On Heterogeneous Model-Preference Default Theories*

Fabrizio Sebastiani†
Istituto di Elaborazione dell'Informazione
Consiglio Nazionale delle Ricerche
Via S. Maria, 46 - 56126 Pisa (Italy)

Abstract

Default systems based on the notion of "model-preference" have recently been proposed by Selman and Kautz to give a semantic account of the phenomena involved in default reasoning and to provide a formal justification for the limited cognitive load that default reasoning seems to require of human beings. In this paper we argue that the way these formal systems have been defined makes them inadequate for the task of reasoning in the presence of both certain information and defeasible information. We propose a modification to the original framework and argue that it formalizes correctly the interaction between these two fundamentally different kinds of information. We then show that the proposed modification has also a positive effect on the complexity of model-preference default reasoning.

1 Introduction

Default reasoning plays an important role in everyday practical reasoning. Agents, be they natural or artificial, typically face situations in which they have to act and make decisions on the basis of a body of knowledge that is far from being an exhaustive description of the domain of discourse; this is a direct consequence both of the limited capacity of their physical repositories of knowledge, and of the fact that the processes involved in the acquisition of knowledge (both from external sources—e.g. books—and internal ones—e.g. speculative reasoning) are computationally demanding and time consuming.

Nevertheless, action and decision-making often require more knowledge than our agents actually possess, thus forcing them to make up for the limited coverage of their knowledge bases by means of "default" assumptions. As the name implies, "assumptions" have an

*This paper appears in the *Proceedings of CSCSI-90, 8th Biennial Conference of the Canadian Society for Computational Studies of Intelligence*, Ottawa, ONT, 1990, pp. 84–91. Edited by Peter F. Patel-Schneider.

†Current address: Department of Computer Science, University of Toronto, MS5 1A4 Toronto, Ontario, Canada. E-mail: fabrizio@ai.toronto.edu

epistemic status that is far from being solid, as they can actually be invalidated by further reasoning or by the future acquisition of empirical data. These phenomena are well-known in cognitive science, and have sometimes been taken to imply that a great deal of human reasoning does not conform to the canons of "logic" and hence escapes attempts at formalization [5]. Doubtless, the overall effectiveness of human reasoning testifies to the effectiveness of this modality of reasoning too: humans are much quicker at finding surrogates of missing knowledge than at actually acquiring this knowledge, either through reasoning or empirical investigation. And, above all, once these surrogates have been found, humans are much quicker at reasoning on the resulting complete, albeit epistemically shakier, description of the domain of discourse than they would be had they to rely on the smaller part of this description that they trust as being accurate *tout court*. This observation is at the heart of the recent interest that the knowledge representation community has shown in *vivid knowledge bases* [1, 8, 9], i.e. exhaustive descriptions of the domain of discourse consisting of collections of atomic formulae[1]. Reasoning on these KBs, which may be considered as "analogues" of the domain being represented, is easily shown to be efficient.

It is precisely in the face of such empirical considerations that the bad computational properties of current formalisms that attempt to formalize default reasoning (such as the ones based on Circumscription [11, 12] or on Autoepistemic Logic [13, 6]) are particularly disturbing: arguably, a formalism for default reasoning not only should account for the conclusions that agents draw in the presence of incomplete information, but it also should possess radically better computational properties than formalisms aimed at reasoning tasks at which humans are notoriously inefficient (such as e.g. first order logic in the case of deductive reasoning). Moreover, it is certainly not plausible that, in order to arrive at knowledge bases upon which fast reasoning can be carried out, humans use a dramatically inefficient reasoning method.

These considerations lead us to look with special interest at formalizations of default reasoning that emphasize computational tractability. In their recent paper "The complexity of model-preference default theories" [18], Selman and Kautz describe $\mathcal{DH}_a^+$, a tractable system for performing inferences on theories of Acyclic Horn Defaults. Their framework has the added appeal of possessing a strong model-theoretic flavour[2], which allows thorough investigations in the mechanisms underlying the inference processes. It is by means of such investigations, however, that one can discover that the systems of [18] have an odd behaviour when confronted with knowledge bases that consist of both certain information and default information. This is especially disturbing, as we surely would like to account for the fact that agents, although making heavy use of default reasoning, normally do also

[1]In formally introducing vivid KBs [8] actually situates his discussion in the framework of the first order predicate calculus; hence, for him a vivid KB is "a collection of ground, function-free atomic sentences, inequalities between all different constants (...), universally quantified sentences expressing closed world assumptions (...) over the domain and over each predicate, and the axioms of equality". As our discussion will be situated in the framework of the propositional calculus, we will take this definition of vivid KB instead.

[2]A semantics for Selman and Kautz's model-preference default systems that fully embraces the model-theoretic credo is described in [17]

possess information upon which they rely with special confidence. Accordingly, a reasoning system should enforce a correct interaction between certain knowledge and default knowledge, and account for the different impact that the two kinds of knowledge have on the overall reasoning process. It is these considerations which inform the attempt, described in this paper, to tune model-preference default systems in such a way as to make them behave correctly with respect to the distinction between certain knowledge and defeasible knowledge.

In order to make this paper self-contained, in Section 2 we give a brief overview of $\mathcal{D}^+$, the most general system described in [18], of which $\mathcal{DH}_a^+$ is a tractable subset[3]. In Section 3 we argue that the two methods proposed in [18] for dealing with the co-presence of certain information and defeasible information in $\mathcal{D}^+$ are, for different reasons, both inadequate, and characterize two interesting classes of reasoning tasks that are mishandled by both methods; we proceed to spell out our modifications to $\mathcal{D}^+$ and to argue that the system so obtained handles well the interaction between certain knowledge and defeasible knowledge (including the two classes of reasoning tasks that had revealed problematic for the original version of $\mathcal{D}^+$). In Section 4 we examine the effects of our modifications on the computational complexity of model-preference default reasoning; such modifications allow us to establish new results for reasoning in the presence of both certain knowledge and defeasible knowledge, and to discover that the presence of certain knowledge has a beneficial effect on the efficiency of the reasoning process. Section 5 concludes.

2 An overview of Selman and Kautz's system $\mathcal{D}^+$

Roughly speaking, the idea around which the systems of [18] revolve is that the import of a default $d \equiv \alpha \rightarrow q$ is to make a model (that is, a complete specification of what the domain of discourse is like) where both α and q are true be *preferred* to another model where α is true but q is not. By combining the effects of the preferences due to the single defaults, a set of defaults identifies a set of "maximally preferred" models; these models, isomorphic as they are to vivid knowledge bases, are meant to represent possible ways in which the agent may "flesh out" his body of certain knowledge by the addition of defeasible knowledge. For instance, according to a set of defaults such as $\{a \rightarrow b,\ b \rightarrow c\}$, the model where a, b and c are all true would be a maximally preferred model. However, the systems in [18] also account for the fact that a more specific default should override a less specific one, and they do so by "inhibiting", where a contradiction would occur, the preference induced by the less specific default; this is meant to prevent a set of defaults such as $\{a \rightarrow b,\ b \rightarrow c,\ ab \rightarrow \neg c,\ a\neg b \rightarrow \neg c\}$ to generate maximally preferred models where a and c are both true.

The first thing we need to do in order to introduce $\mathcal{D}^+$ in detail is to describe what the

[3] In this paper we will implicitly rule out from consideration the system $\mathcal{D}$, as its lack of commitment to any specificity ordering between defaults (see below) makes it less interesting than the other systems of [18]. The other systems discussed in [18], $\mathcal{DH}^+$ and $\mathcal{DH}_a^+$, are restrictions of $\mathcal{D}^+$ to the Horn case and to the Horn Acyclic case, respectively; all modifications that are described in this paper apply straightforwardly to these more restricted systems.

language for representing knowledge in $\mathcal{D}^+$ is. Let $P = \{p_1, p_2, \ldots, p_n\}$ be a finite set of propositional letters, and L be the language of formulae built from P and the connectives $\neg$, $\wedge$ and $\vee$ in the standard way. We define a *default* d to be an expression of the form $\alpha \rightarrow q$, where q is a literal (i.e. a propositional letter p in P or its negation $\neg p$) and α is a set of literals[4]. We will also use the standard definition of a *model* for P as a function $M : P \mapsto \{\texttt{True}, \texttt{False}\}$; accordingly, we will say that M satisfies a theory T of L (written as $M \models T$) iff M assigns `True` to each formula in T, formulae in T being evaluated with respect to M in the standard manner.

The above-mentioned specificity ordering between defaults is captured by stipulating that, given a set of defaults (or *default theory*) D, a default $d \equiv \alpha \rightarrow q$ in D is *blocked* at a model M iff there exists a default d' in D such that $d' \equiv \alpha \cup \beta \rightarrow \neg q$ and $M \models \alpha \cup \beta$. A default $d \equiv \alpha \rightarrow q$ is then said to be *applicable* to a model M iff $M \models \alpha$ and d is not blocked at M. If d is applicable at M, the model $d(M)$ is defined as the model which is identical to M with the possible exception of the truth assignment to the propositional letter occurring in q, which is assigned a truth value such that $d(M) \models q$.

Naturally enough, a preference ordering induced on models by a set of defaults D may at this point be defined. Given a set of defaults D, the relation "$\leq+$" is defined to hold between models M and M' (written $M \leq+ M'$) iff there exists d in D such that d is applicable at M and such that $d(M) = M'$. The relation "$\leq$" is defined as the transitive closure of "$\leq+$"[5]. Finally, we will say that a model M is *maximally preferred* (or *maximal*) with respect to a set of defaults D iff for all models M' either $M' \leq M$ is the case or $M \leq M'$ is not the case. We understand the task of reasoning in $\mathcal{D}^+$ as that of finding an arbitrary model which is maximal with respect to a given default theory D.

We will illustrate the way $\mathcal{D}^+$ works by means of an example[6]. Let $P = \{a, b, c\}, D = \{a \rightarrow b, b \rightarrow c, ab \rightarrow \neg c, a\neg b \rightarrow \neg c\}$. $\neg$`abc`, $\neg$`a`$\neg$`bc`, `ab`$\neg$`c` and $\neg$`a`$\neg$`b`$\neg$`c` are all and the only maximal models. Note that if $b \rightarrow c$ had not been blocked at `ab`$\neg$`c`, then `abc` would have been maximal too, contrary to intuitions. The example is represented graphically in Figure 1.

3 Dealing with heterogeneous theories

In the preceding section we have described the process by which the set of maximal models is singled out from the set of all models of P through the application of a set of defaults

[4]For notational convenience we will omit braces in antecedents of defaults. Hence we will write e.g. $ab \rightarrow \neg c$ instead of $\{a, b\} \rightarrow \neg c$.

[5][18] defines "$\leq$" to be the *reflexive* transitive closure of "$\leq+$"; that this is redundant may be seen by inspecting the way "$\leq$" is used in the definition of maximal model.

[6]In the drawings of the following examples, rectangles will denote models represented in the obvious way (e.g. `a`$\neg$`bc` will represent the function that assigns `True` to a and c and `False` to b). Arrows will represent "$\leq+$" relationships. A slashed arrow will represent what would have been a "$\leq+$" relationship unless a blocking had not occurred. Also, we will omit arrows corresponding to self-loops (i.e. arrows starting and ending in the same model) as they do not contribute in supporting the maximality or non-of a model.

Figure 1: A simple example

D. It is natural, however, to require that a method be also enforced that allows a theory T of certain facts to be brought to bear in the process of maximal model selection. For instance, we might want to represent the situation in which, beside knowing that $a \rightarrow b, b \rightarrow c, ab \rightarrow \neg c$ and $a\neg b \rightarrow \neg c$, the agent also knows for sure that a is the case. In this case $\neg\mathtt{abc}$, $\neg\mathtt{a}\neg\mathtt{bc}$ and $\neg\mathtt{a}\neg\mathtt{b}\neg\mathtt{c}$ should no longer be maximal models, and $\mathtt{ab}\neg\mathtt{c}$ only should be endorsed. There are two methods that are described in [18] for bringing to bear certain knowledge in the process of maximal model selection, and their adequacy to implement a correct interaction between certain and defeasible knowledge will be a central concern of this paper. But in order to do so, we will first need to make a short digression on what we consider an "adequate" method of implementing it.

Until now we have largely proceeded on a formal level only, and are thus in need of providing some empirical justification for the formalism we have chosen. Although identifying the task of generating a vivid KB with that of finding a maximally preferred model sounds fairly intuitive, it is by no means clear why the notion of preference we have formalized should be "the right notion of preference" at all. What we need is a criterion of empirical adequacy which is independent of the formalization itself, a criterion that allows us to judge if our systems actually capture the relevant intuitions behind default reasoning as generation of a vivid KB. Different theorists might obviously have different intuitions concerning this issue; nevertheless, we will lay down our cards and describe the two minimal conditions which *we* think reasonable for a model M to be maximal with respect to a theory T and a default theory D:

1. M satisfies the theory T;

2. if M satisfies the antecedent α of a default $d \equiv \alpha \rightarrow q$ in D, then it also satisfies its consequent q, unless it also satisfies the antecedent $\alpha \cup \beta$ of a default $d' \equiv \alpha \cup \beta \rightarrow \neg q$ in D.

Someone might object that there is a lot of "tester's bias" in this criterion, and that it seems to have been laid out in order to acritically grant a stamp of adequacy to the formal definition we have presented in Section 2. We will see that this is not so, and that this criterion is demanding enough to rule out the definition of maximality enforced by $\mathcal{D}^+$. Quite immodestly, we will call a model satisfying conditions 1 and 2 an *intended model*; we will thus consider a model-preference default system empirically satisfactory iff for every set of defaults D every intended model is also a maximal model and vice-versa. For instance, it is easy to check that this is indeed the case in the example of Section 2.

We may now switch back to the description of the two methods that are proposed in [18] in order to account for the interaction between certain and defeasible knowledge. The first of them consists in encoding certain facts by means of defaults with an empty antecedent; this style of encoding relies on the fact that the ability to reach a conclusion starting from an empty set of premises is usually taken as meaning that the conclusion is a true fact. The situation described above would then be represented by $P = \{a, b, c\}, D = \{\phi \rightarrow a, a \rightarrow b, b \rightarrow c, ab \rightarrow \neg c, a\neg b \rightarrow \neg c\}$; in this case `ab¬ c` is in fact the only maximal model, as shown in Figure 2.

Figure 2: Defaults with empty antecedents

But we feel that this is an unsatisfactory solution, as encoding certain facts as defaults with empty antecedents *exposes them to blockage by more specific defaults*, i.e. by items of knowledge that, although being more specific, have a weaker epistemic status than any item of certain knowledge. That this solution licenses undesired conclusions is shown by the following example. Let $P = \{a, b, c\}, D = \{a \rightarrow b, b \rightarrow c, ab \rightarrow \neg c, a\neg b \rightarrow \neg c, \phi \rightarrow a, \phi \rightarrow \neg b\}$, where the intended interpretation of the last two defaults is "a is certainly the case" and "$\neg b$ is certainly the case". The only intended model is `a¬b¬c` but, as shown in Figure 3, it is not a maximal model. On the contrary, `ab¬c` is maximal (actually, it is the only maximal model) but is not intended. Basically, this happens because $a \rightarrow b$ is allowed to block $\phi \rightarrow \neg b$ at models where a is true. Instead, we think that an item of defeasible knowledge should never be allowed to invalidate an item of certain knowledge; rather, the opposite view should be enforced, with certain knowledge inhibiting the effect of defeasible knowledge when a contradiction would otherwise occur. The preceding example shows that, so to speak, this semantics (or, better, this semantics together with the style of encoding certain knowledge that it encourages) is *neither sound nor complete* with respect to the intuitive, pre-theoretical semantics of default reasoning supposedly embodied by our adequacy criterion.

Figure 3: Improper behaviour

Let us then consider the second solution to the integration of certain and defeasible knowledge that is described in [18]. This solution relies on a different definition of maximality, according to which a model is maximal with respect to what we will call a *heterogeneous theory* $\langle D, T \rangle$ (where D is a default theory and T is a theory) iff $M \models T$ and for all models M' such that $M' \models T$ either $M' \leq M$ is the case, or $M \leq M'$ is not the case. In other words, in order to be maximal a model must first of all satisfy T, and its would-be maximality can only be prevented by another model which itself satisfies T. We can see that the modified system handles the preceding example correctly: let $P = \{a, b, c\}, D = \{a \rightarrow b, b \rightarrow c, ab \rightarrow \neg c, a\neg b \rightarrow \neg c\}, T = \{a, \neg b\}$. As remarked above `a¬b¬c` is the only intended model; as shown in Figure 4, it is now also the only maximal model[7].

Figure 4: Maximality wrt heterogeneous theories

But we feel that this solution too is unsatisfactory, and again we feel that the reason lies in an overestimation of the role of defeasible knowledge in its interaction with certain knowledge. To see what the problems involved are, let us introduce two new definitions. We will say that there is an *internal path* between two models M and M' (written $M \leq^i M'$) iff there exist models $M_1, \ldots, M_n$ (with $M \equiv M_1$ and $M' \equiv M_n$) such that $M_i \leq+ M_{i+1}$ for

[7] In the drawings of this section grey areas represent the set of models that satisfy T.

all $i = 1, \ldots, n-1$ and such that $M_i \models T$ for all $i = 1, \ldots, n$. Conversely, we will say that there is an *external path* between two models M and M' such that $M \models T$ and $M' \models T$ (written $M \leq^e M'$) iff there exist models $M_1, \ldots, M_n$ (with $M \equiv M_1$ and $M' \equiv M_n$) such that $M_i \leq + M_{i+1}$ for all $i = 1, \ldots, n-1$ and $M_j \not\models T$ for some $j = 2, \ldots, n-1$. Intuitively, an external path is a path which goes through at least one model that does not satisfy T.

Our qualms with the (second) definition of maximality described in [18] have to do with the fact that it still allows defeasible knowledge to override certain knowledge that contradicts it, and accomplishes this *by allowing external paths to support either the maximality or non- of a model.* We therefore proceed to give a new definition of maximality, one where external paths are ruled out from consideration, and argue that in all cases in which the original definition differs from the new one, the former yields unintuitive results while the latter does not.

The basic step of this new definition is actually the relativisation of the preference ordering "$\leq +$" with respect to a theory T. We hereby define the relation "$\leq +$" wrt a heterogeneous theory $\langle D, T \rangle$ as the relation which holds between models M and M' iff $M \models T, M' \models T$ and there exists d in D such that d is applicable to M and such that $d(M) = M'$. As in the original version, "$\leq$" is defined as the transitive closure of "$\leq +$", and M is maximal wrt $\langle D, T \rangle$ iff $M \models T$ and for all M' such that $M' \models T$ either $M' \leq M$ is the case or $M \leq M'$ is not the case. It is simple to check that defining maximality this way is equivalent to substituting "$\leq^i$" for "$\leq$" in the original definition. In order to distinguish between the two notions of maximality we will henceforth speak of *$\leq^i$-maximality* and *$\leq$-maximality*, respectively.

In what cases $\leq^i$-maximality and $\leq$-maximality yield different results may be checked by truth table analysis. This analysis, although straightforward, is fairly tedious, and is then confined to Appendix A. The basic result is that there are two cases in which $\leq^i$-maximality and $\leq$-maximality differ: we will call instances of the former case *type-1 theories* and instances of the latter *type-2 theories*. Let us analyze them orderly.

In type-1 theories we have a model M that is not intended and is not $\leq^i$-maximal but is $\leq$-maximal. This is caused by the presence of an external path leading from M' to M and supporting the $\leq$-maximality of M, and by the absence of a similar but internal path to support its $\leq^i$-maximality. The following example is a type-1 theory. Let $P = \{a, b, c\}, D = \{ab \to c, c \to \neg a, \neg ab \to \neg c, b\neg c \to a\}, T = \{a, b\}$. `abc` is the only intended model and, as shown in Figure 5, is also the only $\leq^i$-maximal model. `abc` is also $\leq$-maximal; note, however, that also `ab¬c` is $\leq$-maximal, although it is not intended.
In type-2 theories we have quite the opposite situation, i.e. there is a model M that is intended and is $\leq^i$-maximal but is not $\leq$-maximal. This is caused by the presence of an external path leading from M to a model M' which prevents the $\leq$-maximality of M, and by the absence of a similar but internal path to prevent its $\leq^i$-maximality. The following example is a type-2 theory. Let $P = \{a, b, c\}, D = \{b\neg c \to \neg a, \neg ab \to c, bc \to a\}, T = \{a, b\}$. `abc` and `ab¬c` are the only intended models and, as shown in Figure 6, are also the only $\leq^i$-maximal models; `abc` is also $\leq$-maximal but `ab¬c` is not, although it is intended.
These examples show that also the second solution described in [18] to the problem of model-preference reasoning in heterogeneous default theories is, so to speak, *neither sound*

Figure 5: Improper behaviour. A type-1 theory

.

Figure 6: Improper behaviour. A type-2 theory

(as shown by type-1 theories) *nor complete* (as shown by type-2 theories) with respect to the intuitive, pre-theoretical semantics of default reasoning. Quite apart from this, they also show that the problem lies in paths involving models that do not satisfy the theory T, and that $\leq^i$-maximality, by excluding these paths from consideration, actually implements the correct behaviour.

From an empirical standpoint, the idea that these paths and models should be neglected is informed by the general principle according to which the co-presence of knowledge endowed with a higher epistemic status (or, according to the terminology of [2], knowledge that is more "epistemically entrenched") should have the reasoning process disregard knowledge endowed with a lower epistemic status that contradicts it. In our case, this means that the application of a default to a state of affairs which is inconsistent with the certain knowledge is deemed not only irrelevant but actually misleading, and so is the application of a default that yields such a state of affairs.

It might be legitimate, at this point, to ask ourselves whether $\leq^i$-maximality is sufficient in itself to make $\mathcal{D}^+$ and the other [18] systems comply with the adequacy criterion in Section 3. Somehow surprisingly, the answer is no. This is due to the fact that the adequacy criterion we have laid out can be met only by formalisms that make "$\rightarrow$" enjoy *contraposition*, the property of some conditional notions "$\Rightarrow$" (e.g. material implication in classical logic, strict implication in modal logic) according to which $a \Rightarrow b$ entails $\neg b \Rightarrow \neg a$. We indeed think that the pre-theoretical notion of default implication (if only there exists such a thing) *does* enjoy contraposition, and this is why we did not yield to the temptation of adding "...or unless $\neg q$ belongs to T" to clause 2 of the criterion, an addition that would have exempted a formalism from endorsing contraposition. Neither the systems of [18] nor their modifications as resulting from $\leq^i$-maximality endorse it. Nevertheless, modifying them to this effect would be an easy matter (essentially, this would involve redefining $d(M)$ to be a set of models rather than a single model). More importantly, contraposition is not an issue which is specific to the interrelation of certain and defeasible knowledge (it is surely pertinent also to defeasible knowledge alone), and therefore falls outside the scope of this paper. We may then legitimately consider $\leq^i$-maximality as having successfully accomplished the goal of meeting the adequacy criterion.

4 Complexity issues

As one of the most important contributions of [18] is the establishment of complexity results for reasoning in model-preference default theories, it would be useful to understand whether these results carry over to systems where $\leq^i$-maximality is used, and whether new results can be established for such systems. As $\leq$-maximality and $\leq^i$-maximality coincide when the theory T is empty, we will be interested in results concerning the case when T is non-empty. These two cases, that we will dub the "homogeneous" and the "heterogeneous" case, respectively, have been shown in [18] to have, in general, different complexity. For example, the problem of reasoning in $\mathcal{DH}$ is linear in the homogeneous case, while in the heterogeneous case it is polynomial if T is a set of literals and NP-hard if T is a set of Horn clauses. Also, for some of these systems results are known for the homogeneous case but not for the heterogeneous case; in particular, for the homogeneous case NP-hardness results for $\mathcal{D}$ and $\mathcal{DH}^+$ and a polynomial result for $\mathcal{DH}^+_a$ are reported in [18], but the corresponding results for the heterogeneous case are still unknown.

As remarked in [18], the problem of finding a model which is maximal wrt a pair $\langle D, T\rangle$ is best viewed as a search problem [3]. The main result of this paper is that, for any model-preference default system considered in [18] and variations thereof, if T is a set of literals and $\leq^i$-maximality is used, the search problem in the heterogeneous case is no harder than the homogeneous case. The result is based on the key observation that, in both cases, only internal paths are allowed (internal to the set of all models of the language P in the former case, and to the set of all models satisfying T in the latter), whereas, according to $\leq$-maximality, external paths are allowed in the heterogeneous case but not in the homogeneous case.

Theorem 1 *The search problem for $\mathcal{S}^i$ with HD-theories $\langle D, T\rangle$ (with T a set of literals) belongs to the same complexity class of the search problem for $\mathcal{S}$ with empty T, where $\mathcal{S}$ is any model-preference default system discussed in [18] and $\mathcal{S}^i$ is the version of $\mathcal{S}$ relying on $\leq^i$-maximality.* ∎

In [16] we prove this theorem by giving an algorithm that, given an HD-theory $\langle D, T\rangle$ (where T is a set of literals) defined on some alphabet P, builds a default theory D' defined on $P - P_T$ (where P_T is the set of propositional letters in P that occur in T) such that, whenever a model M is $\leq$-maximal wrt D', the model $M \cup T$ is $\leq^i$-maximal wrt $\langle D, T\rangle$[8]. This algorithm is linear in the number of occurrences of literals in D. Essentially, what the algorithm does is inspecting the set D, eliminating from it any default that does not contribute to the determination of $\leq^i$-maximality (either because its consequent is in T, or because its antecedent or its consequent are inconsistent with T) and eliminating from the antecedents of the remaining defaults all literals that already appear in T. The models of P that satisfy T and the preference relations between them that are induced by D form a graph that is isomorphic to the one formed by the models of $P - P_T$ and by the preference relations between them that are induced by D'. Due to the isomorphism, there is a one-to-one correspondence between $\leq^i$-maximal models endorsed by the former graph and $\leq$-maximal models endorsed by the latter.

Table 1 summarizes complexity results in the heterogeneous case (with T a set of literals) for the main model-preference default systems.

	$\mathcal{D}$	$\mathcal{DH}$	$\mathcal{DH}^+$	$\mathcal{DH}^+_a$
$\leq$	?	P	?	?
$\leq^i$	NP-hard	linear	NP-hard	P

Table 1: The complexity of simple heterogeneous theories

Also, note that although the homogeneous and heterogeneous case have the same theoretical complexity, the heterogeneous case is in practice much simpler (somehow in contrast with what happens for $\leq$-maximality). For example, the search problem for the homogeneous case of $\mathcal{DH}$ is $O(n)$, where n is the number of occurrences of literals in D. Although the heterogeneous case (with $\leq^i$-maximality) of $\mathcal{DH}$ is still $O(n)$, n is now the number of occurrences of literals in D', a subset of D. That by adopting $\leq^i$-maximality the heterogeneous case should be simpler is also apparent from the fact that the presence of a theory T consisting of a set of k literals transforms the problem of searching the set of all models of the propositional language P into the one of searching only the set of those models that satisfy T: the latter graph has 2^k times less nodes than the former. This accounts for the rather intuitive fact that, as the amount of certain information increases, the amount of computational resources required to "flesh out" the knowledge base should proportionally decrease.

[8]For convenience here we think of a model as the largest set of literals that is satisfied by it.

5 Conclusion

We have shown how two methods proposed in [18] fail, for different reasons, to capture what we think the correct interaction between certain and defeasible information should be. The modified definition of maximality ("$\leq^i$-maximality") we have provided has been shown to be the notion of maximality that better represents our intuitions of how default reasoning should accommodate the interaction between these two fundamentally different types of information. When this notion of maximality is plugged in, not only reasoning in the presence of a theory T consisting of a set of literals is provably no harder than reasoning with an empty T, but is likely to be, in practice, much more efficient.

Acknowledgements

This research was carried out in the framework of the TELOS project, dealing with the design and construction of a knowledge representation language supporting the phase of requirements modelling in information systems development. I would like to thank John Mylopoulos for encouraging the work described here and for providing me the environment where this research was carried out. I would also like to thank David Lauzon and Bart Selman for our stimulating discussions about the topic of this paper.

A Comparing $\leq^i$-maximality and $\leq$-maximality

The truth table that establishes type-1 and type-2 theories as all and the only types of theories in which $\leq^i$-maximality and $\leq$-maximality differ is spelled out in Table 2. The $\leq^i$-maximality (column v) and $\leq$-maximality (column vi) of a model M satisfying a propositional theory T is represented as a function of the simple conditions $\forall M'.(M' \leq^i M)$ (column i), $\exists M'.(\neg M' \leq^i M \wedge\ M \leq^i M')$ (column ii), $\forall M'.(M' \leq M)$ (column iii) and $\exists M'.(\neg M' \leq M \wedge\ M \leq M')$ (column iv); all models considered in this Appendix satisfy T. Condition 5 is the conjunction of 1 and the negation of 2, while 6 is the conjunction of 3 and the negation of 4.

Note that 9 of the 16 rows are not even taken into consideration as, for different reasons, they depict impossible situations. For instance, rows 1 through 4 are ruled out on the grounds that the conditions corresponding to columns i and ii cannot be true at the same time, unlike the simultaneous presence of `True` in the first two columns of these rows would imply. Similarly, the simultaneous presence of `True` in columns iii and iv rules out rows 1, 5, 9 and 13; the simultaneous presence of `True` in columns i and iv rules out rows 1, 3, 5 and 7, and the simultaneous presence of `True` in column i and of `False` in column iii rules out rows 3, 4, 7 and 8. In all these cases, the incompatibility of the two conditions corresponding to the two columns at issue is easy to check.

Among the 7 rows that are left, rows 6, 11, 14 and 16 agree on the values assigned to $\leq^i$-maximality and $\leq$-maximality. Instead, rows 10, 12 and 15 assign different values to the two notions.

However, a closer analysis shows that the truth conditions corresponding to rows 10 and 12 actually come down to representing the same situation, as far as the motivations for the dissimilarity between $\leq^i$-maximality and $\leq$-maximality are concerned. In fact, row 10 corresponds to the situation in which $\exists M'.(\neg M' \leq^i M \wedge M \leq^i M')) \wedge \forall M'.(M' \leq M)$ while row 12 corresponds to the situation in which $\exists M'.(\neg M' \leq^i M \wedge M \leq^i M') \wedge \exists M''.(\neg M'' \leq M) \wedge \forall M'''.(M''' \leq M \wedge \neg M \leq M'''))$; the only difference is that in row 10 the existence of an "isolated model" (i.e. one which has no relation of preference whatsoever, either outgoing or incoming, with M) is stated, while the existence of such a model is ruled out in row 10. Since the mere existence of such a model does not have any influence in supporting the truth or falsity of the conditions corresponding to columns v and vi, we may consider the situations corresponding to rows 10 and 12 to be actually the same situation.

Row 10 and 12 correspond then to what we have called type-1 theories, while row 15 corresponds to type-2 theories.

	i	ii	iii	iv	v	vi
1	True	True	True	True		
2	True	True	True	False		
3	True	True	False	True		
4	True	True	False	False		
5	True	False	True	True		
6	True	False	True	False	True	True
7	True	False	False	True		
8	True	False	False	False		
9	False	True	True	True		
10	False	True	True	False	False	True
11	False	True	False	True	False	False
12	False	True	False	False	False	True
13	False	False	True	True		
14	False	False	True	False	True	True
15	False	False	False	True	True	False
16	False	False	False	False	True	True

Table 2: The functional behaviour of $\leq^i$-maximality and $\leq$-maximality.

References

[1] David W. Etherington, Alexander Borgida, Ronald J. Brachman, and Henry A. Kautz. Vivid knowledge and tractable reasoning: preliminary report. In *Proceedings of IJCAI-89, 10th International Joint Conference on Artificial Intelligence*, pages 1146–1152, Detroit, MI, 1989.

[2] Peter Gärdenfors and David Makinson. Revisions of knowledge systems using epistemic entrenchment. In *Proceedings of TARK-88, 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 83–95, Monterey, CA, 1988.

[3] Michael R. Garey and David S. Johnson. *Computers and intractability. A guide to the theory of NP-completeness*. Freeman, New York, NY, 1979.

[4] Matthew L. Ginsberg, editor. *Readings in nonmonotonic reasoning*. Morgan Kaufmann, Los Altos, CA, 1987.

[5] Philip N. Johnson-Laird. *Mental models*. Harvard University Press, Cambridge, MA, 1983.

[6] Kurt Konolige. On the relation between default theories and autoepistemic logic. In *Proceedings of IJCAI-87, 10th International Joint Conference on Artificial Intelligence*, pages 394–401, Milano, Italy, 1987. [a] An extended version appears as [7].

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

[8] Hector J. Levesque. Making believers out of computers. *Artificial Intelligence*, 30:81–108, 1986. [a] Also reprinted in [14], pp. 69–82.

[9] Hector J. Levesque. Logic and the complexity of reasoning. *Journal of Philosophical Logic*, 17:355–389, 1988.

[10] Vladimir Lifschitz, editor. *Programs with common sense - Papers by John McCarthy*. Ablex, Norwood, NJ, 1990.

[11] John McCarthy. Circumscription - a form of nonmonotonic reasoning. *Artificial Intelligence*, 13:27–39, 1980. [a] Also reprinted in [4], pp. 145–151. [b] Also reprinted in [10], pp. 142–155.

[12] John McCarthy. Applications of circumscription to formalizing commonsense knowledge. *Artificial Intelligence*, 28:89–116, 1986. [a] Also reprinted in [4], pp. 153–166. [b] Also reprinted in [10], pp. 198–225.

[13] Robert C. Moore. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25:75–94, 1985. [a] Also reprinted in [4], pp. 127–136.

[14] John Mylopoulos and Michael L. Brodie, editors. *Readings in artificial intelligence and databases*. Morgan Kaufmann, San Mateo, CA, 1989.

[15] Michael Reinfrank, Johan de Kleer, Matthew L. Ginsberg, and Erik Sandewall, editors. *Nonmonotonic reasoning*. Springer, Heidelberg, FRG, 1989.

[16] Fabrizio Sebastiani. A note on the complexity of simple heterogeneous model-preference default theories. Technical Report IEI-B4-37-1990, Istituto di Elaborazione dell'Informazione - CNR, Pisa, Italy, 1990.

[17] Fabrizio Sebastiani. A fully model-theoretic semantics for model-preference default systems. In *Proceedings of WOCFAI-91, First World Conference on the Foundations of Artificial Intelligence*, pages 439–445, Paris, France, 1991.

[18] Bart Selman and Henry A. Kautz. The complexity of model-preference default theories. In *Proceedings of CSCSI/SCEIO-88, 7th Biennial Conference of the Canadian Society for Computational Studies of Intelligence*, pages 102–109, Edmonton, Alberta, 1988. [a] Also reprinted in [15], pp. 115–130. [b] An extended version appears as [19].

[19] Bart Selman and Henry A. Kautz. Model-preference default theories. *Artificial Intelligence*, 45:287–322, 1990.