Prioritized Autoepistemic Logic

Jussi Rintanen *

Helsinki University of Technology
Department of Computer Science
Otakaari 1, 02150 ESPOO, Finland

Abstract. An important problem in data and knowledge representation is the possibility of default rules that conflict. If the application of both of two default rules leads to a contradiction, they cannot both be applied. Systems that support the use of default rules may either remain indifferent or prioritize one rule over the other. In this paper a prioritized version of autoepistemic logic is presented. Priorities determine a subset of all stable expansions of a set, the preferred stable expansions. The priority notion is declarative, unlike e.g. some recent approaches to priorities in default logic that modify the semi-constructive definition of extensions of Reiter. Computationally the new priority notion can nevertheless be seen as a mechanism for pruning search trees in procedures for autoepistemic reasoning, as demonstrated by procedures given in the paper.

1 Introduction

Default rules are used in many areas of knowledge representation, e.g. links in inheritance hierarchies [Touretzky *et al.*, 1987] and correctness assumptions in diagnostic reasoning [Reiter, 1987] can be seen as default rules. A problem in these and other areas of knowledge representation is how to deal with conflicting defaults. A safe way is to resort to skepticism and consider all possible ways of resolving the conflicts between defaults. In diagnostic reasoning this corresponds to the computation of *all* diagnoses. Another way is to bring more knowledge – priority information – to the system so that conflicts can be solved in a principled way: e.g. in single-inheritance the class-inclusion relation directly acts as priority information. Explicitly represented priority information is necessary e.g. for solving conflicts in multiple-inheritance.

Conflicts between defaults, as described above, show up in nonmonotonic logics and other formalizations of defeasible reasoning. Default logic [Reiter, 1980] and autoepistemic logic [Moore, 1985] generate several extensions and stable expansions for many sets of rules and facts. Similarly circumscription [McCarthy, 1980] sanctions distinct classes of minimal models. If conflicts are left unresolved, the convention is to take the intersection of the extensions, the stable expansions,

* This work was funded by Suomen Akatemia. Additional support from the following foundations is gratefully acknowledged: Emil Aaltosen säätiö, Heikki ja Hilma Honkasen säätiö, Alfred Kordelinin yleinen edistys- ja sivistysrahasto.

or the sets of formulae true in minimal models. The priorities can be coded in default rules. In default logic, for example, the justifications in semi-normal defaults can be used for blocking the application of a low-priority rule in favour of a higher-priority rule [Reiter and Criscuolo, 1981]. However, representing such dependencies explicitly in rules is inconvenient, and complicates the maintenance of sets of rules.

Another line of research distinct from autoepistemic logic and other nonmonotonic logics based on consistency uses model preference [Shoham, 1987; Kraus *et al.*, 1990]. Shoham defines a general framework of preferential logics in which both nonmonotonicity and resolution of conflicts between defaults are achieved using model preference. Kraus et al. [1990] use a framework similar to Shoham's for defining logics in which conflicts are solved according to the specificity of the defaults. The problem in these logics is that the addition of irrelevant premises cancels useful conclusions. Geffner and Pearl [1992] avoid these irrelevance problems in their conditional entailment which is obtained by modifying the logics of [Kraus *et al.*, 1990]. Conditional entailment can be seen as a variant of preferred subtheories of Brewka [1989] together with a fixed mechanism for computing priorities. In this paper we concentrate on consistency-based nonmonotonic logics.

There have been several proposals to incorporate priorities in consistency-based nonmonotonic logics in an abstract form, or to introduce related systems that involve explicit priorities, e.g. [Lifschitz, 1985; Konolige, 1989; Brewka, 1989; MacNish, 1991; Ryan, 1992; Tan and Treur, 1992; Brewka, 1992; Baader and Hollunder, 1993]. Some of the earlier proposals concern only simple defaults that correspond to prerequisite-free normal default rules of default logic [Brewka, 1989; Ryan, 1992]. The expressivity of these systems is insufficient. In the propositional case, prioritized circumscription [Lifschitz, 1985] is closely related to preferred subtheories of Brewka [1989], and hence the defaults expressible are essentially prerequisite-free normal. Konolige [1989] and Toyama et al. [1991] give variants of autoepistemic logic in which the syntactic form of defaults is not restricted. MacNish [1991] introduces priorities in default logic, but he encounters serious problems. For example, the extensions his system produces are not default extensions as defined by Reiter [1980]. Tan and Treur [1992] present a very general scheme for bringing priorities in default logic.

Baader and Hollunder [1993] consider arbitrary partial orders of unrestricted default rules of default logic. They give a version of Reiter's semi-constructive definition of default extensions, in which priorities restrict the order of application of default rules. Brewka [1992] has a similar definition. The generality of these systems is comparable to our prioritized autoepistemic logic. However, it seems that these approaches, like those of MacNish, and Tan and Treur, are likely to run into difficulties: the conclusions that are obtained are sometimes unintuitive, or can be explained only on the basis of the mechanism using which they were produced. The semi-constructive definition of extensions is the basis of these proposals, and the principle used is roughly the following: always apply a highest priority rule among all applicable rules that have not already been applied. Consider a default theory consisting of the following formulae and rules (Brewka has given an example similar to this).

$$a \qquad \frac{b : c}{c} \qquad \frac{a : \neg c}{\neg c} \qquad \frac{a : b}{b}$$

Let the three rules be prioritized in the order indicated, i.e. $\frac{b:c}{c}$ has the highest priority. With the above default theory, the second and third default rules are both initially applicable. Hence apply the second one and obtain $\neg c$. Now only the third rule is applicable and b – the prerequisite of the highest priority rule – is obtained. The unintuitivity is that the lowest priority rule is applied in all extensions of the theory, and hence the prerequisite of the highest priority rule $\frac{b:c}{c}$ belongs to all extensions. Despite this fact the highest priority rule does not become applicable because it is blocked by the the rule $\frac{a:\neg c}{\neg c}$. The problem in this example that shows up both in Baader and Hollunder's and Brewka's work, is that the application of high-priority defaults may depend on the application of low-priority rules in unobvious ways. It may be difficult to devise versions of Reiter's semi-constructive definition of extensions that avoid this kind of problems. Furthermore, there does not seem to be alternative, declarative ways of characterizing the preferred extensions in these approaches.

In this paper, we give a definition of priorities in autoepistemic logic. Priorities declaratively select a subset of all stable expansions of a set, the *preferred* stable expansions. Priorities on default rules can be defined in a way that properly fulfills the principle "apply the highest priority default if possible." For example, the problem illustrated in the above example is avoided.

2 Prioritized autoepistemic logic

Prioritized autoepistemic logic is defined as an extension of autoepistemic logic of Moore [1985]. Different versions of autoepistemic logic are obtained by basing it on different monotonic logics, like classical propositional logic or predicate logic. The language of the classical logic on which an autoepistemic logic is based on is denoted by $\mathcal{L}$. To obtain the language $\mathcal{L}_{ae}$ of an autoepistemic logic, $\mathcal{L}$ is extended with the operator L which is read "is believed". The definition of autoepistemic logic extends the notion of a model and the relation $\models$ to cover also formulae beginning with the operator L. These formulae are treated as atomic formulae.

Autoepistemic logics describe the reasoning capabilities of ideally rational agents. An agent is logically omniscient, i.e. it believes all logical consequences of its own beliefs, and it is capable of introspection, i.e. for each proposition ϕ either $L\phi$ or $\neg L\phi$ is its belief depending on whether ϕ is its belief. The beliefs of an autoepistemic agent are based on a set of initial premises. Given a set of initial premises Σ, possible states of belief T of an autoepistemic agent are the solutions of the following equation [Moore, 1985].

$$T = \{\phi \in \mathcal{L}_{ae} | \Sigma \cup \{L\phi | \phi \in T\} \cup \{\neg L\phi | \phi \notin T\} \models \phi\}$$

The sets T are *stable expansions of* Σ. When Σ contains conflicting defaults, there may be several stable expansions. If at least some of the preferences of the autoepistemic agent are available, the number of stable expansions that correspond to possible states of belief can be reduced, and by cautious reasoning, i.e. taking the intersection of the stable expansions, more formulae can be inferred. Prioritized autoepistemic logic is an explication of the preference mechanism used by an autoepistemic agent.

Definition 1. Let $\Phi \subseteq \mathcal{L}_{ae}$ be a set of formulae, and let $\mathcal{P} \subseteq \Phi \times \Phi$ be a transitive and asymmetric relation such that all strict total orders $\mathcal{T} \supseteq \mathcal{P}$ are well-orderings[2]. Then $P = \langle \Phi, \mathcal{P} \rangle$ is a *prioritization*.

The formulae in Φ represent those beliefs that are relevant in determining which state of belief the agent is going to choose. The relation $\phi \mathcal{P} \psi$ expresses that the agent is more reluctant to accept the belief ϕ than the belief ψ, i.e. when having to believe one of them the agent chooses ψ. We have chosen to prefer the absence of formulae of Φ in stable expansions rather than their presence. The opposite, i.e. preferring the presence of formulae in stable expansions, can be easily achieved by allowing formulae $+\phi$ in prioritizations, where $+\phi$ means simply $\neg L\phi$. We call formulae $+\phi$ in prioritizations *positive* and other formulae *negative*.

The relation $\mathcal{P}$ is asymmetric and hence it describes strict preferences. The transitivity property makes the prioritization an ordering relation. The well-ordering property rules out the existence of infinite chains of less and less believable beliefs, and thereby guarantees that any two stable expansions can be ordered. The partial order $\mathcal{P}$ is used lexicographically. Consider two stable expansions E_1 and E_2 given a strict total order $\mathcal{P}$ on Φ. If the least element of Φ belongs to E_1 but not to E_2, then prefer E_2 to E_1, and if it belongs to E_2 but not to E_1, then prefer E_1 to E_2. If the formula belongs to both of them or to none of them, the next formula of Φ is considered and so on. Hence in case of strict total orders a set of most preferred stable expansions naturally arises.

Sometimes not all preferences of an autoepistemic agent are known: the ordering in the prioritization may be properly partial. It seems that the most natural meaning for the partiality is obtained by extending the partial order to a total order, and then performing ordinary lexicographic comparison using the total order. Different notions of most preferred stable expansions are obtained depending on whether only one total order is considered in comparisons or whether different total orders may be used for different stable expansions. The first alternative, a variant of which is used by Brewka [1989], is to take the most preferred stable expansions to be the ones that are according to a single strict total order lexicographically preferred (not necessarily properly) to all other stable expansions. The second alternative, a variant of which is used

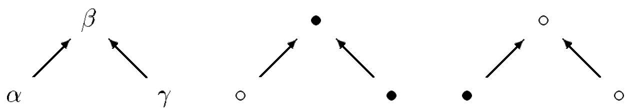
[2] An ordering $\mathcal{P}$ on a set D is a well-ordering if every non-empty subset S of D has an element σ such that $\tau \mathcal{P} \sigma$ for no $\tau \in S$.

by Ryan [1992][3], is to take the most preferred stable expansions to be the ones that are lexicographically preferred (not necessarily properly) to all other stable expansions according to possibly different total orders. The second definition essentially defines an ordering on stable expansions, the maximal elements of which are the most preferred stable expansions. The first definition – which we adopt – in general does not correspond to any such ordering.

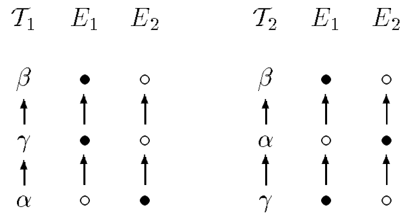
Definition 2 P-preferredness. Let $P = \langle \Phi, \mathcal{P} \rangle$ be a prioritization and Σ a set of formulae. Then E is a *P-preferred* stable expansion of Σ if and only if it is a stable expansion of Σ and there is a strict total order $\mathcal{T}$ of Φ such that $\mathcal{T} \supseteq \mathcal{P}$ and for all stable expansions E' of Σ the following holds.

For all $\phi \in \Phi (\phi \in E \backslash E'$ implies there is $\psi, \psi \mathcal{T} \phi$ and $\psi \in E' \backslash E)$.

Example 1. The diagram below on the left depicts the prioritization $P = \langle \Phi, \mathcal{P} \rangle = \langle \{\alpha, \beta, \gamma\}, \{(\alpha, \beta), (\gamma, \beta)\} \rangle$. Whenever $a \mathcal{P} b$, a is depicted below b. The other two diagrams represent respectively the stable expansions E_1 and E_2 such that $E_1 \cap \Phi = \{\beta, \gamma\}$ and $E_2 \cap \Phi = \{\alpha\}$.



There are two strict total orders that extend $\mathcal{P}$, call them $\mathcal{T}_1$ and $\mathcal{T}_2$, which are shown below.



Because $\alpha \in E_2 \backslash E_1$ and there are no formulae $\psi such that \psi \mathcal{T}_1 \alpha$, E_1 is a P-preferred stable expansion. Similarly, because $\gamma \in E_1 \backslash E_2$ and there are no formulae $\psi, \psi \mathcal{T}_2 \gamma$, E_2 is a P-preferred stable expansion.

The decision problems of prioritized autoepistemic logic are *prioritized cautious reasoning* and *prioritized brave reasoning*, which correspond to the membership of formulae in all P-preferred stable expansions and the membership in some P-preferred stable expansions, respectively.

Theorem 3. *When the underlying logic is the classical propositional logic, prioritized cautious and brave reasoning are respectively Π_2^p-hard and in Π_3^p, and Σ_2^p-hard and in Σ_3^p.*

[3] Ryan does not mention the possibility of stating the meaning of the partiality in his priorities this way.

Proof. Outline: the hardness results of prioritized reasoning are immediate applications of the respective results of ordinary autoepistemic logic [Gottlob, 1992]. The membership in the third level of the polynomial hierarchy is because guessing a stable expansion (not) containing a formula takes nondeterministic polynomial time, and the test that the stable expansion is preferred (i.e. that there are no better stable expansions) uses a Π_2^p-oracle.

3 Relation to other formalisms

Brewka's [1989] preferred subtheories and Ryan's [1992] ordered theory presentations (OTPs) can be seen as prioritized versions of Reiter's default logic restricted to prerequisite-free normal defaults. OTPs are richer in using natural consequences of formulae. There is an exact translation of Brewka's preferred subtheories to prioritized autoepistemic logic. The formulae ϕ in a theory T translate simply into $\neg L\neg\phi \rightarrow \phi$, and the priorities into $\langle\{\neg\phi|\phi \in T\}, \{(\neg\phi, \neg\psi)|\phi < \psi\}\rangle$. The preference notion used by Ryan differs slightly from the one used by us and Brewka. Instead of requiring that a single total ordering of the prioritization partial order exists, it suffices that for every other stable expansion there is some total ordering of the prioritization using which preferredness is achieved. Hence the difference between these two preference notions is that of $\forall\exists$ and $\exists\forall$.

Definition 4. Let E be a stable expansion of Σ and $P = \langle\Phi, \mathcal{P}\rangle$ a prioritization. Then E is *P-maximal* if and only if for all stable expansions E' of Σ, there is a strict total order $\mathcal{T} \supseteq \mathcal{P}$ such that

$$\text{for all } \phi \in \Phi(\phi \in E\backslash E' \text{ implies there is } \psi \text{ such that } \psi\mathcal{T}\phi, \psi \in E'\backslash E).$$

With this preference notion the translation of OTPs to prioritized autoepistemic logic is almost like that of preferred subtheories. A formula ϕ of an OTP is translated to formulae in $\{\neg L\neg\nu \rightarrow \nu|\phi \breve{\models} \nu\}$ where each ν is a *natural consequence* of ϕ. Ryan's idea in defining natural consequences is the conjunctivity in formulae. Ryan can be seen – like Brewka – as constructing maximal consistent subsets of the formulae in an OTP except that if e.g. $p \wedge q$ cannot be consistently included in the set, some of its natural consequences – e.g. p or q – possibly can. Each formula has an infinite number of natural consequences, but Ryan [1992] conjectures that the conjuncts of a suitable conjunctive normal form of ϕ can be used instead of them. Complete proofs of these correspondences are given in [Rintanen, 1993].

Hierarchic autoepistemic logic [Konolige, 1989] and multi-agent autoepistemic logic of Toyama et al. [1991] support the expression of priorities, and do not make syntactic restrictions to the form of defaults. In hierarchic autoepistemic logic the introspection ability of an autoepistemic agent is restricted. Sets of formulae are divided into a number of layers. On layer n, the formula $L_m\phi$ ($m < n$) refers to believing ϕ on layer m. This mechanism effectively resolves conflicts between defaults: there is only one expansion for each layered set of formulae. A drawback of the logic is that priorities are obligatory: if conflicting defaults are on the same layer, an inconsistency results. This logic can be embedded in ordinary autoepistemic logic [Przymusinska, 1989]. Unlike Konolige's logic, our logic allows the expression of partial priorities. Partiality implies the possibility of several preferred stable expansions. Multi-agent autoepistemic logic generalizes autoepistemic logic to several agents. Each agent can access the beliefs of other agents. Toyama et al. demonstrate how specificity in inheritance reasoning can be conveniently represented in this logic. Priorities may be partial, in which case conflicts between defaults produce several expansions. The suitability of the multi-agent logic of Toyama et al. for other uses of priorities than expressing specificity is open. Multi-agent autoepistemic logic can be translated into Moore's autoepistemic logic [Rintanen, 1993].

The logics of Baader and Hollunder [1993] and Brewka [1992] bring priorities to default logic and allow prerequisites in default rules. These logics cannot be easily translated into prioritized autoepistemic logic. Our logic works properly for the example discussed in the introduction.

Example 2. The example given in the introduction can be expressed as formulae of autoepistemic logic as follows.

$$\Sigma = \{a, Lb \wedge \neg L\neg c \rightarrow c, La \wedge \neg Lc \rightarrow \neg c, La \wedge \neg L\neg b \rightarrow b\}$$

Three schemes of using priorities in our logic are obvious. Either give priorities on the justifications, on the conclusions, or on whole application conditions (prerequisites and justifications) of the defaults. That is, the above defaults can be prioritized in the order indicated using either $P = \langle \Phi, \mathcal{P} \rangle$ where $\Phi = \{\neg c, c, \neg b\}$ and $\mathcal{P}$ totally orders Φ in the order $\neg c, c, \neg b$, or P' that orders the formulae $+c, +\neg c, +b$ in this order, or P'' that orders the formulae $+(b \wedge \neg L\neg c), +(a \wedge \neg Lc), +(a \wedge \neg L\neg b)$ in this order. There are two stable expansions for the above set Σ.

$$E_1 = \{a, b, c, \neg L\neg c, \ldots\} \qquad E_2 = \{a, b, \neg c, \neg Lc, \ldots\}$$

According to our definition, E_1 is P-preferred and P'-preferred and P''-preferred, like expected, but E_2 is not. This is opposite to the result given by systems of Baader and Hollunder [1993] and Brewka [1992].

The reason for the unintuitive conclusions in logics of [Baader and Hollunder, 1993; Brewka, 1992] is that the applicability of high-priority rules may depend on the temporally prior application of lower-priority rules. In these logics, the order in which defaults become applicable is a significant factor in determining which extensions the priorities select. Our definition of prioritized autoepistemic logic does not depend on such procedural aspects of application of defaults.

4 Automated reasoning with priorities

In this section we give procedures for automated reasoning with priorities. Our procedures use *full sets* developed by Niemelä for handling stable expansions computationally. Alternatively we could use similar definitions by Shvarts [1990]. The relevant subset of a stable expansion of a set Σ is the set $Sf^L(\Sigma)$ of subformulae of Σ that begin with the L operator. The set $Sf^{qL}(\Sigma)$ is the subset of $Sf^L(\Sigma)$ consisting of those subformulae of Σ that begin with L and are not inside another L in Σ. The following definitions and theorems are from [Niemelä, 1990].

Definition 5. A set Λ is Σ-full if it satisfies the following conditions.

$$\begin{array}{l} \Lambda \subseteq Sf^L(\Sigma) \cup \{\neg L\chi | L\chi \in Sf^L(\Sigma)\} \\ \text{for all } L\chi \in Sf^L(\Sigma),\ L\chi \in \Lambda \text{ iff } \Sigma \cup \Lambda \models \chi \\ \text{for all } L\chi \in Sf^L(\Sigma),\ \neg L\chi \in \Lambda \text{ iff } \Sigma \cup \Lambda \not\models \chi \end{array}$$

Theorem 6. *For a set of sentences Σ there is a bijective mapping from the Σ-full sets to the stable expansions of Σ.*

Definition 7. Given a set of sentences Σ and a sentence ϕ

$$\Sigma \models_L \phi \text{ iff } \Sigma \cup SB_\Sigma(\phi) \models \phi$$

where $SB_\Sigma(\phi) = \{L\chi \in Sf^{qL}(\phi) | \Sigma \models_L \chi\} \cup \{\neg L\chi | L\chi \in Sf^{qL}(\phi), \Sigma \not\models_L \chi\}$.

Theorem 8. *Let Λ be a Σ-full set. Then $\Delta = \{\phi | \Sigma \cup \Lambda \models_L \phi\}$ is the unique stable expansion of Σ for which $\Lambda \subseteq \{L\phi | \phi \in \Delta\} \cup \{\neg L\phi | \phi \notin \Delta\}$.*

Definition 9. Let Σ be a set of formulae and Λ a Σ-full set. Define $SE_\Sigma(\Lambda) = \{\phi | \Sigma \cup \Lambda \models_L \phi\}$.

This way of representing stable expansions immediately suggests a decision procedure: generate all candidate full sets, test whether they are indeed full, and if they are, test the membership of a formula in the respective stable expansion.

The number of candidate full sets for a set of formulae Σ is exponential on the size of Σ: there are $2^{|Sf^L(\Sigma)|}$ such sets. The problem that arises in these decision procedures is how to effectively reduce the size of the search space formed by the candidate full sets. In the stratified case [Gelfond, 1987; Marek and Truszczyński, 1991] there is an efficient way of reducing this search space as shown in [Niemelä and Rintanen, 1992]. In fact, search is completely avoided.

With prioritized autoepistemic reasoning, an obvious algorithm for automated reasoning is to construct all stable expansions (their full sets), test the preferredness condition for each of them, and then test the membership of a formula in the preferred ones. This however is inefficient, and can be improved at least in cases where the prioritizations are on formulae ϕ for which $L\phi \in Sf^L(\Sigma)$. The procedures that use full sets (or something similar) produce search trees where the edges of the tree correspond to the inclusion of a formula $L\phi$ or $\neg L\phi$ in the full set, and the leafs of the tree correspond to candidate full sets, i.e. the set of formulae on the path from the root of the tree to a leaf. The idea is to use priorities for pruning the search tree. Roughly, the basic reduction step is to ignore a subtree T of the search tree that corresponds to full sets containing $L\phi$ (respectively $\neg L\phi$ for a *positive* ϕ), whenever there was a full set containing $\neg L\phi$ (respectively $L\phi$) and all full sets corresponding to T contain exactly the same formulae $L\psi$ or $\neg L\psi$ for formulae ψ that are more important than ϕ. In this case, the subtree T is guaranteed *not to* contain full sets of preferred stable expansions.

It is essential for our procedures that the order in which full sets are generated in our algorithms fulfills the following: the case $L\phi \in \Lambda$ is considered before the case $\neg L\phi \in \Lambda$ for positive formulae $+\phi$, and the case $\neg L\phi \in \Lambda$ before the case $L\phi \in \Lambda$ for negative formulae ϕ. Two subprocedures and the main procedure of our algorithms are shown in Figure 1. For the clarity of presentation, the handling of positive formulae $+\phi$ in prioritizations is not discussed. The modifications required for them are straightforward. The reason for giving them a separate treatment in the procedures instead of using them as a shorthand for $\neg L\phi$, is that this way the constraint $L\phi \in Sf^L(\Sigma)$ works symmetrically for positive $+\phi$.

The main procedure traverses the search space of $2^{|Sf^L(\Sigma)|}$ full sets of a set Σ. The pruning of the search tree is performed by the procedure *extendible*, and the order in which the formulae are chosen in full sets is determined by the procedure *next*. In the main procedure, the variable Q contains the full sets of all preferred stable expansions found so far. The procedure *next* selects elements ϕ of $\Phi \subseteq \{\phi | L\phi \in Sf^L(\Sigma)\}$ in some total order that extends the relation $\mathcal{P}$ in the prioritization P. This total order fulfills a further condition that is taken advantage of in the procedure *extendible*. At each node of the search tree the procedure *extendible* is invoked to detect whether the full sets to be found in the respective subtree can correspond to preferred stable expansions. The procedure attempts to construct a strict total order $\mathcal{T} \supseteq \mathcal{P}$ so that the preferredness condition of Definition 2 would be fulfilled for the stable expansions of the current subtree. It turns out, that for the fulfillment of the condition it suffices to look at the preferred stable expansions found so far (the full sets in Q). The possibility that a preferred stable expansion found later would be "better" than the stable expansions corresponding to the current subtree is ruled out by the way the procedure *next* selects elements: the order in which elements $\phi, L\phi \in Sf^L(\Sigma)$ are chosen is acceptable as the strict total order $\mathcal{T}$. Hence, all full sets Λ found later contain $L\phi$ for the most important formula ϕ (according to $\mathcal{T}$) in which Λ differs from the full sets of the current subtree, and all full sets of the current subtree contain $\neg L\phi$. Therefore the stable expansions of the current subtree are no "worse" than those found later.

We briefly discuss properties of our algorithm. First, priorities are usefully taken advantage of when they are on formulae ϕ such that $L\phi \in Sf^L(\Sigma)\}$ [4]: the number of candidate subsets considered is smaller than in decision procedures of Moore's autoepistemic logic [5]. In cases where prioritizations are strict total

[4] Extending Σ with tautologies $L\phi \rightarrow L\phi$ for formulae ϕ such that $L\phi \notin Sf^L(\Sigma)$ makes our algorithm applicable for all prioritizations.

[5] Improvements to the trivial algorithms that use full sets, as proposed by Niemelä [1994], can be easily incorporated in our algorithm.

```
PROCEDURE extendible(P, Λ, Q)
BEGIN
  ⟨Φ, 𝒫⟩ := P;
  IF Φ = ∅ THEN RETURN true;
  A := {φ ∈ Φ | φ is 𝒫-minimal in Φ, Lφ ∈ Λ implies Lφ ∈ ⋂Q};
  IF A = ∅ THEN RETURN false;
  φ := any member of A;
  P' := ⟨Φ\{φ}, 𝒫 ∩ (Φ\{φ} × Φ\{φ})⟩;
  IF Lφ ∈ Λ THEN RETURN extendible(P', Λ, Q)
  ELSE RETURN extendible(P', Λ, {Λ' ∈ Q | ¬Lφ ∈ Λ'})
END

PROCEDURE next(P, Σ, Λ, Q)
BEGIN
  ⟨Φ, 𝒫⟩ := P;
  A := {φ ∈ Φ | φ is 𝒫-minimal in Φ, Lφ ∈ Λ implies Lφ ∈ ⋂Q};
  IF A = ∅ THEN RETURN any φ such that Lφ ∈ Sf^L(Σ)\Sf^{qL}(Λ);
  B := {φ ∈ A | Lφ ∉ Λ, ¬Lφ ∉ Λ};
  IF B ≠ ∅ THEN RETURN any member of B;
  φ := any member of A;
  P' := ⟨Φ\{φ}, 𝒫 ∩ (Φ\{φ} × Φ\{φ})⟩;
  IF Lφ ∈ Λ THEN RETURN next(P', Σ, Λ, Q)
  ELSE RETURN next(P', Σ, Λ, {Λ' ∈ Q | ¬Lφ ∈ Λ'})
END

PROCEDURE decide(P, Σ, Λ, φ, Q)
BEGIN
  IF extendible(P, Λ, Q) = false THEN RETURN (∅,false);
  IF for some ¬Lχ ∈ Λ, Σ ∪ Λ ⊨ χ THEN RETURN (∅,false);
  IF Sf^L(Σ) ⊆ Sf^{qL}(Λ) THEN
    IF for all Lχ ∈ Λ, Σ ∪ Λ ⊨ χ THEN RETURN ({Λ},test(Σ, Λ, φ))
    ELSE RETURN (∅,false);
  χ := next(P, Σ, Λ, Q);
  (S, b) := decide(P, Σ, Λ ∪ {¬Lχ}, φ, Q);
  (S', b') := decide(P, Σ, Λ ∪ {Lχ}, φ, S ∪ Q);
  RETURN (S ∪ S', b or b')
END
```

Fig. 1. The procedure for prioritized autoepistemic logic

orders on $\{\phi | L\phi \in Sf^L(\Sigma)\}$, our algorithm finds the unique stable expansion and all subsequent computation is avoided. Second, more trivial ways of using priorities in the computation may require considering $n!$ strict total orders for a prioritization with n elements. The definition of Brewka's preferred subtheories [1989] and priorities in default logic [1992] suggest such algorithms. In our algorithms the consideration of $\mathcal{O}(2^n)$ different cases suffices. This seems to be the case also with the algorithm for default logic by Baader and Hollunder [1993].

The running time of our algorithm is exponential on the size of Σ for two reasons: the classical reasoning component (e.g. for propositional logic) takes exponential time, and there may be an exponential number of candidate full sets despite the reduction due to priorities. In the special case of a tractable subset of a classical logic (e.g. propositional Horn clauses), autoepistemic formulae of the form $\neg L\neg\phi \rightarrow \phi$ (prerequisite-free normal defaults), and strict total prioritizations on $\{\phi|L\phi \in Sf^L(\Sigma)\}$, both these causes for exponentiality disappear. The unique preferred stable expansion can be immediately found without search, and the classical theorem-prover – that runs in polynomial time – is called only a polynomial number of times, hence resulting in a tractable decision procedure.

The correctness proof of the algorithm uses the following definitions and lemmata that are proved in [Rintanen, 1993] (In the definitions and lemmata Σ is a finite set of formulae, and P is a finite prioritization).

Definition 10 Preferred in. A stable expansion E of Σ is P*-preferred in* a set of stable expansions $X = \{E_1, \ldots, E_n\}$ of Σ if and only if there is a strict total order $\mathcal{T} \supseteq \mathcal{P}$ on Φ such that for all $E' \in X$,

$$\text{for all } \phi \in \Phi(\phi \in E \backslash E' \text{ implies there is } \psi \in E' \backslash E \text{ such that } \psi \mathcal{T} \phi).$$

Lemma 11. *Let X be a set of stable expansions of Σ such that each $E \in X$ is P-preferred in X, and all P-preferred stable expansions of Σ are in X. Then X is exactly the set of P-preferred stable expansions of Σ.*

***Lemma 12 Procedure* extendible.** *Let $P = \langle \Phi, \mathcal{P} \rangle$ be a prioritization such that $\Phi \subseteq \{\phi|L\phi \in Sf^L(\Sigma)\}$, Λ a Σ-full set, and Q a finite set of Σ-full sets. Let b be the value returned by the procedure call extendible(P, Λ, Q). Then b is* true *if and only if $SE_\Sigma(\Lambda)$ is P-preferred in $SE_\Sigma(Q)$.*

Let Λ' be a subset of a Σ-full set Λ and let the procedure call extendible(P, Λ', Q) return false. *Then $SE_\Sigma(\Lambda)$ is not P-preferred in $SE_\Sigma(Q)$.*

***Lemma 13 Procedure* next.** *Let $P = \langle \Phi, \mathcal{P} \rangle$ be a prioritization such that $\Phi \subseteq \{\phi|L\phi \in Sf^L(\Sigma)\}$, Λ a consistent subset of $Sf^L(\Sigma) \cup \{\neg L\phi|L\phi \in Sf^L(\Sigma)\}$ such that $Sf^L(\Sigma) \backslash Sf^{qL}(\Lambda) \neq \emptyset$, and Q a finite set of Σ-full sets.*

Then next(P, Σ, Λ, Q) returns a formula χ such that $L\chi \in Sf^L(\Sigma) \backslash Sf^{qL}(\Lambda)$ and the following holds. If S is a set of Σ-full sets $\Lambda' \supseteq \Lambda \cup \{\neg L\chi\}$ such that each $SE_\Sigma(\Lambda')$ is P-preferred in $SE_\Sigma(S \cup Q)$, and S' is a set of Σ-full sets $\Lambda' \supseteq \Lambda \cup \{L\chi\}$ such that each $SE_\Sigma(\Lambda')$ is P-preferred in $SE_\Sigma(S \cup S' \cup Q)$, then each $SE_\Sigma(\Lambda'), \Lambda' \in S$ is P-preferred in $SE_\Sigma(S \cup S' \cup Q)$.

***Lemma 14 Procedure* decide.** *Let $P = \langle \Phi, \mathcal{P} \rangle$ be a prioritization such that $\Phi \subseteq \{\phi|L\phi \in Sf^L(\Sigma)\}$, Λ a consistent set such that $\Lambda \subseteq (Sf^L(\Sigma) \cup \{\neg L\phi|L\phi \in Sf^L(\Sigma)\})$, $\phi \in \mathcal{L}_{ae}$ a formula, and Q a finite set of Σ-full sets.*

The procedure call decide($P, \Sigma, \Lambda, \phi, Q$) returns (S,b). The set S consists of Σ-full sets $\Lambda' \supseteq \Lambda$ such that each $SE_\Sigma(\Lambda')$ is P-preferred in $SE_\Sigma(S \cup Q)$, and each P-preferred stable expansion E of Σ such that $E = SE_\Sigma(\Lambda')$ for some $\Lambda' \supseteq \Lambda$, is contained in S. The value b is true *if and only if for some $\Lambda' \in S$, test(Σ, Λ', ϕ) returns* true.

Theorem 15 Correctness. *Let $\phi \in \mathcal{L}_{ae}$ be a formula. The procedure call decide($P, \Sigma \cup \{L\phi \rightarrow L\phi | \phi \in \Phi, L\phi \notin Sf^L(\Sigma)\}, \emptyset, \phi, \emptyset$) returns (S, b), where b is* true *if and only if there is a P-preferred stable expansion E of Σ such that $E = SE_\Sigma(\Lambda)$ and the procedure call test(Σ, Λ, ϕ) returns* true.

Proof. The idea in extending Σ to $\Sigma' = \Sigma \cup \{L\phi \rightarrow L\phi | \phi \in \Phi, L\phi \notin Sf^L(\Sigma)\}$ is to fulfill the requirement that for each formula $\phi \in \Phi$, $L\phi \in Sf^L(\Sigma)$. The procedures *next*, *extendible* and *decide* depend on this requirement. By Lemma 14 $SE_{\Sigma'}(S)$ contains all P-preferred stable expansions. Because every member of $SE_{\Sigma'}(S)$ is P-preferred in $SE_{\Sigma'}(S)$, every member of $SE_{\Sigma'}(S)$ is P-preferred by Lemma 11. The value of b is *true* if and only if for some stable expansion $SE_{\Sigma'}(\Lambda), \Lambda \in S$ test(Σ', Λ, ϕ) returns *true*.

Procedures for different decision problems are obtained by supplying different procedures *test*. For prioritized brave reasoning the procedure *test* simply tests $\Sigma \cup \Lambda \models_L \phi$, and for prioritized cautious reasoning it tests $\Sigma \cup \Lambda \not\models_L \phi$ and the value returned by *decide* is negated. We have implemented all these procedures in an automatic theorem-proving system for autoepistemic and default logics.

Decision procedures for several nonmonotonic modal logics ("nonmonotonic versions" of N, K, T, S4, S4F, KD45, SW5, W5) are given in [Marek *et al.*, 1993]. These procedures use the finite characterization of expansions developed by [Shvarts, 1990]. Similar procedures based on full sets – with some improvements – for Moore's autoepistemic logic and enumeration-based autoepistemic logics are given in [Niemelä, 1994]. Our techniques can be applied in computing expansions for the prioritized versions of these logics. Using prioritized versions of any of the nonmonotonic logics N, K, T, S4, S4F or the L-hierarchic autoepistemic logic of Niemelä [1994] we can avoid the groundedness problems that distinguish default logic from Moore's autoepistemic logic. Using the translation $L\alpha \wedge L\neg L\neg\beta \rightarrow \gamma$ for defaults $\frac{\alpha:\beta}{\gamma}$ [Truszczyński, 1991], a version of prioritized default logic can be embedded in any of these logics.

5 Conclusions

In this work we have presented a formalization of priorities within autoepistemic logic. The work generalizes earlier work on priorities and nonmonotonicity, e.g. [Brewka, 1989; Ryan, 1992], by allowing unrestricted defaults. Comparable generalizations have been proposed by Brewka [1992] and Baader and Hollunder [1993]. We believe that because of the declarativity of our priority notion unintuitive conclusions that are likely to arise in more procedural approaches to default priorities [Baader and Hollunder, 1993; Brewka, 1992; MacNish, 1991; Tan and Treur, 1992] are avoided in our system. For example, logics of Baader and Hollunder as well as Brewka prefer some defaults on the basis that their prerequisites are more directly derivable. In some cases, this preference takes precedence over preference indicated by priorities, as demonstrated by the example in the introduction.

Automated nonmonotonic reasoning with priorities has been investigated earlier in [Baker and Ginsberg, 1989; Junker and Brewka, 1991]. Both approaches essentially restrict to prerequisite-free normal defaults. Baker and Ginsberg work only with layered partial orders. Our work does not make these restrictions. Reasoning with priorities can be more efficient than without. We have demonstrated this for an important class of priorities that are on formulae that occur inside L in the premises: priorities justify a principle for pruning search trees in decision procedures.

Through translations of autoepistemic logic to other formalisms, our priorities and the associated reasoning procedures can be brought to e.g. default logic, justification-based TMSs, and the theory of diagnosis of Reiter [1987].

Future investigations on priorities concern the computation of priorities automatically. An important source of priorities is rule specificity. A method for computing priorities for conditional entailment was presented by Geffner and Pearl [1992]. However, their method involves several prioritizations instead of only one, and hence these priorities are not convenient for automated reasoning. Default rules expressible in autoepistemic logic are more general than in conditional entailment, and Geffner and Pearl's method does not immediately generalize. The problem of rule specificity in the general context of default rules is a challenging subject for further research.

Acknowledgements

I would like to thank Dr I. Niemelä for advice, discussions, and for proposing this research subject.

References

Baader, F. and B. Hollunder: How to prefer more specific defaults in terminological default logic. In Ruzena Bajcsy, editor, *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, volume 1, pages 669–674, Chambery, France, 1993.

Baker, A. B. and M. L. Ginsberg: A theorem prover for prioritized circumscription. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pages 463–467, Detroit, 1989. Morgan Kaufmann Publishers.

Brewka, G.: Preferred subtheories: an extended logical framework for default reasoning. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pages 1043–1048, Detroit, 1989. Morgan Kaufmann Publishers.

Brewka, G.: Specificity in default logic and its application to formalizing obligation. Unpublished, 1992.

Geffner, H. and Judea Pearl: Conditional entailment: bridging two approaches to default reasoning. *Artificial Intelligence*, 53:209–244, 1992.

Gelfond, M.: On stratified autoepistemic theories. In *Proceedings of the 6th National Conference on Artificial Intelligence*, pages 207–211, Seattle, July 1987. American Association for Artificial Intelligence.

Gottlob, G.: Complexity results for nonmonotonic logics. *Journal of Logic and Computation*, 2(3):397–425, June 1992.

Junker, U. and G. Brewka: Handling partially ordered defaults in TMS. In *Symbolic and quantitative approaches to uncertainty. European Conference*, number 548 in Lecture Notes in Computer Science, pages 211–218, Marseille, 1991. Springer-Verlag.

Konolige, K.: Hierarchic autoepistemic theories for nonmonotonic reasoning: Preliminary report. In *Proceedings of the 2nd Workshop on Non-Monotonic Reasoning*, number 346 in Lecture Notes in Artificial Intelligence, pages 42–59, Grassau, Germany, June 1989. Springer-Verlag.

Kraus, S., D. Lehmann, and M. Magidor: Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

Lifschitz, V.: Computing circumscription. In A. Joshi, editor, *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, pages 121–127, Los Angeles, August 1985. Morgan Kaufmann Publishers.

MacNish, C.: Hierarchical default logic. In *Symbolic and quantitative approaches to uncertainty. European Conference*, number 548 in Lecture Notes in Computer Science, pages 246–253, Marseille, 1991. Springer-Verlag.

Marek, W. and M. Truszczyński: Autoepistemic logic. *Journal of the ACM*, 38:588–619, 1991.

Marek, V. W., G. F. Schwarz, and M. Truszczyński: Modal nonmonotonic logics: ranges, characterization, computation. *Journal of the ACM*, 40(4):963–990, September 1993.

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

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

Niemelä, I. and J. Rintanen: On the impact of stratification on the complexity of nonmonotonic reasoning. In B. Nebel, C. Rich, and W. Swartout, editors, *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning*, pages 627–638, Cambridge, MA, October 1992. Morgan Kaufmann Publishers.

Niemelä, I.: A decision method for nonmonotonic reasoning based on autoepistemic reasoning. In *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484, Bonn, May 1994.

Niemelä, I.: Towards automatic autoepistemic reasoning. In J. Van Eijck, editor, *Proceedings of the European Workshop on Logics in Artificial Intelligence—*

JELIA'90, number 478 in Lecture Notes in Artificial Intelligence, pages 428–443, Amsterdam, September 1990. Springer-Verlag.

Przymusinska, H.: The embeddability of hierarchic autoepistemic logic in autoepistemic logic. In Z. W. Ras, editor, *Methodologies for Intelligent Systems, 4: Proceedings of the Fourth International Symposium on Methodologies for Intelligent Systems*, pages 485–493, New York, 1989. North-Holland.

Reiter, R. and G. Criscuolo: On interacting defaults. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pages 270–276, 1981.

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

Reiter, R.: A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–95, 1987.

Rintanen, J.: Priorities and nonmonotonic reasoning. Research report A 28, Helsinki University of Technology, Digital Systems Laboratory, December 1993.

Ryan, M.: Representing defaults as sentences with reduced priority. In B. Nebel, C. Rich, and W. Swartout, editors, *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning*, pages 649–660, Cambridge, MA, October 1992. Morgan Kaufmann Publishers.

Shoham, Y.: Nonmonotonic logics: meaning and utility. In D. McDermott, editor, *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 388–393, Milano, 1987.

Shvarts, G.: Autoepistemic modal logics. In *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 97–109, Pacific Grove, CA, March 1990. Morgan Kaufmann Publishers.

Tan, Y. H. and J. Treur: Constructive default logic and the control of defeasible reasoning. In B. Neumann, editor, *Proceedings of the 10th European Conference on Artificial Intelligence*, pages 299–303, Vienna, 1992. John Wiley & Sons.

Touretzky, D. S., J. F. Horty, and R. H. Thomason: A clash of intuitions: the current state of nonmonotonic multiple inheritance systems. In D. McDermott, editor, *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 476–482, Milano, August 1987. Morgan Kaufmann Publishers.

Toyama, K., Y. Inagaki, and T. Fukumura: Knowledge representation based on autoepistemic logic for multiple agents. Technical report 90-1-06, Chukyo University, School of Computer and Cognitive Sciences, March 1991.

Truszczyński, M.: Embedding default logic into modal nonmonotonic logics. In *Proceedings of the 1st International Workshop on Logic Programming and Non-monotonic Reasoning*, pages 151–165, Washington, DC, July 1991. The MIT Press.

This article was processed using the LaTeX macro package with LLNCS style