Preferred Subtheories:
An Extended Logical Framework for Default Reasoning

Gerhard Brewka
Gesellschaft fur Mathematik und Datenverarbeitung
Postfach 12 40, D-5305 Sankt Augustin, Fed. Rep. of Germany

Abstract

We present a general framework for defining nonmonotonic systems based on the notion of preferred maximal consistent subsets of the premises. This framework subsumes David Poole's THEORIST approach to default reasoning as a particular instance. A disadvantage of THEORIST is that it does not allow to represent priorities between defaults adequately (as distinct from blocking defaults in specific situations). We therefore propose two generalizations of Poole's system: in the first generalization several layers of possible hypotheses representing different degrees of reliability are introduced. In a second further generalization a partial ordering between premises is used to distinguish between more and less reliable formulas. In both approaches a formula is provable from a theory if it is possible to construct a consistent argument for it based on the most reliable hypotheses. This allows for a simple representation of priorities between defaults.

1. Introduction

Intelligent agents have to be able to draw plausible conclusions based on incomplete information, to handle rules with exceptions and to deal with inconsistent information. Classical logic has not much to offer with respect to all of these problems. This was the motivation for the various attempts to define nonmonotonic logics. A variety of approaches have been proposed (Moore 85) (McCarthy 84) (Rciter 80) and their mathematical properties as well as their relative expressiveness and computational aspects have been studied intensively within the last ten years.

The "standard" approaches to formalize nonmonotonic and in particular default reasoning start from a consistent set of premises (otherwise no interesting result at all is obtained) and extend the inference relation to get more than just the classically derivable formulas. Technically, this can, for instance, be achieved by the addition of a second order formula (McCarthy 84) or by the introduction of non-standard inference rules (Reiter 80).

In this paper we will present an approach based on an alternative view. What makes a default a default? What distinguishes it from a fact? Certainly our altitude towards it in case of a conflict, i.e. an inconsistency. If we take this view serious then the idea of default reasoning as a special case of inconsistency handling seems quite natural.[1] There is no problem with inconsistent premises as long as we provide ways to handle the inconsistency adequately (in other words, if we modify the inference relation such that in case of an inconsistency fewer, i.e. not all formulas are derivable). As we will show in this paper, it is possible to specify strategies for inconsistency handling which can-be used for default reasoning.

In the rest of the paper wc will first present a simple general framework for defining nonmonotonic systems. Sect. 3 shows how Poole's approach to default reasoning (Poole 88) fits into this framework and discusses the limitations of his approach which arc due to the inability of representing priorities between defaults. Sect. 4 presents a generalization of Poole's approach which introduces several layers of possible hypotheses representing different degrees of reliability. A second further generalization based on a partial ordering between premises is described in Sect. 5. In both approaches a formula is provable from a theory if it is possible to construct a consistent argument for it based on the most reliable hypotheses. Sect. 6, then, discusses related work.

2. A Framework for Nonmonotonic Systems

A standard way of handling inconsistencies uses maximal consistent subsets of the formulas at hand. The idea behind the "maximal" is clear: we want to modify the available information as few as possible. The notion of maximal consistent subsets per se, however, docs not allow to express, say, that *Tweety flies* should be given up instead of *Tweety is a penguin,* if we know that penguins don't fly. To be able to express such preferences we have to consider not all maximal consistent subsets, but only some of them,

[1] This idea has also been proposed in (Bibel 85).

the <u>preferred</u> maximal consistent subsets, or simpler: *preferred subtheories.*

The notion of a preferred maximal consistent subset is not new: it dates back to (Rescher 64). Rescher has defined a specific ordering of subtheories which will briefly be discussed in Sect 6. The relevance of this idea for default reasoning, however, has - as far as we know - been overlooked so far.

We are now in a position to define a weak and a strong notion of provability:

> A formula p is *weakly provable* from T iff there is a preferred subtheory S of T such that S I- p.
>
> A formula p is *strongly provable* from T iff for all preferred subtheories S of T we have S I- p.

These notions, roughly, correspond to containment in at least one or in all extensions in those approaches to default reasoning which generate multiple extensions in case of conflicting evidence, e.g. Reiter's. Of course, it remains to define what the preferred subtheories are. We will first show how Poole's system can be obtained.

3. Poole's Approach

David Poole (Poole 88) recently presented an approach to default reasoning based on hypothetical reasoning. In Poole's framework it is assumed that the user provides

1) a set F of closed formulas, the facts about the world,
2) a set Δ of, possibly open, formulas, the possible hypotheses.

A *scenario* of F and A then is a set $D \cup F$ where D is a set of ground instances of elements of Δ such that $D \cup F$ is consistent.

g is *explainable* from F and A iff there is a scenario of F and A which implies g.

An *extension* of F and A is the set of logical consequences of a (set inclusion) maximal scenario of F and A.

Poole's system is equivalent to Reiter's default logic with the restriction to normal defaults without prerequisites. These restrictions seem, at first view, too drastic. However, Poole is able to show that many of the standard default reasoning examples from the literature can adequately be dealt with in his simple and elegant approach. In particular, he introduces names for defaults (hypotheses) in the following way: for a formula w(x) ε A with free variables x he introduces a new predicate symbol p_w of the same arity. Poole shows that w(x) can equivalently be replaced by $p_w(x)$, if the formula $\forall x. p_w(x) \supset w(x)$ is added to F. Poole uses the notation Pw(x):w(x) as an abbreviation for that case.

The use of names allows to block the applicability of a default when needed. If we want a default pw(x) to be inapplicable in situation s we simply have to add $\forall x.\ s \supset$ $-p_w(x)$ to our facts. (Poole 88) contains many nice examples of how this technique can be used.

It is easy to sec how Poole's approach can be obtained as one particular instance of our preferred subtheory framework: if we define the preferred subtheories of A' U F (A' is obtained from A by replacing open formulas by all of their ground instances) as those containing F, then weak provability and Poole's explainability coincide.

Poole's approach is simple and elegant, and its expressiveness is astonishing. Moreover, an efficient Prolog-implementation exists (Poole et al. 86). There seems to be an important drawback, however: it is possible to block the applicability of defaults in certain circumstances, but there is no way to express priorities between defaults. We use an example due to Ulrich Junker to illustrate this problem. Assume we have the following commonsense facts:

> *Usually one has to go to a project meeting.*
>
> *This rule does not apply if somebody is sick, unless he only has a cold.*
>
> *The rule is also not applicable if somebody is on vacation.*

In Reiter's default logic (Reiter 80) we can use the following defaults and formulas to represent these facts:

<u>:MR1 A MEETING</u>
1) MEETING

<u>SICK:M-COLD A-IRI</u>
2) -RI

3) VACATION ⊃ -nRI

4) COLD ⊃ SICK

In Poole's system we need, besides formulas 3) and 4), a default

1) RI: MEETING

together with the fact

2) SICK ⊃ -nRI

which blocks the applicability of RI when SICK is known. This blocking cannot be achieved by another default: if we would choose to have Ri: SICK z —iRI instead, then, given SICK, two extensions were generated, one containing -iRi.

As a consequence we have to introduce a <u>new</u> default

5) R2: COLD ⊃ MEETING

to achieve the desired behavior. But this is not sufficient. As a side effect of the inability to use defaults to block defaults we need another fact: since we want to stay home on vacation even if we have a cold, we have to block the applicability of R2 in this case, i.e. we further need

6) VACATION ⊃ ¬R2

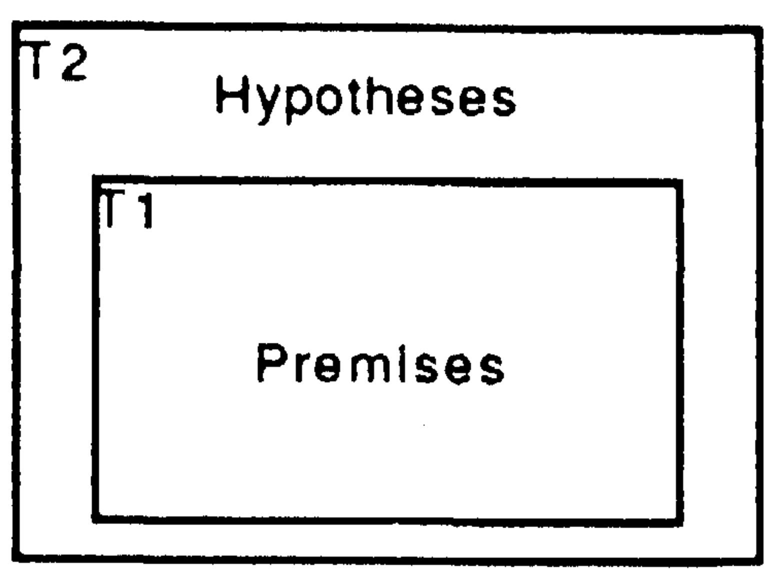
This seems unpleasant, since we have to look "down" in the hierarchy of exceptions and block defaults lower in the hierarchy. It is not difficult to imagine that the number of needed defaults may increase heavily in cases where more exceptions and exceptions of exceptions are involved.

The inability to use defaults to block other defaults seems to be the heart of the problem. It is possible to block a default's applicability in Poole's system, e.g. the default *birds fly* can be blocked for penguins. But it is not possible to express that default dl should have priority over a conflicting default d2 in the sense that d2 is not applicable if dl can be applied. Adding the fact dl ⊃ -,d2 does not help, this is equivalent to d2 ⊃ -,dl. Also the constraint technique proposed by Poole to prevent unwanted consequences of contraposition does not help (see (Poole 88) for the details): adding a default, say d3: dl ⊃ ->d2 together with the constraint d2 ⊃-,d3 to prevent the use of the contrapositive leads to the same problems: d2 still can be applied and its application then blocks dl and d3.

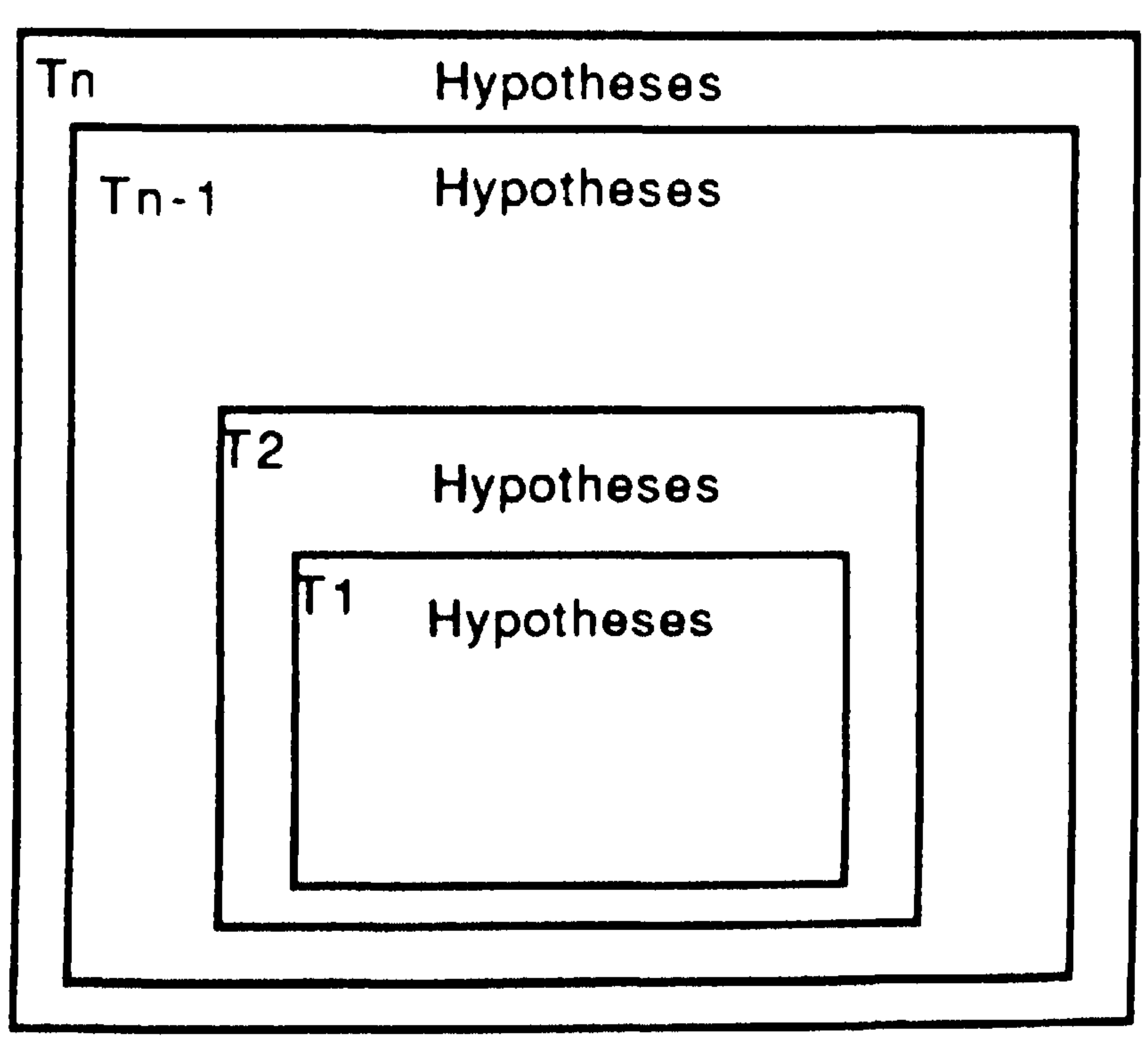
This inability to represent default priorities in Poole's system is the motivation for the generalizations presented in the next sections.

4. First Generalization

The following picture illustrates the basic idea of Poole's approach: we have two levels of theories, the basic level can be seen as premises which must hold (and be consistent), the second level is a level of hypotheses which are less reliable.



We generalize these ideas in two respects. First, we do not require the most reliable formulas (i.e. Tl) to be consistent. In our approach every formula is in principle refutable. And second, we introduce more than just two levels. This can be illustrated by the following graphic:



The idea is that the different levels of a theory represent different degrees of reliability. The innermost part is the most reliable one. If inconsistencies arise the more reliable information is preferred. Intuitively, a formula is provable if we can construct an argument for it from the most reliable available information. Of course, there may be conflicting information with the same reliability. In this case we get something analogous to the multiple extensions of, e.g. Reiter's default logic, i.e. two contradicting formulas can be provable in a weak sense. The fact that there are no in principle unrefutable "premises" makes it possible to treat all levels uniformly. For instance, we can add to any theory information which is even more reliable than the currently innermost level.

We now show how these intuitive ideas can be made precise in the preferred subtheory approach:

> A *default theory* T is a tuple (T1, ..., Tn), where each Ti is a set of classical first order formulas.

Intuitively, information in Ti is more reliable than that in Tj if i<j. A default like *birds fly* can be represented as the set of all ground instances of a schema Bird(x) ⊃ Flies(x). For sake of simplicity we will write Ti = {..., P(x), ...} if we want to express that Ti contains all ground instances of P(x). Note the important difference between universally quantified formulas and schemata containing free variables. It remains to define the preferred subtheories:

> Let T=(Tl,...,Tn) be a default theory.
> S=S1 ∪...∪ Sn is a *preferred subtheory* of T iff for all k (1 < k < n) Sl ∪...∪ Sk is a maximal consistent subset of Tl ∪ ... ∪ Tk.

In other words, to obtain a preferred subtheory of T we have to start with any maximal consistent subset of Tl, add as many formulas from T2 as consistently can be added (in any possible way), and continue this process for T3,..., Tn.

The following simple examples show how the different levels can be used to express priorities between defaults:

1) Good old Tweety:

T1 = {BIRD(TWEETY), ∀X.PENGUIN(X) ⊃ ¬FLIES(X)}

T2 = {BIRD(X) ⊃ FLIES(X)}

FLIES(TWEETY) is strongly provable.

T1 = {BIRD(TWEETY), ∀X.PENGUIN(X) ⊃ ¬FLIES(X), PENGUIN(TWEETY)}

T2 = {BIRD(X) ⊃ FLIES(X)}

-,FLIBS(TWEETY) is strongly provable. This example also illustrates the importance of the distinction between schemata and universally quantified formulas. If we wouldn't use a schema in T2 but instead a quantified formula, then this formula wouldn't be usable if there is a single nonflying bird.

If there is a penguin who does fly we can use the following representation, in which *penguins don't fly* is given higher priority than *birds fly:*

TI = (BIRD(TWEETY),PENGUIN(TWEETY), PENGUIN(TIM), RLES(T1M)}

T2 = {PENGUIN(X) ⊃ -FLIES(X)}

T3 = {BIRDpC) ⊃ FLIES(X)}

2) Nixon example;

TI = {REP(NIXON), QUAK(NIXON)}

T2 = (REP(X) ⊃ -PAC(X), QUAK(X) ⊃ PAC(X)}

Both PAC(NIXON) and -PAC(NIXON) are weakly provable. None of them is strongly provable. If we want to give priority to - say - *Quakers are Pacifists,* this can be achieved as follows:

TI = (REP(NIXON), QUAK(NIXON)}

T2={QUAK(X) ⊃ PAC(X)}

T3 = (REP(X) ⊃ -PAC(X)}

Now PAC(NIXON) is strongly provable.

3) Meeting example:

In this example we use named defaults. R:Q ∈ Ti stands forR ∈ Ti and R ⊃ Q ∈ TI.

T1 = (VACATION ⊃ -RI, COLD ⊃ -R2, COLD ⊃ SICK}

T2={R2:SICK ⊃ -R1}

T3 = (RI: MEETING}

From the above default theory MEETING is strongly provable. If we add VACATION to TI then MEETING is no longer strongly or weakly provable. The same happens if we add SICK. If, however, we add COLD (without VACATION) then again MEETING is strongly provable. And finally, if both COLD and VACATION are added, then again MEETING is not derivable.

In some applications it is possible to generate the levels of reliability automatically. Assume we want to prefer the most specific information (specific here is understood as strictly, undefeasibly more specific. This, of course, makes matters quite simple). Assume the user provides a consistent set of facts F and a set of open defaults D of the form P(x) ⊃ Q(x) where x may be a tuple of variables. To define theoremhood we translate (F,D) into a default theory (TI, T2, ...,Tn) in the following way:

TI = F

T2 = {P(x) ⊃ Q(x) ∈ D |

there is no R(x) ⊃ Z(x) ∈ D with

F |- ∀x.R(x) ⊃ P(x) and not F |- ∀x.P(x) ⊃ R(x)}

Ti+1 = {P(x) ⊃ Q(x) ∈ D\T2∪...∪Ti |

there is no R(x) ⊃ Z(x) ∈ D\T2∪...∪Ti with

F |- ∀x.R(x) ⊃ P(x) and not F |- ∀x.P(x) ⊃ R(x)}

n is the smallest integer such that T2 ∪...∪ Tn = D.

A simple example:

F= {BIRD(TWEETY),PENGUIN(HANSI), ∀x.PENGUIN(x) ⊃ BIRD(x)}

D={BIRD(x) ⊃ FLIES(x), PENGUIN(x) ⊃ -FUES(x)}

The above translation yields

TI = F

T2 = (PENGUIN(x) ⊃ -FLIES(x)}

T3 = {BIRIXx) ⊃ FLIES(x)}

From this theory -FLIES (HANSI) and FLIES(TWEETY) is strongly provable.

5. Second Generalization

For many problems the introduction of levels of reliability as described above is sufficient to express the necessary priorities between defaults. Sometimes, however, we want to leave open whether a formula p is of more, less or the same reliability as another formula q. Consider the following abstract example:

1) A(x) ⊃ P(x)

2) B(x) ⊃ Q(x)

3) C(x) ⊃ R(x)

Assume P, Q and R arc mutually inconsistent. Moreover, let A be a subclass of B, i.e. information about A is more specific than information about B. We certainly want to give 1) priority over 2) in this case. But how about 3)? The approach from the last section forces us to choose exactly one level for each formula, i.e. to specify a priority cither between 1) and 3) or between 2) and 3). There seems to be no reason why we should want this. This problem can be avoided if we allow the degrees of reliability to be represented via an arbitrary partial ordering of the premises instead of the different levels. Again we have to define the preferred subtheories to obtain weak and strong provability based on such a partial ordering:

> Let < be a strict partial ordering on a (finite) set of premises T. S is a preferred subtheory of T iff there exists a strict total ordering (tl, t2, ... , tn) of T respecting < such that S=Sn with
>
> S0 := { }, and for 0≤i<n
>
> Si+1 := if ti+1 consistent with Si then Si ∪ {ti+1} else Si.

If we define in our above example the ordering to be 1) < 2), then P(a) is strongly provable from A(a) and B(a) ("from some formulas" here means that these formulas are smaller than 1), 2) and 3) with respect to <). From A(a) and C(a) both P(a) and R(a) are weakly provable. From B(a) and C(a) we get weakly Q(a) and R(a).

An obvious application of this approach is the formalization of frame systems with multiple inheritance and a strict subclass hierarchy. Frames can be interpreted as unary predicates, slots as two-place predicates. If a frame F has a slot S with (default) value V, this can be represented as $F(x) \supset S(x,V)$. The ordering < has to be defined such that $F1(x) \supset S(x,V1) < F2(x) \supset S(x,V2)$ whenever F1 is below F2 in the frame hierarchy. This formalization of frames is much simpler than the circumscription based approach in (Brewka 87) and does not need reified predicates.

6. Related Work

As mentioned above, Poole's approach is equivalent to Reiter's default logic restricted to prerequisite-free normal defaults. The relation between our approach and Reiter's is - as expected - less simple. In (Brewka 89) we introduce a modification of default logic, called prioritized default logic (PDL). The relation between PDL and the level approach from Sect. 4 is the same as that between default logic and Poole's system:

Let Di (i=1,...,n) be sets of defaults (in the sense of Reiter), W a set of formulas. E is a *PDL-extension* of T = (D1,...,Dn,W) iff there exist sets of formulas E1,...,En such that

E1 is an extension of (D1,W)

E2 is an extension of (D2, E1)

...

E = En is an extension of (Dn, En-1).

With this definition defaults in Di have higher priority than those in Dj if i<j. Every "layer" of defaults in a default theory can produce multiple extensions, each extension is used as basis for the generation of extensions in the next layer, i.e. we get a tree with W as root where every son of a node is a DL-extension of its father and the leaves arc PDL-extensions.

It is not difficult to show that our default theories T=(T1,...,Tn) can be translated to PDL default theories D=(D1,...,Dn,{}), where Di is the set of prerequisite-free normal defaults obtained from Ti (i.e. :p/p € Di iff p € Ti). A formula then is strongly provable from T iff it is contained in all PDL extensions of D, and weakly provable from T iff it is contained in at least one PDL extension of D.

The idea of introducing different levels of subthcories into a nonmonotonic formalism has recently been used by Konolige. His hierarchic autoepistemic logic (Konolige 88) allows a modal operator in one level to refer to lower levels only. This makes it possible to represent priorities between conflicting defaults. Unfortunately, it also forces to do so: conflicting defaults for which no priority has been specified lead to an inconsistency. This major drawback of Konolige's approach does not arise in our framework.

As mentioned in Sect. 3 the idea of preferred subthcories has been developed in (Rescher 64) already. Instead of ordering the premises Rescher introduces an ordering of the whole logical language. Every formula belongs to a modal category Mi. The categories represent different degrees of something like "willingness to accept". They are used to select certain maximal consistent subsets of premises as follows: we start with those premises belonging to the lowest category, add as many formulas of the same category as consistenly possible (they don't have to be theorems of any of the premises!) in all possible ways. Then we consistently add to each of these sets as many of the premises belonging to the next category as possible in all possible ways. After that all formulas of the next category which can consistently be added are added and so on. If the last category has been handled, then each resulting set contains a "preferred" maximal consistent subset of the premises (for the details see (Rescher 64), p.50f).

It is difficult to see, however, how this ordering could be used in a framework for default reasoning. It seems difficult to base the notion of default inference on a given ordering of all formulas of the logical language.

There is also a close relation between our approach, in particular the last generalization, and the notion of epistemic entrenchment in (Gardenfors, Makinson 88). These authors, however, are interested in knowledge states, i.e. deductively closed sets of formulas, and the changes of such states when new information is obtained. They are not interested in deriving plausible conclusions from possibly inconsistent premises. Moreover, they require the new knowledge state after the addition of information to be uniquely determined by the epistemic entrenchment whereas we allow multiple preferred subthcories whenever no priority between conflicting defaults is specified.

It is interesting to compare our framework also with the minimal model approach to default reasoning (McCarthy 84). There you start with a consistent set of formulas (otherwise no interesting results are obtained) and select some of the models of the premises. We start with an inconsistent set of formulas $T1 \cup ... \cup Tn$ (otherwise the approach produces the same results as classical logic) and select some of the maximal consistent subsets of the premises. Our general framework (Sect. 2) generalizes Poole's approach in a similar way as Shoham's preferential logics framework (Shoham 86) generalizes McCarthy's circumscription. The analogy goes even further: the motivation behind the development of prioritized circumscription (Lifschitz 85) and pointwise circumscription (Lifschitz 86) seems to be exactly the same as the motivation for our generalizations from Sect. 4 and 5, respectively.

7. Conclusion

We presented two generalizations of David Poole's approach to default reasoning. The first generalization extends his original approach in two respects: (1) we allow several levels of reliability instead of only two and (2) treat all levels uniformly, i.e. there are no unrefutable premises. The second generalization introduces a partial ordering on the premises instead of the levels. In Poole's theory the applicability of a default can be blocked, but there is no way of representing priorities between defaults in the sense that one of two conflicting defaults is not applied if the other one can be applied. Our systems provide a natural means of representing such priorities.

At the heart of our approach is the notion of preferred subtheory. This notion has - as far as we know - been introduced by Rescher 25 years ago. The different levels of a default theory respectively the partial ordering among premises can be seen as ways to define specific preference orderings on maximal consistent subsets useful for default reasoning. It's a topic for further research whether other interesting orderings can be found (e.g. an ordering corresponding to Shoham's chronological minimality principle (Shoham 86)).

It should be noted that the provable formulas of our theories depend on the syntactic form of the theories. It makes an important difference whether, for instance, a level contains both A and B or the equivalent single formula A & B. This could be avoided by the introduction of a certain normal form for formulas. However, we don't see this unusual behaviour as a drawback. It makes perfect sense to distinguish between situations where A as well as B are possible hypotheses or where A & B is one hypothesis.

There is always a tradeoff between expressiveness and simplicity. Obviously, we had to give up some of the simplicity and elegance of Poole's THEORIST in order to increase expressiveness and to allow for the representation of default priorities. But still we are much closer to classical logic than many other systems: we don't need modal operators, nonstandard inference rules, fixed point constructions, second order logic or abnormality predicates. This should make it simpler to integrate default reasoning with other forms of commonsense reasoning. Take as an example counterfactual reasoning. If we use our reliability level approach, then the truth of a counterfactual A > B, for instance, could be determined by introducing a new level {A} with highest reliability into the default theory representing our world knowledge and checking whether B is strongly provable from the new theory (see (Ginsberg 86) for more on counterfactuals). Moreover, the problem of handling inconsistent information - a problem every commonsense reasoner has to deal with anyway - is implicitly solved. We, therefore, hope that this approach will not just increase the number of proposed formalizations of default reasoning but will find its place as a good compromise between simplicity and expressive power.

Acknowledgments:

Thanks to David Poole and Ulrich Junker for valuable discussions. Ulrich also helped to eliminate some bugs in earlier versions of this paper.

References

(Bibel 85) Bibel, Wolfgang: Methods of Automated Reasoning, in: Bibel, Jorrand (eds): Fundamentals in Artificial Intelligence, Springer, LNCS 232, 1985

(Brewka 87) Brewka, Gerhard: The Logic of Inheritance in Frame Systems, IJCAI 87, 1987

(Brewka 89) Brewka, Gerhard: Nonmonotonic Reasoning - From Theoretical Foundation Towards Efficient Computation, Ph. D. thesis, Cambridge University Press, to appear

(Gardenfors, Makinson 88) Gardenfors, Peter, Makinson, David: Revisions of Knowledge Systems Using Epistemic Entrenchment. In: Vardi, M. (ed): Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge, Morgan Kaufmann, Los Altos, 1988

(Ginsberg 86) Ginsberg, Matthew L.: Counterfactuals, Artificial Intelligence 30, 1986

(Konolige 88) Konolige, Kurt: Hierarchic Autoepistcmic Theories for Nonmonotonic Reasoning, Proc. AAAI 88, 1988

(Lifschitz 85) Lifschitz, Vladimir: Computing Circumscription, IJCAI 85, 1985

(Lifschitz 86) Lifschitz, Vladimir: Pointwise Circumscription, Proc. AAAI-86, 1986

(McCarthy 84) McCarthy, John: Applications of Circumscription to Formalizing Common Sense Knowledge, Proc. AAAI-Workshop Non-Monotonic Reasoning, 1984 (also in Artificial Intelligence 28, 1986)

(Moore 85) Moore, Robert C: Semantical Considerations on Nonmonotonic Logic, Artificial Intelligence 25, 1985

(Poole 88) Poole, D.: A Logical Framework for Default Reasoning, Artificial Intelligence 36, 1988

(Poole ct al. 86) Poole, D.; Goebel, R.; Aleliunas, R.: A Logical Reasoning System for Defaults and Diagnosis, University of Waterloo, Dep. of Computer Science, Research Rep. CS-86-06, 1986

(Reiter 80) Reiter, Raymond: A Logic for Default Reasoning, Artificial Intelligence 13, 1980

(Rescher 64) Rescher, Nicholas: Hypothetical Reasoning, North-Holland Publ., Amsterdam. 1964

(Shoham 86) Shoham, Yoav: Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence, Ph.D. Thesis, Yale University, 1986