

JOHN F. HORTY

A SKEPTICAL THEORY OF MIXED INHERITANCE

1. INTRODUCTION

This paper is concerned with the problem of providing a semantic account for inheritance networks capable of representing both strict and defeasible information. The importance of representing defeasible information in a knowledge base — particularly, in a frame- or network-based inheritance reasoner — has been widely recognized ever since the publication of Minsky's original paper on frames [11]. Although early systems designed to allow defeasible inheritance reasoning, such as FRL [12] and NETL [6], were subject to semantic difficulties in their treatment of cancellation, these problems by now are essentially solved. In fact, there exist today a number of well-defined and intuitively attractive theories of defeasible inheritance, including those of Touretzky [15], Sandewall [13], and Horty et al. [8]. The variety of these theories does not seem to indicate any kind of instability or chaos in our understanding, but instead, the presence of a range of options in the design space for defeasible inheritance reasoners; some of these options are surveyed in Touretzky et al. [16].

It has been suggested, however, by Brachman [3, 4] and Israel [10], that this intense concern with the issue of representing defeasible information has obscured some of the more general problems exhibited by network formalisms capable of representing only defeasible information. Unless a formalism is able to represent strict or indefeasible information as well, it cannot express the important analytic or definitional relations among concepts. A system capable of representing only defeasible information cannot capture, for example, the necessary subsumption of Indian elephants under elephants — a relation which can be represented, without any kind of inconsistency, the fact that some particular Indian elephant might just happen not to be an elephant.

One way of responding to these difficulties is to abandon the project of encoding defeasible information in an inheritance reasoner, concentrating instead on defining a set of relations among concepts. This approach, which has been explored by the author [5] and his descendants, has led to a number of uses of inheritance networks, including the automatic placement of objects in a classification algorithm. However, this approach also carries with it the danger of forcing defeasible information, which must be represented in some way in a knowledge base, to be represented outside of the inheritance reasoner. The problem with this idea, of course, is that it is hard to use the defeasible information if it were represented in this way. The default logic is not a realistic possibility.

An alternative approach is to use a single inheritance network. This paper

Master

**Intelligent solutions
for managing
knowledge assets**
Carnegie Group, Inc.
412/642-6900

presents a first step along these lines: a theory of inheritance for semantic networks containing strict or defeasible, positive or negative IS-A links, all mixed together. The analysis of mixed inheritance described here is itself a mixture, combining the theory of strict inheritance from Thomason et al. [14] with the skeptical theory of defeasible inheritance provided by Horty et al. [8].

2. BASIC CONCEPTS

2.1 Links and nets

Letters from the beginning of the alphabet (a through d) refer only to objects or individuals; letters from the middle of the alphabet (m through t) refer only to properties or kinds. Letters from the end of the alphabet (u through z) range over both objects and properties.

The link types $x \Rightarrow p$ and $x \not\Rightarrow p$ represent positive and negative *strict* statements. If x is a property, these positive and negative strict links are equivalent to quantified conditionals: the link $q \Rightarrow p$ represents a statement of the form 'Every Q is a P '; the link $q \not\Rightarrow p$ represents a statement of the form 'No Q is a P '. If x is an object, these positive and negative strict links are equivalent to atomic and neg-atomic statements from ordinary logic: $a \Rightarrow p$ and $a \not\Rightarrow p$ represent the statements Pa and $\neg Pa$.

The link types $x \rightarrow p$ and $x \not\rightarrow p$ represent positive and negative *defeasible* statements. If x is a property, these defeasible links correspond to ordinary generic statements: $q \rightarrow p$ and $r \not\rightarrow p$, for example, might stand for the statements 'Birds fly' and 'Mammals don't fly'. There is nothing in classical logic very close in meaning to generic statements like these. For example, 'Birds fly' does not mean that all birds fly, since it is true even in the presence of exceptions. Instead, it seems to mean that "typical birds" fly — or that for any given bird a , it is most natural to suppose that a flies. If x is an object, it is more difficult to find a simple reading for these defeasible links; but we will assume that $a \rightarrow p$ and $a \not\rightarrow p$, respectively, mean something along the lines of 'It is most natural to suppose that Pa ' and 'It is most natural to suppose that $\neg Pa$ '.

Capital Greek letters represent *networks* — finite graphs, with nodes and link types as described. Networks are themselves classified as *strict* if they contain only strict links, or *defeasible* if they contain no strict links emanating from property nodes. *Mixed* networks can contain both strict and defeasible links emanating from property nodes.

2.2 Paths

Lower case Greek letters refer to *paths* — special sequences of links. Often, it is convenient to refer to an arbitrary path in a way that displays some of the nodes it passes through without displaying the particular link types connecting those nodes. For this purpose, we adopt a notation according to which ' $\pi(x, \sigma, y)$ ' refers to an arbitrary positive path, and ' $\bar{\pi}(x, \sigma, y)$ ' likewise to an arbitrary negative path, from x through σ to y . As a convention governing this π -notation, we assume that

adjacency
Thus, for
link of any
Paths
negative.
positive o
inductivel

(1)

(2)

(3)

(4)

It follow
initial no
possibly
of the fo
follows f
occur on

2.3 Inher

Intuitive
conclusi
if it is st
 x is a ki
strict or
network
supporte
the state
defining
 Γ permi
argumen

Since on
from [1
these tw

adjacency of node symbols entails adjacency of nodes on the paths symbolized. Thus, for example, ' $\bar{\pi}(x, u, \sigma, y)$ ' refers to a negative path beginning with a direct link of any type from x to u , and then moving through σ to y .

Paths are classified as simple or compound, strict or defeasible, positive or negative. The simple paths are just the direct links — classified as strict or defeasible, positive or negative, along with the links themselves. Compound paths are defined inductively, as follows.

- (1) If $\pi(x, \sigma, p)$ is a strict positive path, then: $\pi(x, \sigma, p) \Rightarrow q$ is a strict positive path; $\pi(x, \sigma, p) \not\Leftarrow q$ is a strict negative path; $\pi(x, \sigma, p) \rightarrow q$ is a defeasible positive path; and $\pi(x, \sigma, p) \not\leftarrow q$ is a defeasible negative path.
- (2) If $\bar{\pi}(x, \sigma, p)$ is a strict negative path, then: $\pi(x, \sigma, p) \Leftarrow q$ is a strict negative path.
- (3) If $\pi(x, \sigma, p)$ is a defeasible positive path, then: $\pi(x, \sigma, p) \Rightarrow q$ is a defeasible positive path; $\pi(x, \sigma, p) \not\Leftarrow q$ is a defeasible negative path; $\pi(x, \sigma, p) \rightarrow q$ is a defeasible positive path; and $\pi(x, \sigma, p) \not\leftarrow q$ is a defeasible negative path.
- (4) If $\bar{\pi}(x, \sigma, p)$ is a defeasible negative path, then: $\pi(x, \sigma, p) \Leftarrow q$ is a defeasible negative path.

It follows from this definition that an individual can occur in a path only as its initial node. Let us define a *negative segment* as a strict or defeasible negative link, possibly followed by a reverse positive strict path — that is, as a link sequence either of the form $x_1 \not\leftarrow x_2 \Leftarrow \dots \Leftarrow x_n$ or of the form $x_1 \not\Leftarrow x_2 \Leftarrow \dots \Leftarrow x_n$. Then it follows from this definition also that if a negative segment occurs in a path, it can occur only at the very end.

2.3 Inheritance

Intuitively, paths represent arguments, which *enable* certain statements as their conclusions. A positive path of the form $\pi(x, \sigma, y)$ enables the statement $x \Rightarrow y$ if it is strict or x is an individual, and the statement $x \rightarrow y$ if it is defeasible and x is a kind; likewise, a negative path of the form $\bar{\pi}(x, \sigma, y)$ enables $x \not\Leftarrow y$ if it is strict or x is an individual, and $x \not\leftarrow y$ if it is defeasible and x is a kind. Given a network Γ , the purpose of an inheritance theory is to specify the set of statements *supported* by Γ — that is, the set of statements we can reasonably conclude from the statements contained in Γ . We arrive at this specification in a roundabout way, defining a statement as supported by Γ just in case it is enabled by a path that Γ permits. It remains only to define the paths *permitted* by Γ — intuitively, the arguments sanctioned in the context of Γ .

3. MOTIVATION

Since our approach to mixed inheritance combines the theory of strict inheritance from [14] with the theory of defeasible inheritance from [8], we first summarize these two theories, and then explain how they fit together.

3.1 Strict inheritance

For strict networks, our definition of permitted paths is very simple. According to the theory of [14], a strict network Γ permits exactly the paths it contains — that is, Γ permits σ just in case σ is a path entirely composed of links contained in Γ . In the case of Γ_1 (Figure 1), for example, the permitted paths include $a \Rightarrow s \Rightarrow r$ and $p \Rightarrow q \not\Leftarrow r \Leftarrow s$. Suppose we interpret the nodes in this net so that $p = \text{starlings}$, $q = \text{birds}$, $r = \text{mammals}$, $s = \text{dogs}$, and $a = \text{Rover}$. Then the first of these paths shows us how Γ_1 supports the conclusion that Rover is a mammal ($a \Rightarrow r$); the second shows how it supports the conclusion that no starlings are dogs ($p \not\Leftarrow s$). The net does not permit, for example, the path $p \Rightarrow s \Rightarrow r$, since the link $p \Rightarrow s$ is not contained in Γ_1 . Likewise, $q \not\Leftarrow r \Leftarrow s \Leftarrow a$ is not permitted, even though all its links are contained in Γ_1 , since it is not a path.

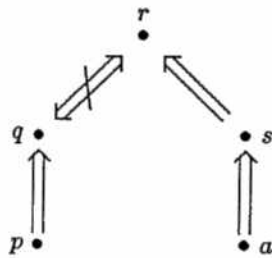


Figure 1: Γ_1

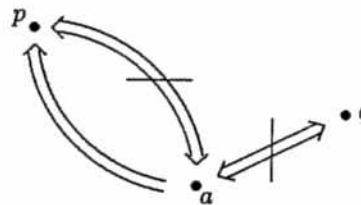


Figure 2: Γ_2

It is important to note that this analysis of strict inheritance, although straightforward, is not the standard view. Strict networks contain only strict links, each of which is equivalent, as explained above, to a formula of classical logic. It may seem natural, then, to use classical logic itself to provide a semantics for such a network — by identifying the network with the set of formulas that translate its links, and then defining a statement as supported by the network just in case it belongs to the deductive closure of that set. This idea, which we take to be the standard view, is due originally to Hayes [7]. To see that it is different from the theory of [14], consider, for example, the net Γ_2 (Figure 2). This network would be translated into the set $\{Pa, \neg Pa, \neg Qa\}$. Since the set is inconsistent, any statement at all belongs to its classical deductive closure; so according to the standard view, Γ_2 should be taken to support every statement — including, say, Qa . According to the analysis of [14], however, Γ_2 does not support Qa , since it permits no positive path from a to q , and in fact provides uncontested evidence that $\neg Qa$.

It is, in some ways, a delicate matter to decide between the analysis of strict inheritance provided by [14] and the traditional analysis of [7]. One is always free to regard a strict network simply as a notational variant of some classical theory, so that the analysis of [7] would be appropriate. Still, there seems to be some value in taking seriously the graph-based nature of inheritance reasoners, which derive conclusions

corresponding
make logical s
by forcing the
This task is ca
style proof the
resulting logic
Dunn [6].

The proo
sequents of the
(a link). Infor
The sequent ca

where A is ato
In addition, we
on the right an

In the rule
not containing

corresponding only to actual paths. The problem is then to see whether we can make logical sense of such a reasoner by designing an appropriate logic, rather than by forcing the reasoner to conform to the standards of an already-existing logic. This task is carried out for strict networks in [14], which provides both a Gentzen-style proof theory for path-based inheritance reasoning and an interpretation of the resulting logic in a four-valued model based on that explored by Belnap [1, 2] and Dunn [6].

The proof theory of [14] is a calculus in the style of Gentzen [8] for proving sequents of the form $\Gamma \vdash A$, where Γ is a set of statements (a net) and A a statement (a link). Informally, such a sequent is supposed to mean that A is derivable from Γ . The sequent calculus contains as its *structural rule* the schema

$$A \vdash A,$$

where A is atomic — that is, of the form $a \Rightarrow p$ or $a \not\Rightarrow p$. This gives us our axioms. In addition, we have the following *logical rules*, for introducing both \Rightarrow and $\not\Rightarrow$, on the right and on the left of the turnstile.

$$\frac{\Gamma^a, a \Rightarrow p \vdash a \Rightarrow q}{\Gamma^a \vdash p \Rightarrow q} \vdash \rightarrow$$

$$\frac{\Gamma \vdash a \Rightarrow p \quad \Delta, a \Rightarrow q \vdash A}{\Gamma, \Delta, p \Rightarrow q \vdash A} \rightarrow \vdash$$

$$\frac{\Gamma \vdash a \not\Rightarrow q \quad \Delta, a \not\Rightarrow p \vdash A}{\Gamma, \Delta, p \Rightarrow q \vdash A} \rightarrow \vdash'$$

$$\frac{\Gamma^a, a \Rightarrow p \vdash a \not\Rightarrow q}{\Gamma^a \vdash p \not\Rightarrow q} \vdash \not\rightarrow$$

$$\frac{\Gamma \vdash a \Rightarrow p \quad \Delta, a \not\Rightarrow q \vdash A}{\Gamma, \Delta, p \not\Rightarrow q \vdash A} \not\rightarrow \vdash$$

$$\frac{\Gamma \vdash a \Rightarrow q \quad \Delta, a \not\Rightarrow p \vdash A}{\Gamma, \Delta, p \not\Rightarrow q \vdash A} \not\rightarrow \vdash'$$

In the rules $\vdash \Rightarrow$ and $\vdash \not\Rightarrow$, Γ^a is supposed to represent a collection of formulas not containing a . We do need both the rules $\Rightarrow \vdash$ and $\Rightarrow \vdash'$ to capture the meaning

NCE

ple. According to
contains — that is,
contained in Γ . In
le $a \Rightarrow s \Rightarrow r$ and
that $p = \text{starlings}$,
first of these paths
mal ($a \Rightarrow r$); the
logs ($p \not\Rightarrow s$). The
link $p \Rightarrow s$ is not
ven though all its



Γ_2

gh straightforward,
ks, each of which
gic. It may seem
or such a network
slate its links, and
e it belongs to the
ndard view, is due
of [14], consider,
slated into the set
t all belongs to its
' Γ_2 should be taken
e analysis of [14],
h from a to q , and

e analysis of strict
e is always free to
ical theory, so that
me value in taking
erive conclusions

of \Rightarrow on the left of the turnstile; neither will do alone. Likewise, both $\nLeftarrow \vdash$ and $\nLeftarrow \vdash'$ are necessary.

We provide here a sample proof, of the sequent $p \Rightarrow q, q \nLeftarrow r \vdash p \nLeftarrow r$, simply in order to illustrate these rules.

$$\frac{\frac{\frac{a \Rightarrow p \vdash a \Rightarrow p \quad a \Rightarrow q \vdash a \Rightarrow q}{a \Rightarrow p, p \Rightarrow q \vdash a \Rightarrow q} \rightarrow \vdash \quad a \nLeftarrow r \vdash a \nLeftarrow r}{p \Rightarrow q, q \nLeftarrow r, a \Rightarrow p \vdash a \nLeftarrow r} \nLeftarrow \vdash'}{p \Rightarrow q, q \nLeftarrow r \vdash p \nLeftarrow r} \vdash \nLeftarrow$$

It is shown in [8] that this logic is both sound and complete with respect to the notion of permission for strict nets: the sequent $\Gamma \vdash A$ is provable just in case Γ enables A .

The interpretation of this logic relies on the set $T = \{\{T\}, \{F\}, \emptyset, \{T, F\}\}$ as truth values. Following Belnap and Dunn, we identify these values with the four information states of a database with respect to a proposition: (i) the state of possessing evidence for the proposition, and no evidence to the contrary; (ii) the state of possessing evidence against the proposition, and no evidence to the contrary; (iii) the state of possessing no evidence either for or against the proposition; (iv) the state of possessing evidence both for the proposition and against it well. These explanations should suggest why it is natural to take the power set of $\{T, F\}$ as the set of truth values: if X is one of the values from T , ' $T \in X$ ' means that there is evidence for any proposition bearing the truth value X , and ' $F \in X$ ' means that there is evidence against such a proposition.

A valuation v on the language of strict links can be defined as follows. Relative to a domain D , the valuation assigns an individual $v(a)$ in D to each individual term a of the language, and a function $v(p)$ from D to T to each generic term p . Where v is a valuation, v^d_a is the valuation like v for all terms other than a , but which assigns the value d to a . The following rules extend v to the entire language.

- $v(a \Rightarrow p) = [v(p)](v(a))$.
- $v(a \nLeftarrow p) = \text{Not}(v(pa))$, where $\text{Not}(\{T\}) = \{F\}$, $\text{Not}(\{F\}) = \{T\}$, $\text{Not}(\emptyset) = \emptyset$, and $\text{Not}(\{T, F\}) = \{T, F\}$.
- $v(p \Rightarrow q) = \{T\}$ if for all $d \in D$, we have $T \in v^d_a(qa)$ if $T \in v^d_a(pa)$ and $F \in v^d_a(pa)$ if $F \in v^d_a(qa)$; and $v(p \Rightarrow q) = \emptyset$ otherwise.
- $v(p \nLeftarrow q) = \{T\}$ if for all $d \in D$, we have $F \in v^d_a(qa)$ if $T \in v^d_a(pa)$ and $F \in v^d_a(pa)$ if $T \in v^d_a(qa)$; and $v(p \nLeftarrow q) = \emptyset$ otherwise.

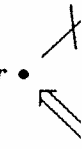
Given this interpretation, the notion of semantic implication is defined in the usual way: Γ *semantically implies* A just in case, for all valuations v , if $T \in v(B)$ for all

$B \in \Gamma$, then characterizes statement A , that the impl above: $\Gamma \vdash A$

3.2 Defeasibi

Defeasible in defeasible ne contain. The networks by inconsistent, can be localiz local inconsis conflicting pa its links repr $a = \text{Nixon}$, consistent, no both of the co would allow a pacifist ($a =$

Theories of conflicting conflicting pa tions of the co (subject to ce the theory of to defeasible with any give *skeptical* app arguments, re other. Applie conflicting pa



Fi

$B \in \Gamma$, then $T \in v(A)$. It is shown in [16] that this kind of four-valued implication characterizes the notion of strict inheritance, in the sense that a net Γ enables a statement A just in case Γ semantically implies A . From this it follows, of course, that the implication relation characterizes also the sequent calculus presented above: $\Gamma \vdash A$ is provable just in case Γ semantically implies A .

3.2 Defeasible inheritance

Defeasible inheritance is more complicated than strict inheritance, primarily because defeasible networks, unlike their strict counterparts, do not permit all the paths they contain. The distinction derives from the different role played in the two kinds of networks by conflicting paths. Any strict network containing conflicting paths is inconsistent, though — as the theory of [16] shows — the effects of the inconsistency can be localized. But defeasible networks can contain conflicting paths without even local inconsistency. Consider, for example, Γ_3 (Figure 3). Although this net contains conflicting paths, it is not inconsistent: it admits interpretations under which all of its links represent true statements — including the well-known interpretation with $a = \text{Nixon}$, $q = \text{Quakers}$, $r = \text{Republicans}$, and $p = \text{pacifists}$. Since the net is consistent, no reasonable theory of defeasible inheritance would say that it permits both of the conflicting paths $a \Rightarrow q \rightarrow p$ and $a \Rightarrow r \not\Rightarrow p$ at once. Any such theory would allow us to draw inconsistent conclusions — for example, that Nixon both is a pacifist ($a \Rightarrow p$) and that he isn't ($a \not\Rightarrow p$) — from consistent information.

Theories of defeasible inheritance differ among themselves in their treatment of conflicting paths. One kind of theory associates with each network containing conflicting paths a number of different extensions, corresponding to different resolutions of the conflicts. Because each extension supports a maximal set of conclusions (subject to certain constraints) we describe these theories as *credulous*; an example is the theory of Touretzky [15]. The present paper is based on an alternative approach to defeasible inheritance, developed in [16], which has the advantage of associating with any given network only a single extension. We describe this alternative as a *skeptical* approach — since it embodies the broadly skeptical idea that conflicting arguments, represented in networks by conflicting paths, tend to neutralize each other. Applied to Γ_3 , for example, the skeptical approach tells us that neither of the conflicting paths should be permitted.

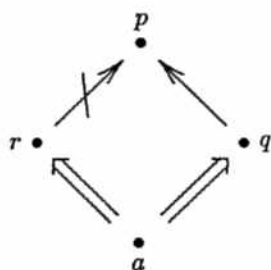


Figure 3: Γ_3

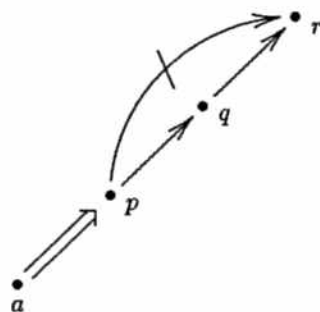


Figure 4: Γ_4

NCE
both $\not\Rightarrow \vdash$ and

$\vdash p \not\Rightarrow r$, simply

$\not\vdash \vdash'$
_____ $\vdash \not\vdash$

with respect to the
role just in case Γ

$\{F\}, \emptyset, \{T, F\}$
values with the
n: (i) the state of
contrary; (ii) the
to the contrary;
proposition; (iv)
ist it well. These
of $\{T, F\}$ as the
means that there is
 $\in X'$ means that

follows. Relative
h individual term
term p . Where v
out which assigns
ge.

$\not\vdash$, $Not(\{F\}) =$

$v_a^d(qa)$ if $T \in$
 $v(p \Rightarrow q) = \emptyset$

$v_a^d(qa)$ if $T \in$
 $v(p \not\Rightarrow q) = \emptyset$

defined in the usual
 $T \in v(B)$ for all

In fact, the theory of [8] is not thoroughly skeptical: its skepticism is restricted to compound paths, and even compound paths can be neutralized only by those conflicting paths that are not themselves, as we say, preempted. The first of these restrictions has the effect that, even in the face of conflicts, any non-compound path contained in a network — that is, any direct link — will be permitted by that network. As explained in [8], this principle is well-motivated, particularly against the background of the four-valued logic; but it is not a crucial feature of the theory.

The second restriction, however, is crucial. Preemption is the mechanism by which, in case of conflicts, arguments based on more specific information are allowed to override arguments based on less specific information. For example, consider Γ_4 (Figure 4) — with, say, $a = \text{Tweety}$, $p = \text{penguins}$, $q = \text{birds}$, and $r = \text{flying things}$. Since this net contains the two conflicting paths $a \Rightarrow p \rightarrow q \rightarrow r$ and $a \Rightarrow p \not\rightarrow r$, an unrestricted skepticism would permit neither. However, it seems in this case that the latter of these paths *should* be permitted, because it represents an argument based on more specific information. The second restriction above reflects this intuition. We say that a path of the form $\pi(x, \tau, v) \rightarrow y$ is *preempted* in a net Γ just in case there is a node z such that (i) Γ permits a path $\pi(x, \tau_1, z, \tau_2, v)$, so that z provides “more specific” information than v about x , and (ii) $z \not\rightarrow y \in \Gamma$, so that z gives us “direct” information contrary to that provided by v . (The definition of preemption for negative paths is symmetrical.) According to the theory of [8], even a conflicted path will be permitted if the only paths with which it conflicts are themselves preempted; so, for example, $a \Rightarrow p \not\rightarrow r$ will be permitted by Γ_4 , since $a \Rightarrow p \rightarrow q \rightarrow r$ is preempted.

3.3 Mixed inheritance

The theory of [14] tells us, then, that a strict network permits exactly the paths it contains; the theory of [8] tells us that a defeasible network permits a path it contains just in case that path is either a direct link, or any other path with which it conflicts is itself preempted. Now, to combine these two theories into an account of inheritance for mixed networks, we first carry over entirely the analysis of strict inheritance from [14], and then modify the analysis of defeasible inheritance from [8] to accommodate the presence of strict links. Since it incorporates the analysis of [14], the resulting theory tells us that a mixed network permits exactly the strict paths it contains. Likewise, since it is based on the analysis of [8], the resulting theory also embodies the skeptical idea that a compound defeasible path is neutralized by any conflicting path that is not itself preempted. However, in order to develop this idea in a mixed context, we need to modify slightly our conception of the kind of paths that represent conflicts, as well as our understanding of the preemption relation among conflicting paths.

In defeasible networks, all conflicts share a simple form: they involve paths with identical initial nodes, identical end nodes, and opposite polarity. But the presence of strict links introduces the possibility of less direct conflicts, even among defeasible paths. As an illustration, consider Γ_5 (Figure 5). Here it seems reasonable, in light of the strict segment $r \Rightarrow s \Rightarrow t$, to regard $p \rightarrow q \rightarrow r$ and $p \rightarrow u \rightarrow v \not\rightarrow t$ themselves as conflicting paths, even though they do not share an end node. Imagine,

for exampl
tells us tha
which repr
the conclu
represents
What
arguments,
a strict con
be classifie

$\kappa_\Gamma(x)$
 $\bar{\kappa}_\Gamma(x)$

so that $\kappa_\Gamma(x)$
attributed t
and the set
conception
of clashes,
with any p
path of the
regarding c
well. In Γ_5 ,
since each i



Just as the p
however, it
see this, con

for example, that $r = \text{dogs}$, $s = \text{mammals}$, and $t = \text{animals}$, so that the strict segment tells us that *all* dogs are animals. In the context of Γ_5 , then, the path $p \rightarrow q \rightarrow r$, which represents an argument to the effect that p 's are dogs, carries with equal force the conclusion that p 's are animals; so it conflicts with $p \rightarrow u \rightarrow v \not\rightarrow t$, which represents an argument that p 's are not animals.

What this example shows is that two defeasible paths can represent conflicting arguments, even if they have different end nodes, when one of the paths clashes with a strict consequence of the other. Of course, such strict consequences can themselves be classified as positive or negative. Let us define

$$\kappa_\Gamma(x) = \{x\} \cup \{y : \Gamma \text{ contains a strict positive path from } x \text{ to } y\},$$

$$\bar{\kappa}_\Gamma(x) = \{y : \Gamma \text{ contains a strict negative path from } x \text{ to } y\},$$

so that $\kappa_\Gamma(x)$ and $\bar{\kappa}_\Gamma(x)$ represent the positive and negative strict consequences attributed to x by Γ — the set of properties that x must possess, according to Γ , and the set of properties that x cannot possess. It is then natural to extend our conception of conflicting defeasible paths so that, in addition to the ordinary kinds of clashes, a path of the form $\pi(x, \sigma, u) \rightarrow y$ will be said to conflict in a net Γ with any path of the form $\pi(x, \tau, v) \not\rightarrow m$ where $m \in \kappa_\Gamma(y)$, and also with any path of the form $\pi(x, \tau, v) \rightarrow m$ where $m \in \bar{\kappa}_\Gamma(y)$. Our general skeptical attitude regarding conflicting paths will then have to apply to these new kinds of conflicts as well. In Γ_5 , for instance, neither $p \rightarrow q \rightarrow r$ nor $p \rightarrow u \rightarrow v \not\rightarrow t$ will be permitted, since each is neutralized by its conflict with the other.

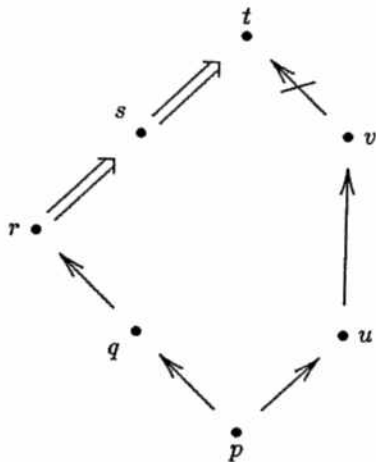


Figure 5: Γ_5

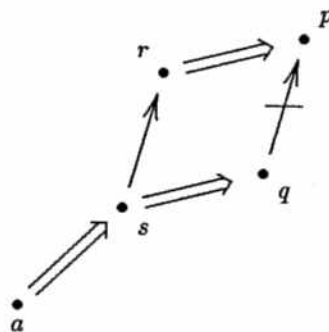


Figure 6: Γ_6

Just as the presence of strict links allows for the possibility of new kinds of conflicts, however, it provides also for the possibility of new relations of preemption. To see this, consider the network Γ_6 (Figure 6), supplied with an interpretation under

icism is restricted
ed only by those
The first of these
y non-compound
permitted by that
rticularly against
ure of the theory.
s the mechanism
: information are
n. For example,
= birds, and $r =$
 $p \rightarrow q \rightarrow r$ and
ever, it seems in
e it represents an
on above reflects
empted in a net
, τ_1, z, τ_2, v), so
i) $z \not\rightarrow y \in \Gamma$, so
(The definition
he theory of [8],
h it conflicts are
tted by Γ_4 , since

xactly the paths
ermits a path it
path with which
into an account
analysis of strict
heritance from
s the analysis of
y the strict paths
resulting theory
s neutralized by
to develop this
n of the kind of
emption relation

y involve paths
larity. But the
cts, even among
ems reasonable,
 $\rightarrow u \rightarrow v \not\rightarrow t$
node. Imagine,

which a = Hermann, p = persons born in America, q = native speakers of German, r = persons born in Pennsylvania, and s = native speakers of Pennsylvania Dutch. Under this interpretation, Γ_6 tells us that Hermann is a particular speaker of Pennsylvania Dutch, that every speaker of Pennsylvania Dutch speaks German (since Pennsylvania Dutch is a dialect of German), that German speakers tend not to be born in America, that speakers of Pennsylvania Dutch tend to be born in Pennsylvania, and that everyone born in Pennsylvania is born in America.

According to our new, extended conception, the paths $a \Rightarrow s \rightarrow r$ and $a \Rightarrow s \Rightarrow q \not\vdash p$ now represent conflicting arguments in the context of Γ_6 , since $p \in \kappa_{\Gamma_6}(r)$. Of course, we would not want to remain skeptical in this case. The path $a \Rightarrow s \Rightarrow q \not\vdash p$, representing the argument that Hermann was not born in America since he is a native speaker of German, should be preempted in Γ_6 : the fact that his dialect is Pennsylvania Dutch provides a more specific argument to the contrary. Without modification, however, the treatment of preemption from [8] does not give us this result. A path can be preempted only if there is more specific and direct information to the contrary; and, although s does provide "more specific" information than q , the path $s \rightarrow r \Rightarrow p$ does not represent "direct" information to the contrary — at least, not according to the standards of [8], which holds that direct information can be carried only by direct links.

Evidently, it is this last requirement concerning the nature of direct information that needs to be modified in the present context. In the context of defeasible networks, it makes good sense to say that direct information can be carried only by direct links: any compound path represents an argument that can itself be undermined. In the context of mixed nets, however, certain kinds of compound paths *can* legitimately be thought to carry direct information — namely, compound paths consisting of a single defeasible link followed by a strict end segment, of any length. In Γ_6 , for example, the path $s \rightarrow r \Rightarrow p$ should be thought of as telling us directly that speakers of Pennsylvania Dutch are born in America: for even by the standards of [8], $s \rightarrow r$ counts as a direct statement of the fact that speakers of Pennsylvania Dutch are born in Pennsylvania, and the strict extension $r \Rightarrow p$ simply tells us that *everyone* born in Pennsylvania is born in America.

Adjusting our definition of preemption to account for this new notion of direct information, we say now that a negative path $\pi(x, \tau, v) \not\vdash m$ is preempted in a mixed network Γ if there exist nodes z and n such that Γ permits a path $\pi(x, \tau_1, z, \tau_2, v)$ with $z \rightarrow n \in \Gamma$ and $m \in \kappa_{\Gamma}(n)$. This new definition allows us to conclude, as it should, that $a \Rightarrow s \Rightarrow q \not\vdash p$ is preempted in Γ_6 ; so the net does end up supporting the conclusion that Hermann was born in America. It is a bit more complicated to formulate mixed preemption for positive paths, although no new ideas are involved, simply because direct information to the contrary can now take the form either of a positive defeasible link followed by a negative strict extension, or of a negative defeasible link followed by a reverse positive strict extension. Formally, we say that a positive path $\pi(x, \tau, v) \rightarrow m$ is preempted in a mixed network Γ if there exist nodes z and n such that Γ permits a path $\pi(x, \tau_1, z, \tau_2, v)$ with either (i) $z \rightarrow n \in \Gamma$ and $m \in \bar{\kappa}_{\Gamma}(n)$ or (ii) $z \not\vdash n \in \Gamma$ and $n \in \kappa_{\Gamma}(m)$.

In this section
permission
permission

4.1. Mixed

Like that of
measure of
whether $\Gamma \vdash$
than σ itself

In order
through a n
links contain
that a gener
it can contain
 $p \not\vdash q \Leftarrow r$
segment p
length of a
strict initial
links in the
its defeasible
(Example:)
since it contains
path $p \Rightarrow q$
along with t

Using t
 $\text{deg}_{\Gamma}(\sigma)$ —
from the ini
the acyclic ξ
 $p \rightarrow u \rightarrow v$
the assignm
we need to r
kinds of def
at least one
severely the
contain only

The no
notion defin
complexity
context, the
Basically, w
degree, whe
the mixed d
the pair tells
us, simply, v

4. DEFINING MIXED INHERITANCE

In this section, we assemble our motivational ideas into a rigorous definition of the permission relation for mixed networks; we use the symbol ' \vdash ' to stand for the permission relation, so that ' $\Gamma \vdash \sigma$ ' means that the net Γ permits the path σ .

4.1. Mixed degree

Like that of [8], the present definition is inductive. Our first step, then, is to assign a measure of "complexity" to each path σ in a net Γ in such a way that it can be decided whether $\Gamma \vdash \sigma$ once it is known whether $\Gamma \vdash \sigma'$ for each path σ' less complex in Γ than σ itself.

In order to arrive at the appropriate notion of path complexity, we proceed through a number of auxiliary ideas. As we recall, a path is a joined sequence of links containing a negative segment, if at all, only at the very end. Let us say, then, that a *generalized path* is a sequence of links joined like an ordinary path, except that it can contain negative segments anywhere, and perhaps more than one. (Example: $p \nrightarrow q \leftarrow r \nrightarrow s \leftarrow t$ is a generalized path, but it is not a path, since its negative segment $p \nrightarrow q \leftarrow r$ is not an end segment.) Next, let us define the *defeasible length* of a generalized path as follows: if the generalized path does not contain a strict initial segment, then its defeasible length is simply the number of defeasible links in the path; if the generalized path does contain a strict initial segment, then its defeasible length is the number of defeasible links in the path augmented by one. (Example: the generalized path $r \rightarrow s \Rightarrow t \rightarrow u$ has a defeasible length of two, since it contains two defeasible links and no strict initial segment; the generalized path $p \Rightarrow q \Rightarrow r \rightarrow s \Rightarrow t \rightarrow u$ is three, since it contains a strict initial segment along with two defeasible links.)

Using these ideas, we can now define the *degree* of a path σ in a net Γ — written, $deg_{\Gamma}(\sigma)$ — as the greatest defeasible length of any acyclic generalized path in Γ from the initial node of σ to its end node. (Example: $deg_5(p \rightarrow q \rightarrow r) = 3$, since the acyclic generalized path from p to r in Γ_5 whose defeasible length is greatest is $p \rightarrow u \rightarrow v \nrightarrow t \leftarrow s \leftarrow r$, with a defeasible length of 3.) In order to insure that the assignment of degree to the paths in a network has the appropriate properties, we need to restrict the application of our theory, as in [8], to paths free from certain kinds of defeasible cycles (a defeasible cycle is a cyclic generalized path containing at least one defeasible link); for the present, we limit will our attention, even more severely than necessary, to networks which are either entirely acyclic, or which contain only strict cycles.

The notion of degree defined here is a straightforward generalization of the notion defined in [8]. However, it is not quite appropriate as a measure of path complexity for an inductive definition of the permission relation; in the present context, the measure of complexity needs to carry just a bit more information. Basically, we want our measure of a path's complexity to tell us, in addition its degree, whether or not the path possesses a strict end segment. Therefore, we define the *mixed degree* of a path σ in a net Γ as a pair $\langle n, v \rangle$. The first component of the pair tells us the degree of σ in Γ : $n = deg_{\Gamma}(\sigma)$. The second component tells us, simply, whether or not σ possesses a strict end segment: by convention, we let

$v = 0$ if σ does not possess a strict end segment, and $v = 1$ if it does. We define a lexical ordering on the mixed degrees by giving priority to the first component: $\langle n, v \rangle < \langle n', v' \rangle$ iff either $n < n'$ or $n = n'$ and $v < v'$. The idea behind this ordering is that degree is the primary measure of path complexity — but of two paths identical in degree, one with and one without a strict end segment, the path lacking the strict end segment is classified as less complex.

4.2. The definition

Our definition of the permission relation has the overall structure of a definition by cases. Any path σ from a mixed network can be divided into the subpaths $\mu(\sigma)$ and $\delta(\sigma)$, where $\mu(\sigma)$ is the maximal strict end segment of σ , and $\delta(\sigma)$ is the result of truncating $\mu(\sigma)$ from σ . (Example: if σ is $x \Rightarrow y \rightarrow p \not\Leftarrow r \Leftarrow s$, then $\mu(\sigma)$ is $p \not\Leftarrow r \Leftarrow s$ and $\delta(\sigma)$ is $x \Rightarrow y \rightarrow p$.) Using this notation, then, we specify the conditions under which $\Gamma \vdash \sigma$ in three separate cases, depending on the form of σ . Our first case deals with defeasible paths possessing strict end segments.

Case A: $\sigma \neq \delta(\sigma)$ and $\sigma \neq \mu(\sigma)$. Then $\Gamma \vdash \sigma$ iff $\Gamma \vdash \delta(\sigma)$ and $\Gamma \vdash \mu(\sigma)$.

The next case deals with strict paths.

Case B: $\sigma = \mu(\sigma)$. Then $\Gamma \vdash \sigma$ iff each link in σ is contained in Γ .

Finally, we deal with the case of paths ending in defeasible links — which itself divides into subcases, as such paths may be simple or compound.

Case C-I: $\sigma = \delta(\sigma)$ and σ is a direct link. Then $\Gamma \vdash \sigma$ iff $\sigma \in \Gamma$.

Case C-II: $\sigma = \delta(\sigma)$ and σ is a compound path. Two subcases to consider.

1. σ is a positive path, of the form $\pi(x, \sigma_1, u) \rightarrow y$. Then $\Gamma \vdash \sigma$ iff
 - (a) $\Gamma \vdash \pi(x, \sigma_1, u)$;
 - (b) $u \rightarrow y \in \Gamma$;
 - (c) For $m \in \kappa_\Gamma(y)$, $x \not\rightarrow m \notin \Gamma$ and $m \notin \bar{\kappa}_\Gamma(x)$;
 - (d) For $m \in \bar{\kappa}_\Gamma(y)$, $x \rightarrow m \notin \Gamma$ and $m \notin \kappa_\Gamma(x)$;
 - (e) For all v, m, τ such that $\Gamma \vdash \pi(x, \tau, v)$ with $v \not\rightarrow m \in \Gamma$ and $m \in \kappa_\Gamma(y)$, there exist z, n, τ_1, τ_2 such that either (i) $z = x$ or (ii) $\Gamma \vdash \pi(x, \tau_1, z, \tau_2, v)$, with (iii) $z \rightarrow n \in \Gamma$ and $m \in \kappa_\Gamma(n)$;
 - (f) For all v, m, τ such that $\Gamma \vdash \pi(x, \tau, v)$ with $v \rightarrow m \in \Gamma$ and $m \in \bar{\kappa}_\Gamma(y)$, there exist z, n, τ_1, τ_2 such that either (i) $z = x$ or (ii) $\Gamma \vdash \pi(x, \tau_1, z, \tau_2, v)$, with either (iii) $z \rightarrow n \in \Gamma$ and $m \in \bar{\kappa}_\Gamma(n)$ or (iv) $z \not\rightarrow n \in \Gamma$ and $n \in \kappa_\Gamma(m)$.
2. σ is a negative path, of the form $\pi(x, \sigma_1, u) \not\rightarrow y$. Then $\Gamma \vdash \sigma$ iff
 - (a) $\Gamma \vdash \pi(x, \sigma_1, u)$;
 - (b) $u \not\rightarrow y \in \Gamma$;
 - (c) For m such that $y \in \kappa_\Gamma(m)$, $x \rightarrow m \notin \Gamma$ and $m \notin \kappa_\Gamma(x)$;
 - (d) For all v, m, τ , such that $\Gamma \vdash \pi(x, \tau, v)$ with $v \rightarrow m \in \Gamma$ and $y \in \kappa_\Gamma(m)$, there exist z, n, τ_1, τ_2 such that either (i) $z = x$ or (ii) $\Gamma \vdash \pi(x, \tau_1, z, \tau_2, v)$, with either (iii) $z \rightarrow n \in \Gamma$ and $m \in \bar{\kappa}_\Gamma(n)$ or (iv) $z \not\rightarrow n \in \Gamma$ and $n \in \kappa_\Gamma(m)$.

It sho
cases, is p
path σ of r
the path μ
are basis c
and $\langle 1, 0 \rangle$.
with $n >$
or 1, but n'

It is e
the treatme
treatment c
this fact is
of any path
defined in [
If Γ is pure
permission
[8]. Since
and $\bar{\kappa}_\Gamma(x)$
disappear (I
C-II.2 are I
[8].

By combin
of defeasibl
attractive th
defeasible I
satisfy the c
this point, th
concerns th
such as Brac
man." In o
taxonomic r
for the expre
such a theory
concepts wil

This materia
Grant No. IR
03-88-K008

It should be clear that this definition, although structured as a definition by cases, is properly an induction on mixed degree. Case A defines permission for a path σ of mixed degree $\langle n, 1 \rangle$ in terms of the path $\delta(\sigma)$ of mixed degree $\langle n, 0 \rangle$ and the path $\mu(\sigma)$ of mixed degree $\langle 1, 1 \rangle$ — both inductively simpler. Cases B and C-I are basis cases, defining permission respectively for paths of mixed degree $\langle 1, 1 \rangle$ and $\langle 1, 0 \rangle$. Finally, Case C-II defines permission for paths of mixed degree $\langle n, 0 \rangle$ with $n > 1$ in terms of paths of mixed degree $\langle n', v' \rangle$ — where v' may be either 0 or 1, but $n' < n$ so that the overall measure of mixed degree is simpler.

It is evident from Case B that our treatment of mixed inheritance agrees with the treatment of [14] when it is applied to strict networks. It agrees also with the treatment of [8] when it is applied to purely defeasible networks — and the proof of this fact is reassuringly simple. If Γ is a defeasible network, then the mixed degree of any path in Γ as defined here is simply $\langle n, 0 \rangle$, where n is its ordinary degree as defined in [8]; so the two inductive definitions move through paths in the same order. If Γ is purely defeasible, then only the Cases C-I and C-II are used in determining permission. Case C-I is identical to the corresponding clause in the definition from [8]. Since Γ is defeasible, it turns out that for every node x we have $\kappa_{\Gamma}(x) = \{x\}$ and $\bar{\kappa}_{\Gamma}(x) = \emptyset$. Under these conditions, the extra clauses (d) and (f) in Case C-II.1 disappear (become logical truths), and the remaining clauses in both Case C-II.1 and C-II.2 are logically equivalent to the corresponding clauses in the definition from [8].

5. CONCLUSION

By combining the analysis of strict inheritance from [14] with the skeptical analysis of defeasible inheritance from [8], we have developed a well-defined and intuitively attractive theory of inheritance for semantic networks containing both strict and defeasible links. Although these matters are not discussed here, this theory does satisfy the crucial properties of *soundness* and *atomic stability* defined in [8]. At this point, the topic that stands out as the most important area for further research concerns the application of this work to the representation of complex concepts, such as Brachman and Israel's "three-legged elephant" or the traditional "unmarried man." In order to represent such concepts along with defeasible information in a taxonomic reasoner, it is necessary, first, to develop a theory of inheritance allowing for the expression of both strict and defeasible relations. The present paper presents such a theory — but it does not go on to address the problem of dealing with complex concepts within this framework.

ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under Grant No. IRI-8700705, and by the Army Research Center under Grant No. DAAL-03-88-K0087.

The account of mixed inheritance described here was worked out in conjunction with Richmond Thomason, who has developed an equivalent approach to the same problem; it was first presented in [10]. I am very grateful to Thomason, and also to David Touretzky, both for sparking my interest in the topic of inheritance, and for many illuminating discussions.

This paper is dedicated with respect and affection to Nuel Belnap.

- [1] Belnap, N. *Philosophical Studies*, 71 (1982), pp. 121-136.
- [2] Belnap, N. *Journal of Philosophy*, 79 (1982), pp. 1-16.
- [3] Brachman, R. *Artificial Intelligence*, 17 (1982), pp. 179-190.
- [4] Brachman, R. *Artificial Intelligence*, 17 (1982), pp. 191-203.
- [5] Brachman, R. *Artificial Intelligence*, 17 (1982), pp. 204-216.
- [6] Dunn, J. M. *Philosophical Studies*, 71 (1982), pp. 137-148.
- [7] Fahlman, S. *Artificial Intelligence*, 17 (1982), pp. 217-231.
- [8] Gentzen, G. *Journal of Philosophy*, 79 (1982), pp. 17-39.
- [9] Hayes, P. (ed.) *Artificial Intelligence and Representation*. New York: Academic Press, 1985.
- [10] Horty, J. and Thomason, R. *AAAI-88, Proceedings of the AAAI Conference on Artificial Intelligence*, 1988, pp. 101-110.
- [11] Horty, J. *Artificial Intelligence*, 17 (1982), pp. 232-246.
- [12] Israel, D. *Artificial Intelligence*, 17 (1982), pp. 247-261.
- [13] Minsky, M. *Artificial Intelligence*, 17 (1982), pp. 262-276.
- [14] Roberts, R. *Artificial Intelligence*, 17 (1982), pp. 277-291.
- [15] Sandewall, E. *Artificial Intelligence*, 17 (1982), pp. 292-306.
- [16] Thomason, R. *Artificial Intelligence*, 17 (1982), pp. 307-321.
- [17] Touretzky, D. *Artificial Intelligence*, 17 (1982), pp. 322-336.
- [18] Touretzky, D. *Artificial Intelligence*, 17 (1982), pp. 337-351.

BIBLIOGRAPHY

- [1] Belnap, N. (1977): How a computer should think. In G. Ryle (ed.), *Contemporary Aspects of Philosophy*. Oriel Press, pp. 30-56.
- [2] Belnap, N. (1977): A useful four-valued logic. In J. Dunn and G. Epstein (eds.), *Modern Uses of Multiple-valued Logic*. D. Reidel, pp. 8-37.
- [3] Brachman, R. (1983): What IS-A is and isn't: an analysis of taxonomic links in semantic networks. *Computer*, vol. 16 (October, 1983), pp. 67-73.
- [4] Brachman, R. (1985): "I lied about the trees" or, defaults and definitions in knowledge representation. *The AI Magazine*, vol. 6 (Fall, 1985), pp. 80-93.
- [5] Brachman, R. and Schmolze, J. (1985): An overview of the KL-ONE knowledge representation system. *Cognitive Science*, vol. 9, pp. 171-216.
- [6] Dunn, J. M. (1976): Intuitive semantics for first degree entailments and "coupled trees." *Philosophical Studies*, vol. 29.
- [7] Fahlman, S. (1979): *NETL: a System for Representing and Using Real-world Knowledge*. The MIT Press.
- [8] Gentzen, G. (1934): Untersuchungen über das Logische Schliessen. *Mathematische Zeitschrift* 39, pp. 176-210, 405-431. Translated in M. Szabo (ed.), *The collected papers of Gerhard Gentzen*, North-Holland, 1969, pp. 68-131.
- [9] Hayes, P. (1979): The logic of frames. In *Frame Conceptions and Text Understanding*, D. Metzinger (ed.), Walter de Gruyter and Co., pp. 46-61. Reprinted in *Readings in Knowledge Representation*, R. Brachman and H. Levesque (eds.), Morgan Kaufmann (1985), pp. 287-297.
- [10] Horty, J. and Thomason, R. (1988): Mixing strict and defeasible inheritance. In *Proceedings of AAAI-88*, Morgan Kaufmann.
- [11] Horty, J., Thomason, R. and Touretzky, D. (1990): A skeptical theory of inheritance in non-monotonic semantic networks. In *Artificial Intelligence*, vol. 42, pp. 311-348. A preliminary version appears in *Proceedings of AAAI-87*, Morgan Kaufmann, pp. 358-363.
- [12] Israel, D. and Brachman, R. (1984): Some remarks on the semantics of representation languages. In *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*, M. Brodie, J. Mylopoulos, and J. Scholze (eds.), Springer-Verlag.
- [13] Minsky, M. (1974): A framework for representing knowledge. MIT Artificial Intelligence Laboratory Memo No. 306. Reprinted without appendix in *The Psychology of Computer Vision*, P. Winston (ed.), McGraw-Hill (1975), pp. 211-277.
- [14] Roberts, R. and Goldstein, I. (1977): *The FRL Manual*. AI Memo No. 409, MIT Artificial Intelligence Laboratory.
- [15] Sandewall, E. (1986): Non-monotonic inference rules for multiple inheritance with exceptions. *Proceedings of the IEEE*, vol. 74, pp. 1345-1353.
- [16] Thomason, R., Horty, J., and Touretzky, D. (1986): A calculus for inheritance in monotonic semantic nets. Technical Report CMU-CS-86-138, Computer Science Department, Carnegie Mellon University, 24 pp.
- [17] Touretzky, D. (1986): *The Mathematics of Inheritance Systems*. Morgan Kaufmann.
- [18] Touretzky, D., Horty, J., and Thomason, R. (1987): A clash of intuitions: the current state of nonmonotonic multiple inheritance systems. In *Proceedings of IJCAI-87*, Morgan Kaufmann, pp. 476-482.