Model-Preference Default Theories[*]

Bart Selman
Dept. of Computer Science
University of Toronto
Toronto, Canada M5S 1A4
bart@ai.toronto.edu

Henry Kautz
AT&T Bell Laboratories
AI Principles Research Dept.
Murray Hill, NJ 07974
kautz@allegra.att.com

October 25, 2005

Abstract

Most formal theories of default inference have very poor computational properties, and are easily shown to be intractable, or worse, undecidable. We are therefore investigating limited but efficiently computable theories of default reasoning. This paper defines systems of Propositional Model-Preference Defaults, which provide a true model-theoretic account of default inference with exceptions.

The most general system of Model-Preference Defaults is decidable but still intractable. Inspired by the very good (linear) complexity of propositional Horn theories, we consider systems of Horn Defaults. Surprisingly, finding a most-preferred model in even this very limited system is shown to be NP-hard. Tractability can be achieved in two ways: by eliminating the "specificity ordering" among default rules, thus limiting the system's expressive power; and by restricting our attention to systems of Acyclic Horn Defaults. These acyclic theories can encode acyclic defeasible inheritance hierarchies, but are strictly more general.

This analysis suggests several directions for future research: finding other syntactic restrictions which permit efficient computation; or more

[*] This is a much longer version of a paper entitled "The Complexity of Model-Preference Default Theories" presented at the Seventh Biennial Conference of the Canadian Society for Computational Studies of Intelligence, Edmonton, June 1988.

daringly, investigation of default systems whose implementations do not require checking global consistency – that is, fast "approximate" inference.

1 Introduction

An agent need not, indeed cannot, have absolute justification for all of his or her beliefs. For example, an agent often assumes that a member of a particular kind (*e.g.*, Tweety the bird) has a particular property (*e.g.*, the ability to fly) simply because it is typically true that entities of that kind have that property. When formulating a plan of action, an agent often assumes that certain acts will lead to certain consequences, when in fact those consequences are not guaranteed because the world maybe in some unusual state. In order to assimilate information about its environment, an agent will often use a strategy of "hypothesize and test", and adopt a particular model of those inputs, rather than maintaining a representation of all logically possible interpretations.

Such default reasoning seems to offer several advantages. It allows an agent to come to a decision and act in the face of incomplete information. It provides a way of cutting off the possibly endless amount of reasoning and observation that the agent might have to perform in order to gain perfect confidence in its beliefs. And, as Levesque (1986) argues, default reasoning may greatly reduce the complexity of regular deduction. Defaults can be used to "flesh out" an incomplete knowledge base to a *vivid* one; that is, a set of atomic formulas which completely characterize a domain. Once a vivid knowledge base is obtained, deduction reduces to standard database lookup.

A satisfactory formal theory of default reasoning should therefore both model what an agent could come to believe on the basis of given facts and default assumptions, and precisely characterize the very real efficiency of default reasoning over pure deduction. While there is some dispute (Hanks and McDermott 1986) as to the representational adequacy of such proposed formal systems as Default Logic (Reiter 1980) or Circumscription (McCarthy 1980), no one is prepared to defend their abysmal computational properties. All are easily shown to be undecidable in the first-order case, and badly intractable in the propositional case.

We are therefore investigating limited but efficiently computable theories of default reasoning. Such results are of interest even if one intends to implement the default reasoning system on a massively parallel machine. As Levesque (1986) points out, the processing requirements of an exponentially-hard problem can quickly overwhelm even enormous arrays of processors, equal in size to the number of neurons of the brain.

Our interest in using defaults to generate vivid models is a particular reason for our concern with complexity results. It is hardly of interest to eliminate the exponential component of deductive reasoning by introducing an even more costly process of transforming the representation into a vivid form. A number of encouraging results have been developed for non-default vivification, which eliminates explicit disjunctive information through the use of abstraction (Borgida and Etherington, 1989). At some stage, however, it not sufficient to either hide incompleteness through abstraction or by making arbitrary choices; default information must be applied to produce a reasonable and useful vivid model (Etherington *et al.* 1989; Selman 1989).

The number and variety of formal default systems presents an immediate obstacle to the problem of determining the complexity of the *task* of default inference itself. Who is to say, for example, that a problem which is intractable when formulated in theory A is not tractable when formulated in theory B? Etherington (1986) has demonstrated that one should not simply lump all default theories together, as they differ significantly in both their expressive power and the kinds of conclusions they justify. Part of the problem in comparing default theories is their primarily syntactic characterization; indeed, even the semantic accounts provided in the literature retain a strong syntactic flavor (Etherington 1987).

This paper defines a straightforward way of encoding defaults by stating a *preference ordering* over the space of all possible models. This ordering is defined by statements of the form, “a model where α holds is to be preferred over one where β holds." The details of this system of Model-Preference Defaults are spelled out below. The task of the default inference process is to find a most preferred model.

This theory provides a true semantic characterization of default inference; it is important to note that it is *not* a "semantics" which simply mimics the sequential application of syntactic rules. One benefit of this model-theoretic foundation is the ease with which one can incorporate a general *specificity ordering* over defaults. As will be seen, this ordering allows more specific defaults (such as the default that penguins don't fly) to override a less specific one (such as the default that birds fly). This notion of specificity is an important part of practically all known systems of defeasible and uncertain reasoning, including probability theory (Kyburg 1983).

The propositional version of Model-Preference Default theory is decidable but still intractable. Inspired by the very good (linear) complexity of propositional Horn theories, we next consider systems of specificity ordered Horn Defaults over initially-empty knowledge bases. Surprisingly, finding a most-preferred model in even this very limited system is shown to be NP-hard. Tractability is finally achieved by restricting our attention to systems of Acyclic Horn Defaults. These acyclic theories can encode acyclic defeasible inheritance hierarchies, but are strictly more general. Following our complexity analysis we will compare our model-preference default formalism with default logic. It will be shown how model-preference default rules of a certain form can be translated into semi-normal default logic rules.

The final section of this paper considers the consequences of this complexity analysis. One reaction may be to search for other syntactic restrictions on default theories which permit efficient computation. A more daring venture would be to investigate default systems which do not require the existence of a single model of the entire theory. Such systems might be able to perform fast "approximate" inference.

2 Model-Preference Defaults

What is the meaning of a default rule? A common approach (*e.g.*, Reiter's default logic) is to take it to be similar to a deductive rule, but with the odd property of possessing a global (and perhaps non-computable) applicability condition. The conclusions of such a system can only be defined by examining the syntactic structure of particular proofs. There is a very different interpretation of default rules, however, with a natural and intuitive semantics, which is independent of the details of the proof theory. This approach is to use rules to define constraints on the set of *preferred* (or *most likely*) *models* of a situation. The goal of default inference is then to find a most preferred model (of which there may be many), but the details of the syntactic processes employed are separate from the model's semantic characterization.

Unlike previous approaches, the result of Model-Preference Default inference is always a complete model; an appropriate result given our goal of obtaining a vivid representation as described above. By contrast, a default logic proof arrives at an extension, that is, a set of formulas which only partially characterizes a situation.

The model theory for Circumscription is similar to that for Model-Preference Defaults, in that it involves considering models which are maximal w.r.t. some order relation. They differ, however, in that the conclusions of a circumscriptive proof must hold in *all* maximal models, and in the fact that the order relation in a circumscriptive theory is defined solely in terms of minimizing predicates. The first difference makes circumscriptive theory (perhaps too) cautious, while the second leads, at times, to unnatural complexity in encoding default knowledge in terms of predicate minimization. The work of Shoham (1986) on default reasoning involving time and his unifying framework for nonmonotonic reasoning (Shoham 1987) appear to be quite similar to our own, in the emphasis on a semantic theory based on partially-ordered models. While we have studied systems which arbitrarily choose one of the most preferred models, Shoham has concentrated on tightly-constrained domains which have a unique most preferred model. It remains to be seen how comparable our systems are in expressive power.

We hope that model-preference defaults will allow us to construct a precise semantic account of the vivification process described above. A rough characterization would be that the vivid model is simply a most preferred model of the non-vivid theory. The use of abstraction complicates the situation, since the loss of information by abstraction may introduce new models. This issue is currently under investigation.

Model-preference default systems may have applications beyond reasoning with uncertainty. As their name implies, they may prove useful for expressing an agent's desires and preferences, and thus provide the basis for a non-numeric utility theory. Other potential applications are in problems of design and configuration, where default rules express favored design heuristics, which are not absolute constraints on the final solution.

We define a series of default systems, beginning with a general but weak system $\mathcal{D}$, add a specificity ordering over defaults to obtain $\mathcal{D}^+$, then restrict to Horn defaults to yield $\mathcal{DH}$ and $\mathcal{DH}^+$, and finally consider acyclic sets of default rules $\mathcal{DH}^+_{\mathrm{a}}$. This paper considers only purely propositional systems; a later paper will provide a straightforward extension to include propositional schemas.

Definitions

Let $P = \{p_1, p_2, ...p_n\}$ be a set of propositional letters, and $\mathcal{L}$ be a propositional language built up in the usual way ¿from the letters in P and the connectives $\neg$ and $\wedge$ (the others will also be used freely as syntactic abbreviations). Also, let x and x_i be single literals (a literal is either a propositional letter $p \in P$, called a positive literal, or its negation $\neg p$ written as $\overline{p}$, called a negative literal), and α and β be (possibly empty) sets of literals.

Definition: Model
A model (or truth assignment) M for P is a function $t : P \rightarrow \{\mathrm{T}, \mathrm{F}\}$ (T for *true*; F for *false*). M satisfies a set S of formulas of $\mathcal{L}$ (written as $M \models S$) iff M assigns T to each formula in the set. Complex formulas are evaluated with respect to M in the standard manner.

A model is represented by the complete set of literals that is satisfied by the model. For example, if $P = \{p_1, p_2\}$, then the model represented by the set $\{\overline{p_1}, p_2\}$ assigns F to p_1 and T to p_2. Note that the mapping from models to complete sets of literals is one-to-one and onto.

Let γ be a single literal or a set of literals. We will use the notation $M|\gamma$

to denote a model identical to M with the possible exception of the truth assignment for the letters in γ; the truth assignment of those letters is such that $M|\gamma \models \gamma$.

Definition: Default Rule
A default rule d is an expression of the form $\alpha \rightarrow x$. The rule d is a Horn default rule iff α contains only positive literals. Default rules that have a positive literal on the right-hand side will be called a positive default rules, the other rules will be referred to as negative default rules.

Definition: Applicability
A default rule d, of the form $\alpha \rightarrow x$, is applicable at a model M iff

1. $M \models \alpha$, and

2. d is not blocked at M. (For the definition of blocking see the description of the Specificity Condition given below.)

If d is applicable at M, then the application of rule d at M leads to a model M', we will write $M \overset{d}{\rightarrow} M'$. The model M' is identical to M with the possible exception of the truth assignment to the letter corresponding to the literal x; this letter is assigned a truth value such that $M' \models x$.

Definition: Model-Preference Relation
Given a set of default rules D, we will write $M \rightarrow_D M'$ if there exists some rule d in D such that $M \overset{d}{\rightarrow} M'$. The model-preference relation $\leq_D$ is the reflexive, transitive closure of $\rightarrow_D$. When the set of defaults to which we refer is obvious, we write $M \rightarrow M'$ instead of $M \rightarrow_D M'$ and $M \leq M'$ instead of $M \leq_D M'$.

Given a set of defaults, we will say that model M' is *preferred* over M iff $M \leq M'$, and that M' is *strictly preferred* over M iff $M \leq M'$ and $\neg(M' \leq M)$. Two models M and M' are called equivalent w.r.t. the model-preference relation iff $M \leq M'$ and $M' \leq M$. Note that the preference relation induces a partial order on equivalence classes of models.

Definition: Maximal Model
M is a maximal model w.r.t. a set of defaults D iff there does *not* exist a model that is strictly preferred over M.

Definition: Default System $\mathcal{D}$.
In default system $\mathcal{D}$ we consider problems of the the following form: given a set of defaults and a set of propositional letters P that includes those in the default rules, find an arbitrary maximal model for P w.r.t. the set of defaults, temporarily ignoring condition 2 of the definition of applicability.

For example, suppose that P is $\{student, adult, employed\}$, with the intended interpretations "*this person* is a university student", "*this person* is an adult", and "*this person* is employed" (example from Reiter and Criscuolo 1983). Then the defaults "Typically university students are adults", "Typically adults are employed", and "Typically university students are not employed" can be captured as follows:[1]

1) $student \rightarrow adult$
2) $adult \rightarrow employed$
3) $student \rightarrow \overline{employed}$

So, for example, rule 1 says that when given two models that assign T to *student* and that differ only in the truth assignment of *adult*, give preference to the model with *adult* assigned T. The default which says *this person* is a university student can be encoded by:[2]

4) $\emptyset \rightarrow student$

Figure 1 gives the preference ordering on the models as defined by these defaults rules. We use the obvious abbreviations for the propositional letters in P. Thus, for example, $sa\overline{e}$ stands for the model in which both *student* and *adult* are assigned T and *employed* is assigned F. A path from a model M to a model M' indicates that $M \leq M'$. The numbers alongside the directed edges indicate the corresponding default rules.

[1] We omit the set braces in the left-hand side of the default rules.

[2] Instead of adding default rule 4 to the set of defaults, one can express the fact that *this person* is a university student by having the propositional formula *student* in the theory. Below, we define a maximal model w.r.t. a set of defaults and a non-empty theory.

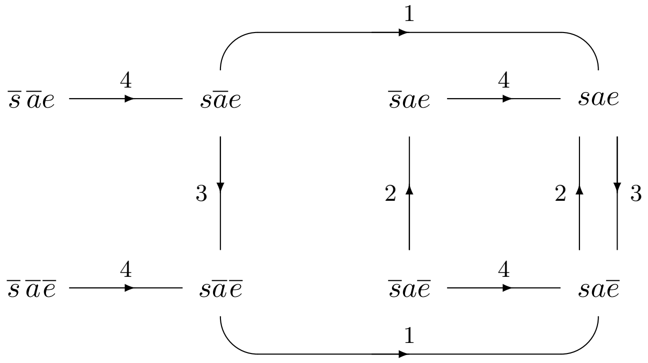
1
$\overline{s}\,\overline{a}e$
4
$s\overline{a}e$
$\overline{s}ae$
4
sae
3
2
2
3
$\overline{s}\,\overline{a}\overline{e}$
4
$s\overline{a}\overline{e}$
$\overline{s}a\overline{e}$
4
$sa\overline{e}$
1

Figure 1: The preference ordering on models as given by the default rules 1) – 4).

We see that the model $sa\overline{e}$ is maximal, since there is no model that is strictly preferred over this model (as a matter of fact, for all other models M, such as for example $s\overline{a}e$, we have $M \leq sa\overline{e}$).

There is a maximal model in this system, however, that does not correspond to our intuitive understanding of the situation. This model is related to the "multiple extension" problem which has created much trouble in previous work on default reasoning (Hanks and McDermott 1986). Because $\mathcal{D}$ does not capture the notion that the third rule above should override the second, the model *sae* is also maximal.

Therefore we define a stronger default system which includes the notion that a more specific default overrides a less specific one.

Definition: Specificity Condition
Given a set of defaults D, a default rule d of the form $\alpha \rightarrow x$ is *blocked* at M iff $\exists d' \in D$ of the form $(\beta \cup \alpha) \rightarrow \overline{x}$ and $M \models (\beta \cup \alpha)$.

Definition: Default System $\mathcal{D}^+$
In default system $\mathcal{D}^+$ we consider problems of the following form: find an arbitrary model for a given set of propositional letters P which is a maximal model according to a given set of defaults, where rules may be blocked by the specificity condition (*i.e.*, both conditions of the definition of applicability are taken into account).

The first example is now more completely captured in $\mathcal{D}^+$ as follows.

$$\begin{aligned}&\mathit{student} \rightarrow \mathit{adult}\\&\mathit{adult} \rightarrow \mathit{employed}\\&\mathit{student}, \mathit{adult} \rightarrow \overline{\mathit{employed}}\\&\emptyset \rightarrow \mathit{student}\end{aligned}$$

The *only* maximal model is now $sa\bar{e}$.[3] (The graph representing the preference ordering is identical to the one in figure 1, without the arc labeled 2 from $sa\bar{e}$ to sae and the arc labeled 3 from $s\bar{a}e$ to $s\overline{ae}$.)

While $\mathcal{D}^+$ appears to have adequate expressive power to handle the standard examples of default reasoning, we will see that it does not succumb to a tractable algorithm. Therefore we define the following restricted classes of default problems.

Definition: Default Systems $\mathcal{DH}$ and $\mathcal{DH}^+$
In $\mathcal{DH}$ we are concerned with the set of problems in system $\mathcal{D}$ involving only Horn default rules; and likewise for $\mathcal{DH}^+$ w.r.t. $\mathcal{D}^+$.

Definition: Acyclic Defaults
Define the directed graph $G(D) = (V, E)$ associated with a set of default rules D as follows:[4] the V contains a vertex labeled p_i for each propositional letter p_i in P, and $E = \{(p_i, p_j) \mid \exists d \in D$ of the form $\alpha \rightarrow x$ s.t. $\{[(p_i \in \alpha) \vee (\bar{p_i} \in \alpha)] \wedge [(p_j = x) \vee (\bar{p_j} = x)]\}\}$. A set of defaults D is called acyclic iff the $G(D)$ is an acyclic directed graph.

The two sets of defaults discussed above are examples of sets of acyclic defaults. They encode defeasible inheritance hierarchies (Touretzky 1986). Acyclic theories can encode such hierarchies, but are strictly more general.

[3] In this example a different solution to the the problem of multiple extensions that does not rely on specificity ordering would be to instead replace rule 2 by *adult*, $\overline{\mathit{student}} \rightarrow \mathit{employed}$. However, specificity ordering captures nicely the intuition behind property inheritance, namely that properties inherited from more general concepts can be overridden by properties inherited from more specific concepts (more generally: more specific defaults should override less specific ones).

[4] This graph should not be confused with a graph like the one in figure 1 which makes explicit the ordering on the models.

Note, however, that we encode exceptions explicitly using more specific defaults. This is similar to the use of semi-normal defaults in the default logic encoding of defeasible inheritance reasoning (Etherington 1986).

Note that there are also natural examples that do not fall into the class of acyclic default systems, such as those obtained by adding the default rule $adult \rightarrow \overline{student}$ to the sets of defaults given above.

Definition: Default System $\mathcal{DH}_{\mathrm{a}}^{+}$
In $\mathcal{DH}_{\mathrm{a}}^{+}$ we are concerned with the set of problems in system $\mathcal{DH}^{+}$ involving only acyclic sets of defaults.

While problems of property inheritance fall within $\mathcal{DH}_{\mathrm{a}}^{+}$, they do not completely circumscribe it.

Finally, we consider the case in which we have apart from a set of defaults D also a non-empty set of facts Th.

Definition: Maximal model w.r.t. D and Th
Let D be a set of defaults and Th a set of propositional formulas. A model M is maximal w.r.t. D and Th iff $(M \models Th) \wedge \neg \exists M' \, ((M' \models Th) \wedge (M'$ is strictly preferred over $M))$.

3 Computational Complexity

We defined a notion of default reasoning based on a model-preference ordering. As stated above, the goal of default inference is to find a maximal model given a set of facts and a partial ordering on the models as defined by a set of default rules. Because of our interest in tractable forms of default reasoning, a central question is: what is the computational cost of finding such a model?

Whenever there are only finitely many models, the problem of finding a maximal model is clearly decidable, since one can simply scan the directed graph representing the partial order on models for a maximal model w.r.t. the defaults and the set of facts. We proceed by analyzing the computational complexity of finding such a model. First we consider the general system $\mathcal{D}$.

We are interested in the complexity of algorithms that handle arbitrary problems in $\mathcal{D}$. Therefore, we consider the search problem (Garey and Johnson 1979) associated with $\mathcal{D}$. A search problem Π is defined as a set of finite objects S_Π called instances, and for each instance $I \in S_\Pi$ a set of finite objects $S[I]$ called solutions of I. An algorithm is said to solve a search problem if it returns the answer "no" whenever $S[I]$ is empty and otherwise returns some arbitrary solution belonging to $S[I]$.

With each system of defaults $\mathcal{X}$ defined in section 2 one can associate in a straightforward manner a search problem $\mathcal{X}_\mathrm{s}$. *E.g.*, an instance I of the search problem $\mathcal{D}_\mathrm{s}$ associated with Problem Class $\mathcal{D}$ is a set of propositional letters P and a set of default rules D. $S[I]$ is the set of maximal models for P w.r.t. D (ignoring condition 2 in the definition of applicability).

The following theorem shows that there does not exist a polynomial algorithm, provided $P \neq NP$, that, given as input a set of defaults D, finds an arbitrary maximal model (ignoring the specificity ordering):

Theorem 1 *The search problem $\mathcal{D}_\mathrm{s}$ is NP-hard.*

In the proof of this theorem (Borgida 1987) we use the following definition and lemma.

Definition: f_D

The function f_D maps a formula in 3CNF (conjunctive normal form with exactly three literals per clause) to a set of default rules in the following manner. If c is a single clause $\{x_i, x_j, x_k\}$,[5] then the set $f_D(c)$ contains the following defaults:

$$\overline{x_i}\,\overline{x_j}{\rightarrow}x_k,\;\; \overline{x_j}\,\overline{x_k}{\rightarrow}x_i,\;\; \overline{x_k}\,\overline{x_i}{\rightarrow}x_j.$$[6]

If γ is a propositional formula in 3CNF consisting of n clauses $c_1, c_2, \ldots c_n$, then $f_D(\gamma) = \bigcup_{i=1,..n} f_D(c_i)$.

Lemma 1 *For any satisfiable 3CNF formula γ, M is a maximal model of $f_D(\gamma)$ iff $M \models \gamma$.*

[5] A clause is a disjunction of literals, and is represented by the set of literals occurring in it.

[6] We use a simplified notation, *e.g.*, $\overline{x_i}\,\overline{x_j}{\rightarrow}x_k$ stands for the default rule $\{\overline{x_i}, \overline{x_j}\}{\rightarrow}x_k$.

Proof: (if) Let M be a model such that $M \models \gamma$. It follows from the definition of $f_D(\gamma)$ that none of the default rules in this set will lead to a truth assignment different from M (note that each clause in γ is satisfied). Therefore, M is a maximal model of $f_D(\gamma)$.
(only if) Let M be a model such that $M \not\models \gamma$. We will show that M is not a maximal model of $f_D(\gamma)$. Since M does not satisfy γ, it follows that there is at least one clause $c = \{x_1, x_2, x_3\}$ such that $M \not\models c$. Let M_{sat} be a model such that $M_{sat} \models \gamma$. Thus, $M_{sat} \models c$, and therefore, M_{sat} satisfies at least one literal in c. Without loss of generality we assume that $M_{sat} \models x_1$. Since $M \not\models c$, we have $M \not\models x_i$ for $i = 1, 2, 3$. It follows that the rule $d_1 : \overline{x_2}\ \overline{x_3} \rightarrow x_1$ in $f_D(\gamma)$ is applicable at M, leading to a model M_1 such that $M_1 \models x_1$. Note that M_1 agrees with M_{sat} on the truth assignment of at least one propositional letter, namely the one in the literal x_1. If $M_1 \not\models \gamma$, then, by a similar argument, there exists a rule d_2 in $f_D(\gamma)$ that leads from M_1 to a model M_2 that agrees on the truth assignment of at least two letters with M_{sat}. In general, in k steps we can reach from M a model M_k such that M_k agrees on at least k letters with M_{sat}. And thus, for some $k \leq n$ (n the number of distinct letters in γ) we have $M_k \models \gamma$. Now, as argued above, there does not exist a rule in $f_D(\gamma)$ that leads from this model to a different one. Thus, we have $M \leq M_k$ and $\neg(M_k \leq M)$ for some $k \leq n$. Therefore, M is not a maximal model. ■

Proof of theorem 1: The proof is based on a Turing reduction from 3-satisfiability. Consider an algorithm that takes as input a formula γ in 3CNF (*i.e.*, an instance of 3-Satisfiability) and constructs $f_D(\gamma)$ (note that this can be done in polynomial time), then calls an oracle that returns in constant time a maximal model M of this set of defaults, and, finally, returns "yes" if $M \models \gamma$ and "no" otherwise. If γ is satisfiable it follows from lemma 1 that the algorithm will return "yes". Otherwise, the algorithm returns "no", as can be seen directly from the algorithm. So, the algorithm returns "yes" iff γ is satisfiable. Moreover, it runs in polynomial time. Therefore, finding a maximal model is NP-hard. ■

Given the very good complexity (linear, Dowling and Gallier 1984) of propositional Horn theories, we now turn our attention to the default system $\mathcal{DH}$. According to the following theorem such defaults can indeed be handled efficiently:

Theorem 2 *Let D be a set of Horn defaults, P be a set of propositional letters that includes those in D, and M_0 be a model such that $M_0 \models \{\overline{p} \mid$*

```
procedure POS($M, D$)
if exists $d : (\alpha \rightarrow p) \in D$ such that $M \models (\alpha \cup \overline{p})$
   then return POS($M|p, D$)
   else  return $M$
```

Figure 2: A procedure for finding a maximal model of a set of Horn defaults (no specificity ordering). Note that the procedure ignores negative default rules, *i.e.*, rules of the form $\alpha \rightarrow \overline{p}$.

$p \in P\}$. *With parameters* M_0 *an* D *the procedure POS (figure 2) returns a maximal model for* D *in time* $O(nk)$, *where* n *is the number of literals*[7] *in* D *and* k *is the number of letters in* P.

The correctness proof of the procedure *POS* is rather tedious, and does not provide much additional insight concerning the complexity of model-preference default theories. We therefore placed the proof of theorem 2 and subsequent correctness proofs of algorithms in appendix A.

Theorem 2 shows that there is an polynomial time algorithm that finds a maximal model of a set of Horn default rules. We now consider the problem of finding a maximal model w.r.t. such defaults and a theory consisting of a set of literals. The following theorem shows that also in this case a maximal model can be found in polynomial time.

Theorem 3 *Let DH be a set of Horn defaults, Th be a consistent set of literals,*[8] *and* P *be a set of propositional letters that includes those in DH and Th. The Max-Model-DH algorithm (figure 3) finds a maximal model of DH and Th in time* $O(nk^2)$, *where* n *is the number of literals in DH and* k *is the number of letters in* P.

In figure 3 we use the notation $\overline{\beta}$, where β is a set of literals, to denote the set $\{\overline{x} \mid x \in \beta\}$. See appendix A for the correctness proof of this algorithm.

[7] Counting each occurrence of a literal separately.

[8] A set of literals *Th* is consistent iff it does not contain a pair of complementary literals such as p and $\overline{p}$.

```
Max-Model-DH Algorithm
Input: a set of propositional letters $P$, set of Horn defaults $DH$,
       and a consistent set of literals $Th$ (the theory).
Output: a maximal model $M_{max}$ of $DH$ and $Th$.
begin
    $M_0 \leftarrow Th \cup \{\overline{p} \mid p \in P$ and $p \notin Th\}$ ; $\delta \leftarrow \emptyset$
    loop
        $D \quad\ \leftarrow DH - \{d \in DH \mid d : \alpha \rightarrow p$ with $p \in \delta\}$
        $M_{pos} \leftarrow \mathrm{POS}(M_0, D)$
        $\beta \quad\ \leftarrow \{p \mid \overline{p} \in Th$ and $M_{pos} \models p\}$
        $\gamma \quad\ \leftarrow \beta - \mathrm{NEG}(\beta, M_{pos}, D)$
        if $(\gamma = \emptyset)$ then $M_{max} \leftarrow M_{pos}|\overline{\beta}$; exit
                     else $\delta \leftarrow \delta \cup \gamma$
    end loop
end
procedure NEG$(\beta, M, D)$
if exists $d : (\alpha \rightarrow \overline{p}) \in D$ such that $[(p \in \beta) \wedge (M \models \alpha) \wedge (\alpha \cap (\beta - \{p\})) = \emptyset]$
  then return $\{p\} \cup \mathrm{NEG}(\beta - \{p\}, M, D)$
  else return $\emptyset$
```

Figure 3: A polynomial algorithm for the search problem $\mathcal{DH}_{\mathrm{s}}$. The algorithm allows for a non-empty theory *Th* consisting of a set of literals.

We will now consider the influence of the specificity condition (used to handle exceptions properly in default reasoning). This leads to the following surprising result:

Theorem 4 *The search problem* $\mathcal{DH}_{\mathrm{s}}^{+}$ *is NP-hard.*[9]

The essence of the proof lies in transforming the set of default rules as used in the proof of theorem 1 into a set of Horn defaults. We therefore replace negative literals by new letters, *e.g.*, $\overline{p}$ is replaced by p'. We then add extra sets of Horn default rules that guarantee that when the original formula α

[9] As a direct consequence it follows that the search problem $\mathcal{D}_{\mathrm{s}}^{+}$ is NP-hard.

is satisfiable, no maximal model will assign the same truth value to a pair of corresponding letters, such as p and p'. The details of this process are spelled out below. The following notation, definitions, and lemmas will be used in the proof of theorem 4.

Let γ be a 3CNF formula containing the set of propositional letters $P = \{p_1, p_2, ..., p_n\}$. W.l.o.g., we assume that no clause in γ contains a pair of complementary literals, such as p and $\overline{p}$, and each clause contains only distinct literals.

Definition: f_{DH^+}

The function f_{DH^+} maps a formula in 3CNF to a set of Horn default rules in the following manner. The set $f_{DH^+}(\gamma)$ contains the rules from the following groups:

Group A. The rules obtained from $f_D(\gamma)$ by replacing each occurrence of $\overline{p_i}$ by a new letter p'_i (for $i = 1, 2, ...n$).
Group B. The rules: $p_i \rightarrow \overline{p'_i}$ (for $i = 1, 2, ...n$).
Group C. The rules: $p'_i \rightarrow \overline{p_i}$ (for $i = 1, 2, ...n$).
Group D. The rules: p'_i (for $i = 1, 2, ...n$).

Let P_{ext} be the set of propositional letters $\{p_1, p'_1, p_2, p'_2, ...p_n, p'_n\}$.

Definition: Consistent Model

The truth assignment for the pair of letters p_i, p'_i $(1 \leq i \leq n)$ in a model M for P_{ext} is consistent iff either $M \models (p_i \wedge \overline{p'_i})$, or $M \models (\overline{p_i} \wedge p'_i)$. A model M for P_{ext} is called consistent iff each pair of letters p_i, p'_i in P_{ext} is assigned consistently. If M is not consistent, then the model is inconsistent.

In the lemmas 2 to 5 and the proof of theorem 4, the models are truth assignments for P_{ext}, and the preference relation is w.r.t. $f_{DH^+}(\gamma)$. Note that, since we are dealing with problems in default system $\mathcal{DH}^+$, we have to consider the possibility of default rules being blocked by other defaults (see the definition of applicability in section 2).

Lemma 2 *If M is inconsistent, then* $\exists M'$ $((M'$ *is consistent*$) \wedge (M \leq M'))$.

Proof: Let M be an inconsistent model for P_{ext}. Therefore there are k $(1 \leq k \leq n)$ pairs of corresponding letters p_i and p'_i inconsistently assigned in M. Without loss of generality, we assume that the pair p_1, p'_1 is assigned inconsistently in M. We will show how one can reach a model M_1 via a default rule in $f_{DH^+}(\gamma)$ such that M_1 is identical to M except for the truth assignment of the pair p_1 and p'_1. This pair will have a consistent truth assignment in M_1. Thus, M_1 will have $k-1$ inconsistently assigned pairs. Therefore, after k default rule applications one can reach, starting from M, a consistent model.

Let the pair of letters p_1 and p'_1 be inconsistently assigned in M. We have to consider the following two cases. Case a) — $M \models (\overline{p_1} \wedge \overline{p'_1})$. In this case, rule p'_1 in group D will apply. Leading to a model M_1 such that $M_1 \models (\overline{p_1} \wedge p'_1)$, *i.e.*, consistent w.r.t. this pair of letters. Note that this rule cannot be blocked, since only the rule $p_1 \rightarrow \overline{p'_1}$ in group B could potentially block this rule. However, this rule is not applicable at M. Case b) — $M \models (p_1 \wedge p'_1)$. In this case, both rule $p_1 \rightarrow \overline{p'_1}$ in group B and rule $p'_1 \rightarrow \overline{p_1}$ in group C will lead to a consistent truth assignment for p_1. Note that neither of these rules can be blocked since there are no rules of the form $(\beta \cup \{p_1\}) \rightarrow p'_1$ or of the form $(\beta \cup \{p'_1\}) \rightarrow p_1$, where β is an arbitrary set of literals, in $f_{DH^+}(\gamma)$. ■

Lemma 3 *If M is consistent and $M \models \gamma$, then $\neg\exists M'\ ((M \neq M') \wedge (M \leq M'))$.*

Proof: Let M be a consistent model that satisfies γ. We will show that none of the rules in $f_{DH^+}(\gamma)$ leads to a model different from M. Since M is consistent, we only have to consider rules in group A. This can be seen as follows. Let M be a consistent model. We will show that none of the rules in the groups B, C, or D will lead to another model. Consider a rule $d : p_i \rightarrow \overline{p'_i}$ in group B. If this rule is applicable at M, then $M \models p_i$. And therefore, since M is consistent, $M \models \overline{p'_i}$. So, this rule will not lead to a model different from M. By a similar argument it follows that none of the rules in group C will lead to a model different from M. Finally, consider a rule $d : p'_i$ in group D. If, $M \models p'_i$, then rule d does not lead to a model different from M. Otherwise, if $M \not\models p'_i$, then $M \models p_i$, since M is consistent. Therefore, the the rule d will be blocked by rule $p_i \rightarrow \overline{p'_i}$ in group B.[10]

We will now consider the rules in group A. These rules are obtained from those in $f_D(\gamma)$ with occurrences of $\overline{p_i}$ replaced by p'_i (for $i =$

[10] Note the fact that the rules in group D can be blocked by more specific ones in group B is essential here.

$1, 2, ..., n$). Since $M \models \gamma$, the truth assignment of the letters $p_1, p_2, ..., p_n$ will satisfy γ. And thus, as argued in the proof of lemma 1, none of the rules in $f_D(\gamma)$ leads to a different truth assignment to those letters. Now, since M is consistent, we have for each letter p_i' ($1 \leq i \leq n$) that $M \models p_i'$ iff $M \models \overline{p_i}$. And thus, by the definition of the rules in group A, none of these rules will lead to a truth assignment different from M. ■

Lemma 4 *For any satisfiable 3CNF formula γ, if M is a consistent model and $M \not\models \gamma$, then $\exists M'$ (M' is strictly preferred over M).*

Proof: Let γ be a satisfiable 3CNF formula and M be a consistent model such that $M \not\models \gamma$. Since M does not satisfy γ, there exists at least one clause c such that $M \not\models c$. Without loss of generality, we assume that $c = \{p_1, \overline{p_2}, p_3\}$. Let M_{sat} be a consistent model such that $M_{sat} \models \gamma$. So, M_{sat} satisfies at least one literal in c. Without loss of generality, we assume that $M_{sat} \models p_1$. Since M does not satisfy c, we have $M \models (\overline{p_1} \wedge p_2 \wedge \overline{p_3})$. Since M is consistent, it follows that the rule $d_1 : p_2 p_3' \rightarrow p_1$ in $f_{DH^+}(\gamma)$ is applicable at M, *i.e.*, $M \xrightarrow{d_1} M|p_1$. (Note that this rule cannot be blocked since rules in group A are the most specific ones, and moreover, they cannot block each other, since all of them are positive.) ¿From the inconsistent model $M|p_1$ we can reach a consistent one via the application of the rule $d_2 : p_1 \rightarrow \overline{p_1'}$ in group B, *i.e.*, $M|p_1 \xrightarrow{d_2} M|\{p_1, \overline{p_1'}\}$. (Note this rule cannot be blocked, as argued in the proof of of lemma 2.) So now, we have obtained a consistent model that agrees on the truth assignment of at least two letters with M_{sat}. If this model does not satisfy γ, it follows, by the above argument, that one can reach a consistent model that agrees in the truth assignment of at least four letters with M_{sat}. And thus, after at most $2n$ default rule applications we obtain a consistent model M' such that $M \leq M'$ and $M' \models \gamma$. And, by lemma 3, it follows that $\neg(M' \leq M)$. ■

We can now state the analogue of lemma 1 for the set of defaults $f_{DH^+}(\gamma)$.

Lemma 5 *For any satisfiable 3CNF formula γ, M is a maximal model of $f_{DH^+}(\gamma)$ iff M is consistent and $M \models \gamma$.*

Proof: (if) Let γ be a satisfiable 3CNF formula and M be a consistent model that satisfies γ. Therefore, by lemma 3, M is maximal.

(only if) Let γ be a satisfiable 3CNF formula and M be a maximal model of $f_{DH^+}(\gamma)$. Assume that M is inconsistent. ¿From lemma 2, it follows that there exists a consistent M' such that $M{\leq}M'$. If $M' \models \gamma$, then, by lemma 3, it follows that $\neg(M'{\leq}M)$, so M is not maximal, a contradiction. Otherwise, by lemma 4, there will exist a model M'' such that $M'{\leq}M''$ and $\neg(M''{\leq}M')$. Since, $M{\leq}M'$, it follows that $M{\leq}M''$ and $\neg(M''{\leq}M)$. Thus, M is not maximal, a contradiction. Finally, assume that M is consistent and $M \not\models \gamma$. It follows, by lemma 4, that M is not a maximal model, a contradiction. So, M is a consistent model that satisfies γ. ■

Proof of theorem 4: According to lemma 5 the set of Horn defaults $f_{DH^+}(\gamma)$ has a property similar the one stated in lemma 1 for the set $f_D(\gamma)$, namely for any satisfiable 3CNF formula γ, M is a maximal model of $f_{DH^+}(\gamma)$ iff M is consistent and $M \models \gamma$. Therefore, the fact that the search problem $\mathcal{DH}^+_{\mathrm{s}}$ is NP-hard follows from a Turing reduction from 3SAT as given in the proof of theorem 1 with $f_D(\gamma)$ replaced by $f_{DH^+}(\gamma)$ and an oracle that takes the second condition of the applicability definition into account. ■

Theorems 2 and 4 show how a relatively small change in expressive power of a tractable representation can lead to a computationally intractable system. Results like this show the importance of a detailed analysis of the computational properties of knowledge representation and reasoning systems (Levesque and Brachman 1985). The following result is another illustration of the tradeoff between expressiveness and tractability:

Theorem 5 *Given a set of Horn defaults DH and a theory TH consisting of a set of Horn formulas, the problem of finding a maximal model w.r.t. DH and TH while ignoring condition 2 of the definition of applicability is NP-hard.*

This result is of interest because of the fact that both propositional Horn defaults without specificity ordering (theorem 2) and Horn theories by themselves are linear. We will use the following notation, definition, and lemmas in the proof of theorem 5.

Let γ be a 3CNF formula containing the set of propositional letters $P = \{p_1, p_2, ..., p_n\}$, and $f_{DH}(\gamma)$ be the set of default rules containing exactly the same rules as $f_{DH^+}(\gamma)$ defined above. (In applying these defaults we will now ignore condition 2 of the definition of applicability.)

Definition: f_{TH}

The function f_{TH} maps a formula in 3CNF to a set of Horn formulas in the following manner. The set $f_{TH}(\gamma)$ is the union of the following groups of formulas:

Group A. From the set containing all clauses in γ we obtain a set of Horn clauses[11] by replacing each occurrence of a positive literal p_i by the negation of a new letter p_i', *i.e.*, $\overline{p_i'}$ (for $i = 1, 2, ...n$). (Thus, we obtain a set of Horn clauses containing only negative literals.)

Group B. The formulas: $p_i \rightarrow \overline{p_i'}$ (for $i = 1, 2, ...n$).

Let P_{ext} again be the set of propositional letters $\{p_1, p_1', p_2, p_2', ...p_n, p_n'\}$. The definition of a consistent model M for P_{ext} is as given above.

Lemma 6 *For any consistent model M for P_{ext}, $M \models \gamma$ iff $M \models f_{TH}(\gamma)$.*

Proof: (only if) Let M be a consistent model such that $M \models \gamma$. Since M is consistent, it will satisfy the formulas in group B. Let $c = \{\overline{p_i'}, \overline{p_j'}, \overline{p_k}\}$ be an arbitrary clause in $f_{TH}(\gamma)$ (the particular choice of literals is not relevant). By the definition of f_{TH} it follows that c is obtained from the clause $c' = \{p_i, p_j, \overline{p}_k\}$ in γ. It can easily be seen that $M \models c$ because $M \models c'$ and M is consistent. It follows that $M \models f_{TH}(\gamma)$.

(if) If M for P_{ext} is a consistent model and $M \models f_{TH}(\gamma)$, it follows, by an argument similar to the one given for the only if direction, that $M \models \gamma$ (Note that the rules in group A are essential here.) ■

Lemma 7 *For any satisfiable 3CNF formula γ, if M is a maximal model of $f_{DH}(\gamma)$ and $f_{TH}(\gamma)$, then $M \models \gamma$.*

Proof: Let γ be a satisfiable 3CNF formula and M be a maximal model of $f_{DH}(\gamma)$ and $f_{TH}(\gamma)$. We will show that $M \models \gamma$.

Since M is a maximal model of $f_{DH}(\gamma)$ and $f_{TH}(\gamma)$, M will satisfy $f_{TH}(\gamma)$. Therefore, because of the formulas in group B, M cannot contain a pair of corresponding letters, such as p_1 and p_1', with both letters assigned T. Moreover, as we will show below, M cannot contain a pair of corresponding letters that are both assigned F. Therefore, M

[11] A Horn clause is clause containing at most one positive literal.

is a consistent model. Since $M \models f_{TH}(\gamma)$ and is consistent, it follows, by lemma 6, that $M \models \gamma$.

We will now show that M cannot contain a pair of corresponding letters both assigned F. Assume that M assigns F to p_k and p'_k ($1 \leq k \leq n$). According to lemma 2, there exists a consistent model M' such that $M{\leq}M'$. Moreover, since there are no rules in $f_{DH}(\gamma)$ that can lead ¿from M' to a truth assignment with both p_k and p'_k assigned F, we have $\neg(M'{\leq}M)$. If M' satisfies $f_{TH}(\gamma)$, then M is not a maximal model of $f_{DH}(\gamma)$ and $f_{TH}(\gamma)$, contradiction. Otherwise, assume that M' does not satisfy $f_{TH}(\gamma)$. Now, as argued in the proof of lemma 4, there will exist a consistent model M^* such that $M'{\leq}M^*$ and $M^* \models \gamma$. By lemma 6, it follows that $M^* \models f_{TH}(\gamma)$. Since there are no rules that can lead from M^* to an assignment of F to both p_k and p'_k, we have $\neg(M^*{\leq}M)$. Therefore, M is not a maximal model of $f_{DH}(\gamma)$ and $f_{TH}(\gamma)$, contradiction. Thus, if γ is satisfiable, then there does not exist a maximal model for $f_{DH}(\gamma)$ and $f_{TH}(\gamma)$ that assigns F to both letters of a corresponding pair of letters. ■

Proof of theorem 5: The proof is based on a Turing reduction from 3SAT similar to the one given in the proof of theorem 1. Consider an algorithm that takes as input a 3CNF formula γ and constructs in polynomial time a set of defaults $f_{DH}(\gamma)$ and a Horn theory $f_{TH}(\gamma)$, then calls an oracle that returns in constant time a maximal model M for the set of defaults and the theory (temporarily ignoring the specificity ordering);[12] and finally, returns "yes" if M satisfies γ, and "no" otherwise. From lemma 7 it follows that the algorithm returns "yes" iff γ is satisfiable. Therefore, since it runs in polynomial time, it follows that the search problem for $\mathcal{DH}$ and a Horn theory is NP-hard. ■

Finally, we consider again specificity ordered Horn defaults. We can obtain a tractable system by limiting our default systems to acyclic ones:

Theorem 6 *Let DH_a^+ be a set of acyclic Horn defaults, and P be a set of propositional letters that includes those in DH_a^+. The Max-Model-DH_{a}^+ algorithm (figure 4) finds a maximal model of DH in time $O(kn^2)$, where n is the number of literals in DH_a^+ and k is the number of letters in P.*

[12] Note that for each input γ there exists at least one maximal model since $f_{TH}(\gamma)$ is satisfiable, *e.g.*, the model that assigns F to all propositional letters in P_{ext} is a satisfiable assignment.

```
Max-Model-DH$_a^+$ Algorithm
Input: A set of acyclic Horn defaults $DH_a^+$ and a set of propositional letters $P = \{p_1, p_2, ...p_n\}$ that includes those in $DH_a^+$.
Output: a maximal model $M_{max}$ of $DH_a^+$
begin
    $p\_remain \leftarrow$ ORDER($P$, $DH_a^+$)
    $M_{part} \leftarrow \emptyset$
    loop
        if (ELEM($p\_remain$) $= \emptyset$) then exit
        $p \leftarrow$ HEAD($p\_remain$) ; $p\_remain \leftarrow$ TAIL($p\_remain$)
        $set\_t \leftarrow$ false ; $set\_f \leftarrow$ false
        for all $d : (\alpha \rightarrow x)$ with $x = p$ or $\overline{p}$ do
            $blocked \leftarrow$ false
            if ($M_{part}$ satisfies $\alpha$) then
                $M_{min} \leftarrow \{x \mid x =$ (if $(p \in M_{part} \cup \alpha)$ then $p$ else $\overline{p}$), $p \in P\}$
                for all $r : (\beta \cup \alpha \rightarrow \overline{x})$do
                    if ($M_{min}$ satisfies $\beta$) then $blocked \leftarrow$ true
                end for
                if ($\neg blocked$) then
                    if $(x = p)$ then $set\_t \leftarrow$ true
                                  else $set\_f \leftarrow$ true
        end for
        if $\neg set\_t$ then $M_{part} \leftarrow M_{part} \cup \{\overline{p}\}$
                else if $\neg set\_f$ then $M_{part} \leftarrow M_{part} \cup \{p\}$
    end loop
    $M_{max} \leftarrow \{x \mid x =$ ( if $(p \in M_{part})$ then $p$ else $\overline{p}$) , $p \in P\}$
end

procedure ORDER($P, D$)
Input: A set of acyclic Horn defaults $DH_a^+$ and a set of propositional letters $P = \{p_1, p_2, ...p_n\}$ that includes those in $DH_a^+$.
Output: a list of letters in $P$, $(p_{i_1}, p_{i_2}, ..., p_{i_n})$, such that for each
        pair of letters $p_{i_k}, p_{i_l}$ in $P$, if $k < l$ then there does not exist
        a path from $p_{i_l}$ to $p_{i_k}$ in the graph $G(DH_a^+)$ associated
        with $DH_a^+$ (see definition of acyclic defaults).
```

Figure 4: A polynomial algorithm for the search problem $\mathcal{DH}^+_{\text{a,s}}$.

In figure 4 the functions ELEM, HEAD, and TAIL each take as argument a list and return, respectively, the set of elements in the list, the first element of the list, and the list obtained after removing the first element. M_{part} is a *partial model* for P, defined as follows:

Definition: Partial Model
A partial model (or partial truth assignment) M_{part} for P is a partial function $t : P \rightarrow \{\mathrm{T}, \mathrm{F}\}$.

In the Max-Model-DH$_a^+$ algorithm we represent a partial model by a set S of literals in the following manner: if S contains the literal p, then p is assigned T in M_{part}, if S contains the literal $\overline{p}$, then p is assigned F in M_{part}, and, if S contains neither p nor $\overline{p}$, then M_{part} does not assign a truth value to p. A partial model represented by the set S satisfies a literal x iff $\overline{x}$ is not an element of S. See appendix A for the proof of theorem 6.

The algorithm can be adapted to handle non-empty theories consisting of a set of literals.[13]

4 A Comparison to Default Logic

In this section we compare model-preference defaults to default logic (Reiter 1980). More specifically, given a model-preference default theory, we will consider whether there exists a set of propositional default logic rules such that there is a one-to-one and onto correspondence between maximal models and extensions. Since an extension of a default logic theory need not contain a complete set of literals, it is clear that given such a theory, there does in general not exist a corresponding model-preference default theory. However, as we will see below, for a large class of model-preference theories one can find a corresponding default logic theory. In our analysis, we assume an empty set of facts. For a comparison to circumscription see Boddy *et al.* (1989).

We will first discuss an example. Consider the set of defaults D used in the example illustrating specificity ordering (section 2). The corresponding

[13] Although we expect that theories consisting of Horn formulas can also be handled efficiently, we have yet to find a polynomial-time algorithm for this case.

set of default logic rules D_{dl} contains two groups of rules. The first group consists of rules that correspond directly to the model-preference defaults:[14]

$$\left\{ \frac{: a \wedge s}{a}, \frac{: e \wedge a \wedge \overline{s}}{e}, \frac{: \overline{e} \wedge s \wedge a}{\overline{e}}, \frac{: s}{s} \right\}.$$

Note how $\overline{s}$ in second rule enforces the specificity ordering. The second group of rules guarantees that the only extensions of D_{dl} are complete sets of literals:

$$\left\{ \frac{: a}{a}, \frac{: \overline{a} \wedge \overline{s}}{\overline{a}}, \frac{: e \wedge \overline{(s \wedge a)}}{e}, \frac{: \overline{e} \wedge \overline{a}}{\overline{e}} \right\}.$$

These defaults can be viewed as a set of closed world assumptions (Reiter 1978) that "force" the system to decide on the truth assignment of each letter, but unlike other cases of closed world assumptions, no preference is given to negative information. D_{dl} has only one extension,[15] $Th\{s, a, \overline{e}\}$, corresponding to the only maximal model of D. Thus, we have a one-to-one and onto mapping between the maximal models and the extensions.

In general, the translation into default logic is given by a function f_{DL} which maps a set of specificity ordered Horn defaults DH^+ to a set of default logic rules.

Definition: f_{DL}

The set $f_{DL}(DH^+)$ contains the rules from the following groups:

Group A. For each model-preference default $\alpha \rightarrow q$ a default logic rule:

$$\frac{: q \wedge \alpha \wedge \overline{\beta_1} \wedge ... \wedge \overline{\beta_k}}{q}$$

where each β_i corresponds to a model-preference default $(\alpha \cup \beta_i) \rightarrow \overline{q}$.[16]

[14] Again, we use the obvious abbreviations.

[15] Here Th denotes closure under logical consequence.

[16] We use semi-normal defaults without prerequisites. It might be more natural to have α as a prerequisite instead of as a justification. To do so, however, we would need more complicated closure rules (group B) because extensions in default logic must be "grounded," *i.e.*, there must be some sequence of rules that "constructs" the extension.

Group B. For each literal q such that $(\emptyset \rightarrow \overline{q}) \notin DH^+$ a rule:

$$\frac{: q \wedge \overline{\delta_1} \wedge ... \wedge \overline{\delta_l}}{q}$$

where each δ_i corresponds to a model-preference default $\delta_i \rightarrow \overline{q}$.

The translation from model-preference defaults into default logic rules is somewhat complicated by the fact that a maximal model is defined as a model for which there does not exist a model that is *strictly* preferred. Therefore, the models on a cycle in the graph corresponding to the preference relation on models can all be maximal. In that case, a translation like the one given above may lead to a set of semi-normal defaults with no extensions. However, given some, relatively weak, restrictions on our model-preference theories, we do obtain a correspondence between maximal models and extensions as shown by the following theorem.

Theorem 7 *Let DH^+ be a set of specificity ordered model-preference defaults that does not contain a mutually blocking pair of defaults such as $\alpha \rightarrow q$ and $\alpha \rightarrow \overline{q}$ or a self-supporting default such as $(\alpha \cup q) \rightarrow q$. If the preference relation induced by DH^+ is a partial order, then M is a maximal model of DH^+ iff $Th(M)$ is an extension of the default logic theory $< f_{DL}(DH^+), \emptyset >$. (In $Th(M)$, M is taken to be a propositional theory; in this case, the set of literals representing the maximal model.)*

See appendix B for the proof of this theorem.

The main restriction is the requirement that the preference relation is a partial ordering; self-supporting defaults and mutually blocking pairs could be allowed but would require a more complicated translation procedure. The correspondence between sets of model-preference defaults and default logic rules can be used to show the intractability of certain classes of semi-normal default rules by reductions ¿from intractable model-preference systems. We will not explore this here since direct methods provide more general complexity results for default logic theories, as demonstrated in Kautz and Selman (1989).

In conclusion, our analysis shows that a partial ordering induced by a set of model-preference defaults can be captured by a semi-normal default logic theory by introducing defaults that force the system to decide on the truth assignment of each letter.

5 Conclusions

We introduced a system for default reasoning based on a model-preference relation. A maximal model in such a system is a complete description of a preferred or most likely state of affairs, based on incomplete information and a set of defaults. Unlike most other approaches to default reasoning, ours is purely semantic, and is defined independent of syntactic notions.

The goal of our work is to develop tractable methods of default reasoning, for use in fast reasoning systems which represent knowledge in a vivid form. Therefore, we only allow complete models as default conclusions. Model-preference theories seem to be of interest, however, beyond this one application. The specificity ordering on defaults, a crucial component of any kind of default reasoning, is neatly captured in $\mathcal{D}^+$ and its subtheories. Another natural application for model-preference theories is to encode a logic of choice, whereby an an agent chooses which of his goal states is most preferred.

We presented a detailed analysis of complexity properties of the various model-preference default systems. The analysis indicates that only systems with quite limited expressive power lead to tractable reasoning (*e.g.*, $\mathcal{DH}$ and $\mathcal{DH}^+_{\mathrm{a}}$). We also gave an example of how a relatively small change in the expressive power of a tractable system can lead to intractability (from $\mathcal{DH}$ to the intractable $\mathcal{DH}^+$).

Acyclic inheritance hierarchies can be represented in the tractable system $\mathcal{DH}^+_{\mathrm{a}}$. Classes of acyclic rules have also been singled out by others (*e.g.*, Touretzky (1986) on acyclic inheritance hierarchies, and related work by Etherington (1986) on ordered default theories) for their relatively good computational properties. A direct comparison with our approach is complicated by the fact that we do not allow for partial models. Selman and Levesque (1989) give a complexity analysis of highly specialized forms of defeasible reasoning such as Touretzky's inheritance reasoner.

The nature of model-preference defaults was further illustrated by a comparison with default logic (Reiter 1980). In this comparison we showed how a set of semi-normal rules, which can be viewed as representing a special form of closed world assumptions, can be added to a set of default logic rules in a way that guarantees the extensions to be complete models.

This work suggests several directions for future research. One is the development of a first-order version of model-preference defaults. Another is to allow for more expressive power and introduce some form of "approximate reasoning" to keep the system tractable. A search for other tractable sub-classes would be in order. And, finally, to determine the usefulness of the tractable systems we have identified, a further study of the forms of defaults necessary in real world domains, *e.g.*, conventions in cooperative conversation (Perrault 1987), is needed.

Acknowledgments

This work originated from research meetings which were organized by Ron Brachman and were financially supported by AT&T Bell Laboratories. It was supported in part by a Government of Canada Award to the first author. We would like to thank Hector Levesque for providing the definition of model-preference defaults and for introducing us to the complexity questions associated with these systems, and Alex Borgida for the complexity result for $\mathcal{D}$. We also thank Ron Brachman, Jim des Rivières, Dave Etherington, Armin Haken, Gerhard Lakemeyer, and David Poole for useful discussions and comments.

References

Boddy, M.; Goldmann, R.P.; Stein, L.A.; and Kanazawa, K. (1989). Investigations of Model-Preference Default Theories. Forthcoming Technical Report, Department of Computer Science, Brown University, 1989.

Borgida, A. (1986) Personal communication, September 1986.

Borgida, A. and David W. Etherington, D.W. (1989). Hierarchical Knowledge Bases and Tractable Disjunction. *Proceedings of Knowledge Representation and Reasoning '89*, Toronto, Canada, May 1989.

Dowling, W.F. and Gallier, J.H. (1984). Linear-Time Algorithms for Testing the Satisfiability of Propositional Horn Formulae. *Journal of Logic Programming*, **3**, 1984, 267–284.

Etherington, D.W. (1986). Reasoning With Incomplete Information: Investigations of Non-Monotonic Reasoning. Ph.D. Thesis, University of British Columbia, Department of Computer Science, Vancouver, BC, Canada, 1986. Revised version to appear as: *Reasoning With Incomplete Information.* London: Pitman / Los Altos, CA: Morgan Kaufmann.

Etherington, D.W. (1987). A Semantics for Default Logic. *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, 1987, 495 – 498.

Etherington, D., Borgida, A., Brachman, R.J., and Kautz, H. (1989). Vivid Knowledge and Tractable Reasoning: Preliminary Report. *Submitted for publication.*

Garey, M.R. and Johnson, D.S. (1979). *Computers and Intractability, A Guide to the Theory of NP-Completeness.* New York: W.H. Freeman, 1979.

Hanks, S. and McDermott, D. (1986). Default Reasoning, Non-Monotonic Logics, and the Frame Problem. *Proceedings of the Fifth National Conference on Artificial Intelligence*, Philadelphia, PA, 1986, 328–333.

Kautz, H.A., and Selman, B. (1989). Hard Problems for Simple Default Logics. *Proceedings of Knowledge Representation and Reasoning '89*, Toronto, Canada, May 1989.

Kyburg, H. (1983). The Reference Class. *Philosophy of Science*, **50**, 1983, 374–397.

Levesque, H.J. (1986). Making Believers out of Computers. *Artificial Intelligence*, **30**, 1986, 81–108.

Levesque, H.J. and Brachman, J.R. (1985). A Fundamental Tradeoff in Knowledge Representation and Reasoning (Revised Version). In *Readings in Knowledge Representation* by R.J. Brachman and H.J. Levesque (Eds.), Los Altos, CA: Morgan Kaufmann, 1985, 41–70.

McCarthy, J. (1980). Circumscription – A Form of Non-Monotonic Reasoning. *Artificial Intelligence*, **13**, 1980, 27–38.

Perrault, C.R. (1987). An Application of Default Logic to Speech Act Theory. Technical Report, SRI International, Artificial Intelligence Center, Palo Alto, CA, 1987.

Reiter, R. (1980). On Closed World Data Bases. In *Logic and Data Bases*, Gallaire, H. and Minker, J. (eds.), New York: Plenum Press, 1978.

Reiter, R. (1980). A Logic for Default Reasoning. *Artificial Intelligence*, **13**, 1980, 81–132.

Reiter, R. and Criscuolo, G. (1983). Some Representational Issues in Default Reasoning. *Computers & Mathematics with Applications*, (Special Issue on Computational Linguistics), **9** (1), 1983, 1 – 13.

Selman, B. (1989). Tractable Default Reasoning. Ph.D. thesis, Department of Computer Science, University of Toronto, Toronto, Ont. (forthcoming).

Selman, B. and Levesque, H.J. (1989). The Tractability of Path-Based Inheritance. *Submitted for publication.*

Shoham, Y. (1986). Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence. Ph.D. Thesis, Yale University, Computer Science Department, New Haven, CT, 1986.

Shoham, Y. (1987). Nonmonotonic Logics: Meaning and Utility. *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, 1987, 388– 393.

Touretzky, D.S. (1986). *The Mathematics of Inheritance Systems.* Research Notes in Artificial Intelligence. London: Pitman / Los Altos, CA: Morgan Kaufmann, 1986.

A Correctness proofs and complexity of algorithms and procedures

Theorem 2: *Let D be a set of Horn defaults, P be a set of propositional letters that includes those in D, and M_0 be a model such that $M_0 \models \{\overline{p} \mid p \in P\}$. With parameters M_0 an D the procedure POS (figure 2) returns a maximal model for D in time $O(nk)$, where n is the number of literals in D and k is the number of letters in P.*

We will use the following definition and lemma in the proof of theorem 2.

Definition: $\preceq^+$

The $\preceq^+$ relation defines a partial order on models: for any two models M and M' for a set of propositional letters P, $M \preceq^+ M'$ iff $\{p \in P \mid M \models p\} \subseteq \{p \in P \mid M' \models p\}$. (Informally, $M \preceq^+ M'$ iff M' is *as true as* M.)

The lemma states two properties of the procedure POS. These properties are more general than strictly needed for the proof of theorem2. We will use them in their full generality in the correctness proof of an algorithm that finds a maximal model for a set of Horn defaults and a non-empty theory (theorem 3).

Lemma 8 *Let D be a set of Horn defaults, P be a set of propositional letters that includes those in D, and M be a truth assignment for P. With parameters M and D the procedure POS returns a truth assignment with the following properties:*

1. $M \leq POS(M, D)$
2. $\forall M' \; (POS(M, D) \leq M') \wedge (M \preceq^{+} M')) \Longrightarrow (M' \leq POS(M, D))$

Proof: First we will show that POS does return a truth assignment on all inputs. In each recursive call the number of letters assigned T in the truth assignment will be increased by one. Therefore, after at most |P| recursive calls the algorithm will return a truth assignment. We now consider the two properties of POS(M, D).

Property 1: By a straightforward finite induction on the number of letters assigned F in M. We omit the details of this induction. (Note that the number of letters assigned F in M decreases by one in each recursive call, and in each call, except for the final one, a rule is found that satisfies the condition of the if statement. A sequence of these rules leads from M to POS(M, D).)

Property 2: We first consider the special case in which there does not exist a rule in D that satisfies the condition in the if statement. It follows that POS$(M, D) = M$ and that any model M' reachable from M will be such that $M' \preceq^{+} M$. Now, if M' also satisfies $M \preceq^{+} M'$, then $M' = M$. Thus, $M' =$ POS(M, D), and therefore, $M' \leq$POS(M, D). We now prove property 2 by finite induction on the number of letters assigned F in M. Base: Let M be the model such that $M \models P$. It follows that there does not exist a rule in D that satisfies the condition in the if-statement. Therefore, from the case discussed above, property 2 holds. Induction Step: Assume property 2 holds for all M with k $(0 \leq k < |P|)$ letters assigned F. We will show that property 2 holds for all M with $k + 1$ letters assigned F. Let M be a model with $k + 1$ letters assigned F. If there does not exist a rule in D that satisfies the condition in the if statement, then property 2 follows from the case discussed above. Otherwise, let $d : \alpha \rightarrow p$ be the first rule found that satisfies the condition of the if statement. Therefore, we have that POS$(M, D) =$ POS$(M|p, D)$ and $M \xrightarrow{d} M|p$. ¿From the induction hypothesis it follows that $\forall M'$ ((POS$(M, D) \leq M') \wedge (M|p \preceq^{+} M')) \Longrightarrow (M' \leq$POS$(M, D))$. Now, let M' be a model such that POS$(M, D) \leq M'$ and $M \preceq^{+} M'$. Application of rule d to M' leads to the model $M'|p$, *i.e.*, $M' \xrightarrow{d} M'|p$ (note that d is applicable at M' since $M \preceq^{+} M'$ and d is applicable at M). It follows that $M'|p$ is such that POS$(M, D) \leq M'|p$. We also have that $M|p \preceq^{+} M'|p$, since $M \preceq^{+} M'$. Thus, by the induction hypothesis, $M'|p \leq$POS(M, D), and therefore, since $M' \xrightarrow{d} M'|p$, we have $M' \leq$POS(M, D). ■

Proof of theorem 2: Let M_0 be the model such that $M_0 \models \{\overline{p} \mid p \in P\}$. Note that for any model M we have $M_0 \preceq^+ M$. We will now show that there does not exist a model that is strictly preferred over $\mathrm{POS}(M_0, D)$. Let M be a model such that $\mathrm{POS}(M_0, D) \leq M$. It follows, directly form lemma 8 property 2, that $M \leq \mathrm{POS}(M_0, D)$. Thus there does not exist a model M such that $\mathrm{POS}(M_0, D) \leq M$ and $\neg(M \leq \mathrm{POS}(M_0, D))$. And therefore, $\mathrm{POS}(M_0, D)$ is a maximal model.

Finally, we show that POS finds a maximal model in time $O(nk)$, where n is the number of literals in D and k the number of letters in P. This follows directly from the fact that there will be at most k recursive calls, as argued in the proof of lemma 8, and checking the condition of the if statement may require checking each rule which takes time $O(n)$. ■

Theorem 3: *Let DH be a set of Horn defaults, Th be a consistent set of literals,*[17] *and P be a set of propositional letters that includes those in DH and Th. The Max-Model-DH algorithm (figure 3) finds a maximal model of DH and Th in time $O(nk^2)$, where n is the number of literals in DH and k is the number of letters in P.*

In figure 3 we use the notation $\overline{\beta}$, where β is a set of literals, to denote the set $\{\overline{x} \mid x \in \beta\}$. Before giving a correctness proof of the Max-Model-DH algorithm, we will first discuss, in general terms, the method used in the algorithm to find a maximal model. The basic idea is the same as for the procedure POS. *I.e.*, we start with a minimal model M_0 w.r.t. $\preceq^+$ that satisfies the theory and search for a maximal model by applying positive defaults as often as possible using the procedure POS. If the model returned by POS, M_{pos}, satisfies the theory *Th*, then this is indeed a maximal model w.r.t. the defaults and the theory. This follows from lemma 8 property 2. However, in case the theory contains negative literals, the procedure POS may return a model that does not satisfy the theory. (Note that since M_0 satisfies the theory, the model returned by POS will satisfy at least the positive literals in the theory.) If so, the algorithm proceeds by testing, using the procedure NEG, whether negative rules can lead from the model returned by POS to a model that does satisfy the theory. If such a model can be reached, then that model is again maximal w.r.t. the defaults and the theory. Otherwise, consider the case in which no model that satisfies the theory can be reached from the model M_{pos}. This means that there is a non-empty set of letters γ such that $\overline{\gamma} \subseteq \mathit{Th}$, $M_{pos} \models \gamma$, and there is no sequence of rules that leads from M_{pos} to a model that satisfies $\overline{\gamma}$. The algorithm now continues the search for a maximal model by repeating the above process starting from M_0 and applying a proper subset of the default rules.

[17] A set of literals *Th* is consistent iff there does not exist a propositional letter p such that both p and $\overline{p}$ are elements of *Th*.

This subset is obtained from the original set of rules by removing all positive rules that have a letter from γ on the right-hand side (note that in the algorithm δ is the union of the γ's obtained each time through the loop). That we can safely restrict ourselves to this subset in the search for a maximal model, starting from M_0, follows from the following argument. The application of a positive rule with a letter in γ on the right-hand side leads to a model M such that $M \not\models \overline{\gamma}$. It follows from lemma 10 below, that from the model M one cannot reach a model that satisfies the theory (note that $\overline{\gamma} \subseteq Th$). Therefore, we can ignore such rules in our search for a maximal model. After at most $|\,Th\,|$ alternations of POS and NEG the algorithm will find a maximal model w.r.t. the defaults and the theory. (Note that since M_0 satisfies the theory, at least one maximal model w.r.t. the defaults and the theory is reachable from this model.) Before we prove theorem 3 we first consider several lemmas. The first one states a basic property of the procedure NEG.

Lemma 9 *Let M be a truth assignment for the set of propositional letters P, β be a subset of P, and D be a set of Horn defaults. The set η =NEG(β, M, D) is a subset of P with the following property:*

$$M {\leq} M|\overline{\eta}$$

Proof: By a straightforward induction on $|\beta|$. We omit the details of this induction. (Note that $|\beta|$ decreases by one in each recursive call, and in each call, except for the final one, a rule is found that satisfies the condition of the if statement. A sequence of these rules leads from M to $M|\overline{\eta}$.) ■

The following lemma is central to the correctness proof of the algorithm Max-Model-DH, it concerns the set of letters returned by NEG when applied to a model M_{pos} obtained by POS with parameters M_0 and D. In the algorithm, β is the set of letters that occur negated in the theory and are assigned T in the model M_{pos}. The procedure NEG with parameters β , M_{pos}, and D will return a subset of letters in β for which the truth assignment can be changed via a sequence of application of default rules (lemma 9). Lemma 10 shows how the search for a maximal model starting from the model M_0 using the procedure POS can now be pruned by removing all default rules that lead to models that assign T to any of the letters in the set $\gamma = \beta - \text{NEG}(\beta, M_{pos}, D)$.

Lemma 10 *Let M_0 be truth assignments for the set of propositional letters P, D be a set of Horn defaults, M_{pos} be the model returned by POS(M_0, D), β be a subset of P. The set $\gamma = \beta -$NEG(β, M_{pos}, D) has the following property:*

$$\forall M\ ((M_0{\leq}M) \wedge (M \not\models \overline{\gamma})) {\Longrightarrow} \neg\exists M'\ ((M{\leq}M') \wedge (M' \models \overline{\gamma}))$$

Proof: By finite induction on $|\beta|$. Base: $|\beta| = 0$, thus $\gamma = \emptyset$, and the property holds vacuously. Induction Step: Assume that the property holds for all β with $|\beta| = k$ $(0 \leq k < |P|)$. We will show that the property holds for for all β with $|\beta| = k+1$. Let β be a subset of P with $|\beta| = k+1$. We first consider the case in which there exists at least one rule that satisfies the condition of the if statement in the procedure NEG. Let $d : \alpha{\to}\overline{p}$ be the first such rule found. ¿From the procedure NEG it follows that $\text{NEG}(\beta, M_{pos}, D) = \{p\} \cup \text{NEG}(\beta - \{p\}, M_{pos}, D)$. Let $\beta' = \beta - \{p\}$ and $\gamma' = \beta' - \text{NEG}(\beta', M_{pos}, D)$. Since $\gamma = \beta - \text{NEG}(\beta, M_{pos}, D)$, we have $\gamma = \gamma'$, and the property for γ follows directly from the induction hypothesis.

Finally, we consider the case in which there does not exist a rule $d : \alpha{\to}\overline{p}$ in D such that $[(p \in \beta \wedge (M_{pos} \models \alpha) \wedge (\alpha \cap (\beta - \{p\})) = \emptyset)]$. It follows that $\gamma = \beta$. Let M be a model such that $M_0{\leq}M$ and $M \not\models \overline{\gamma}$. We will show that there does not exist a model M' such that $M{\leq}M'$ and $M' \models \overline{\gamma}$. Assume that such a model M' does exist. Since $M{\leq}M'$, there exists a sequence of default rule applications that leads from M to M'. Let M'' be the first model in this sequence such that $M'' \models \overline{\gamma}$, M''_{prec} the preceding one, and d be the rule that leads from $M''_{\text{prec}}$to M'', *i.e.*, $M''_{\text{prec}} \xrightarrow{d} M''$. The rule d must be of the form $\alpha{\to}\overline{p}$ with $p \in \gamma$ and $(\alpha - (\gamma - \{p\})) = \emptyset$, since $M''_{\text{prec}} \models \alpha$ and M''_{prec} satisfies $\overline{\gamma}$ except for the letter p. Since $M_0{\leq}M$ and $M{\leq}M''_{\text{prec}}$, we have $M_0{\leq}M''_{\text{prec}}$. And therefore, $M''_{\text{prec}} {\preceq^+} M_{pos}$. (This can easily be shown by induction on the length of the sequence of rule applications that leads from M_0 to M''_{prec}.) Thus, $M_{pos} \models \alpha$. Now, since $\gamma = \beta$ for this case, the rule $d : \alpha{\to}\overline{p}$ is such that $p \in \beta$, $M_{pos} \models \alpha$, and $(\alpha \cap (\beta - \{p\})) = \emptyset$, a contradiction. Therefore, such an M' does not exist. ∎

The following lemma gives two properties of the set δ in the Max-Model-DH algorithm. The first property follows directly from the algorithm, and the second one extends the property for γ, as stated in lemma 10, to the set δ.

Lemma 11 *Let DH be a set of Horn defaults, Th be a consistent set of literals, P be a set of propositional letters that includes those in DH and Th, and* $M_0 = Th \cup \{\overline{p} \mid p \in P \text{ and } p \notin Th\}$. *The set* δ *in the Max-Model-DH algorithm, after being initialized, has the following properties:*

1. $\overline{\delta} \subseteq Th$
2. $\forall M\ ((M_0{\leq}M) \wedge (M \not\models \overline{\delta})) {\Longrightarrow} \neg\exists M'\ ((M{\leq}M') \wedge (M' \models \overline{\delta}))$

Proof: By a course of values induction on $|\delta|$.

Property 1. Trivial.

Property 2. Induction assumption: property 2 holds for δ with $|\delta| < k$ $(0 \leq k)$. We will show that property 2 holds for δ with $|\delta| = k$. If, $k = 0$ than $\delta = \emptyset$ and property 2 holds vacuously. Otherwise, if $k > 0$ then, from the Max-Model-DH algorithm, it follows that δ is given by $\delta = \delta_{prev} \cup \gamma$ with δ_{prev} the previous value of δ, $|\delta_{prev}| < k$, and $\gamma = \beta -$ $\text{NEG}(\beta, M_{pos}, D) \neq \emptyset$, in which $M_{pos} = \text{POS}(M_0, D)$, $D = \mathit{DH} - \{d \in \mathit{DH} \mid d : \alpha \rightarrow p \text{ with } p \in \delta_{prev}\}$, and $\beta = \{p \mid \overline{p} \in \mathit{Th} \text{ and } M_{pos} \models p\}$. Let M be a model such that $M_0 \leq M$ and $M \not\models \overline{\delta}$. We distinguish between the following two cases.

Case a) — Assume that $\neg(M_0 \leq_D M)$. Then the sequence of rule applications that leads from M_0 to M will contain a rule in $\mathit{DH} - D$. In the sequence, this rule leads to a model M' such that $M' \not\models \overline{\delta_{prev}}$. Therefore, by the induction assumption and the fact that $\delta_{prev} \subseteq \delta$ we have $\neg\exists M''\ ((M' \leq M'') \wedge (M'' \models \overline{\delta}))$. And thus, since $M' \leq M$, $\neg\exists M''\ ((M \leq M'') \wedge (M'' \models \overline{\delta}))$.

Case b) — Assume that $M_0 \leq_D M$. Then, because $M_0 \models \overline{\delta_{prev}}$ (by the definition of M_0 and property 1) and by the definition of D, we have $M \models \overline{\delta_{prev}}$. And thus, $M \not\models \overline{\gamma}$, since $M \not\models \overline{\delta}$ and $\overline{\delta} = \overline{\gamma} \cup \overline{\delta_{prev}}$. Now, from lemma 10, it now follows that $\neg\exists M'\ ((M \leq_D M') \wedge (M' \models \overline{\delta}))$. Finally, we have to consider the possibility that there exists a model M' that satisfies $\overline{\delta}$ and is reachable from M via a sequence of rules that includes at least one rule in $\mathit{DH} - D$. Assume that such a model exists. Consider the sequence of rule applications that leads from M to M'. In this sequence, a rule in $\mathit{DH} - D$ will lead to a model M'' such that $M'' \not\models \overline{\delta_{prev}}$. ¿From the induction hypothesis and the fact that $\delta_{prev} \subseteq \delta$, it follows that there does not exists a model M''' such that $M'' \leq M'''$ and $M''' \models \overline{\delta}$. Therefore, since $M'' \leq M'$, we have $M' \not\models \overline{\delta}$, a contradiction. ■

Proof of theorem 3: Firstly, we will show that the algorithm halts on all inputs, and that M_{max} is a maximal model w.r.t. *DH* and *Th*.

The algorithm halts on all inputs and returns a truth assignment because δ monotonically increases each time through the loop, as can be seen ¿from the algorithm, and $|\overline{\delta}| \leq |\mathit{Th}|$ (lemma 11 property 1). (Note that at the execution of the assignment statement for β, β is assigned a set of letters that do not yet occur in γ.)

We now show that M_{max} is indeed a maximal model w.r.t. *DH* and *Th*. Let M_0, D, M_{pos}, M_{max}, δ, and β be as in the Max-Model-DH algorithm just before execution of the exit statement in the loop statement. Thus, we have $M_0 = \mathit{Th} \cup \{\overline{p} \mid p \in P \text{ and } p \notin \mathit{Th}\}$, $D = \mathit{DH} - \{d \in \mathit{DH} \mid d : \alpha \rightarrow p \text{ with } p \in \delta\}$, $M_{pos} = \text{POS}(M_0, D)$, $M_{max} = M_{pos}|\overline{\beta}$, and $\beta = \{p \in P \mid \overline{p} \in \mathit{Th} \text{ and } M_{pos} \models p\} = \text{NEG}(\beta, M, D)$ (since $\gamma = \emptyset$).

We first show that $M_{max} \models \mathit{Th}$. Let A be the set $\{\overline{p} \mid \overline{p} \in \mathit{Th}\}$ and B be the set $\{p \mid p \in \mathit{Th}\}$. It follows directly that $M_{max} \models A$. Also, from the procedure POS and the fact that $M_0 \models B$, we have $M_{pos} \models B$. Therefore, since $(\beta \cap B) = \emptyset$ because *Th* is consistent, it follows that $M_{pos}|\overline{\beta} \models B$. Thus, $M_{max} \models \mathit{Th}$.

We now proceed to show that $\neg\exists M'((M' \models \mathit{Th}) \wedge (M_{max} \leq M') \wedge \neg(M' \leq M_{max}))$. Let M' be a truth assignment such that $((M' \models \mathit{Th}) \wedge (M_{max} \leq M'))$. We will show that it follows that $(M' \leq M_{max})$. By lemma 8 property 1 and lemma 9 we have $M_0 \leq M_{pos} \leq M_{max} \leq M'$. If M' is obtained via a sequence of rules in D, it follows ¿from lemma 8 property 2 that $M' \leq_D M_{pos}$, since $M_{pos} \leq_D M'$ and $M_0 \preceq^+ M'$ because $M' \models \mathit{Th}$ and $M_0 = \mathit{Th} \cup \{\overline{p} \mid p \in P \text{ and } p \notin \mathit{Th}\}$. Therefore, we have that $M' \leq M_{max}$, since $D \subseteq \mathit{DH}$ and $M_{pos} \leq M_{max}$. (Note that when we do not indicate a particular set of rules, the preference relation is w.r.t. *DH*.) Otherwise, assume that M' is obtained from M_{max} via a sequence s of rules that contains at least one rule $d : \alpha \rightarrow p$ with $p \in \delta$. Therefore, in the corresponding sequence of truth assignments there is a model M'' such that $M'' \not\models \overline{\delta}$. Since $M_0 \leq M_{max}$, we have $M_0 \leq M''$. Therefore, by lemma 11 property 2 and the fact that $\overline{\delta} \subseteq Th$ by lemma 11 property 1, it follows that $M' \not\models Th$, a contradiction. Thus, such an M' does not exist. It follows that M_{max} is a maximal model w.r.t. *DH* and *Th*.

Finally, we show that Max-Model-DH algorithm finds M_{max} in time $O(nk^2)$, where n is the number of literals in *DH* and k the number of letters in P. Like the procedure POS, the procedure NEG will take time $O(nk)$. The body of the loop statement will be executed at most k times, since δ monotonically increases and is bounded by *Th*. The procedures POS and NEG are the most time consuming steps in the body of the loop, so executing the loop will take time $O(nk)$. Thus, the Max-Model-DH algorithm will find a maximal model in time $O(nk^2)$. ■

Theorem 6: *Let DH_a^+ be a set of acyclic Horn defaults, and P be a set of propositional letters that includes those in DH_a^+. The Max-Model-DH_a^+ algorithm (figure 4) finds a maximal model of DH in time $O(kn^2)$, where n is the number of literals in DH_a^+ and k is the number of letters in P.*

In figure 4 the functions ELEM, HEAD, and TAIL each take as argument a list and return, respectively, the set of elements in the list, the first element of the list, and the list obtained after removing the first element. M_{part} is a *partial model* for P, defined as follows:

Definition: Partial Model
A partial model (or partial truth assignment) M_{part} for P is a partial function $t : P \rightarrow \{\mathrm{T}, \mathrm{F}\}$.

In the Max-Model-DH_a^+ algorithm we represent a partial model by a set S of literals in the following manner: if S contains the literal p, then p is assigned T in M_{part}, if S contains the literal $\overline{p}$, then p is assigned F in M_{part}, and, if S contains neither p nor $\overline{p}$, then M_{part} does not assign a truth value to p.

A partial model represented by the set S satisfies a literal x iff $\overline{x}$ is not an element of S. We will say that two (partial) models agree on the truth assignment of a letter p iff they both assign the same truth value to p *or* at least one of the models does not assign a truth value to p. A (partial) model M for P agrees with a (partial) model M_R for $R \subseteq P$ iff M and M_R agree on the truth assignment of the letters in P. Note that M and M_R by definition agree on letters in the set $P - R$, since M_R does not assign a truth value to those letters. Unless explicitly stated otherwise, models are not partial.

The following lemma states properties of of the set of models for P that agree with the partial model M_{part} as constructed in the algorithm and with an arbitrary, fixed truth assignment for the letters in the set $R = \text{ELEM}(\textit{p\_remain})$. Note that the size of R decreases by one each time through the main loop.

Lemma 12 *In the Max-Model-DH$_a^+$ algorithm upon entering the loop-statement, M_{part} has the following properties (let $R = \text{ELEM}(\textit{p\_remain})$):*

1. *Let M_R be an arbitrary model for R. If the model M for P agrees with both M_{part} and M_R, and the model M' for P agrees with M_R, then $(M \leq M')$ iff M' agrees with M_{part}*
2. *If $R = \emptyset$, then any M for P that agrees with M_{part} is a maximal model.*

Proof: Property 1: Proof by finite induction on $|P - R|$. Base: $|P - R| = 0$, therefore $P = R$. Let M_R be an arbitrary truth assignment for R. ¿From the algorithm it follows that $M_{part} = \emptyset$. Since $P = R$, it follows that M_R is the only model for P that agrees with M_{part} and M_R, and M_R also is the only model for P that agrees with M_R. Thus, by the reflexivity of $\leq$, property 1 follows. Induction Step: Assume property 1 holds upon entering the loop statement with $|P - R| = k$ $(0 \leq k < |P|)$. We will show that property 1 holds upon the next entrance of the loop statement with $|P - R| = k + 1$. Below, we will use the subscript *prev* to refer to the previous value of a variable in the algorithm.

Let $|P - R| = k + 1$ and consider M_{part} upon entering the loop statement. M_{part} is given by:

1. If *set_t* is assigned **false**, then $M_{part} = M_{part,prev} \cup \{\overline{p}\}$
2. If *set_t* is assigned **true** and *set_f* is assigned **false**, $M_{part} = M_{part,prev} \cup \{p\}$
3. If both *set_t* and *set_f* are assigned **true**, then $M_{part} = M_{part,prev}$

¿From the algorithm we also have that $R_{prev} = R \cup \{p\}$ (note $R_{prev} = \text{ELEM}(\textit{p\_remain,prev})$). Let M_R be an arbitrary truth assignment for R, $M^+_{R_{prev}}$ be a model for R_{prev} that agrees with M_R and assigns T to p and $M^-_{R_{prev}}$ be a model for R_{prev} that agrees with M_R and assigns F to p. By the induction hypothesis, we have that if M is a model for P that agrees with both $M_{part,prev}$ and $M^+_{R_{prev}}$, and M' a model for P that agrees with $M^+_{R_{prev}}$, then $(M \leq M')$ iff M' agrees with $M_{part,prev}$. And a similar property w.r.t. $M^-_{R_{prev}}$. We will now show that property 1 holds for M_{part} given by each of the cases listed above.

Consider the first case. It follows form the notion of "agrees with," and the fact that $M_{part} = M_{part,prev} \cup \{\overline{p}\}$ that the set of models for P that agree with M_{part} and M_R is identical to the set of models that agree with $M_{part,prev}$ and $M^-_{R_{prev}}$. Now let M be a model for P that agrees with both M_{part} and

M_R ; so, M agrees with $M_{part,prev}$ and $M^-_{R_{prev}}$. And, let M' be a model that agrees with M_R. We will show that $(M{\leq}M')$ iff M' agrees with M_{part}.

(if) Assume that M' agrees with M_{part}. So, M' agrees with $M_{part,prev}$. Also, since M' assigns F to p and M' agrees with M_R, it follows that M' agrees with $M^-_{R_{prev}}$. Now because, M agrees with both $M_{part,prev}$ and $M^-_{R_{prev}}$, it follows from the induction hypothesis that $M{\leq}M'$.

(only if) Assume that $M{\leq}M'$. If M' assigns F to p, then M' agrees with $M^-_{R_{prev}}$. So, since M agrees with both $M_{part,prev}$ and $M^-_{R_{prev}}$, it follows from the induction hypothesis that M' agrees with $M_{part,prev}$. And since, $M_{part} = M_{part,prev} \cup \{\overline{p}\}$, M' agrees with M_{part}. Otherwise, assume that M' assigns T to p. We will show that this assumption is inconsistent with the assumption that $M{\leq}M'$. ¿From $M{\leq}M'$ it follows that there exists a sequence s of rules that leads from M to M'. Let M'' be the first model in this sequence that assigns T to p, and let $d : \alpha{\rightarrow}p$ be the rule in s that leads to this model from a model M''' (note that $\alpha \in P - R - \{p\}$). If $M_{part,prev}$ satisfies α and d is not blocked at $M_{min,prev}$ (see algorithm), then the algorithm will assign **true** to *set_t*, contradiction. If $M_{part,prev}$ satisfies α and d is blocked at $M_{min,prev}$ by a rule $d_0 : (\beta \cup \alpha){\rightarrow}\overline{p}$, then there is at least one letter $p_0 \in \beta$ such that p_0 is assigned T in $M_{min,prev}$, while p must be assigned F in M''' because we can assume w.l.o.g. that d is not blocked at M'''. Since p_0 must be an element of $P - R - \{p\}$, it follows that M''' does not agree with $M_{part,prev}$. And therefore, by the induction hypothesis, $\neg(M{\leq}M''')$, contradiction. Finally, if $M_{part,prev}$ does not satisfy α, then there exists a letter $p_0 \in \alpha$ assigned F in $M_{part,prev}$. Let M'' be the first model in the sequence s that assigns T to p_0. Since M'' does not agree with $M_{part,prev}$ it follows by the induction hypothesis that $\neg(M{\leq}M'')$, contradiction.

We now consider the second case. Property 1 follows by an argument similar to the one given for the first case. Note that the condition that *set_t* is assigned **true** prevents a potential overlap with the first case. An algorithm that assigns p to T instead of F in case both *set_t* and *set_f* are assigned **false** will also converge on a maximal model.

Finally, we consider the third case: assume that both *set_t* and *set_f* are assigned **true**, then $M_{part} = M_{part,prev}$. When we ignore potential blocking for a moment, it follows form the algorithm that given a model M that agrees on $M_{part,prev}$, there exists a rule $d : \alpha{\rightarrow}p$ such that d leads to a model $M' = M|p$. In general, taking blocking into account, it can be shown that there exists a sequence of rules that leads from M to M'. Similarly, one can reach a model $M'' = M|\overline{p}$ from M.

Let M be a model for P that agrees with $M_{part} = M_{part,prev}$ and with M_R. W.l.o.g. we will assume that M assigns p to T. So, M agrees with $M^+_{R_{prev}}$. Let M' be a model that agrees with M_R. We will show that $(M{\leq}M')$ iff M' agrees with M_{part}. Let $M'' = M|\overline{p}$. Thus, as argued above, we have $M''{\leq}M$ and $M{\leq}M''$.

(if) Assume that M' agrees with $M_{part} = M_{part,prev}$. Now if M' assigns T to p, it follows that M' agrees with $M^+_{R_{prev}}$, and therefore, by the induction hypothesis we have $(M \leq M')$. Otherwise, if M' assigns F to p, it follows that M' agrees with $M^{-}_{R_{prev}}$. Thus, by the induction hypothesis, $M'' \leq M'$. And, because $M \leq M''$, we have $M \leq M'$.

(only if) Assume that $M \leq M'$. If M' assigns T to p, then M' agrees with $M^{+}_{R_{prev}}$. Therefore, by the induction hypothesis M' agrees with $M_{part,prev} = M_{part}$. Otherwise, if M'assigns F to p, then M' agrees with $M^{-}_{R_{prev}}$. Now, since $M'' \leq M$, we have $M'' \leq M'$. And therefore, again by the induction hypothesis, it follows that M' agrees with $M_{part,prev} = M_{part}$.

Property 2. Let $R = \emptyset$. Property 1 becomes: given an arbitrary model M' for P, if M is a model for P that agrees on M_{part}, then $(M \leq M')$ iff M' agrees with M_{part}. Now, let M be a model for P that agrees on M_{part}, and let M'' be a model such that $M \leq M''$. By property 1, it follows M'' agrees with M_{part}. And therefore, again by property 1, $M'' \leq M$. So, M is a maximal model. A direct consequence of this property is that we can arbitrarily assign truth values to the letter not assigned in M_{part} when the main loop is exited; any model thus obtained will be a maximal model. ■

Proof of theorem 6: The algorithm will terminate when $R = \emptyset$. M_{max} returned by the algorithm agrees with M_{part}. So, by property 2 of lemma 12, M_{max} is a maximal model of DH^{+}_{a}.

Finally, we show that Max-Model-DH^{+}_{a} finds a maximal model in time $O(kn^2)$, where n is the number of literals in D and k the number of letters in P. This follows directly from the fact that the loop statement is executed at most k times, and each of the for statements requires at most time $O(n)$. ■

B Proof of theorem on the relation to default logic

Theorem 7: *Let DH^{+} be a set of specificity ordered model-preference defaults that does not contain a mutually blocking pair of defaults such as $\alpha \rightarrow q$ and $\alpha \rightarrow \overline{q}$ or a self-supporting default such as $(\alpha \cup q) \rightarrow q$. If the preference relation induced by DH^{+} is a partial order, then M is a maximal model of DH^{+} iff $Th(M)$ is an extension of the default logic theory $< f_{DL}(DH^{+}), \emptyset >$. (In $Th(M)$, M is taken to be a propositional theory; in this case, the set of literals representing the maximal model.)*

We will use the following lemmas in the proof of this theorem. For the definitions and terminology regarding default logic we refer to Reiter (1980) and Etherington (1986).

Lemma 13 *Let $\alpha_1, ..., \alpha_k$, and ϕ be conjunctions of literals. If $\alpha_1 \wedge ... \wedge \alpha_k$ does not contain a complementary pair of literals such as q and $\overline{q}$, then $\phi \models (\alpha_1 \vee ... \vee \alpha_k)$ iff $\exists i. \, \phi \models \alpha_i$.*

Proof: (if) Trivial. (only if) Proof by contradiction. Let $\phi \models (\alpha_1 \vee ... \vee \alpha_k)$. Assume that $\phi \not\models \alpha_i$ for $1 \leq i \leq k$. ¿From $\phi \not\models \alpha_i$ it follows that there exists a truth assignment that assigns T to each literal in ϕ and to at least one literal in $\overline{\alpha_i}$. Since $\alpha_1 \wedge ... \wedge \alpha_k$ does not contain a pair of complementary literals, it follows that there exists a model M that assigns T to each literal in ϕ and T to at least one literal in each $\overline{\alpha_i}$. So, $M \models \phi \wedge \overline{\alpha_1} \wedge ... \wedge \overline{\alpha_k}$. Therefore, $\phi \not\models \alpha_1 \vee ... \vee \alpha_k$, contradiction. Thus, $\exists i.\ \phi \models \alpha_i$. ■

Lemma 14 *Let d be a default logic rule of the following form:*

$$\frac{: q \wedge \alpha \wedge \overline{\beta_1} \wedge ... \wedge \overline{\beta_k}}{q}$$

where q is a single literal, α is a conjunction of literals, each β_i is a conjunction of positive literals none of which share a letter with literals in q or α, and $E = Th\{\phi\}$ where ϕ is a set of literals. The following properties hold:

(a) If d is a generating default for E w.r.t. $\emptyset$, then $\phi \not\models \overline{q}$, $\phi \not\models \overline{\alpha}$, and $\forall i.E \not\models \beta_i$.

(b) If d is not a generating default for E w.r.t. $\emptyset$, then $\phi \models \overline{q}$, $\phi \models \overline{\alpha}$, or $\exists i.\ E \models \beta_i$.

Proof: (a) Given that d is a generating default for E w.r.t. $\emptyset$, it follows that $\neg(q \wedge \alpha \wedge \overline{\beta_1} \wedge ... \wedge \overline{\beta_k}) \notin E$. Thus, $\neg[\phi \models (\overline{q} \vee \overline{\alpha} \vee \beta_1 \vee ... \vee \beta_k)]$. Therefore, $\phi \not\models \overline{q}$, $\phi \not\models \overline{\alpha}$, and $\forall i.E \not\models \beta_i$.

(b) Given that d is not a generating default for E w.r.t. $\emptyset$, it follows that $\neg[\neg(q \wedge \alpha \wedge \overline{\beta_1} \wedge ... \wedge \overline{\beta_k}) \notin E]$. Thus, $\phi \models (\overline{q} \vee \overline{\alpha} \vee \beta_1 \vee ... \vee \beta_k)$, and thus by lemma 13, it follows that $\phi \models \overline{q}$, $\phi \models \overline{\alpha}$, or $\exists i.\ E \models \beta_i$. ■

Proof of theorem 7: (if) Let M be an extension of of $< f_{DL}(DH^+), \emptyset >$. ¿From the form of the defaults it follows that any extension can be represented as the deductive closure of a set of literals. So, let $E = Th\{M\}$ with M a set of literals represent an extension. We will show that (a) M is a complete set of literals, and, subsequently, that (b) M is a maximal model of DH^+.

(a) We will show by contradiction that M is a complete set of literals. Assume that neither p nor $\overline{p}$ occurs in M. Consider the following cases: 1) There does not exists a rule for p in group B. Therefore, $(\emptyset \rightarrow \overline{p}) \in DH^+$ with corresponding default rule $d_1 = [: \overline{p} \wedge \overline{\beta_1} \wedge ... \wedge \overline{\beta_k}/\overline{p}]$ in group A. This rule cannot be a generating default for E w.r.t. $\emptyset$ because $\overline{p}$ is not in E. So, by lemma 14 and since $p \notin E$ it follows that there exists a b_i such that $M \models \beta_i$. Now, consider the default logic rule $d_2 = [: p \wedge \beta_i \wedge \gamma_1 \wedge ... \wedge \gamma_l/p]$ corresponding to the model-preference default $\beta_i \rightarrow p$. Again, this rule cannot be a generating default, so there must exist a default $\beta_i \wedge \gamma_j \rightarrow \overline{p}$ with $M \models \beta_i \wedge \gamma_j$. By repeating this argument one will obtain a model-preference default $\delta \rightarrow p$ (or $\overline{p}$) for which no more specific default exists. The associated default logic rule will be a generating default for E w.r.t. $\emptyset$, bringing in p (or $\overline{p}$) (Note that this argument relies on the fact that rules can be blocked only by a rule with a more specific left-hand side, *i.e.*, on the absence of pairs of mutually blocking rules such as $\alpha \rightarrow q$ and $\alpha \rightarrow \overline{q}$.) 2) There does exist a rule for p in group B. By again considering a sequence of default logic rules corresponding to more and more specific model-preference rules, just as argued in case 1), we obtain a contradiction. ¿From 1) and 2), it follows that M must be a complete set of literals.

(b) We will now show that M corresponds to a maximal model of DH^+. Let $E = Th\{M\}$ be an extension of $f_{DL}(DH^+)$. By (a), it follows that M is a complete set of literals. Assume that M is not a maximal model of DH^+. Therefore, there exists a rule $d : (\alpha \rightarrow \overline{q}) \in DH^+$ that is applicable at M and leads to model M' distinct from M. Since M is an extension of $f_{DL}(DH^+)$, it follows that q is supported by some generating default. Assume that q is supported by a rule in group B [: $q \wedge \overline{\gamma_1} \wedge ... \wedge \overline{\gamma_k}/q$]. From lemma 14 it follows that $M \not\models \gamma_i$ for $1 \leq i \leq k$. Now, since $\alpha = \gamma_i$ for some i, it follows $M \not\models \alpha$, contradiction. Therefore, q is supported by some generating default in group A. Assume q is supported by the rule $d_{DL} = [: q \wedge \beta \wedge \overline{\delta_1} \wedge ... \wedge \overline{\delta_o}]$ corresponding to the model-preference default $d' : \beta \rightarrow q$. Since d_{DL} is a generating default and M is a complete set of literals, it follows using lemma 14 that $M \models q \wedge \beta \wedge \overline{\delta_1} \wedge ... \wedge \overline{\delta_o}$. Consider the model $M' = M|\overline{q}$. It follows that d' is applicable at M' (note that q does not occur in β since we do not allow self-supporting rules such as $\alpha \wedge q \rightarrow q$). So, we have $M \leq M'$ (by rule d) and $M' \leq M$ (by rule d'). Contradiction, since the preference ordering is a partial ordering. So, M is a maximal model of DH^+.

(only if) Let M be a maximal model of DH^+. We will show that each literal is supported by a generating default. Since the rules have empty prerequisites, this means that there exists a converging sequence of default rule applications that brings in all the literals in M. Moreover, since M is a complete set of literals no other literals can be brought in, and therefore $Th(M)$ is an extension $f_{DL}(DH^+)$.

We will now show that each literal M is supported by some generating default for $Th\{M\}$ w.r.t. $\emptyset$. Let q be an arbitrary literal in M. Consider the following possibilities: 1) There exists a rule $d : \alpha \rightarrow q \in DH^+$ and d is applicable at M. Consider the corresponding default logic rule $d_{DL} = [: q \wedge \alpha \ \wedge \overline{\beta_1} \wedge ... \wedge \overline{\beta_k}/q]$ in group A. Since d is applicable at M, it follows that d_{DL} is a generating default for $Th\{M\}$ and supports q. 2) There does not exist a rule in DH^+ that is applicable at M. Consider the following rule in group B: $d_{DL} = [: q \wedge \overline{\gamma_1} \wedge ... \wedge \overline{\gamma_l}/q]$ in which for each γ_i there exists a rule $\gamma_i \rightarrow \overline{q}$ in DH^+. (Below we will consider the case where no such rule exists.) We will show by contradiction that d_{DL} is a generating default for $Th\{M\}$. Assume that d_{DL} is not a generating default for $Th\{M\}$. Therefore, by lemma 14, it follows that there exists a γ_i such that $M \models \gamma_i$. Now, if $d : \gamma_i \rightarrow \overline{q}$ is applicable at M then M is not a maximal model (note that preference ordering is a partial ordering), contradiction. So, d is blocked at M by a more specific default $d' : \gamma_i \wedge \delta \rightarrow q$ and $M \models (\gamma_i \wedge q)$. Since, by assumption, no rule of the form $\alpha \rightarrow q$ is applicable at M, the rule d' must be blocked by a more specific default in DH^+. After, a certain number of repetitions of the above argument, one will encounter a most specific rule which cannot be blocked by a more specific rule. Since by assumption no rule of the form $\alpha \rightarrow q$ is applicable at M, the rule must be of the form $\delta \rightarrow \overline{q}$ with $M \models \delta$. But, this would imply that M is not maximal since the preference relation is a partial ordering, contradiction. (As in the (if) direction, argument relies on the absence of pairs of mutually blocking defaults.) Finally, assume there does not exist a default for q in group B. Therefore, $\emptyset \rightarrow \overline{q} \in DH^{+}$. Since M is maximal this rule must be blocked by a rule of the form $\beta \rightarrow q$. Now, by a similar argument as used above, we obtain a contradiction. ¿From 1) and 2) it follows that each literal in M is supported by some generating default for $Th\{M\}$ w.r.t. $\emptyset$. ■