Nonmonotonicity and the Scope of Reasoning: Preliminary Report

David W. Etherington
Artificial Intelligence Principles Research Department
AT&T Bell Laboratories
Murray Hill, NJ 07974
ether@research.att.com

Sarit Kraus*
Institute for Advanced Computer Studies
Department of Computer Science
University of Maryland
College Park, MD 20742
sarit@umiacs.umd.edu

Donald Perlis*
Department of Computer Science
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742
perlis@cs.umd.edu

Abstract

Existing formalisms for default reasoning capture some aspects of the nonmonotonicity of human commonsense reasoning. However, Perlis has shown that one of these formalisms, circumscription, is subject to certain counterintuitive limitations. Kraus and Perlis suggested a partial solution, but significant problems remain. In this paper, we observe that the unfortunate limitations of circumscription are even broader than Perlis originally pointed out. Moreover, these problems are not confined to circumscription; they appear to be endemic in current nonmonotonic reasoning formalisms. We develop a much more general solution than that of Kraus and Perlis, involving restricting the scope of nonmonotonic reasoning, and show that it remedies these problems in a variety of formalisms.

Introduction

The search for theories of *nonmonotonic reasoning*—theories of how to reach reasonable conclusions that are not *strictly* entailed by what is known, and hence are subject to retraction—has yielded many promising formal systems. While these formalisms provide many useful insights, each has some persistent problems that have, thus far, resisted solution. In many naturally-occurring cases, the straightforward encoding of a situation either leads these commonsense-reasoning formalisms to quite unintuitive conclusions or prevents the derivation of intuitively-obvious conclusions.

We discuss several such significant problems, and show their manifestations in each of the major formalisms. We argue that these problems are actually aspects of a single, more general, problem, having more to do with the underlying understanding of the function of nonmonotonic reasoning than with the particular details of existing frameworks. We then show that a simple idea, simple in its realization, solves these problems. This not only greatly enhances the usefulness of the theories, but seems to bring them into much closer harmony with an intuitive understanding of commonsense reasoning.

*Supported in part by ARO research contract no. DAAL03-88-K0087.

"Paradoxes" of Nonmonotonic Reasoning

A study of the problems with existing theories of nonmonotonic reasoning—Default Logic [Reiter 1980b], Circumscription [McCarthy 1980; 1986], and Autoepistemic Logic [Moore 1985]—presupposes at least some familiarity with those formalisms. Space limitations preclude reintroducing the formalisms here; the unfamiliar reader is referred to [Etherington 1988] for a detailed introduction. Familiarity with the basics of nonmonotonic reasoning should suffice for most purposes in this paper.

Different variants of these formalisms have been studied for many years. For most of that time it was believed that they (or at least some of them) captured the essential ideas of nonmonotonic reasoning and that it would only be a matter of time before they could be adapted to practical reasoning systems. Recently, problems have been noticed that seem to shake these optimistic projections. Some of these—such as the "Yale Shooting Problem" [Hanks and McDermott 1986]—seem more indicative of the difficulty of adequately axiomatizing even a relatively simple world; others seem more paradoxical, since the formalisms' basic mechanisms block the conclusions they were, intuitively, designed to produce.

We briefly recount four such "paradoxes" of nonmonotonic reasoning, and show how they affect the various formalisms. We then argue that the observed problems can be viewed as stemming from a common root—a misapprehension, common to all the approaches, of the principles underlying this type of reasoning. Once identified, this deficiency is readily corrected with simple tools whose benefits, we believe, easily outweigh their cost.

The Lottery Paradox

The first problematic example is the "Lottery Paradox" [Kyburg 1961; Perlis 1986]. The lottery paradox arises in situations in which the conjunction of a set of assumptions, each reasonable individually, is inconsistent with what is known about the world. For example, in the paradigmatic case, it is usually safe to assume that any particular ticket in a lottery will not win—given the overwhelming odds against it. Assuming the lottery is "fair", however, the conjunction of such an assumption for each ticket with the fact that some ticket must win is inconsistent.[1]

To maintain consistency, some (or all) of the assumptions about tickets not winning must be foregone. Since there is no basis for determining *which* assumptions to forego, however, any is as good as any other, and none are unequivocally sanctioned. There are as many preferred models (or extensions) as there are tickets, each with a different ticket chosen as the winner. Since nonmonotonic formalisms generally license conjectures based on what is true in all preferred models (extensions), nothing can be assumed about the individual tickets. The most that can be assumed is that if some particular ticket wins, it will be the only one.

Counterexample Axioms

Problems also occur when there are *counterexample axioms* [Perlis 1986] that assert that there are exceptions to defaults. Counterexample axioms specify the existence of individuals lacking some default property, without specifying their identities. For example, given the "birds fly" default, a counterexample axiom might look like $\exists x.\ Bird(x) \wedge \neg Flies(x)$. Circumscription has trouble with such axioms because it stipulates that there are as few exceptions as possible, without necessarily determining *which* individuals are exceptional. Thus, any of a number of individuals might be exceptional without changing the number of exceptions. For example, if we minimize the set of flightless birds in the theory $\{Bird(Tweety),\ \exists x.\ Bird(x) \wedge \neg Flies(x)\}$, we cannot conjecture that Tweety flies, since there is a minimal model in which Tweety is the only bird, and hence the flightless bird stipulated by the counterexample axiom. Even if we posit the existence of other birds different from Tweety,[2] circumscription has no way to prefer Tweety's flying to that of any other bird.

The obvious patch is to try to somehow distinguish Tweety from the existentially-specified flightless bird, for instance by naming the latter (say Opus), and replacing the original counterexample axiom by a Skolemized version such as: $Bird(Opus) \wedge \neg Flies(Opus)$. However, $Flies(Tweety)$ still does not follow by circumscription unless the further axiom that $Tweety \neq Opus$ is adopted. But this amounts to assuming that Tweety is not the exceptional bird—which seems to obviate the circumscription.

Default logic and autoepistemic logic are less susceptible to counterexample axioms, since their conclusions can affect the ontology, but they are not immune. The peculiar conclusions that sometimes arise, especially in the context of domain closure axioms[3] or axioms restricting the reference class (e.g., *Bird*), are discussed in [Etherington *et al.* 1990].

Everything is Abnormal

Yet another inappropriate result occurs when there are defaults describing the typical values of a variety of (possibly orthogonal) properties for some class. If that class consists of several subclasses, each but one of which is atypical with respect to a different property, then current nonmonotonic formalisms will conjecture that individuals known to belong to the class must belong to the completely typical subclass [Poole 1989].

For example, imagine that birds typically fly, sing, are drab, and build nests, except that penguins don't fly, swans don't sing, and mynahs don't build nests. Now if birds must be penguins, swans, mynahs, or canaries, default reasoners of the type envisioned in the literature will assume that arbitrary birds are canaries, in order to minimize the violation of defaults!

Even more counterintuitively, if it turns out that *all* subclasses are atypical (e.g., canaries are found to be abnormal by virtue of being brightly coloured), then nonmonotonic formalisms will suddenly no longer be able to make any normality conjectures: different atypicalities will hold in different minimal models (extensions); the theory entails that some abnormality holds in each. Thus, e.g., learning that canaries are not drab blocks the assumption that Tweety flies.

Poole [1989] and others have noticed that situations in which *everything* is abnormal in some way occur frequently in practice. This suggests that the problem is not an isolated baroque instance where the formalisms do not perform well but is, rather, symptomatic of fundamental difficulties.

[1] We assume the set of tickets is fixed and finite. Other, related, problems arise if not.

[2] Since circumscription cannot generate new equality facts without resorting to variable terms [Etherington *et al.* 1985], explicit inequalities are needed to rule out models where only Tweety is a bird, but she goes by various aliases.

[3] A *domain-closure axiom* (DCA) [Reiter 1980a] is a formula of the form $\forall x.\ x = t_1 \vee \ldots \vee x = t_n$, for some set of ground terms, $t_1, \ldots, t_n$.

There's Nobody Here But Us Chickens[4]

Another counterintuitive aspect of some nonmonotonic formalisms is that, in their efforts to maximize typicality, they conjecture that exceptional classes are empty. Since belonging to an exceptional class entails violating a default, they naturally infer that exceptional classes have as few members as possible. This is both reasonable and nonsensical: reasonable because default reasoning *does* seem to involve assuming things are as normal as possible; nonsensical because the assumption that some object of interest is typical should not necessarily rest on the absence of atypicality elsewhere in the world.

Circumscription is particularly susceptible, explicitly stating that there is no less exceptional world than this; its semantics explicitly prefers those models where all exceptions are forced. For example, if we are told that penguins are flightless birds, that birds normally fly, and that Tweety is a bird:

$$\forall x.\ Penguin(x) \supset Bird(x) \wedge \neg Flies(x)$$
$$\forall x.\ Bird(x) \wedge \neg Abnormal(x) \supset Flies(x)$$
$$Bird(Tweety)$$

and minimize the set of abnormal individuals, we conclude that Tweety flies, and hence is not a penguin—but also that there are no penguins! Conversely, if the objection to this conclusion is made explicit, by asserting $\exists x.\ Penguin(x)$, the enriched theory implies the counterexample axiom, $\exists x.\ Bird(x) \wedge \neg Flies(x)$, and $Flies(Tweety)$ is no longer conjectured.

The obvious answer, including *Penguin* in the set of fixed predicates, prevents the conclusion that there are no penguins, at the expense of the ability to conclude that Tweety flies. With *Penguin* fixed, the strongest conjecture that can be made about Tweety is that she flies unless she is a penguin, which seems unsatisfactory.

The problem is more subtle in default logic, since the effects of default reasoning are conditioned by the provability of the prerequisites of defaults, and the form of the default plays a greater role. For example, the default: $\dfrac{Bird(x)\ :\ Flies(x)}{Flies(x)}$ will sanction the conjecture that none of the known birds are penguins, but not that there are no penguins at all. The former seems more innocuous, although perhaps less so as the number of known birds becomes very large. If all birds are known, the conclusion that there are no penguins follows. Other popular default representations (e.g., "abnormality" theories) can exaggerate the problem. This is discussed in detail in [Etherington *et al.* 1990], where similar problems with autoepistemic logic are also outlined.

[4] Or whatever class of birds is quintessentially prototypical.

A Common Thread

Each of the above difficulties with existing theories of nonmonotonic reasoning can be attributed to a single cause—overzealousness. In the attempt to capture default reasoning, a subtle twist has been introduced. The commonsense notion that such reasoning is essentially the *elimination of unforced abnormalities* has become the notion of the *introduction of forced normalities.*

Assumptions are necessary in everyday reasoning because what follows from what we *know* about the world leaves too many questions undecided. Paradoxically, the mechanisms developed to redress this shortcoming leave too *few* questions undecided. Using such tools to decide whether Tweety flies is akin to cracking walnuts with a cannon—not only are there likely to be undesired side-effects, but the meat of the matter may be much harder to find among the irrelevant fragments.

We frequently know *that* there are exceptional individuals without knowing *who* they are. If defaults are applied injudiciously, paradoxes are bound to arise—yet paradoxes rarely arise in people's default reasoning. It seems clear that defaults are usually not broadly applied.

The *directed* nature of reasoning seems to have been ignored. We contend that the intention of default reasoning is generally not to determine the properties of every individual in the domain, but rather those of some particular individual(s) of interest. Incorporating uncertain beliefs into a belief system when those beliefs are not of direct interest is likely to be counterproductive, simply increasing the probability that some beliefs will have to be retracted.

Reconsider the paradoxes discussed above. In each case, problems arise because something atypical must exist and default reasoning might encompass it. In the case of the lottery paradox, by considering the fate of every ticket, we face the problem that some ticket must win—giving rise to numerous "preferred" models. If we could reason about only the small set of tickets we might consider buying, there would be no problem with assuming that none of them would win, and we would find ourselves safely past the lottery vendor. Similarly, faced with a counterexample axiom, so long as there was no expectation that the posited counterexample was among the individuals of interest, one could make assumptions about the interesting cases without wrestling with the identity of the counterexample. Analogously, when everything is abnormal in some aspect or other, it should be possible to reason about a few aspects of interest, and ignore all the others. Finally, when the scope of interest does not cover whole domains, conjectures to the effect that atypical classes are empty would not arise.

The risk associated with making any particular conjecture on the basis that it is supported by all extensions of a scoped theory is generally higher than the correspond risk for standard default reasoning. However, provided the scope of interest is sufficiently narrow vis à vis the antecedent class(es) for the defaults, the risk does not seem disproportionate to that of doing default reasoning in the first place. Intuitively, since fewer substantive default conclusions are made, it is reasonable to believe that the net result is more probable. Of course, if the scope is too broad, or there is evidence that exceptional cases are within the scope, the advisability of making assumptions decreases proportionally.

Scope in Nonmonotonic Reasoning

At the conceptual level, then, it is clear that making the default reasoning processes dependent on the scope of interest enables intuitively-desirable conclusions in otherwise intransigent cases. We next show that this can be done easily for the existing formalisms, that more powerful conjectures obtain, and that appropriate notions of consistency are preserved.

As a methodological point, we require that the scope of reasoning be narrow. We do not attempt to define or enforce this, beyond noting that the scope of interest should not include a "significant fraction" of whatever reference class we are drawing default conclusions about. Our approach to limiting the scope of reasoning ensures that—even when this requirement is violated—performance and consistency will be at least as good as that of the unscoped approaches, however.

The technical requirements for limiting the scope of default reasoning are methodological rather than structural. The contribution of this work is not sophisticated new versions of the formalisms—developing yet another nonmonotonic formalism is unnecessary. The important result is that a simple, *uniform*, representational technique provides significant leverage on a variety of problems across a variety of formalisms.

Scoped Circumscription

Circumscription can accommodate scope by minimizing only within the extent of a predicate representing the scope of interest. Specifically, we minimize $W[P,y] \wedge Scope(y)$ rather than just $W[P,y]$, resulting in the scoped circumscription schema, $CIRC_{Scope}$:[5]

$$A[P'] \wedge [\forall y. W[P',y] \wedge Scope'(y) \rightarrow W[P,y] \wedge Scope(y)] \rightarrow [\forall y. W[P,y] \wedge Scope(y) \rightarrow W[P',y] \wedge Scope'(y)].$$

Scoped circumscription overcomes many of the limitations of its unscoped counterpart. For example, it provides a solution to the counterexample problem. Given a nontrivial domain with Tweety in the scope of concern, it is possible to conclude that Tweety flies from $Bird(Tweety)$, despite the presence of a counterexample axiom. To see this, consider the following axioms, $A[Scope, Bird, Flies]$:

$$Bird(Tweety)$$
$$\exists x.\ Bird(x) \wedge \neg Flies(x)$$
$$Charlie \neq Tweety$$
$$Scope(Tweety).$$

We introduce Charlie here to ensure an ontology rich enough to allow the formation of various interpretations. In particular, we need an object other than Tweety that we can at least imagine to be a potential flightless bird, to let Tweety off the hook. However, Charlie's role as "scapebird" is quite limited—we do not conclude $\neg Flies(Charlie)$ nor even $Bird(Charlie)$. It would even suffice to have simply $\exists x.\ x \neq Tweety$ instead of $Charlie \neq Tweety$. Since, in general, we expect any realistic ontology to provide many individuals, this requirement presents no particular hardship.

From A and $CIRC_{Scope}$, with $W[Bird, Flies, y]$ being $\neg Flies(y)$, $Flies(Tweety)$ follows. The necessary substitutions are $x = x$ for $Bird'(x)$, and $x = Tweety$ for $Scope'(x)$ and $Flies'(x)$. We get $\forall x.\ Flies(x) \vee \neg Scope(x)$—all non-fliers are outside the scope of reasoning—and so $Flies(Tweety)$, since we have $Scope(Tweety)$.

More generally, even given some ***known*** scoped exceptions, scoped circumscription can frequently preclude ***unknown*** exceptions in the scope, as the following theorem shows.

> **Theorem 1** If $A \vdash W(P,\alpha_i) \wedge Scope(\alpha_i)$, for ground terms $\alpha_i \in \{\alpha_1, ..., \alpha_n\}$, and no consistent extension of A by ground (in)equalities entails $\exists x.\ x \neq \alpha_1 \wedge ... \wedge x \neq \alpha_n \wedge W(P,x) \wedge Scope(x)$, then $CIRC_{Scope}[A] \vdash \forall x.\ [x \neq \alpha_1 \wedge ... \wedge x \neq \alpha_n \wedge Scope(x)] \supset \neg W(P,x)$, provided all predicates are variable, and A entails a domain-closure axiom. ∎

It is easily seen that scoped circumscription is similarly effective in the other paradoxical cases.

As the example just above shows, the restrictions in Theorem 1 are stronger than necessary. Essentially, what is required is an ontology with "enough" distinct individuals, but in which exceptions and the scope do not depend on the ontology of the model. Thus, for example, the result cannot be generalized to cover theories such as:

$$a \neq b$$
$$[\forall x.\ x = a \vee x = b] \supset P(a) \wedge Scope(a),$$

where P is to be minimized since, in models with domain $\{a,\ b\}$, a must be a scoped exception, even though a need not be exceptional (nor scoped) in general. The need for "domain independence", captured in the conditions imposed on equality in the theorem,

[5] Notice that this is not a new form of circumscription. Rather, the circumscription is made relative to the *Scope* predicate. This approach can be used independently of which major variant of circumscription is chosen.

is a consequence of circumscription's inability (without use of variable terms) to produce conjectures entailing new facts about the ontology [Etherington 1988]. It may be possible to relax this requirement by allowing variable terms, or using "Equality Circumscription" [Rathmann and Winslett 1989]. This remains to be investigated.

Although the necessary conditions for effective scoped circumscription are difficult to make precise, the problematic cases do not seem particularly troublesome. It seems likely that a realistic theory of a reasonably-complex problem domain will have an abundance of individuals known to be distinct from those known to be in the scope. Similarly, predicating exceptionalness on what exists or what things are identical seems inappropriate for commonsense theories.

It is crucial that the theory not entail that the unknown exceptional individuals claimed to exist are also in the scope; otherwise the problem resurfaces. We argue that it is unreasonable for an agent to use a default while believing that an anonymous object of concern is a counterexample to that default. Notice that there is no problem, however, in believing that there are ***known*** exceptions in the scope (e.g., $Bird(Opus) \wedge \neg Flies(Opus) \wedge Scope(Opus)$).

Is scoped circumscription consistent, however? This question is important because inconsistency has plagued certain applications of circumscription [Etherington *et al.* 1985]. Etherington [1988] shows that theories without existential quantifiers have consistent circumscriptions, but counterexample axioms take us out from under this umbrella of safety. Nevertheless, scoped circumscription is consistent, regardless of the form of the original theory, provided the scope is finite.

Theorem 2 If A has a model in which *Scope* is finite, then $CIRC_{Scope}[A]$ is consistent. ∎

We consider other cases that are "well-behaved", and what can be said about them, in [Etherington *et al.* 1990].

Scoped Default Logic

The greater expressive power of default logic [Etherington 1987] means there are many more candidate methods for restricting the scope of reasoning in default logic than were available in circumscription. In [Etherington *et al.* 1990], we study a variety of possibilities, and compare their representational power. Here, we restrict our attention to one particular representation, and say the ***scoped representation*** of a normal default, $\frac{\alpha x\ :\ \beta x}{\beta x}$, is $\frac{\alpha x \wedge Scope(x)\ :\ \beta x}{\beta x}$.[6] The latter default says that individuals known to be α's in the scope can be assumed to be β's.

[6] α and β may be arbitrary formulae in which x occurs free.

The introduction of scope to default logic is sufficient to circumvent the lottery paradox, as the following example shows. Imagine a lottery with 10,000 tickets, $t_1, \ldots, t_{10,000}$, and imagine we are considering buying one of the tickets, t_{100}–t_{175}, available at the corner store. This corresponds to the theory with the axioms:

$$\forall t.\ Ticket(t) \equiv t = t_1 \vee \ldots \vee t = t_{10,000}$$
$$Scope(t_{100}), \ldots, Scope(t_{175})$$
$$\exists t.\ Ticket(t) \wedge Wins(t)$$

and the default:

$$\frac{Ticket(t) \wedge Scope(t)\ :\ \neg Wins(t)}{\neg Wins(t)}.$$

This theory has a unique extension in which $\neg Wins(t_{100}), \ldots, \neg Wins(t_{175})$, but the fate of the remaining tickets is undecided. Conversely, the unscoped theory has 10,000 extensions, including 76 in which one of the tickets of interest wins.

It is no accident that the desired result holds; we have proved that ground terms in the scope are conjectured to be unexceptional whenever possible.

Theorem 3 If $D = \left\{ \frac{\Phi x \wedge Scope(x)\ :\ \Psi x}{\Psi x} \right\}$ and $W \not\vdash \exists x.\ \Phi x \wedge \neg \Psi x \wedge Scope(x)$, then any extension, E, for $\Delta = (D, W)$ has no scoped exceptions. Specifically, if $E \vdash \Phi\alpha \wedge Scope(\alpha)$ then $E \vdash \Psi\alpha$, for any ground term, α. ∎

Analogous results hold for the other representations, and the results generalize to cases where there are known exceptions, and/or multiple defaults.

It can be shown, in many cases, that every extension of the scope-limited theory is a subset of an extension of the unscoped version. These results are comforting, since they mean that narrowly-scoped reasoning does not lead in directions that would be rejected as unreasonable if the scope of reasoning were broader.

Theorem 4 Let $\Delta = (D, W)$ be a normal default theory, and let D' be the result of replacing each default in D with its scoped counterpart. Then every extension of $\Delta' = (D', W)$ is contained in an extension of Δ. ∎

Scoped Autoepistemic Logic

Unscoped reasoning also presents problems in autoepistemic logic which are ameliorated by restricting the scope of reasoning. However, since a fully-quantificational first-order autoepistemic logic has not yet been formalized (but see [Konolige 1988; Levesque 1987] for suggestions), we restrict our discussion to a propositional version that approximates quantification by grounding variables over a closed domain.

First, consider the Lottery Paradox again. As in the previous section, suppose we have 10,000 lottery tickets, and wish to buy one among the 76 from

$t_{100} \dots t_{175}$. The only change required for autoepistemic logic is that, instead of a default rule, we use the schema:

$$\mathbf{L}Ticket(t) \wedge \mathbf{L}Scope(t) \wedge \neg\mathbf{L}Wins(t) \rightarrow \neg Wins(t)$$

where t ranges over the 10,000 ticket constants. In such a case, as with default logic, we get only one extension, in which we have $\neg Wins(t)$ for the 76 scoped tickets but not for the rest.

To see how multiple, orthogonal, properties can be handled in this framework, suppose there are only three kinds of bird, *Canary*, *Mynah*, and *Penguin*, and that canaries are typical but mynahs and penguins are not, since mynahs do not build nests and penguins do not fly, as in [Poole 1989]:

$$\forall x.\ Mynah(x) \supset \neg Nests(x)$$
$$\forall x.\ Penguin(x) \supset \neg Flies(x)$$
$$\forall x.\ Bird(x) \equiv Mynah(x) \vee Penguin(x) \vee Canary(x).$$

We certainly do not want to conclude that all birds are canaries, although that is the result of straightforward application of autoepistemic logic. Specifically, from the above axioms and the defaults that birds typically fly and build nests:

$$\mathbf{L}Bird(b) \wedge \neg\mathbf{L}\neg Flies(b) \rightarrow Flies(b)$$
$$\mathbf{L}Bird(b) \wedge \neg\mathbf{L}\neg Nests(b) \rightarrow Nests(b)$$

(where again b ranges over the finite set of constants), we get that there are no mynahs or penguins—i.e., all birds are canaries. Scope can help if we employ the two scope-limited schemata:

$$\mathbf{L}Bird(b) \wedge \mathbf{L}Scope(b) \wedge \mathbf{L}Scope(flying) \wedge \neg\mathbf{L}\neg Flies(b) \rightarrow Flies(b)$$
$$\mathbf{L}Bird(b) \wedge \mathbf{L}Scope(b) \wedge \mathbf{L}Scope(nesting) \wedge \neg\mathbf{L}\neg Nests(b) \rightarrow Nests(b).$$

The new constants, *flying* and *nesting*, represent particular aspects of the descriptions of birds that might be of interest at a particular time (see, for example, [McCarthy 1986]). Provided scope is narrow and includes *nesting* and *flying*, there will be only one extension, in which all scoped birds are canaries, but unscoped birds are indeterminate as to species (as well as flying and nesting behaviours). If *Scope* only includes *flying*, we conclude that birds in the scope fly and are not penguins, but remain agnostic on their nesting behaviour.

The examples suggest that scoped autoepistemic reasoning provides an intuitively-plausible solution to the paradoxes. Obviously, general results would be better, even if based on strong restrictions. For the case of strongly-grounded autoepistemic extensions (see [Konolige 1988]), we can provide such results. We begin with a sufficiency result.

Theorem 5 If W entails a domain closure axiom, $W \not\vdash \exists x.\ \Phi x \wedge \neg\Psi x \wedge Scope(x)$, and $D = \{\mathbf{L}\Phi(c) \wedge \mathbf{L}Scope(c) \wedge \neg\mathbf{L}\neg\Psi(c) \supset \Psi(c)\}$ is a schema over all the constants, c, of W, then no strongly grounded autoepistemic extension of $W \cup D$ contains any scoped exceptions. ∎

The obvious generalizations to n-ary predicates and multiple scope terms follow directly. Similarly, we get a consistency result analogous to Theorem 4.

Theorem 6 Suppose W entails a DCA and is $\mathbf{L}$-free, and D consists of schemata of the form $\mathbf{L}\Phi(c) \wedge \neg\mathbf{L}\neg\Psi(c) \supset \Psi(c)$. Let D' be the result of replacing each schema in D by $\mathbf{L}\Phi(c) \wedge \mathbf{L}Scope(c) \wedge \neg\mathbf{L}\neg\Psi(c) \supset \Psi(c)$. Then the $\mathbf{L}$-free subtheory of any strongly-grounded autoepistemic extension of $W \cup D'$ is contained in the $\mathbf{L}$-free subtheory of a strongly-grounded autoepistemic extension of $W \cup D$. ∎

These results are not as broad as those above for circumscription or default logic; however, they suggest the same trend, indicating that a *Scope* predicate can be useful in treating the "paradoxes" of overzealousness (forced normalities) surveyed above.

Related Work

Kraus and Perlis [1988] suggest restricting default reasoning to "named" individuals (individuals for whom the reasoner has a standard name) in order to solve the counterexample problem in a particular variant of circumscription. This approach does not seem to generalize to the other problems we addressed here, nor has it been worked out for the other formalisms we treat. Furthermore, the notion of limiting the scope of reasoning seems to be more flexible and intuitive than that of restricting reasoning to named individuals.

Poole's [1988] THEORIST system provides for goal-directed default reasoning by searching for *explanations* for goals. An explanation consists of a set of defaults which are mutually consistent with the known facts and jointly entail the goal. In paradoxical situations such as those we have discussed, however, THEORIST can generally explain both a goal (e.g., $\neg Wins(ticket_1)$) and its negation ($Wins(ticket_1)$), depending on which defaults it chooses to apply. Based on the correspondence between THEORIST's defaults and those of default logic [Poole 1988], it appears that our notion of scope can be added directly to THEORIST, providing both more tightly focused reasoning and an alternative to paradox. Similarly, Ginsberg's [1988] circumscriptive theorem prover provides facilities for goal-directed nonmonotonic reasoning, but the conclusions it reaches are circumscriptively sound and hence subject to paradox. It seems, therefore that Ginsberg's system might also benefit from our approach.

Conclusions and Future Work

We have pointed out common roots underlying four significant problems with existing approaches to nonmonotonic reasoning. We showed that these problems visit all the major current approaches, and argued that they were real impediments to using these formalisms for commonsense reasoning.

We then introduced the idea of restricting the scope of reasoning, providing powerful leverage on the problems. This idea has direct application in the variants of circumscription, in default logic, and in autoepistemic logic; it is similarly effective in each. Even more satisfying, we showed that what is required to achieve these benefits involves simple methodological changes, rather than development of new formalisms or new variants of existing formalisms.

We outlined how restricting the scope of nonmonotonic reasoning provides acceptable, commonsensical, solutions to the problems in question. These include the lottery paradox, the problem of anonymous exceptions to defaults, the problems arising when almost everything is atypical in some respect, and the tendency to conjecture typicality by rejecting the existence of atypicality.

For the formalisms in question, we showed that the conclusions sanctioned by our strengthened, scope-limited, approach are generally in accord with (some subset of) the preferred models of the original theory. This is comforting, since it means that we have strengthened the theories, rather than simply subverting them. We also showed that appropriate notions of consistency are preserved.

Our framework not only avoids paradox, but also adapts naturally to goal-directed reasoning. Assumptions are sanctioned only about objects of interest; this appears to be much more natural than current maximal-consistent-set approaches. This focussing offers promise for the development of practical nonmonotonic reasoning systems.

The most obvious outstanding question concerns the nature of the scope theory. Ideally, it should be possible to determine scope from the current context, attention, and goals of the agent, although we have not yet worked on this. Among other things, we imagine that the individuals mentioned in a query or goal statement, or attended to as the result of recent discourse or experience will be scoped. We suspect, too, that work such as [Halpern and Rabin 1987], [Halpern and McAllester 1989], [Halpern and Moses 1984], [Drapkin *et al.* 1987], and [Nutter 1983] will be relevant. In particular the notion of an awareness set seems to have a similar spirit. We imagine "scope" to be slightly different, however—more like "of concern" or "relevant to making a decision". In this respect, it is encouraging that the approach seems robust enough to tolerate fairly gross determinations of scope.

In this paper, we have skirted some of the difficult issues of equality and domain closure that face theories of nonmonotonic reasoning. Some of these are taken up in [Etherington *et al.* 1990]—in particular, we consider the effects of various ontological assumptions on the relationship between scoped and unscoped nonmonotonic reasoning. Much work, however, remains to be done in pursuit of a commonsense theory of ontology.

Acknowledgements

We thank Kurt Konolige and Matt Ginsberg for helpful discussions and useful comments about this work. Kurt independently observed that changes to the circumscription schema employed in an early draft were unnecessary.

References

Drapkin, J., Miller, M., and Perlis, D. 1987. Life on a desert island. In *Proc. Workshop on The Frame Problem in Artificial Intelligence*, pages 349–357. American Association for Artificial Intelligence.

Etherington, D. W., Mercer, R. E., and Reiter, R. 1985. On the adequacy of predicate circumscription for closed-world reasoning. *Computational Intelligence*, 1:11–15.

Etherington, D. W., Kraus, S., and Perlis, D. 1990. Nonmonotonicity and the scope of reasoning. Technical Report UMIACS TR-90-56 and CS-TR-2457, University of Maryland. Submitted for publication.

Etherington, D. W. 1987. Relating default logic and circumscription. In *Proc. Tenth Int'l Joint Conf. on Artificial Intelligence*, pages 489–494, Milan, Italy. Int'l Joint Conferences on Artificial Intelligence, Inc.

Etherington, D. W. 1988. *Reasoning with Incomplete Information*. Research Notes in Artificial Intelligence. Pitman Publishing, London, GB.

Ginsberg, M. L. 1988. A circumscriptive theorem prover: Preliminary report. In *Proc. Seventh National Conference on Artificial Intelligence*, pages 470–474, St. Paul, MN. American Association for Artificial Intelligence.

Halpern, J. Y. and McAllester, D. A. 1989. Likelihood, probability and knowledge. *Computational Intelligence*, 5:151–160.

Halpern, J. Y. and Moses, Y. 1984. Towards a theory of knowledge and ignorance: preliminary report. In *Proc. Workshop on Nonmonotonic Reasoning*, pages 125–143, New Paltz, NY. American Association for Artificial Intelligence.

Halpern, J. Y. and Rabin, M. O. 1987. A logic to reason about likelihood. *Artificial Intelligence*, 32:379–405.

Hanks, S. and McDermott, D. 1986. Default reasoning, nonmonotonic logic, and the frame problem. In *Proc. Fifth National Conference on Artificial Intelligence*, pages 328–333, Philadelphia, PA. American Association for Artificial Intelligence.

Konolige, K. 1988. On the relation between default and autoepistemic logic. *Artificial Intelligence*, 35:343–382.

Kraus, S. and Perlis, D. 1988. Names and nonmonotonicity. Technical Report UMIACS TR-88-84 and CS-TR-2140, University of Maryland.

Kyburg, H. 1961. *Probability and the Logic of Rational Belief.* Wesleyan University Press, Middletown, CT.

Levesque, H. 1987. All I know: An abridged report. In *Proc. Sixth National Conference on Artificial Intelligence*, pages 426–431, Seattle, WA. American Association for Artificial Intelligence.

McCarthy, J. 1980. Circumscription–a form of nonmonotonic reasoning. *Artificial Intelligence*, 13:27–39.

McCarthy, J. 1986. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28:89–118.

Moore, R. 1985. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25:75–94.

Nutter, J. T. 1983. Default reasoning using monotonic logic: a modest proposal. In *Proc. Third National Conference on Artificial Intelligence*, pages 297–300. American Association for Artificial Intelligence.

Perlis, D. 1986. On the consistency of commonsense reasoning. *Computational Intelligence*, 2:180–190.

Poole, D. 1988. A logical framework for dft reasoning. *Artificial Intelligence*, 36:27–47.

Poole, D. 1989. What the lottery paradox tells us about nonmonotonic reasoning. In *Proc. First Int'l Conf. on Principles of Knowledge Representation and Reasoning*, pages 333–340.

Rathmann, P. K. and Winslett, M. 1989. Circumscribing equality. In *Proc. Eleventh Int'l Joint Conf. on Artificial Intelligence*, pages 468–473, Detroit, MI. Int'l Joint Conferences on Artificial Intelligence, Inc.

Reiter, R. 1980. Equality and domain closure in first-order databases. *J.ACM*, 27:235–249.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence*, 13:81–132.