason about the additional properties that that instance may possess. The fact that the majority of instances of that kind possess a property is surely

[4] In fact, if the set of Q's is much larger than the set of P's we could have 99% of all P's being Q's and 99% of all Q's being S's and still have *no* P's being S's (try 100 P's and 10,000 Q's).

[5] This does not necessarily mean that probabilities are not useful for expressing notions of prototypicality. For example, although the majority of birds do not lay eggs the probability of a bird laying an egg is much greater than the probability of, say, a mammal laying an egg. So, perhaps prototypical properties can be expressed by ratios of probabilities.

a good reason to conjecture that that particular instance also possesses that property.

In fact, it is not clear that inheritance systems are appropriate for reasoning about prototypical properties. The problem is that it is not usually reasonable to conclude that an arbitrary individual is prototypical. For example, if prototypical properties like "birds lay eggs" were encoded in the inheritance net, then one would be lead to rather counter-intuitive conclusions like "Tweety lays eggs," once it was known that Tweety is a bird.

4 The Inheritance Reasoner

The formal details of the inheritance reasoner are presented in this section. Formal semantics for this reasoner are provided. The semantics gives meaning to the edges and vertices contained in the inheritance net. It also allows an examination of the validity of the conclusions supported by the net.

4.1 Semantics

A semantic model, $\mathcal{M}$, of an inheritance net consists of the following triple:

$$\mathcal{M} = \langle \mathcal{D}, \mathcal{R}, \mu \rangle$$

$\mathcal{D}$ is a set of individuals, $\mathcal{R}$ is a collection of sets of individuals. Each set in $\mathcal{R}$ represents a set of individuals which share a certain property, e.g., the set of birds, the set of flying objects. Finally, μ represents a probability distribution over the field of subsets generated from the collection of sets in $\mathcal{R}$.[6]

4.2 The Inheritance Net and its Interpretation

The inheritance net, N, is a graph, $N = \langle C, L \rangle$, where C is a set of nodes and L a set of edges (links). There are two types of nodes, individual nodes, e.g., *Tweety, Fido, Nixon*, and unary predicate nodes,[7] e.g., *Bird, Fly, Elephant*,

[6] This field of subsets is the smallest collection of sets which contains $\mathcal{R}$, is closed under intersection, union, and complementation with respect to $\mathcal{D}$, and contains $\mathcal{D}$. Such a field is the minimum structure over which a probability distribution can be defined.

[7] Relations are not handled by this reasoner.

and four types of links, $\Rightarrow$ (IS-A), $\not\Rightarrow$ (IS-NOT-A), $\rightarrow$ (Probably-IS-A), and $\not\rightarrow$ (Probably-IS-NOT-A).

We also have interpretations which map the expressions in the net to a semantic model, i.e., they assign meaning to the expressions in the net. Each interpretation σ maps every individual node, c, to a particular individual, $u \in \mathcal{D}$, every predicate node, P, to a particular set of individuals, $R \in \mathcal{R}$. Symbolically, we have $c^\sigma = u$ and $P^\sigma = R$.

Given an interpretation σ with domain $\mathcal{M}$ (a model), each edge in the net makes an assertion about $\mathcal{M}$. Let P and Q represent any predicate node, and c represent any individual node. Any edge $c \Rightarrow P$ asserts that c has property P, i.e., $c^\sigma \in P^\sigma$ in $\mathcal{M}$; $P \Rightarrow Q$ asserts that *all* P's are Q's, i.e., $P^\sigma \subset Q^\sigma$; $P \rightarrow Q$ asserts that at least 100c% of the P's are Q's, i.e., $\mu(Q^\sigma \cap P^\sigma) \div \mu(P^\sigma) > \mathbf{c}$. Similarly we have the negated assertions: $c \not\Rightarrow P$: asserts $c^\sigma \notin P^\sigma$; $P \not\Rightarrow Q$ asserts $P^\sigma \subset \neg Q^\sigma$; $P \not\rightarrow Q$ asserts $\mu(Q^\sigma \cap P^\sigma) \div \mu(P^\sigma) < 1 - \mathbf{c}$.

Note, $c \rightarrow P$ is not a valid link in the net. Probable links cannot emanate from nodes representing individuals in the inheritance net, i.e., all of the initial properties of an individual are assumed to be known with certainty.

Given an inheritance net N, the models of N, $M(N)$, is a set of models. For each model $\mathcal{M} \in M(N)$ there exist an interpretation (could be more than one) whose domain is $\mathcal{M}$ and under which all of the assertions made by N are true in $\mathcal{M}$. Generally a net will have many different models. In particular, the defeasible formulas (edges) in N only place constraints on the probability distribution μ, and there may be many different distributions which satisfy these constraints. In other words, the probability distribution is not fully specified; all that is known are certain majority relationships which hold between the properties.

It is required that the net N satisfy the consistency constraint that $M(N)$ be nonempty; i.e., there is some model with associated interpretation under which all of the assertions of N are true.

Reasoning in the net about an individual is accomplished by following certain types of paths emanating from the node representing that individual. The property nodes reachable from the node through legal non-preempted paths are possible properties of that individual.

The next section defines what constitutes legal paths, as well as what it means for a path to be preempted. It also discuss the meaning of traversing the strict and defeasible links.

4.3 Reasoning

Some of the inferences generated by the inheritance system are simply deductive consequences of the assertions in the net. For example, if there is a path $c \Rightarrow P \Rightarrow Q$ in the net, then it is easy to see that the inference that c is a Q is a valid deduction. The link specifies strict set containment; hence, every P, including c, must be a Q.

The situation is different when the path $c \Rightarrow P \rightarrow Q$ is in the net. The probable link specifies that the majority of P's are Q's, but it says nothing about any particular $c \in P$ being a Q. If c is assumed to be an arbitrary or in no way distinguished member of P then one can justifiably claim that c is more likely a Q than not. This claim involves an inductive assumption of randomization; it is not a strictly deductive consequence of the knowledge.

This inductive assumption has a long history among empirical interpreters of probability (see Kyburg [13]). It is similar to the manner in which sense is made out of statements like "The probability that a coin flip will yield heads is 50%." It is clear that on any particular flip the outcome will be either heads or tails. Yet, when that particular flip is undistinguished from any other it is reasonable to assume that it will yield heads with probability 0.5. The particular event is assigned a probability equal to the underlying relative frequency.[8] In a similar manner, an individual is assumed to possess properties with probability equal to the relative frequency of those properties in the class to which the individual is a member.

Of course, an individual may be a member of more than one class, and in each class the relative frequency of particular properties may be very different. In this case it is necessary to chose which class is the correct one for inducing the probability of the property in question. This is the problem of choosing the correct reference class; it also has a long history (see Kyburg [14] also [15]). However, for the relative simple case of inheritance nets there is a reasonable solution to this problem.

The fundamental intuition for property inheritance is that the properties of subclasses should override superclasses. This notion is also justifiable in probability theory, and is used to chose the correct reference class. Say we have the following knowledge: *Tweety* is a penguin and a bird, penguins are

[8] Kyburg [14] gives a very reasonable argument that this approach to probabilities is not be limited to events which are inherently repeatable.

birds, birds probably fly, and penguins probably do not fly.[9] If *Tweety* was considered to be an arbitrary bird then, through the inductive assumption, he would probably be able to fly. On the other hand if *Tweety* was considered to be an arbitrary penguin then he probably would not be able to fly. Actually, given what is known about *Tweety*, he is not an arbitrary bird nor an arbitrary penguin; rather he is, to the best of our knowledge, an arbitrary member of the intersection of these two classes; i.e., the best reference class is the intersection of these two classes. However, since penguins are a subset of birds, this intersection is equal to the set of penguins. Hence, Tweety should be considered to be an arbitrary penguin and thus probably not able to fly. Formally, this is a well known property of probabilities, i.e., $Penguin^{\sigma} \subset Bird^{\sigma}$ implies that

$$\frac{\mu(Fly^{\sigma} \cap Bird^{\sigma} \cap Penguin^{\sigma})}{\mu(Bird^{\sigma} \cap Penguin^{\sigma})} = \frac{\mu(Fly^{\sigma} \cap Penguin^{\sigma})}{\mu(Penguin^{\sigma})}.$$

These observations serve as the intuitive basis for the definitions which follow. First, we define the legal paths and the conclusions they support, then path preemption, and finally, once the interaction between paths is taken into consideration, the conclusions supported by the net as a whole.

P is used to represent a property node, a superscripted plus sign (i.e., $^{+}$) indicates one or more iterations of the link it is applied to, and a superscripted star (i.e., *) indicates zero or more iterations of the link it is applied to.

Definition 4.1 (Paths) *Positive and Negative, Necessary and Probable Paths are defined as follows:*

Positive Necessary 1 A path $c \Rightarrow^{+} P$ supports the conclusion *c is certainly a P.*

Negative Necessary 1 A path $c \Rightarrow^{+} \bullet \not\Rightarrow \bullet \Leftarrow^{*} P$ supports the conclusion *c is certainly not a P.*

Negative Necessary 2 A path $c \Rightarrow^{+} \bullet \not\Leftarrow \bullet \Leftarrow^{*} P$ supports the conclusion *c is certainly not a P.*

[9] Actually, if we are reasoning about birds and penguins in their natural state a more intuitive rendering would be that penguins do not fly, i.e., no uncertainty is involved. In this case other inferences become possible, see figure 4.

Positive Probable 1 A path $c \Rightarrow^+ \lambda \rightarrow \bullet \Rightarrow^* P$ supports the conclusion *probably c is a P founded on* λ. That is, this conclusion is based on c being an arbitrary member of the reference class λ.

Negative Probable 1 A path $c \Rightarrow^+ \lambda \not\rightarrow \bullet \Leftarrow^* P$ supports the conclusion *probably c is not a P founded on* λ.

Negative Probable 2 A path $c \Rightarrow^+ \lambda \rightarrow \alpha \cdots P$, where $\cdots$ represents a negative necessary path (either type 1 or 2) from α to P, supports the conclusion *probably c is not a P founded on* λ.

To return briefly to the requirement of net consistency, it is worth mentioning that for N to have a nonempty set of models at least requires that the net not contain both a necessary positive and a necessary negative path from any individual to the same property node. Also, it requires that the net not contain both a probable positive and a probable negative path founded on the same reference class from any individual c to the same property node. *It is not required, however, that the net be acyclic.* In fact two properties may often be cross correlated (i.e., existence for either may provide evidence for the other). The requirement of acyclicity has been cited as a deficiency of previous inheritance reasoners (Geffner and Pearl [6]). See figure 3 for an example of a cyclic net representing a reasonable set of knowledge.

As long as the net is consistent the necessary paths will sanction inferences which can never be wrong; thus, such paths can never be preempted. Probable paths can, however, be preempted both by necessary paths and by probable paths of the opposite polarity.

Definition 4.2 (Path Preemption) *Probable paths are preempted if any of the following conditions hold:*

1. A probable path from c to P is *preempted* if there is a necessary path from c to P. In this case the polarity of the two paths is irrelevant.

2. A probable path from c to P founded on λ_1 is *preempted* if there is a probable path of the opposite polarity from c to P founded on λ_2 such that there exists a path $\lambda_2 \Rightarrow^+ \lambda_1$ in the net, i.e., if the net supports the conclusion that λ_2 is a subset of λ_1.

3. A probable path from c to P is *preempted* if any of its subpaths which originate at c is preempted.

Definition 4.3 (Conclusions Supported by the Net) *Given an individual c, the net supports the following conclusions:*

1. *c is certainly a P* if a positive necessary path exists from c to P.
2. *c is certainly not a P* if a negative necessary path exists from c to P.
3. *c is evidently a P* if there exists a non-preempted positive probable path from c to P, and there does not exist any non-preempted negative probable paths from c to P.
4. *c is evidently not a P* if there exists a non-preempted negative probable path from c to P, and there does not exist any non-preempted positive probable paths from c to P.
5. The net is *ambiguous* about c being a P if there exists both a non-preempted positive probable and a non-preempted negative probable path from c to P.

4.4 On the Validity of the Conclusions

All of the certain conclusions supported by the net are valid in the sense of classical logical validity. That is, for every interpretation σ, with model $\mathcal{M}$, if all of the assertions of N are true under that interpretation then the certain conclusions are also true. To see that this claim is true just use some simple set theory, the semantic interpretation of the edges, and the definition of the necessary paths.

The evidential conclusions are justifiable in the sense that there exists knowledge in the net about that particular individual which supports the conclusions with probability greater than **c**. Knowledge about a particular individual is any property that the individual is known to have with certainty, i.e., any property P which is a certain conclusion supported by the net. An individual's known properties are the only possible reasons for conjecturing its defeasible properties. For example, the only reason for supposing that *Tweety* can fly is that he is know to be a bird, or that she is known to be an aeroplane, etc.

From the construction of the probable paths it can be seen that λ, the founding or reference class, is a known property of the individual. It is easy to show that for all positive probable paths more than 100**c**% of the members of the reference class, λ, possess the concluded property, P. Hence, under the inductive assumption that c is an arbitrary member of λ, the probability that c is a member of P, i.e., has property P, is greater that **c**.

Similarly, it can be shown that for all negative probable paths less than $100(1 - \mathbf{c})\%$ of the members of the reference class possess the concluded property. Or equivalently, more than 100**c**% possess the negated property.

The conclusions supported by these paths are not always the best conclusions because they do not consider all of the knowledge available. In the Nixon diamond (figure 2) the conclusion that *Nixon* is a pacifist founded on him being a Quaker does not consider the contradictory evidence that he is also a Republican. When both a positive and a negative probable path (non-preemptied) exists from an individual to a property, there is both positive and negative evidence for claiming that the individual possesses that property. Hence, the net is ambiguous about this claim.

A probable path preempts a probable path of the opposite polarity if its reference class, λ_1, is a subset of the reference class of the preempted probable path, λ_2. As was discussed before, this has a formal justification in probability theory. The probability of the concluded property given both pieces of knowledge, λ_1 and λ_2, is equal to its probability given just λ_1. Hence, the preempting path is founded on knowledge which already considers the knowledge used in the preempted path. In a similar manner, if a subpath of a path is preempted the path loses its foundation and should not further influence any other paths (i.e., it should not preempt any other path). For a more detailed discussion on this point, and on validity in general see Bacchus [9].

The existence of non-preemptied probable paths of only one polarity from an individual c to a property P indicates that all of the knowledge (evidence) in the net about c's P-ness is of that polarity. It is in this sense that the evidential conclusions supported by the net are valid.

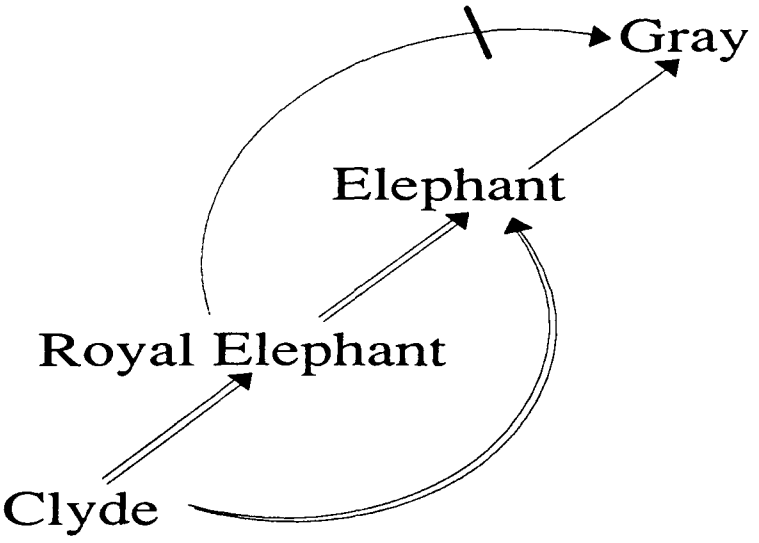


Figure 1: Redundant Information

5 Behavior of the Reasoner

This section examines the behavior of the reasoner presented in the previous section through the use of some examples. All of the examples have appeared previously in the literature. However, since this reasoner is heterogeneous, the examples have been altered. In particular, some of the edges in these examples have been changed to strict IS-A links. In the examples presented here it is fairly clear which edges should be strict.

First, we show that the reasoner satisfies the main desiderata for inheritance reasoners (Touretzky [1])—the ability to deal with redundancy and ambiguity. The grey elephant net is shown in figure 1. In this net a non-preempted negative probable path exists from *Clyde* to the property node *Grey* founded on the property *Royal Elephant*. The opposing positive probable path founded on the node *Elephant* is preempted. Hence, the net sanctions the conclusion that *Clyde* is probably not grey, based of *Clyde* being an arbitrary royal elephant. An interesting point is that the plausible conclusion generated by the system carries with it explicit mention of the inductive assumption upon which it is based.

In figure 2 there is true ambiguity. In this net, called the Nixon diamond,

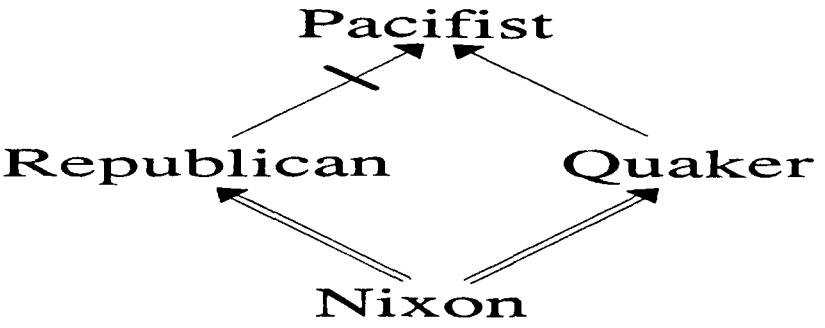


Figure 2: Ambiguous Information

there is both a positive and a negative probable non-preempted path to the node *Pacifist*, sanctioning the conclusion that the net is ambiguous. The ambiguity corresponds to the fact that neither of the founding sets for the two paths, *Republican* and *Quaker*, is a subset of the other.

The next example, in figure 3, is due to M. Ginsberg. This net supports the conclusion that *Nixon* is probably politically motivated, but is at the same time ambiguous about *Nixon* being either a dove or a hawk. The knowledge contained in the net is quite reasonable, and the conclusions it supports intuitive. However, it contains a cycle and thus cannot be dealt with by either Touretzky's system [1] nor Horty et al.'s skeptical reasoner [2].

The negative paths generated by backward IS-A links can sometimes generate interesting conclusions, as in figure 4, although, many times the conclusions are uninteresting negative facts.[10] This net supports the conclusion that *Tweety* is probably not a penguin founded on *Tweety* being an arbitrary bird (through a negative probable 2 path). Semantically, the conclusion that most ($> 100c\%$) birds are not penguins is entailed by the knowledge in this net. Geffner and Pearl [6] have argued that the ability to make such inferences, which essentially are a consequence of the properties of probabilities, represents an argument in favor of using probabilities as a semantic foundation for defeasible inference.

[10] As George Bernard Shaw once said "An intelligent man wants to know what you believe, not what you don't believe"

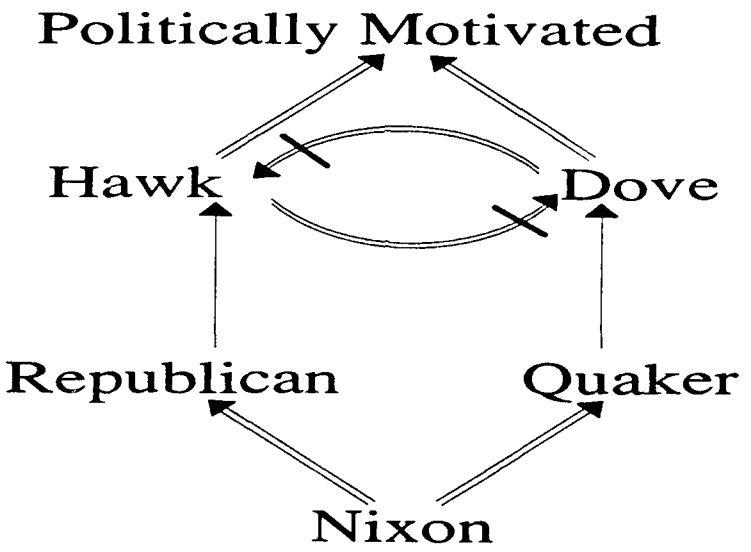


Figure 3: (M. Ginsberg) Is Nixon Politically motivated?

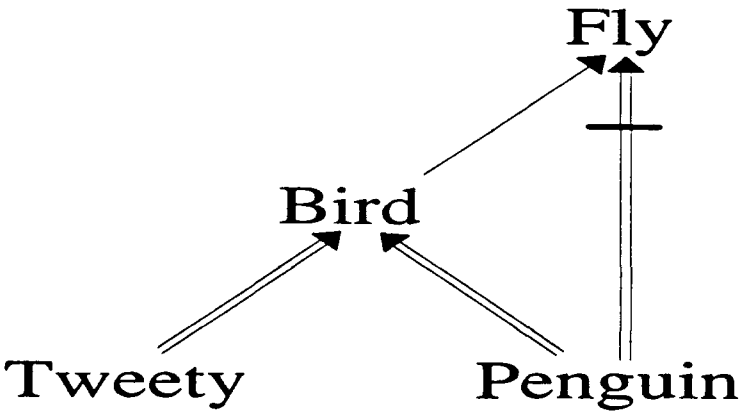


Figure 4: Tweety is probably not a penguin

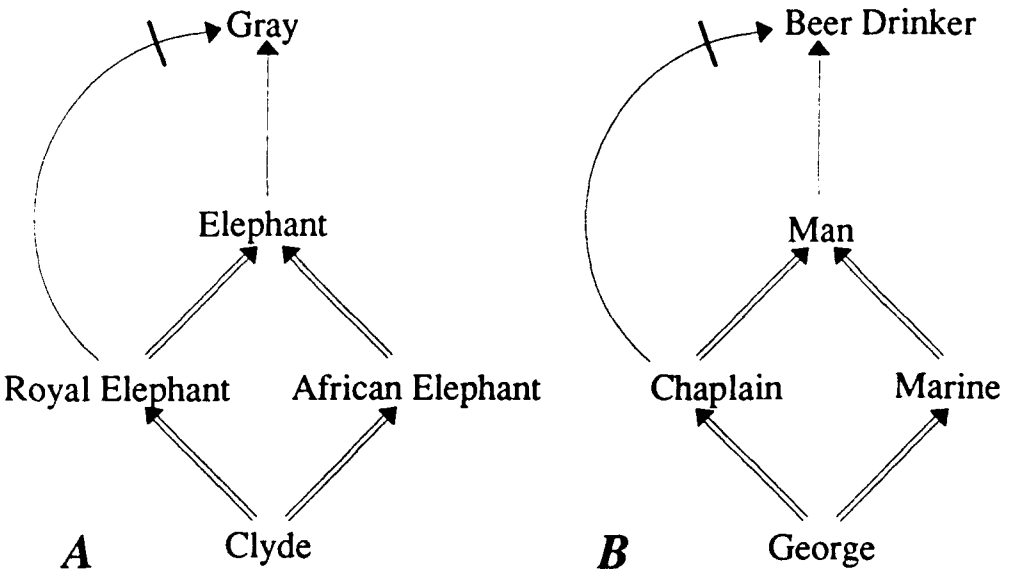


Figure 5: On-path vs. Off-path preemption

6 Ambiguity and Generalizations

The semantics resolves the ambiguities present in inheritance systems by giving a particular interpretation to the knowledge expressed in the net. There is insufficient space to discuss all of the ambiguities presented by Touretzky et al. [4], but it can be demonstrated how the on-path/off-path ambiguity is resolved.

Touretzky's credulous reasoner [1] using on-path preemption concludes that *Clyde's* grayness is ambiguous in the homogeneous version of the net in figure 5A (just change all of the strict links to defeasible ones). Sandewall [7], on the other hand, claims that this net should be unambiguous about *Clyde's* grayness and advocates the use of off-path preemption.

With these semantics it can be seen that this net contains no information about the grayness of African elephants. African elephants inherit their grayness from the superclass elephant; thus, as Sandewall argued, the fact that *Clyde* is a Royal elephant, which are known to be normally non-gray, should override the normal grayness of average elephants. If we have specific knowledge about the grayness of African elephants, then this knowledge could be encoded in the net through an explicit probable link from *African elephant*

to *Gray*. In that case there would also be a non-preempted positive probable path from *Clyde* to *Gray* and the net would be ambiguous about *Clyde's* grayness.

The same reasoning applies to the net in figure 5B. The homogeneous version of this net was cited by Touretzky et al. as being a possible counter-example to Sandewall's argument. Touretzky et al. argue that *George's* beer drinking habits should be ambiguous in this net, because, although he is a chaplain he is also a marine. They do, however, acknowledge that possibly the node *Marine* should have its own link to *Beer Drinker*, if we have explicit knowledge about the beer drinking habits of marines. Again, when the knowledge in this net is examined through the semantics provided, it can be seen that the net contains no information about the beer drinking properties of marines. Hence, with these semantics there is no choice but to add an explicit probable link from *Marine* to *Beer Drinker*. If this is done the net will be ambiguous about *George's* beer drinking habits.

The semantic model presented here can be expanded to include functions and n-place relations. The expanded model can serve as the semantics for a full logic capable of expressing a wide range of statistical knowledge. Such a logic is described in Bacchus [8].

This logic, called **Lp** also allows for the expression of comparative, interval, functional, as well as point valued probabilities. The logic is an extension of first order logic; thus, it can express composite classes formed with logical connectives. It also possesses a sound and complete proof theory.

In [8] it is shown how this deductive proof theory, along with an inductive assumption similar to the one used here, gives a general treatment of inheritable properties. This treatment is not limited to primitive classes and one place relations. It can handle complex classes, like the class of Republicans who are also Quakers, and n-place relations, like elephants loving zookeepers. Exceptions to the n-place relations, like Fred—a zookeeper who the elephants hate, are also handled by this logical system.

The advantage of the graph based reasoner is that since it deals with a limited subset of the knowledge expressible in **Lp** (only one-place relations and probabilisic majority relations) arbitary theorem proving can be replaced by path finding in the net. The author is currently working on an implementation which uses both the logic **Lp** and an inheritance graph. The aim is that the inheritance graph can be used to produce some inferences rapidly and the underlying logic can be used to deal with more complex problems.

7 Conclusions

A sound inheritance reasoner with a natural formal semantics has been presented. The semantics gives meaning to all of the pieces of the system. The reasoner represents a departure from the homogeneous inheritance reasoners previously developed in that it makes an explicit differentiation between strict IS-A links and defeasible ones. In this way it allows a more intuitive rendering of the knowledge usually encoded in inheritance nets.

References

[1] David S. Touretzky. *The Mathematics of Inheritance Systems. Research Notes in Artificial Intelligence*, Pitman, London, 1986.

[2] John F. Horty, Richmond H. Thomason, and David S. Touretzky. A skeptical theory of inheritance in nonmonotonic semantic networks. In *AAAI-87*, pages 358–363, 1987.

[3] Judea Pearl. *Probabilistic Semantics for Inheritance Hierarchies with Exceptions.* Technical Report CSD870052, Cognitive Systems Laboratory, U.C.L.A., Los Angeles, CA. 90024–1596, U.S.A., 1987.

[4] David S. Touretzky, John F. Horty, and Richmond H. Thomason. A clash of intuitions: the current state of nonmonotonic multiple inheritance systems. In *IJCAI-87*, pages 476–482, 1987.

[5] David W. Etherington. Formalizing nonmonotonic reasoning systems. *Artificial Intelligence*, 31:41–85, 1987.

[6] Hector Geffner and Judea Pearl. *Sound Defeasible Inference.* Technical Report CSD870058, Cognitive Systems Laboratory, U.C.L.A., Los Angeles, CA. 90024–1596, U.S.A., 1987.

[7] Erik Sandewall. Nonmonotonic inference rules for multiple inheritance with exceptions. *Proceedings of the IEEE*, 74(10):1345–1353, October 1986.

[8] Fahiem Bacchus. *Statistically Founded Degrees of Belief.* Technical Report 87-02, Alberta Center for Machine Intelligence, University of Alberta, Edmonon, Alberta, Canada. T6G-2E9. 1987.

[9] Fahiem Bacchus. *A Heterogeneous Inheritance System Based on Probabilities.* Technical Report 87-03, Alberta Center for Machine Intelligence, University of Alberta, Edmonon, Alberta, Canada. T6G-2E9, 1987.

[10] Ronald J. Brachman. I lied about the trees. *AI Magazine*, 6(3):80-93, 1985.

[11] G. Carlson. *Reference to Kinds in English.* PhD thesis, University of Massachusetts, 1977. unpublished—available from Indiana University Linguistics Club.

[12] Jane Terry Nutter. Uncertainty and probability. In *Proceedings of the 10th IJCAI*, pages 373-379, 1987.

[13] Henry E. Kyburg, Jr. *The Logical Foundations of Statistical Inference.* D. Reidel, 1974.

[14] Henry E. Kyburg, Jr. Objective probabilites. In *IJCAI-87*, pages 902-904, 1987.

[15] Henry E. Kyburg, Jr. The reference class. *Philosophy of Science*, 50(3):374-397, September 1983.