Nonmonotonic Inheritance and Generic Reflexives

David S. Touretzky
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213

Richmond H. Thomason
Intelligent Systems Program
University of Pittsburgh
Pittsburgh, PA 15260

Abstract. Generic reflexive statements such as *Elephants love themselves* have traditionally been formalized using some variant of predicate logic, with variables to mark coreferentiality. We present a radically different semantics for reflexives, based on nonmonotonic inheritance and an extension to Touretzky's inferential distance ordering. Our system can derive new generic reflexive statements as well as statements about individuals. And unlike the leading predicate logic-based approaches, our formalism does not use variables; this brings it closer in structure to actual human languages. The significance of this work for AI is its demonstration of the benefits of a non-classical knowledge representation for analyzing commonsense reasoning phenomena.

1 Motivation

Reflexive constructions are common in the world's languages. Contemporary linguistic theories subsume reflexivization under *anaphora,* treating these constructions (along with expressions like *each other*) as expressions that have no independent meaning, but are bound in some way to other expressions, thereby contributing to larger units that are meaningful.

The leading current linguistic theories of the semantics of reflexives[1] use variables to interpret reflexives, so—since no human language uses variables at surface level—variables or indices marking "co-referentiality" must be introduced in the course of parsing a sentence. Though theories differ on how these variables are introduced, they agree in producing "logical forms" that contain them.

Human languages that mark reflexive, however, generally do so either with special pronominal forms like the English *themselves,* or in the verb morphology, as in English *This watch is self-winding.*

It is hard to say whether the discrepancy between the logical form and the way in which human languages encode reflexives is a deep linguistic discovery or an artifact of our only having one semantic theory of reflexivization—a logical theory that was originally designed to explain mathematical notations rather than natural language. For purposes of comparison, it would be useful to have alternative semantic theories of reflexive. One such alternative is presented here.

The need for alternative theories is intensified by the limited ability of first-order logic to cope with the phenomena of natural language. Sentences like *Most politicians are honest* can't be formalized using forms such as

$$(Many\ x)(Politician\ x \rightarrow honest\ x)$$

and such difficulties have led to a theory of generalized quantifiers.[2]

However, relatively conservative extensions such as a *Many* quantifier are totally inadequate for handling what linguists call *generic plural,* and so can't deal with sentences like *Elephants are gray.*[3]

Nonmonotonic semantic networks can't be used as they stand as an alternative to logical formalisms in interpreting natural language because they are so limited in expressive power. This paper doesn't offer a solution to the general problem, but we do show that we can account for systematic interactions between generics and reflexives, using techniques from nonmonotonic inheritance theory. This suggests an alternative representation according to which reflexives—though still anaphoric because they must be bound to a relation by appearing in a path containing a single relational symbol R—seem to resemble individual nodes in many ways, and to have a greater measure of semantic independence than the variables of logical representations.

2 Structure of the Paper

Figures 1 and 2 contain all the network primitives that appear in the paper. Our graphical notation is a variant of NETL [Fahlman 1979]. There are several kinds of nodes, representing individuals (Clyde), classes (elephant), instances of the term "self" (denoted by $\odot$), and instances of "other" (denoted by $\oslash$). There are also several kinds of links. IS-A and IS-NOT-A links ($\rightarrow$ and $\not\rightarrow$) express taxonomic information, such as that elephants are gray and Clyde is not a herbivore. Positive and negative relational links ($\stackrel{R}{\rightarrow}$ and $\stackrel{R}{\not\rightarrow}$) represent binary relations between

[1] See [Thomason 1976] and [Thomason 1983] for the Montague Grammar approach, and [May 1985] for the Government-Binding approach.

[2] For example, the work of Altham [1971], and van Benthem & ter Meulen [1985].

[3] See [Carlson 1982] for detailed arguments.

classes or individuals, such as *Herbivores like gray things.* The asertion in Figure 1 that royal elephants do not like themselves is expressed by a $\overset{R}{\not\rightarrow}$ link from the royal elephant node to a $\odot$ (read "self") node. A fifth type of link, drawn as a dashed line in network notation, connects an "other" node to its referent. For example, Figure 2 indicates that opera stars admire other celebrities (i.e., celebrities other than themselves.) In path notation the term "other celebrities" is written $\oslash{:}c$.

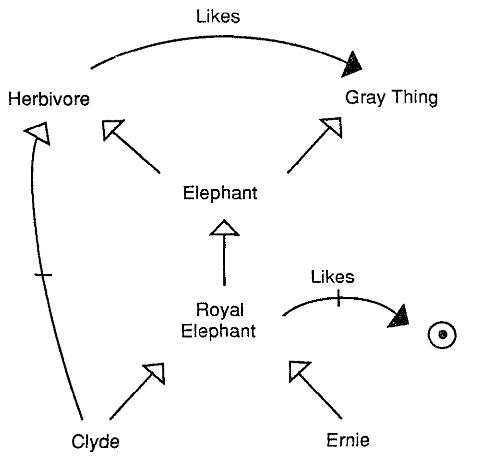

Figure 1: Herbivores like gray things, but royal elephants don't like themselves.

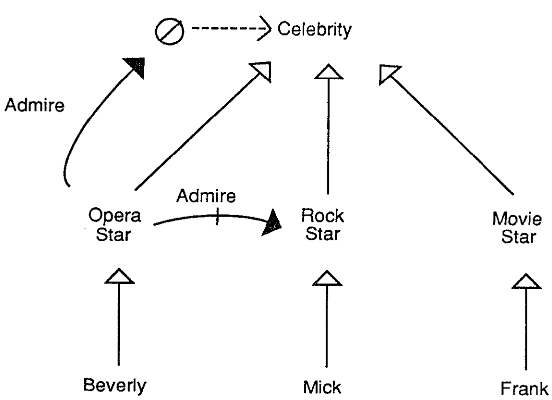

Figure 2: Opera stars admire other celebrities, but they don't admire rock stars.

The next few sections introduce a notation for inheritance paths and extensions (nonmonotonic theories), followed by axioms for nonmonotonic multiple inheritance with relations. We will then extend this system to handle reflexive and irreflexive statements. The paper concludes with an evaluation of the inheritance-based approach to reflexives, and some linguistic observations.

3 Notation

Let $\rightarrow$ and $\not\rightarrow$ denote nonmonotonic IS-A and IS-NOT-A links. Let $\overset{R}{\rightarrow}$ and $\overset{R}{\not\rightarrow}$ denote nonmonotonic positive and negative relational links. A network Γ is a collection of these four types of links, plus the links that bind "other" nodes to their referents.

Taxonomic paths are sequences of abutting IS-A and/or IS-NOT-A links, such as $x_1 \rightarrow x_2 \rightarrow x_3 \not\rightarrow x_4$. Positive paths are composed purely of IS-A links, while negative paths include an IS-NOT-A link at the end. Taxonomic paths contain only individual and class nodes; they contain no instances of $\odot$ or $\oslash$.

Lowercase Greek letters such as σ and τ will range over taxonomic paths, or, in the degenerate case, single individual or class nodes, or the null path. If σ is null then $x \rightarrow \sigma \rightarrow y$ should be read as $x \rightarrow y$, and $x \rightarrow \sigma \not\rightarrow y$ should be read as $x \not\rightarrow y$.

The notation $x_1 \rightarrow \ldots \rightarrow x_n$ refers to a path of length n whose ith element is x_i. Other occurrences of subscripted variables do not imply a continuous chain of x_i's; for example, the path $x_1 \rightarrow \sigma \rightarrow x_n$ denotes a path whose first and last elements are x_1 and x_n; it is not necessarily the case that the subsequence σ has any nodes in common with $x_2 \rightarrow \ldots \rightarrow x_{n-1}$.

We define $\tau = y_1 \rightarrow \ldots \rightarrow y_m$ and $\bar{\tau} = y_m \leftarrow \ldots \leftarrow y_1$ to be the forward and backward notations, respectively, for the same path.

Relational paths are paths of form $\sigma \overset{R}{\rightarrow} \bar{\tau}$ or $\sigma \overset{R}{\not\rightarrow} \bar{\tau}$. They are often written in expanded form as $x_1 \rightarrow \ldots \rightarrow x_n \overset{R}{\rightarrow} y_m \leftarrow \ldots \leftarrow y_1$, where the $\overset{R}{\rightarrow}$ link may be replaced by $\overset{R}{\not\rightarrow}$. The components σ and τ must be positive taxonomic paths, and $x_n \overset{R}{\rightarrow} y_m$ (or $x_n \overset{R}{\not\rightarrow} y_m$) must be in Γ. Figure 1 generates the relational path $E \rightarrow r \rightarrow e \rightarrow h \overset{R}{\rightarrow} g \leftarrow e \leftarrow r \leftarrow C$, which says that Ernie is a royal elephant, royal elephants are elephants, elephants are herbivores, herbivores like gray things, elephants are gray, royal elephants are elephants (again), and Clyde is a royal elephant, so *Ernie likes Clyde.*

The definition of *inheritability* governs the way paths may be extended to form new paths. A set of paths Φ is *perfect* iff every one of its elements is inheritable in Φ and no path not in Φ is inheritable in Φ. We define the *extensions* of a network Γ to be the minimal perfect supersets of Γ.

Finally, if κ is a path, then ButFirst(κ) is κ with the first link removed, and ButLast(κ) is κ with the last link removed.

4 Taxonomic Inheritance

We now present a nonmonotonic multiple inheritance system for classes and individuals. The definition is similar to that of [Touretzky, 1986]. In the terminology of [Touretzky *et al.* 1987], the system is a *credulous, downward reasoner*, with *coupled* extensions. But it uses *off-*

path preemption, which Sandewall [1986] has proposed as an improvement on Touretzky's original definition. We state only the axioms for inheritability of positive paths; the negative path axioms can be derived from these by appropriate substitution of "$\rightarrow$" and "$\not\rightarrow$" links.

Basis step: $x \rightarrow y$ is inheritable in Φ iff $x \rightarrow y \in \Phi$.

Induction: The path $\kappa = x_1 \rightarrow \ldots \rightarrow x_n$ is inheritable in Φ iff:

T1. ButFirst(κ) $\in \Phi$.

T2. ButLast(κ) $\in \Phi$.

T3. There is no path $x_1 \rightarrow \tau \not\rightarrow x_n \in \Phi$. (Contradiction)

T4. There is no w such that $x_1 \rightarrow \tau_1 \rightarrow w \rightarrow \tau_2 \rightarrow x_{n-1} \in \Phi$ and $w \not\rightarrow x_n \in \Phi$, where τ_1 or τ_2 may be null, and w may equal x_1 or x_{n-1}, respectively. (Preemption)

The notions of contradiction and preemption are the heart of the nonmonotonic inheritance definition. Contradiction keeps paths with conflicting conclusions from both being present in the same extension, as in Reiter's classic Nixon/pacifist example. Preemption is what allows subclasses to override the properties they would inherit from superclasses, even in the presence of redundant links. See [Touretzky *et al.* 1987] and [Horty *et al.* 1987] for more details and examples.

5 Relational Inheritance

We next present inheritance axioms for binary relations. Again, this definition is similar to [Touretzky, 1986], except for the use of off-path preemption. David Etherington, who with Ray Reiter was the first to translate taxonomic inheritance into default logic [Etherington 1987a], recently produced a version of this system in default logic as well [Etherington 1987b].

Basis step: $x \xrightarrow{R} y$ is inheritable in Φ iff $x \xrightarrow{R} y \in \Phi$.

Induction step: The path $\kappa = x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} y_m \leftarrow \ldots \leftarrow y_1$ is inheritable in Φ iff:

R1. ButFirst(κ) $\in \Phi$.

R2. ButLast(κ) $\in \Phi$.

R3. There is no path $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau} \leftarrow y_1 \in \Phi$. (Contradiction)

R4. There are no w, w' such that $x_1 \rightarrow \tau_1 \rightarrow w \rightarrow \tau_2 \rightarrow x_n \in \Phi$, $y_1 \rightarrow \tau_1' \rightarrow w' \rightarrow \tau_2' \rightarrow y_m \in \Phi$, and $w \overset{R}{\not\rightarrow} w' \in \Phi$, where τ_1 or τ_2 may be null and w may equal x_1 or x_n, respectively, and similarly τ_1' or τ_2' may be null and w' may equal y_1 or y_m, respectively. (Preemption)

An example of contradictory relational paths is: *Fred likes animals but Fred dislikes gray things.* In one extension Fred will like elephants because they're animals; in the other he won't like them because they're gray. These two reasoning paths can never appear in the same extension because they contradict each other.

An example of preemption of a relational path is *Citizens dislike crooks, but gullible citizens don't dislike elected crooks.* If Fred is a gullible citizen and Dick an elected crook, there is only one extension, and in it Fred does not dislike Dick.

6 Reflexive Statements

We are now ready to introduce reflexive statements. Readers who are not yet comfortable with the preceding definitions are advised to skim this and the following section the first time through, proceed to the discussion section, and then return here to study the definitions in greater detail.

There will be two kinds of reflexive paths in our system. Explicit reflexive paths are derived from statements that mention "self" directly, such as *John is a philosopher, and philosophers confuse themselves.* These paths are of form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} \odot$. Implicit reflexive paths, on the other hand, are derived from ordinary relational paths that double back on themselves. For example, in Figure 1, since elephants are herbivores and are gray, from *Herbivores like gray things* we can derive *Elephants like elephants*, and hence *Elephants like themselves.* The latter conclusion would be written $e \rightarrow h \xrightarrow{R} g \leftarrow e \leftarrow \odot$. Implicit reflexive paths take the general form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} y_m \leftarrow \ldots \leftarrow y_1 \leftarrow \odot$, where the doubling back means $y_1 = x_i$ for some i, $1 \leq i \leq n$.

6.1 Implicit Reflexive Paths

The following rule creates a new implicit reflexive path which can be inherited by lower nodes:

Let κ be a relational path of form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} y_m \leftarrow \ldots \leftarrow y_1$ where $x_1 = y_1$. Then the implicit reflexive path $\kappa \leftarrow \odot$ is inheritable in Φ iff:

SR1. ButFirst(κ) $\in \Phi$.

SR2. ButLast(κ) $\in \Phi$.

SR3. There is no path $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau} \leftarrow y_1 \in \Phi$. (Contradiction)

SR4. There is no path $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau} \leftarrow \odot \in \Phi$. (Contradiction)

Notice that the rule does not require κ itself to be present in Φ. The reason is that κ can be preempted by an "other" statement. For example, given *Parrots like green things, Amazon parrots are parrots and are green,* and *Amazon parrots don't like other Amazon parrots,* we can't infer *Amazon parrots like Amazon parrots,* but we can still infer *Amazon parrots like themselves.*

6.2 Inheritance of Reflexive Paths

Let κ be a reflexive path, i.e. a path of form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} y_m \leftarrow \ldots \leftarrow y_1 \leftarrow \odot$. If $m = 0$ then κ is an explicit reflexive path; otherwise it is an implicit path. The rule for inheritability of reflexive paths appears below. To allow a reflexive path to be preempted by an ordinary relational link, we require the head and tail nodes of the preempting relational link to be on the same path $x_1 \rightarrow \sigma \rightarrow x_n \in \Phi$. This is reflected in clause S5 below. See [Touretzky & Thomason, *forthcoming*] for an explanation of why this is necessary.

Basis step: $x \xrightarrow{R} \odot$ is inheritable in Φ iff $x \xrightarrow{R} \odot \in \Phi$.

Induction step: The path $\kappa = x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} y_m \leftarrow \ldots \leftarrow y_1 \leftarrow \odot$ (with m possibly 0, in which case there is a direct link $x_n \xrightarrow{R} \odot$; otherwise, with $x_1 \neq y_1$) is inheritable in Φ iff:

S1. ButFirst(κ) $\in \Phi$.

S2. ButLast(κ) $\in \Phi$.

S3. There is no $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau} \leftarrow \odot \in \Phi$. (Contradiction)

S4. There is no w such that $x_1 \rightarrow \tau_1 \rightarrow w \rightarrow \tau_2 \rightarrow x_n \in \Phi$, $x_1 \rightarrow \tau_1' \rightarrow w \rightarrow \tau_2' \rightarrow y_m \in \Phi$ (or $m = 0$), and $w \overset{R}{\not\rightarrow} \odot \in \Phi$. (Preemption by explicit reflexive statement.)

S5. There are no w_1, w_2 such that $x_1 \rightarrow \tau_1 \rightarrow w_1 \rightarrow \tau_2 \rightarrow w_2 \rightarrow \tau_3 \rightarrow x_n \in \Phi$, $x_1 \rightarrow \tau_1' \rightarrow w_1 \rightarrow \tau_2' \rightarrow w_2 \rightarrow \tau_3' \rightarrow y_m \in \Phi$ (or $m = 0$), and either $w_1 \overset{R}{\not\rightarrow} w_2 \in \Phi$ or $w_2 \overset{R}{\not\rightarrow} w_1 \in \Phi$. (Preemption by more specific ordinary relation.)

6.3 Statements About Individuals

In order to make the individual a's statements about "self" agree with its statements about a, we add the following axiom. Note that it is an implication, not an equivalence:

SI. If $a \rightarrow \sigma \xrightarrow{R} \bar{\tau} \leftarrow \odot \in \Phi$, then $a \rightarrow \sigma \xrightarrow{R} \bar{\tau} \leftarrow a$ is inheritable in Φ.

In Figure 1, this axiom derives *Clyde does not like Clyde* from the inherited path *Clyde does not like himself.*

6.4 Modification to Ordinary Relations

A reflexive statement should block inheritance of a contradictory ordinary relation. Thus, if *Herbivores like gray things* but *Royal elephants do not like themselves*, we should not infer *Royal elephants like royal elephants.* To achieve this behavior we modify the rule for inheriting ordinary relations by adding an additional restriction, R5. But we are still free to infer the slightly more restricted statement, *Royal elephants like other royal elephants.*

R5. There is no w such that $x_1 \rightarrow \tau_1 \rightarrow w \rightarrow \tau_2 \rightarrow x_n \in \Phi$, $x_1 \rightarrow \tau_1' \rightarrow w \rightarrow \tau_2' \rightarrow y_m \in \Phi$, and $w \overset{R}{\not\rightarrow} \odot \in \Phi$. (Preemption by explicit reflexive statement.)

7 Irreflexive Statements

Let $x \xrightarrow{R} \oslash : y$ mean "x's are in relation R to other y's." Node y must be a class, not an individual, for this construct to make sense. An *explicit* irreflexive path is of form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} \oslash : y_m \leftarrow \ldots \leftarrow y_1$. For example, if opera stars admire other celebrities, ($o \xrightarrow{R} \oslash : c$), Beverly is an opera star, and Frank is a movie star (hence a celebrity), we may conclude that Beverly admires Frank ($B \rightarrow o \xrightarrow{R} \oslash : c \leftarrow m \leftarrow F$).

An *implicit* irreflexive path is generated when an ordinary relation doubles back on itself. Implicit paths take the form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} z_p \leftarrow \ldots \leftarrow z_1 \leftarrow \oslash : y_m \leftarrow \ldots \leftarrow y_1$, with $p > 0$. In Figure 1, since herbivores love gray things, and elephants are gray herbivores, we generate implicit paths for both *Elephants love themselves* and *Elephants love other elephants.* The latter path is written $e \rightarrow h \xrightarrow{R} g \leftarrow \oslash : e$.

7.1 Implicit Irreflexive Paths

Let κ be a path of form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} z_p \leftarrow \ldots \leftarrow z_1 \leftarrow y$, where $y = x_1$ and y is a class rather than an individual. Then the implicit irreflexive path $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} z_p \leftarrow \ldots \leftarrow z_1 \leftarrow \oslash : y$ is inheritable in Φ iff:

OR1. ButFirst(κ) $\in \Phi$.

OR2. ButLast(κ) $\in \Phi$.

OR3. There is no path $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau} \leftarrow y \in \Phi$. (Contradiction)

OR4. There is no path $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau}_1 \leftarrow \oslash : \bar{\tau}_2 \leftarrow y \in \Phi$. (Contradiction)

As was the case with implicit reflexive paths, we do not require κ to be present in Φ; it could be preempted.

7.2 Inheritance of Irreflexive Paths

The $\oslash$ node never stands alone; it always appears connected to a node indicating the referent of the word "other." To simplify the definition below, we will treat the structure $\oslash : y$ as a single node. In particular, ButLast($\sigma \xrightarrow{R} \bar{\tau} \leftarrow \oslash : y$) is $\sigma \xrightarrow{R} \bar{\tau}$.

Let κ be a path of form $x_1 \rightarrow \ldots \rightarrow x_n \xrightarrow{R} z_p \leftarrow \ldots \leftarrow z_1 \leftarrow \oslash : y_m \leftarrow \ldots \leftarrow y_1$. (If $p = 0$ there is a direct link $x_n \xrightarrow{R} \oslash : y_m$.) The path κ is inheritable in Φ iff:

O1. ButFirst(κ) $\in \Phi$.

O2. ButLast(κ) $\in \Phi$.

O3. There is no path $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau} \leftarrow y_1 \in \Phi$. (Contradiction)

O4. There is no path $x_1 \rightarrow \sigma \overset{R}{\not\rightarrow} \bar{\tau}_1 \leftarrow \oslash : \bar{\tau}_2 \leftarrow y_1 \in \Phi$. (Contradiction)

O5. There are no w, w' such that $x_1 \rightarrow \tau_1 \rightarrow w \rightarrow \tau_2 \rightarrow x_n \in \Phi$, $y_1 \rightarrow \tau_1' \rightarrow w' \rightarrow \tau_2' \rightarrow z_p \in \Phi$ (or, if $p = 0$, then $y_1 \rightarrow \tau_1' \rightarrow w' \rightarrow \tau_2' \rightarrow y_m \in \Phi$), and $w \overset{R}{\not\rightarrow} \oslash : w' \in \Phi$, where τ_1 or τ_2 may be null and w may equal x_1 or x_n, respectively, and similarly τ_1' or τ_2' may be null and w' may equal y_1 or z_p (or y_m if $p = 0$), respectively. (Preemption by explicit irreflexive statement.)

O6. There are no w, w' such that $x_1 \rightarrow \tau_1 \rightarrow w \rightarrow \tau_2 \rightarrow x_n \in \Phi$, $y_1 \rightarrow \tau_1' \rightarrow w' \rightarrow \tau_2' \rightarrow z_p \in \Phi$ (or, if $p = 0$, then $y_1 \rightarrow \tau_1' \rightarrow w' \rightarrow \tau_2' \rightarrow y_m \in \Phi$), and $w \overset{R}{\not\rightarrow} w' \in \Phi$, where τ_1 or τ_2 may be null and w may equal x_1 or x_n, respectively, and similarly τ_1' or τ_2' may be null and w' may equal y_1 or z_p (or y_m if $p = 0$), respectively. (Preemption by more specific ordinary relation.)

O7. If x_1 and y_1 are individuals then $x_1 \neq y_1$. (Non-coreferentiality.)

Since opera stars do not admire rock stars, Beverly does not admire Mick; this is an instance of preemption due to O6. We do no derive *Beverly admires Beverly* because of the non-coreferentiality constraint, O7.

7.3 Modification to Ordinary Relations

Irreflexive statements can also block the inheritance of contradictory ordinary relations. For example, if *Herbivores like gray things*, but *Wild elephants do not like other elephants*, we want to block the inference that *Wild elephants like elephants*. This is accomplished by R6 below. We may still infer *Wild elephants like themselves.*

R6. There are no w, w' such that $x_1 \rightarrow \tau_1 \rightarrow w \rightarrow \tau_2 \rightarrow x_n \in \Phi$, $y_1 \rightarrow \tau_1' \rightarrow w' \rightarrow \tau_2' \rightarrow y_m \in \Phi$, and $w \overset{R}{\not\rightarrow} \oslash : w' \in \Phi$.

8 Discussion

Since generics admit exceptions, they cannot be expressed in classical first order logic. We therefore started with a nonmonotonic inheritance system that allowed us to represent generic statements such as *Elephants are gray.* We then extended the system by adding axioms for reflexive and irreflexive statements. Although there are some subtleties in the phrasing of the new axioms which space does not permit us to go into, the general nature of the extended system should be clear.

One thing we have not yet done is prove the constructibility (or at least the existence) of extensions. However, a constructibility proof for networks containing only ordinary relations was given in [Touretzky, 1986]. We are confident that addition of reflexive and irreflexive relations presents no obstacle to constructibility.

Our system can derive new statements about classes as well as about individuals. Inheritance systems based on default logic cannot. This difference becomes more apparent when reflexives are added to the language, because relational paths that double back on themselves can generate reflexive paths even when a network contains no explicit reflexive statements. From *Herbivores like gray things*, for example, we can derive the generic conclusion *Elephants like themselves*, even if there are no instances of elephants in the network.

Some researchers may still prefer to operate within a default logic framework, since default logic has greater expressive power than current semantic network formalisms. Our formulation will be valuable for them as well, since we have solved the problem of extending the inferential distance ordering (the determiner of preemption) to reflexive and irreflexive statements. Etherington's default logic formulation of ordinary relational inheritance, which replaced our path-based notation with default rules, still relied on our inferential distance definition to filter the set of extensions [Etherington 1987b]. This was necessary to ensure that subclasses did indeed override superclasses. A similar translation of our new system into default logic would appear to be straightforward.

Another advantage of our path-based formulation is that it does not require the use of variables to constrain co-referentiality. There is a natural mapping between inheritance paths and surface structure which does not exist for predicate logic-based treatments of reflexives. Reflexive pronouns map to $\odot$ nodes, and the phrase "other y's" maps to $\oslash : y$. Inheritance paths may be translated to English sentences by extracting the first node, the relation, and the last node, as when we read $e \rightarrow h \overset{R}{\rightarrow} g \leftarrow e \leftarrow \odot$ as *Elephants like themselves.* The interior of the path serves as an argument or justification for the statement.

In conclusion, there is no *a priori* reason why the rich structure of human language should map conveniently to a predicate logic-based representation. Logic was originally developed to describe mathematics. One can increase the expressive power of classical logic by adding nonstandard quantifiers, modal operators, and extra truth values, but other formalisms may in some cases prove more natural. We find path-based formalisms convenient for inheritance reasoning, and their treatment of reflexives more natural than logic-based formalisms.

9 Some Linguistic Observations

There are two possible interpretations of the sentence *Rock stars detest other celebrities*, depending on the scope of the word "other." We have so far been using the narrow interpretation of "other," which is that each rock star detests all celebrities other than himself or herself. This is implemented by clause O7 in the definition of inheritability for irreflexive paths. The alternate, broad interpretation of *Rock stars detest other celebrities* is that rock stars detest celebrities other than rock stars. We will not formalize this second interpretation here, but it appears straightforward to handle. The two uses of "other" can even be intermixed by introducing a new node type to denote broadly scoped "other." Note that the distinction between narrow and broad scope disappears when the origin of the relational link is an individual, e.g., *Harry is jealous of other musicians* can only mean "musicians other than himself," while *Trombone players are jealous of other musicians* is ambiguous.

Similarly, for relational links whose head and tail reference the same class, only the narrow interpretation makes sense, e.g., $e \xrightarrow{R} \oslash : e$ could only mean that elephants love elephants other than themselves.

In English one can substitute the expression "each other" when the first and last nodes of a relational path are identical: compare *Elephants love other mammals* ($e \xrightarrow{R} \oslash : m$) with *Elephants love each other* ($e \xrightarrow{R} \oslash : e$). This substitution is mandatory for some speakers.

One aspect of the use of "other" in English that is not part of the formal system presented here is that it usually requires a subset membership. For example, *Politicians intimidate other crooks* cannot be true unless *Politicians are crooks* is true. With our current set of axioms, the link $p \xrightarrow{R} \oslash : c$ means "politicians intimidate crooks in general, but do not conclude from this that an individual politician who is also a crook intimidates himself." It doesn't imply that any politicians actually are crooks. We can get the true English semantics by imposing a restriction on networks to require that any link of form $x \xrightarrow{R} \oslash : y$ be accompanied by a link $x \longrightarrow y$, unless $x = y$.

Finally, we acknowledge that our account of "other" is far from complete. For example, "other" has an existential interpretation as well as the universal one we have been using. A sentence like *Roger fools around with other women* means some women other than his wife, not every woman who is not his wife. Semantics mainly determines which sense is appropriate, but there may also be syntactic cues. For some speakers, *Elephants love other elephants* is preferentially understood as an existential because they expect the universal interpretation to be expressed *Elephants love each other.*

Acknowledgements

This research was supported by the National Science Foundation under grant number IRI-8700705. We are grateful to Jeff Horty for helpful discussions, and to Scott Fahlman, whose NETL system provided the original impetus for our investigation of reflexives.

References

[Altham 1971] Altham, J. E. J. *The Logic of Plurality.* Methuen and Company, Ltd., London, 1971.

[Aoun 1985] Aoun., J., *A Grammar of Anaphora.* MIT Press, Cambridge, Massachusetts, 1985.

[Carlson 1982] Carlson, G., "Generic terms and generic sentences." *Journal of Philosophical Logic* **11**, (1982), pp. 145-182.

[Etherington 1987a] Etherington, D. W. "Formalizing nonmonotonic reasoning systems." *Artificial Intelligence* **31**, pp. 41-85.

[Etherington 1987b] Etherington, D. W. "More on inheritance hierarchies with exceptions: default theories and inferential distance." *Proc. AAAI-87*, Seattle, 352-357.

[Fahlman 1979] Fahlman, S. E., *NETL: a System for Representing and Using Real-World Knowledge.* MIT Press, Cambridge, Massachusetts, 1979.

[Horty *et al.* 1987] Horty, J. F., Thomason, R. H., and Touretzky, D. S. "A skeptical theory of inheritance in nonmonotonic semantic networks." *Proc. AAAI-87*, pp. 358-363.

[May 1985] May., R., *Logical Form: Its Structure and Derivation.* MIT Press, Cambridge, Massachusetts, 1985.

[Sandewall 1986] Sandewall, E. "Non-monotonic inherence rules for multiple inheritance with exceptions." *Proc. IEEE* vol. 74, pp. 1345-1353.

[Thomason 1976] Thomason, R., "Some extensions of Montague grammar." In *Montague Grammar*, B. Partee, ed., Academic Press, New York, 1976, pp. 77-117.

[Thomason 1983] Thomason, R., "On the semantic interpretation of the Thomason 1972 fragment." Indiana University Linguistics Club, 1979.

[Touretzky, 1986] Touretzky, D. S. *The Mathematics of Inheritance Systems.* Morgan-Kauffman, 1986.

[Touretzky *et al.* 1987] Touretzky, D. S., Thomason, R. H., and Horty, J. F. "A clash of intuitions: the current state of nonmonotonic multiple inheritance systems." *Proc. IJCAI-87*, Milan, pp. 476-482.

[Touretzky & Thomason, *forthcoming*] Touretzky, D. S., and Thomason, R. H. "An inheritance-based theory of generic reflexives." Forthcoming technical report, Carnegie Mellon University.

[van Benthem & ter Meulen 1985] van Benthem, J. and A. ter Meulen, eds., *Generalized Quantifiers in Natural Language.* Foris Publications, Dordrecht.