HELSINKI UNIVERSITY OF TECHNOLOGY
DIGITAL SYSTEMS LABORATORY
Series **A:** Research Reports | ISSN 0783–5396
No. **41**; December 1996 | ISBN 951-22-3369-X

LEXICOGRAPHIC ORDERING AS A BASIS OF PRIORITIES IN DEFAULT REASONING

Jussi Rintanen

Digital Systems Laboratory
Department of Computer Science
and Engineering
Helsinki University of Technology
Otaniemi, FINLAND

Helsinki University of Technology
Department of Computer Science
and Engineering
Digital Systems Laboratory
Otaniemi, Otakaari 1
FIN–02150 ESPOO, FINLAND

HELSINKI UNIVERSITY OF TECHNOLOGY
DIGITAL SYSTEMS LABORATORY
Series **A:** Research Reports
No. **41**; December 1996

ISSN 0783–5396
ISBN 951-22-3369-X

Lexicographic Ordering as a Basis of Priorities in Default Reasoning

JUSSI RINTANEN

Abstract: Nonmonotonic reasoning with priorities is investigated. First a version of default logic with priorities is defined. This logic extends Reiter's default logic and uses a priority mechanism that is based on lexicographic comparison. Questions concerning the existence of extensions, monotonicity of default reasoning with respect to priorities, alternative interpretations of partiality in priorities, and computational complexity of prioritized default logic are investigated. An algorithm for automated reasoning in prioritized default logic is presented and proved correct. Two translational inheritance theories are presented. The first is based on normal default theories and the second is based on prerequisite-free normal default theories. Results on the connection between these translational theories and translational theories presented in earlier work are given. Applicability of lexicographic ordering as a basis of priorities for default reasoning in general is discussed.
Keywords: nonmonotonic logics, default logic, preferences, priorities, lexicographic ordering, computational complexity, inheritance networks.

Dissertation for the degree of Doctor of Technology to be presented with due permission for public examination and debate in Auditorium S4 at Helsinki University of Technology (Espoo, Finland) on the 20th of December, 1996, at 12 o'clock noon.

Printing: TKK Monistamo; Otaniemi 1996

Helsinki University of Technology
Department of Computer Science
and Engineering
Digital Systems Laboratory
Otaniemi, Otakaari 1
FIN–02150 ESPOO, FINLAND

Phone: $\frac{09}{+358\ 9}$ 4511

Telex: 125 161 htkk fi
Telefax: +358 9 451 3369
E-mail: lab@saturn.hut.fi

Preface

This thesis is the result of my work from January 1993 to June 1995 at the Helsinki University of Technology, and from July 1995 to May 1996 at the University of Texas at Austin. During the work I have been given advice and otherwise been influenced and guided by several people for whom I would like to express my gratitude.

Professor Leo Ojala has taken good care of my graduate studies in Helsinki by being my supervisor and arranging funding that made this work financially possible. Professor Vladimir Lifschitz kindly hosted my visit to Austin which proved to be vital for finishing the work.

The thesis subject emerged from discussions with Ilkka Niemelä in fall and winter 1992-93. Without him my thesis would not have been on this subject, and discussions with him and the advice I got from him during the work have turned out to be useful without exception. My current and former colleagues at the Helsinki University of Technology have maintained a friendly working atmosphere that has made it possible to proceed in my studies and research. Especially I want to thank Johan Lilius for his enthusiasm that strengthened my interest in theoretical computer science and logic and encouraged me to pursue my undergraduate and graduate studies in that area.

My visit to Austin greatly helped in acquiring a wider perspective on the methods and problems of nonmonotonic reasoning. The visit was very pleasant and not the least because of the excellent computer science department and the logic in artificial intelligence group at UT Austin. Many thanks to Professor Lifschitz, Norman McCain and Hudson Turner.

Preliminary results from this thesis have been published in the proceedings of and presented in the European Workshop on Artificial Intelligence JELIA'94 in York, England in August 1994 and in the International Joint Conference on Artificial Intelligence in Montréal in August 1995. The funding for this research mainly came from Professor Ojala's project funds granted to him by the Academy of Finland. Additional financial support for the thesis work has been provided by Alfred Kordelinin yleinen edistys- ja sivistysrahasto, Heikki ja Hilma Honkasen säätiö, Emil Aaltosen säätiö, and Suomen Kulttuurirahasto. My visit to Austin was funded by grants from the Academy of Finland and Teknillisen korkeakoulun tukisäätiö.

Espoo, Finland, December 1996

Contents

List of Figures

List of symbols and notations

$\times$	cartesian product of sets
$\cup$	union of sets
$\cap$	intersection of sets
$\setminus$	difference of sets
$\vartriangle$	symmetric difference of sets, $S_1 \vartriangle S_2 = (S_1 \setminus S_2) \cup (S_2 \setminus S_1)$
$\in$	set membership
$\subseteq$	set inclusion
$\subset$	proper set inclusion
2^S	the power set of S, i.e. the set $\{S' \mid S' \subseteq S\}$
S^n	the n-fold cartesian product $S \times S \times \cdots \times S$
$\circ$	relation composition, $R \circ R' = \{\langle a, c\rangle \mid \langle a, b\rangle \in R, \langle b, c\rangle \in R'\}$
$\sum_{i=a}^{b} E(i)$	the sum $E(a) + E(a+1) + \cdots + E(b-1) + E(b)$
$\bigcup_{S \in U}$	the union $S_1 \cup S_2 \cup \cdots$ of the sets in $U = \{S_1, S_2, \ldots\}$
$\mathcal{O}(P)$	an upper bound of complexity
$\leftrightarrow$	equivalence
$\rightarrow$	implication
$\wedge$	conjunction
$\vee$	disjunction
$\neg$	negation
$\perp$	constant *false*
$\top$	constant *true*
$\models$	logical consequence, truth in a model
$\overline{l}$	the complement of the literal l: $\overline{p} = \neg p$ and $\overline{\neg p} = p$.

Chapter 1

Introduction

This work is an investigation concerning the incorporation of priorities to systems for knowledge representation that support the expression of default rules. Default rules are expressions that allow drawing a conclusion whenever the conclusion does not contradict other information that is available. An example default rule would say that it does not rain in Austin, Texas. Because one's knowledge is incomplete, that is, it may rain in Austin even though there is no knowledge about that, the conclusion obtained by using the default rule may have to be retracted if more knowledge, about a thunderstorm for example, becomes available. This is a characteristic property of reasoning with default rules: increase in the premises does not necessarily increase the number of conclusions. Because of this property, logics that formalize default rules are called nonmonotonic logics as distinguished from the majority of logics that have the monotonicity property.

Research on default rules is motivated by the possibility to use computers for representing knowledge and reason about it. It would be a great benefit if database systems could be extended to handle default rules. Default rules provide a means for expressing information about typical states of affairs concisely. More generally, the project of building intelligent computers and robots, artificial intelligence, seems to require default rules because human beings, the creatures these intelligent systems are meant to imitate, can be seen to use default rules all the time. In fact, it seems that there are very few rules used by people every day that universally hold. Notable exceptions are rules about linguistic conventions (all penguins are birds, all bachelors are unmarried.)

Nonmonotonic logics that are used in investigations on default reasoning are based on monotonic logics. The most important (monotonic) logics are the classical propositional logic and the predicate logic. These logics formalize

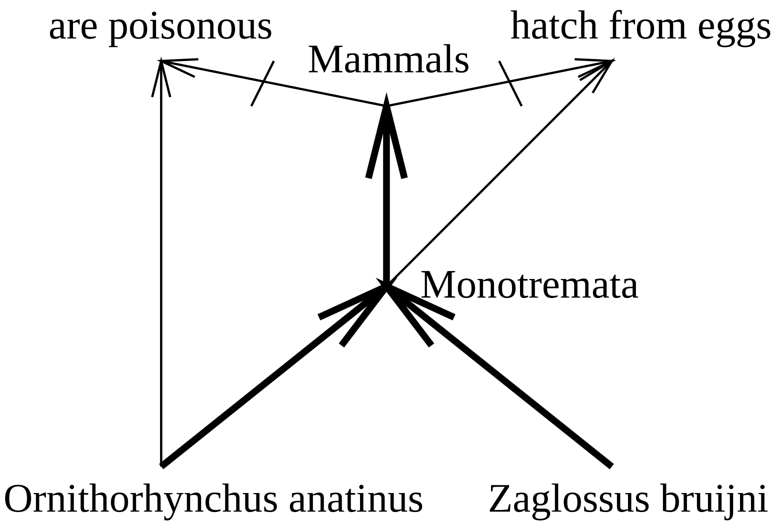


Figure 1.1: A part of a biological knowledge base

deductive reasoning, that is, inferences that guarantee the truth of conclusions given that the premises are true. Nonmonotonic logics do not have this property, and therefore reasoning in nonmonotonic logics is not in general deductive. The combination of a monotonic logic together with default rules is useful because deductive inferences are often needed too.

Consider the biological knowledge base in Figure 1.1. The knowledge base contains information concerning two species of animals in the order Monotremata [Encyclopædia Britannica, 1995]. Thick arrows express strict laws that do not allow exceptions. For example, the arrow from Monotremata to Mammals means that Monotremata are Mammals. The lack of exceptions is because by definition all species of animals in the order Monotremata belong to the class Mammals. In contrast, thin arrows express default rules, that is, laws that may have exceptions. For example, the arrow from Mammals to "hatch from eggs" expresses the fact that most Mammals do not hatch from eggs.

The same knowledge base can be written as sentences of a database language similar to Datalog¬ [Kanellakis, 1990] as follows.

mammal(x) ← monotremata(X)
monotremata(x) ← zaglossus-bruijni(x)
monotremata(x) ← ornithorhynchus-anatinus(x)
¬poisonous(x) ← mammal(x), not poisonous(x)
¬hatch-from-eggs(x) ← mammal(x), not hatch-from-eggs(x)
poisonous(x) ← ornithorhynchus-anatinus(x), not ¬poisonous(x)
hatch-from-eggs(x) ← monotremata(x), not ¬hatch-from-eggs(x)

The ¬ operator corresponds to the "not" in English, and the *not* operator roughly corresponds to "is not known that." Hence if the knowledge base does not specify a fact P, then *not* P can be inferred. The fourth rule in the knowledge base can be read as "mammals, if not known to be poisonous, are not poisonous." Rules of this kind are default rules in the sense discussed earlier.

The following are some of the conclusions we would like to obtain from the knowledge base.

1. If x belongs to the species Ornithorhynchus anatinus, then x is a Mammal.
2. If x belongs to the species Ornithorhynchus anatinus or x belongs to the species Zaglossus bruijni, then x is a member of the order Monotremata.
3. If x belongs to Zaglossus bruijni, then x hatched from an egg.
4. If x belongs to Ornithorhynchus anatinus, then x is poisonous.

The first conclusion is sanctioned by deductive inferences, because the arrows involved are thick and therefore express laws with no exceptions. This is also the case in the second conclusion that involves the word "or." If x is either a Ornithorhynchus anatinus or a Zaglossus bruijni, and all members of these species are members of Monotremata, then x clearly is a member of Monotremata. The third is based on the thick arrow from Zaglossus bruijni to Monotremata and the thin arrow from Monotremata to "hatch from eggs." The inferences made are not deductive because x being a Zaglossus bruijni and hence Monotremata and the fact that Monotremata usually hatch from eggs do not give a guarantee to the fact that x hatched from an egg. In fact, in this case like in the last one there are sequences of arrows that seem to justify a contradicting conclusion.

The possibility of alternative inferences that are mutually conflicting is a characteristic property of default reasoning. If no such conflicts existed, deductive reasoning would suffice for the representation of knowledge bases like the above.

1.1 Priorities

When there are two default rules such that applying them both would lead to a contradiction, the question of which default rule to apply arises. There is no general answer to this question. The strategy adopted in nonmonotonic logics is that reasoning is performed with both defaults in turn and a sentence is a conclusion if it is obtained in both cases. In the example presented above, this strategy would not give conclusions concerning the egg laying ability or possession of poison of Ornithorhynchus anatinus.

However, in many cases one of the default rules in conflict is more important than others or pertains more closely to the subject at hand than others, and therefore other default rules should be rejected. In the example presented above, the rule about hatching from eggs would override the rule about not hatching from eggs as far as members of Monotremata are concerned, because the former rule is specifically about Monotremata and the latter is not. To express the information that certain default rules are more specific or more important than others, we need *priorities* on the default rules. Each default rule is associated with a priority, and the priorities of the default rules can be compared. If there is a conflict and the priority of one of the default rules is higher than the priorities of the others, then this default rule should be chosen and others should be rejected. Priorities increase the number of conclusions because they decrease the number of conflicts that are left unresolved.

This work concentrates on priorities in nonmonotonic logics. The two main motivations for investigating nonmonotonic logics are default reasoning as discussed in the example above, and the frame problem in temporal reasoning [McCarthy and Hayes, 1969]. Default reasoning is important because of its applications in data and knowledge bases and in artificial intelligence. The possibility of conflicts between default rules is inherent in default reasoning, and hence priorities or other ways of solving the specificity problem are of great importance. To date, no general solutions to this problem have been presented. Research on inheritance reasoning, which is a restricted special case of default reasoning, has focused on giving direct characterizations of correct inheritance reasoning [Horty, 1994] without the help of nonmonotonic logics. Unlike in inheritance reasoning, in default reasoning no restriction to atomic formulae is made, and for arbitrary formulae the path-based approach used in inheritance reasoning is no longer feasible because graph-theoretic notions like edge and path are not defined. It is believed that default reasoning in general is better investigated in the framework of nonmonotonic logics like default logic [Reiter, 1980], autoepistemic logic [Moore, 1985], or circumscription [McCarthy, 1980], or, more recently, in argument-based frameworks of nonmonotonic reasoning [Simari and Loui, 1992; Pollock, 1994].

1.2 Contributions of the work

In this work we investigate the use of priorities in nonmonotonic logics, the computational properties of of nonmonotonic reasoning with priorities, and the applicability of nonmonotonic logics with priorities to representation of knowledge in settings like the biological knowledge base discussed earlier. The contributions of this work to the research area are the following.

- A new proposal for the use of priorities in Reiter's default logic [1980].

  Our proposal is based on lexicographic comparison. Earlier proposals by Baader and Hollunder [1993], Marek and Truszczyński [1993], Brewka [1991; 1994], MacNish [1991] are based on the idea of using priorities as the order in which the application of defaults is attempted. In these approaches priorities essentially act as control information for a decision procedure for default logic. Interestingly, both the prioritized default logic proposed by us and the prioritized default logics proposed in [Baader and Hollunder, 1993; Marek and Truszczyński, 1993; Brewka, 1994; MacNish, 1991] can be seen as generalizations of Brewka's preferred subtheories framework [1989]. The preferred subtheories framework is equivalent to the prerequisite-free normal subclass of prioritized default theories.

- Results on interconnections between different formalizations of priorities in default reasoning.

  We give embeddings of the preferred subtheories framework [Brewka, 1989], ordered theory presentations [Ryan, 1992a], prioritized circumscription [McCarthy, 1986; Lifschitz, 1985], and minimal model semantics used in conditional entailment [Geffner and Pearl, 1992] in our framework of prioritized default logic. The similarity between prioritized prerequisite-free default logic and preferred subtheories and prioritized circumscription seems to be a part of the folklore of the research area of nonmonotonic reasoning, but there are no published accounts of these connections, nor can precise statements about these connections be found in the literature. The similarity between ordered theory presentations and preferred subtheories is hinted at in [Ryan, 1992b].

  We are not aware of other studies that compare different kinds of lexicographic and other priority mechanisms used in nonmonotonic reasoning. The fact that there are two natural ways of reducing lexicographic comparison with partially ordered positions to ordinary lexicographic comparison (with totally ordered positions) has not been recognized in earlier research in the research area. The fact that Geffner and Pearl's [1992] and Brewka's [1989] preference notions are not the same has been noticed by Junker and Brewka [1991].

- Proposal of the notion of a classification scheme for defaults rules. This notion is a basis for recognizing the existence of a range of different variants of prioritized default logic.

  Baader and Hollunder [1993], Marek and Truszczyński [1993], and Brewka [1994] classify default rules to those for which the prerequisite is in an extension and the negations of the justifications are not. This corresponds to our classification to applied and non-applied defaults. We also use a classification to defeated and non-defeated defaults, where a defeated default is one for which the prerequisite is in an extension and at least one of the negations of the justifications is in the extension. There are many other classification schemes besides these two.

- An analysis of the computational complexity of nonmonotonic reasoning with lexicographic priorities: identification of the boundary between tractable and intractable inference, location of the decision problems of prioritized default logic in the polynomial hierarchy.

  Our main complexity results are the membership of the decision problems of prioritized default logic on the third level of the polynomial hierarchy, and a hardness result that suggests the inexistence of polynomial time translations of the decision problems of prioritized default logic to the decision problems of default logic without priorities. Furthermore, we analyze the complexity of prioritized default logic in syntactically restricted cases. As default logic is a special case of prioritized default logic (prioritized default theories with the empty ordering on defaults), the latter is intractable with all restrictions where default logic is. Hence we have analyzed the complexity of prioritized default logic with priorities that seem most promising in achieving tractability, this is the case when priorities totally order the defaults. We show that for most of the syntactically restricted classes there is a translation from the decision problem of brave reasoning [Kautz and Selman, 1991; Stillman, 1990], and hence in these cases prioritized default reasoning is intractable whenever brave reasoning is intractable. We give polynomial time algorithms for prerequisite-free normal default theories with total priorities and where the objective parts of default theories are formulae in any subclass of propositional logic with a polynomial time satisfiability test and the conclusions of the defaults are formulae in the same class. Also, we give a polynomial time decision procedure for the class of Horn default theories [Kautz and Selman, 1991] with total priorities.

  There are no prior analyses of the complexity of nonmonotonic logics with priorities. Earlier complexity results [Gottlob, 1992; Stillman, 1992; Kautz and Selman, 1991; Stillman, 1990] are on nonmonotonic logics without priorities. Bylander et al. [1991] and Eiter and Gottlob [1995] and Eiter, Gottlob and Leone [1995] analyze the complexity of abduction, including cases where the hypotheses are ordered. Abduction can be seen as a form of nonmonotonic reasoning.

- A procedure for automated reasoning in prioritized default logic.

Analyses of computational complexity of nonmonotonic logics [Kautz and Selman, 1991; Stillman, 1990] have identified the conflicts between defaults as a major source of complexity. Priorities often uniquely determine how conflicts are resolved, and the case analysis that corresponds to the resolution of a conflict in alternative ways is avoided. Therefore the idea of using priorities for reducing the search space in procedures for prioritized nonmonotonic reasoning arises. We present an algorithm for automated reasoning in prioritized default logic. This algorithm is based on binary search in the space of all extensions of a default theory. The priorities prune the search tree so that subtrees that contain only extensions that are not preferred are not traversed. The algorithm runs in polynomial space whenever the number of preferred extensions is polynomial.

Baader and Hollunder [1993] present an algorithm for automated reasoning in their version of prioritized default logic. Brewka and Junker [1991] present an algorithm for automated reasoning in preferred subtheories with arbitrary strict partial orders. Preferred subtheories correspond to the subclass of prerequisite-free normal default theories of prioritized default logic. Their algorithm is based on the use of truth-maintenance systems [Doyle, 1979] and they state that the algorithm can be used for finding only one preferred subtheory in a straightforward way. Baker and Ginsberg [1989] present an algorithm for reasoning in prioritized circumscription. Prioritized circumscription corresponds to the prerequisite-free subclass of prioritized default logic. Baker and Ginsberg's algorithm is defined for layered priorities only. Algorithms for unprioritized nonmonotonic logics have been presented for example by [Marek *et al.*, 1993; Niemelä, 1994]. Like our algorithm, these algorithms are based on binary search over the set of all candidate extensions.

- Results establishing the equivalence of different translational theories of inheritance.

  We show that an inheritance theory presented by Gelfond and Przymusinska [1990] coincides with a translational theory presented in this work. Our translational theory can be formalized with two different classification schemes for defaults proposed in this work. The translational theory presented in this work can be represented as prerequisite-free normal defaults with implicational conclusions. The atomic consequences with this translation are the same. An inheritance theory by Brewka [1987] coincides with the translational theory based on prerequisite-free defaults presented in this work for the subclass of inheritance networks for which Brewka's theory is defined.

Chapter 2

Preliminaries

The basic definitions related to default logic [Reiter, 1980] and complexity classes related to deterministic and nondeterministic polynomial time reductions are given in this chapter.

Default logic is used as the basic framework of this work because it subsumes many of the alternative, less general formalizations of nonmonotonic reasoning like Poole's Theorist [Poole, 1985] and Brewka's preferred subtheories [Brewka, 1989], and because it provides a syntax that directly allows the expression of default rules.

Concerning the difficulty of various computational problems, the most significant boundary is the one between the problems that are solvable in polynomial time, and the problems that are not. Some of the results of this work demonstrate that in restricted cases nonmonotonic reasoning can be performed in polynomial time. Claims of that kind are proved by giving an algorithm for the problem in question. Other results indicate cases in which nonmonotonic reasoning very likely, that is assuming P$\neq$NP, cannot be performed in polynomial time. These results and others presenting a more fine-grained analysis of the difficulty of the problems use the complexity classes P, NP, Δ_2^p, Σ_2^p, Π_2^p and so on, that are defined in terms of polynomial time translations between computational problems.

2.1 Default logic

Default logic [Reiter, 1980] was one of the first formalizations of nonmonotonic reasoning and has been the basis of a lot of research in the research area. The formalization of default logic has a property that drastically separates it from any monotonic logic: some of the inferences are based on the prior *underivability* of certain formulae. In proof theories of monotonic logics, the condition for applying an inference rule α/γ is simply that the antecedent α has been derived. In default logic, a *default rule* $\alpha{:}\beta_1, \ldots, \beta_n/\gamma$ can be applied, if (informally) α has been derived, and none of the formulae $\neg\beta_1, \ldots, \neg\beta_n$ can be derived[1]. A typical default rule is $a{:}b/b$ that allows to conclude b whenever a has been derived and $\neg b$ cannot be derived. Given the fact a, the application of this default rule would be alternative to the application of the default rule $a{:}\neg b/\neg b$: at most one of them can be applied according to the informal conditions stated above. As derivability and underivability are mutually dependent, there is circularity in the definition of what is derivable in default logic. Unlike in monotonic logics, where the consequences of a set of formulae is defined as the set of formulae that can be derived by using the axioms of the logic and the inference rules, the conclusions of a *default theory* $\Delta = \langle D, W\rangle$, where D is a set of default rules and W is a set of formulae, are defined as fixpoints of a nonmonotonic operator. The operator may have several fixpoints, none of which is the least fixpoint, and different fixpoints can be seen as a result of applying different mutually incompatible sets of default rules. The fixpoints of the operator for a default theory are the *extensions* of the default theory. The notion of extension is sometimes given an explanation in terms of a story about an agent the beliefs of which are based on a set of facts, expressed by the formulae in W that is the *objective part* of a default theory, and a set of defaults D that express the reasoning of the agent on the basis of its beliefs. Now extensions can be thought of as the states of belief the agent may be in.

Any given monotonic logic can be extended to a default logic, because the definition of default logic relies only on basic properties of consequence relations and formulae that are satisfied by all commonly used monotonic logics. In this work we concentrate on propositional default logic that is based on the classical propositional logic. The language of classical propositional logic is denoted by the symbol $\mathcal{L}$. We assume that there is a countably infinite number of propositional variables in $\mathcal{L}$. The set of all default rules $\{\alpha{:}\beta_1, \ldots, \beta_n/\gamma | n \geq 0, \{\alpha, \beta_1, \ldots, \beta_n, \gamma\} \subseteq \mathcal{L}\}$ is denoted by the symbol $\mathcal{D}$. The closure of a set $S \subseteq \mathcal{L}$ under logical consequence is defined as $Cn(S) = \{\phi \in \mathcal{L} | S \models \phi\}$.

Definition 2.1.1 ([Reiter, 1980]) *Let $\Delta = \langle D, W\rangle$ be a default theory. For any set of formulae $S \subseteq \mathcal{L}$, let $\Gamma(S)$ be the smallest set satisfying the following three properties.*

1. $W \subseteq \Gamma(S)$

[1] Corollary 2.1 in [Reiter, 1980] suggests that Reiter considers only defaults $\alpha{:}\beta_1, \ldots, \beta_n/\gamma$ with $n \geq 1$. We allow all $n \geq 0$. The proofs of those of Reiter's theorems that we use can easily be extended to this more general case.

2. *$Cn(\Gamma(S)) = \Gamma(S)$*

3. *If $\alpha{:}\beta_1, \ldots, \beta_n/\gamma \in D$ and $\alpha \in \Gamma(S)$ and $\{\neg\beta_1, \ldots, \neg\beta_n\} \cap S = \emptyset$, then $\gamma \in \Gamma(S)$.*

A set of formulae $E \subseteq \mathcal{L}$ is an extension *for Δ if and only if $\Gamma(E) = E$.*

The above definition is not convenient to use in many proofs, and other definitions of extensions that are more directly applicable in proofs based on mathematical induction are often used instead. The next theorem presents the *semiconstructive* definition of extensions, so called because it suggests a nondeterministic procedure for computing extensions of default theories.

Theorem 2.1.2 ([Reiter, 1980]) *Let $E \subseteq \mathcal{L}$ be a set of formulae, and let $\Delta = \langle D, W\rangle$ be a default theory. Define for all $i \geq 0$*

$$\begin{aligned} E_0 &= W \\ E_{i+1} &= Cn(E_i) \cup \left\{ \gamma \,\middle|\, \frac{\alpha : \beta_1, \ldots, \beta_n}{\gamma} \in D, \alpha \in E_i, \{\neg\beta_1, \ldots, \neg\beta_n\} \cap E = \emptyset \right\}. \end{aligned}$$

Then E is an extension for Δ if and only if $E = \bigcup_{i \geq 0} E_i$.

A closely related definition of extensions is presented by Gelfond, Lifschitz, Przymusinska, and Truszczyński[2].

Theorem 2.1.3 ([Gelfond *et al.*, 1991]) *Let $\langle D, W\rangle$ be a default theory. Let E be a set of formulae. Define*

$$D^E = \left\{ \frac{\alpha : \beta_1, \ldots, \beta_n}{\gamma} \in D \,\middle|\, \{\neg\beta_1, \ldots, \neg\beta_n\} \cap E = \emptyset \right\}.$$

Define $Cn_W(D)$ as $\bigcup_{i \geq 0} Cn(C_i)$ where for all $i \geq 0$,

$$\begin{aligned} C_0 &= W \\ C_{i+1} &= C_i \cup \left\{ \gamma \,\middle|\, \frac{\alpha : \beta_1, \ldots, \beta_n}{\gamma} \in D, C_i \models \alpha \right\}. \end{aligned}$$

Then E is an extension of $\langle D, W\rangle$ if and only if $E = Cn_W(D^E)$.

The notion of *generating defaults* of an extension is useful in many proofs because by using it the extensions of a default theory can be represented as a finite set of formulae whenever the set of defaults and the set of objective formulae are finite.

[2] Actually, unlike Gelfond et al., the definition we give here does not eliminate the justifications of defaults when constructing D^E and makes explicit what it means to be "closed under inference rules."

Definition 2.1.4 ([Reiter, 1980]) *Suppose $\Delta = \langle D, W\rangle$ is a default theory and E is an extension for Δ. The set of* generating defaults *for E with respect to Δ is defined as*

$$GD(E,\Delta) = \left\{ \frac{\alpha : \beta_1, \ldots, \beta_n}{\gamma} \in D \middle| \alpha \in E \textit{ and } \{\neg\beta_1, \ldots, \neg\beta_n\} \cap E = \emptyset \right\}.$$

Theorem 2.1.5 ([Reiter, 1980]) *Suppose E is an extension of a default theory $\Delta = \langle D, W\rangle$. Then $E = Cn(W \cup \{\gamma | \alpha{:}\beta_1, \ldots, \beta_n/\gamma \in GD(E,\Delta)\})$.*

The standard consequence relation of default logic is *cautious reasoning* $\models_c$.

Definition 2.1.6 *Let $\Delta = \langle D, W\rangle$ be a default theory and $\phi \in \mathcal{L}$ a formula. Then $\Delta \models_c \phi$ if and only if $\phi \in E$ for all extensions E of Δ.*

The following terminology is used in referring to default rules of certain syntactic forms. Defaults of the form $\alpha{:}\beta/\beta$, $\alpha{:}\beta \wedge \gamma/\gamma$, $\top{:}\beta_1, \ldots, \beta_n/\gamma$ are respectively *normal*, *seminormal*, and *prerequisite-free*. Prerequisite-free defaults are often written without prerequisites as $:\beta_1, \ldots, \beta_n/\gamma$. We shall often denote justifications $\beta_1, \ldots, \beta_n$ of defaults by letters $\sigma, \sigma', \sigma_1$ and so on.

A property that makes normal default theories, that is, default theories in which all defaults are normal, well-behaved is that they have at least one extension.

Theorem 2.1.7 ([Reiter, 1980]) *Every normal default theory has an extension.*

In the special case of prerequisite-free default theories the definition of extensions reduces to a particularly simple form because the applicability of a default does not depend on the prior application of other defaults, and hence it is unnecessary to view the construction of an extension as consisting of multiple steps.

Theorem 2.1.8 *Let $\langle D, W\rangle$ be a default theory where all defaults in D are prerequisite-free. Then for all $E \subseteq \mathcal{L}$, E is an extension of $\langle D, W\rangle$ if and only if*

$$E = Cn(W \cup \left\{ \gamma \middle| \frac{: \beta_1, \ldots, \beta_n}{\gamma} \in D, \{\neg\beta_1, \ldots, \neg\beta_n\} \cap E = \emptyset \right\}).$$

Proof: The proof is by using the definition of extensions given in Theorem 2.1.3. Let $E \subseteq \mathcal{L}$ be a set of formulae and $\Delta = \langle D, W\rangle$ a default theory. By definition

$$\begin{aligned}
D^E &= \left\{ \frac{:\beta_1,\ldots,\beta_n}{\gamma} \in D \,\middle|\, \{\neg\beta_1,\ldots,\neg\beta_n\} \cap E = \emptyset \right\}, \\
C_0 &= W, \\
C_1 &= C_0 \cup \left\{ \gamma \,\middle|\, \frac{\top:\beta_1,\ldots,\beta_n}{\gamma} \in D^E, C_0 \models \top \right\}, \\
&= W \cup \left\{ \gamma \,\middle|\, \frac{\top:\beta_1,\ldots,\beta_n}{\gamma} \in D, \{\neg\beta_1,\ldots,\neg\beta_n\} \cap E = \emptyset \right\}, \text{ and} \\
C_i &= C_1 \text{ for all } i \geq 2.
\end{aligned}$$

By Theorem 2.1.3, E is an extension of Δ if $E = \bigcup_{i\geq 0} Cn(C_i)$, that is, if $E = Cn(C_1) = Cn(W \cup \{\gamma | {:}\beta_1,\ldots,\beta_n/\gamma \in D, \{\neg\beta_1,\ldots,\neg\beta_n\} \cap E = \emptyset\})$. □

A number of lemmata on default logic that are used in proofs in the body of this work are included in the appendix.

2.2 Computational complexity

There are two kinds of complexity results in this work. For some computational problems it is demonstrated that they can be solved in polynomial time. These results are based on giving an algorithm for the problem in question, showing that the algorithm is correct, and deriving an upper-bound on the running time of the algorithm for all input values. For other problems it is demonstrated that they cannot be solved in polynomial time. These results are given in the framework of the theory of NP-completeness or more generally structural complexity [Balcázar *et al.*, 1995]. The proofs are typically translations from other computationally difficult problems to the problems at hand, and they rely on some widely accepted conjectures on the difficulty of certain problems, like that the satisfiability of formulae of classical propositional logic cannot be tested in polynomial time (that is, the complexity classes NP and P do not coincide.)

The complexity of the computationally difficult problems discussed in this work is mainly analyzed with respect to the complexity classes NP, co-NP and the polynomial hierarchy PH. Here we briefly introduce some basic concepts. For a full account see [Balcázar *et al.*, 1995].

P is the class of decision problems that are solvable in polynomial time by a deterministic Turing machine. NP is the class of decision problems that are solvable in polynomial time by a nondeterministic Turing machine. The class co-NP consists of those problems the complements of which are in NP. In general, the class co-C consists of problems whose complements are in the class C. The polynomial hierarchy PH is an infinite hierarchy of complexity classes Σ_i^p, Π_i^p, and Δ_i^p for all $i \geq 0$ that is defined by using oracle Turing machines in the following way.

$$\begin{array}{rclrclrcl} \Sigma_0^p & = & P & \Pi_0^p & = & P & \Delta_0^p & = & P \\ \Sigma_{i+1}^p & = & \mathrm{NP}^{\Sigma_i^p} & \Pi_{i+1}^p & = & \text{co-}\Sigma_{i+1}^p & \Delta_{i+1}^p & = & P^{\Sigma_i^p} \end{array}$$

Note that it is not known whether classes in PH differ from each other. $C_1^{C_2}$ denotes the class of problems that is defined like the class C_1 except that oracle Turing machines that use an oracle for a problem in C_2 are used instead of Turing machines without oracles. Oracle Turing machines with an oracle for a problem B are like Turing machines except that they may perform tests for membership in B with constant cost. A problem L is *Turing reducible* to a problem L', if there is an oracle Turing machine that solves L and has an oracle for L'. The problem is Turing reducible in polynomial time, if the oracle Turing machines solves L with a polynomial number of execution steps.

A problem L is *C-hard* (where C may be NP, co-NP or any of the classes in the polynomial hierarchy), if all problems in the class C are polynomial time *many-one reducible* to it, that is, for all problems $L' \in C$ there is a function $f_{L'}$ that can be computed in polynomial time on the size of its input and $f_{L'}(x) \in L$ if and only if $x \in L'$. We say that the function $f_{L'}$ is a translation from L' to L. A problem is *C-complete*, if it belongs to the class C and is C-hard. The most famous NP-complete problem is SAT, the satisfiability of Boolean formulae, that is, formulae of the classical propositional logic [Cook, 1971]. An example of a problem that is co-NP-hard is the dual problem of SAT, testing whether a formula is valid (a tautology.)

The boundary between feasible and infeasible computational problems is often considered to reside between problems solvable in deterministic polynomial time (the class P) and those that are not solvable in deterministic polynomial time. Problems solvable in polynomial time are *tractable*. Problems that are not in P are *intractable*. The term intractable usually refers also to problems that are widely believed not to be in P (for example the NP-complete problems), even if there is no certainty of the inexistence of polynomial time algorithms for them.

The membership of a problem in P can be demonstrated by presenting a polynomial time algorithm for solving the problem. An upper-bound on the time complexity of an algorithm is indicated by using the notation $\mathcal{O}(f)$ where f is a function on x. This notation means that there are constants c and a so that for all input values of size x for the algorithm, the algorithm terminates after at most $c + af(x)$ execution steps.

Chapter 3

Lexicographic priorities for default logic

In this chapter we introduce our framework of default reasoning with priorities. Priorities express the plausibility of pieces of knowledge. As discussed in Section 2.1, for a given knowledge base expressed in default logic there may be several extensions each of which is a possible meaning of the knowledge base. However, an extension can be unacceptable because in it a conflict between defaults is resolved in an unintended way. The simplest example of this is the penguin triangle.

Example 3.0.1 Consider the inheritance network in Figure 3.1[1]. The letters p, q and r respectively express the statements that "x is a penguin", "x is a bird", and "x is able to fly". The arrows can be understood as stating that

[1] As in this example, it is often natural to depict default theories pictorially as directed graphs with two kinds of arrows. Ordinary arrows going from p to q represent defaults $p{:}q/q$, and broken arrows represent defaults $p{:}\neg q/\neg q$.

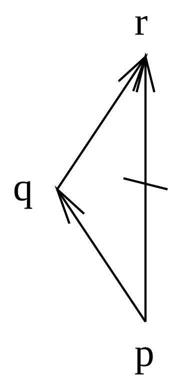


Figure 3.1: The penguin triangle

penguins are usually birds, birds are usually able to fly, and penguins are not usually able to fly. In default logic, these defaults are formalized as $D = \{p{:}q/q, p{:}\neg r/\neg r, q{:}r/r\}$. It would be correct to conclude that a given penguin is not able to fly. For the default theory $\langle D, \{p\}\rangle$, however, default logic gives the extensions $Cn(\{p, q, r\})$ and $Cn(\{p, q, \neg r\})$ that respectively correspond to the application of the defaults $q{:}r/r$ and $p{:}\neg r/\neg r$. The conclusion that penguins do not fly is not supported by the extension $Cn(\{p, q, r\})$. Hence for this example default logic gives too many extensions and therefore the logic is too weak. ■

There are many different ways to strengthen the conclusions sanctioned by default logic. One possibility is to modify the default theories so that conflicts between defaults are resolved in the intended way. This approach has been proposed by Reiter and Criscuolo [1981]. For the above example Reiter and Criscuolo suggest that the default $q{:}r/r$ is replaced by $q{:}\neg p \wedge r/r$. Now the default about the flying ability of birds cannot be used if the bird is known to be a penguin. This approach will be discussed in more detail in Section 7.3.5. The justification for modifying some of the defaults is based on the idea that some extensions are more important or more plausible than others. This suggests that each extension should be assigned a measure of plausibility, the plausibility of different extensions can be compared, and the intended extensions are the most plausible ones. Reiter and Criscuolo could be seen as proposing a way of selecting the most plausible extensions by having the default theories syntactically modified. In this work, we concentrate on a different approach in which default theories are left intact and the selection of most plausible extensions is performed by means that are outside the definition of extensions of default logic.

What is a priority?

The plausibility of an extension of a default theory is determined by the formulae in it. If the dependence between the formulae and the plausibility is arbitrary, the formalization of plausibility would simply be a function f_Δ from the set X of extensions of a default theory Δ to a set P that represent different degrees of plausibility. We can abstract this proposal by choosing P to be $\{0, 1\}$, where plausible extensions are mapped to 1 and other extensions to 0. This proposal is not very useful because it completely leaves open the origin of the plausibility, and to use it one has to know what the extensions of the default theory are. Alternatively, for each default theory a set of formulae relevant to the plausibility of its extensions is identified. This set is some (not necessarily proper) subset S of $\mathcal{L}$. Now the plausibility of each extension E can be evaluated in terms of $E \cap S$. Because not all formulae in S are equally important – some are more important than others – a priority value is associated with each member of S, and extensions E and E' could be compared according to the priorities the formulae in $E \cap S$ and $E' \cap S$ have. The priorities are members of some set, either an arbitrary set for which the properties of the members are not important or for example the set of integers or real numbers. In the latter case, the mathematical operations on the numbers could be used in determining the plausibility of extensions. The question of what subsets S of $\mathcal{L}$ are used in determining the plausibility of an extension is postponed later. For the time being we shall assume that for each default theory there is a designated set $S \subseteq \mathcal{L}$ each member of which is assigned a priority and the plausibility of an extension E is a function of the priorities of the members of $E \cap S$.

Nonmonotonic logics have been proposed as solutions to problems in application areas that differ considerably. Hence it also seems that the information that determines the intuitively acceptable inferences varies between the application areas, and therefore there most likely is no one single framework of expressing priorities and choosing the most plausible extensions of default theories that would work in all applications. Consider model-based diagnosis [Reiter, 1987] where a diagnosis is a set of components the assumed failure of which explains the observations that contradict the correct behavior of the system to be diagnozed. In this application there may be information on the probability in which each component fails. With the assumption that the failure of each component is probabilistically independent of the failure of others, the probability of each diagnosis can be computed by multiplying the probabilities of the component failures it assumes. Hence the problem of model-based diagnosis suggests a certain kind of priority mechanism for default logic and in fact it seems that other kinds of priority mechanisms would be either of questionable or of restricted utility.

Another problem area that has received a lot of attention by the logic in artificial intelligence community is inheritance reasoning. It seems that priority mechanisms that are based on numeric priorities are not suitable for inheritance reasoning. This is because a conflict between defaults should always be resolved in favour of the most specific default, no matter how many less specific defaults are violated. This idea, that the default with the highest priority wins, leads naturally to lexicographic comparison. If the highest priority default is applied in some extensions, then no extension that does not apply it is a preferred extension. For extensions that are not discriminated by the highest priority default the same analysis is repeated for the default with the second highest priority and so on. Lexicographic comparison has been used as a basis for priorities in nonmonotonic reasoning for example by McCarthy and Lifschitz [1986; 1985], Brewka [1989], Geffner and Pearl [1992], and Ryan [1992a]. A weaker priority mechanism is used by Pearl [1990]. He orders two propositional models according to the highest priority formula that is not satisfied by both of them. This ordering policy is weaker than lexicographic ordering, and leads to problems in many inheritance networks because as soon as a default is violated, it is of no significance whether defaults with a lower priority are violated or not.

In our investigation we concentrate on priorities based on lexicographic ordering and regard inheritance reasoning as a major application area of this kind of priorities.

Which formulae are priorities assigned to?

The question of which subsets $S \subseteq \mathcal{L}$ are reasonable bases for priorities was left open in the earlier discussion. As shown in Theorem 2.1.5 the intersection of an extension with the objective part and the conclusions of the defaults of a default theory uniquely identify the extension. From this point of view an expressive enough subset S of $\mathcal{L}$ would be the set of conclusions of the defaults of a default theory, as for any set of extensions with a designated extension there is an ordering on S so that the designated extension is lexicographically better than the rest. This essentially is the choice for S for example in the preferred subtheories approach of Brewka [1989] and in the minimal model semantics of conditional entailment [1992] because of the correspondence between prerequisite-free normal defaults and these approaches (see Sections 6.2 and 6.5.)

There is the possibility that different defaults have the same conclusion – this is indeed common in default theories that represent inheritance networks – but that the conclusion should be given a different priority depending on which default was used in obtaining it in a given extension. Hence a slightly more expressive choice for S is the set of conjunctions of prerequisites and conclusions of the defaults in a default theory. This way priorities are assigned to defaults and not to conclusions of defaults. This difference is of importance for example in the inheritance network represented by the default theory $\langle\{a_1{:}b/b, a_2{:}\neg b/\neg b, a_3{:}b/b\}, \{a_1 \rightarrow a_2, a_2 \rightarrow a_3\}\rangle$. If the priorities are on b and $\neg b$, we cannot have both the set $Cn(\{a_1, a_2, a_3, b\})$ as a preferred extension if the objective part of the default theory is extended with a_1, and $Cn(\{a_2, a_3, \neg b\})$ if it is extended with a_2. This impossibility depends on the assumption that the priorities on the defaults do not depend on facts in the sense that priorities need to be modified if facts are added or removed.

A framework of priorities for default logic in which priorities are assigned to defaults instead of the conclusions of the defaults has more expressivity as demonstrated by the above example. Still more expressivity could be obtained if priorities were assigned to sets or chains of defaults (this is essentially

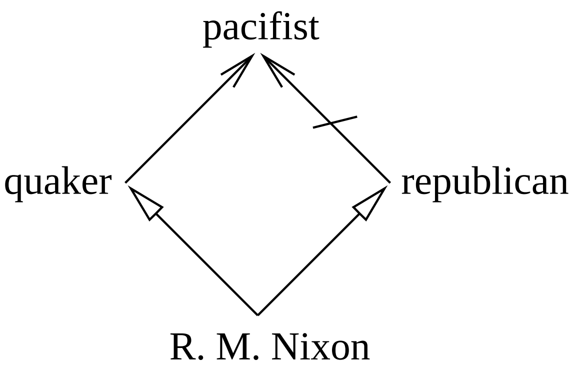


Figure 3.2: The Nixon diamond

what happens in path-based theories of inheritance [Horty, 1994], where the preemption relation is the priority relation,) or whole extensions. In this work we restrict to the case in which priorities are assigned to individual defaults.

The idea of basing nonmonotonic reasoning on arbitrary orderings on $\mathcal{L}$ has been investigated in a setting inspired by the research on belief revision [Gärdenfors and Makinson, 1994; Gärdenfors, 1988]. An immediate practical problem is that such an ordering in general cannot be represented finitely. This problem has been addressed in [Fariñas del Cerro *et al.*, 1994].

Partiality in priorities

When sequences of letters are ordered lexicographically there is a unique lexicographically best sequence. However, in inheritance reasoning, there in general is no unique best extension. The standard example of this phenomenon is the Nixon diamond [Reiter and Criscuolo, 1981] in Figure 3.2. In this case the defaults should not be totally ordered: the default stating that quakers usually are pacifists does not have a higher priority than the default stating that republicans usually are not pacifists, or vice versa. Lexicographic comparison is usually understood only in the setting where the criteria according to which two entities are ordered form a totally ordered set. In our setting the criteria are not totally ordered, and hence a generalization of lexicographic comparison to the case where the criteria are ordered by any strict partial order is needed.

A natural way to generalize lexicographic comparison to partially ordered sets is to reduce it to lexicographic comparison with totally ordered criteria. Hence a candidate is better than another candidate if the former is lexicographically better than the other according to all total orders that extend the partial order on the criteria. Now the best candidates are those for which there are no better candidates.

However, there is another natural definition for lexicographically best candidates. According to this definition a candidate is best if there is a total order on the criteria that extends the partial order and the candidate is better than any other candidate according to lexicographic comparison by using this total order. Hence, instead of requiring that a candidate is pairwise not worse than any other candidate by using lexicographic comparison with some total order, it is required that the candidate is better than any other candidate by using the same total order for all pairwise comparisons. Clearly, this second definition is stronger than the first definition in the sense that the set of best candidates is a subset of the best candidates in the first case.

These two definitions of lexicographically best candidates stem from different ways of utilizing the partiality in partial orders. A definition that we show equivalent to the first definition has been used for example by Geffner and Pearl [1992] and Ryan [1992a]. The second definition has been used by Brewka [1989]. Neither Geffner and Pearl nor Ryan explicitly do the reduction to lexicographic comparison with total orders. Brewka's definition, on the other hand, does not mention lexicographic comparison as the lexicographically best candidates are selected implicitly by a procedure.

3.1 Prioritized default logic

The significance of defaults is represented by strict partial orders, that is, transitive and asymmetric relations $\mathcal{P}$. If defaults δ and δ' are related as $\delta\mathcal{P}\delta'$, then the rejection of δ' is more acceptable than the rejection of δ. The default δ is more significant or has a higher priority. Informally, extensions are compared according to which defaults are in them. Extensions are sets of formulae and as default rules are not formulae the above statement needs clarification. There may be several different interpretations of what it means for a default to be in an extension, and a meaning we adopt is that the default is one of the generating defaults of the extension (see Theorem 2.1.5.) We say that a generating default of an extension is *applied* in the extension. Hence we classify defaults to those that are applied and to those that are not.

Definition 3.1.1 (Application) *A default $\alpha{:}\beta_1,\ldots,\beta_n/\gamma$ is* applied *in $E \subseteq \mathcal{L}$, if $\alpha \in E$ and $\{\neg\beta_1,\ldots,\neg\beta_n\} \cap E = \emptyset$. For this we use the notation $appl(\alpha{:}\beta_1,\ldots,\beta_n/\gamma, E)$.*

Many other schemes for classifying defaults are possible. A property of a useful scheme is that for any two different extensions the respective classifications of defaults are different. This property guarantees that any two extensions can be distinguished on the basis of the classification of the defaults. In the case of application, Theorem 2.1.5 shows that this is indeed the case. Another classification scheme introduced below is to defaults that are *non-defeated* and *defeated* (justified by Theorem 2.1.3 and Lemma A.1.5.) A third scheme, suggested by Theorem 2.1.3, is to classify defaults to those for which the negation of one of the justifications is in the extension, and to those for which none is. A fourth scheme, that is ternary unlike the afore-mentioned binary ones, is to classify defaults to those that are applied, those that are not applied nor defeated, and to those that are defeated.

Only two of the possibly large number of interesting classification schemes are used in this work. The classification to applied and non-applied defaults is the basis of most of our results. In particular, the basic results on prioritized default logic in this section are given only for priorities that use application. Similar results can be obtained for priorities that are based on a classification to defaults that are defeated and to defaults that are not. In the section on the computational complexity of prioritized default reasoning, results on the complexity of both kinds of prioritized default logic are given. In Sections 7.3.1 and 7.2.3 priorities based on the notion of defeat seem more useful in establishing some results. We have not discovered fundamental differences between these two ways of classifying defaults, and for prerequisite-free normal default theories they coincide.

Definition 3.1.2 (Defeat) *A default $\alpha{:}\beta_1, \ldots, \beta_n/\gamma$ is* defeated in $E \subseteq \mathcal{L}$*, if* $\alpha \in E$ *and* $\{\neg\beta_1, \ldots, \neg\beta_n\} \cap E \neq \emptyset$*. This is denoted by* $\mathit{def}(\alpha{:}\beta_1, \ldots, \beta_n/\gamma, E)$*.*

For convenience, we introduce the notation $\mathrm{appl}(\delta, E, E')$ that abbreviates $\mathrm{appl}(\delta, E)$ and not $\mathrm{appl}(\delta, E')$. This notation is useful because the unabbreviated form is rather clumsy, and it would be frequently used in the body of this work. Similarly, $\mathrm{def}(\delta, E, E')$ means that $\mathrm{def}(\delta, E)$ and not $\mathrm{def}(\delta, E')$.

Definition 3.1.3 (Preferredness) *Let $\Delta = \langle D, W\rangle$ be a default theory and $\mathcal{P}$ a strict partial order on D. Let E be an extension of Δ. Then E is a* $\mathcal{P}$-preferred extension *of Δ, if there is a strict total order $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$ and for all extensions E' of Δ and $\delta \in D$,*

$$\mathit{appl}(\delta, E', E) \text{ implies that for some } \epsilon \in D, \epsilon\mathcal{T}\delta \text{ and } \mathit{appl}(\epsilon, E, E').$$

Such a strict total order is a $\Delta, \mathcal{P}$-characteristic ordering *for E.*

The above definition may seem unfamiliar because of the way lexicographic comparison is formulated in it. For finite sets of defaults D the statement "$\mathrm{appl}(\delta, E', E)$ implies that for some $\epsilon \in D, \epsilon\mathcal{T}\delta$ and $\mathrm{appl}(\epsilon, E, E')$" can be equivalently rephrased as "either $E = E'$ or there is $\delta \in D$ such that $\mathrm{appl}(\delta, E, E')$ and there is no $\epsilon \in D$ such that $\epsilon\mathcal{T}\delta$ and $\mathrm{appl}(\epsilon, E', E)$." The reason why we have decided to use the first formulation is that a variant of it has been used in most of the prior work on priorities and lexicographic comparison.

Definition 3.1.4 *The consequence relation* $\models \subseteq ((2^{\mathcal{D}} \times 2^{\mathcal{L}}) \times 2^{\mathcal{D}\times\mathcal{D}}) \times \mathcal{L}$ *is defined as the relation such that* $\Delta \models_{\mathcal{P}} \phi$ *if and only if the formula* ϕ *is in all* $\mathcal{P}$*-preferred extensions of* Δ*.*

We shall use the notation $\models_{\mathcal{P}}$ for the consequence relation of prioritzed default logic to distinguish it from the logical consequence relation $\models$ of the classical propositional logic.

Example 3.1.1 The default theory $\langle D, W\rangle$ with $D = \{p{:}q/q, p{:}\neg r/\neg r, q{:}r/r\}$ and $W = \{p\}$ is depicted in Figure 3.1. Define the strict partial order $\mathcal{P}$ on D as

$$\mathcal{P} = \left\{ \left\langle \frac{p:\neg r}{\neg r}, \frac{q:r}{r} \right\rangle \right\}.$$

The default theory has two extensions, $E_1 = Cn(\{p, q, \neg r\})$ where the defaults $p{:}\neg r/\neg r$ and $p{:}q/q$ are applied, and $E_2 = Cn(\{p, q, r\})$ where $p{:}q/q$ and $q{:}r/r$ are applied. These extensions and all strict total orders $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$ are depicted below. The most significant defaults are the lowest. The symbol $\bullet$ signifies that the default is applied and $\circ$ that it is not applied.

$$\begin{array}{ccc@{\qquad}ccc@{\qquad}ccc}
 & E_1 & E_2 & & E_1 & E_2 & & E_1 & E_2 \\
\frac{p:q}{q} & \bullet & \bullet & \frac{q:r}{r} & \circ & \bullet & \frac{q:r}{r} & \circ & \bullet \\
\frac{q:r}{r} & \circ & \bullet & \frac{p:q}{q} & \bullet & \bullet & \frac{p:\neg r}{\neg r} & \bullet & \circ \\
\frac{p:\neg r}{\neg r} & \bullet & \circ & \frac{p:\neg r}{\neg r} & \bullet & \circ & \frac{p:q}{q} & \bullet & \bullet
\end{array}$$

The extension E_1 is a $\mathcal{P}$-preferred extension because for example the leftmost strict total order $\mathcal{T}_1$ is a $\Delta, \mathcal{P}$-characteristic ordering for E_1: $q{:}r/r$ is the only default δ such that $\mathrm{appl}(\delta, E_2, E_1)$, and $\mathrm{appl}(p{:}\neg r/\neg r, E_1, E_2)$ and $p{:}\neg r/\neg r \mathcal{T}_1 q{:}r/r$. The extension E_2 is not a $\mathcal{P}$-preferred extension because none of the three strict total orders $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$ is a $\Delta, \mathcal{P}$-characteristic ordering for E_2: for all $i \in \{1, 2, 3\}$, there is the default $p{:}\neg r/\neg r$ such that $\mathrm{appl}(p{:}\neg r/\neg r, E_1, E_2)$ and there is no default δ such that $\delta \mathcal{T}_i p{:}\neg r/\neg r$ and $\mathrm{appl}(\delta, E_2, E_1)$. ■

The alternative way of interpreting the partiality in priorities is used in the following definition.

Definition 3.1.5 (Preferredness$_2$) *Let* $\Delta = \langle D, W\rangle$ *be a default theory, and* $\mathcal{P}$ *a strict partial order on* D*. Let* E *be an extension of* Δ*. Then* E *is* *a* $\mathcal{P}$-preferred$_2$ extension *of* Δ*, if for all extensions* E' *of* Δ *there is a strict total order* $\mathcal{T}$ *on* D *such that* $\mathcal{P} \subseteq \mathcal{T}$ *and for all* $\delta \in D$*,*

$$appl(\delta, E', E) \textit{ implies that for some } \epsilon \in D, \epsilon \mathcal{T} \delta \textit{ and } appl(\epsilon, E, E').$$

Such a strict total order is a $\Delta, \mathcal{P}$-characteristic ordering against E' *for* E*.*

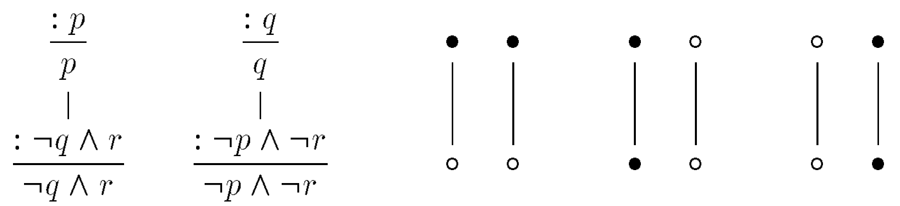
Definition 3.1.6 *The consequence relation* $\models^2 \subseteq ((2^{\mathcal{D}} \times 2^{\mathcal{L}}) \times 2^{\mathcal{D} \times \mathcal{D}}) \times \mathcal{L}$ *is defined as the relation such that* $\Delta \models^2_{\mathcal{P}} \phi$ *if and only if the formula* ϕ *is in all* $\mathcal{P}$*-preferred*$_2$ *extensions of* Δ*.*

By replacing def(δ, E, E') by appl(δ, E', E) and def(ϵ, E', E) by appl(ϵ, E, E') in the above definitions, definitions of $\mathcal{P}$-preferred$_d$ and $\mathcal{P}$-preferred$_{2d}$ extensions are obtained. Similarly we obtain the consequence relations $\models^d_{\mathcal{P}}$ and $\models^{2d}_{\mathcal{P}}$. We use the term characteristic$_d$ when referring to characteristic orderings of preferred$_{2d}$ extensions.

Lemma 3.1.7 (Inclusion) *Let* $\Delta = \langle D, W \rangle$ *be a default theory. Let* $\mathcal{P}$ *be a strict partial order on* D*. If* E *is a* $\mathcal{P}$*-preferred extension of* Δ*, then* E *is a* $\mathcal{P}$*-preferred*$_2$ *extension of* Δ*. The converse does not hold.*

Proof: Let $\Delta = \langle D, W \rangle$ be a default theory and $\mathcal{P}$ a strict partial order on D. If an extension E of Δ is $\mathcal{P}$-preferred, then there is a $\Delta, \mathcal{P}$-characteristic ordering $\mathcal{T}$ for E. Now $\mathcal{T}$ fulfills the condition that for all extensions E' of Δ and all defaults δ, appl(δ, E', E) implies that for some $\epsilon \in D$, $\epsilon \mathcal{T} \delta$ and appl(ϵ, E, E'). Let E' be any extension of Δ. By definition of characteristic orderings against extensions, $\mathcal{T}$ is also a $\Delta, \mathcal{P}$-characteristic ordering against E' for E. Because this holds for any extension E', E is a $\mathcal{P}$-preferred$_2$ extension of Δ.

The following example demonstrates that the converse does not hold. Let $\Delta = \langle D, \emptyset \rangle$ be a default theory with four defaults in D and the strict partial order $\mathcal{P} = \{\langle :\neg q \wedge r / \neg q \wedge r, :p/p \rangle, \langle :\neg p \wedge \neg r / \neg p \wedge \neg r, :q/q \rangle\}$ on D as depicted in the diagram below on the left. The default theory Δ has the three extensions $E_1 = Cn(\{p, q\}), E_2 = Cn(\{p, \neg q \wedge r\}), E_3 = Cn(\{q, \neg p \wedge \neg r\})$ shown below.



The extension E_1 is $\mathcal{P}$-preferred$_2$ because of the $\Delta, \mathcal{P}$-characteristic ordering $:\neg q \wedge r / \neg q \wedge r \mathcal{T}_3 :p/p \mathcal{T}_3 :\neg p \wedge \neg r / \neg p \wedge \neg r \mathcal{T}_3 :q/q$ against E_2, and the $\Delta, \mathcal{P}$-characteristic ordering $:\neg p \wedge \neg r / \neg p \wedge \neg r \mathcal{T}_2 :q/q \mathcal{T}_2 :\neg q \wedge r / \neg q \wedge r \mathcal{T}_2 :p/p$ against E_3.

However, there are no $\Delta, \mathcal{P}$-characteristic orderings for E_1. In particular, $\mathcal{T}_2$ is a $\Delta, \mathcal{P}$-characteristic ordering for E_2, and $\mathcal{T}_3$ is a $\Delta, \mathcal{P}$-characteristic ordering for E_3. Brewka and Junker [1991] use a similar example to demonstrate that the preference notions of Brewka [1989] and Geffner and Pearl [1992] differ. $\square$

Lemma 3.1.8 *Let $\Delta = \langle D, W \rangle$ be a default theory, ϕ a formula, and $\mathcal{P}$ a strict partial order on D. If $\Delta \models_c \phi$, then $\Delta \models^2_{\mathcal{P}} \phi$. If $\Delta \models^2_{\mathcal{P}} \phi$, then $\Delta \models_{\mathcal{P}} \phi$.*

Proof: $\Delta \models_c \phi$ means that ϕ is in all extensions of Δ. Because $\mathcal{P}$-preferred$_2$ extensions of Δ are a subset of all extensions of Δ, ϕ is in all $\mathcal{P}$-preferred$_2$ extensions of Δ. By Lemma 3.1.7 the $\mathcal{P}$-preferred extensions of Δ are a subset of the $\mathcal{P}$-preferred$_2$ extensions of Δ, and therefore the membership in all $\mathcal{P}$-preferred$_2$ extensions of Δ implies the membership in all $\mathcal{P}$-preferred extensions of Δ. $\square$

Our definitions of preferred and preferred$_2$ extensions directly determine a subclass of all the extensions of a default theory. This is in contrast with some earlier work in prioritized nonmonotonic reasoning where priorities impose an ordering on propositional models or extensions, and the preferred models or extensions are the minimal or maximal elements. It seems that there is no natural definition of preferred extensions that uses orderings of extensions in that way. However, for preferred$_2$ extensions the priorities impose an ordering on the extensions so that the preferred extensions are the minimal elements.

Definition 3.1.9 *Let $\Delta = \langle D, W \rangle$ be a default theory and let E and E' be extensions of Δ. Let $\mathcal{P}$ be a strict partial order on D. Then $E \sqsubset^{\Delta}_{\mathcal{P}} E'$ if and only if $E \neq E'$ and for all $\delta \in D$ and strict total orders $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$,*

$$appl(\delta, E', E) \text{ implies that for some } \epsilon \in D, \epsilon \mathcal{T} \delta \text{ and } appl(\epsilon, E, E').$$

The assumption concerning $\mathcal{P}$ in the next theorem is to guarantee that all strict total orders on D that extend $\mathcal{P}$ are well-orderings, that is, for any subset of D there is a $\mathcal{T}$-least element.

Theorem 3.1.10 *Let $\Delta = \langle D, W \rangle$ be a default theory. Let $\mathcal{P}$ be a strict partial order on D such that for each $\delta \in D$ the set $\{\delta' \in D | \text{not } \delta \mathcal{P} \delta'\}$ is finite. Then the relation $\sqsubset^{\Delta}_{\mathcal{P}}$ is transitive and asymmetric.*

Proof: Let $\mathcal{T}$ be a strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$, and let δ be any member of D. By assumption the set $B = \{\delta' \in D | \text{not } \delta \mathcal{P} \delta'\}$ is finite. Because the set $B' = \{\delta' \in D | \text{not } \delta \mathcal{T} \delta'\}$ is a subset of B, it is finite. Hence there are no infinite descending chains from any default in D and all strict total orders extending $\mathcal{P}$ are well-orderings.

Transitivity: Suppose that $E_1 \sqsubset_{\mathcal{P}}^{\Delta} E_2$ and $E_2 \sqsubset_{\mathcal{P}}^{\Delta} E_3$. We have to show that $E_1 \sqsubset_{\mathcal{P}}^{\Delta} E_3$. Let $\mathcal{T}$ be a strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$. For each $\delta \in D$ such that $\mathrm{appl}(\delta, E_3, E_1)$ we find $\epsilon \in D$ such that $\mathrm{appl}(\epsilon, E_1, E_3)$ and $\epsilon \mathcal{T} \delta$. We consider the following two cases.

Suppose $\mathrm{appl}(\delta, E_3, E_1)$ and not $\mathrm{appl}(\delta, E_2)$. By the definition of $E_2 \sqsubset_{\mathcal{P}}^{\Delta} E_3$ there is $\delta' \in D$ such that $\mathrm{appl}(\delta', E_2, E_3)$ and $\delta' \mathcal{T} \delta$. Let δ' be the $\mathcal{T}$-least such default (because $\mathcal{T}$ is a well-ordering, there is such a default.) If $\mathrm{appl}(\delta', E_1)$ then we have found a default $\epsilon = \delta'$ such that $\mathrm{appl}(\epsilon, E_1, E_3)$. If not $\mathrm{appl}(\delta', E_1)$ then by the definition of $E_1 \sqsubset_{\mathcal{P}}^{\Delta} E_2$ there is $\delta'' \in D$ such that $\mathrm{appl}(\delta'', E_1, E_2)$ and $\delta'' \mathcal{T} \delta'$. Now not $\mathrm{appl}(\delta'', E_3)$, because if $\mathrm{appl}(\delta'', E_3, E_2)$ and $E_2 \sqsubset_{\mathcal{P}}^{\Delta} E_3$, there has to be a default δ''' such that $\mathrm{appl}(\delta''', E_2, E_3)$ and $\delta''' \mathcal{T} \delta''$, but this is impossible because $\delta''' \mathcal{T} \delta'$ by the transitivity of $\mathcal{T}$ and δ' is the $\mathcal{T}$-least default such that $\delta' \mathcal{T} \delta$ and $\mathrm{appl}(\delta', E_2, E_3)$. Hence $\mathrm{appl}(\delta'', E_1, E_3)$ and $\delta'' \mathcal{T} \delta$.

Suppose $\mathrm{appl}(\delta, E_3, E_1)$ and $\mathrm{appl}(\delta, E_2)$. By the definition of $E_1 \sqsubset_{\mathcal{P}}^{\Delta} E_2$ there is $\delta' \in D$ such that $\mathrm{appl}(\delta', E_1, E_2)$ and $\delta' \mathcal{T} \delta$. Let δ' be the $\mathcal{T}$-least default such that $\delta' \mathcal{T} \delta$ and $\mathrm{appl}(\delta', E_1, E_2)$. If not $\mathrm{appl}(\delta', E_3)$ then we have found a default such that $\mathrm{appl}(\delta', E_1, E_3)$ and $\delta' \mathcal{T} \delta$. Otherwise $\mathrm{appl}(\delta', E_3, E_2)$. By the definition of $E_2 \sqsubset_{\mathcal{P}}^{\Delta} E_3$ there is $\delta'' \in D$ such that $\mathrm{appl}(\delta'', E_2, E_3)$ and $\delta'' \mathcal{T} \delta'$. Now it has to be the case that $\mathrm{appl}(\delta'', E_1, E_3)$, because if not $\mathrm{appl}(\delta'', E_1)$, then $\mathrm{appl}(\delta'', E_2, E_1)$ and there has to be δ''' such that $\delta''' \mathcal{T} \delta''$ and $\mathrm{appl}(\delta''', E_1, E_2)$, and this contradicts the assumption that δ' is the $\mathcal{T}$-least such default as $\delta''' \mathcal{T} \delta'' \mathcal{T} \delta'$.

Hence for all $\delta \in D$ such that $\mathrm{appl}(\delta, E_3, E_1)$ there is $\epsilon \in D$ such that $\epsilon \mathcal{T} \delta$ and $\mathrm{appl}(\epsilon, E_1, E_3)$.

Asymmetry: Suppose $E_1 \sqsubset_{\mathcal{P}}^{\Delta} E_2$ and $E_2 \sqsubset_{\mathcal{P}}^{\Delta} E_1$ for some extensions E_1 and E_2 of Δ. By definition of $\sqsubset_{\mathcal{P}}^{\Delta}$ E_1 and E_2 are different extensions, and by [Reiter, 1980, Theorems 2.4 and 2.5] there are defaults δ_0 and ϵ_0 such that $\mathrm{appl}(\delta_0, E_2, E_1)$ and $\mathrm{appl}(\epsilon_0, E_1, E_2)$. Let $\mathcal{T}$ be any strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$. By assumption it is a well-ordering. Therefore for all $\delta \in D$, $\mathrm{appl}(\delta, E_2, E_1)$ implies that there is $\epsilon \in D$ such that $\epsilon \mathcal{T} \delta$ and $\mathrm{appl}(\delta, E_1, E_2)$, and similarly for all $\epsilon \in D$, $\mathrm{appl}(\epsilon, E_1, E_2)$ implies that there is $\delta \in D$ such that $\delta \mathcal{T} \epsilon$ and $\mathrm{appl}(\delta, E_2, E_1)$. Starting from ϵ_0 we get an infinite sequence of defaults $\epsilon_0, \delta_1, \epsilon_1, \delta_2, \ldots$ such that for all $i \geq 0$, $\mathrm{appl}(\epsilon_i, E_1, E_2)$, and $\mathrm{appl}(\delta_i, E_2, E_1)$ and $\delta_{i+1} \mathcal{T} \epsilon_i$ and $\delta_{i+1} \mathcal{T} \epsilon_{i+1}$. This contradicts the assumption that there can be no infinite descending sequences of defaults in D because $\mathcal{T}$ is a well-ordering. Hence the assumption that $E_1 \sqsubset_{\mathcal{P}}^{\Delta} E_2$ and $E_2 \sqsubset_{\mathcal{P}}^{\Delta} E_1$ is false, and $\sqsubset_{\mathcal{P}}^{\Delta}$ is asymmetric. $\square$

The following lemmata and theorem establish the fact that the minimal elements of $\sqsubset^{\Delta}_{\mathcal{P}}$ are exactly the $\mathcal{P}$-preferred$_2$ extensions of Δ.

Lemma 3.1.11 ($\sqsubset^{\Delta}_{\mathcal{P}}$ monotonicity) *Let $\Delta = \langle D, W\rangle$ be a default theory, and $\mathcal{P}$ and $\mathcal{P}'$ strict partial orders on D such that $\mathcal{P}' \subseteq \mathcal{P}$. Then $\sqsubset^{\Delta}_{\mathcal{P}'} \subseteq \sqsubset^{\Delta}_{\mathcal{P}}$.*

Proof: Let E and E' be extensions of Δ and $E \sqsubset^{\Delta}_{\mathcal{P}'} E'$. Let $\mathcal{T}$ be any strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$. Because $\mathcal{P}' \subseteq \mathcal{P}$, also $\mathcal{P}' \subseteq \mathcal{T}$. Because $E \sqsubset^{\Delta}_{\mathcal{P}'} E'$, for all $\delta \in D$, $\mathrm{appl}(\delta, E', E)$ implies that for some $\epsilon \in D, \epsilon\mathcal{T}\delta$ and $\mathrm{appl}(\epsilon, E, E')$. As this holds for any strict total order $\mathcal{T}$ such that $\mathcal{P} \subseteq \mathcal{T}$, by definition of $\sqsubset^{\Delta}_{\mathcal{P}}$ finally $E \sqsubset^{\Delta}_{\mathcal{P}} E'$. Therefore $\sqsubset^{\Delta}_{\mathcal{P}'} \subseteq \sqsubset^{\Delta}_{\mathcal{P}}$. □

Lemma 3.1.12 *Let $\Delta = \langle D, W\rangle$ be a default theory, $\mathcal{P}$ a strict partial order on D, E and E' extensions of Δ such that $E \neq E'$, and $\mathcal{T}$ a $\Delta,\mathcal{P}$-characteristic ordering for E. Then $E \sqsubset^{\Delta}_{\mathcal{T}} E'$.*

Proof: Because $\mathcal{T}$ is a $\Delta, \mathcal{P}$-characteristic ordering for E and E' is an extension of Δ, for all $\delta \in D$, $\mathrm{appl}(\delta, E', E)$ implies that for some $\epsilon \in D, \epsilon\mathcal{T}\delta$ and $\mathrm{appl}(\epsilon, E, E')$. Because $\mathcal{T}$ is the only strict total order on D that extends $\mathcal{T}$, $E \sqsubset^{\Delta}_{\mathcal{T}} E'$ by definition of $\sqsubset^{\Delta}_{\mathcal{T}}$. □

Theorem 3.1.13 *Let $\Delta = \langle D, W\rangle$ be a default theory, E an extension of Δ, and $\mathcal{P}$ a strict partial order on D such that for each $\delta \in D$ the set $\{\delta' \in D | \text{not } \delta\mathcal{P}\delta'\}$ is finite. Then E is a $\mathcal{P}$-preferred$_2$ extension of Δ if and only if there is no E' such that $E' \sqsubset^{\Delta}_{\mathcal{P}} E$.*

Proof: Assume that E is a $\mathcal{P}$-preferred$_2$ extension of Δ. By definition of $\mathcal{P}$-preferredness$_2$, there is a $\Delta, \mathcal{P}$-characteristic ordering $\mathcal{T}$ against E' for E. By Lemma 3.1.12 $E \sqsubset^{\Delta}_{\mathcal{T}} E'$. Because by Theorem 3.1.10 $\sqsubset^{\Delta}_{\mathcal{T}}$ is asymmetric, not $E' \sqsubset^{\Delta}_{\mathcal{T}} E$. By Lemma 3.1.11 and the fact that $\mathcal{P} \subseteq \mathcal{T}$ not $E' \sqsubset^{\Delta}_{\mathcal{P}} E$.

Assume E is an extension of Δ that is not $\mathcal{P}$-preferred$_2$. Then for some extension E' there are no $\Delta, \mathcal{P}$-characteristic orderings against E' for E. Now for all strict total orders $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$ there is $\delta_{\mathcal{T}} \in D$ such that $\mathrm{appl}(\delta_{\mathcal{T}}, E', E)$ and there is no $\delta'\mathcal{T}\delta_{\mathcal{T}}$ such that $\mathrm{appl}(\delta', E, E')$. Take any strict total order $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$. Let $\delta \in D$ be such that $\mathrm{appl}(\delta, E, E')$. Now $\delta_{\mathcal{T}}\mathcal{T}\delta$. Because $\mathrm{appl}(\delta_{\mathcal{T}}, E', E)$, finally $E' \sqsubset^{\Delta}_{\mathcal{P}} E$. □

The relation $\sqsubset^{\Delta}_{\mathcal{P}}$ can be defined directly on the basis of $\mathcal{P}$ without reducing lexicographic comparison with partial orders to lexicographic comparison with total orders. This kind of definition is used by Geffner and Pearl [1992] and Ryan [1992a].

Theorem 3.1.14 *Let E and E' be extensions of a default theory $\Delta = \langle D, W\rangle$. Let $\mathcal{P}$ be a strict partial order on D. Then $E \sqsubset^{\Delta}_{\mathcal{P}} E'$ if and only if $E \neq E'$ and for all $\delta \in D$,*

$$appl(\delta, E', E) \text{ implies that for some } \epsilon \in D, \epsilon\mathcal{P}\delta \text{ and } appl(\epsilon, E, E').$$

Proof: Assume that for all $\delta \in D$, appl(δ, E', E) implies that for some $\epsilon \in D$, $\epsilon\mathcal{P}\delta$ and appl(ϵ, E, E'). Let $\mathcal{T}$ be any strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$. Let δ be any member of D. If appl(δ, E', E), then by the above assumption there is $\epsilon \in D$ such that $\epsilon\mathcal{P}\delta$ and appl(ϵ, E, E'). Because $\mathcal{P} \subseteq \mathcal{T}$, also $\epsilon\mathcal{T}\delta$. Because this holds for all $\mathcal{T}$ such that $\mathcal{P} \subseteq \mathcal{T}$, by definition of $\sqsubset^{\Delta}_{\mathcal{P}}$, $E \sqsubset^{\Delta}_{\mathcal{P}} E'$.

Assume that it is not the case that for all $\delta \in D$, appl(δ, E', E) implies that for some $\epsilon \in D$, $\epsilon\mathcal{P}\delta$ and appl(ϵ, E, E'). That is, for some $\delta \in D$, appl(δ, E', E) and there is no $\epsilon \in D$ such that $\epsilon\mathcal{P}\delta$ and appl(ϵ, E, E'). Take any strict total order $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$ and $\delta\mathcal{T}\mu$ for all $\mu \in D$ such that not $\mu\mathcal{P}\delta$. Hence there is no $\epsilon \in D$ such that $\epsilon\mathcal{T}\delta$ and appl(ϵ, E, E'). By definition of $\sqsubset^{\Delta}_{\mathcal{P}}$ we conclude that not $E \sqsubset^{\Delta}_{\mathcal{P}} E'$. □

A subclass of strict partial orders that arises for example from the view of priorities as integer valued functions on a set of elements, is that of *layered* strict partial orders. With layered strict partial orders there can be no two defaults one of which has a higher priority than the other, and both of which have a priority incomparable to a third default. Brewka [1989] and Benferhat et al. [1993] investigate this class of priorities in the context of maximally consistent subsets of formulae.

Theorem 3.1.15 (Layered priorities) *Let $\Delta = \langle D, W\rangle$ be a default theory and $\mathcal{P}$ a strict partial order on D such that for $D_1 \cup D_2 \cup \cdots \cup D_n$,*

$$\mathcal{P} = \{\langle \delta, \delta'\rangle | \delta \in D_i, \delta' \in D_j, i < j\}.$$

Then for all extensions E of Δ, E is a $\mathcal{P}$-preferred extension if and only if E is a $\mathcal{P}$-preferred$_2$ extension.

Proof: By Lemma 3.1.7 each $\mathcal{P}$-preferred extension of Δ is $\mathcal{P}$-preferred$_2$. We have to show that for layered $\mathcal{P}$, each $\mathcal{P}$-preferred$_2$ extension of Δ is $\mathcal{P}$-preferred. Let E be a $\mathcal{P}$-preferred$_2$ extension of Δ. Let $\mathcal{T}$ be a strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$ and $\delta\mathcal{T}\delta'$ whenever appl(δ, E) and not appl(δ', E) and neither $\delta\mathcal{P}\delta'$ nor $\delta'\mathcal{P}\delta$. Such an ordering can be constructed by taking defaults in D layer by layer starting from the most significant one, and taking the ones applied in E before the ones in the same layer that are not applied in E. Let E' be an extension of Δ such that $E \neq E'$. Because E is $\mathcal{P}$-preferred$_2$, by Theorem 3.1.13 it is not the case that $E' \sqsubset^{\Delta}_{\mathcal{P}} E$. Therefore by

Theorem 3.1.14 for some $\delta \in D$ such that appl(δ, E, E') there is no $\delta'\mathcal{P}\delta$ such that appl(δ', E', E). By definition of $\mathcal{T}$, all δ' such that $\delta'\mathcal{T}\delta$ have either the property that $\delta'\mathcal{P}\delta$ or appl(δ', E). Therefore appl(δ', E', E) for no $\delta'\mathcal{T}\delta$. Now $\mathcal{T}$ is a $\Delta, \mathcal{P}$-characteristic ordering for E, and E is a $\mathcal{P}$-preferred extension of Δ. □

Reiter's default logic is a special case of prioritized default logic. The embedding is the restriction to empty priorities.

Lemma 3.1.16 *Let Δ be a default theory and $\mathcal{P}$ the empty relation. Then all extensions of Δ are $\mathcal{P}$-preferred and $\mathcal{P}$-preferred$_2$.*

Proof: Take any extension E of $\Delta = \langle D, W \rangle$. Take any strict total order $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$ and for all $\delta \in D$ and $\delta' \in D$, $\delta\mathcal{T}\delta'$ whenever appl(δ, E) and not appl(δ', E). Assume that there is an extension E' such that for some $\delta' \in D$, appl(δ', E', E) but there is no $\delta \in D$ such that $\delta\mathcal{T}\delta'$ and appl(δ, E, E'). By definition of $\mathcal{T}$, $\delta\mathcal{T}\delta'$ for all δ such that appl(δ, E, E'). This means that $\{\delta \in D | \text{appl}(\delta, E')\} \subset \{\delta \in D | \text{appl}(\delta, E)\}$. But this contradicts [Reiter, 1980, Theorems 2.4 and 2.5]. Hence $\mathcal{T}$ is a $\Delta, \mathcal{P}$-characteristic ordering for E and E is a $\mathcal{P}$-preferred extension of Δ. □

Intuitively, the purpose of priorities is to indicate which ways of resolving a conflict between defaults are acceptable. If priorities do not determine how a certain conflict should be resolved, a form of case analysis is performed by resolving conflicts in all possible ways. If more priorities are given, the number of possible ways of resolving the conflicts decreases, and the number of conclusions increases. Our definition of priorities has this monotonicity property.

Lemma 3.1.17 (Monotonicity) *Let $\Delta = \langle D, W \rangle$ be a default theory. Let E be a $\mathcal{P}$-preferred ($\mathcal{P}$-preferred$_2$) extension of Δ for some strict partial order $\mathcal{P}$ on D. Let $\mathcal{P}'$ be a strict partial order on D such that $\mathcal{P}' \subseteq \mathcal{P}$. Then E is a $\mathcal{P}'$-preferred ($\mathcal{P}'$-preferred$_2$) extension of Δ.*

Proof: If E is a $\mathcal{P}$-preferred extension of Δ, there is a $\Delta, \mathcal{P}$-characteristic ordering $\mathcal{T}$ for E. Because $\mathcal{P}' \subseteq \mathcal{P}$, $\mathcal{T}$ is also a $\Delta, \mathcal{P}'$-characteristic ordering for E, and hence E is a $\mathcal{P}'$-preferred extension. Similarly if E is a $\mathcal{P}$-preferred$_2$ extension of Δ, for all extensions E' of Δ there is a $\Delta, \mathcal{P}$-characteristic ordering $\mathcal{T}$ against E' for E. Because $\mathcal{P}' \subseteq \mathcal{P}$, $\mathcal{T}$ is also a $\Delta, \mathcal{P}'$-characteristic ordering against E' for E, and hence E is a $\mathcal{P}'$-preferred$_2$ extension. □

Results characterizing preferred$_2$ extensions as minimal elements of strict partial orders on extensions indirectly demonstrate that for default theories with a finite non-empty set of extensions the existence of preferred$_2$ extensions is guaranteed for any strict partial order on the defaults. The next results answer the question of the existence of preferred extensions under same assumptions. Moreover, these results establish the existence of preferred$_2$ extensions in an alternative way. The restriction to finite sets of defaults made in Lemma 3.1.18 and Theorem 3.1.20 is justified by Example 3.1.2.

Lemma 3.1.18 *Let $\Delta = \langle D, W\rangle$ be a default theory where D is finite, and let $\mathcal{P}$ be a strict total order on D. Let Δ have at least one extension. Then there is exactly one $\mathcal{P}$-preferred extension of Δ.*

Proof: We show that there is an extension E of Δ such that $\mathcal{P}$ is the $\Delta, \mathcal{P}$-characteristic ordering for E. Let $\delta_1, \ldots, \delta_n$ be the ordering $\mathcal{P}$ of D. Define $D_i = \{\delta_1, \ldots, \delta_i\}$ for all $i \in \{0, \ldots, n\}$. Define for all $i \in \{1, \ldots, n-1\}$,

$$\begin{aligned} X_0 &= \{E \subseteq \mathcal{L} | E \text{ is an extension of } \Delta\}, \text{ and} \\ X_{i+1} &= \begin{cases} \{E \in X_i | \mathrm{appl}(\delta_{i+1}, E)\} & \text{if } \mathrm{appl}(\delta_{i+1}, E) \text{ for some } E \in X_i, \\ X_i & \text{otherwise.} \end{cases} \end{aligned}$$

Induction hypothesis: For $j \in \{0, \ldots, i\}$,

1. the set X_j is non-empty,
2. for all $E \in X_j$ and $E' \in X_j$ and $\delta \in D_j$, $\mathrm{appl}(\delta, E)$ iff $\mathrm{appl}(\delta, E')$, and
3. for all $E \in X_j$ and $E' \in X_0 \backslash X_j$ there is $\delta \in D_j$ such that $\mathrm{appl}(\delta, E, E')$ and there is no $\delta' \in D$ such that $\delta' \mathcal{P} \delta$ and $\mathrm{appl}(\delta', E', E)$.

Base case $i = 0$:

1. X_0 is non-empty by the assumption of the lemma.
2. Because $D_0 = \emptyset$, for all $\delta \in D_0$ everything holds.
3. Because $X_0 \backslash X_0 = \emptyset$, for all $E' \in X_0 \backslash X_0$ everything holds.

Inductive case $i \geq 1$:

1. For $j \in \{1, \ldots, i-1\}$ the claim is directly by the induction hypothesis. If there is an extension $E \in X_{i-1}$ such that $\mathrm{appl}(\delta_i, E)$, then $X_i = \{E \in X_{i-1} | \mathrm{appl}(\delta_i, E)\}$ and X_i is non-empty. Otherwise $X_i = X_{i-1}$ which by induction hypothesis is non-empty.

2. For $j \in \{1, \ldots, i-1\}$ the claim is directly by the induction hypothesis. Let E and E' be any extensions in X_i. Because $X_i \subseteq X_{i-1}$, by induction hypothesis $\mathrm{appl}(\delta, E)$ iff $\mathrm{appl}(\delta, E')$ for all $\delta \in D_{i-1}$. Assume $\mathrm{appl}(\delta_i, E)$. Now $X_i = \{E \in X_{i-1} | \mathrm{appl}(\delta_i, E)\}$. Because $E' \in X_i$, also $\mathrm{appl}(\delta_i, E')$. Therefore $\mathrm{appl}(\delta_i, E)$ iff $\mathrm{appl}(\delta_i, E')$.

3. For $j \in \{1, \ldots, i-1\}$ the claim is directly by the induction hypothesis. So assume $E \in X_i$ and $E' \in X_0 \backslash X_i$. If $E' \in X_{i-1} \backslash X_i$, then $\mathrm{appl}(\delta_i, E, E')$. By item 2 of induction hypothesis $\mathrm{appl}(\delta, E', E)$ for no $\delta \in D_{i-1}$. Therefore $\mathrm{appl}(\delta, E', E)$ for no $\delta \mathcal{P} \delta_i$. If $E' \in X_0 \backslash X_{i-1}$, then by induction hypothesis and the facts that $E \in X_{i-1}$ and $X_{i-1} \subseteq X_i$, there is $\delta \in D_{i-1}$ such that there is no $\delta' \in D$ such that $\delta' \mathcal{P} \delta$ and $\mathrm{appl}(\delta', E', E)$. Because $D_{i-1} \subseteq D_i$, there is $\delta \in D_i$ such that there is no $\delta' \in D$ such that $\delta' \mathcal{P} \delta$ and $\mathrm{appl}(\delta', E', E)$.

The claim of the lemma is obtained from the facts established in the induction proof as follows. By 1 the set X_n is non-empty. By 2 and [Reiter, 1980, Theorems 2.4 and 2.5] $|X_n| \leq 1$. Hence X_n is a singleton $\{E\}$. Let E' be any extension of Δ. Assume that there is $\delta' \in D$ such that $\mathrm{appl}(\delta, E', E)$. Hence $E' \neq E$ and $E' \in X_0 \backslash X_n$. Now by 3 there is $\delta \in D_n = D$ such that $\mathrm{appl}(\delta, E, E')$ and there is no δ'' such that $\delta'' \mathcal{P} \delta$ and $\mathrm{appl}(\delta'', E', E)$. Therefore not $\delta' \mathcal{P} \delta$, and because $\mathcal{P}$ is a strict total order, it is the case that $\delta \mathcal{P} \delta'$. Because this holds for all $\delta' \in D$ and all extension E' of Δ, $\mathcal{P}$ is a $\Delta, \mathcal{P}$-characteristic ordering for E. Therefore E is a $\mathcal{P}$-preferred extension of Δ. Let E' be any extension such that $E \neq E'$. Now $E \in X_n$ and $E' \notin X_n$, and therefore by 3 there is $\delta \in D$ such that $\mathrm{appl}(\delta, E, E')$ and there is no $\delta' \in D$ such that $\delta' \mathcal{P} \delta$ and $\mathrm{appl}(\delta', E', E)$. Hence E is the only $\mathcal{P}$-preferred extension of Δ. □

Corollary 3.1.19 *Let $\Delta = \langle D, W \rangle$ be a default theory where D is finite, and let $\mathcal{P}$ be any strict total order on D. Let Δ have at least one extension. Then there is exactly one $\mathcal{P}$-preferred$_2$ extension of Δ.*

Proof: Because the strict total order $\mathcal{P}$ is layered (more specifically total), by Theorem 3.1.15 $\mathcal{P}$-preferredness coincides with $\mathcal{P}$-preferredness$_2$. □

Existence results easily follow for strict partial orders in general.

Theorem 3.1.20 *Let $\Delta = \langle D, W \rangle$ be a default theory where D is finite. If Δ has extensions, it has both $\mathcal{P}$-preferred and $\mathcal{P}$-preferred$_2$ extensions for any strict partial order $\mathcal{P}$ on D.*

Proof: Let $\mathcal{P}$ be any strict partial order on D. Let $\mathcal{T}$ be a strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$. By Lemma 3.1.18 there is a $\mathcal{T}$-preferred and $\mathcal{T}$-preferred$_2$

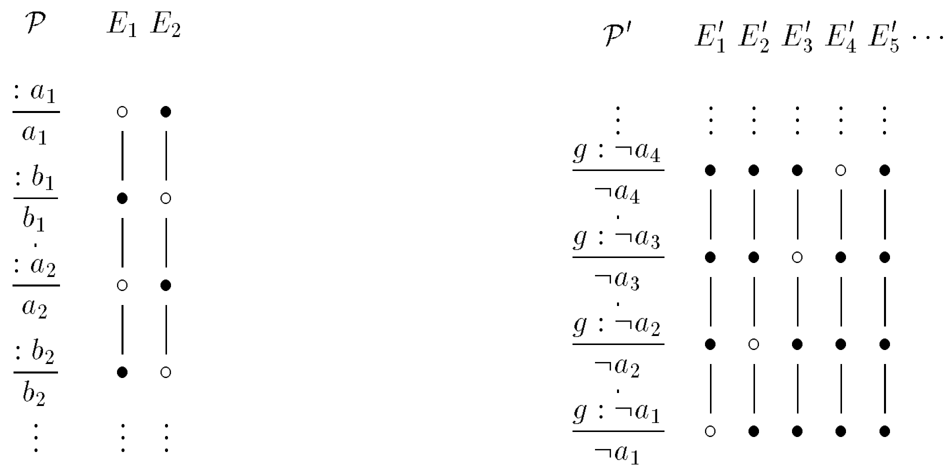


Figure 3.3: Two pathological prioritizations and sets of extensions

extension E of Δ. By Lemma 3.1.17 E is also $\mathcal{P}$-preferred and $\mathcal{P}$-preferred$_2$. □

The following two examples demonstrate two kinds of anomalous priorities $\mathcal{P}$ such that there are no $\mathcal{P}$-preferred extensions. In the first case, the first extension is better than the second and vice versa, and in the second case an ordering can be imposed but there are nevertheless no preferred extensions because for each extension there is an even better one. The restriction in Theorem 3.1.10 that the number of defaults of lower or incomparable priority should be finite is motivated by the first example. The second example demonstrates that in the case of an infinite number of defaults (and infinite $\mathcal{P}$) $\mathcal{P}$-preferred extensions do not necessarily exist even though the strict partial order $\mathcal{P}$ is a well-ordering. This is the justification for the restrictions made in Lemma 3.1.18 and in Theorem 3.1.20.

Example 3.1.2 Let $\Delta = \langle\{:a_i/a_i|i \geq 0\} \cup \{:b_i/b_i|i \geq 0\}, \{(a_i \leftrightarrow \neg b_i) \wedge (a_i \leftrightarrow a_{i+1})|i \geq 1\}\rangle$. The strict total order $\mathcal{P}$ and the two extensions E_1 and E_2 of Δ are depicted in Figure 3.3 on the left. There is no least default in this ordering. For each $:a_i/a_i$ such that appl$(:a_i/a_i, E_1, E_2)$ there is $:b_j/b_j$ such that appl$(:b_j/b_j, E_2, E_1)$ (take any $j \geq i$) and $:b_j/b_j\mathcal{P}:a_i/a_i$, and for each $:b_j/b_j$ such that appl$(:b_j/b_j, E_2, E_1)$ there is $:a_i/a_i$ such that appl$(:a_i/a_i, E_1, E_2)$ (take any $i > j$) and $:a_i/a_i\mathcal{P}:b_j/b_j$. Hence both extensions are $\mathcal{P}$-preferred. ■

Example 3.1.3 Let

$$D = \left\{\frac{:a_i}{a_i}\middle| i \geq 1\right\} \cup \left\{\frac{g : \neg a_i}{\neg a_i}\middle| i \geq 1\right\},$$

$$W = \left\{\begin{array}{l} a_i \rightarrow ua_{i+1} \wedge da_{i-1} \wedge g \\ ua_i \rightarrow ua_{i+1} \wedge \neg a_i \\ da_i \rightarrow da_{i-1} \wedge \neg a_i \end{array}\middle| i \geq 1\right\}, \text{ and}$$

$$\mathcal{P}' = \left\{ \left\langle \frac{g:\neg a_j}{\neg a_j}, \frac{g:\neg a_i}{\neg a_i} \right\rangle \middle| \, i > j \geq 1 \right\} \cup \left\{ \left\langle \frac{g:\neg a_i}{\neg a_i}, \frac{:a_j}{a_j} \right\rangle \middle| \, i \geq 1, j \geq 1 \right\}.$$

The extensions of the default theory $\langle D, W \rangle$ are depicted in Figure 3.3 on the right. In each extension, exactly one of defaults $:a_i/a_i$ is applied, and each $:a_i/a_i$ is applied in exactly one extension. Each extension contains exactly one of a_i and $\neg a_i$ for every $i \geq 1$. The purpose of the formulae in W is to guarantee that whenever the default $:a_i/a_i$ is applied in an extension E, the negations of the justifications of all defaults $:a_k/a_k, k \neq i$ are in E as well, and that the defaults $g{:}a_i/a_i, i \geq 1$ can be applied only if at least one of $:a_i/a_i, i \geq 1$ has been applied. ■

Chapter 4

Complexity of the decision problems

The decision problems of most nonmonotonic logics are at least as complex as the decision problems of the classical logics on which the nonmonotonic logics are based. For example testing the membership of a formula in all extensions of a default theory is obviously co-NP-hard as it has the co-NP-hard problem of propositional entailment as its subproblem.

The decision problems of propositional default logic are located on the second level of the polynomial hierarchy [Gottlob, 1992], and hence they do not belong to the class NP unless the polynomial hierarchy collapses to its first level. This result does not show that automated reasoning in default logic inherently needs more time or other resources than problems in NP, but the proof of the result demonstrates that reasoning in default logic cannot be performed in polynomial time simply by restricting the formulae in default theories to a tractable subclass of propositional logic like 2-literal clauses or Horn clauses. This fact can be seen as a consequence of the problem of conflicting defaults. A conflict between defaults gives rise to multiple extensions because a case analysis on alternative ways of resolving the conflict is required. For a default theory of size n there may be of order n conflicts between defaults, and each conflict may double the number of extensions. This results in the order of 2^n extensions.

Priorities in many cases uniquely determine how a conflict between defaults is resolved, and hence the case analyses that lead to an exponential number of extensions is avoided. This gives raise to the interesting question whether priorities are a computational advantage in the sense that for interesting classes of default theories with priorities the decision problems could be solved in polynomial time.

In this chapter we first analyze the complexity of decision problems of prioritized default logic in the general case. We show that the decision problems are at most one level higher in the polynomial hierarchy than the decision problems of default logic without priorities. To investigate the idea of priorities as a computational advantage we analyze the complexity of decision problems of prioritized default logic in restricted cases. Of great interest are priorities that are strict total orders. With such priorities all conflicts are resolved in a unique way.

In Section 3.1 a number of different variants of prioritized default logic were presented. Our results are given for the consequence relation $\models_{\mathcal{P}}$ but most of them can be easily established for any variant. The most fine-grained results are the ones that concern the tractability and the intractability of restricted classes of prioritized default theories. In a small number of cases there may be differences that we have not investigated. For instance, the Horn class analyzed in Theorem 4.2.11 is tractable for $\models_{\mathcal{P}}$ but we do not know whether it is tractable for $\models^{d}_{\mathcal{P}}$. These special cases for which we do not have definite results do not however change the overall picture of intractability and tractability of prioritized default logic.

In earlier work on the complexity of nonmonotonic logics [Gottlob, 1992; Marek and Truszczyński, 1993] three kinds of decision problems have been considered: the existence of extensions, cautious reasoning, and brave reasoning. As the existence of extensions coincides with the existence of preferred extensions for finite default theories (Theorem 3.1.20), the complexity of the extension existence problems in Reiter's default logic and our prioritized default logic are the same. The consequence relations $\models_{\mathcal{P}}, \models^{2}_{\mathcal{P}}, \models^{d}_{\mathcal{P}}$ and $\models^{2d}_{\mathcal{P}}$ represent the problem of cautious reasoning with priorities: membership of a formula in all preferred extensions of a default theory. The problem of brave reasoning, the membership of a formula in at least one extension, has a natural counterpart in prioritized default logic: the membership of a formula in at least one preferred extension.

The *size* of a default theory $\langle D, W\rangle$ is the sum of the lengths of the formulae in W and in defaults in D. A default theory is finite, if it is of finite size. The language that corresponds to the consequence relation $\models$ of the classical propositional logic is denoted by CN and is defined as the set of pairs $\langle \Sigma, \phi\rangle \in 2^{\mathcal{L}} \times \mathcal{L}$ such that $\Sigma \models \phi$.

Some of the complexity results in this chapter use the procedure in Figure 4.1 that is directly based on the semiconstructive definition of extensions given in Theorem 2.1.2. The next lemma states its correctness for which we do not give a proof.

Lemma 4.0.21 *For the procedure in Figure 4.1, the call extension(D, W, E)*

$PROCEDURE$ extension(D, W, E);
$BEGIN$
  $E' := W$;
  $REPEAT$
    $E'' := E'$;
    $FOR\ EACH\ \alpha{:}\beta_1, \ldots, \beta_n/\gamma \in D\ DO$
      $IF\ \langle E', \alpha\rangle \in$ CN and $\langle E, \neg\beta\rangle \notin$ CN for all $\beta \in \{\beta_1, \ldots, \beta_n\}$
      $THEN\ E' := E' \cup \{\gamma\}$;
    END
  $UNTIL\ E' = E''$;
  $IF\ \langle E, \phi\rangle \in$ CN for all $\phi \in E'$ and $\langle E', \phi\rangle \in$ CN for all $\phi \in E$
  $THEN\ RETURN$ true $ELSE\ RETURN$ false
END

Figure 4.1: A procedure for recognizing extensions

returns true *if and only if $Cn(E)$ is an extension of the default theory $\langle D, W\rangle$. If it does not return* true, *it returns* false. *Excluding the tests of membership in CN the procedure runs in polynomial time on the size of $\langle D, W\rangle$.*

Proof: The procedure runs in polynomial time because of the following. Let n be the number of defaults in D. At each iteration of the repeat-until loop, the set E' is extended with zero or more conclusions of defaults in D. If E' is not extended, the execution of the loop terminates. Because there are n defaults, E' can be extended at most n times. Hence $n+1$ is an upper bound on the number of iterations of the loop. The number of iterations of the for-each loop inside the repeat-until loop is n. Other computation in the procedure *extension* is constant time assuming that membership tests for CN are constant time, and hence the procedure runs in quadratic time. □

4.1 Complexity in the general case

In this section we give results that characterize the complexity of prioritized default logic in the general case with unrestricted propositional logic as the base logic. First we give decision procedures for the decision problems of prioritized default logic. These decision procedures are based on nondeterministic polynomial time reductions to the language ENC, which in turn can be reduced to CN in nondeterministic polynomial time. There are three kinds of questions ENC can answer: is the logical closure of a set of formulae an extension of a default theory, is a strict total order a characteristic ordering for an extension, and is a formula a logical consequence of a set of formulae.

The language ENC$\subseteq \{0,1,2\} \times 2^{\mathcal{D}} \times 2^{\mathcal{L}} \times 2^{\mathcal{L}} \times (2^{\mathcal{D}\times\mathcal{D}} \cup \mathcal{L})$ is defined as the set of quintuples

- $\langle 0, D, W, E, \mathcal{T}\rangle$ where D is a set of defaults, $W \subseteq \mathcal{L}$ is a set of formulae, and $Cn(E)$ is an extension of $\langle D, W\rangle$,
- $\langle 1, D, W, E, \mathcal{T}\rangle$ where D is a set of defaults, $W \subseteq \mathcal{L}$ is a set of formulae, E is a set of formulae, $\mathcal{T}$ is a strict total order on D, and $\mathcal{T}$ is not a $\langle D, W\rangle, \mathcal{T}$-characteristic ordering for $Cn(E)$, and
- $\langle 2, D, W, E, \phi\rangle$ where $\langle E, \phi\rangle \in$ CN.

Lemma 4.1.1 *ENC is Turing reducible to CN in nondeterministic polynomial time.*

Proof: The language ENC is accepted by the Turing machine given as the procedure in Figure 4.2. The Turing machine uses an oracle for the language CN. The Turing machine runs in nondeterministic polynomial time for the following reasons. The nondeterministic part is the guessing of an extension $Cn(E')$ that potentially shows that $\mathcal{T}$ is not a $\langle D, W\rangle, \mathcal{T}$-characteristic ordering for $Cn(E)$. The size of the set E' is polynomial to the size of $\langle D, W\rangle$. Rest of the computation in the body of the procedure is constant time with the exception of the procedure calls.

The procedure *extension* runs in polynomial time in the size of D by Lemma 4.0.21, assuming that calls to CN take constant time. The procedure *compare* considers each element of the total order $\mathcal{T}$ one at a time and clearly runs in polynomial time. The call compare$(E, E', \mathcal{T})$ returns *true* if and only if for all $\delta \in D$, appl$(\delta, Cn(E'), Cn(E))$ implies there is $\delta'\mathcal{T}\delta$ such that appl$(\delta', Cn(E), Cn(E'))$.

The correctness of the procedure is as follows. There are three cases, one for each value of $n \in \{0,1,2\}$. If $n \notin \{0,1,2\}$ the procedure does not accept. Assume $n = 0$. Now $Cn(E)$ should be an extension of $\langle D, W\rangle$. This is by Lemma 4.0.21. Assume $n = 1$. Now $\mathcal{T}$ should not be a $\langle D, W\rangle, \mathcal{T}$-characteristic ordering for $Cn(E)$. There should be $\delta \in D$ such that appl$(\delta, E', Cn(E))$ for some extension E' of $\langle D, W\rangle$, and there is no $\delta'\mathcal{T}\delta$ such that appl$(\delta', Cn(E), E')$. The procedure guesses a set E'_b, and tests whether $E' = Cn(E'_b)$ fulfills these properties, and accepts if $\mathcal{T}$ is not a $\langle D, W\rangle, \mathcal{T}$-characteristic ordering for $Cn(E)$, and rejects otherwise. Assume $n = 2$. Now $\mathcal{T} = \phi \in \mathcal{L}$ should be a formula such that $\langle E, \phi\rangle \in$ CN. This is immediate. □

Theorem 4.1.2 *The problem of testing whether a formula belongs to at least one preferred extension of a default theory is Turing reducible to ENC in nondeterministic polynomial time.*

```
PROCEDURE ENC(n, D, W, E, T);
BEGIN
  CASE n OF
  0: IF extension(D, W, E) THEN accept ELSE reject
  1:
    guess nondeterministically a subset E' of conclusions of D;
    E' := E' ∪ W;
    IF not extension(D, W, E') or not extension(D, W, E)
      THEN reject
    ELSE
      IF compare(E, E', T) = true THEN reject
      ELSE accept
  2: IF ⟨E, T⟩ ∈ CN THEN accept ELSE reject
  otherwise: reject
END

PROCEDURE applied(α:β1, ..., βn/γ, E);
BEGIN
  IF ⟨E, α⟩ ∈ CN and ⟨E, ¬β⟩ ∉ CN for all β ∈ {β1, ..., βn}
  THEN RETURN true ELSE RETURN false
END

PROCEDURE compare(E, E', T);
let T be the ordering δ0 T δ1 T ··· T δn;
BEGIN
  FOR i := 0 TO n DO
    IF not applied(δi, E) and applied(δi, E') THEN RETURN false;
    IF applied(δi, E) and not applied(δi, E') THEN RETURN true
  END;
  RETURN true
END
```

Figure 4.2: A decision procedure for ENC

PROCEDURE brave$(D, W, \mathcal{P}, \phi)$;
BEGIN
  guess nondeterministically a subset E of conclusions of D;
  guess nondeterministically a strict total order $\mathcal{T}$ on D;
  IF $\mathcal{P} \not\subseteq \mathcal{T}$ *THEN* reject;
  IF $\langle 0, D, W, E \cup W, \emptyset \rangle \in$ ENC
    AND $\langle 1, D, W, E \cup W, \mathcal{T} \rangle \notin$ ENC
    AND $\langle 2, \emptyset, \emptyset, E \cup W, \phi \rangle \in$ ENC
  THEN accept
  ELSE reject
END

Figure 4.3: A decision procedure for prioritized brave reasoning

Proof: The question is answered by a Turing machine described by the procedure in Figure 4.3. The Turing machine uses an oracle for the language ENC. First the machine guesses a candidate extension $Cn(E \cup W)$ of $\langle D, W \rangle$ and a strict total order $\mathcal{T}$ on D that extends the relation $\mathcal{P}$. Then the machine verifies that $Cn(E \cup W)$ in fact is an extension (the first consultation of the oracle for ENC), that $\mathcal{T}$ is a $\langle D, W \rangle, \mathcal{P}$-characteristic ordering for $Cn(E \cup W)$ (the second consultation), and that $\phi \in Cn(E \cup W)$ (the third consultation.) If all these verifications succeed, the Turing machine accepts. Because the guesses of E and $\mathcal{T}$ are nondeterministic, the machine accepts whenever there is an extension that contains ϕ and has a $\langle D, W \rangle, \mathcal{P}$-characteristic ordering.

The only parts of the procedure that do not run in constant time are the nondeterministic guessing of the set E, guessing of a strict total order $\mathcal{T}$ and verifying that it extends $\mathcal{P}$. Because the sizes of E and $\mathcal{T}$ are polynomial in $\langle D, W \rangle$, the number of steps needed to guess them is polynomial in $\langle D, W \rangle$. Verification of $\mathcal{P} \subseteq \mathcal{T}$ takes polynomial time. □

Theorem 4.1.3 *The complement of the problem of testing whether a formula belongs to all preferred extensions of a default theory is Turing reducible to ENC in nondeterministic polynomial time.*

Proof: Proof is similar to that of Theorem 4.1.2. The procedure is shown in Figure 4.4. □

We give upper and lower bounds for the location of the decision problems in the polynomial hierarchy. Because ENC is Turing reducible to CN in nondeterministic polynomial time, we immediately get the following lemma.

Lemma 4.1.4 *The language ENC is in* Σ_2^p.

PROCEDURE co-cautious($D, W, \mathcal{P}, \phi$);
BEGIN
  guess nondeterministically a subset E of conclusions of D;
  guess nondeterministically a strict total order $\mathcal{T}$ on D;
  IF $\mathcal{P} \not\subseteq \mathcal{T}$ *THEN* reject;
  IF $\langle 0, D, W, E \cup W, \emptyset \rangle \in$ ENC
    AND $\langle 1, D, W, E \cup W, \mathcal{T} \rangle \notin$ ENC
    AND $\langle 2, \emptyset, \emptyset, E \cup W, \phi \rangle \notin$ ENC
  THEN accept
  ELSE reject
END

Figure 4.4: A decision procedure for prioritized cautious reasoning

By Theorems 4.1.2 and 4.1.3 we get upper bounds for the location of the prioritized decision problems in the polynomial hierarchy. These results are immediate by the definition of the polynomial hierarchy.

Theorem 4.1.5 *The problem of whether a formula belongs to at least one preferred extension of a default theory is in Σ_3^p, and the problem of whether a formula belongs to all preferred extensions of a default theory is in Π_3^p.*

Same upper bounds can be obtained for the decision problems for preferred$_2$ extensions. Note that the definition of preferredness$_2$ in Definition 3.1.5 easily leads to a result indicating the membership of the problems in Π_4^p and Σ_4^p, respectively. Theorem 3.1.14 demonstrates how the reduction of lexicographic comparison with partial orders to lexicographic comparison with total orders is avoided. Using it the membership in Π_3^p and Σ_3^p can be obtained easily.

Next we show that the decision problems are Δ_3^p-hard, thereby obtaining a lower bound on the location of the problems in the polynomial hierarchy. This result is based on the Δ_3^p-hard problem in the following theorem is polynomial time many-one reducible to the decision problems of prioritized default logic.

Theorem 4.1.6 ([Krentel, 1992]) *For any $k \geq 1$, the problem of computing a given bit of X^1 is Δ_{k+1}^p-complete, where X^1 is defined as follows.*

MAXIMUM k-QUANTIFIED FORMULA
Instance: *Boolean formula $C[X^1, X^2, \ldots, X^k]$ where X^i is an abbreviation for an n-tuple of Boolean variables that corresponds to a valuation of $x_1^i, \ldots, x_n^i$.*
Output: *Maximum $X^1 \in \{0,1\}^n$ that satisfies*

$$\exists X^1 \forall X^2 \exists X^3 \ldots QX^k (C[X^1, \ldots, X^k] = 1)$$

where Q is either $\forall$ or $\exists$ depending on the parity of k.

We state this in a slightly simpler form for $k = 2$.

Theorem 4.1.7 *The problem of computing a given bit of X^1 is Δ_3^p-complete, where X^1 is defined as follows.*

> MAXIMUM 2-QUANTIFIED FORMULA
> Instance: *Boolean formula $C[X^1, X^2]$ where X^i is an abbreviation for an n-tuple of Boolean variables that corresponds to a valuation of $x_1^i, \ldots, x_n^i$.*
> Output: *Maximum $X^1 \in \{0,1\}^n$ that satisfies*
> $$\exists X^1 \forall X^2 (C[X^1, X^2] = 1).$$

The following lemma points out a connection between the validity of quantified Boolean formulae with a universal quantifier as the outermost quantifier and the logical consequence relation.

Lemma 4.1.8 *Let X be a set that contains exactly one of x and $\neg x$ for every $x \in \{x_1^1, \ldots, x_n^1\}$. Let $X^1 = \langle x_1'^1, \ldots, x_n'^1\rangle$ where $x_i'^1 = 1$ if $x_i^1 \in X$ and $x_i'^1 = 0$ otherwise. Then $\forall X^2(C[X^1, X^2])$ if and only if $X \models C[X^1, X^2]$.*

Proof: Assume X^1 satisfies $\forall X^2(C[X^1, X^2])$. Hence for the choice of truth values for $x_1^1, \ldots, x_n^1$ indicated by X^1 and any choice of truth values for $x_1^2, \ldots, x_n^2$, $C[X^1, X^2]$ is true. Now let M be any propositional model such that $M \models X$. Hence M assigns the same truth values to $x_1^1, \ldots, x_n^1$ as X^1. Now $M \models C[X^1, X^2]$, and by the definition of logical consequence $X \models C[X^1, X^2]$.

Assume $X \models C[X^1, X^2]$. Hence for all models M such that $M \models X$ and M assigns any truth-values to $x_1^2, \ldots, x_n^2$, $M \models C[X^1, X^2]$. By definition of quantified Boolean formulae, for X^1 and any X^2 the formula $C[X^1, X^2]$ is true, that is, X^1 satisfies $\forall X^2(C[X^1, X^2])$. □

The proof of the next lemma is an extension of a proof by Gottlob [1992] that demonstrates Π_2^p-hardness of cautious reasoning in default logic without priorities.

Lemma 4.1.9 *The problem of computing a given bit of X^1 in* MAXIMUM 2-QUANTIFIED FORMULA *is polynomial time many-one reducible to the problem of testing whether a formula belongs to all preferred extensions of a default theory.*

Proof: We construct a default theory $\Delta = \langle D, \emptyset\rangle$ and define a strict partial order $\mathcal{P}$ on D so that the kth bit of X^1 is 1 if and only if $\Delta \models_{\mathcal{P}} x_k^1$. The value of a given bit of X^1 in Maximum 2-quantified formula is defined only if $\exists X^1 \forall X^2 (C[X^1, X^2])$ is satisfiable, so assume it is. Let a be a propositional variable that is not in $\{x_1^1, \ldots, x_n^1, x_1^2, \ldots, x_n^2\}$. Define

$$
\begin{aligned}
D &= \left\{ \frac{: x_1^1}{x_1^1}, \frac{: \neg x_1^1}{\neg x_1^1}, \ldots, \frac{: x_n^1}{x_n^1}, \frac{: \neg x_n^1}{\neg x_n^1}, \frac{C[X^1, X^2] : a}{a} \right\}, \\
\mathcal{P} &= \left\{ \frac{C[X^1, X^2] : a}{a} \right\} \times \left(D \backslash \left\{ \frac{C[X^1, X^2] : a}{a} \right\} \right) \\
&\cup \left\{ \left\langle \frac{: x_i^1}{x_i^1}, \frac{: x_j^1}{x_j^1} \right\rangle \middle| i \in \{1, \ldots, n-1\}, j \in \{i+1, \ldots, n\} \right\} \\
&\cup \left\{ \frac{: x_1^1}{x_1^1}, \ldots, \frac{: x_n^1}{x_n^1} \right\} \times \left\{ \frac{: \neg x_1^1}{\neg x_1^1}, \ldots, \frac{: \neg x_n^1}{\neg x_n^1} \right\}.
\end{aligned}
$$

The sizes of the default theory $\langle D, \emptyset \rangle$ and the strict partial order $\mathcal{P}$ are linearly proportional to the size of $C[X^1, X^2]$, and they can be constructed in polynomial time.

First we show that for each extension E of $\langle D, \emptyset \rangle$ that contains a there is X^1 such that X^1 satisfies $\forall X^2 (C[X^1, X^2])$ and for all $i \in \{1, \ldots, n\}$, $p_i(X^1) = 1$ iff $x_i^1 \in E$ (the function p_i selects the ith element of an n-tuple.) Let E be an extension of $\langle D, \emptyset \rangle$ and $a \in E$. Because defaults in D are normal, E is consistent. Hence appl($C[X^1, X^2]$:$a/a, E$) because the only default where a occurs in the conclusion is $C[X^1, X^2]$:a/a. Hence $C[X^1, X^2] \in E$, that is, $C[X^1, X^2]$ is a logical consequence of the conclusions of generating defaults of E. Because a does not occur in $C[X^1, X^2]$, $X \models C[X^1, X^2]$ for $X = E \cap \{x_1^1, \neg x_1^1, \ldots, x_n^1, \neg x_n^1\}$. By Lemma 4.1.8 X^1, defined as $p_i(X^1) = 1$ if $x_i^1 \in X$ and $p_i(X^1) = 0$ otherwise, satisfies $\forall X^2 (C[X^1, X^2])$.

Then we show that for each X^1 that satisfies $\forall X^2 (C[X^1, X^2])$, there is an extension E of $\langle D, \emptyset \rangle$ such that $a \in E$ and $p_i(X^1) = 1$ iff $x_i^1 \in E$. Assume that X^1 satisfies $\forall X^2 (C[X^1, X^2])$. Let $X = \{x_i^1 | p_i(X^1) = 1, 1 \leq i \leq n\} \cup \{\neg x_i^1 | p_i(X^1) = 0, 1 \leq i \leq n\}$. We use Theorem 2.1.3 to show that $E = Cn(X \cup \{a\})$ is an extension of $\langle D, \emptyset \rangle$.

$$
\begin{aligned}
D^E &= \left\{ \frac{\alpha : \beta}{\beta} \in D \middle| \neg \beta \notin E \right\} \\
&= \left\{ \frac{: x_i^1}{x_i^1} \in D \middle| \neg x_i^1 \notin E \right\} \cup \left\{ \frac{: \neg x_i^1}{\neg x_i^1} \in D \middle| \neg\neg x_i^1 \notin E \right\} \\
&\quad \cup \left\{ \frac{C[X^1, X^2] : a}{a} \middle| \neg a \notin E \right\} \\
&= \left\{ \frac{: x}{x} \in D \middle| x \in X \right\} \cup \left\{ \frac{C[X^1, X^2] : a}{a} \right\}
\end{aligned}
$$

The equations in Theorem 2.1.3 give the following.

$$C_0 = \emptyset$$

$$\begin{aligned}
C_1 &= C_0 \cup \left\{ \beta \,\middle|\, \frac{\alpha : \beta}{\beta} \in D^E, C_0 \models \alpha \right\} \\
&= X \cup \{a | C_0 \models C[X^1, X^2]\} \\
C_2 &= X \cup \{a | C_0 \models C[X^1, X^2]\} \cup X \cup \{a | C_1 \models C[X^1, X^2]\} \\
&= X \cup \{a\} \\
C_i &= C_2 \text{ for all } i \geq 3
\end{aligned}$$

The simplification of C_2 is because $X \subseteq C_1$ and by Lemma 4.1.8 $X \models C[X^1, X^2]$. Now $E = \bigcup_{i \geq 0} Cn(C_i) = Cn(X \cup \{a\})$. Because the extension E is consistent, appl$(C[X^1, X^2]{:}a/a, E)$. Because $C[X^1, X^2]{:}a/a$ is the $\mathcal{P}$-least default in D, all $\mathcal{P}$-preferred extensions apply $C[X^1, X^2]{:}a/a$.

The counterpart of the $\mathcal{T}$-preferred extension is the lexicographically maximal X^1, and vice versa. This is by the construction of $\mathcal{T}$, where the defaults ${:}x_1^1/x_1^1, \ldots, {:}x_n^1/x_n^1$ are ordered as ${:}x_i^1/x_i^1 \mathcal{T} {:}x_j^1/x_j^1$ whenever $i < j$. Therefore $p_i(X^1) = 1$ for the maximum X^1 satisfying $\forall X^2(C[X^1, X^2])$ if and only if x_i^1 belongs to the $\mathcal{T}$-preferred extension of $\langle D, \emptyset \rangle$. □

Theorem 4.1.10 *For default theories $\Delta = \langle D, W \rangle$, strict partial orders $\mathcal{P}$ on D, and formulae ϕ, the problem of testing $\Delta \models_{\mathcal{P}} \phi$ is Δ_3^p-hard.*

Proof: All problems in Δ_3^p are polynomial time many-one reducible to MAXIMUM 2-QUANTIFIED FORMULA, which in turn is polynomial time many-one reducible to our problem in polynomial time as shown by Lemma 4.1.9. □

In the special case where priorities are strict total orders the decision problems of prioritized default logic are in Δ_3^p. The proof of this fact is based on an idea for computing the unique preferred extensions in a sequence of steps, at each step determining whether there is an extension in which one of the defaults is applied.

Lemma 4.1.11 *For the procedure in Figure 4.5, the call exists(D, W, A) returns true if and only if there is an extension E of $\langle D, W \rangle$ such that appl(δ, E) for all $\delta \in A$. The procedure runs in nondeterministic polynomial time given an NP oracle for CN.*

Proof: First the procedure guesses a set of conclusions G, that possibly are the generating defaults of the possible extension $Cn(G \cup W)$. The call to extension(D, W, E) with $E = G \cup W$ tests if this is indeed the case (Lemma 4.0.21.) Next the procedure tests whether all defaults in A are applied in $Cn(G \cup W)$. The procedure returns true if and only if all members of A are applied in $Cn(G \cup W)$. Because the procedure represents a nondeterministic Turing machine, the procedure returns true if and only if it is possible to

```
PROCEDURE exists(D,W,A);
BEGIN
  guess a subset G of conclusions of defaults in D;
  E := G ∪ W;
  IF not extension(D, W, E) THEN RETURN false;
  FOR EACH α:β_1, ..., β_n/γ ∈ A DO
    IF ⟨E, α⟩ ∉ CN or ⟨E, ¬β⟩ ∈ CN for some β ∈ {β_1, ..., β_n}
    THEN RETURN false
  END;
  RETURN true
END
```

Figure 4.5: A procedure that tests for the existence of extensions

guess G so that $Cn(G \cup W)$ is an extension where members of A is applied. Hence *exists*(D, W, A) returns true if and only if there is an extension E of $\langle D, W\rangle$ such that $\mathrm{appl}(\delta, E)$ for all $\delta \in A$.

The procedure runs in non-deterministic polynomial time and uses an NP oracle for propositional satisfiability, hence the problem solved by it is in Σ_2^p. □

Theorem 4.1.12 *For default theories $\Delta = \langle D, W\rangle$, formulae ϕ, and strict total orders $\mathcal{P}$ on D, the problem of testing $\Delta \models_{\mathcal{P}} \phi$ is in Δ_3^p.*

Proof: We give a decision procedure for the problem that runs in deterministic polynomial time and uses an oracle for a problem that belongs to Σ_2^p. This demonstrates that the problem is in Δ_3^p. The procedure *decide* is given in Figure 4.6 and the Σ_2^p oracle procedure *exists* is given in Figure 4.5 and its correctness is stated in Lemma 4.1.11. The correctness of the procedure *decide* is as follows. Let $\langle D, W\rangle$ be a finite default theory with $|D| = n$, $\mathcal{T}$ a strict total order on D, and ϕ a formula. Assume $\langle D, W\rangle$ has no extensions. In this case the first statement in the procedure *decide* returns true which is correct because ϕ trivially belongs to all extensions of $\langle D, W\rangle$. Assume $\langle D, W\rangle$ has at least one extension. The correctness proof in this case proceeds by induction on $i \in \{1, \ldots, n\}$. The value of the program variable A after the ith iteration is denoted by A_i. Let $\delta_1, \ldots, \delta_n$ be the ordering $\mathcal{T}$. Because $\mathcal{T}$ is a strict total order on D, by Lemma 3.1.18 there is exactly one $\mathcal{T}$-preferred extension E of $\langle D, W\rangle$.

Induction hypothesis: for all $j \in \{1, \ldots, i\}$, $\mathrm{appl}(\delta_j, E)$ if and only if $\delta_j \in A_i$.

Base case $i = 0$: Now $A_0 = \emptyset$. There are no $j \in \{1, \ldots, i\}$ and hence the hypothesis is true.

PROCEDURE decide(D,W,$\mathcal{T}$,ϕ);
BEGIN
  IF not exists($D, W, \emptyset$) *THEN RETURN* true;
  $A := \emptyset$;
  let $\delta_1, \ldots, \delta_n$ be the ordering $\mathcal{T}$;
  FOR $i := 1$ *TO* n *DO*
    IF exists($D, W, A \cup \{\delta_i\}$) *THEN* $A := A \cup \{\delta_i\}$
  END;
  $E := W \cup \{\gamma | \alpha{:}\sigma/\gamma \in A\}$;
  IF $E \models \phi$ *THEN RETURN* true *ELSE RETURN* false
END

Figure 4.6: A decision procedure for total priorities

Inductive case $i \geq 1$: We analyze the value b returned by exists($D, W, A_{i-1} \cup \{\delta_i\}$) by cases. Assume $b =$ true. Hence $A_i = A_{i-1} \cup \{\delta_i\}$. By the correctness of the procedure *exists*, there is an extension E' such that appl(δ, E') for all $\delta \in A_{i-1} \cup \{\delta_i\}$. Assume E is not such an extension. Hence appl(δ_i, E', E). By induction hypothesis appl(δ_j, E) iff $\delta_j \in A_{i-1}$ for all $j \in \{1, \ldots, i-1\}$. Now because appl(δ, E) implies appl(δ, E') for all $\delta \mathcal{T} \delta_i$ there are no $\delta'' \mathcal{T} \delta_i$ such that appl(δ'', E, E'). Hence E could not be $\mathcal{T}$-preferred, and the assumption that not appl(δ_i, E) was wrong. Assume $b =$ false. As there is no extension E' such that for all $\delta \in A_{i-1} \cup \{\delta_i\}$ appl(δ, E'), and appl(δ, E) for all $\delta \in A_{i-1}$, not appl(δ_i, E). Now $A_i = A_{i-1}$, and the induction hypothesis is fulfilled. This finishes the induction proof.

Now A_n is the set of generating defaults of the unique $\mathcal{T}$-preferred extension E of $\langle D, W \rangle$. The execution of the program continues at the last statement of the procedure. By Theorem 2.1.5, $E = Cn(W \cup \{\gamma | \alpha{:}\sigma/\gamma \in A_n\})$. Hence $\phi \in E$ if and only if $W \cup \{\gamma | \alpha{:}\sigma/\gamma \in A_n\} \models \phi$. □

Corollary 4.1.13 *For default theories* $\Delta = \langle D, W \rangle$*, strict total orders* $\mathcal{P}$ *on* D*, and formulae* ϕ*, the problem of testing* $\Delta \models_{\mathcal{P}} \phi$ *is* Δ_3^p*-complete.*

Proof: Directly by Theorems 4.1.12 and 4.1.10. □

For the problems of membership of formulae in some or all preferred extensions of a default theory without the restriction to total priorities, the membership on the second level of the polynomial hierarchy Δ_3^p as well as their Σ_3^p-hardness and Π_3^p-hardness respectively remain open.

4.2 Complexity in restricted cases

In this section we investigate the complexity of prioritized default logic in syntactically restricted cases. This research problem is closely related to the research on the boundary between tractability and intractability of default logic without priorities, most notably by Kautz and Selman [1991] and Stillman [1990]. We give complexity results for prioritized default logic for classes of default theories that correspond to the ones investigated by them.

The decision problems of nonmonotonic logics are computationally very complex also in cases where reasoning in the classical logic on which default logic is based on can be performed efficiently. This is demonstrated in general by results of Gottlob [1992] and Stillman [1992] that show that even if the calls to a theorem-prover for classical propositional logic in a decision procedure for default logic did not consume any resources, the decision procedure would not run in polynomial time (assuming that P≠NP.) Therefore it is not surprising that for many simple classes of default theories reasoning is intractable. For example decision problems of default logic with Horn clauses and default rules of very restricted forms are NP-hard [Stillman, 1990]. The complexity in these simple cases seems to be due to the conflicts between defaults, and conflicts between defaults are the reason why a default theory has more than one extension. Priorities often determine how conflicts are resolved, thereby reducing the number of relevant extensions. Hence the question arises what impact do priorities have on complexity of nonmonotonic reasoning.

4.2.1 Restrictions investigated by Kautz, Selman and Stillman

Kautz and Selman [1991] analyze default theories where the objective parts are sets of literals and the defaults are of the forms shown below. A default theory is *ordered* [Etherington, 1987], if its dependency graph does not contain cycles. Ordered default theories, like normal default theories and unlike seminormal default theories, have at least one extension. The letters a, b, c – possibly subscripted – denote literals and the letters p, q, r denote propositional variables.

Disjunction free	$\dfrac{a_1 \wedge \cdots \wedge a_l : b_1 \wedge \cdots \wedge b_n \wedge c_1 \wedge \cdots \wedge c_m}{b_1 \wedge \cdots \wedge b_n}$
Unary	$\dfrac{p : q}{q} \quad \dfrac{p : q \wedge \neg r}{q} \quad \dfrac{p : \neg q}{\neg q}$
Disjunction free ordered	$\dfrac{a_1 \wedge \cdots \wedge a_l : b_1 \wedge \cdots \wedge b_n \wedge c_1 \wedge \cdots \wedge c_m}{b_1 \wedge \cdots \wedge b_n}$
Ordered unary	$\dfrac{p : q}{q} \quad \dfrac{p : q \wedge \neg r}{q} \quad \dfrac{p : \neg q}{\neg q}$
Disjunction free normal	$\dfrac{a_1 \wedge \cdots \wedge a_l : b_1 \wedge \cdots \wedge b_n}{b_1 \wedge \cdots \wedge b_n}$
Horn	$\dfrac{p_1 \wedge \cdots \wedge p_n : q}{q} \quad \dfrac{p_1 \wedge \cdots \wedge p_n : \neg q}{\neg q}$
Normal unary	$\dfrac{p : q}{q} \quad \dfrac{p : \neg q}{\neg q}$

Stillman [1990] extends the analysis to rules without prerequisites.

Prerequisite free	$\dfrac{: b_1 \wedge \cdots \wedge b_n \wedge c_1 \wedge \cdots \wedge c_m}{b_1 \wedge \cdots \wedge b_n}$
Prerequisite free ordered	$\dfrac{: b_1 \wedge \cdots \wedge b_n \wedge c_1 \wedge \cdots \wedge c_m}{b_1 \wedge \cdots \wedge b_n}$
Prerequisite free unary	$\dfrac{: q \wedge \neg r}{q} \quad \dfrac{: \neg q}{\neg q} \quad \dfrac{: q}{q}$
Prerequisite free ordered unary	$\dfrac{: q \wedge \neg r}{q} \quad \dfrac{: \neg q}{\neg q} \quad \dfrac{: q}{q}$
Prerequisite free normal	$\dfrac{: b_1 \wedge \cdots \wedge b_n}{b_1 \wedge \cdots \wedge b_n}$
Prerequisite free normal unary	$\dfrac{: q}{q} \quad \dfrac{: \neg q}{\neg q}$
Prerequisite free positive normal unary	$\dfrac{: q}{q}$

The inclusion relation between the above classes is depicted in Figure 4.7. Stillman's [1990] definition of prerequisite-free unary and prerequisite-free ordered unary classes does not include normal rules with positive conclusions. We, however, include these rules in our definition in order to have a closer correspondence with the Kautz and Selman definition of unary classes. This change does not sacrifice generality, as the default $p{:}q/q$ corresponds to a default $p{:}q \wedge \neg r/q$ whenever r does not occur elsewhere in the default theory.

Kautz and Selman [1991] analyze the complexity of determining existence of extensions, cautious reasoning, and brave reasoning for default theories with defaults of the above form (excluding the prerequisite-free defaults) and objective parts that are sets of literals. Stillman [1990] considers only the problem of brave reasoning, but considers a larger range of classes. He analyzes classes

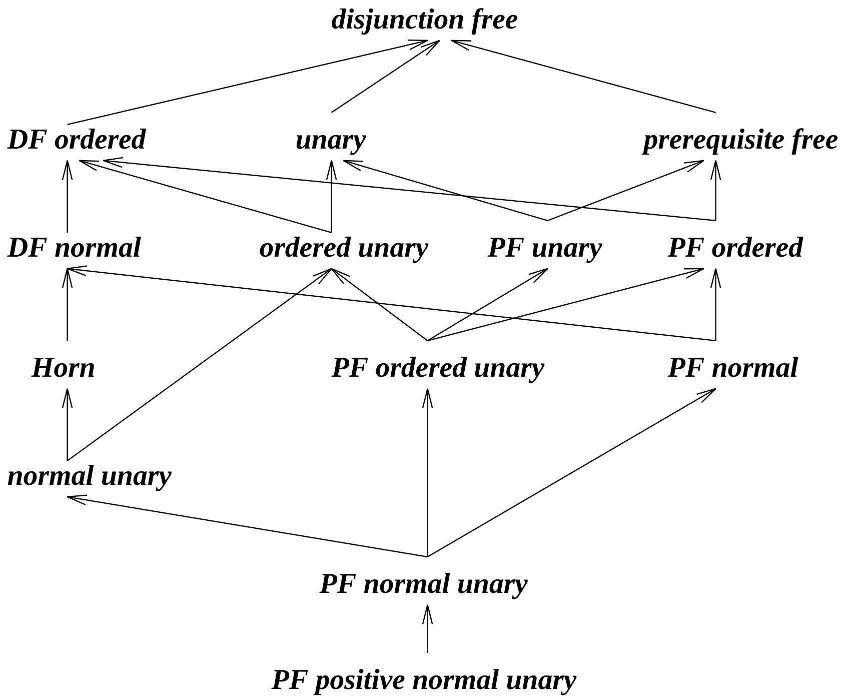


Figure 4.7: A hierarchy of classes of default theories

of default theories where objective parts are sets of Horn clauses or sets of two literal clauses, and also the case where defaults are prerequisite-free. According to Kautz and Selman, and Stillman, the most general tractable classes in default logic without priorities are the following.

- If the objective part W may contain Horn clauses, none of the restrictions on the form of default rules is sufficient for achieving tractability in brave reasoning [Stillman, 1990]. Cautious reasoning has not been investigated.

- If W consists of 2-literal clauses, brave reasoning for the prerequisite-free normal class is tractable [Stillman, 1990]. Cautious reasoning has not been investigated.

- If W consists of literals, brave reasoning for the prerequisite-free normal class and the Horn class are tractable [Kautz and Selman, 1991; Stillman, 1990]. Cautious reasoning is tractable for the normal unary class [Kautz and Selman, 1991]. Cautious reasoning for the prerequisite-free classes has not been investigated.

We present a similar analysis for the same classes of prioritized default logic. Because default logic without priorities is a special case of prioritized default logic by Lemma 3.1.16, prioritized default logic with no restriction on priorities is at least as complex as default logic without priorities. Hence there is no hope of having tractability for prioritized default logic with no restrictions on priorities in cases where default logic without priorities is intractable. Therefore we restrict to the case where priorities are a total ordering on the defaults, because in this case priorities uniquely determine how conflicts between defaults are resolved. We present for each class either an NP-hardness result or a polynomial time decision procedure. Most of the NP-hardness results are reductions from the intractable classes of default reasoning without priorities: we show that the problems of brave reasoning is in many cases polynomial time many-one reducible to the decision problems of prioritized default logic. We consider only the problem of cautious reasoning. Because for total priorities there is at most one extension, and because for most of the classes (excluding the non-ordered seminormal ones) there is at least one extension, brave and cautious reasoning often coincide. Because in these cases cautious reasoning and brave reasoning coincide, a complexity result for cautious reasoning implies the same result for brave reasoning.

4.2.2 Restricted intractable classes

A property of our definition of preferred extensions is that whenever there are finitely many defaults, the existence of preferred extensions is equivalent to the existence of extensions in general (Theorem 3.1.20.) Hence, the problem of the existence of preferred extensions is computationally as complex as the problem of existence of extensions.

Theorem 4.2.1 *The problem of existence of preferred extensions of default theories with defaults of the form* $:\neg p/q$ *(simple* K_1*-theories) is NP-complete.*

Proof: In [Marek and Truszczyński, 1991] the class of simple K_1-theories of autoepistemic logic is defined. The formulae in theories of this class are of the form $\neg Lp \rightarrow q$ where p and q are propositional variables. Marek and Truszczyński show that the problem of existence of stable expansions of simple K_1-theories is NP-complete. Their proof is based on the NP-completeness of the problem of determining whether a graph has a kernel [Garey and Johnson, 1979]. We adapt this proof to default logic and give more details. Extensions of default theories with defaults of the form $:\neg p/q$ correspond to vertex covers with an additional property, just like stable expansions of autoepistemic theories with formulae of the form $\neg Lp \rightarrow q$. The complements of these vertex covers are kernels.

A *vertex cover* of a directed graph $G = \langle V, E\rangle$ is a set $H \subseteq V$ such that for all $\langle p, q\rangle \in V$, $H \cap \{p, q\} \neq \emptyset$. A *kernel* of G is a set $V' \subseteq V$ such that $\langle p, q\rangle \in E$

for no $\{p,q\} \subseteq V'$ and for all $v \in V\backslash V'$ there is $u \in V'$ such that $\langle u,v\rangle \in E$. We show that for a given directed graph $G = \langle V,E\rangle$, $K \subseteq V$ is a kernel of G if and only if $H = V\backslash K$ is a vertex cover of G such that for every $q \in H$ there is $p \in V\backslash H$ such that $\langle p,q\rangle \in E$ if and only if $Cn(H)$ is an extension of Δ, where $\Delta = \langle D,\emptyset\rangle$ and $D = \{:\neg q/p | \langle q,p\rangle \in E\}$. The first equivalence is by definitions of a kernel and a vertex cover. We prove the second equivalence.

Assume H is a vertex cover for $G = \langle V,E\rangle$ such that for every $v \in H$ there is $q \in V\backslash H$ such that $\langle q,v\rangle \in E$. First we show that $H \subseteq \{q|\langle p,q\rangle \in E, p \notin H\}$. Assume $q \in H$. Therefore there is $p \in V\backslash H$ such that $\langle p,q\rangle \in E$. Hence $q \in \{q|\langle p,q\rangle \in E, p \notin H\}$. Then we show that $\{q|\langle p,q\rangle \in E, p \notin H\} \subseteq H$. Assume $q \in \{q|\langle p,q\rangle \in E, p \notin H\}$. Hence there is $\langle p,q\rangle \in E$ and $p \notin H$. Because H is a vertex cover and $\langle p,q\rangle \in E$, it has to be the case that $q \in H$. Hence $H = \{q|\langle p,q\rangle \in E, p \notin H\}$. By definition of D, this is equivalently $H = \{q|:\neg p/q \in D, p \notin H\}$. Because H is a set of propositional variables, this is equivalently $Cn(H) = Cn(\{q|:\neg p/q \in D, p \notin Cn(H)\})$. By Theorem 2.1.8 now $Cn(H)$ is an extension of $\langle D,\emptyset\rangle$.

Let $Cn(H)$ be an extension of $\langle D,\emptyset\rangle$ and H the set of conclusions of the generating defaults of $Cn(H)$. Hence $Cn(H) = Cn(\{q|:\neg p/q \in D, p \notin Cn(H)\})$. By the same simplifications that were used in the first part of this proof, $H = \{q|\langle p,q\rangle \in E, p \notin H\}$. Now H is a vertex cover, because if $\langle p,q\rangle \in E$ and $p \notin H$, then by the above equation $q \in E$, and hence it is not possible that $\langle p,q\rangle \in E$ and $\{p,q\} \cap H = \emptyset$. Because $q \in H$ if and only if $\langle p,q\rangle \in E$ for some $p \in V$ and $p \notin H$, for all $q \in H$ there is $p \in V\backslash H$ such that $\langle p,q\rangle \in E$. □

Corollary 4.2.2 *Prioritized cautious reasoning for simple K_1-theories is co-NP-hard.*

Proof: By translating from the complement of the NP-complete problem in Theorem 4.2.1. The extensions of simple K_1-theories are consistent because the set of conclusions of defaults is consistent. Let p be a propositional variable that does not occur in the simple K_1-theory in question. The complement of the NP-complete existence question is polynomial time many-one reducible to the problem of whether all preferred extensions contain p. Specifically, for any strict partial order $\mathcal{P}$ on the defaults of a simple K_1 default theory Δ, there are extensions of Δ if and only if p does not belong to all $\mathcal{P}$-preferred extensions of Δ. Hence the problem is co-NP-hard. □

The proofs of the above theorem and corollary do not depend on the properties of priorities. Hence tractability cannot be achieved merely by restricting to some subclass of prioritizations, for example priorities that totally order all defaults, and restrictions on the form of defaults are necessary as well. The fact that the intractability result above is based on a translation to the extension existence problem directly hints at possible tractable classes: tractable subclasses of prioritized default logic are based on subclasses of default theories for which the existence of extensions can be tested in polynomial time. Examples of such possibly tractable classes are normal and seminormal ordered theories of default logic based on a tractable class of a classical logic, like those analyzed by Kautz and Selman [1991] and Stillman [1990].

Kautz and Selman and Stillman identify many intractable classes of default logic without priorities. Many of these intractability results can be used in finding intractable subclasses of prioritized default logic. The next theorems are based on translations from intractable problems of brave reasoning.

Theorem 4.2.3 *Let $\mathcal{F}$ be a class of formulae and $\mathcal{C}$ a class of finite default theories such that*

- *if $\langle D, W\rangle \in \mathcal{C}$, then $\langle D \cup \{\phi{:}p/p\}, W\rangle \in \mathcal{C}$ where $\phi \in \mathcal{F}$ and p is a propositional variable that does not occur in D or W, and*
- *each member of $\mathcal{C}$ has at least one extension.*

The problem of testing $\Delta \models_b \phi$ for $\Delta \in \mathcal{C}$ and $\phi \in \mathcal{F}$ is polynomial time many-one reducible to the problem of testing $\Delta' \models_{\mathcal{P}} \phi'$ where $\Delta' \in \mathcal{C}$, $\phi' \in \mathcal{F}$ and $\mathcal{P}$ is a strict total order on defaults in Δ'.

Proof: Let $\langle D, W\rangle \in \mathcal{C}$ be a default theory, $\phi \in \mathcal{F}$ a formula, and k a propositional variable that does not occur in $\langle D, W\rangle$ or ϕ. Let $\mathcal{P}$ be any strict partial order on $D \cup \{\phi{:}k/k\}$ such that $\phi{:}k/k$ is the $\mathcal{P}$-least element. We claim that $\langle D, W\rangle \models_b \phi$ if and only if $\langle D \cup \{\phi{:}k/k\}, W\rangle \models_{\mathcal{P}} k$. The reduction is clearly polynomial time.

Let $\langle D, W\rangle \models_b \phi$, that is, there is at least one extension E of $\langle D, W\rangle$ that contains ϕ. If E is inconsistent, then by Lemma A.1.9 E is an extension of $\langle D \cup \{\phi{:}k/k\}, W\rangle$, and hence k is contained in all $\mathcal{P}$-preferred extensions of $\langle D \cup \{\phi{:}k/k\}, W\rangle$ as E is the only such extension by Lemma A.1.2. Otherwise E is consistent, $\neg k \notin E$ and by Lemma A.1.8 $E' = Cn(E \cup \{k\})$ is an extension of $\langle D \cup \{\phi{:}k/k\}, W\rangle$. Let $\mathcal{P}$ be any strict total order on $D \cup \{\phi{:}k/k\}$ such that $\phi{:}k/k\mathcal{P}\delta$ for all $\delta \in D$. Assume there is a $\mathcal{P}$-preferred extension E'' of $\langle D \cup \{\phi{:}k/k\}, W\rangle$ such that $k \notin E''$. Now $\mathrm{appl}(\phi{:}k/k, E', E'')$. By definition of $\mathcal{P}$-preferredness, there is a strict total order $\mathcal{T}$ on $D \cup \{\phi{:}k/k\}$ such that $\mathcal{P} \subseteq \mathcal{T}$ and for all extensions E' of Δ and all $\delta \in D$, if $\mathrm{appl}(\delta, E', E'')$ then there is $\delta'\mathcal{T}\delta$ such that $\mathrm{appl}(\delta', E'', E')$. Now $\mathrm{appl}(\phi{:}k/k, E', E'')$, but as $\phi{:}k/k\mathcal{P}\delta$ for all $\delta \in D$, there is no $\delta'\mathcal{T}\phi{:}k/k$ such that $\mathrm{appl}(\delta', E'', E')$. Therefore E'' could not be $\mathcal{P}$-preferred, which contradicts our assumption. Therefore all $\mathcal{P}$-preferred extensions of Δ contain k.

Assume $\langle D \cup \{\phi{:}k/k\}, W\rangle \models_{\mathcal{P}} k$, that is, all $\mathcal{P}$-preferred extensions contain k. By the assumptions of the lemma, there is at least one extension of $\langle D \cup \{\phi{:}k/k\}, W\rangle$, and by Theorem 3.1.20 there is at least one $\mathcal{P}$-preferred extension E of $\langle D \cup \{\phi{:}k/k\}, W\rangle$. If E is inconsistent, then by Lemma A.1.2 E is the only extension. Now $\neg k \in E$, and hence by Lemma A.1.9 E is an extension of $\langle D, W\rangle$. Because $\phi \in E$, $\langle D, W\rangle \models_b \phi$. If E is consistent, then appl$(\phi{:}k/k, E)$ because k does not occur in D or W. Hence $\phi \in E$. Finally, $E' = Cn(E \cap (W \cup \{\gamma | \alpha{:}\sigma/\gamma \in D\}))$ is an extension of $\langle D, W\rangle$ by Lemma A.1.12. Because k does not occur in ϕ, $\phi \in E'$. Therefore $\langle D, W\rangle \models_b \phi$. □

Theorem 4.2.3 may be applied to all classes in Figure 4.7 except for the prerequisite-free ones. The translations are from the intractable classes mentioned in the end of Section 4.2.1. Hence we obtain the following intractability result.

Corollary 4.2.4 *Determining $\langle D, W\rangle \models_{\mathcal{P}} \phi$ for strict total orders $\mathcal{P}$ in any of the classes of default theories below is NP-hard.*

D	W	ϕ
normal unary	*Horn clauses*	*literal*
normal unary	*2-literal clauses*	*literal*
DF normal	*literals*	*literal*
ordered unary	*literals*	*literal*

The translation from $\models_b$ to $\models^d_{\mathcal{P}}$ is more complicated because the membership of a formula ϕ in an extension is translated to the non-defeat of a default, that is, the absence of the prerequisite or the negation of a justification.

Theorem 4.2.5 *Let $\mathcal{F}$ be a class of formulae and $\mathcal{C}$ a class of finite default theories such that*

- *if $\langle D, W\rangle \in \mathcal{C}$, then $\langle D \cup \{\phi{:}l/l\}, W\rangle \in \mathcal{C}$ where l is a literal and $\phi \in \mathcal{F}$ or ϕ is a propositional variable,*
- *if $\langle D, W\rangle \in \mathcal{C}$, then $\langle D, W \cup \{p\}\rangle \in \mathcal{C}$ where p is a propositional variable, and*
- *each member of $\mathcal{C}$ has at least one extension.*

The problem of testing $\Delta \models_b \phi$ for $\Delta \in \mathcal{C}$ and $\phi \in \mathcal{F}$ is polynomial time many-one reducible to the problem of testing $\Delta' \models^d_{\mathcal{P}} \phi'$ where $\Delta' \in \mathcal{C}$, $\phi' \in \mathcal{F}$ and $\mathcal{P}$ is a strict total order on defaults in Δ'.

Proof: Let $\Delta = \langle D, W\rangle \in \mathcal{C}$ be a default theory and $\phi \in \mathcal{F}$ a formula. Let c and k be propositional variables that do no occur in D or W or ϕ. Let

$D' = D \cup \{k{:}\neg c/\neg c, c{:}k/k, \phi{:}\neg k/\neg k\}$ and $\Delta' = \langle D', W \cup \{c\}\rangle$. Clearly $\Delta' \in \mathcal{C}$. Let $\mathcal{P}$ be any strict partial order on D' such that $k{:}\neg c/\neg c$ is the $\mathcal{P}$-least element of D'. We claim that $\Delta \models_b \phi$ if and only if $\Delta' \models^d_{\mathcal{P}} \neg k$.

Assume $\Delta \models_b \phi$, that is, there is an extension E of Δ such that $\phi \in E$. If E is inconsistent, then by Lemma A.1.9 and Lemma A.1.2 E is the only extension of Δ'. Hence $\Delta' \models^d_{\mathcal{P}} \neg k$. Otherwise assume E is consistent. By Lemma A.1.10 $E_1 = Cn(E \cup \{c\})$ is an extension of $\langle D, W \cup \{c\}\rangle$. By Lemma A.1.8 $E_2 = Cn(E' \cup \{\neg k\})$ is an extension of $\langle D \cup \{\phi{:}\neg k/\neg k\}, W \cup \{c\}\rangle$. By Lemma A.1.9 E_2 is an extension of $\langle D \cup \{c{:}k/k, \phi{:}\neg k/\neg k\}, W \cup \{c\}\rangle$. By Lemma A.1.9 E_2 is an extension of $\langle D \cup \{k{:}\neg c/\neg c, c{:}k/k, \phi{:}\neg k/\neg k\}, W \cup \{c\}\rangle$. Now $\neg k \in E_2$ and not $\mathrm{def}(k{:}\neg c/\neg c, E_2)$. Assume there is an extension E' of Δ' such that $\neg k \notin E'$. Because $c \in E'$ and $c{:}k/k \in D'$, $k \in E'$. Hence $\mathrm{def}(k{:}\neg c/\neg c, E', E_2)$. Because $k{:}\neg c/\neg c$ is the $\mathcal{P}$-least default of D', E' cannot be a $\mathcal{P}$-preferred$_d$ extension of Δ'. Hence all $\mathcal{P}$-preferred$_d$ extensions contain $\neg k$, and $\Delta' \models^d_{\mathcal{P}} \neg k$.

Assume $\Delta' \models^d_{\mathcal{P}} \neg k$. By assumption Δ' has at least one extension and by Theorem 3.1.20 there is at least one $\mathcal{P}$-preferred extension of Δ'. Therefore there is at least one $\mathcal{P}$-preferred extension E' of Δ' such that $\neg k \in E'$. If E' is inconsistent, then by Lemma A.1.9 and Lemma A.1.2 E' is the only extension of Δ, and $\Delta \models_b \phi$. Otherwise E' is consistent. By Lemma A.1.12 $E = Cn(W \cup \{\gamma | \alpha{:}\sigma/\gamma \in \mathrm{GD}(E, \Delta)\})$ is an extension of Δ. Because the propositional variables $\{k, c\}$ do not occur in ϕ, $\phi \in E$. Therefore $\Delta \models_b \phi$. □

Although the restrictions of Theorem 4.2.5 are slightly stronger than those of Theorem 4.2.3, the applicability of the theorem to the classes in Figure 4.7 is the same.

Corollary 4.2.6 *Determining $\langle D, W\rangle \models^d_{\mathcal{P}} \phi$ for strict total orders $\mathcal{P}$ in any of the classes of default theories below is NP-hard.*

D	*W*	ϕ
normal unary	*Horn clauses*	*literal*
normal unary	*2-literal clauses*	*literal*
DF normal	*literals*	*literal*
ordered unary	*literals*	*literal*

The above theorems do not apply to prerequisite-free default theories. Next we give a similar theorem for prerequisite-free default theories. As application and non-defeat coincide for prerequisite-free defaults, there is no need for separate theorems for the consequence relations $\models_{\mathcal{P}}$ and $\models^d_{\mathcal{P}}$.

Theorem 4.2.7 *Let $\mathcal{F}$ be a class of formulae and $\mathcal{C}$ a class of finite default theories such that*

- *if* $\langle D, W\rangle \in \mathcal{C}$*, then* $\langle D \cup \{:k \wedge \neg l/k\}, W\rangle \in \mathcal{C}$ *where* k *and* l *are propositional variables,*
- *if* $\langle D, W\rangle \in \mathcal{C}$*, then* $\langle D \cup \{:k \wedge \neg\phi/k\}, W\rangle \in \mathcal{C}$ *where* $\phi \in \mathcal{F}$ *and* p *is a propositional variable, and*
- *each member of* $\mathcal{C}$ *has at least one extension.*

The problem of testing $\Delta \models_b \phi$ *for* $\Delta \in \mathcal{C}$ *and* $\phi \in \mathcal{F}$ *is polynomial time many-one reducible to the problem of testing* $\Delta' \models_{\mathcal{P}'} \phi'$ *where* $\Delta' \in \mathcal{C}$ *and* $\phi' \in \mathcal{F}$ *and* $\mathcal{P}'$ *is a strict total order on defaults in* Δ'*.*

Proof: Let $\Delta = \langle D, W\rangle \in \mathcal{C}$ be a default theory. Let $\Delta' = \langle D \cup \{:k \wedge \neg\phi/k, :k' \wedge \neg k/k'\}, W\rangle$ where k and k' are propositional variables that do not occur in $\langle D, W\rangle$ or in ϕ. Let $\mathcal{P}$ be a strict partial order on $D \cup \{:k \wedge \neg\phi/k, :k' \wedge \neg k/k'\}$ with $:k' \neg k/k'$ as the least element. We claim that $\Delta \models_b \phi$ if and only if $\Delta' \models_{\mathcal{P}} k'$. The construction of Δ' from Δ is polynomial time.

Assume $\Delta \models_b \phi$, that is, there is an extension E of Δ such that $\phi \in E$. If E is inconsistent, then by Lemma A.1.2 and Lemma A.1.9 E is the only extension of Δ'. Hence $\Delta' \models_{\mathcal{P}} k'$. Otherwise E is consistent. By Lemma A.1.8 $E' = Cn(E \cup \{k'\})$ is an extension of $\langle D \cup \{:k' \wedge \neg k/k'\}, W\rangle$. By Lemma A.1.9 E' is an extension of $\langle D \cup \{:k \wedge \neg\phi/k, :k' \wedge \neg k/k'\}, W\rangle$. Assume that E'' is an extension of Δ' such that not appl$(:k' \wedge \neg k/k', E'')$. Now appl$(:k' \wedge \neg k/k', E', E'')$ but there is a default $\delta\mathcal{T}:k' \wedge k/k'$ for no strict total order $\mathcal{T}$ such that $\mathcal{P} \subseteq \mathcal{T}$. Therefore E'' is not a $\mathcal{P}$-preferred extension of Δ', and appl$(:k' \wedge \neg k/k', E'')$ for all $\mathcal{P}$-preferred extensions E''. By the definition of application, now $k' \in E''$ for all $\mathcal{P}$-preferred extensions of Δ' and therefore $\Delta' \models_{\mathcal{P}} k'$.

Assume $\Delta' \models_{\mathcal{P}} k'$, that is, k' belongs to all $\mathcal{P}$-preferred extensions of Δ'. By assumption there is at least one extension of Δ' and by Theorem 3.1.20 there is at least one $\mathcal{P}$-preferred extension E' of Δ', and hence $k' \in E'$. If E' is inconsistent, then by Lemma A.1.9 E' is an extension of Δ and $\Delta \models_b \phi$. Otherwise E' is consistent. Now appl$(k' \wedge \neg k{:}k'/, E')$ because E' is consistent and k' does not occur in W or in any default in D. Now $\neg(k' \wedge \neg k) \notin E'$ and hence $k \notin E'$. Therefore not appl$(:k \wedge \neg\phi/k, E')$, and $\neg(k \wedge \neg\phi) \in E'$. As E' is consistent and k does not occur in any other default in Δ', $\phi \in E'$. By Lemma A.1.12 $E = Cn(W \cup \{\gamma | \alpha{:}\sigma/\gamma \in D \cap \mathrm{GD}(E', \Delta')\})$ is an extension of Δ. Because propositional variables k and k' do not occur in ϕ, $\phi \in E$. Therefore there is an extension E of Δ that contains ϕ, and hence $\Delta \models_b \phi$. $\Box$

Corollary 4.2.8 *Determining* $\langle D, W\rangle \models_{\mathcal{P}} \phi$ *for strict total orders* $\mathcal{P}$ *in any of the classes of default theories below is NP-hard.*

D	W	ϕ
prerequisite-free ordered	*literals*	*literal*
prerequisite-free ordered unary	*literals*	*positive literal*

By Theorem 3.1.15 for layered strict partial orders, of which strict total orders are a special case, $\mathcal{P}$-preferred$_2$ and $\mathcal{P}$-preferred$_{2d}$ extensions respectively coincide with $\mathcal{P}$-preferred and $\mathcal{P}$-preferred$_d$ extensions. Because all the above intractability results are for strict total orders, they imply the intractability of $\models^2_{\mathcal{P}}$ and $\models^{2d}_{\mathcal{P}}$ by Theorem 3.1.15.

The remaining classes in the Kautz, Selman and Stillman classification without an intractability result or a polynomial time decision procedure are the prerequisite-free unary classes, the prerequisite-free normal classes, and the Horn and unary classes with literals. Our first intractability result for a restricted subclass of prioritized default logic (Corollary 4.2.2) can be used for establishing the intractability of the prerequisite-free unary class. The remaining classes are discussed in the next section.

Theorem 4.2.9 *Let $\mathcal{C}$ be a class of default theories $\langle D, W\rangle$ where W is a finite set of formulae and D is a finite set of defaults that may be of the form $:q\wedge\neg r/q$ where q and r are propositional variables. Testing $\Delta \models_{\mathcal{P}} l$ for $\Delta \in \mathcal{C}$, literals l, and strict total orders $\mathcal{P}$ is co-NP-hard.*

Proof: The result is based on the fact that the extensions used in the intractability proof of Marek and Truszczyński (Theorem 4.2.1) contain only positive literals. Hence the conjuncts q in the justifications $q \wedge \neg r$ of the prerequisite-free unary defaults $:q \wedge \neg r/q$ do not figure in the definition of extensions.

Let D be a set of defaults of the form $:\neg p/q$ where p and q are propositional variables. Define the operator $F(D) = \{:p\wedge\neg q/p | :\neg q/p \in D\}$. We claim that every set $E \subseteq \mathcal{L}$ fulfills one of the following equations if and only if it fulfills the other.

$$E = Cn(\left\{p \,\middle|\, \frac{:p\wedge\neg q}{p} \in F(D), \neg(p\wedge\neg q) \notin E\right\})$$
$$E = Cn(\left\{p \,\middle|\, \frac{:\neg q}{p} \in D, q \notin E\right\})$$

By Theorem 2.1.8 sets E that fulfill the above equations are extensions of $\langle F(D), \emptyset\rangle$ and $\langle D, \emptyset\rangle$, respectively. Now $\neg(p\wedge\neg q) \in E$ if and only if $\neg p \vee q \in E$ if and only if $q \in E$. The first equivalence is simply the logical equivalence of the formulae, and the second equivalence is as follows. Let q be in E. Because E is closed under logical consequence and $q \models \neg p \vee q$, also $\neg p \vee q \in E$. Let q not be in E. Let G be the set of consequents of the generating defaults of

E, and hence $E = Cn(G)$. Let M be a model that makes all propositional variables except q true. Now $M \models G$ because G is a set of propositional variables and $q \notin G$. Now $M \not\models \neg p \wedge q$ and $G \not\models \neg p \wedge q$ and $\neg p \wedge q \notin E$.

Therefore E is an extension of $\langle D, \emptyset \rangle$ if and only if E is an extension of $\langle F(D), \emptyset \rangle$. By Corollary 4.2.2 the problem of testing $\Delta \models_{\mathcal{P}} \phi$ for default theories $\Delta = \langle D, \emptyset \rangle$, where D consists of defaults $:\neg p/q$ with propositional variables p and q, and literals ϕ is co-NP-hard. Because the translation from $\langle D, \emptyset \rangle$ to $\langle F(D), \emptyset \rangle$ is polynomial time, testing $\langle F(D), \emptyset \rangle \models_{\mathcal{P}} \phi$ – the problem stated in this theorem – is co-NP-hard as well. □

4.2.3 Restricted tractable classes

Next we give a restricted tractable class of prioritized default theories for which the unique $\mathcal{P}$-preferred extensions can be computed by using $\mathcal{P}$ as the order in which the application of defaults is attempted. Exhaustive search is avoided because there is no need to retract a decision to apply a default. The theorem is given in a general form for any tractable subset of classical propositional logic, like Horn clauses, 2-literal clauses, or 1-literal clauses.

Theorem 4.2.10 *Let $\mathcal{F}$ be a class of formulae such that satisfiability testing for finite $S \subseteq \mathcal{F}$ takes polynomial time in the size of S. Let $\mathcal{C}$ be a set of default theories $\langle D, W \rangle$ such that $D = \{:\phi/\phi | \phi \in F\}$ for some finite set $F \subseteq \mathcal{F}$ and $W \subseteq \mathcal{F}$ is finite. Then for $\Delta \in \mathcal{C}$, $\phi \in \mathcal{F}$, and strict total orders $\mathcal{P}$ on the defaults in Δ, the problem of testing $\Delta \models_{\mathcal{P}} \neg\phi$ is solvable in polynomial time on the size of $D \cup W \cup \{\neg\phi\}$. If satisfiability testing in $\mathcal{F}$ is linear time, then the problem is solvable in quadratic time.*

Proof: By Lemma 3.1.18 there is at most one $\mathcal{P}$-preferred extension for each member of $\mathcal{C}$, and because the defaults in members of $\mathcal{C}$ are normal, there is at least one extension for each member of $\mathcal{C}$ [Reiter, 1980, Theorem 3.1]. Therefore there is exactly one $\mathcal{P}$-preferred extension for each member of $\mathcal{C}$. A decision procedure for the class of default theories stated in the theorem is given in Figure 4.8. The procedure computes a set E that consists of W and the conclusions of the generating defaults of the extension $Cn(E)$ of $\langle D, W \rangle$, and returns *true* if and only if $E \models \phi$.

First we prove the correctness of the algorithm, and then we give a quadratic time upper bound for the computation in the case where satisfiability testing is linear time. Let $:\phi_1/\phi_1, \ldots, :\phi_n/\phi_n$ be the ordering $\mathcal{P}$ of defaults in D. If W is inconsistent the procedure returns *true*. In this case $Cn(W) = \mathcal{L}$ is the only extension of $\langle D, W \rangle$ which agrees with the statement of the lemma. Otherwise extensions of $\langle D, W \rangle$ are consistent. Define $D_i = \{:\phi_1/\phi_1, \ldots, :\phi_i/\phi_i\}$. Define

$\Pi_i = \{\phi_1, \ldots, \phi_i\}$. The correctness proof is by induction on i and we obtain the $\mathcal{P}$-preferredness of $Cn(E)$ as the case $i = n$.

Induction hypothesis: for all $j \in \{0, \ldots, i\}$, if W is consistent, E_j is a maximal consistent subset of $\Pi_j \cup W$ such that $W \subseteq E_j$ and $Cn(E_i)$ is a $\mathcal{P} \cap (D_i \times D_i)$-preferred extension of $\langle D_i, W\rangle$.

Base case $i = 0$: Because $D_0 = \emptyset$, there is exactly one extension $Cn(W)$ of $\langle D_0, W\rangle$ that is $\emptyset$-preferred by Theorem 3.1.20. Furthermore W is a maximal consistent subset of W whenever W is consistent.

Inductive case $i \geq 1$: Assume $E_{i-1} \cup \{\phi_i\}$ is consistent. Now $E_i = E_{i-1} \cup \{\phi_i\}$. Because by induction hypothesis $E_{i-1} \cup \{\phi\}$ is inconsistent for any $\phi \in \Pi_{i-1} \backslash E_{i-1}$ and $E_{i-1} \subseteq E_i$, also $E_i \cup \{\phi\}$ is inconsistent for any $\phi \in \Pi_{i-1} \backslash E_{i-1}$. Because $\Pi_i \backslash E_i = \Pi_{i-1} \backslash E_{i-1}$, $E_i \cup \{\phi\}$ is inconsistent for any $\phi \in \Pi_i \backslash E_i$, and therefore E_i is a maximal consistent subset of Π_i. Assume $E_{i-1} \cup \{\phi_i\}$ is inconsistent. Now $E_i = E_{i-1}$. Because by induction hypothesis $E_{i-1} \cup \{\phi\}$ is inconsistent for any $\phi \in \Pi_{i-1} \backslash E_{i-1}$ and $E_{i-1} \subseteq E_i$, also $E_i \cup \{\phi\}$ is inconsistent for any $\phi \in \Pi_{i-1} \backslash E_{i-1}$. Because $\Pi_{i-1} \backslash E_i = \Pi_{i-1} \backslash E_{i-1}$ and $E_i \cup \{\phi_i\}$ is inconsistent, $E_i \cup \{\phi\}$ is inconsistent for any $\phi \in \Pi_i \backslash E_i$, and therefore E_i is a maximal consistent subset of Π_i.

By Theorem 2.1.8 $Cn(E_i)$ is an extension of $\langle D_i, W\rangle$ if and only if $Cn(E_i) = Cn(\{\phi | {:}\phi/\phi \in D_i, \neg\phi \notin Cn(E_i)\})$. The right hand side of the equivalence is shown by the following equations.

$$\begin{array}{rcl}
Cn(E_i) & = & Cn(\{\phi \in \Pi_i | \phi \in E_i\}) \\
& = & Cn(\{\phi \in \Pi_i | E_i \cup \{\phi\} \text{ is consistent}\}) \\
& = & Cn(\{\phi \in \Pi_i | E_i \not\models \neg\phi\}) \\
& = & Cn(\left\{\phi \,\middle|\, \dfrac{:\phi}{\phi} \in D_i, \neg\phi \notin Cn(E_i)\right\})
\end{array}$$

The first equality is because $E_i \subseteq \Pi$. The second equality is because E_i is a maximal consistent subset of Π_i, and therefore $E_i \cup \{\phi\}$ for $\phi \in \Pi_i$ is consistent if and only if $\phi \in E_i$. The third equality is because the consistency of $E_i \cup \{\phi\}$ is equivalent to $E_i \not\models \neg\phi$. The last equality is by the definitions of D_i, Π_i and Cn.

We show that $Cn(E_i)$ is a $\mathcal{P} \cap (D_i \times D_i)$-preferred extension of $\langle D_i, W\rangle$. Let $Cn(E')$ be any extension of $\langle D_i, W\rangle$. Assume there is default ${:}\phi_k/\phi_k \in D_i$ for some $k \in \{1, \ldots, i\}$ such that $\text{appl}(\delta, E', Cn(E_i))$ and there is no $\delta' \in D_i$ such that $\delta' \mathcal{P} {:}\phi_k/\phi_k$ and $\text{appl}(\delta', Cn(E_i), E')$. By induction hypothesis, $W \cup (E_i \cap \Pi_k)$ is a maximal consistent subset of $W \cup \Pi_k$. This, however, contradicts the fact that $W \cup (E_i \cap \Pi_k) \subseteq W \cup (E' \cap \Pi_k)$ and $W \cup (E' \cap \Pi_k)$ is consistent. Therefore the assumption that E' exists is wrong, and E_i is a $\mathcal{P} \cap (D_i \times D_i)$-preferred extension of $\langle D_i, W\rangle$.

$PROCEDURE$ decide$(D, W, \mathcal{P}, \phi)$
$BEGIN$
  LET $:\phi_1/\phi_1, :\phi_2/\phi_2, \ldots, :\phi_n/\phi_n$ be the ordering $\mathcal{P}$ on D;
  $E_0 := W$;
  $FOR\ i := 1\ TO\ n\ DO$
    $IF\ E_{i-1} \cup \{\phi_i\}$ is consistent (equivalently, $E_{i-1} \not\models \neg\phi_i$)
    $THEN\ E_i := E_{i-1} \cup \{\phi_i\}$
  END;
  $E := E_n$;
  $IF\ E \models \phi$
  $THEN\ RETURN$ true $ELSE\ RETURN$ false
END

Figure 4.8: A decision procedure for a tractable class of prerequisite-free normal default theories

Finally, we show that the computation of E is $\mathcal{O}(s^2)$ where s is the sum of the lengths of formulae in W and the conclusions of defaults in D, and f is a function that represents the time complexity of making a satisfiability test. For example, for a subclass of classical propositional logic for which there is a linear time satisfiability testing algorithm, f is a function $f(s) = c + as$ for some c and a, and in general it is a polynomial on s. Let $s_1, \ldots, s_n$ be the sizes of $\phi_1, \ldots, \phi_n$, and $s = \sum_{i=1}^{n} s_i$. An upper bound for the time taken by the satisfiability test at step i is $f(w + s_i + \sum_{j=1}^{i-1} s_j)$ where w is the size of W. Taking $f(s) = c + as$ we obtain

$$\begin{aligned}
\sum_{i=1}^{n} f(w + s_i + \sum_{j=1}^{i-1} s_j) &= \sum_{i=1}^{n}(c + a(w + \sum_{j=1}^{i} s_j)) \\
&\leq \sum_{i=1}^{n}(c + a(w + \sum_{j=1}^{n} s_j)) \\
&= n(c + aw + as) \\
&\leq (c + aw + as)s \leq (c + aw + a)s^2.
\end{aligned}$$

The consequence test $E \models \phi$ takes linear time on the size of $E \cup \{\neg\phi\}$. Hence the whole computation is $\mathcal{O}(n^2)$ where n is the size of $D \cup W \cup \{\phi\}$. Clearly, the reconstruction of the above chain of equations for some other polynomial f would also result in a polynomial. Thus any polynomial time subset $\mathcal{C}$ of classical propositional logic would produce a tractable class of default theories. □

Note that reasoning with prerequisite-free normal default rules and Horn clauses without priorities is NP-complete [Stillman, 1990]. The tractability of the above class of default theories is due to the the absence of prerequi-

$PROCEDURE$ exists(D, W, A)
$BEGIN$
  $IF\ A \neq \emptyset$ and $W \cup \{\pi | \alpha{:}\pi/\pi \in A\} \models \bot\ THEN\ RETURN$ false;
  $D' := A \cup \{\alpha{:}p/p \in D | \neg p \notin W, \alpha'{:}\neg p/\neg p \notin A, p$ is atomic$\}$;
  $E := W$;
  $REPEAT$
    $E' := E$;
    $FOR\ EACH\ \alpha{:}\pi/\pi \in D'\ DO$
      $IF\ E' \models \alpha\ THEN\ E := E \cup \{\pi\}$
    END
  $UNTIL\ E = E'$;
  $IF\ E \not\models \alpha$ for some $\alpha{:}\pi/\pi \in A\ THEN\ RETURN$ false
  $ELSE\ RETURN$ true
END

Figure 4.9: A subprocedure of a decision procedure

sites. For normal defaults, tractability can be obtained also for defaults with prerequisites if suitable restrictions are imposed on the prerequisites, conclusions, and the objective parts of the default theories. Next we present such a class that is neither subsumed by nor subsumes the tractable prerequisite-free classes. The idea behind the result is the same as the one used in the proof of Theorem 4.1.12. Theorem 4.2.3 indicates that the result cannot be generalized (assuming P≠NP) to any of the classes higher in the hierarchy of default theories presented by Kautz and Selman [1991].

Theorem 4.2.11 *Let $\mathcal{C}$ be a class of default theories $\langle D, W\rangle$ where W is a finite set of literals and defaults in D are of the form $p_1 \wedge \cdots \wedge p_n{:}l/l$ where l is a literal and $p_1, \ldots, p_n$ are propositional variables. For $\Delta \in \mathcal{C}$, negated Horn clauses ϕ, and strict total orders $\mathcal{P}$ on the defaults in Δ, the problem of testing $\Delta \models_{\mathcal{P}} \neg\phi$ is solvable in polynomial time on the size of $D \cup W \cup \{\neg\phi\}$.*

Proof: The body of the decision procedure is given in Figure 4.6 and its correctness proof is included in the proof of Theorem 4.1.12. Here we use the subprocedure *exists* in Figure 4.9. The procedure *exists* runs in polynomial time on the size of $D \cup W \cup A$ because the number of iterations in the repeat-until loop is at most $|D'|$, and the logical consequence tests in the procedure can be performed in polynomial time because the formulae involved are propositional Horn clauses. The proof that the procedure *exists* called with arguments (D, W, A) returns *true* if and only if there is an extension E of $\Delta = \langle D, W\rangle$ such that $A \subseteq \mathrm{GD}(E, \Delta)$, is given next.

First we show that *exists* returns true if and only if there is an extension E' of $\Delta' = \langle D', W\rangle$ such that $A \subseteq \mathrm{GD}(E', \Delta')$ where $D' = A \cup \{\alpha{:}p/p \in D | \neg p \notin$

$W, \alpha'{:}\neg p/\neg p \notin A, p$ is atomic$\}$. If $W \cup \{\pi|\alpha{:}\pi/\pi \in A\}$ is inconsistent and $A \neq \emptyset$, then there are no extensions that apply A, and hence it is correct to return *false*. So assume $W \cup \{\pi|\alpha{:}\pi/\pi \in A\}$ is consistent. Now the set $W \cup G'$, where $G' = \{\pi|\alpha{:}\pi/\pi \in D'\}$, is consistent by definition of D'. Any extension E' of Δ' satisfies $E' \subseteq Cn(W \cup G')$. Because ${D'}^{Cn(W \cup G')} = D'$, ${D'}^{E'} = D'$ for any extension E' of Δ'. That the procedure computes a set E such that $Cn(E) = Cn_W(D')$ is immediate by the similarity of the definition of $Cn_W(D')$ and the statements in the procedure. Hence $E = Cn_W(D')$ is the unique extension of Δ'. If $W \cup \{\pi|\alpha{:}\pi/\pi \in A\}$ is consistent, the procedure returns false if and only if $\alpha \notin E$ for some $\alpha{:}\pi/\pi \in A$. This is equivalent to the fact that a default in A is not applied in the unique extension E of Δ'.

Then we show that there is an extension E' of Δ' such that $A \subseteq \mathrm{GD}(E', \Delta')$ if and only if there is an extension E of Δ such that $A \subseteq \mathrm{GD}(E, \Delta)$. The "only if" direction is by [Reiter, 1980, Theorem 3.2], because $D' \subseteq D$ and defaults in D are normal. Assume that there is an extension E of Δ such that $A \subseteq \mathrm{GD}(E, \Delta)$. Let $D_b = \{p{:}\neg q/\neg q \in D|p{:}\neg q/\neg q \notin A\}$. Let $E' = Cn_W(D^E \backslash D_b)$. Now $\neg\pi \notin E'$ for all $p{:}\pi/\pi \in A$ because $Cn_W(D^E \backslash D_b) \subseteq E = Cn_W(D^E)$ by Lemma A.1.3. We show by induction that $p \in E'$ for all $p{:}\pi/\pi \in A$. Let $C_0, C_1, \ldots$ be the sets in Theorem 2.1.3 for $Cn_W(D^E)$, and $C'_0, C'_1, \ldots$ the sets for $Cn_W(D^E \backslash D_b)$.

Induction hypothesis: $C_i \cap P \subseteq C'_i \cap P$ where P is the set of all propositional variables.

Base case $i = 0$: Immediate as $C'_0 = W = C_0$.

Inductive case $i \geq 1$: Assume $p \in C_i \cap P$. If $p \in C_{i-1}$ then by induction hypothesis $p \in C'_{i-1} \cap P$. If $p \in (C_i \cap P) \backslash (C_{i-1} \cap P)$, then $C_{i-1} \models q$ for some $q{:}p/p \in D^E$. Because defaults in D_b have a non-atomic conclusion, $q{:}p/p \in D^E \backslash D_b$. By induction hypothesis $q \in C'_{i-1}$ and hence $p \in C'_i \cap P$.

Hence for all $p{:}\pi/\pi \in A$, $p \in Cn_W(D^E \backslash D_b)$, and because $Cn_W(D^E \backslash D_b) \subseteq Cn_W(D^E)$, $\overline{\pi} \notin Cn_W(D^E \backslash D_b)$. Let $D^* = \{\alpha{:}\beta/\beta \in D^E \backslash D_b | \alpha \in E'\}$. By Lemma A.1.7 E' is an extension of $\langle D^*, W\rangle$. Clearly $A \subseteq \mathrm{GD}(E', \langle D^*, W\rangle)$. By [Reiter, 1980, Theorem 3.2] there is an extension E'' of $\langle D, W\rangle$ such that $\mathrm{GD}(E', \langle D^*, W\rangle) \subseteq \mathrm{GD}(E'', \Delta)$. Hence $A \subseteq \mathrm{GD}(E'', \Delta)$.

Therefore the procedure returns *true* if and only if there is an extension E of $\langle D, W\rangle$ such that $A \subseteq \mathrm{GD}(E, \Delta)$. □

As a comparison, cautious reasoning for a more narrow class of default theories without priorities is NP-hard [Kautz and Selman, 1991]. The tractability result is related to the fact that for the class of default theories in question the problem of brave reasoning is tractable [Kautz and Selman, 1991, Lemma 6.4]. This fact is taken advantage of in the computation of the unique preferred extension. Kautz and Selman's tractability result indicates a way of performing the computation of *exists* in linear time, thereby implying a quadratic time algorithm for this class of prioritized default theories.

4.2.4 Summary

The results on the boundary between tractability and intractability for $\models_{\mathcal{P}}$ and $\models^2_{\mathcal{P}}$ for strict total orders $\mathcal{P}$ are summarized in Figure 4.10. We have listed the references to the theorems and corollaries where the results are stated. Clearly, all superclasses of an intractable class are intractable as well, and similarly all subclasses of a tractable class are tractable. We use the notation $n \subseteq$ to indicate that the intractability result of the class is based on the intractability of the subclass n, and the notation $\subseteq n$ to indicate that the tractability result is based on the tractability of the superclass n.

The tractable classes with prerequisite-free normal defaults $:b_1 \wedge \cdots \wedge b_n/b_1 \wedge \cdots \wedge b_n$ have more general tractable superclasses. Instead of a conjunction of literals $b_1 \wedge \cdots \wedge b_n$, the justifications and the conclusions of the defaults may be conjunctions of clauses $c_1 \wedge \cdots \wedge c_n$ so that each clause c_i is of the same form as allowed in W, that is, Horn clauses, 2-literal clauses, or 1-literal clauses. With this modification, all satisfiability tests needed in the algorithm of Theorem 4.2.10 are for a tractable class of propositional formulae.

The idea that the reduction in the number of unresolved conflicts between defaults due to priorities would reduce the complexity of nonmonotonic reasoning seems in general unfounded. In all the cases we have investigated in Section 4.2.2, the classical reasoning involved is polynomial time (typically linear time), due to total priorities there is a unique way of solving each conflict, and the decision problems are still intractable. Only in the case of prerequisite-free normal and Horn defaults some of the classes are tractable with priorities and intractable without them. A polynomial-time test for the existence of extensions does not yield a polynomial-time procedure for constructing the extensions.

	form of defaults	clauses in W		
		Horn	2-literal	1-literal
1	$\frac{a_1\wedge\cdots\wedge a_l : b_1\wedge\cdots\wedge b_n\wedge c_1\wedge\cdots\wedge c_m}{b_1\wedge\cdots\wedge b_n}$	NP-hard	NP-hard	NP-hard
2	$\frac{p:q}{q}$ $\frac{p:q\wedge\neg r}{q}$ $\frac{p:\neg q}{\neg q}$	NP-hard	NP-hard	NP-hard
3	$\frac{a_1\wedge\cdots\wedge a_l : b_1\wedge\cdots\wedge b_n\wedge c_1\wedge\cdots\wedge c_m}{b_1\wedge\cdots\wedge b_n}$	NP-hard	NP-hard	NP-hard
4	$\frac{p:q}{q}$ $\frac{p:q\wedge\neg r}{q}$ $\frac{p:\neg q}{\neg q}$	NP-hard	NP-hard	NP-hard
5	$\frac{a_1\wedge\cdots\wedge a_l : b_1\wedge\cdots\wedge b_n}{b_1\wedge\cdots\wedge b_n}$	NP-hard	NP-hard	NP-hard
6	$\frac{p_1\wedge\cdots\wedge p_n : q}{q}$ $\frac{p_1\wedge\cdots\wedge p_n : \neg q}{\neg q}$	NP-hard	NP-hard	PTIME
7	$\frac{p:q}{q}$ $\frac{p:\neg q}{\neg q}$	NP-hard	NP-hard	PTIME
8	$\frac{: b_1\wedge\cdots\wedge b_n\wedge c_1\wedge\cdots\wedge c_m}{b_1\wedge\cdots\wedge b_n}$	NP-hard	NP-hard	NP-hard
9	$\frac{: b_1\wedge\cdots\wedge b_n\wedge c_1\wedge\cdots\wedge c_m}{b_1\wedge\cdots\wedge b_n}$	NP-hard	NP-hard	NP-hard
10	$\frac{:q}{q}$ $\frac{:q\wedge\neg r}{q}$ $\frac{:\neg q}{\neg q}$	co-NP-hard	co-NP-hard	co-NP-hard
11	$\frac{:q}{q}$ $\frac{:q\wedge\neg r}{q}$ $\frac{:\neg q}{\neg q}$			
12	$\frac{: b_1\wedge\cdots\wedge b_n}{b_1\wedge\cdots\wedge b_n}$	PTIME	PTIME	PTIME
13	$\frac{:q}{q}$ $\frac{:\neg q}{\neg q}$	PTIME	PTIME	PTIME
14	$\frac{:q}{q}$	PTIME	PTIME	PTIME

	form of defaults	reference to result		
		Horn	2-literal	1-literal
1	disjunction free	$4 \subseteq$	$4 \subseteq$	$4 \subseteq$
2	unary	$4 \subseteq$	$4 \subseteq$	$4 \subseteq$
3	disjunction free ordered	$5 \subseteq$	$5 \subseteq$	$5 \subseteq$
4	ordered unary	C4.2.4	C4.2.4	C4.2.4
5	disjunction free normal	C4.2.4	C4.2.4	C4.2.4
6	Horn	$7 \subseteq$	$7 \subseteq$	T4.2.11
7	normal unary	C4.2.4	C4.2.4	$\subseteq 6$
8	prerequisite free	$9 \subseteq$	$9 \subseteq$	$9 \subseteq$
9	prerequisite free ordered	C4.2.8	C4.2.8	C4.2.8
10	prerequisite free unary	T4.2.9	T4.2.9	T4.2.9
11	prerequisite free ordered unary			
12	prerequisite free normal	T4.2.10	T4.2.10	T4.2.10
13	prerequisite free normal unary	$\subseteq 12$	$\subseteq 12$	$\subseteq 12$
14	prerequisite free positive normal unary	$\subseteq 12$	$\subseteq 12$	$\subseteq 12$

Figure 4.10: Summary of the complexity results in restricted cases

Chapter 5

Automated reasoning in prioritized default logic

In this section we give an algorithm for the decision problem of cautious reasoning in prioritized default logic. Some of the decision procedures for tractable classes of prioritized default logic given in Section 4.2.3 suggest that priorities have a natural computational interpretation as an ordering in which defaults are considered by a decision procedure. The algorithm is based on the close connection between lexicographic comparison and binary search. Consider the binary tree in Figure 5.1. The leafs of tree can be seen to represent the sequences 000, 001, 010, 011, 100, 101, 110, 111. If the tree is traversed depth-first from left to right, the order in which the leaf nodes corresponding to the sequences are encountered is exactly the lexicographic ordering for the sequences. This suggests a way of computing the unique $\mathcal{P}$-preferred extension of a default theory when $\mathcal{P}$ is a total order. Let $\delta_0, \ldots, \delta_n$ be the total ordering $\mathcal{P}$ on the defaults. Nodes at distance i from the root node correspond to the default δ_i. The leaf nodes do not correspond to any default. For a default δ, the leafs in the leftmost subtree (labelled 0) correspond to extensions where δ is applied, and the leafs in the rightmost subtree (labelled

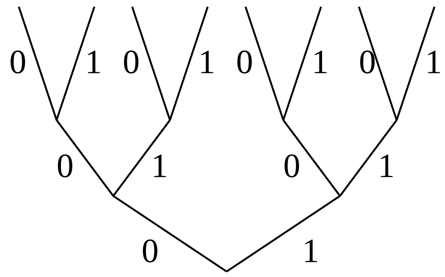


Figure 5.1: Relation between binary search and lexicographic comparison

1) correspond to extensions where δ is not applied. The computation now proceeds as follows. The binary tree is traversed from left to right, and at each leaf node it is determined whether the set of defaults that correspond the 0s are the generating defaults of an extension of the default theory. The first extension that is found is the unique $\mathcal{P}$-preferred extension of the default theory.

When $\mathcal{P}$ is not a strict total order there are several strict total orders that extend $\mathcal{P}$ and they may be characteristic orderings for different extensions. An obvious procedure would consider each strict total order $\mathcal{T}$ that extends $\mathcal{P}$ and perform the computation described above. However, for a strict partial order $\mathcal{P}$ on n elements there may be $n!$ strict total orders $\mathcal{T}$ that extend $\mathcal{P}$. Performing the computation for each of these of order $n!$ strict total orders is not efficient because $n!$ grows very fast. Even in the case $n = 10$, when the number of nodes in the binary tree is only $2^{10} = 1024$, the number of strict total orders $n!$ is 3628800. The function $n!$ grows a lot faster than 2^n. A procedure that considers all strict total orders has in the order of $n!$ phases, and at each phase it traverses a binary search tree that has in the order of 2^n leaf nodes. A more promising approach is to try to extend the basic procedure to cover all strict partial orders. This way the traversal of the tree with 2^n nodes only once would suffice.

It turns out that the implementation of this idea is not straightforward. It seems that for determining whether an extension found during the traversal of the binary tree is a preferred extension it is necessary to carry along the set of extensions found earlier during the traversal. Also, it seems that it is necessary to alter the correspondence between links and the defaults during the traversal, that is, the default that corresponds to the ith node on a path from the root to a leaf node may be different for different leaf nodes.

The need to carry along the set of extensions already found can be motivated by the following example. Consider a default theory $\Delta = \langle\{\delta_1, \delta_2\}, W\rangle$ and the priorities $\mathcal{P} = \{\langle\delta_1, \delta_2\rangle\}$. The search tree corresponds to the following sets of potential sets of generating defaults read from left to right: $G_1 = \{\delta_1, \delta_2\}, G_2 = \{\delta_1\}, G_3 = \{\delta_2\}$, and $G_4 = \emptyset$. Let us be in the leaf node corresponding to the set G_3, and let G_3 be the generating defaults of an extension of Δ. Now we have to know whether G_1 or G_2 is the set of generating defaults of an extension of Δ. If one of them is, then G_3 is not a preferred extension. If they are not, then G_3 is a preferred extension. In the case of total priorities it suffices to maintain information on whether one of the preceding leaf nodes correspond to an extension, but for arbitrary priorities this does not suffice.

The demonstration for the need to change the correspondence between defaults and links of the binary tree during the traversal is more complicated.

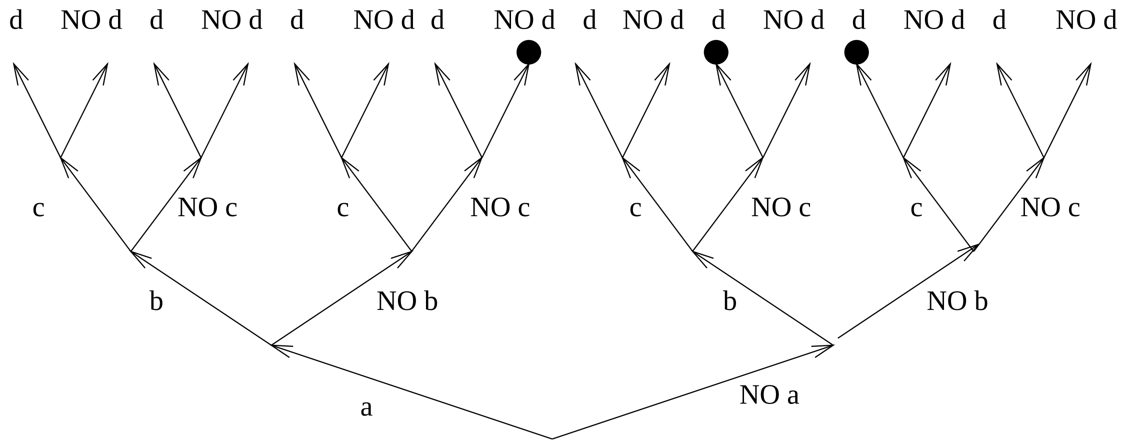
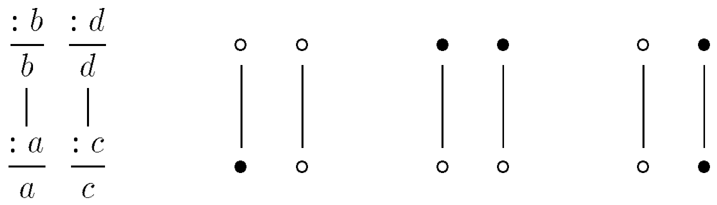


Figure 5.2: A search tree not traversed by the decision procedure

Consider a default theory $\Delta = \langle D, W\rangle$ with the following priorities $\mathcal{P}$ on the four defaults in D, and the three extensions.



The three extensions respectively apply the defaults $\mathrm{GD}(E_1, \Delta) = \{:a/a\}$, $\mathrm{GD}(E_2, \Delta) = \{:b/b, :d/d\}$, and $\mathrm{GD}(E_3, \Delta) = \{:c/c, :d/d\}$. The search tree for finding the preferred extensions of Δ is generated with a fixed ordering on the defaults $:a/a$ followed by $:b/b$, $:c/c$ and $:d/d$. The search tree with the extensions depicted as $\bullet$ is depicted in Figure 5.2. Branch with the label *NO* x means that the default $:x/x$ is not applied and x means that the defaults is applied in the extensions represented by the leaf nodes of the current subtree. Now the extensions E_1 and E_3 are $\mathcal{P}$-preferred. How would it be possible to detect that E_2 is not a $\mathcal{P}$-preferred extension before finding out that E_3 is an extension of Δ? By looking at the set $Q = \{E_1\}$ of extensions found earlier it would seem that any strict total order $\mathcal{T}$ on D with the property $:d/d\mathcal{T}:a/a$ would be a $\Delta, \mathcal{P}$-characteristic ordering for E_2 but this is not the case. This example suggests that generating the search tree with a fixed ordering on defaults does not work.

However, the problem may be solved by varying the ordering on defaults during the traversal. In the above example the crucial point is the selection of the next default δ for the subtree for *NO* a. Let Q be the set of extensions found in the subtree for a. Let S be the set of extensions E in the subtree for δ that are $\mathcal{P}$-preferred in $S \cup Q$ (here we recursively assume that the generation of this subtree of the search tree was successful.) Let S' be the set of preferred extensions in the subtree for *NO* δ. We must choose δ so that the $S \cup Q, \mathcal{P}$-characteristic ordering $\mathcal{T}_E$ for each E in S is also a $S' \cup S \cup Q', \mathcal{P}$-

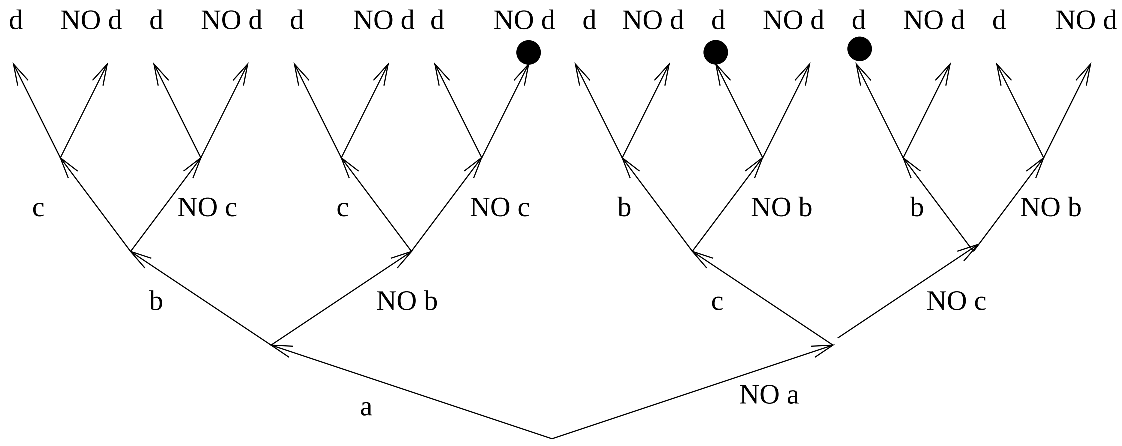


Figure 5.3: A search tree traversed by the decision procedure

characteristic ordering for E. For this purpose we find an ordering $\mathcal{T}'$ for the defaults between the root and the current node so that there is a $S \cup Q, \mathcal{P}$-characteristic ordering $\mathcal{T}$ for each extension E in S with $\mathcal{T}'$ as a prefix of $\mathcal{T}$. Now $\mathcal{T}$ is also a $S' \cup S \cup Q, \mathcal{P}$-characteristic ordering for E because the highest default where E and extensions in S' differ is δ, and $\mathrm{appl}(\delta, E)$ and for no $E' \in S'$ $\mathrm{appl}(\delta, E')$. This implies that after taking the *NO a* branch of the search tree, the default $:c/c$ instead of $:b/b$ should be chosen next. The new search tree is depicted in Figure 5.3. Now when the extension with generating defaults $:b/b$ and $:d/d$ is found, it can be seen in terms of the extensions already found that the extension is not preferred.

The decision procedure is given in Figure 5.6 and two auxiliary procedures in Figures 5.4 and 5.5. We need some lemmata for the correctness proofs. At each node of the search tree it is determined whether the extensions corresponding to the leaves in the current subtree could be preferred extensions as far as the extensions found so far are concerned. We introduce the notion *preferred in* for this purpose. As a special case, an extension is preferred exactly when it is preferred in the set of all extensions of the default theory.

Definition 5.0.12 (Preferred in) *Let $\Delta = \langle D, W\rangle$ be a default theory and $\mathcal{P}$ be a strict partial order on D. An extension E of Δ is $\mathcal{P}$-*preferred in* a set of extensions X of Δ if and only if there is a strict total order $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$ and for all $E' \in X$ and $\delta \in D$,*

$$\mathit{appl}(\delta, E', E) \text{ implies that for some } \epsilon \in D, \epsilon\mathcal{T}\delta \text{ and } \mathit{appl}(\epsilon, E, E').$$

*Then $\mathcal{T}$ is a $X, \mathcal{P}$-*characteristic ordering *for E.*

Lemma 5.0.13 *Let $\Delta = \langle D, W\rangle$ be a default theory where D is finite. Let $\mathcal{P}$ be a strict partial order on D. Let X be a set of extensions of Δ such that every $E \in X$ is $\mathcal{P}$-preferred in X and all $\mathcal{P}$-preferred extensions of Δ are* *in X. Then X is exactly the set of $\mathcal{P}$-preferred extensions of Δ. If E is a $\mathcal{P}$-preferred extension of Δ, then E is $\mathcal{P}$-preferred in any subset of extensions of Δ.*

Proof: For the first claim it suffices to show that all members of X are $\mathcal{P}$-preferred extensions of Δ. Assume $E \in X$ is not $\mathcal{P}$-preferred. Then there is an $X, \mathcal{P}$-characteristic ordering $\mathcal{T}$ for E. By Lemma 3.1.18 there is exactly one $\mathcal{T}$-preferred extension E' of Δ, and $\mathcal{T}$ is a $\Delta, \mathcal{T}$-characteristic ordering for E'. By assumption $E' \in X$. Now by definition of $\Delta, \mathcal{T}$-characteristic orderings and $X, \mathcal{P}$-characteristic orderings, for all $\delta \in D$,

$$\mathrm{appl}(\delta, E, E') \text{ implies that for some } \epsilon \in D, \epsilon \mathcal{T} \delta \text{ and } \mathrm{appl}(\epsilon, E', E),$$

and for all $\delta \in D$,

$$\mathrm{appl}(\delta, E', E) \text{ implies that for some } \epsilon \in D, \epsilon \mathcal{T} \delta \text{ and } \mathrm{appl}(\epsilon, E, E').$$

By [Reiter, 1980, Theorems 2.4 and 2.5] for any two extensions there is a default that is applied in the first but not in the second. Starting from any δ such that $\mathrm{appl}(\delta, E, E')$ the above conditions induce an infinite sequence of defaults $\delta_1, \delta_2, \delta_3, \ldots$, where $\delta_1 = \delta$ and $\delta_{i+1} \mathcal{T} \delta_i$ and $\mathrm{appl}(\delta_{2i-1}, E, E')$ and $\mathrm{appl}(\delta_{2i}, E', E)$ for all $i \geq 1$. Infinite descending sequences of defaults are impossible because D is finite. Hence the assumption is wrong and all extensions that are $\mathcal{P}$-preferred in X are $\mathcal{P}$-preferred.

The proof of the second claim is as follows. Assume E is a $\mathcal{P}$-preferred extension of Δ. Then there is a $\Delta, \mathcal{P}$-characteristic ordering $\mathcal{T}$ for E. Obviously, $\mathcal{T}$ is a $X, \mathcal{P}$-characteristic ordering for E where X is the set of all extensions of Δ. Let X' be any set such that $X' \subseteq X$. Now the quantification over X in the definition of $X, \mathcal{P}$-characteristic orderings is restricted to a subset of X, and therefore $\mathcal{T}$ is a $X', \mathcal{P}$-characteristic ordering for E. □

The assumptions of the next lemma are the same as of Lemma 5.0.15.

Lemma 5.0.14 *Let $\Delta = \langle D, W \rangle$ be a default theory, $\mathcal{P}$ a strict partial order on D, and $\Phi \subseteq D$ a set such that $D \backslash \Phi \times \Phi \subseteq \mathcal{P}$. Let $Q \cup \{E\}$ be a set of extensions of Δ such that for all $\delta \in D \backslash \Phi$ and $F \in Q$, appl(δ, E) if and only if appl(δ, F). Let $\mathcal{T}$ be a $Q, \mathcal{P}$-characteristic ordering for E. Let δ be a $\mathcal{P}$-minimal element of Φ such that appl(δ, E) if appl(δ, F) for some $F \in Q$. Then there is a $Q, \mathcal{P}$-characteristic ordering $\mathcal{T}'$ for E such that $\{\delta\} \times \Phi \backslash \{\delta\} \subseteq \mathcal{T}'$.*

Proof: Let

$$\mathcal{T}' = (\mathcal{T} \cap (D \backslash \Phi)^2) \cup (D \backslash \Phi \times \Phi) \cup (\mathcal{T} \cap (\Phi \backslash \{\delta\})^2) \cup (\{\delta\} \times \Phi \backslash \{\delta\}).$$

We show that $\mathcal{T}'$ is a strict total order and $\mathcal{P} \subseteq \mathcal{T}'$.

1. $\mathcal{P} \subseteq \mathcal{T}'$.

   Assume $\langle\delta', \delta''\rangle \in \mathcal{P}$. Case 1. If $\{\delta', \delta''\} \subseteq D\backslash\Phi$, then $\langle\delta', \delta''\rangle \in \mathcal{T}'$ because $\mathcal{P} \subseteq \mathcal{T}$ and $\mathcal{T} \cap (D\backslash\Phi)^2 \subseteq \mathcal{T}'$. Case 2. If $\{\delta', \delta''\} \subseteq \Phi\backslash\{\delta\}$, then $\langle\delta', \delta''\rangle \in \mathcal{T}'$ because $\mathcal{P} \subseteq \mathcal{T}$ and $\mathcal{T} \cap (\Phi\backslash\{\delta\})^2 \subseteq \mathcal{T}'$. Case 3a. If $\delta' = \delta$ and $\delta'' \in \Phi\backslash\{\delta\}$, then $\langle\delta', \delta''\rangle \in \mathcal{T}'$ because $\{\delta\} \times \Phi\backslash\{\delta\} \subseteq \mathcal{T}'$. Case 3b. If $\delta' \in \Phi\backslash\{\delta\}$ and $\delta'' = \delta$, then $\langle\delta', \delta''\rangle \notin \mathcal{P}$ because by assumption δ is $\mathcal{P}$-minimal in Φ. Case 4. If $\delta' \in \Phi$ and $\delta'' \in D\backslash\Phi$, then it is not possible that $\langle\delta', \delta''\rangle \in \mathcal{P}$ as $D\backslash\Phi \times \Phi \subseteq \mathcal{P}$ and $\mathcal{P}$ is asymmetric. Case 5. If $\delta' \in D\backslash\Phi$ and $\delta'' \in \Phi$, then $\langle\delta', \delta''\rangle \in \mathcal{T}'$ because by definition $D\backslash\Phi \times \Phi \subseteq \mathcal{T}'$.

2. $\mathcal{T}'$ is asymmetric, transitive, and total.

   We show that if $\langle\delta', \delta''\rangle \in \mathcal{T}'$, then $\langle\delta'', \delta'\rangle \notin \mathcal{T}'$. Because $\mathcal{T}$ is asymmetric and $\mathcal{T}'\backslash\mathcal{T} \subseteq \{\delta\} \times \Phi\backslash\{\delta\}$, it suffices to show that $\langle\delta', \delta\rangle \in \mathcal{T}'$ for no $\delta' \in \Phi$. This is directly by the definition of $\mathcal{T}'$, because the only set in the union that contains a pair with δ and another member of Φ is $\{\delta\} \times \Phi\backslash\{\delta\}$.

   We show that if $\langle\delta', \delta''\rangle \in \mathcal{T}'$ and $\langle\delta'', \delta'''\rangle \in \mathcal{T}'$, then $\langle\delta', \delta'''\rangle \in \mathcal{T}'$. So assume that $\langle\delta', \delta''\rangle \in \mathcal{T}'$ and $\langle\delta'', \delta'''\rangle \in \mathcal{T}'$. Case 1. If $\{\delta', \delta'', \delta'''\} \subseteq D\backslash\Phi$, then $\{\langle\delta', \delta''\rangle, \langle\delta'', \delta'''\rangle\} \subseteq \mathcal{T} \cap (D\backslash\Phi)^2$. Because $\mathcal{T}$ is transitive, $\langle\delta', \delta'''\rangle \in \mathcal{T} \cap (D\backslash\Phi)^2 \subseteq \mathcal{T}'$. Case 2. If $\{\delta', \delta'', \delta'''\} \subseteq \Phi\backslash\{\delta\}$, then $\{\langle\delta', \delta''\rangle, \langle\delta'', \delta'''\rangle\} \subseteq \mathcal{T} \cap (\Phi\backslash\{\delta\})^2$. Because $\mathcal{T}$ is transitive, $\langle\delta', \delta'''\rangle \in \mathcal{T} \cap (\Phi\backslash\{\delta\})^2 \subseteq \mathcal{T}'$. Case 3. If $\delta' \in D\backslash\Phi$ and $\delta''' \in \Phi$, then $\langle\delta', \delta'''\rangle$ because $D\backslash\Phi \times \Phi \subseteq \mathcal{T}'$. Case 4. Other cases are not possible because $D\backslash\Phi \times \Phi \subseteq \mathcal{T}'$.

   We show that if $\langle\delta', \delta''\rangle \notin \mathcal{T}'$ and $\delta' \neq \delta''$, then $\langle\delta'', \delta'\rangle \in \mathcal{T}'$. So assume $\langle\delta', \delta''\rangle \notin \mathcal{T}'$. Case 1. If $\{\delta', \delta''\} \subseteq D\backslash\Phi$ or $\{\delta', \delta''\} \subseteq \Phi\backslash\{\delta\}$, then because $\mathcal{T}$ is a strict total order, $\langle\delta'', \delta'\rangle \in \mathcal{T}$. Because $\mathcal{T} \cap (D\backslash\Phi)^2 \subseteq \mathcal{T}'$ and $\mathcal{T} \cap (\Phi\backslash\{\delta\})^2 \subseteq \mathcal{T}'$, $\langle\delta'', \delta'\rangle \in \mathcal{T}'$. Case 2. It is not possible that $\delta' = \delta$ and $\delta'' \in \Phi\backslash\{\delta\}$ because $\{\delta\} \times \Phi\backslash\{\delta\} \subseteq \mathcal{T}'$. Case 3. If $\delta' \in \Phi\backslash\{\delta\}$ and $\delta'' = \delta$, then obviously $\langle\delta'', \delta'\rangle \in \mathcal{T}'$ because $\{\delta\} \times \Phi\backslash\{\delta\} \subseteq \mathcal{T}'$. Case 4. If $\delta' \in \Phi$ and $\delta'' \in D\backslash\Phi$, then obviously $\langle\delta'', \delta'\rangle \in \mathcal{T}'$ because $D\backslash\Phi \times \Phi \subseteq \mathcal{T}'$. Case 5. It is not possible that $\delta' \in D\backslash\Phi$ and $\delta'' \in \Phi$, because $D\backslash\Phi \times \Phi \subseteq \mathcal{T}'$ and $\mathcal{T}'$ is asymmetric.

We show that $\mathcal{T}'$ is a $Q, \mathcal{P}$-characteristic ordering for E. Assume $\mathrm{appl}(\delta', F, E)$ for some $\delta' \in D$ and $F \in Q$. Because $\mathcal{T}$ is a $Q, \mathcal{P}$-characteristic ordering for E, there is $\delta''\mathcal{T}\delta'$ such that $\mathrm{appl}(\delta'', E, F)$. Because $\mathrm{appl}(\delta', E)$ iff $\mathrm{appl}(\delta', F)$ for all $\delta \in D\backslash\Phi$, $\{\delta', \delta''\} \subseteq \Phi$. The ordering of $\Phi\backslash\{\delta\}$ is the same in $\mathcal{T}'$ and $\mathcal{T}$, and therefore $\delta''\mathcal{T}'\delta'$ if $\delta'' \neq \delta$ and $\delta' \neq \delta$. It is not possible that $\delta' = \delta$ because the assumptions that $\mathrm{appl}(\delta', F, E)$ and that $\mathrm{appl}(\delta, F)$ implies $\mathrm{appl}(\delta, E)$ contradict it. If $\delta'' = \delta$, then $\delta''\mathcal{T}'\delta'$ because $\{\delta\} \times \Phi\backslash\{\delta\} \subseteq \mathcal{T}'$.

PROCEDURE continue($\Phi, \mathcal{P}, O, Q$)
BEGIN
  IF $\Phi = \emptyset$ *THEN RETURN* true
  ELSE
    BEGIN
      A := $\{\delta \in \Phi | \delta' \mathcal{P} \delta$ for no $\delta' \in \Phi, \delta \in O$ implies $\delta \notin \bigcup Q\}$;
      IF $A = \emptyset$ *THEN RETURN* false;
      δ := any member of A;
      Φ' := $\Phi \backslash \{\delta\}$;
      $\mathcal{P}'$:= $\mathcal{P} \cup (\{\delta\} \times \Phi')$;
      IF $\delta \in O$ *THEN RETURN* continue($\Phi', \mathcal{P}', O, Q$)
      ELSE RETURN continue($\Phi', \mathcal{P}', O, \{F \in Q | \delta \in F\}$)
    END
END

Figure 5.4: Subprocedure of the decision procedure

Therefore in all cases $\delta'' \mathcal{T}' \delta'$, and hence $\mathcal{T}'$ is a $Q, \mathcal{P}$-characteristic ordering for E. □

If F is a set of defaults and W a set of formulae, $\mathrm{Ext}_W(F)$ denotes the set $Cn(W \cup \{\gamma | \alpha{:}\sigma/\gamma \in F\})$. The idea is that F is the set of generating defaults of an extension $E = \mathrm{Ext}_W(F)$. If Q is a set of sets of defaults, then $\mathrm{Ext}_W(Q)$ denotes $\{\mathrm{Ext}_W(F) | F \in Q\}$.

Next we give the procedure *continue* and a lemma that characterizes its properties. The purpose of the procedure is to test whether it is possible that some of the extensions in the subtree of the current node of the search tree are $\mathcal{P}$-preferred. One of the arguments to the procedure is a set Q such that all extensions in $\mathrm{Ext}_W(Q)$ are $\mathcal{P}$-preferred. A sufficient condition for the extensions in the current subtree to not be $\mathcal{P}$-preferred is that they are not $\mathcal{P}$-preferred in $\mathrm{Ext}_W(Q)$. This is the test the procedure performs.

It is necessary to perform the test for preferredness only in the leaf nodes of the search tree, but it is useful to perform the test also in other nodes in order to prune the search tree. Consider a default theory $\Delta = \langle D, W \rangle$ and a strict partial order $\mathcal{P}$ on D such that $\delta \in D$ is the $\mathcal{P}$-least element. The root node of the search tree for Δ has two subtrees, one of which corresponds to extensions of Δ that apply δ and the other corresponds to extensions of Δ that do not apply δ. Let S be the set of preferred extensions in the first subtree, and let S be non-empty. Now the procedure *continue* detects that there can be no preferred extensions in the second subtree, because all strict total orders on D have δ as the least element. Hence the traversal of the second subtree is unnecessary.

The arguments to the procedure *continue* are the following. The procedure is called at each node of the search tree, and the defaults in $I \cup O \subseteq D$ represent the edges of the search tree between the current node and the root. If a default is in I it means that the default is applied in all extensions in the subtree for which the current node is the root, and if it is in O, it is applied in no extension in the subtree. The computation performed by *continue* treats defaults in $D\backslash(I \cup O)$ and defaults in I in an indistinguishable way, so the procedure does not take I as an argument. When *continue* is first called, the set Q contains the sets of generating defaults of the preferred extensions that correspond to the leaf nodes of the search tree left from the current node. The relation $\mathcal{P}$ is a strict partial order on a set of defaults D. The set Φ is the set of defaults D.

In the following, we restrict to the case when $D = I \cup O$, that is, when the procedure is called in a leaf node. In this case $D\backslash O$ is the set of generating defaults of an extension E. The computation performed by *continue* tries to find a $Q, \mathcal{P}$-characteristic ordering for E. Each recursive call to *continue* removes an element δ from Φ, possibly extends $\mathcal{P}$, and possibly removes some elements from Q. The following relations hold between $\mathcal{P}$, Φ and Q in all recursive calls to *continue*. The ordering $\mathcal{P}$ restricted to $D\backslash\Phi$ is a strict total order and all elements in $D\backslash\Phi$ are related to all elements in Φ by $\mathcal{P}$. Hence $\mathcal{P}$ is a strict total order on D when $\Phi = \emptyset$. For all $\delta \in D\backslash\Phi$ and $E' \in \mathrm{Ext}_W(Q)$, $\mathrm{appl}(\delta, E)$ iff $\mathrm{appl}(\delta, E')$. Whenever an element F is removed from Q, it means that the default δ is applied in E but not in $\mathrm{Ext}_W(F)$. This guarantees that if Φ becomes empty, $\mathcal{P}$ is a $Q, \mathcal{P}_0$-characteristic ordering for E, where $\mathcal{P}_0 \subseteq \mathcal{P}$ is a strict partial order on D and Q is the original value of Q. If there are no characteristic orderings for E, the computation of *continue* terminates before Φ becomes empty. In this case *false* is returned.

Lemma 5.0.15 (Procedure *continue*) *Let*

- $\Delta = \langle D, W\rangle$ *be a finite default theory,*
- $\mathcal{P}$ *a strict partial order on* D*,*
- $O \subseteq D$ *a set of defaults,*
- Q *a set of sets* F *such that* $Ext_W(F)$ *is an extension of* Δ *and* $F = GD(Ext_W(F), \Delta)$*,*
- $\Phi \subseteq D$ *such that* $\delta \in D\backslash O$ *if and only if* $\delta \in F$ *for all* $\delta \in D\backslash\Phi$ *and* $F \in Q$*, and*
- $D\backslash\Phi \times \Phi \subseteq \mathcal{P}$*.*

Then the following holds. ***I.*** *If for some $I \subseteq D\backslash O$, $E = Ext_W(I)$ is an extension of Δ that is $\mathcal{P}$-preferred in $Ext_W(Q)$ and $I = GD(E, \Delta)$, then the procedure call continue($\Phi, \mathcal{P}, O, Q$) returns* true. ***II.*** *If $Ext_W(D\backslash O)$ is an extension of Δ, $D\backslash O = GD(Ext_W(D\backslash O), \Delta)$, and $Ext_W(D\backslash O)$ is not $\mathcal{P}$-preferred in $Ext_W(Q)$, then the procedure call continue($\Phi, \mathcal{P}, O, Q$) returns* false.

Proof: The proof is by induction on $|\Phi|$. The base case corresponds to the *then*-part of the outermost *if*-statement of the procedure, and the inductive case to the *else*-part.

Base case $|\Phi| = 0$: In this case the procedure returns *true*. The first claim of the lemma is immediate. For the second claim assume that $\mathrm{Ext}_W(D\backslash O)$ is an extension of Δ. Because $\delta \in D\backslash O$ iff $\delta \in F$ for all $\delta \in D\backslash\Phi$ and $F \in Q$, $Q \subseteq \{\mathrm{Ext}_W(D\backslash O)\}$. Hence $\mathrm{Ext}_W(D\backslash O)$ is $\mathcal{P}$-preferred in Q because $\mathrm{appl}(\delta, \mathrm{Ext}_W(D\backslash O), F)$ for no $\delta \in D$ and $F \in \mathrm{Ext}_W(Q)$.

Inductive case $|\Phi| \geq 1$: First we show that the recursive calls fulfill the conditions stated in the lemma, and hence the induction assumption can be used. A recursive call is made if there is a $\mathcal{P}$-minimal element in Φ such that $\delta \in O$ implies $\delta \notin \bigcup Q$.

1. For all $\delta \in D\backslash\Phi'$, $\delta \in D\backslash O$ iff $\delta \in F$.

   By assumption $\delta' \in D\backslash O$ iff $\delta' \in F$ for all $F \in Q$. If $\delta \in O$, $\delta \notin F$ for all $F \in Q' = Q$, and if $\delta \notin O$, $\delta \in F$ for all $F \in Q' = \{F \in Q | \delta \in F\}$. Hence $\delta' \in D\backslash O$ iff $\delta' \in F$ for all $\delta' \in D\backslash\Phi'$ and $F \in Q'$.

2. $D\backslash\Phi' \times \Phi' \subseteq \mathcal{P}'$.

   Because $(D\backslash\Phi' \times \Phi')\backslash(D\backslash\Phi \times \Phi) \subseteq \{\delta\} \times \Phi'$ and $\{\delta\} \times \Phi' \subseteq \mathcal{P}'$ and $\mathcal{P} \subseteq \mathcal{P}'$ and by assumption $D\backslash\Phi \times \Phi \subseteq \mathcal{P}$, $D\backslash\Phi' \times \Phi' \subseteq \mathcal{P}'$.

I. Assume that continue($\Phi', \mathcal{P}', O, Q$) returns false. If $A = \emptyset$, then for all $\mathcal{P}$-minimal elements δ of Φ, $\delta \in O$ and $\delta \in F$ for some $F \in Q$. Let $\mathcal{T}$ be a strict total order such that $\mathcal{P} \subseteq \mathcal{T}$ and let E be an extension of Δ such that $\mathrm{GD}(E, \Delta) \subseteq D\backslash O$. Let δ be the $\mathcal{T}$-least element in Φ. Hence $\mathrm{appl}(\delta, \mathrm{Ext}_W(F), E)$ for some $F \in Q$. Let $\delta' \in D$ be any default such that $\delta' \mathcal{T} \delta$. Clearly $\delta' \in D\backslash\Phi$. By assumption $\delta' \in D\backslash O$ iff $\delta' \in F$. Hence if $\mathrm{appl}(\delta', E)$ then $\mathrm{appl}(\delta, \mathrm{Ext}_W(F))$. Hence $\mathrm{appl}(\delta', E, \mathrm{Ext}_W(F))$ for no $\delta' \mathcal{T} \delta$, and there are no $Q, \mathcal{P}$-characteristic orderings for E. Hence E is not $\mathcal{P}$-preferred in Q.

If $A \neq \emptyset$, then the procedure call continue($\Phi', \mathcal{P}', O, Q'$), where $Q' = Q$ if $\delta \in O$ and $Q' = \{F \in Q | \delta \in F\}$ if $\delta \notin O$, returns false. By the induction hypothesis, there is no extension E such that $\mathrm{GD}(E, \Delta) \subseteq D\backslash O$ and E is $\mathcal{P}'$-preferred in Q'. Because $Q' \subseteq Q$, there is no extension E such that $\mathrm{GD}(E, \Delta) \subseteq D\backslash O$ and E is $\mathcal{P}'$-preferred in Q.

Assume that E is an extension such that $\mathrm{GD}(E, \Delta) \subseteq D \backslash O$ and E is $\mathcal{P}$-preferred in Q. Hence there is a $Q, \mathcal{P}$-characteristic ordering $\mathcal{T}$ for E. Assume $\mathrm{appl}(\delta', \mathrm{Ext}_W(F), E)$ for some $\delta' \in D \backslash \Phi$ and $F \in Q$. Because $\mathrm{appl}(\delta', E)$ implies $\mathrm{appl}(\delta', \mathrm{Ext}_W(F))$ there is no $\delta'' \mathcal{T} \delta'$ such that $\mathrm{appl}(\delta', E, \mathrm{Ext}_W(F))$, and therefore $\mathcal{T}$ could not be a $Q, \mathcal{P}$-characteristic ordering for E. Hence $\mathrm{appl}(\delta', E)$ iff $\mathrm{appl}(\delta', \mathrm{Ext}_W(F))$ for all $\delta' \in D \backslash \Phi$ and $F \in Q$. Let δ be the $\mathcal{T}$-least element in Φ. Assume $\delta \notin \mathrm{GD}(E, \Delta)$ implies $\delta \notin \bigcup Q$. Then by Lemma 5.0.14 there is a $Q, \mathcal{P}$-characteristic ordering $\mathcal{T}'$ for E such that $\{\delta\} \times \Phi \backslash \{\delta\} \subseteq \mathcal{T}'$. Because $\mathcal{P} \subseteq \mathcal{T}'$ and $\{\delta\} \times \Phi \backslash \{\delta\} \subseteq \mathcal{T}'$, $\mathcal{P}' \subseteq \mathcal{T}$. This contradicts the inexistence of $Q, \mathcal{P}'$-characteristic orderings for E.

Assume $\delta \in F \backslash \mathrm{GD}(E, \Delta)$ for some $F \in Q$. Because for all $\delta' \in D \backslash \Phi$ $\mathrm{appl}(\delta', E)$ iff $\mathrm{appl}(\delta', \mathrm{Ext}_W(F))$, $\mathrm{appl}(\delta', E, \mathrm{Ext}_W(F))$ for no $\delta' \mathcal{T} \delta$. Then there are no $Q, \mathcal{P}$-characteristic orderings for E.

In all cases it was shown that there are no $Q, \mathcal{P}$-characteristic orderings for E. Hence E is not $\mathcal{P}$-preferred in Q.

II. Assume that $\mathrm{Ext}_W(D \backslash O)$ is an extension of Δ and $\mathrm{continue}(\Phi, \mathcal{P}, O, Q)$ returns *true*. Hence $A \neq \emptyset$, and there is a $\mathcal{P}$-minimal δ in Φ such that $\delta \in O$ implies $\delta \notin \bigcup Q$, and the execution proceeds after the first *if*-statement in the *else*-part of the outermost *if*-statement of the procedure. We analyze by cases the last innermost *if*-statement.

Assume $\delta \in O$. By the induction hypothesis and because $\mathrm{continue}(\Phi', \mathcal{P}', O, Q)$ returns *true*, $\mathrm{Ext}_W(D \backslash O)$ is $\mathcal{P}'$-preferred in $\mathrm{Ext}_W(Q)$. Hence there is an $\mathrm{Ext}_W(Q), \mathcal{P}'$-characteristic ordering $\mathcal{T}$ for $\mathrm{Ext}_W(D \backslash O)$. Now $\mathcal{P} \subseteq \mathcal{T}$ because $\mathcal{P} \subseteq \mathcal{P}'$ and $\mathcal{P}' \subseteq \mathcal{T}$, and hence $\mathcal{T}$ is an $\mathrm{Ext}_W(Q), \mathcal{P}$-characteristic ordering for $\mathrm{Ext}_W(D \backslash O)$. Therefore $\mathrm{Ext}_W(D \backslash O)$ is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(Q)$.

Assume $\delta \notin O$. Because $\mathrm{continue}(\Phi', \mathcal{P}', O, \{F \in Q | \delta \in F\})$ returns *true* and by the induction hypothesis, the extension $\mathrm{Ext}_W(D \backslash O)$ is $\mathcal{P}'$-preferred in $\mathrm{Ext}_W(\{F \in Q | \delta \in F\})$. Hence there is an $\mathrm{Ext}_W(\{F \in Q | \delta \in F\}), \mathcal{P}'$-characteristic ordering $\mathcal{T}$ for $\mathrm{Ext}_W(D \backslash O)$. Now $\mathcal{P} \subseteq \mathcal{T}$ because $\mathcal{P} \subseteq \mathcal{P}'$, and $\mathcal{T}$ is an $\mathrm{Ext}_W(\{F \in Q | \delta \in F\}), \mathcal{P}$-characteristic ordering for $\mathrm{Ext}_W(D \backslash O)$. Assume that for some $\delta' \in D$ and $F \in Q$, $\mathrm{appl}(\delta', \mathrm{Ext}_W(F), \mathrm{Ext}_W(D \backslash O))$. If $F \in \{F \in Q | \delta \in F\}$, then $\delta \in F$ and as $\mathcal{T}$ is $\mathrm{Ext}_W(\{F \in Q | \delta \in F\}), \mathcal{P}$-characteristic ordering for $\mathrm{Ext}_W(D \backslash O)$, there is $\delta'' \in D$ such that $\delta'' \mathcal{T} \delta'$ and $\mathrm{appl}(\delta'', \mathrm{Ext}_W(D \backslash O), \mathrm{Ext}_W(F))$. If $F \notin \{F \in Q | \delta \in F\}$, that is, $\delta \notin F$, then $\mathrm{appl}(\delta, \mathrm{Ext}_W(D \backslash O), \mathrm{Ext}_W(F))$. Because δ is $\mathcal{P}'$-least in Φ by definition of $\mathcal{P}'$, and $\delta' \in F$ iff $\delta' \in D \backslash O$ for all $\delta' \in D \backslash \Phi$, $\delta \mathcal{T} \delta'$. Therefore $\mathrm{Ext}_W(D \backslash O)$ is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(Q)$. $\square$

When going to a child node from a parent node during the traversal of the search tree, the default that corresponds to the child node is chosen by the subprocedure *next*. The variables used in the procedure are the same as the

$PROCEDURE$ next$(\Phi, \mathcal{P}, D, I, O, Q)$
$BEGIN$
  A := $\{\delta \in \Phi | \delta' \mathcal{P} \delta$ for no $\delta' \in \Phi, \delta \in O$ implies $\delta \notin \bigcup Q\}$;
  $IF\ A = \emptyset\ THEN\ RETURN$ any $\mathcal{P}$-minimal member of $D \backslash (I \cup O)$
  $ELSE$
    $BEGIN$
      $B := A \backslash (I \cup O)$;
      $IF\ B \neq \emptyset\ THEN\ RETURN$ any member of B;
      $\delta :=$ any member of A;
      $\Phi' := \Phi \backslash \{\delta\}$;
      $\mathcal{P}' := \mathcal{P} \cup (\{\delta\} \times \Phi')$;
      $IF\ \delta \in O\ THEN\ RETURN$ next$(\Phi', \mathcal{P}', D, I, O, Q)$
      $ELSE\ RETURN$ next$(\Phi', \mathcal{P}', D, I, O, \{F \in Q | \delta \in F\})$
    END
END

Figure 5.5: Another subprocedure for the decision procedure

ones used in the procedure *continue* (with the exception of $I \subseteq D \backslash O$), and the arguments in the recursive calls are the same. The procedure returns a $\mathcal{P}$-minimal default in $D \backslash (I \cup O)$. The arguments to *next* and the idea behind the way defaults are chosen have been explained earlier.

Lemma 5.0.16 (Procedure *next*) *Let*

1. *$\Delta = \langle D, W \rangle$ be a finite default theory,*

2. *$\mathcal{P}$ a strict partial order on D,*

3. *$Q \subseteq 2^D$ a finite set of sets $F \subseteq D$ such that $Ext_W(F)$ is an extension of Δ and $F = GD(Ext_W(F), \Delta)$,*

4. *I and O sets of defaults such that $I \cap O = \emptyset$ and $I \cup O \subset D$ and $\delta' \in I \cup O$ whenever $\delta' \mathcal{P} \delta''$ and $\delta'' \in I \cup O$,*

5. *$\Phi \subseteq D$ a non-empty set such that $D \backslash \Phi \times \Phi \subseteq \mathcal{P}$,*

6. *for all $\delta \in D \backslash \Phi$, $\delta \in I$ if and only if $\delta \in F$, and $\delta \in O$ if and only f $\delta \notin F$, and*

7. *$\delta \in D \backslash (I \cup O)$ the value returned by next$(\Phi, \mathcal{P}, D, I, O, Q)$.*

Then all $E \in Ext_W(S)$ are $\mathcal{P}$-preferred in $Ext_W(S \cup S' \cup Q)$, where $S \subseteq 2^D$ is a set of sets $F \subseteq D$ such that $I \cup \{\delta\} \subseteq F$ and $F \cap O = \emptyset$ and $Ext_W(F)$ is an extension of Δ that is $\mathcal{P}$-preferred in $Ext_W(S \cup Q)$ and $F = GD(Ext_W(F), \Delta)$,

$S' \subseteq 2^D$ is a set of sets $F \subseteq D$ such that $I \subseteq F$ and $F \cap (O \cup \{\delta\}) = \emptyset$ and $Ext_W(F)$ is an extension of Δ that is $\mathcal{P}$-preferred in $Ext_W(S \cup S' \cup Q)$ and $F = GD(Ext_W(F), \Delta)$.

Proof: **Lemma C** The set Φ is upward closed, that is, whenever $\delta' \in \Phi$ and $\delta'\mathcal{P}\delta''$, also $\delta'' \in \Phi$.
Proof: Assume $\delta'' \in D \backslash \Phi$. Because $D\backslash\Phi \times \Phi \subseteq \mathcal{P}$, now $\delta''\mathcal{P}\delta'$. This contradicts the asymmetry of $\mathcal{P}$. □

The proof is by induction on the cardinality of Φ. The base case corresponds to the *then*-part of the outermost *if*-statement, and the inductive case to the *else*-part. Our induction begins from sets Φ with cardinality 1.

Base case $|\Phi| = 1$: In this case A is a singleton $\{\delta\}$ or the empty set. If A is non-empty, $\{\delta\} = D\backslash(I \cup O)$, because $D\backslash\{\delta\} \subseteq I \cup O$ and $I \cup O \subset D$. Hence in this case $B = A = \{\delta\}$. The proof for the base case continues as in i and ii in the inductive case (i if $A = \emptyset$, and ii otherwise.)

Inductive case $|\Phi| \geq 1$: We analyze the three cases where A and $B \subseteq A$ are empty or non-empty. The cases correspond to the *if* statements of the procedure.

(i) Let A be the empty set. Assume that E is an extension such that $\text{appl}(\delta', E)$ for all $\delta' \in I$ and no $\delta' \in O$. Let $\mathcal{T}$ be any strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$. Let δ be the $\mathcal{T}$-least default in Φ, and hence $\delta \in D\backslash\Phi$ for all $\delta'\mathcal{T}\delta$. Because $A = \emptyset$, $\delta' \in O$ implies $\delta' \notin \bigcup Q$ for no $\mathcal{P}$-minimal default δ' in Φ. Therefore $\delta \in O$ and $\delta \in F$ for some $F \in Q$, and consequently $\text{appl}(\delta, \text{Ext}_W(F), E)$. By assumptions $\text{appl}(\delta', \text{Ext}_W(F))$ iff $\text{appl}(\delta', E)$ for all $\delta'\mathcal{T}\delta$. Hence there is no default δ' such that $\delta'\mathcal{T}\delta$ and $\text{appl}(\delta', E, \text{Ext}_W(F))$. This shows that there are no $\text{Ext}_W(Q), \mathcal{P}$-characteristic orderings for E. Hence no extension E such that $\text{appl}(\delta', E)$ for all $\delta' \in I$ and not $\text{appl}(\delta', E)$ for all $\delta' \in O$ can be $\mathcal{P}$-preferred in $\text{Ext}_W(Q)$. Therefore the set S is empty and any member of $D\backslash(I \cup O)$ can be returned without violating the claim of the lemma.

(ii) Let $B \subseteq A$ be non-empty. In this case the default δ returned belongs to B. Let E be any member of $\text{Ext}_W(S)$. By assumption, E is $\mathcal{P}$-preferred in $\text{Ext}_W(S \cup Q)$. Let $\mathcal{T}$ be an $\text{Ext}_W(S \cup Q), \mathcal{P}$-characteristic ordering for E. By Lemma 5.0.14 there is a $\text{Ext}_W(S \cup Q), \mathcal{P}$-characteristic ordering $\mathcal{T}'$ for E such that $\{\delta\} \times \Phi\backslash\{\delta\} \subseteq \mathcal{T}'$. We show that $\mathcal{T}'$ is an $\text{Ext}_W(S \cup S' \cup Q), \mathcal{P}$-characteristic ordering for E. Let E' be a member of $\text{Ext}_W(S')$ and δ' a member of D and $\text{appl}(\delta', E', E)$. By assumption $\text{appl}(\delta'', E)$ iff $\text{appl}(\delta'', E')$ for all $\delta'' \in D\backslash\Phi$. Hence $\delta' \in \Phi$, and because $\text{appl}(\delta, E)$, $\delta' \neq \delta$. Because $\{\delta\} \times (\Phi\backslash\{\delta\}) \subseteq \mathcal{T}'$, $\delta\mathcal{T}'\delta'$. Therefore $\mathcal{T}'$ is an $\text{Ext}_W(S \cup S' \cup Q), \mathcal{P}$-characteristic ordering for E.

(iii) Let B be empty and A non-empty. We analyze by cases the last *if*-statement of the procedure. In both branches of the *if* statement a recursive call to *next* is made. We verify that the assumptions of this lemma hold for this recursive call, so that we can use the induction hypothesis. The arguments that are the same in the recursive call are D, I and O, and the arguments that are different are Q, Φ and $\mathcal{P}$. Let Q', Φ' and $\mathcal{P}'$ be the values of these arguments the recursive call of *next* receives.

- $\mathcal{P}' = \mathcal{P} \cup (\{\delta\} \times \Phi \backslash \{\delta\})$ is a strict partial order on D.

  The relation $\mathcal{P}'$ is asymmetric because $\mathcal{P}$ is asymmetric and as δ is $\mathcal{P}$-minimal in Φ $\langle \delta', \delta \rangle \in \mathcal{P}$ for no $\delta' \in \Phi \backslash \{\delta\}$.

  The relation $\mathcal{P}'$ is transitive because of the following. Clearly, any $\langle \delta', \delta'' \rangle \in \mathcal{P}'$ and $\langle \delta'', \delta''' \rangle \in \mathcal{P}'$ such that $\delta \notin \{\delta', \delta'', \delta'''\}$ are in $\mathcal{P}$, and hence $\langle \delta', \delta'' \rangle \in \mathcal{P}'$ as $\mathcal{P} \subseteq \mathcal{P}'$ and $\mathcal{P}$ is transitive.

  Assume $\langle \delta, \delta' \rangle \in \mathcal{P}'$ and $\langle \delta', \delta'' \rangle \in \mathcal{P}'$. Because $\delta \in \Phi$ and by Lemma **C** Φ is upward closed, $\{\delta', \delta''\} \in \Phi \backslash \{\delta\}$ and hence $\langle \delta', \delta'' \rangle \in \mathcal{P}$. Because $\{\delta\} \times \Phi \backslash \{\delta\} \subseteq \mathcal{P}'$, $\langle \delta, \delta'' \rangle \in \mathcal{P}'$.

  Assume $\langle \delta', \delta \rangle \in \mathcal{P}'$ and $\langle \delta, \delta'' \rangle \in \mathcal{P}'$. By Lemma **C** $\delta'' \in \Phi$. Clearly $\langle \delta', \delta \rangle \in \mathcal{P}$. Because δ is $\mathcal{P}$-minimal in Φ, $\delta' \in D \backslash \Phi$. By assumption $D \backslash \Phi \times \Phi \subseteq \mathcal{P}$, hence $\langle \delta', \delta'' \rangle \in \mathcal{P} \subseteq \mathcal{P}'$. Assume $\langle \delta', \delta'' \rangle \in \mathcal{P}'$ and $\langle \delta'', \delta \rangle \in \mathcal{P}'$. Clearly $\langle \delta', \delta'' \rangle \in \mathcal{P}$ and $\langle \delta'', \delta \rangle \in \mathcal{P}$. Hence $\langle \delta', \delta \rangle \in \mathcal{P} \subseteq \mathcal{P}'$.

- $Q' \subseteq 2^D$ is a finite set of sets $F \subseteq D$ such that $E = \mathrm{Ext}_W(F)$ is an extension of Δ and $F = \mathrm{GD}(E, \Delta)$, where $Q' = Q$ if $\delta \in O$ and otherwise $Q' = \{F \in Q | \delta \in F\}$.

  This is simply because Q' is a subset of Q and Q satisfies the same properties.

- $\Phi' = \Phi \backslash \{\delta\}$ is a non-empty subset of D such that $D \backslash \Phi' \times \Phi' \subseteq \mathcal{P}'$.

  By assumption $|\Phi| \geq 2$. Assume $\langle \delta', \delta'' \rangle \in D \backslash \Phi' \times \Phi'$. If $\delta' \in D \backslash \Phi$ and $\delta'' \in \Phi$, then because $D \backslash \Phi \times \Phi \subseteq \mathcal{P} \subseteq \mathcal{P}'$, also $\langle \delta', \delta'' \rangle \in \mathcal{P}'$. Otherwise, because $\Phi' = \Phi \backslash \{\delta\}$, $\delta' = \delta$. Because $\delta \in \Phi$ and Φ is upward closed by Lemma **C**, $\delta'' \in \Phi$. Because $\mathcal{P}'$ is irreflexive (as a consequence of being transitive and asymmetric), $\delta'' \in \Phi \backslash \{\delta\}$. Because $\{\delta\} \times \Phi \backslash \{\delta\} \subseteq \mathcal{P}'$, $\langle \delta', \delta'' \rangle \in \mathcal{P}'$.

- For all $\delta' \in D \backslash \Phi'$ and $F \in Q'$, $\delta' \in I$ iff $\delta' \in F$, and $\delta' \in O$ iff $\delta' \notin F$.

  For $\delta' \in D \backslash \Phi$ this is by assumption, and it remains to show the equivalences for δ. By construction $\delta \in O$ implies $\delta \notin \bigcup Q$. Now either $\delta \in O$ and $\delta \notin \bigcup Q$ and $Q' = Q$, or $\delta \notin O$ and $Q' = \{F \in Q | \delta \in F\}$. Hence for all $F \in Q'$, $\delta \in O$ iff $\delta \notin F$.

PROCEDURE decide($\mathcal{P}, D, W, I, O, \phi, Q$)
BEGIN
  IF continue($D, \mathcal{P}, O, Q$) $=$ false *THEN RETURN* ($\emptyset$,false);
  $E := W \cup \{\gamma | \alpha{:}\sigma/\gamma \in I\}$;
  IF $I \cup O = D$ *THEN*
    IF extension(D, W, E) $=$ false
      or $E \models \neg\beta_i$ for some $\alpha{:}\beta_1, \ldots, \beta_n/\gamma \in I$ and $i \in \{1, \ldots, n\}$
    THEN RETURN ($\emptyset$,false)
    ELSE RETURN ($\{I\}$,test(E, ϕ));
  $\delta :=$ next($D, \mathcal{P}, D, I, O, Q$);
  $(S, b) :=$ decide($\mathcal{P}, D, W, I \cup \{\delta\}, O, \phi, Q$);
  $(S', b') :=$ decide($\mathcal{P}, D, W, I, O \cup \{\delta\}, \phi, S \cup Q$);
  RETURN ($S \cup S', b$ or b')
END

Figure 5.6: The decision procedure for prioritized default logic

By construction δ is a member of the set $A \subseteq D$. Now $A \backslash (I \cup O) = \emptyset$. Otherwise the procedure *next* would have returned a member of $A \backslash (I \cup O)$. Hence $A \subseteq I \cup O$ and $\delta \in I \cup O$. Because $\delta \in I \cup O$ and $I \cap O = \emptyset$, $\delta \in O$ iff $\delta \notin I$. Because $\delta \in O$ iff $\delta \notin F$ as shown above, finally $\delta \in I$ iff $\delta \in F$.

Assume $\delta \notin O$. The value δ' returned is obtained from the recursive call to next($\Phi', \mathcal{P}', D, I, O, \{F \in Q | \delta \in F\}$). Take any E in $\mathrm{Ext}_W(S)$. Now by induction hypothesis E is $\mathcal{P}'$-preferred in $\mathrm{Ext}_W(S \cup S' \cup \{F \in Q | \delta \in F\})$, that is, there is an $\mathrm{Ext}_W(S \cup S' \cup \{F \in Q | \delta \in F\}), \mathcal{P}'$-characteristic ordering $\mathcal{T}$ for E. Because $\mathcal{P} \subseteq \mathcal{P}'$, also $\mathcal{P} \subseteq \mathcal{T}$. By definition of $\mathcal{P}'$, δ is the $\mathcal{T}$-least element of Φ. We show that E is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S \cup S' \cup Q)$. The missing part is that $\mathcal{T}$ is an $\mathrm{Ext}_W(Q \backslash \{F \in Q | \delta \in F\}), \mathcal{P}$-characteristic ordering for E. The rest is directly by induction hypothesis. Let E' be any member of $\mathrm{Ext}_W(Q \backslash \{F \in Q | \delta \in F\})$. Let δ'' be any default in D such that appl(δ'', E', E). Because appl(δ, E, E') and for all defaults $\delta''' \mathcal{T}' \delta$ appl(δ''', E) iff appl(δ''', E'), $\delta \mathcal{T} \delta''$ holds. Hence $\delta \mathcal{T} \delta''$ and appl(δ, E, E'). Therefore E is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S \cup S' \cup Q)$.

Assume $\delta \in O$. The value δ' returned is from next($\Phi', \mathcal{P}', D, I, O, Q$). By induction hypothesis, every $E \in \mathrm{Ext}_W(S)$ is $\mathcal{P}'$-preferred in $\mathrm{Ext}_W(S \cup S' \cup Q)$. Hence there is an $\mathrm{Ext}_W(S \cup S' \cup Q), \mathcal{P}'$-characteristic ordering $\mathcal{T}$ for E. Because $\mathcal{P} \subseteq \mathcal{P}'$, also $\mathcal{P} \subseteq \mathcal{T}$. Hence E is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S \cup S' \cup Q)$. □

Lemma 5.0.17 (Procedure *decide*) *Let $\Delta = \langle D, W \rangle$ be a finite default theory, $\mathcal{P}$ a strict partial order on D, I and O subsets of D such that $I \cap O = \emptyset$,*

$\phi \in \mathcal{L}$ a formula, and Q a finite set of sets of defaults such that for all $F \in Q$ $Ext_W(F)$ is an extension of Δ and $F = GD(Ext_W(F), \Delta)$. Let test *be a procedure that takes a set of formulae and a formula as its arguments.*

Then the procedure call decide($\mathcal{P}, D, W, I, O, \phi, Q$) returns (S, b). The set $Ext_W(S)$ is the set of extensions of Δ that are $\mathcal{P}$-preferred in $Ext_W(S \cup Q)$, and for each $F \in S$, $I \subseteq F$ and $F \cap O = \emptyset$ and $F = GD(Ext_W(F), \Delta)$. The value b is true *if and only if for some $E \in Ext_W(S)$, test($\{\gamma | \alpha{:}\sigma/\gamma \in GD(E, \Delta)\}, \phi$) returns* true.

Proof: The proof is by induction on $|D \backslash (I \cup O)|$.

Base case $|D \backslash (I \cup O)| = 0$: In this case the procedure returns ($\emptyset$,false) if and only if the procedure call extension(D, W, E) returns false or $E \models \neg\beta$ for some $\alpha{:}\sigma/\gamma \in I$ and β in σ, or the procedure call continue($D, \mathcal{P}, O, Q$) returns false. By Lemma 4.0.21 extension(D, W, E) returns true if and only if $Cn(E)$ is an extension of the default theory $\langle D, W \rangle$. By Lemma 5.0.15, if $D = I \cup O$ the procedure call continue($D, \mathcal{P}, O, Q$) returns *true* if and only if $\mathrm{Ext}_W(I)$ is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(Q)$. Hence the procedure *decide* returns $(\{I\}, b)$ if and only if $E = \mathrm{Ext}_W(I)$ is a $\mathcal{P}$-preferred extension of $\langle D, W \rangle$ and $I = \mathrm{GD}(E, \Delta)$.

Inductive case $|D \backslash (I \cup O)| \geq 1$: If the call to the procedure *continue* returns false, then by Lemma 5.0.15 there are no extensions E of $\langle D, W \rangle$ such that $I \subseteq \mathrm{GD}(E, \Delta)$ and $O \cap \mathrm{GD}(E, \Delta) = \emptyset$ and E is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(Q)$. Therefore it is correct for the procedure *decide* to return ($\emptyset$,false). The condition of the second *if*-statement in the procedure is not false because $D \backslash (I \cup O)$ is non-empty, and hence the execution continues at the assignment statement that calls the procedure *next*. Consider the values assigned to the variables δ, S, S', b, b' in the three assignment statements and the values $S \cup S'$ and (b or b') returned by the procedure.

By induction hypothesis all $\mathcal{P}$-preferred extensions $E = \mathrm{Ext}_W(B)$ of Δ such that $I \cup \{\delta\} \subseteq B$ and $B \cap O = \emptyset$, are in S. Similarly all extensions $E' = \mathrm{Ext}_W(B')$ of Δ such that $I \subseteq B'$ and $B' \cap (O \cup \{\delta\}) = \emptyset$ are in S'. Hence all extensions $\mathrm{Ext}_W(B)$ such that $B \subseteq I$ and $B \cap O = \emptyset$, are in $S \cup S'$.

By induction hypothesis extensions in $\mathrm{Ext}_W(S)$ are $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S \cup Q)$ and extensions in $\mathrm{Ext}_W(S')$ are $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S' \cup S \cup Q)$. By Lemma 5.0.16 extensions in $\mathrm{Ext}_W(S)$ are $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S \cup S' \cup Q)$. Therefore all extensions in $\mathrm{Ext}_W(S \cup S')$ are $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S \cup S' \cup Q)$.

By induction hypotheses b is *true* if and only if for some $F \in S$, test($\{\gamma | \alpha{:}\sigma/\gamma \in F\}, \phi$) returns *true*, and b' is *true* if and only if for some $F \in S'$, test($\{\gamma | \alpha{:}\sigma/\gamma \in F\}, \phi$) returns *true*. Therefore (b or b') is *true* if and only if for some $F \in S \cup S'$ test($\{\gamma | \alpha{:}\sigma/\gamma \in F\}, \phi$) returns *true*. □

Theorem 5.0.18 (Correctness) *Let $\Delta = \langle D, W\rangle$ be a finite default theory, $\mathcal{P}$ a strict partial order on D, and $\phi \in \mathcal{L}$ a formula. The procedure call decide($\mathcal{P}, D, W, \emptyset, \emptyset, \phi, \emptyset$) returns (S, b), where b is* true *if and only if there is a $\mathcal{P}$-preferred extension E of Δ such that the procedure call test($\{\gamma|\alpha{:}\sigma/\gamma \in GD(E, \Delta)\}, \phi$) returns* true.

Proof: By Lemma 5.0.17 $\mathrm{Ext}_W(S)$ contains all $\mathcal{P}$-preferred extensions. Because each member of $\mathrm{Ext}_W(S)$ is $\mathcal{P}$-preferred in $\mathrm{Ext}_W(S)$, all its members are $\mathcal{P}$-preferred by Lemma 5.0.13. By Lemma 5.0.17 the value of b is *true* if and only if for an extension E in $\mathrm{Ext}_W(S)$ $\mathrm{test}(\{\gamma|\alpha{:}\sigma/\gamma \in \mathrm{GD}(E, \Delta)\}, \phi)$ returns *true*. $\square$

The procedure *decide* can be used for testing $\Delta \models_{\mathcal{P}} \phi$ by calling it with the argument *test* assigned to a procedure that tests $E \not\models \phi$. Then *decide* returns true if there is an extension of Δ that does not contain ϕ. Hence if *decide* returns false, ϕ is contained in all extensions of Δ.

During the traversal of the space of all extensions, the procedure maintains the set Q of sets of generating defaults of the preferred extensions found so far. Because the number of preferred extensions of a default theory may be exponential on the size of the default theory, the size of the list may be exponential in the worst case. Hence without further modifications the procedure is practical only in cases where there are only a polynomial number of preferred extensions or the number of preferred extensions is small. More research should be invested on finding ways to reduce the size of Q.

Chapter 6

Related work on priorities

Priorities have earlier been introduced to default logic, autoepistemic logic, and different forms of minimal model semantics for nonmonotonic reasoning. The first formalization of priorities in default logic is due to Brewka [1991]. He views priorities as an ordering in which default rules are applied. For arbitrary default theories not all orderings lead to extensions, and therefore Brewka considers only prerequisite-free normal default theories. Default logic restricted to prerequisite-free normal default theories is equivalent to a simpler formalization of nonmonotonic reasoning based on maximal consistent subsets of sets of formulae by Poole [1985]. Brewka's definition of *preferred subtheories* [1989] extends Poole's work with priorities. Later, the restriction to prerequisite-free normal defaults has been removed by Baader and Hollunder [1993], Marek and Truszczyński [1993], Brewka [1994], and MacNish [1991]. These formalizations coincide with our prioritized default logic for prerequisite-free normal default theories but otherwise they fundamentally differ.

Many other formalizations of explicit priorities based on minimal models are equivalent or closely related to prioritized default logic with prerequisite-free normal defaults. This is the case – as pointed out above – with preferred subtheories [Brewka, 1989], ordered theory presentations [Ryan, 1992a], propositional prioritized circumscription [McCarthy, 1986; Lifschitz, 1985], and the model minimization used in conditional entailment [Geffner and Pearl, 1992]. For all these systems we present an embedding in prioritized default logic. The main differences between these systems lie in the way partiality in priorities is interpreted, and the class of strict partial orders supported. Of lesser importance are the differences in the way they are formalized.

6.1 Priorities in default logic

6.1.1 Hierarchical default logic

MacNish [1991] defines a hierarchical version of default logic. The basic idea of hierarchical default logic resembles the idea behind Konolige's hierarchic autoepistemic logic [Konolige, 1988]. MacNish considers a totally ordered set of sets of default rules, instead of partially ordered sets of sets of formulae. The definition of extensions is based on iterating Reiter's [1980] definition of extensions.

Definition 6.1.1 *Let D be a set of default rules. Then a* transition relation *R_D between sets of formulae is defined as WR_DW' if and only if W' is an extension of $\langle D, W\rangle$. A* hierarchical parameter set *is a tuple $H_D = \langle\{D_1, \ldots, D_n\}, \preceq\rangle$ where D_i are sets of defaults and $\preceq$ is a total order on $\{D_1, \ldots, D_n\}$.*

A hierarchical parameter set expresses the priorities of the default rules in $D = \bigcup_{i\in\{1,\ldots,n\}} D_i$: whenever $D \preceq D'$, the default rules in D' have a higher priority than the default rules in D.

Definition 6.1.2 *Let $H_D = \langle\{D_1, \ldots, D_n\}, \preceq\rangle$ be a hierarchical parameter set, and let $D = \bigcup_{i\in\{1,\ldots,n\}} D_i$. Prioritized extensions of $\langle D, W\rangle$ are extension E such that WR_DE where $R_D = R_{D_n} \circ R_{D_{n-1}} \circ \cdots \circ R_{D_2} \circ R_{D_1}$.*

Unlike in hierarchic autoepistemic logic, in hierarchical default logic conflicting defaults of an equal priority do not cause an inconsistency. However, MacNish's extensions are not extensions as defined by [Reiter, 1980], and this may be a disadvantage in some cases.

Example 6.1.1 Consider $\Delta = \langle D, W\rangle$ where $D = D_1 \cup D_2$, $D_1 = \{a{:}b/b\}$, $D_2 = \{b{:}c/c\}$, $W = \{a\}$, and $H_D = \langle\{D_1, D_2\}, \preceq\rangle$ is a hierarchical parameter set where $D_1 \preceq D_2$. Now $WR_{D_2}W_1$ where $W_1 = W = Cn(\{a\})$, and $W_1R_{D_1}W_2$ where $W_2 = Cn(\{a, b\})$, and $R_D = R_{D_2} \circ R_{D_1} = \{\langle Cn(\{a\}), Cn(\{a, b\})\rangle\}$. Now WR_DW_2 but W_2 is not an extension of Δ because the default rule $b{:}c/c$ – although applicable – is not applied. ■

The problem illustrated in the above example arises whenever the prerequisite of a default is derived from a default of a lower priority. This is not possible if the default rules are prerequisite-free. Therefore for prerequisite-free normal default theories MacNish's extensions fulfill Reiter's definition of extensions.

6.1.2 Prioritized default logics by Brewka, Baader and Hollunder, and Marek and Truszczyński

Brewka [1994], Baader and Hollunder [1993], and Marek and Truszczyński [1993] introduce priorities to default logic by modifying the semiconstructive definition of extensions (Theorem 2.1.2.) Priorities are expressed as strict partial orders on defaults, and the semiconstructive definition is modified to select defaults in an order that agrees with the partial order. The definition of Baader and Hollunder is given below. The definitions by Brewka and Marek and Truszczyński are similar but work differently in some cases.

Definition 6.1.3 ([Baader and Hollunder, 1993]) *Let $\Delta = \langle D, W\rangle$ be a default theory and $\mathcal{B}$ a strict partial order on D. Let E be a set of formulae. Define E_i for all $i \geq 0$ as follows.*

$$
\begin{aligned}
E_0 &= W \\
E_{i+1} &= E_i \cup \left\{ \gamma \,\middle|\, \frac{\alpha : \beta_1, \ldots, \beta_n}{\gamma} \in D, E_i \models \alpha, \{\neg\beta_1, \ldots, \neg\beta_n\} \cap E = \emptyset, \right. \\
&\qquad E_i \models \alpha' \text{ and } E_i \not\models \gamma' \text{ and } E_i \not\models \neg\beta' \text{ for no} \\
&\qquad \left. \frac{\alpha' : \beta'_1, \ldots, \beta'_{n'}}{\gamma'} \mathcal{B} \frac{\alpha : \beta'_1, \ldots, \beta'_n}{\gamma} \text{ and } \beta' \in \{\beta'_1, \ldots, \beta'_{n'}\} \right\}
\end{aligned}
$$

Then E is an S-extension if $E = \bigcup_{i \geq 0} Cn(E_i)$.

The above definition can be summarized by saying that the preferred extensions are computed in a forward chaining manner, and whenever there is more than one default rule that can be applied, those with a highest priority are applied. Note that the way extensions are computed does not guarantee that the highest priority default – if it exists – is applied at some point of the computation. The following example shows this.

Example 6.1.2 ([Brewka, 1994]) Consider the default theory $\Delta = \langle D, W\rangle$ where $W = \{a\}$ and $D = \{b{:}c/c, a{:}\neg c/\neg c, a{:}b/b\}$. Define the relation

$$
\mathcal{B} = \left\{ \left\langle \frac{b : c}{c}, \frac{a : \neg c}{\neg c} \right\rangle, \left\langle \frac{a : \neg c}{\neg c}, \frac{a : b}{b} \right\rangle, \left\langle \frac{b : c}{c}, \frac{a : b}{b} \right\rangle \right\}.
$$

According to Definition 6.1.3, the following sets E_i are obtained.

$$
\begin{aligned}
E_0 &= \{a\} \\
E_1 &= \{a, \neg c\} \\
E_2 &= \{a, \neg c, b\} \\
E_i &= E_2 \text{ for all } i \geq 3
\end{aligned}
$$

In other words, initially the highest priority default $b{:}c/c$ is not applicable, and hence the second default $a{:}\neg c/\neg c$ is applied, and $\neg c$ is obtained. The highest priority default is still not applicable, and hence the third default $a{:}b/b$ is applied, and b is obtained. Now $b{:}c/c$ were applicable if the contradicting default $b{:}c/c$ would not have been applied first. Arguably, this is unintuitive, as the prerequisite of the highest priority default $b{:}c/c$ belongs to all extensions, but despite this fact the highest priority default does not become applicable because it is blocked by a default with a lower priority. ■

The property demonstrated by the above example is possessed by all three versions of prioritized default logic mentioned [Baader and Hollunder, 1993; Marek and Truszczyński, 1993; Brewka, 1994]. The application of high priority defaults may depend on the application of low priority defaults in unobvious ways. Whether this is a serious problem depends on the application at hand. In inheritance reasoning highest priority defaults are usually the ones for which the prerequisite is more directly derivable. However, as will be shown in Section 7.4, this is not always the case. On the other hand, the application of the highest priority default – whenever possible – would seem a useful declarative property for nonmonotonic reasoning with priorities. The satisfaction of this property naturally leads to lexicographic priorities as advocated elsewhere in this work.

Note that Definition 6.1.3, like the definition in [Brewka, 1994], can be seen as an extension of the definition of preferred subtheories of Brewka in Section 6.2. Preferred subtheories framework is equivalent to the subclass of default theories with prerequisite-free normal defaults, and it can be embedded to our prioritized default logic that uses a lexicographic priority mechanism. Hence the work of Baader and Hollunder, Marek and Truszczyński, and Brewka can be seen as an alternative way of extending the preferred subtheories framework to defaults that are not prerequisite-free normal.

6.2 Priorities and maximal consistent subsets

Poole [1988] presents a framework for default reasoning that is based on the use of maximal consistent subsets of sets of formulae. Brewka [1989] extends this framework with priorities. This generalization is called *preferred subtheories*. He considers two kinds of priorities. In the first version, instead of a set of formulae there is an n-tuple of sets of formulae. The formulae in elements of the n-tuples with a lower index have a higher priority. Clearly, this kind of priorities correspond to assigning an integer to each member of a set of formulae. In the second, more general version, the formulae are ordered by a strict partial order.

Definition 6.2.1 ([Brewka, 1989]) *Let $T_1, \ldots, T_n$ be sets of formulae. The set $S = S_1 \cup \cdots \cup S_n$ is a* preferred subtheory *of $\langle T_1, \ldots, T_n\rangle$ if and only if for all $k \in \{1, \ldots, n\}$, $S_1 \cup \cdots \cup S_k$ is a maximal consistent subset of $T_1 \cup \cdots \cup T_k$.*

In some applications it is necessary that a formula can have a priority that is incomparable to the priorities of two formulae the priorities of which differ. In the above definition this is not possible. The generalization that is needed is to allow the ordering of the formulae to be any strict partial order.

Definition 6.2.2 ([Brewka, 1989]) *Let $\mathcal{B}$ be a strict partial order on a finite set of formulae $T \subseteq \mathcal{L}$. Let $O \subseteq \mathcal{L}$ be a set of formulae. The set S is* a preferred subtheory *of $\langle T, O, \mathcal{B}\rangle$ if and only if there is a strict total order $t_1, t_2, \ldots, t_n$ on T that extends $\mathcal{B}$ and $S = S_n$ where*

$$\begin{array}{rcl} S_0 & = & O \\ S_{i+1} & = & \left\{ \begin{array}{ll} S_i \cup \{t_{i+1}\} & \textit{if } S_i \cup \{t_{i+1}\} \textit{ is consistent} \\ S_i & \textit{otherwise} \end{array} \right. \quad \textit{for } i, 0 \leq i < n. \end{array}$$

Actually, we have taken the liberty to extend Brewka's definition slightly. Brewka does not have the second component O. Whenever O is consistent, O could just be taken to be the formulae with the highest priority and exactly the same preferred sutheories would be obtained. The reason for our modification is that it makes the presentation of some results in later sections more convenient.

We show in Lemma 6.2.6 that Definition 6.2.2 is a generalization of Definition 6.2.1. Brewka [1989] does not give a proof for this fact. The next assumption does not lead to a loss of generality as shown by the lemma that follows.

Assumption 6.2.3 *For $T = \langle T_1, \ldots, T_n\rangle$, $T_i \cap T_j = \emptyset$ for all $i \in \{1, \ldots, n\}$ and $j \in \{i+1, \ldots, n\}$.*

Lemma 6.2.4 *Let $T = \langle T_1, \ldots, T_n\rangle$ be a level default theory. Let $T' = \langle T'_1, \ldots, T'_n\rangle$ where $T'_i = T_i \setminus \bigcup_{j=1}^{i-1} T_j$ for all $i \in \{1, \ldots, n\}$. For any set $S \subseteq \mathcal{L}$, S is a preferred subtheory of T if and only if S is a preferred subtheory of T'.*

Proof: By definition $S_1 \cup \cdots \cup S_n$ is a preferred subtheory of $\langle T_1, \ldots, T_n\rangle$ if and only if $S_1 \cup \cdots \cup S_i$ is a maximal consistent subset of $T_1 \cup \cdots \cup T_i$ for all $i \in \{1, \ldots, n\}$. Because $T_1 \cup \cdots \cup T_i = T'_1 \cup \cdots \cup T'_i$ for all $i \in \{1, \ldots, n\}$, equivalently $S_1 \cup \cdots \cup S_n$ is a preferred subtheory of $\langle T'_1, \ldots, T'_n\rangle$. $\square$

Lemma 6.2.5 *Let O be a consistent set of formulae and T a set of formulae. The sets $S_i, i \in \{1, \ldots, n\}$ in the definition of preferred subtheories of $\langle T, O, \mathcal{B}\rangle$ in Definition 6.2.2 are maximal consistent subsets of $\{t_1, \ldots, t_i\} \cup O$.*

Proof: Proof is by induction on $i \in \{0, \ldots, n\}$.

Induction hypothesis: S_i is a maximal consistent subset of $\{t_1, \ldots, t_i\} \cup O$.

Base case $i = 0$: By assumption $S_0 = O$ is consistent, and obviously it is a maximal subset of O.

Inductive case $i \geq 1$: By induction hypothesis S_{i-1} is a maximal consistent subset of $\{t_1, \ldots, t_{i-1}\} \cup O$. If $S_{i-1} \cup \{t_i\}$ is inconsistent, then $S_i = S_{i-1}$, and hence S_i is consistent. The set S_i is maximal because $S_i \cup \{t_i\}$ is inconsistent, and $S_i = S_{i-1}$ and by induction hypothesis $S_{i-1} \cup \{t\}$ is inconsistent for all $t \in \{t_1, \ldots, t_{i-1}\} \backslash S_{i-1}$. If $S_{i-1} \cup \{t_i\}$ is consistent, then $S_i = S_{i-1} \cup \{t_i\}$ and S_i is consistent. By induction hypothesis $S_{i-1} \cup \{t\}$ is inconsistent for all $t \in \{t_1, \ldots, t_{i-1}\} \backslash S_{i-1}$ and therefore $S_{i-1} \cup \{t\} \subseteq S_i \cup \{t\}$ is inconsistent for all $t \in \{t_1, \ldots, t_{i-1}\} \backslash S_{i-1}$. □

The next lemma shows that Brewka's level default theories are a special case of his preferred subtheories with arbitrary strict partial orders.

Lemma 6.2.6 *Let $\langle T_1, \ldots, T_n \rangle$ be a level default theory. Define $\mathcal{B} = \{\langle t, t' \rangle | t \in T_i, t' \in T_j, i < j\}$. Then $Z \subseteq \mathcal{L}$ is a preferred subtheory of $\langle T_1 \cup \cdots \cup T_n, \emptyset, \mathcal{B} \rangle$ if and only if it is a preferred subtheory of $\langle T_1, \ldots, T_n \rangle$.*

Proof: ($\Rightarrow$) With Assumption 6.2.3 $\mathcal{B}$ can be shown to be a strict partial order. Assume Z is a preferred subtheory of $\langle T_1 \cup \cdots \cup T_n, \emptyset, \mathcal{B} \rangle$. Hence there is a strict total order $t_1, \ldots, t_m$ on $T_1 \cup \cdots \cup T_n$ that extends $\mathcal{B}$ as defined in Definition 6.2.2. Let $l_0, \ldots, l_n$ be integers such that $\{t_{(l_{i-1})+1}, \ldots, t_{l_i}\} = T_i$ for all $i \in \{1, \ldots, n\}$. Let $S_0, \ldots, S_m$ be sets as defined in Definition 6.2.2. By Lemma 6.2.5 $S_1 \cup \cdots \cup S_{l_i}$ is a maximal consistent subset of $T_1 \cup \cdots \cup T_i$ for all $i \in \{1, \ldots, n\}$. Therefore by definition of preferred sutheories S_{l_n} is a preferred subtheory of $\langle T_1, \cdots, T_n \rangle$.

($\Leftarrow$) Assume Z is a preferred subtheory of $\langle T_1, \cdots, T_n \rangle$. Define $Z_i = T_i \cap Z$ for all $i \in \{1, \ldots, n\}$. Let $\mathcal{P}$ be a strict total order on $T_1 \cup \cdots \cup T_n$ such that $t\mathcal{P}t'$ whenever $t\mathcal{B}t'$ or $t \in Z_i$ and $t' \in T_i \backslash Z_i$ for some $i \in \{1, \ldots, n\}$. Let $t_1, \ldots, t_m$ be the ordering $\mathcal{P}$. We show that $Z = S_m$ where the sets S_i for $i \in \{0, \ldots, m\}$ are as in the equations of Definition 6.2.2.

Induction hypothesis: $S_i = Z \cap \{t_1, \ldots, t_i\}$.

Base case $i = 0$: Obviously $S_0 = \emptyset = Z \cap \emptyset$.

Inductive case $i \geq 1$: By induction hypothesis $S_{i-1} = Z \cap \{t_1, \ldots, t_{i-1}\}$. Assume $t_i \in Z$. Because Z is consistent and $S_{i-1} \cup \{t_i\} \subseteq Z$, $S_{i-1} \cup \{t_i\}$ is consistent and hence by definition, $S_i = S_{i-1} \cup \{t_i\}$. Assume $t_i \notin Z$, that is, $t_i \in T_j \backslash Z_j$ for some $j \in \{1, \ldots, n\}$. By definition of $\mathcal{P}$, $t_i \mathcal{P} t_k$ for no $t_k \in Z_j$, and hence $\{t_1, \ldots, t_{i-1}\} \subseteq Z_j$. Because by induction hypothesis

$Z \cap \{t_1, \ldots, t_{i-1}\} = S_{i-1}$, $Z \cap Z_j \subseteq S_{i-1}$ and hence $Z_j \subseteq S_{i-1}$. Because $Z_1 \cup \cdots \cup Z_j$ is a maximal consistent subset of $T_1 \cup \cdots \cup T_j$, $Z_1 \cup \cdots \cup Z_j \cup \{t_i\}$ is inconsistent. Therefore $S_{i-1} \cup \{t_i\}$ is inconsistent, and $S_i = S_{i-1}$ by definition of S_i. $\Box$

The following lemmata are used in the proofs of theorems that establish the connection between prioritized default logic and preferred subtheories.

Definition 6.2.7 *Let $\Theta = \langle T, O, \mathcal{B}\rangle$ where $T \subseteq \mathcal{L}$ is a set of formulae, $O \subseteq \mathcal{L}$ is a set of formulae, and $\mathcal{B}$ is a strict partial order on T. Then define*

$$
\begin{aligned}
tr_{ps}(\Theta) &= \left\langle \left\{ \frac{:\phi}{\phi} \middle| \phi \in T \right\}, O \right\rangle, \text{ and} \\
p_{ps}(\Theta) &= \left\{ \left\langle \frac{:\phi}{\phi}, \frac{:\psi}{\psi} \right\rangle \middle| \phi \mathcal{B} \psi \right\}.
\end{aligned}
$$

Lemma 6.2.8 *Let $T \subseteq \mathcal{L}$ and $O \subseteq \mathcal{L}$ be sets of formulae and O is consistent. Let S be a maximal consistent subset of $T \cup O$ that contains O. Then $Cn(S)$ is an extension of the default theory $tr_{ps}(\langle T, O, \emptyset\rangle)$.*

Proof: Define $\langle D, O\rangle = tr_{ps}(\langle T, O, \emptyset\rangle)$. We show that $Cn(S)$ fulfills the equation in Theorem 2.1.8.

$$
\begin{aligned}
Cn(O \cup \left\{ \phi \middle| \frac{:\phi}{\phi} \in D, \neg\phi \notin Cn(S) \right\}) &= Cn(O \cup \{\phi \in T \,|\neg\phi \notin Cn(S)\}) \\
&= Cn(O \cup \{\phi \in T \,|S \cup \{\phi\} \not\models \bot\}) \\
&= Cn(O \cup S) \\
&= Cn(S)
\end{aligned}
$$

The first equality is directly by the definition of D. The second equality is by elementary properties of the symbols involved. As $O \subseteq S$ and S is a maximal consistent subset of $T \cup O$, for all $\phi \in T$, $S \cup \{\phi\}$ is consistent if and only if $S \not\models \neg\phi$ if and only if $\phi \in S$. Hence we get the third equality. The fourth equality is because $O \subseteq S$. Hence $Cn(S)$ is an extension of $\langle D, O\rangle$. $\Box$

Lemma 6.2.9 *Let $T \subseteq \mathcal{L}$ be a set of formulae, $O \subseteq \mathcal{L}$ a consistent set of formulae, and E an extension of $tr_{ps}(\langle T, O, \emptyset\rangle)$. Then $S = E \cap (T \cup O)$ is a maximal consistent subset of $T \cup O$ and $E = Cn(S)$.*

Proof: Assume that S is not a maximal consistent subset of $T \cup O$. Because Δ is a normal default theory and the objective part of Δ is empty, E and consequently also S are consistent. Then S is not maximal, that is, there is a set S' such that $S \subseteq S'$ and $\phi \in S'$ for some $\phi \in T\backslash S$. Let S' be a set-inclusion maximal such set. Now by Lemma 6.2.8 $Cn(S')$ is an extension of $tr_{ps}(\langle T, O, \emptyset\rangle)$. Clearly $\phi \notin E$ because $\phi \in T$ and $\phi \notin E \cap (T \cup O)$. Now $Cn(S') \subset Cn(S)$ and this contradicts [Reiter, 1980, Theorem 2.4]. Therefore $E \cap (T \cup O)$ is a maximal consistent subset of $T \cup O$. □

Lemma 6.2.10 *Let $T \subseteq \mathcal{L}$ be a set of formulae and $O \subseteq \mathcal{L}$ a consistent set of formulae. Then for all $\phi \in T$ and all extensions E of $tr_{ps}(\langle T, O, \emptyset\rangle)$, $\neg\phi \notin E$ if and only if $\phi \in E$.*

Proof: By Lemma 6.2.9, for every extension E of $tr_{ps}(\langle T, O, \emptyset\rangle)$ there is a maximal consistent subset S of $T \cup O$ such that $E = Cn(S)$. Let ϕ be any member of T and let E be any extension of $tr_{ps}(\langle T, O, \emptyset\rangle)$.

($\Rightarrow$) Assume $\neg\phi \notin E$. Now $E = Cn(S)$ for some maximal consistent subset S of $T \cup O$. Because $\neg\phi \notin E$, $S \not\models \neg\phi$ and hence $S \cup \{\phi\}$ is consistent. Because S is a maximal consistent subset of $T \cup O$, $\phi \in S \subseteq E$.

($\Leftarrow$) Assume $\phi \in E$. Because $E = Cn(S)$ for some maximal consistent subset S of $T \cup O$, E is consistent. If $\neg\phi$ were in E, E would not be consistent. Therefore $\neg\phi \notin E$. □

Lemma 6.2.11 *If S is a preferred subtheory of $\langle T, O, \mathcal{B}\rangle$, then $Cn(S)$ is a $p_{ps}(\langle T, O, \mathcal{B}\rangle)$-preferred extension of $tr_{ps}(\langle T, O, \mathcal{B}\rangle)$.*

Proof: Assume O is inconsistent. Then O is the only preferred subtheory of $\langle T, O, \mathcal{B}\rangle$ and $Cn(O) = \mathcal{L}$ is the unique extension of $tr_{ps}(\langle T, O, \mathcal{B}\rangle)$. By Theorem 3.1.20 $Cn(O)$ is a $p_{ps}(\langle T, O, \mathcal{B}\rangle)$-preferred extension. Assume O is consistent. Let S be a preferred subtheory of $\langle T, O, \mathcal{B}\rangle$ and let $t_1, \ldots, t_n$ be a total ordering of T that is mentioned in Definition 6.2.2. Take $\mathcal{T}$ to be the total ordering $:t_1/t_1, \ldots, :t_n/t_n$. Clearly $p_{ps}(\langle T, O, \mathcal{B}\rangle) \subseteq \mathcal{T}$. We show that $\mathcal{T}$ is a $tr_{ps}(\langle T, O, \mathcal{B}\rangle), p_{ps}(\langle T, O, \mathcal{B}\rangle)$-characteristic ordering for $Cn(S)$.

Suppose there is an extension E' of $tr_{ps}(\langle T, O, \mathcal{B}\rangle)$ such that for some $t_k \in T$, appl$(:t_k/t_k, E', Cn(S))$ and appl$(:t_l/t_l, Cn(S), E')$ for no $l \leq k$. Therefore $Cn(S) \cap \{t_1, \ldots, t_{k-1}\} \subseteq E' \cap \{t_1, \ldots, t_{k-1}\}$ and $t_k \in E' \backslash Cn(S)$. By Lemma 6.2.5 $Cn(S) \cap (\{t_1, \ldots, t_k\} \cup O)$ is a maximal consistent subset of $\{t_1, \ldots, t_k\} \cup O$. Because $Cn(S) \cap (\{t_1, \ldots, t_k\} \cup O) \subset E' \cap (\{t_1, \ldots, t_k\} \cup O)$, must $E' \cap (\{t_1, \ldots, t_k\} \cup O)$ be inconsistent. However, this contradicts the consistency of E' by Lemma 6.2.9. Hence we must conclude that no such E' exists, and therefore $\mathcal{T}$ must be a $tr_{ps}(\langle T, O, \mathcal{B}\rangle), p_{ps}(\langle T, O, \mathcal{B}\rangle)$-characteristic ordering for $Cn(S)$, and $Cn(S)$ is a $p_{ps}(\langle T, O, \mathcal{B}\rangle)$-preferred extension of $tr_{ps}(\langle T, O, \mathcal{B}\rangle)$. □

Lemma 6.2.12 *If E is a $p_{ps}(\langle T, O, \mathcal{B}\rangle)$-preferred extension of $tr_{ps}(\langle T, O, \mathcal{B}\rangle)$, then $E \cap (T \cup O)$ is a preferred subtheory of $\langle T, O, \mathcal{B}\rangle$ and $E = Cn(E \cap (T \cup O))$.*

Proof: Let E be a $p_{ps}(\langle T,O,\mathcal{B}\rangle)$-preferred extension of $tr_{ps}(\langle T,O,\mathcal{B}\rangle)$. Assume E is inconsistent. Then O is inconsistent, and O is the only preferred subtheory of $\langle T,O,\mathcal{B}\rangle$. Assume E is consistent. Then there is a $tr_{ps}(\langle T,O,\mathcal{B}\rangle), p_{ps}(\langle T,O,\mathcal{B}\rangle)$-characteristic ordering $\mathcal{T}$ for E. Let the ordering $\mathcal{T}$ be $:t_1/t_1, \ldots, :t_n/t_n$. The proof that $E \cap (T \cup O)$ is a preferred subtheory of $T \cup O$ is by induction. In the proof we refer to the sets S_i in the equations of Definition 6.2.2.

Induction hypothesis: $S_i = E \cap (\{t_1, \ldots, t_i\} \cup O)$.

Base case $i = 0$: $S_0 = O = E \cap (\emptyset \cup O)$ by definition of S_0 and because $O \subseteq E$.

Inductive case $i \geq 1$: By induction hypothesis $S_{i-1} = E \cap (\{t_1, \ldots, t_{i-1}\} \cup O)$. Assume $S_{i-1} \cup \{t_i\}$ is consistent. Now $S_i = S_{i-1} \cup \{t_i\}$ and $t_i \in S_i$ and there is a maximal consistent subset S of $T \cup O$ such that $S_i \subseteq S$. Hence $Cn(S)$ is an extension of $tr_{ps}(\langle T,O,\mathcal{B}\rangle)$ by Lemma 6.2.8. Now $t_i \in E$ because if $t_i \notin E$ it would be the case that appl$(:t_i/t_i, Cn(S), E)$ and appl$(\delta, E, Cn(S))$ for no $\delta \mathcal{T} :t_i/t_i$. Hence $\mathcal{T}$ would not be a $tr_{ps}(\langle T,O,\mathcal{B}\rangle), tr_{ps}(\langle T,O,\mathcal{B}\rangle)$-characteristic ordering for E. Assume $S_{i-1} \cup \{t_i\}$ is inconsistent. Hence $S_i = S_{i-1}$ and $t_i \notin S_i$. Because $S_i \subseteq E$ and E is consistent by A.1.1, $t_i \notin E$. Hence $S_i = E \cap (\{t_1, \ldots, t_i\} \cup O)$. □

The consequence relations defined by Brewka are strong provability and weak provability. A formula is *strongly provable* from a theory if it is a logical consequence of all preferred subtheories of the theory. A formula is *weakly provable* from a theory if it is a logical consequence of at least one preferred subtheory of the theory. These notions respectively correspond to cautious and brave reasoning in default logic.

Theorem 6.2.13 *A formula $\phi \in \mathcal{L}$ is strongly provable from $\langle T,O,\mathcal{B}\rangle$ if and only if ϕ is in all $p_{ps}(\langle T,O,\mathcal{B}\rangle)$-preferred extensions of $tr_{ps}(\langle T,O,\mathcal{B}\rangle)$, and ϕ is weakly provable from $\langle T,O,\mathcal{B}\rangle$ if and only if ϕ is in some $p_{ps}(\langle T,O,\mathcal{B}\rangle)$-preferred extension of $tr_{ps}(\langle T,O,\mathcal{B}\rangle)$.*

Proof: With Assumption 6.2.3 $p_{ps}(\langle T,O,\mathcal{B}\rangle)$ can be shown to be a strict partial order. A formula $\phi \in \mathcal{L}$ is strongly provable from $\langle T,O,\mathcal{B}\rangle$ if and only if $S \models \phi$ for all preferred subtheories S of $\langle T,O,\mathcal{B}\rangle$ if and only if ϕ is in all $p_{ps}(\langle T,O,\mathcal{B}\rangle)$-preferred extensions of $tr_{ps}(\langle T,O,\mathcal{B}\rangle)$. The first *iff* is by the definition of strong provability. The second *iff* is because the intersection of the closures of the preferred subtheories of $\langle T,O,\mathcal{B}\rangle$ under logical consequence is the same set as the intersection of the $p_{ps}(\langle T,O,\mathcal{B}\rangle)$-preferred extensions of $tr_{ps}(\langle T,O,\mathcal{B}\rangle)$. This is because by Lemmata 6.2.12 and 6.2.11 the collection of the closures under logical consequence of all preferred subtheories of $\langle T,O,\mathcal{B}\rangle$ is the same as the collection of all $p_{ps}(\langle T,O,\mathcal{B}\rangle)$-preferred extensions of $tr_{ps}(\langle T,O,\mathcal{B}\rangle)$.

A formula $\phi \in \mathcal{L}$ is weakly provable from $\langle T, O, \mathcal{B}\rangle$ if and only if $S \models \phi$ for some preferred subtheory S of $\langle T, O, \mathcal{B}\rangle$ if and only if ϕ is in some $p_{ps}(\langle T, O, \mathcal{B}\rangle)$-preferred extension of $tr_{ps}(\langle T, O, \mathcal{B}\rangle)$. The first *iff* is by definition and the second is directly by Lemmata 6.2.11 and 6.2.12. □

6.3 Ordered theory presentations

Ryan [1992a] defines ordered theory presentations (OTPs) as partial orders on sets of formulae. The ordering expresses the plausibility of the formulae and the maximal elements are the most plausible. An OTP determines a partial order on the models of the language $\mathcal{L}$. The consequences of an OTP are the formulae that are true in all the maximal models. An interesting novelty in OTPs is the use of natural consequences of formulae. Natural consequences are logical consequences that (as truth-functions) fulfill certain monotonicity conditions.

Definition 6.3.1 ([Ryan, 1992a]) *A formula ϕ is* monotonic *in p (written $p \in \phi^+$), if for all models M and N such that $N \models p$ and for all propositional variables $q \neq p$ $M \models q$ if and only if $N \models q$, $M \models \phi$ implies $N \models \phi$.*

A formula ϕ is anti-monotonic *in p (written $p \in \phi^+$), if for all models M and N such that $N \not\models p$ and for all propositional variables $q \neq p$ $M \models q$ if and only if $N \models q$, $M \models \phi$ implies $N \models \phi$.*

A logical consequence ψ of a formula ϕ is a natural consequence of ϕ, *written $\phi \models^\natural \psi$, if $\phi \models \psi$ and $\phi^+ \subseteq \psi^+$ and $\phi^- \subseteq \psi^-$.*

The motivation for introducing natural consequences is that even when a formula cannot be satisfied simultaneously with some other formulae, some "part" of it can.

The set of natural consequences of a formula is infinite. If ϕ is a natural consequence of a formula, then also $\neg\neg\phi, \neg\neg\neg\neg\phi$ and so on, are natural consequences of the formula. In the propositional case the set of natural consequences can be replaced by a finite subset of it without affecting the consequence relations induced by OTPs.

Theorem 6.3.2 *For OTPs based on classical propositional logic with a finite number of propositional variables, the relation $\models^\natural$ in Ryan's definition can be replaced by a relation $\models^\circ$ that is a subset of $\models^\natural$ such that $\{\psi|\phi \models^\circ \psi\}$ is finite for all $\phi \in \mathcal{L}$. This change preserves all properties of $\models$ in $\Gamma \models \phi$, where Γ is an OTP and $\phi \in \mathcal{L}$ a formula.*

Proof: In Ryan's definitions natural consequences of formulae are treated syntax-independently, that is, each formula could be replaced by a logically equivalent but syntactically different formula without affecting the definitions. Therefore it suffices to consider for each formula ϕ only one natural consequence from each equivalence class of logically equivalent natural consequences.

In the propositional case, when there are only a finite number of propositional variables, the number of equivalence classes of logically equivalent natural consequences is finite. Therefore there is a finite set of natural consequences so that for each equivalence class one of its members is in the set. $\Box$

The following is the formal definition of OTPs and the ordering relations on models induced by an OTP.

Definition 6.3.3 ([Ryan, 1992a]) *Let X be a set, $\leq$ a partial order (an antisymmetric and transitive relation) on X and $F : X \rightarrow \mathcal{L}$ a function that maps elements of X to formulae. Then $\Gamma = \langle X, \leq, F\rangle$ is an OTP.*

Let M and N belong to $\mathcal{M}$ (the collection of all models.) Then $M \sqsubseteq_\phi N$, if for all $\psi \in \mathcal{L}$, $\phi \models^\circ \psi$ implies ($M \models \psi$ implies $N \models \psi$).

Let $\Gamma = \langle X, \leq, F\rangle$ be an OTP, and let $M, N \in \mathcal{M}$. Then $M \sqsubseteq^\Gamma N$ if for each $x \in X$, if $M \not\sqsubseteq_{F(x)} N$ then there is $y \leq x$ such that $M \sqsubset_{F(y)} N$ (that is, $M \sqsubseteq_{F(y)} N$ and $N \not\sqsubseteq_{F(y)} M$.)

Let Γ be an OTP, and M an element of $\mathcal{M}$. Then $M \models \Gamma$ if M is $\sqsubseteq^\Gamma$-maximal.

Let Γ be an OTP. Then $\Gamma \models \phi$, if for each $M \in \mathcal{M}, M \models \Gamma$ implies $M \models \phi$.

The following translation of OTPs to prioritized default logic resembles the translation of preferred subtheories in Section 6.2. By $<$ we mean the strict (asymmetric) partial order $\leq \backslash \{\langle x, x\rangle | x \in X\}$.

Definition 6.3.4 *Let $\Gamma = \langle X, \leq, F\rangle$ be an OTP. The translation of Γ to prioritized default logic is defined by using the following functions.*

$$
\begin{aligned}
tr_o(\Gamma) &= \left\langle \left\{ \frac{:\phi}{\phi} \middle| F(x) \models^\circ \phi, x \in X \right\}, \emptyset \right\rangle \\
p_o(\Gamma) &= \left\{ \left\langle \frac{:\phi}{\phi}, \frac{:\psi}{\psi} \right\rangle \middle| F(x) \models^\circ \phi, F(y) \models^\circ \psi, x < y, x \in X, y \in X \right\}
\end{aligned}
$$

Because Ryan allows multiple occurrences of the same formula in an OTP and distinct formulae may share natural consequences, there is a minor problem with our definition of priorities where there may be only one occurrence of each default in the partial order that expresses the priorities. It can be solved by renaming each propositional variable p occurring in a natural consequence of a formula $F(x)$ to p_x and extending the translation of an OTP with equivalences $p \leftrightarrow p_x$. This transformation causes a linear increase in the size of the translation. Our discussion of OTPs from now on assumes that this transformation has been made.

The main result about OTPs is the following theorem, which we prove later after giving some lemmata.

Theorem 6.3.5 *Let $\Gamma = \langle X, \leq, F\rangle$ be an OTP where $\leq$ is a well-ordering. Then $\Gamma \models \phi$ if and only if ϕ belongs to all $p_o(\Gamma)$-preferred$_2$ extensions of $tr_o(\Gamma)$.*

Lemma 6.3.6 *Let $\Gamma = \langle X, \leq, F\rangle$ be an OTP where $\leq$ is a well-ordering. Let E and E' be extensions of $tr_o(\Gamma)$ and let M and M' be models of E and E', respectively. Then $E \sqsubset^{tr_o(\Gamma)}_{p_o(\Gamma)} E'$ if and only if $M' \sqsubset^{\Gamma} M$.*

Proof: In this proof we use the alternative definition of $\sqsubset_{\mathcal{P}}$ that is given in Theorem 3.1.14. Let $\mathcal{P} = p_o(\Gamma)$ and $\Delta = \langle D, \emptyset\rangle = tr_o(\Gamma)$.

($\Rightarrow$) Assume $E \sqsubset^{\Delta}_{\mathcal{P}} E'$. Assume $M' \not\sqsubseteq_{F(x)} M$ for some $x \in X$. That is, for some ϕ, $F(x) \models^{\circ} \phi$ and $M' \models \phi$ and $M \not\models \phi$. By Lemma 6.2.10 $\phi \in E' \backslash E$ and appl($:\phi/\phi, E', E$). Because $E \sqsubset^{\Delta}_{\mathcal{P}} E'$ there is $:\psi/\psi \in D$ such that $:\psi/\psi\mathcal{P}{:}\phi/\phi$ and appl($:\psi/\psi, E, E'$). Take $:\psi/\psi$ to be a $\mathcal{P}$-minimal element of D such that $:\psi/\psi\mathcal{P}{:}\phi/\phi$ and appl($:\psi/\psi, E, E'$). Such an element exists because $\mathcal{P}$ is a well-ordering. Because $\psi \in E\backslash E'$, by Lemma 6.2.10 $M \models \psi$ and $M' \not\models \psi$, and because $:\psi/\psi\mathcal{P}{:}\phi/\phi$ there is $y \in X$ such that $F(y) \models^{\circ} \psi$ and $y < x$. Hence $M \not\sqsubseteq_{F(y)} M'$. It remains to show that $M' \sqsubseteq_{F(y)} M$. Assume the opposite, that is, for some ω, $F(y) \models^{\circ} \omega$ and $M' \models \omega$ and $M \not\models \omega$. This implies by Lemma 6.2.10 that $\omega \in E'\backslash E$. Now because $E \sqsubset^{\Delta}_{\mathcal{P}} E'$ there should be $:\theta/\theta \in D$ such that $:\theta/\theta\mathcal{P}{:}\omega/\omega$ and appl($:\theta/\theta, E, E'$), and consequently there is $z \in X$ such that $F(z) \models^{\circ} \theta$. Because $F(y) \models^{\circ} \omega$ and $:\theta/\theta\mathcal{P}{:}\omega/\omega$, $z < y$. Hence $:\theta/\theta\mathcal{P}{:}\psi/\psi\mathcal{P}{:}\phi/\phi$. This contradicts the assumption that $:\psi/\psi$ is a $\mathcal{P}$-minimal element such that $:\psi/\psi\mathcal{P}{:}\phi/\phi$ and appl($:\psi/\psi, E, E'$). Therefore for all ω such that $F(y) \models^{\circ} \omega$, $M' \models \omega$ implies $M \models \omega$. Therefore $M' \sqsubseteq_{F(y)} M$. Because $M \not\sqsubseteq_{F(y)} M'$ and $M' \sqsubseteq_{F(y)} M$, $M' \sqsubset_{F(y)} M$. This completes the proof of $M' \sqsubseteq^{\Gamma} M$.

To prove that $M \not\sqsubseteq^{\Gamma} M'$, we show that there is $x \in X$ such that $M \not\sqsubseteq_{F(x)} M'$ and there is no $y \in X$ such that $y \leq x$ and $M \sqsubset_{F(y)} M'$. Take $:\phi/\phi$ to be such that appl($:\phi/\phi, E, E'$) and there is no δ such that $\delta\mathcal{P}{:}\phi/\phi$ and appl(δ, E, E'). Such a δ exists because $\mathcal{P}$ is a well-ordering and as $E \sqsubset^{\Delta}_{\mathcal{P}} E'$, $E \neq E'$ and by [Reiter, 1980, Theorems 2.4 and 2.5] there must be some $:\phi/\phi$ such that appl($:\phi/\phi, E, E'$). Now $M \models \phi$ and $M' \not\models \phi$. By definition of Δ there is

$x \in X$ such that $F(x) \models^{\circ} \phi$. Now $M \not\sqsubseteq_{F(x)} M'$. Assume there is y such that $y \leq x$ and $M \sqsubset_{F(y)} M'$. Now there is ψ such that $F(y) \models^{\circ} \psi$ and $M \not\models \psi$ and $M' \models \psi$, and therefore $\psi \in E' \backslash E$ by Lemma 6.2.10 and $:\psi/\psi \mathcal{P} :\phi/\phi$ by the definition of $\mathcal{P}$. Because $E \sqsubset^{\Delta}_{\mathcal{P}} E'$ there is $:\chi/\chi \mathcal{P} :\psi/\psi$ such that appl$(:\chi/\chi, E, E')$. Now $:\chi/\chi \mathcal{P} :\psi/\psi \mathcal{P} :\phi/\phi$. This, however, contradicts our assumption that $:\phi/\phi$ is a $\mathcal{P}$-minimal default such that appl$(:\phi/\phi, E, E')$. Hence there is no $y \in X$ such that $y \leq x$ and $M \sqsubseteq_{F(y)} M'$. Therefore $M' \sqsubset^{\Gamma} M$.

($\Leftarrow$) Assume $M' \sqsubset^{\Gamma} M$, that is, $M' \sqsubseteq^{\Gamma} M$ and $M \not\sqsubseteq^{\Gamma} M'$. We show that $E \sqsubset^{\Delta}_{\mathcal{P}} E'$, that is, for all $\delta \in D$, if appl(δ, E', E) then there is $\delta' \mathcal{P} \delta$ such that appl(δ', E, E'). Assume appl$(:\phi/\phi, E', E)$. Now $M \not\models \phi$ and $M' \models \phi$ and $F(x) \models^{\circ} \phi$ for some $x \in X$. Hence $M' \not\sqsubseteq_{F(x)} M$. Because $M' \sqsubseteq^{\Gamma} M$, $M' \not\sqsubseteq_{F(x)} M$ implies the existence of $y \in X$ such that $y \leq x$ and $M' \sqsubset_{F(y)} M$. Now it must be the case that $y \neq x$ and hence $y < x$. Further, for some ω, $F(y) \models^{\circ} \omega$ and $M' \not\models \omega$ and $M \models \omega$. By definition of $\mathcal{P}$, $:\omega/\omega \mathcal{P} :\phi/\phi$. By Lemma 6.2.10, $\omega \in E \backslash E'$, and thus appl$(:\omega/\omega, E, E')$. Hence there is $:\omega/\omega \in D$ such that appl$(:\omega/\omega, E, E')$ and $:\omega/\omega \mathcal{P} :\phi/\phi$, and finally $E \sqsubset^{\Delta}_{\mathcal{P}} E'$. $\square$

Lemma 6.3.7 *Let $\Gamma = \langle X, \leq, F\rangle$ be an OTP, and $\phi \in \mathcal{L}$ a formula. If $\phi \notin E$ for some $p_o(\Gamma)$-preferred$_2$ extension E, then there is a $\sqsubseteq^{\Gamma}$-maximal model M such that $M \not\models \phi$.*

Proof: Let E be a $p_o(\Gamma)$-preferred$_2$ extension of $tr_o(\Gamma)$ such that $\phi \notin E$. Hence there is a model M such that $M \models E \cup \{\neg\phi\}$. Assume there is N such that $M \sqsubset^{\Gamma} N$. The set $\Omega = \{\phi | N \models \phi, F(x) \models^{\circ} \phi, x \in X\}$ is not necessarily a maximal consistent subset of $T = \{\phi | F(x) \models^{\circ} \phi, x \in X\}$. Let Ω' be a maximal consistent set of T such that $\Omega \subseteq \Omega' \subseteq T$. Because Ω' is consistent, it has a model L. Now $N \not\sqsubseteq_{F(x)} L$ for no $x \in X$, because for all ϕ and $x \in X$ such that $F(x) \models^{\circ} \phi$, $N \models \phi$ implies $L \models \phi$. Therefore $N \sqsubseteq^{\Gamma} L$ and by the transitivity of $\sqsubseteq^{\Gamma}$, $M \sqsubset^{\Gamma} L$. Because Ω' is a maximal consistent subset of T, there is by Lemma 6.2.8 an extension $E' = Cn(\Omega')$ of $tr_o(\Gamma)$. By Lemma 6.3.6 $E' \sqsubset_{p_o(\Gamma)} E$. This contradicts the $\mathcal{P}$-preferredness$_2$ of E, and hence M is $\sqsubseteq^{\Gamma}$-maximal. $\square$

Lemma 6.3.8 *Let $\Gamma = \langle X, \leq, F\rangle$ be an OTP. Let $\phi \in \mathcal{L}$ be a formula. If $M \not\models \phi$ for some $\sqsubseteq^{\Gamma}$-maximal model M, then there is a $p_o(\Gamma)$-preferred$_2$ extension E of $tr_o(\Gamma)$ such that $\phi \notin E$.*

Proof: Let $T = \{\phi | F(x) \models^{\circ} \phi, x \in X\}$. Let M be a $\sqsubseteq^{\Gamma}$-maximal model. Clearly, $\Pi = \{\phi \in T | M \models \phi\}$ is a consistent subset of T. Assume Π is not maximal, that is, there is a consistent set $\Omega \subseteq T$ such that $\Pi \subset \Omega$. Now

$M' \models \Omega$ for some model M'. Clearly, for no $\chi \in T$, $M \models \chi$ and $M' \not\models \chi$. Hence $M \sqsubseteq^\Gamma M'$. Take any $\chi \in \Omega \backslash \Pi$. Now $M' \models \chi$ and $M \not\models \chi$. Hence $M' \not\sqsubseteq^\Gamma M$, and finally $M \sqsubset^\Gamma M'$. This contradicts the maximality of M, so Π must be a maximal consistent subset of T. Hence by Lemma 6.2.8 $Cn(\Pi)$ is an extension of $tr_o(\Gamma)$. Because $M \models \Pi$ and $M \not\models \phi$, $\phi \notin Cn(\Pi)$. Assume $E = Cn(\Pi)$ is not $p_o(\Gamma)$-preferred$_2$. By Lemma 6.3.6, there would be a model N for which $M \sqsubset^\Gamma N$. This would contradict the maximality of M. Hence E is $p_o(\Gamma)$-preferred$_2$ and $\phi \notin E$. □

Proof of Theorem 6.3.5:

$$\begin{array}{lll}
\Gamma \models \phi & \text{iff} & \text{for all } \sqsubseteq^\Gamma\text{-maximal models } M, M \models \phi \\
 & \text{iff} & \text{there is no } \sqsubseteq^\Gamma\text{-maximal model } M \text{ such that } M \not\models \phi \\
 & \text{iff} & \text{there is no } p_o(\Gamma)\text{-preferred}_2 \text{ extension } E \text{ such that } \phi \notin E \\
 & \text{iff} & \phi \in E \text{ for all } p_o(\Gamma)\text{-preferred}_2 \text{ extensions } E.
\end{array}$$

The third equivalence is by Lemmata 6.3.7 and 6.3.8: $\phi \notin E$ for some $p_o(\Gamma)$-preferred$_2$ extension E of $tr_o(\Gamma)$ if and only if $M \not\models \phi$ for some $\sqsubseteq^\Gamma$-maximal model M. □

The use of natural consequences in OTPs is the first difference between preferred subtheories of Brewka and Ryan's ordered theory presentations. However, this difference is independent of the kind of model minimization used, and natural consequences of formulae can be used in conjunction of preferred subtheories as well. The second difference is that Ryan's preference notion corresponds to preferred$_2$ extensions in prioritized default logic, and Brewka uses in preferred subtheories a preference notion that corresponds to preferred extensions.

In some applications, like belief revision, there may be advantages in using natural consequence of formulae instead of the formulae themselves [Ryan, 1992b]. Surprisingly, in some computationally favourable cases like Horn clauses or 2-literal clauses, these advantages disappear: there is only one equivalence class of natural consequences for disjunctions of literals (with respect to logical consequence) which is represented by the disjunction itself. With this kind of disjunctions as formulae OTPs come very close to preferred subtheories because only the difference between preferred and preferred$_2$ extensions remains.

6.4 Prioritized circumscription

Prioritized circumscription is a variant of circumscription [McCarthy, 1980] proposed by McCarthy [1986] and Lifschitz [1985]. A definition of circumscription with varied predicates, on which prioritized circumscription is based, is given in [Lifschitz, 1994]. The purpose of circumscription is to form a formula Φ_{circ} that represents the models of a formula or a set of formulae Φ where the interpretation of a predicate P is as small as possible (in the set-inclusion sense.) Here the predicate P is an abnormality predicate, and the idea is that the models of Φ containing as few abnormalities as possible are the interesting ones.

Let $Z_1, \ldots, Z_n$ be the predicates to be varied while circumscribing P in Φ. The result Φ_{circ} of circumscribing P in Φ with other predicates varied is logically equivalent to the second-order formula $\Phi \wedge \neg\exists p z_1 \cdots z_n(\Phi(p, z_1, \ldots, z_n) \wedge p < P)$ where $\Phi(p, z_1, \ldots, z_n)$ is like Φ but occurrences of P have been replaced by p and occurrences of Z_i have been replaced by z_i for all $i \in \{1, \ldots, n\}$, and $p < P$ is defined as $(p \leq P) \wedge \neg(p = P)$ where $p = P$ stands for $\forall x(p(x) \equiv P(x))$ and $p \leq P$ stands for $\forall x(p(x) \rightarrow P(x))$.

Circumscription can be seen as defining an ordering on models of Φ on the basis of the interpretation of P and other predicates. That a predicate Z is varied means that the interpretation of Z is not taken into account when determining whether two models are ordered; the ordering is determined by the interpretation of P and non-varied ordinary predicates: a model M is better than a model M' if the interpretation of P in M is smaller than in M' (in the set-inclusion sense), and the interpretations of other predicates – excluding Z – are the same in M and M'. The formula Φ_{circ} is true in M if and only if there is no model that is better than M.

Lifschitz [1994] gives a definition of priorities for circumscription. The abnormality predicates occurring in a set of formulae are ordered by a strict partial order. An abnormality predicate is given a higher priority than another abnormality predicate by circumscribing the former predicate with the latter predicate varied. When the latter predicate is circumscribed the former predicate is kept fixed. This policy informally used by Lifschitz may seem to work for arbitrary strict partial orders but in fact it is restricted to layered strict partial orders.

We show that prioritized circumscription corresponds to preferred subtheories of Brewka with layered strict partial orders and hence is essentially the same as prioritized default logic with prerequisite-free normal defaults. First we define the class of circumscriptive theories used in our results.

Definition 6.4.1 *Let $\mathcal{P}$ be a strict partial order on a set $A = \{ab_1, \ldots, ab_n\}$*

of abnormality predicates. Let Σ *be* $\{\forall x(\neg ab_1(x) \rightarrow \phi_1), \ldots, \forall x(\neg ab_n(x) \rightarrow \phi_n), \psi_1, \ldots, \psi_m\}$ *so that members of* A *do not occur in* ϕ_i *or in* ψ_j *for all* $i \in \{1, \ldots, n\}$ *and* $j \in \{1, \ldots, m\}$*. Then* $\langle \Sigma, A, \mathcal{P} \rangle$ *is a* prioritized circumscriptive theory.

Above we made a restriction to formulae where the predicates to be circumscribed occur positively as first disjuncts of disjunctive formulae. That this is not a serious restriction is shown next. Informally, for any set of formulae, there is another set of formulae, that is of the same form as prioritized circumscriptive theories and has exactly the same logical consequences as the former set, as far as formulae with no occurrences of certain auxiliary predicates are concerned.

Lemma 6.4.2 *Let* Ψ *be the result of circumscribing* ab *in* Φ*, and* Ψ' *the result of circumscribing* ab' *in* $\Phi \cup \{\forall x(\neg ab'(x) \rightarrow \neg ab(x))\}$*. Then for all formulae* ϕ *with no occurrences of* ab'*,* $\Psi \models \phi$ *if and only if* $\Psi' \models \phi$*.*

Proof: By definition, the result of circumscribing ab in Φ is C_1 and the result of circumscribing ab$'$ in $\Phi \cup \{\forall x(\neg \mathrm{ab}(x) \rightarrow \neg \mathrm{ab}'(x))\}$ is C_2, where

$$\begin{aligned} C_1 &= \bigwedge \Phi \wedge \neg \exists p z_1 \cdots z_n (\bigwedge \Phi(p, z_1, \ldots, z_n) \wedge p < \mathrm{ab}), \text{and} \\ C_2 &= \bigwedge \Phi \wedge \forall x(\neg \mathrm{ab}'(x) \rightarrow \neg \mathrm{ab}(x)) \wedge \neg \exists p z_1 \cdots z_{n+1} (\bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \\ &\quad \wedge \forall x(\neg p(x) \rightarrow \neg z_{n+1}(x)) \wedge p < \mathrm{ab}'). \end{aligned}$$

In C_2 the predicate variable z_{n+1} corresponds to the varied predicate ab. The claim of this lemma can be stated in the following two equivalent ways: for $\phi \in \mathcal{L}$ with no occurrences of ab$'$,

1. $C_1 \models \phi$ if and only if $C_2 \models \phi$, and
2. there is a model M such that $M \models C_1 \wedge \neg \phi$ if and only if there is a model M' such that $M' \models C_2 \wedge \neg \phi$.

We prove the second statement.

($\Rightarrow$) Assume that $M \models C_1 \wedge \neg \phi$ for some model M. Let M' be a model such that $M'(\mathrm{ab}') = M(\mathrm{ab})$ and $M'(P) = M(P)$ for all other predicates P[1]. Take any interpretations for predicate variables $p, z_1, \ldots, z_{n+1}$. The second conjunct of C_1 is logically equivalent to $\forall p z_1 \cdots z_n (\neg \bigwedge \Phi(p, z_1, \ldots, z_n) \vee \neg(p < \mathrm{ab}))$. Hence $M' \models \neg \bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \vee \neg(z_{n+1} < \mathrm{ab}')$. Case 1. If $M' \not\models \forall x(\neg p(x) \rightarrow \neg z_{n+1}(x)) \wedge p < \mathrm{ab}'$, then clearly $M' \models \neg(\bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \wedge$

[1] Note that the assumption $M \models C_1 \wedge \neg \phi$ does not restrict the way how ab$'$ is interpreted in M because ab$'$ does not occur in Φ or in ϕ.

$\forall x(\neg p(x) \rightarrow \neg z_{n+1}(x)) \wedge p < \mathrm{ab}'$. Case 2. Otherwise $M' \models \forall x(\neg p(x) \rightarrow \neg z_{n+1}(x)) \wedge p < \mathrm{ab}'$. Hence $z_{n+1} < \mathrm{ab}$. Now because $M' \models C_1$, we get $M' \models \neg \bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \vee \neg(z_{n+1} < \mathrm{ab})$, and $M' \models \neg \bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n)$. Because the derivation was independent of the $p, z_1, \ldots, z_{n+1}$ chosen, we get $M' \models \neg\exists p z_1 \cdots z_{n+1}(\bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \wedge \forall x(\neg p(x) \rightarrow \neg z_{n+1}(x)) \wedge p < \mathrm{ab}')$. Clearly $M' \models \Phi \cup \{\forall x(\neg \mathrm{ab}'(x) \rightarrow \neg \mathrm{ab}(x)), \neg\phi\}$ because ab and ab$'$ have the same interpretation in M' and ϕ has no occurrences of ab$'$. Therefore $M' \models C_2 \wedge \neg\phi$.

($\Leftarrow$)

Lemma A If $M' \models C_2$, then $M' \models \forall x(\mathrm{ab}(x) \leftrightarrow \mathrm{ab}'(x))$.
Proof: Because $M' \models C_2$, $M'(\mathrm{ab}) \subseteq M'(\mathrm{ab}')$. Assume $M'(\mathrm{ab}) \subset M'(\mathrm{ab}')$. Let $z_1, \ldots, z_{n+1}$ be the interpretations of the varied predicates in M' and let p be $M'(\mathrm{ab})$. Now $M' \models \bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \wedge \forall x(\neg p(x) \rightarrow \neg z_{n+1}(x)) \wedge p < \mathrm{ab}'$. This contradicts $M' \models \neg\exists p z_1 \cdots z_{n+1}(\bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \wedge \forall x(\neg p(x) \rightarrow \neg z_{n+1}(x)) \wedge p < \mathrm{ab}')$. Therefore $M'(\mathrm{ab}) \not\subset M'(\mathrm{ab}')$, and finally $M'(\mathrm{ab}) = M'(\mathrm{ab}')$ and $M' \models \forall x(\mathrm{ab}(x) \leftrightarrow \mathrm{ab}'(x))$. □

Assume that $M' \models C_2 \wedge \neg\phi$ for some model M'. Let $M = M'$. Take any interpretations for predicate variables $p, z_1, \ldots, z_n$ and let $z_{n+1} = p$. Because $M' \models C_2$, $M \models \neg(\bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \wedge p < \mathrm{ab}')$. By Lemma **A**, $M \models \forall x(\mathrm{ab}(x) \leftrightarrow \mathrm{ab}'(x))$. Hence $M \models \neg(\bigwedge \Phi(z_{n+1}, z_1, \ldots, z_n) \wedge p < \mathrm{ab})$. Because the derivation was independent of the interpretations chosen for $p, z_1, \ldots, z_n$, we get $M \models \neg\exists p z_1 \cdots z_n(\bigwedge \Phi(p, z_1, \ldots, z_n) \wedge p < \mathrm{ab})$. Therefore $M \models \bigwedge \Phi \wedge \neg\exists p z_1 \cdots z_n(\bigwedge \Phi(p, z_1, \ldots, z_n) \wedge p < \mathrm{ab}) \wedge \neg\phi$. □

The following results will be presented in a propositional version of circumscription, that is, we restrict to one-element universes.

Lifschitz's circumscription policy for prioritized circumscription determines which abnormality predicates to vary when circumscribing, but it leaves open which ordinary predicates to vary. The purpose of our results is to establish a link between circumscription and preferred subtheories of Brewka, and this goal determines our circumscription policy: all ordinary predicates are varied.

Theorem 6.4.3 *Let $\langle \Sigma, A, \mathcal{P} \rangle$ be a prioritized circumscriptive theory where $A = \{ab_1, \ldots, ab_n\}$. Define $I_i = \{ab_j | 1 \leq j \leq n, not\ ab_i \mathcal{P} ab_j\}$ for all $i \in \{1, \ldots, n\}$. For all $i \in \{1, \ldots, n\}$, let α_i be the definition of ab_i obtained by circumscribing ab_i with everything else varied except abnormality predicates in I_i. For a given $i \in \{1, \ldots, n\}$, let $\{C^i_1, \ldots, C^i_{n_i}\}$ be the set of minimal sets such that $C^i_j \subseteq \{\neg\phi | \phi \in I_i\}$ and $\Sigma \cup C^i_j \cup \{\neg ab_i\}$ is inconsistent. Then for all* *models M, $M \models \Sigma \cup \{\alpha_i\}$ if and only if*

$$M \models \Sigma \cup \left\{ ab_i \leftrightarrow \bigvee_{j \in \{1,\ldots,n_i\}} \bigwedge_{\phi \in C^i_j} \phi \right\}.$$

Proof: The result of circumscribing ab_i with everything else varied except members of I_i is by definition

$$\bigwedge \Sigma(\mathrm{ab}_i, z_1, \ldots, z_m) \wedge \neg\exists a z'_1 \cdots z'_m (\bigwedge \Sigma(a, z'_1, \ldots, z'_m) \wedge \neg a \wedge \mathrm{ab}_i)$$

where $z_1, \ldots, z_m$ are the propositional variables occurring in Σ but not in I_i and $\Sigma(p, p_1, \ldots, p_m)$ denotes Σ with p in place of ab_i and p_i in place of z_i for all $i \in \{1, \ldots, m\}$. This is equivalent to

$$\bigwedge \Sigma(\mathrm{ab}_i, z_1, \ldots, z_m) \wedge \neg\exists z'_1 \cdots z'_m (\bigwedge \Sigma(\bot, z'_1, \ldots, z'_m) \wedge \mathrm{ab}_i)$$

and

$$\bigwedge \Sigma(\mathrm{ab}_i, z_1, \ldots, z_m) \wedge (\mathrm{ab}_i \rightarrow \neg\exists z'_1 \cdots z'_m (\bigwedge \Sigma(\bot, z'_1, \ldots, z'_m))).$$

We denote the second conjunct by α'_i.

($\Rightarrow$) Let M be a model such that $M \models \Sigma \cup \{\alpha'_i\}$. Assume $M \models \mathrm{ab}_i$. Hence $M \models \neg\exists z'_1 \cdots z'_m (\Sigma(\bot, z'_1, \ldots, z'_m))$. We show that $M \models \bigwedge C^i_1 \vee \cdots \vee \bigwedge C^i_{n_i}$. Because $\Sigma \cup \{\mathrm{ab}_i\}$ is consistent and $\bigvee \emptyset = \bot$, $n_i \geq 1$. Assume $M \not\models \bigwedge C^i_j$ for all $j \in \{1, \ldots, n_i\}$. Clearly for all $j \in \{1, \ldots, n_i\}$, $\Sigma \cup \{\neg\mathrm{ab} \in C^i_j | M \models \neg\mathrm{ab}\} \cup \{\neg\mathrm{ab}_i\}$ is consistent. Assume $\Sigma \cup \{\neg\mathrm{ab} \in C^i_j | M \models \neg\mathrm{ab}, j \in \{1, \ldots, n_i\}\} \cup \{\neg\mathrm{ab}_i\}$ is inconsistent. Let C be a minimal subset of $\{\neg\mathrm{ab} \in C^i_j | M \models \neg\mathrm{ab}, j \in \{1, \ldots, n_i\}\}$ such that $\Sigma \cup C \cup \{\neg\mathrm{ab}_i\}$ is inconsistent. Because all members of C are true in M, $C \neq C^i_j$ for all $j \in \{1, \ldots, n_i\}$. This contradicts the fact that $\{C^i_1, \ldots, C^i_{n_i}\}$ is the set of all minimal C such that $\Sigma \cup C \cup \{\neg\mathrm{ab}_i\}$ is inconsistent. Clearly $\Pi = \Sigma \cup \{\neg\mathrm{ab} \in C^i_j | M \models \neg\mathrm{ab}, j \in \{1, \ldots, n_i\}\} \cup \{\neg\mathrm{ab}_i\}$ is consistent. Now $\Pi \cup \{\mathrm{ab} | \neg\mathrm{ab} \in C^i_j, j \in \{1, \ldots, n_i\}, M \not\models \neg\mathrm{ab}\}$ is consistent too, because all occurrences of ab in Σ with $\neg\mathrm{ab} \in C^i_j, j \in \{1, \ldots, n_i\}$, are positive. Now M can be modified to a model of $\Pi \cup \{\mathrm{ab} | \neg\mathrm{ab} \in C^i_j, j \in \{1, \ldots, n_i\}, M \not\models \neg\mathrm{ab}\}$ by changing the truth-values of propositional variables in I_i only, and this contradicts the assumption that $M \models \neg\exists z'_1 \cdots z'_m (\Sigma(\bot, z'_1, \ldots, z'_m))$. Therefore $M \models \bigwedge C^i_1 \vee \cdots \vee \bigwedge C^i_{n_i}$. Assume $M \not\models \mathrm{ab}_i$. Because $\Sigma \cup C^i_j \cup \{\neg\mathrm{ab}_i\}$ is inconsistent for all $j \in \{1, \ldots, n_i\}$, no disjunct of $\bigwedge C^i_1 \vee \cdots \vee \bigwedge C^i_{n_i}$ is true in M. Hence $M \models \alpha_i$.

($\Leftarrow$) Let M be a model such that $M \models \Sigma \cup \{\alpha_i\}$. If $M \not\models \mathrm{ab}_i$, then clearly $M \models \alpha'_i$. If $M \models \mathrm{ab}_i$, then $M \models C^i_j$ for some $j \in \{1, \ldots, n_i\}$. By definition of C^i_j, $\Sigma \cup C^i_j \cup \{\neg\mathrm{ab}_i\}$ is inconsistent. Now $M \not\models \exists z'_1 \cdots z'_m (\bigwedge \Sigma(\bot, z'_1, \ldots, z'_m))$

because $\Sigma(\perp, z'_1, \ldots, z'_m)$ has the same truth value in M as $\Sigma(\mathrm{ab}_i, z'_1, \ldots, z'_m)$, because ab_i does not occur in $z'_1, \ldots, z'_m$ and $M \not\models \neg\mathrm{ab}_i$, and propositional variables in C^i_j belong to I_i and are therefore distinct from $z'_1, \ldots, z'_m$ in α'_i, and $\Sigma(\mathrm{ab}_i, z'_1, \ldots, z'_m) \cup C^i_j \cup \{\neg\mathrm{ab}_i\}$ is inconsistent. Hence $M \models \alpha'_i$. □

First we show that each model of a preferred subtheory is also a model of the theory and the circumscription axioms. The inclusion of models of preferred subtheories in models of prioritized circumscriptive theories in this direction is not subject to the restriction to layered strict partial orders.

Lemma 6.4.4 *Let $\langle \Sigma, A, \mathcal{P}\rangle$ be a prioritized circumscriptive theory. Let $\mathcal{B} = \{\langle \neg\phi, \neg\phi'\rangle | \langle \phi, \phi'\rangle \in \mathcal{P}\}$. Let M be a model of a preferred subtheory S of $\langle \{\neg\phi | \phi \in A\}, \Sigma, \mathcal{B}\rangle$. Then $M \models \Sigma \cup \{\alpha_1, \ldots, \alpha_n\}$, where α_i is the definition of the abnormality predicate $ab_i \in A$.*

Proof: We show that $M \models \{\alpha_1, \ldots, \alpha_n\}$. Because S is a preferred subtheory of $\langle \{\neg\phi | \phi \in A\}, \Sigma, \mathcal{B}\rangle$, there is a total ordering $\mathcal{T}$ on $\{\neg\phi | \phi \in A\}$ such that $\mathcal{B} \subseteq \mathcal{T}$ and a sequence $S_0, \ldots, S_n$ of sets such that $S = S_n$ as given in Definition 6.2.2. Let $\alpha_i = \mathrm{ab}_i \leftrightarrow \bigwedge C^i_1 \vee \cdots \vee \bigwedge C^i_{n_i}$ be the definition of the ith abnormality predicate in $\mathcal{T}$, where C^i_j are the sets defined in Theorem 6.4.3.

Assume $\neg\mathrm{ab}_i \in S_i$. Hence $S_i = S_{i-1} \cup \{\neg\mathrm{ab}_i\}$ and $S_{i-1} \cup \{\neg\mathrm{ab}_i\}$ is consistent. By definition of sets C^i_j, $\Sigma \cup C^i_j \cup \{\neg\mathrm{ab}_i\}$ is inconsistent for all $j \in \{1, \ldots, n_i\}$. Hence $M' \models C^i_j$ for no model M' such that $M' \models \Sigma \cup \{\neg\mathrm{ab}_i\}$. Because $M \models S$ and $S_i \subseteq S$, $M \models S_i$. Because $M \models \Sigma \cup \{\neg\mathrm{ab}_i\}$ and $\Sigma \cup \{\neg\mathrm{ab}_i\} \subseteq S_i$, none of the disjuncts $\bigwedge C^j_i$ may be true in M, and hence both sides of the equivalence $\mathrm{ab}_i \leftrightarrow \bigwedge C^i_1 \vee \cdots \vee \bigwedge C^i_{n_i}$ are false in M. Therefore $M \models \alpha_i$.

Assume $\neg\mathrm{ab}_i \notin S_i$. Hence $S_i = S_{i-1}$ and $S_{i-1} \cup \{\neg\mathrm{ab}_i\}$ is inconsistent. By assumption $M \models S$ and $S_i \subseteq S$. Let J be a minimal subset of S_{i-1} such that $\Sigma \subseteq J$ and $J \cup \{\neg\mathrm{ab}_i\}$ is inconsistent. Now the conjunction of the members of $J \backslash \Sigma$ is by definition equivalent to one of the disjuncts $\bigwedge C^i_j$. Because $M \models J$, $M \models \bigwedge C^i_1 \vee \cdots \vee \bigwedge C^i_{n_i}$. Because $S_{i-1} \cup \{\neg\mathrm{ab}_i\}$ is inconsistent, $S_{i-1} \models \mathrm{ab}_i$. Therefore $M \models \mathrm{ab}_i$. Because both sides of the equivalence $\mathrm{ab}_i \leftrightarrow \bigwedge C^i_1 \vee \cdots \vee \bigwedge C^i_{n_i}$ are true in M, $M \models \alpha_i$.

Because $\Sigma \subseteq S$ and $M \models \alpha_i$ for all $i \in \{1, \ldots, n\}$, $M \models \Sigma \cup \{\alpha_1, \ldots, \alpha_n\}$. □

Then we show that each model of the theory and the circumscription axioms is also a model of a corresponding preferred subtheory. The inclusion in this direction relies on the layeredness condition on the strict partial orders.

Lemma 6.4.5 *Let $\langle \Sigma, A, \mathcal{P}\rangle$ be a prioritized circumscriptive theory. Let $\mathcal{P}$ be a layered strict partial order on $A = \{ab_1, \ldots, ab_n\}$. Let M be a model of $\Sigma \cup \mathcal{A}$ where $\mathcal{A}$ consists of the definitions of the abnormality predicates in A.*

Then there is a preferred subtheory S of $\langle\{\neg\phi|\phi \in A\}, \Sigma, \mathcal{B}\rangle$ such that $M \models S$, where $\mathcal{B} = \{\langle\neg\phi, \neg\phi'\rangle|\langle\phi, \phi'\rangle \in \mathcal{P}\}$.

Proof: Define the ordering $\mathcal{T}$ on A as follows: $\mathcal{P} \subseteq \mathcal{T}$ and $\mathrm{ab}\mathcal{T}\mathrm{ab}'$ whenever neither $\mathrm{ab}\mathcal{P}\mathrm{ab}'$ nor $\mathrm{ab}'\mathcal{P}\mathrm{ab}$ and $M \models \neg\mathrm{ab}$ and $M \not\models \neg\mathrm{ab}'$. This is the same construction that is used in the proof of Theorem 3.1.15. Let $\neg\mathrm{ab}_1, \ldots, \neg\mathrm{ab}_n$ be the ordering of $\{\neg\phi|\phi \in A\}$ according to $\mathcal{T}$. We show by induction that $S = \Sigma \cup \{\neg\phi|\phi \in A, M \models \neg\phi\}$ is a preferred subtheory of $\langle\{\neg\phi|\phi \in A\}, \Sigma, \mathcal{B}\rangle$.

Induction hypothesis: $S_i = S \cap (\Sigma \cup \{\neg\mathrm{ab}_1, \ldots, \neg\mathrm{ab}_i\})$.

Base case $i = 0$: Immediate as $S_0 = \Sigma$.

Inductive case $i \geq 1$: By induction hypothesis $S_{i-1} = S \cap (\Sigma \cup \{\neg\mathrm{ab}_1, \ldots, \neg\mathrm{ab}_{i-1}\})$. Assume $\neg\mathrm{ab}_i \in S$ and therefore $M \models \neg\mathrm{ab}_i$. Because $M \models \alpha_i$, $M \not\models C_j^i$ for all $j \in \{1, \ldots, n_i\}$, where the sets C_j^i are defined in Theorem 6.4.3. Therefore for all subsets C of $\{\neg\mathrm{ab}_1, \ldots, \neg\mathrm{ab}_n\}$ such that C is consistent and $C \cup \{\neg\mathrm{ab}_i\}$ is inconsistent, at least one member ϕ of C is false in M. Therefore no such set C is contained in S_{i-1}. Therefore $S_{i-1} \cup \{\neg\mathrm{ab}_i\}$ is consistent and $S_i = S_{i-1} \cup \{\neg\mathrm{ab}_i\}$.

Assume $\neg\mathrm{ab}_i \notin S$ and therefore $M \models \mathrm{ab}_i$. Because $M \models \alpha_i$, $M \models C_j^i$ for at least one $j \in \{1, \ldots, n_i\}$. We have to show that $C_j^i \subseteq S_{i_1}$. By definition of C_j^i, $\mathrm{ab}\mathcal{P}\mathrm{ab}_i$ for no abnormality predicate ab occurring in C_j^i. By definition of $\mathcal{T}$ and because $M \models \mathrm{ab}_i$, $\mathrm{ab}\mathcal{T}\mathrm{ab}_i$ for all ab that occur in C_j^i and $M \models \neg\mathrm{ab}$. By induction hypothesis now $\neg\mathrm{ab} \in S_{i-1}$ for all $\neg\mathrm{ab} \in C_j^i$. Because $C_j^i \subseteq S_{i-1}$ and $\Sigma \cup C_j^i \subseteq S_{i-1}$ is inconsistent with $\neg\mathrm{ab}_i$, $S_i = S_{i-1}$ and therefore $\neg\mathrm{ab}_i \notin S_i$.
□

The restriction in the above lemma to layered strict partial orders is justified by the following example that shows that for non-layered strict partial orders there are models of the theory and the circumscription axioms that are not models of any preferred subtheory.

Example 6.4.1 Let $\Sigma = \{\neg\mathrm{ab}_a \rightarrow a, \neg\mathrm{ab}_b \rightarrow b, \neg\mathrm{ab}_{\neg a} \rightarrow \neg a, \neg\mathrm{ab}_{\neg b} \rightarrow \neg b\}$. Let $\mathcal{P}$ be the following strict partial order on the abnormality predicates.

$$\begin{tikzcd} \mathrm{ab}_b \arrow[d, dash] & \mathrm{ab}_a \arrow[d, dash] \\ \mathrm{ab}_{\neg a} & \mathrm{ab}_{\neg b} \end{tikzcd}$$

By circumscribing the abnormality predicates according to Lifschitz's policy we get $C = \{\mathrm{ab}_{\neg a} \leftrightarrow \neg\mathrm{ab}_a, \mathrm{ab}_{\neg b} \leftrightarrow \neg\mathrm{ab}_b, \mathrm{ab}_b \leftrightarrow \neg\mathrm{ab}_{\neg b}, \mathrm{ab}_a \leftrightarrow \neg\mathrm{ab}_{\neg a}\}$. In addition to the intended models M_1, M_2, M_3 that respectively correspond to the preferred subtheories $S_1 = \Sigma \cup \{\neg\mathrm{ab}_{\neg a}, \neg\mathrm{ab}_{\neg b}\}$, $S_2 = \Sigma \cup \{\neg\mathrm{ab}_{\neg a}, \neg\mathrm{ab}_b\}$

and $S_3 = \Sigma \cup \{\neg \mathrm{ab}_{\neg b}, \neg \mathrm{ab}_a\}$ there is the model M' such that $M' \models \Sigma \cup C$ and $M' \models a \wedge b$. ■

The definitions of prioritized circumscription in [Lifschitz, 1985; Lifschitz, 1987] make a restriction to layered strict partial orderings of abnormality predicates. Grosof [1991] proposes a generalization of prioritized circumscription to arbitrary strict partial orders. He, however, considers only the obvious model-theoretic definition of model minimization that is similar to the definition of preferred subtheories with layered strict partial orders, and he does not address the question of constructing the circumscription axioms. One potential solution to this problem is to find a reduction to circumscription without priorities, but our example above shows that the obvious way of doing it fails.

6.5 Minimal models in conditional entailment

Geffner and Pearl [1992] define a system for default reasoning that is based on an ordering on propositional models that is induced by priorities on a set of formulae. They call it *conditional entailment*. In this section we show how their minimal model semantics can be embedded in prioritized default logic.

Default theories $T = \langle K, V\rangle$ consist of a background $K = \langle L, D\rangle$ and a set of evidence $V \subseteq \mathcal{L}$. Here $L \subseteq \mathcal{L}$ and D is a set of conditionals that are not used in the definition of minimal models. The formulae in L may have occurrences of members of a designated set of propositional variables $\Delta_{\mathcal{L}} = \{\delta_1, \delta_2, \ldots\}$ called *assumptions*. For a model M, $\Delta[M]$ denotes the set of members δ of $\Delta_{\mathcal{L}}$ such that $M \not\models \delta$.

Definition 6.5.1 ([Geffner and Pearl, 1992]) *A* prioritized preferential structure *is a quadruple $\langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, \prec\rangle$ where $\mathcal{I}_{\mathcal{L}}$ stands for the set of interpretations over the underlying language $\mathcal{L}$, $\Delta_{\mathcal{L}}$ stands for the set of assumptions in $\mathcal{L}$, $\prec$ stands for an irreflexive and transitive priority relation over $\Delta_{\mathcal{L}}$, and $<$ is a binary relation over $\mathcal{I}_{\mathcal{L}}$ such that for two interpretations M and M', $M < M'$ holds if and only if $\Delta[M] \neq \Delta[M']$ and for every assumption δ in $\Delta[M] \backslash \Delta[M']$ there exists an assumption δ' in $\Delta[M'] \backslash \Delta[M]$ such that $\delta \prec \delta'$.*

Definition 6.5.2 ([Geffner and Pearl, 1992]) *A* preferred model of $T = \langle K, V\rangle$ *in a prioritized preferential structure $\langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, \prec\rangle$ is a model M such that $M \models L$ and $M \models V$ and there is no model M' such that $M' < M$ and $M' \models L$ and $M' \models V$.*

In the following we give an embedding of the Geffner and Pearl model minimization in prioritized default logic and show that the consequence relation used by Geffner and Pearl coincides with the consequence relation in prioritized default logic. The translation and the correspondence proofs are variants of those presented in Section 6.3 for ordered theory presentations.

Definition 6.5.3 *Let $T = \langle K, V \rangle$ be a default theory of [Geffner and Pearl, 1992] where $K = \langle L, D \rangle$. Let $S = \langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, \prec \rangle$ be a prioritized preferential structure. Let $\Delta_{\mathcal{L}} = \{\delta_1, \ldots, \delta_n\}$. Define the translation of T to default logic as $\Delta_T = \langle D_T, L \cup V \rangle$ where $D_T = \{:\delta/\delta | \delta \in \Delta_{\mathcal{L}}\}$, and the translation of the strict partial order on the assumptions in S as $\mathcal{P} = \{\langle :\delta/\delta, :\delta'/\delta' \rangle | \delta' \prec \delta\}$.*

Lemma 6.5.4 *Let S be a maximal consistent subset of T. Let M be a model such that $M \models S$. Let $\phi \in T \backslash S$. Then $M \not\models \phi$.*

Proof: Assume $M \models S \cup \{\phi\}$. Then $S \cup \{\phi\}$ is a consistent subset of T which contradicts the maximality of S. □

Lemma 6.5.5 *Let $\mathcal{S} = \langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, \prec \rangle$ be a prioritized preferential structure. Let E and E' be extensions of Δ_T and $M \in \mathcal{I}_{\mathcal{L}}$ and $M' \in \mathcal{I}_{\mathcal{L}}$ such that $M \models E$ and $M' \models E'$. Then $E \sqsubset_{\mathcal{P}}^{\Delta_T} E'$ if and only if $M < M'$.*

Proof: Assume $E \sqsubset_{\mathcal{P}}^{\Delta_T} E'$. Assume that $M \not\models \delta$ and $M' \models \delta$ for some $\delta \in \Delta_{\mathcal{L}}$. Hence $\delta \in \Delta[M] \backslash \Delta[M']$. Now $\delta \notin E$ and $\neg\delta \notin E'$ and because $E' \cap (L \cup V \cup \Delta_{\mathcal{L}})$ is by Lemma 6.2.9 a maximal consistent subset of $L \cup V \cup \Delta_{\mathcal{L}}$, $\delta \in E'$. Hence appl$(:\delta/\delta, E', E)$. By the assumption $E \sqsubset_{\mathcal{P}}^{\Delta_T} E'$ and Theorem 3.1.14, appl$(:\delta/\delta, E', E)$ implies that for some $:\delta'/\delta' \in D_T, :\delta'/\delta' \mathcal{P} :\delta/\delta$ and appl$(:\delta'/\delta', E, E')$. Now $M \models \delta'$ and by Lemma 6.5.4 $M' \not\models \delta'$. Because $\delta \prec \delta'$, $M < M'$.

Assume $M < M'$. Assume that appl$(:\delta/\delta, E', E)$ for some default $:\delta/\delta \in D_T$. Now $M' \models \delta$ and by Lemma 6.5.4 $M \not\models \delta$. Therefore by definition of $<$ there is δ' such that $M \models \delta'$ and $M' \not\models \delta'$ and $\delta \prec \delta'$. Now by Lemma 6.5.4 $\delta' \in E$ and $\delta' \notin E'$. Hence appl$(:\delta'/\delta', E, E')$ and $\delta' \mathcal{P} \delta$. Therefore $E \sqsubset_{\mathcal{P}}^{\Delta_T} E'$. □

Lemma 6.5.6 *Let $T = \langle K, V \rangle$ be a default theory and $\phi \in \mathcal{L}$ a formula. If $\phi \notin E$ for some $\mathcal{P}$-preferred$_2$ extension E of Δ_T, then there is a preferred model M of T such that $M \models L \cup V$ and $M \not\models \phi$.*

Proof: Let E be a $\mathcal{P}$-preferred$_2$ extension of Δ_T such that $\phi \notin E$. Hence there is a model M such that $M \models E \cup \{\neg\phi\}$. Assume there is N such that $N < M$. The set $\Omega = \{\phi \in L \cup V \cup \Delta_{\mathcal{L}} | N \models \phi\}$ is a consistent subset of $X = L \cup V \cup \Delta_{\mathcal{L}}$ but not necessarily maximal. Let Ω' be a maximal consistent set of X such that $\Omega \subseteq \Omega' \subseteq X$. Because Ω' is consistent, it has a model N'. Now $N \models \delta$ and $N' \not\models \delta$ for no $\delta \in \Delta_{\mathcal{L}}$ because for all $\delta \in \Delta_{\mathcal{L}}$, $N \models \delta$

implies $N' \models \delta$. Therefore $N' < N$ and by the transitivity of $<$ [Geffner and Pearl, 1992, Lemma 3.2] $N' < M$. Because Ω' is a maximal consistent subset of X, there is by Lemma 6.2.8 an extension $E' = Cn(\Omega')$ of Δ_T. By Lemma 6.5.5 $E' \sqsubset_{\mathcal{P}}^{\Delta_T} E$. This contradicts the $\mathcal{P}$-preferredness$_2$ of E. Hence there is no $N < M$, and M is a preferred model of T. □

Lemma 6.5.7 *Let $T = \langle K, V\rangle$ be a default theory and $\phi \in \mathcal{L}$ a formula. If $M \not\models \phi$ for some preferred model M of T, then there is a $\mathcal{P}$-preferred$_2$ extension E of Δ_T such that $\phi \notin E$.*

Proof: Let M be a preferred model of T such that $M \not\models \phi$. By Lemma [Geffner and Pearl, 1992, Lemma 3.3] $\Pi = L \cup V \cup \{\delta \in \Delta_{\mathcal{L}} | M \models \delta\}$ is a maximal consistent subset of $L \cup V \cup \Delta_{\mathcal{L}}$. Hence by Lemma 6.2.8 $Cn(\Pi)$ is an extension of Δ_T. Because $M \models \Pi$ and $M \not\models \phi$, $\Pi \not\models \phi$. $\Pi \not\models \phi$ implies $\phi \notin Cn(\Pi)$. Assume $E = Cn(\Pi)$ is not a $\mathcal{P}$-preferred$_2$ extension, that is, there is an extension E' of Δ_T such that $E' \sqsubset_{\mathcal{P}}^{\Delta_T} E$. Let N be a model such that $N \models E'$. By Lemma 6.5.5 $N < M$. This contradicts the fact that M is a preferred model. Hence E is $\mathcal{P}$-preferred$_2$ extension of Δ_T and $\phi \notin E$. □

Theorem 6.5.8 *Let $T = \langle K, V\rangle$ be a default theory where $K = \langle L, D\rangle$. Let $\Delta_{\mathcal{L}} = \{\delta_1, \ldots, \delta_n\}$. Let $\Delta_T = \langle D_T, L \cup V\rangle$ where $D_T = \{:\delta/\delta | \delta \in \Delta_{\mathcal{L}}\}$, and $\mathcal{P} = \{\langle :\delta/\delta, :\delta'/\delta'\rangle | \delta' \prec \delta\}$. Let ϕ be a formula. Then ϕ is true in all preferred models of T if and only if $\Delta_T \models_{\mathcal{P}}^2 \phi$.*

Proof:

For all preferred models M of $L \cup V$, $M \models \phi$
  iff there is no preferred model M' such that $M' \not\models \phi$
  iff there is no $\mathcal{P}$-preferred$_2$ extension E such that $\phi \notin E$
  iff $\phi \in E$ for all $\mathcal{P}$-preferred$_2$ extensions E.

The second equivalence is by Lemmata 6.5.6 and 6.5.7: $\phi \notin E$ for some $\mathcal{P}$-preferred$_2$ extension E of Δ_T if and only if $M \not\models \phi$ for some preferred model M of $L \cup V$. □

Chapter 7

Inheritance reasoning in nonmonotonic logics

In this chapter we discuss lexicographic priority mechanisms in connection with translational theories of inheritance reasoning. We first investigate the problem of computing priorities that make default logic satisfy specificity. Then we present a translation of inheritance networks to default logic with priorities, and compare the translation to earlier translational theories of defeasible inheritance. Finally, we discuss the obstacles that are encountered in trying to use priorities in a setting without the restriction made in inheritance networks, that is, without an acyclicity condition and the restriction to non-disjunctive prerequisites and conclusions in defaults.

7.1 Priorities induced by specificity

7.1.1 Background

We take defeasible inheritance networks [Horty, 1994] as the paradigmatic framework of defeasible reasoning that addresses the problem of specificity. The basic form of inheritance networks [Hayes, 1979] does not support the expression of defeasible information. Inheritance networks are a formalism for representing information about a set of individuals that are classified to a number of classes. Only some of the classes an individual belongs to are stated explicitly, and the inclusion relation between classes is stated only partially. The idea is that the explicit sentences about class membership and inclusion can be used in inferring more sentences about membership and inclusion. The utility of this kind of representation is that the modification of the information concerning to a whole class of individuals can be made to the class, instead of all individuals separately. Inheritance networks contain less redundancy than representations where information concerning an individual is always directly associated with the individual.

Extensions of the above kind of inheritance networks have been proposed to support the expression of defeasible information [Horty, 1994]. Now there are sentences that express that members of a class are or are not *usually* members of another class. These sentences allow exceptions so that sentences stating that members of A are usually members of B and an individual is a member of A and not a member of B do not lead to an inconsistency. Without defeasible sentences inheritance can be completely understood in terms of classical monotonic logics. The introduction of defeasible information takes inheritance reasoning outside the realm of classical logics. The problem is which defeasible sentences to reject in cases of conflict. The research on defeasible inheritance concentrates on this question.

7.1.2 A principle of specificity

A goal of this work is to find representations of inheritance networks in prioritized default logic and related nonmonotonic formalisms. It will turn out that in the framework of default logic and lexicographic priorities the question of what is the correct definition of preemption that is used in path-based theories of inheritance [Horty, 1994] does not arise. This is because in our framework, default rules are prioritized, and a conflict between two defaults is always resolved in the same way irrespective of which paths (in the sense of path-based inheritance theories) are involved. Because the degrees of freedom in the framework of prioritized default logic are more restricted, it is more straightforward to arrive to a definition of inheritance reasoning. We will see that a simple and intuitive principle that is satisfied by most path-based inheritance theories and nonmonotonic logics that respect specificity gives us a good basis for specificity-based priorities. This principle leads to a definition of priorities for a subclass of prioritized default logic that corresponds to acyclic inheritance networks. It has not been earlier used in the context of default logic. The requirement that default theories with priorities satisfy this principle pose certain requirements to priorities. We discuss the computation of priorities so that the principle is satisfied. Our analysis suggests that there are two quite separate justifications for giving one default a higher priority than to another, and therefore priorities in inheritance reasoning serve two separate purposes. The first reason to impose an ordering on two defaults, well-known from earlier work on specificity, is to resolve conflicts between conflicting defaults when some of the defaults are more specific than others. The second reason to order two defaults is to indicate that the way conflicts related to the first default are resolved does not depend on whether the second default is affected. Later we will be able to trace some of the problems encountered in earlier work to the fact that the need for the second reason to order defaults has not been recognized.

Poole [1991] argues for the distinction between background facts and contingent facts to be made in defeasible reasoning with specificity. Informally, background facts express general knowledge about the problem domain a knowledge base represents, and contingent facts are knowledge about particular problem instances. For example in the penguin domain, the fact that Tweety is a penguin is contingent, and the rules that penguins are birds, penguins do not usually fly, and birds usually do, constitute the background. The same or similar distinction is used in conditional entailment [Geffner and Pearl, 1992], preferential logics [Kraus *et al.*, 1990], probabilistic semantics for nonmonotonic reasoning [Pearl, 1988], and probability theory [Pearl, 1989]. We adopt this distinction between background and contingent facts, and it has a role in the definition of priorities for default theories. The priorities on the defaults of a given default theory are independent of the contingent facts in the default theory. Hence given priorities $\mathcal{P}$ for default theories $\langle D, B \cup C \rangle$ with background B are the same irrespective of what the set C is. Hence we compute priorities for $\langle D, B \rangle$ and use them with reasoning with arbitrary default theories $\langle D, B \cup C \rangle$.

Many formalizations of defeasible reasoning that address specificity employ the following principle [Kraus *et al.*, 1990; Lehmann and Magidor, 1992; Adams, 1975; Pearl, 1988; Pearl, 1990; Goldszmidt *et al.*, 1990; Geffner and Pearl, 1992]. If the default $\alpha \rightharpoonup \beta$ is in D, then from the contingent facts $\{\alpha\}$ the formula β can be inferred irrespective of what other defaults there are in D. This principle resembles the rule of modus ponens in classical logics, but is more restricted in the respect that the conclusion β is not sanctioned if the contingent facts are a superset of $\{\alpha\}$. Intuitively the principle can be understood as saying that if α is all the information available, and if typically β whenever α, then conclude β. In the penguin triangle, if p is read as "x is a penguin", b as "x is a bird", and f as "x is able to fly", and if b is all the information available, conclude f. If p is all the information available, conclude $\neg f$.

The standard consequence relation for default logic is cautious reasoning $\Delta \models_c \phi$, which means that the formula ϕ is belongs to all extensions of the default theory Δ. A counterpart of the above principle is not fulfilled by this consequence relation. For normal default theories $\langle D, B \rangle$ the principle can be expressed as follows.

$$\langle D, B \cup \{\alpha\} \rangle \models_x \beta \text{ for all } \frac{\alpha : \beta}{\beta} \in D \tag{7.1}$$

Here D is the set of default rules of a default theory, B is a set of sentences

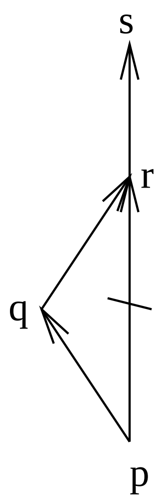


Figure 7.1: An extended penguin triangle

that forms the background, and $\models_x$ is a consequence relation.

Example 7.1.1 Let $D = \{p{:}b/b, b{:}f/f, p{:}\neg f/\neg f\}$. The default theory $\Delta = \langle D, \{p\}\rangle$ has the extensions $E_1 = Cn(\{p, b, \neg f\})$ and $E_2 = Cn(\{p, b, f\})$. Because $\neg f \notin E_2$, $\Delta \not\models_c \neg f$. ■

We investigate consequence relations for default logic that treat specificity properly in the sense that they fulfill Condition 7.1. As shown above, $\models_c$ is not such a relation. A natural way to strengthen it is to consider only a subset of the extensions, those in which conflicts are resolved without violating specificity, and take their intersection. Such a subset of extensions can be selected by priorities.

Note that arguments against Condition 7.1 can be presented. For example following Lifschitz [1993], the propositions p and b in the preceding example can be considered *parameters* of the default theory in question and it would be required that our inferences concerning the default theory are monotonic with respect to the parameters. Given the fact b we would not be allowed to conclude f, because the addition of the fact p would force us to retract this conclusion. This is a question of what kind of conventions are used in representing knowledge. Most researchers have adopted notions of specificity that are compatible with Condition 7.1.

7.1.3 Computation of priorities according to the principle

The next example, although simple, demonstrates some of the important properties that priorities have to have in order to satisfy Condition 7.1.

Example 7.1.2 For the default theory $\langle D, \emptyset \rangle$ depicted in Figure 7.1 the consequence relation $\models_{\mathcal{P}}$ satisfies Condition 7.1, where

$$\mathcal{P} = \left\{ \left\langle \frac{p : \neg r}{\neg r}, \frac{q : r}{r} \right\rangle, \left\langle \frac{p : \neg r}{\neg r}, \frac{r : s}{s} \right\rangle \right\}.$$

The default theory $\Delta = \langle D, \{p\} \rangle$ has the extensions E_1 and E_2 such that $\mathrm{GD}(E_1, \Delta) = \{p{:}\neg r/\neg r, p{:}q/q\}$ and $\mathrm{GD}(E_2, \Delta) = \{p{:}q/q, q{:}r/r, r{:}s/s\}$. Only E_1 is $\mathcal{P}$-preferred. The two extensions and all strict total orders $\mathcal{T}$ on the defaults that extend $\mathcal{P}$ are depicted below on the left (the default $p{:}q/q$ is omitted because it is applied in both extensions and therefore does not figure in the definition of preferred extensions.) If $\langle p{:}\neg r/\neg r, r{:}s/s \rangle$ were not in $\mathcal{P}$, the strict total order $r{:}s/s, p{:}\neg r/\neg r, p{:}q/q, q{:}r/r$ would extend $\mathcal{P}$, and consequently E_2 would be a $\mathcal{P}$-preferred extension in which $p{:}\neg r/\neg r$ is not applied. The diagram on the right depicts the unwanted total order.

$$\begin{array}{ccc} & E_1 & E_2 \\ \frac{r : s}{s} & \circ & \bullet \\ & | & | \\ \frac{q : r}{r} & \circ & \bullet \\ & | & | \\ \frac{p : \neg r}{\neg r} & \bullet & \circ \end{array} \qquad \begin{array}{ccc} & E_1 & E_2 \\ \frac{q : r}{r} & \circ & \bullet \\ & | & | \\ \frac{r : s}{s} & \circ & \bullet \\ & | & | \\ \frac{p : \neg r}{\neg r} & \bullet & \circ \end{array} \qquad\qquad \begin{array}{ccc} & E_1 & E_2 \\ \frac{q : r}{r} & \circ & \bullet \\ & | & | \\ \frac{p : \neg r}{\neg r} & \bullet & \circ \\ & | & | \\ \frac{r : s}{s} & \circ & \bullet \end{array}$$

■

The fact that $p{:}\neg r/\neg r$ should be given a higher priority than $r{:}s/s$ leads to a general rule concerning priorities that satisfy Condition 7.1. A default $\alpha{:}\beta/\beta \in D$ should be given a higher priority than default δ that is applied in an extension E of $\langle D, B \cup \{\alpha\} \rangle$ that does not contain β, and is not applied in extensions that contain β. Here the extension E should not be a preferred extension of $\langle D, B \cup \{\alpha\} \rangle$. The rule for giving priorities on the defaults guarantees that whenever there is a trade-off between $\alpha{:}\beta/\beta$ and δ, priorities indicate that δ should be traded for $\alpha{:}\beta/\beta$. Later it will be obvious that Condition 7.1 alone is in general far too weak, but that the informal rule stated above indicates a way to strengthen the priorities in an appropriate manner in those cases as well.

Some general remarks on priorities are worth noticing at this point. First, violation of Condition 7.1 is sometimes unavoidable, for example when $D = \{a{:}b/b, a{:}\neg b/\neg b\}$ and the background B together with $\{a\}$ is consistent. The default theory does not have extensions that contain both b and $\neg b$. Second, priorities that fulfill Condition 7.1 are not in general unique. Consider Example 7.1.2. Because of the monotonicity of default reasoning with respect to priorities (Lemma 3.1.17) any strict partial order that extends $\mathcal{P}$ fulfills Condition 7.1. Hence the goal we set does not uniquely determine what kind of priorities should be computed unless some further principles are adopted.

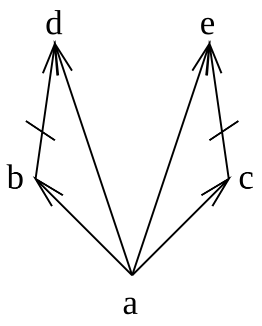


Figure 7.2: An inheritance network

In Theorem 7.1.1 we give a sufficient but not necessary condition for priorities to satisfy Condition 7.1. Of course, conditions that are sufficient but not necessary may be completely useless, for example by being too strong. The following example illustrates what is involved in ensuring that the condition the priorities have to satisfy is not too strong.

Example 7.1.3 Let $\Delta = \langle D, \{a\}\rangle$ be a default theory where the defaults D are depicted in Figure 7.2. The following are the extensions of Δ.

$$\begin{array}{rclrcl} E_1 & = & Cn(\{a,b,c,d,e\}) & E_2 & = & Cn(\{a,b,c,\neg d,e\}) \\ E_3 & = & Cn(\{a,b,c,d,\neg e\}) & E_4 & = & Cn(\{a,b,c,\neg d,\neg e\}) \end{array}$$

Consider the default $a{:}d/d$. Good extensions for $a{:}d/d$ are E_1 and E_3, and bad extensions are E_2 and E_4. Now the defaults $b{:}\neg d/\neg d$ and $a{:}e/e$ would seem to become applicable – they are in $\mathrm{GD}(E_2,\Delta)\backslash\mathrm{GD}(E_3,\Delta)$ – if the conflict on d is resolved in the wrong way. Therefore it would seem that $a{:}d/d$ should be preferred to both of these defaults. A similar reasoning concerning $a{:}e/e$ shows that $a{:}e/e$ should be given a higher priority than $a{:}d/d$ and $c{:}\neg e/\neg e$. This is impossible, because now $a{:}d/d\mathcal{P}a{:}e/e$ and $a{:}e/e\mathcal{P}a{:}d/d$. There is, however, the strict partial order $\mathcal{P}' = \{\langle a{:}d/d, b{:}\neg d/\neg d\rangle, \langle a{:}e/e, c{:}\neg e/\neg e\rangle\}$ on D that fulfills Condition 7.1. Hence this example shows that the rule "a default that becomes applied when a conflict is resolved in the wrong way" should be applied cautiously. The application of $a{:}e/e$ in E_3 and not in E_2 is not related to the way the conflict concerning d is resolved because each conflict can be resolved in two ways independently of the other. ■

We do not have a general solution to the problem pointed out in the example above, and consequently there is the possibility that for some useful default theories the condition given in Theorem 7.1.1 cannot be satisfied. The problem of conflicts between defaults that are unrelated is reduced by requiring that extensions that incorrectly do not apply a default are compared to extensions that apply the default and otherwise differ minimally. In the above example the computation of priorities related to the default $a{:}d/d$ would be based on comparing the extensions E_1 and E_3 on one hand and on comparing the extensions E_2 and E_4 on the other. Here each pair differs on the application of $a{:}d/d$ and differ minimally with respect to the application of other defaults.

The priorities defined in Theorem 7.1.1 for a set of defaults D and a background B are based on a function f that gives for each δ the set $f(\delta, D, B)$ that can be thought of as the set of defaults that may become applied if a conflict is resolved so that Condition 7.1 is violated for δ. To avoid this, δ has to be given a higher priority than any of the defaults in $f(\delta, D, B)$.

Theorem 7.1.1 *Let $\Delta = \langle D, B\rangle$ be a default theory where D is finite. Define $f(\alpha{:}\beta/\beta, D, B)$ as the set*

$$\begin{aligned}\{\delta \in D|\ & E \textit{ and } E' \textit{ are extensions of } \langle D, B \cup \{\alpha\}\rangle,\\ & \mathit{appl}(\delta, E', E), \mathit{appl}(\alpha{:}\beta/\beta, E, E'),\\ & \textit{for no extension } E'' \textit{ of } \langle D, B \cup \{\alpha, \beta\}\rangle\\ & (\mathit{GD}(E', \Delta) \vartriangle \mathit{GD}(E'', \Delta)) \subset (\mathit{GD}(E', \Delta) \vartriangle \mathit{GD}(E, \Delta))\}.\end{aligned}$$

Let $\mathcal{P}$ be a strict partial order on D such that $\delta\mathcal{P}\delta'$ for all $\delta \in D$ and $\delta' \in f(\delta, D, B)$. Then $\models_{\mathcal{P}}$ fulfills Condition 7.1 for $\langle D, B\rangle$.

Proof: Assume E' is an extension of $\langle D, B \cup \{\alpha\}\rangle$ so that $\neg\beta \in E'$. Let E be an extension of $\langle D, B \cup \{\alpha\}\rangle$ so that $\beta \in E$ and $(\mathrm{GD}(E', \Delta) \vartriangle \mathrm{GD}(E'', \Delta)) \subset (\mathrm{GD}(E', \Delta) \vartriangle \mathrm{GD}(E, \Delta))$ for no extension E'' such that $\mathrm{appl}(\alpha{:}\beta/\beta, E'')$. Such an extension clearly exists because the sets $\mathrm{GD}(E'', \Delta) \subseteq D$ are finite for all extensions E'' of Δ. Because $\beta \in E\backslash E'$, $\mathrm{appl}(\alpha{:}\beta/\beta, E, E')$. Now for all δ such that $\mathrm{appl}(\delta, E', E)$, $\delta \in f(\alpha{:}\beta/\beta, D, B)$. Let $\mathcal{T}$ be any strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$. Now $\mathrm{appl}(\alpha{:}\beta/\beta, E, E')$ but because $\alpha{:}\beta/\beta\mathcal{P}\delta$ for all δ such that $\mathrm{appl}(\delta, E', E)$ there is no $\delta\mathcal{T}\alpha{:}\beta/\beta$ such that $\mathrm{appl}(\delta, E', E)$. Therefore there are no $\Delta, \mathcal{P}$-characteristic orderings for E' and E' is not a $\mathcal{P}$-preferred extension of $\langle D, B \cup \{\alpha\}\rangle$. Therefore β is in all $\mathcal{P}$-preferred extensions of $\langle D, B \cup \{\alpha\}\rangle$ and $\langle D, B \cup \{\alpha\}\rangle \models_{\mathcal{P}} \beta$. □

Condition 7.1 imposes only a lower bound on the conclusions that should be sanctioned by a consequence relation. Reasoning with lexicographic priorities is monotonic with respect to the priorities: the more priorities there are, the more conclusions can be obtained. Consider the Nixon Diamond. No priorities on the defaults violate Condition 7.1. Some priorities prefer "quakers are pacifists" to "republicans are not pacifists". There is no reason to do so. It would seem that demanding that priorities are minimal would solve this problem. This is not the case. Let

$$D = \left\{\frac{p : q}{q}, \frac{p : a \vee b}{a \vee b}, \frac{q : \neg a}{\neg a}, \frac{q : \neg b}{\neg b}\right\}.$$

If the second default is preferred over one or both of the last two, Condition 7.1 is satisfied. However, minimal such priorities give arbitrarily preference to one of the last two defaults when reasoning with $\langle D, \{p\}\rangle$. In conditional entailment [Geffner and Pearl, 1992], this problem is avoided by considering all (minimal) priorities that satisfy a counterpart of Condition 7.1.

Two rules for specificity-based priorities

The principle expressed as Condition 7.1 is too weak to capture many intended inferences simply because all it states is that whenever the premises consist of the prerequisite of a single default, then the conclusion of that default should be inferred. The principle says nothing about the inferences concerning other defaults, and nothing about the case when the contingent facts are something else than the prerequisite of a single default. Now it is not surprising that the priorities given in Theorem 7.1.1 are not in general sufficient for inheritance reasoning or default reasoning in general. The weaknesses are pointed out in Example 7.1.5. However, the prior analysis of priorities for lexicographic priority mechanisms has been useful in uncovering two informal rules that guide the construction of priorities that have more general applicability. Each rule gives an intuitive justification for a part of the elements in the priorities that are needed in inheritance reasoning.

1. The first rule states that among directly conflicting defaults, that is defaults that have mutually contradictory conclusions, more specific defaults should be given a higher priority. In inheritance networks, a default is more specific than another default if there is a path consisting of positive links from the prerequisite of the former to the prerequisite of the latter.

2. The second rule states that more specific defaults involved in a conflict should be given a higher priority than any of the defaults that may become applied if the conflict is incorrectly resolved in favour of less specific defaults, even if these defaults are not directly involved in the conflict. In the inheritance reasoning setting these defaults are those that are higher in an inheritance network.

The priorities required by the second rule have not been earlier discussed in the nonmonotonic reasoning literature. A form of priorities implicitly used in the context of "explanations" in the translational inheritance theory by Gelfond and Przymusinska [1990] obey also the second rule. As an example of the application of the second rule for a more complicated default theory we give a variant of Example 7.1.2.

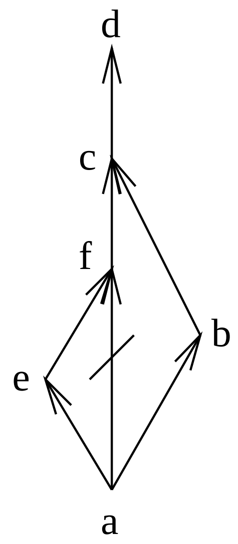


Figure 7.3: Another inheritance network

Example 7.1.4 Let the defaults be $D = \{p{:}\neg r/\neg r, p{:}q/q, q{:}r/r, r \wedge a{:}s/s\}$ and the background is empty. The set-inclusion minimal strict partial order that satisfies the condition in Theorem 7.1.1 is $\mathcal{P} = \{\langle p{:}\neg r/\neg r, q{:}r/r\rangle\}$, but $\langle D, \{p, a\}\rangle \not\models_{\mathcal{P}} \neg r$. The conclusions with $\mathcal{P}$ are too weak. ■

The second rule suggests that $p{:}\neg r/\neg r$ should have a higher priority than $r \wedge a{:}s/s$ because for the contingent facts $\{p, a\}$ the default $r \wedge a{:}s/s$ becomes applied if the conflict on r is resolved by rejecting $p{:}r/r$. The insufficiency of Condition 7.1 for acyclic inheritance networks is demonstrated by the following example.

Example 7.1.5 Let $D = \{a{:}e/e, e{:}f/f, a{:}\neg f/\neg f, a{:}b/b, f{:}c/c, b{:}c/c, c{:}d/d\}$ and $\Delta = \langle D, \emptyset\rangle$. The set of defaults D is depicted in Figure 7.3. Theorem 7.1.1 gives the following priorities for Δ.

$$\mathcal{P} = \left\{\left\langle \frac{a:\neg f}{\neg f}, \frac{e:f}{f}\right\rangle, \left\langle \frac{a:\neg f}{\neg f}, \frac{f:c}{c}\right\rangle\right\}$$

Now consider $\Delta = \langle D, \{a, \neg b\}\rangle$. There are two extensions of Δ, $E_1 = Cn(\{a, \neg b, e, f, c, d\})$ and $E_2 = Cn(\{a, \neg b, e, \neg f\})$. There is a strict total order $\mathcal{T}$ such that $\mathcal{P} \subseteq \mathcal{T}$ and $c{:}d/d\mathcal{T}a{:}\neg f/\neg f$. Now $\mathcal{T}$ is a $\Delta, \mathcal{P}$-characteristic ordering for E_1 because $a{:}\neg f/\neg f$ is the only default that is applied in E_2 but not in E_1, and $c{:}d/d$ is applied in E_1 but not in E_2 and $c{:}d/d\mathcal{T}a{:}\neg f/\neg f$. Therefore E_1 is a $\mathcal{P}$-preferred extension of Δ. This conflicts with our intuition because the way the conflict between the defaults $a{:}\neg f/\neg f$ and $e{:}f/f$ is resolved should not be affected by the presence of b or $\neg b$. ■

7.2 Translations of inheritance networks to prioritized default logic

In this section we formally define what we mean by an inheritance network, and give two translations of inheritance networks to default logic. We prove that the first translation works in an equivalent way irrespective of which of the consequence relations $\models_{\mathcal{P}}$, $\models^{2}_{\mathcal{P}}$, $\models^{d}_{\mathcal{P}}$, and $\models^{2d}_{\mathcal{P}}$ is used. We prove that the second translation is equivalent to the first translation in the sense that it sanctions exactly the same positive literals as consequences. Finally, the translations are illustrated by means of benchmark inheritance networks.

7.2.1 Formal definition of inheritance networks

We define an inheritance network as a quintuple $G = \langle V, E, I, P, N\rangle$. Here V is a set of *nodes* that corresponds to classes and properties, and $E \subseteq V \times V \times \{0,1,2,3\}$ is a set of *links*. The *polarity* of a link $\langle v_1, v_2, p\rangle$ is p. If the polarity is 0 or 1, the link is *defeasible*. If the polarity is 2 or 3, the link is *strict*. If the polarity is 0 or 2, the link is *positive*. If the polarity is 1 or 3, the link is *negative*. The set I is a set of individuals, and $P \subseteq I \times V$ and $N \subseteq I \times V$ are relations that respectively represent the membership and non-membership of the individuals in the classes. We sometimes talk about links and nodes in G when we respectively refer to members of E and V.

A sequence $\langle v_1, w_1, p_1\rangle, \ldots, \langle v_n, w_n, p_n\rangle$ of links in E is *a path* from v_1 to w_n, if $w_i = v_{i+1}$ for all $i \in \{1, \ldots, n-1\}$. The path is *positive*, if p_i is 0 to 2 for all $i \in \{1, \ldots, n\}$. The *empty path* from any node to itself is also a path. An inheritance network is *acyclic* if there is a function f from V to the integers such that $f(v_1) < f(v_2)$ whenever there is a link $\langle v_1, v_2, p\rangle \in E$. This is equivalent to the inexistence of a non-empty path from a node to itself. The *rank* $r(p)$ of a node $p \in V$ is defined as the length of a longest path in G that ends in p. This notion was introduced in [Gelfond and Przymusinska, 1990].

7.2.2 The translation

For simplicity of presentation all results in this section assume that inheritance networks have exactly one individual. This simplification does not sacrifice generality because inheritance for each individual is performed separately from other individuals. The translation of an inheritance network to a default theory is as follows.

Definition 7.2.1 *Assume* $G = \langle V, E, I, P, N\rangle$ *is an inheritance network where* $I = \{x\}$*. The translation of* G *into default logic is* $\Delta_G = \langle D, W\rangle$*, where*

$$\begin{array}{rcl} D & = & \left\{ \frac{p:q}{q} \middle| \langle p,q,0\rangle \in E \right\} \cup \left\{ \frac{p:\neg q}{\neg q} \middle| \langle p,q,1\rangle \in E \right\} \\ & \cup & \left\{ \frac{p:}{q} \middle| \langle p,q,2\rangle \in E \right\} \cup \left\{ \frac{p:}{\neg q} \middle| \langle p,q,3\rangle \in E \right\}, \textit{ and} \\ W & = & \{p | \langle x,p\rangle \in P\} \cup \{\neg p | \langle x,p\rangle \in N\}. \end{array}$$

We adopt the terminology concerning paths and ranks of nodes in inheritance networks also to the setting of default theories. If Δ is the translation of an acyclic inheritance network G, then there is a path from p to q in Δ if there is a path from p to q in G. Also, there is a path from $\neg p$ or p to $\neg q$ or q in Δ, if there is a path from p to q in G. The rank $r(l)$ of a literal $l = p$ or $l = \neg p$ in a default theory Δ is the rank $r(p)$ of the node p in G. For a default δ with the conclusion l we define $r(\delta) = r(l)$.

Because our translation of inheritance networks to default logic uses non-normal defaults, it is not obvious that extensions always exist. Indeed, in a default theory there may be two defaults with contradicting conclusions and no justifications. In this case there may either be an inconsistent extension or no extensions at all. We do not consider inheritance networks that translate to this kind of default theories.

Definition 7.2.2 (Good) *Let* $G = \langle V, E, I, P, N\rangle$ *be an inheritance network. Then* G *is* good *if*

1. G *is acyclic,*
2. $V \cup I$ *is finite,*
3. $P \cap N = \emptyset$,
4. *there is no node* $p \in V$ *such that* $\langle q, p, 2\rangle \in E$ *and* $\langle q', p, 3\rangle \in E$ *for some* $q \in V$ *and* $q' \in V$,
5. *there is no node* $p \in V$ *such that* $\langle q, p, 2\rangle \in E$ *and* $\langle i, p\rangle \in N$ *for some* $i \in I$ *and* $q \in V$,
6. *there is no node* $p \in V$ *such that* $\langle q, p, 3\rangle \in E$ *and* $\langle i, p\rangle \in P$ *for some* $i \in I$ *and* $q \in V$,
7. $s = s'$ *whenever* $\langle p, q, s\rangle \in E$ *and* $\langle p, q, s'\rangle \in E$, *and*
8. *there is no defeasible link* $\langle p, q, n\rangle \in E$ *and strict link* $\langle p', q', n'\rangle \in E$ *such that* $q = p'$.

The first condition in the above definition is motivated by problems associated with cyclic inheritance networks as discussed in Section 7.4. Conditions 3, 4, 5 and 6 ensure that there are no unresolvable conflicts, that is, every conflict in an inheritance network can be resolved by rejecting a defeasible link. Condition 8 is not motivated by consistency or existence of extensions. Here we follow Gelfond and Przymusinska [1990] who give examples of problematic inheritance networks where a strict link follows a defeasible link. The problem is that if the defeasible link is indirectly (via the strict link) involved in a conflict, it is not in general possible to reject it because there are no defaults directly in conflict with it.

Lemma 7.2.3 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Then the set $W \cup \{\beta | p{:}/\beta \in D\}$ is consistent.*

Proof: For good $G = \langle V, E, I, P, N\rangle$, $P \cap N = \emptyset$, and hence W is consistent. Because there are no pairs of links $\langle p, q, 2\rangle \in E$ and $\langle p', q, 3\rangle \in E$, $C = \{\beta | p{:}/\beta \in D\}$ is consistent. And $W \cup C$ is consistent because $\langle x, p\rangle \notin N$ whenever $\langle p, q, 2\rangle \in E$ and $\langle x, p\rangle \notin E$ whenever $\langle p, q, 3\rangle \in P$ and the complement of no member of C is in W. $\Box$

Lemma 7.2.4 *Let G be an acyclic inheritance network and $\Delta_G = \langle D \cup D', W\rangle$ the translation of G as given in Definition 7.2.1, and the propositional variables in conclusions of D' do not occur in any prerequisite of a member of D, and defaults in D' are normal. Let E be an extension of $\langle D, W\rangle$. Then there is an extension E' of Δ_G such that $E \subseteq E'$ and def(δ, E') implies def(δ, E) for all $\delta \in D$.*

Proof: Let $\delta_1, \ldots, \delta_n$ be a total ordering of D' such that if $r(\delta_i) < r(\delta_j)$ then $i < j$. Such an ordering exists because G is acyclic. Let $p_i{:}\beta_i/\beta_i = \delta_i$ for all $i \in \{1, \ldots, n\}$. Define $D_i = \{\delta_1, \ldots, \delta_i\} \cup D$ for all $i \in \{0, \ldots, n\}$. We construct a sequence of sets $B_0, \ldots, B_n$ and prove by induction that $Cn(B_i)$ is an extension of $\langle D_i, W\rangle$. As the case $i = n$ we obtain the fact that $Cn(B_n)$ is an extension of $\langle D \cup D', W\rangle$. Define

$$
\begin{array}{rcl}
B_0 & = & W \cup \{\gamma | \gamma \text{ is the conclusion of a default in } GD(E, \Delta)\} \\
B_i & = & \left\{ \begin{array}{ll} B_{i-1} \cup \{\beta_i\}, & \text{if } B_{i-1} \models p_i \text{ and } B_{i-1} \not\models \neg\beta_i \\ B_{i-1}, & \text{otherwise.} \end{array} \right.
\end{array}
$$

Induction hypothesis: $Cn(B_i)$ is an extension of $\langle D_i, W\rangle$, and for all $\delta \in D$, def$(\delta, Cn(B_i))$ implies def$(\delta, Cn(B_0))$.

Base case $i = 0$: By assumption E is an extension of $\langle D, W\rangle$, and by Theorem 2.1.5 $Cn(B_0) = E$, and by definition $D_0 = D$.

Inductive case $i \geq 1$: Assume $B_{i-1} \models \neg\beta_i$. In this case $B_i = B_{i-1}$. By Lemma A.1.9 $Cn(B_i)$ is an extension of $\langle D_i, W\rangle = \langle D_{i-1} \cup \{\delta_i\}, W\rangle$. Assume $B_{i-1} \not\models p_i$. In this case $B_i = B_{i-1}$. By Lemma A.1.11 $Cn(B_i)$ is an extension of $\langle D_i, W\rangle = \langle D_{i-1} \cup \{\delta_i\}, W\rangle$. Assume $B_{i-1} \models p_i$ and $B_{i-1} \not\models \neg\beta_i$. In this case $B_i = B_{i-1} \cup \{\beta_i\}$. By induction hypothesis $Cn(B_{i-1})$ is an extension of $\langle D_{i-1}, W\rangle$. By Lemma A.1.14 $Cn(B_i)$ is an extension of $\langle D_i, W\rangle$ because by assumption the conclusion of δ_i is not the prerequisite of any default in D, and by definition of rank and D_i the conclusion of δ_i is not the prerequisite of any default in D_i.

Assume $\mathrm{def}(p{:}\beta/\beta, Cn(B_i))$ and $p{:}\beta/\beta \in D$. If $B_i = B_{i-1}$, then clearly $\mathrm{def}(p{:}\beta/\beta, Cn(B_{i-1}))$ and by induction hypothesis $\mathrm{def}(p{:}\beta/\beta, Cn(B_0))$. So assume $B_i \neq B_{i-1}$ and hence $B_i = B_{i-1} \cup \{\beta_i\}$. Hence $B_{i-1} \models p_i$ and $B_{i-1} \not\models \overline{\beta_i}$ and B_i is consistent. Because B_i is a consistent set of literals, β_i is the only literal in $Cn(B_i)\backslash Cn(B_{i-1})$. We show that $\overline{\beta} \neq \beta_i$ and $p \neq \beta_i$, and hence $\mathrm{def}(p{:}\beta/\beta, Cn(B_{i-1}))$. By assumption there are no defaults $\beta_i{:}\beta'/\beta' \in D_{i-1}$. Hence $p \neq \beta_i$. Because B_i is consistent and $B_i \models \beta_i \wedge \overline{\beta}$, $\beta \neq \beta_i$. Because by definition $B_{i-1} \not\models \overline{\beta_i}$ and hence $B_i = B_{i-1} \cup \{\beta_i\} \not\models \overline{\beta_i}$, it is not possible that $\beta = \overline{\beta_i}$. The remaining case is that the propositional variables in β and β_i are distinct. Clearly $\overline{\beta} \neq \beta_i$. This finishes the induction proof.

Hence $E' = Cn(B_n)$ is an extension of $\langle D \cup D', W\rangle$ and $\mathrm{def}(\delta, E')$ implies $\mathrm{def}(\delta, E)$ for all $\delta \in D$. □

Lemma 7.2.5 *Let G be a good inheritance network. Then the default theory Δ_G given in Definition 7.2.1 has at least one extension, and all extensions of Δ_G are consistent.*

Proof: Let $D_0 = \{p{:}/\beta \in D|\}$ where $\Delta_G = \langle D, W\rangle$. Let $E_0 = Cn_W(D_0)$. Because G is good, by Lemma 7.2.3 $E_0 \subseteq Cn(W \cup \{\beta|p{:}/\beta \in D\})$ is consistent. By Lemma A.1.7 E_0 is an extension of $\langle D', W\rangle$ where $D' = \{p{:}/\beta \in D_0|p \in E_0\}$. By repeated application of Lemma A.1.11 E_0 is an extension of $\langle D_0, W\rangle$. By Lemma 7.2.4 there is an extension E of $\langle D, W\rangle$.

Assume there is an inconsistent extension E of $\langle D, W\rangle$. Hence none of the normal defaults belongs to D^E. Therefore $D^E = D_0$. Above we showed that $Cn_W(D_0) = E$ is consistent, which contradicts our assumption. Hence no extension of $\langle D, W\rangle$ is inconsistent. □

In addition to the translation given in Definition 7.2.1, in Section 7.2.3 we consider a closely related translation where defaults are prerequisite-free normal. If our goal were to give as weak priorities as possible, then priorities for these translations would be different. We however present only one definition that works for both translations. We do not discuss the differences in priorities in more detail because the differences are small and do not seem to be of great importance. The motivation for our definition of priorities comes from the discussion in Section 7.1.3.

Definition 7.2.6 (Priorities) *Let G be an inheritance network. Let $\Delta_G = \langle D, W\rangle$ be the translation of G as given in Definition 7.2.1. Let $\mathcal{P}_G$ be the smallest transitive relation such that the following holds. For defaults $p{:}\pi/\pi, r{:}\overline{\pi}/\overline{\pi}$ and $s{:}\sigma/\phi$ in D, $p{:}\pi/\pi\mathcal{P}_G r{:}\overline{\pi}/\overline{\pi}$ whenever there is a non-empty positive path from p to r, and $p{:}\pi/\pi\mathcal{P}_G s{:}\sigma/\phi$ whenever there is a path from π to s. Here p, r and s are propositional variables, π and ϕ are literals, and σ is the empty sequence or a sequence that consists of ϕ.*

Lemma 7.2.7 *Let G be an acyclic inheritance network. Then $\mathcal{P}_G$, as given in Definition 7.2.6, is a strict partial order.*

Proof: By definition $\mathcal{P}_G$ is transitive. We show that it is asymmetric. If two defaults δ and δ' are related by $\mathcal{P}_G$, it is either because they fulfill one of the two conditions stated in the definition of $\mathcal{P}_G$ or because of transitivity. We show that the following property holds for all pairs of defaults δ and δ'. This property, call it Property A, implies asymmetry.

Whenever $\alpha{:}\sigma/\beta\mathcal{P}_G\alpha'{:}\sigma'/\beta'$, either $r(\alpha) < r(\alpha')$ and $r(\beta) = r(\beta')$ or $r(\beta) < r(\beta')$.

The induction is by a notion of distance between defaults. If $\delta\mathcal{P}_G$ and there is no default δ'' such that $\delta\mathcal{P}_G\delta''\mathcal{P}_G\delta'$, then the distance between δ and δ' is 1. If there is a chain of defaults $\delta, \delta_1, \ldots, \delta_{n-1}, \delta'$ with each two consecutive defaults related by $\mathcal{P}_G$ and there is no longer such chain between δ and δ', then the distance between δ and δ' is n. The induction proof uses the distance between defaults as follows. If the defaults δ and δ' are related by transitivity and the distance between δ and δ' is n, then there is default δ'' such that $\delta\mathcal{P}_G\delta''$ and $\delta''\mathcal{P}_G\delta'$ and the distance between δ and δ'' is less than n and the distance between δ'' and δ' is less than n.

Induction hypothesis: Property A holds for defaults δ and δ' with distance i.

Base case $i = 1$: Assume $\alpha{:}\sigma/\beta\mathcal{P}_G\alpha'{:}\sigma'/\beta'$ because $\beta = \overline{\beta'}$ and there is a non-empty positive path from α to α'. Clearly $r(\beta) = r(\beta')$. Let m be the length of the longest path in G that ends in α. Now $r(\alpha) = m$. The longest path that ends in α' has to be longer than m (by at least the length of the path from α to α'), and therefore $r(\alpha') > r(\alpha)$.

Assume $\alpha{:}\sigma/\beta\mathcal{P}_G\alpha'{:}\sigma'/\beta'$ because there is a path from β to α'. Let $m = r(\beta)$. As above, $r(\beta)$ is at least m plus the length of the path from β to β' which is at least one, and therefore $r(\beta) < r(\beta')$.

Inductive case $i > 1$: Assume $\alpha{:}\sigma/\beta\mathcal{P}_G\alpha'{:}\sigma'/\beta'$ because of transitivity, that is, $\alpha{:}\sigma/\beta\mathcal{P}_G\delta''$ and $\delta''\mathcal{P}_G\alpha'{:}\sigma'/\beta'$ for some $\delta'' = \alpha''{:}\sigma''/\beta''$. By induction hypothesis, either **1a)** $r(\alpha) < r(\alpha'')$ and $r(\beta) = r(\beta'')$ or **1b)** $r(\beta) < r(\beta'')$, and either **2a)** $r(\alpha'') < r(\alpha')$ and $r(\beta'') = r(\beta')$ or **2b)** $r(\beta'') < r(\beta')$.

If 1a and 2a, then by transitivity of $<$, $r(\alpha) < r(\alpha')$, and by transitivity of $=$, $r(\beta) = r(\beta')$. If 1a and 2b, then by $r(\beta) = r(\beta'')$ and $r(\beta'') < r(\beta')$, we get $r(\beta) < r(\beta')$. If 1b and 2a, then by $r(\beta) < r(\beta'')$ and $r(\beta'') = r(\beta')$, we get $r(\beta) < r(\beta')$. If 1b and 2b, then by transitivity of $<$ we get $r(\beta) < r(\beta')$. This covers all the cases, and establishes the induction hypothesis for the inductive case, thereby completing the induction proof.

Now assume that for some $\delta = \alpha{:}\beta/\beta$ and $\delta' = \alpha'{:}\beta'/\beta'$, both $\delta\mathcal{P}_G\delta'$ and $\delta'\mathcal{P}_G\delta$. This contradicts Property A, because now either $r(\beta) < r(\beta')$ and $r(\beta') < r(\beta)$ or $r(\alpha) < r(\alpha')$ and $r(\alpha') < r(\alpha)$, and $<$ is asymmetric. Therefore $\mathcal{P}_G$ is asymmetric. $\square$

Next we show that preferredness coincides with preferredness$_d$ for default theories and priorities that are translations of good inheritance networks.

Lemma 7.2.8 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let $D' \subseteq D$ be a set of defaults such that $p'{:}\sigma'/\beta \in D'$ whenever $p{:}\sigma/\beta \in D'$ and $p'{:}\sigma'/\beta \in D$, and $p'{:}\sigma'/p \in D'$ whenever $p{:}\sigma/\beta \in D'$ and $p'{:}\sigma'/p \in D$. Let E and E' be extensions of Δ_G such that for all $\delta \in D'$, def(δ, E) if and only if def(δ, E'). Then for all $p{:}\sigma/\beta \in D'$, $\beta \in E$ if and only if $\beta \in E'$.*

Proof: Because G is good, by Lemma 7.2.5 the extensions E and E' are consistent. Because the members of W and the conclusions of defaults in D are literals, a literal l belongs to E (or E') only if $l \in W$ or there is a generating default of E (respectively E') with l as the conclusion.

Induction hypothesis: For all $\beta \in \mathcal{L}$ such that $r(\beta) = i$ and $p{:}\sigma/\beta \in D'$ for some p and σ, $\beta \in E$ if and only if $\beta \in E'$.

Base case $i = 0$: Assume $\beta \in E$. Because $r(\beta) = 0$, $\beta \in W$. Therefore $\beta \in E'$ because $W \subseteq E'$.

Inductive case $i \geq 1$: Let $\beta \in E$ be a conclusion of a default in D' and $r(\beta) = i$. Now either $\beta \in W$ or there is a generating default of E with β as the conclusion. If $\beta \in W$, then $\beta \in E'$ because $W \subseteq E'$. Otherwise there is $\delta = p{:}\sigma/\beta \in D$ such that appl(δ, E). By definition of D', $\delta \in D'$. Because appl(δ, E), $p \in E$. By definition of rank, $r(p) < r(\beta)$. As $p \in E$, either $p \in W$ or p is the conclusion of a generating default of E that is in D' by definition of D'. In the former case $p \in E$ because $W \subseteq E'$. In the latter case $p \in E'$ by induction hypothesis. Because not def(δ, E) and $p \in E'$ and by assumption def(δ, E) iff def(δ, E'), now appl(δ, E'). Hence $\beta \in E'$. $\square$

Lemma 7.2.9 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let $D' \subseteq D$ be a set of defaults such that whenever $p{:}\sigma/\beta \in D'$ and $p'{:}\sigma'/\beta' \in D$, then $p'{:}\sigma'/\beta' \in D'$, and whenever $p{:}\sigma/\beta \in D'$ and $p'{:}\sigma'/p \in D$, then $p'{:}\sigma'/p \in D'$. Let E and E' be extensions of G such that for all defaults $\delta \in D'$, appl(δ, E) if and only if appl(δ, E'). Then for all $p{:}\sigma/\beta \in D'$, $\beta \in E$ if and only if $\beta \in E'$.*

Proof: The proof is almost identical to the one of Lemma 7.2.8. The key to both results is that non-defeat and application coincide for defaults the prerequisite of which is in the extension. □

Lemma 7.2.10 *Let G be a good inheritance network. Let $\Delta_G = \langle D, W\rangle$ and $\mathcal{P}_G$ as defined in Definitions 7.2.1 and 7.2.6. Let E be an extension of Δ_G. Then E is a $\mathcal{P}_G$-preferred$_d$ extension if and only if E is a $\mathcal{P}_G$-preferred extension.*

Proof: For all $\delta \in D$ define $B_\delta = \{\delta' \in D | \delta' \mathcal{P}_G \delta, r(\delta') < r(\delta)\}$. For defaults $\delta = \alpha{:}\sigma/\gamma \in D$, if there is a default $\alpha'{:}\sigma/\alpha \in D$, then $\alpha'{:}\sigma/\alpha \in B_\delta$ directly by definition of $r(\delta)$, $\mathcal{P}_G$ and B_δ. The sets B_δ satisfy the assumptions concerning D' in Lemmata 7.2.8 and 7.2.9. Let E' be an extension of Δ_G. Hence if def(δ', E) iff def(δ', E') for all $\delta' \in B_\delta$ (or appl(δ', E) iff appl(δ', E') for all $\delta' \in B_\delta$), then $\alpha \in E$ if and only if $\alpha \in E'$ where $\delta = \alpha{:}\sigma/\beta$. If there is no default $\alpha'{:}\sigma'/\alpha \in D$, then $\alpha \in E$ iff $\alpha \in E'$ because of the following. If $\alpha \in E$, then $\alpha \in W$. Because $W \subseteq E'$, $\alpha \in E'$.

Because E is a $\mathcal{P}_G$-preferred$_d$ extension of Δ_G, there is a $\Delta_G, \mathcal{P}_G$-characteristic$_d$ ordering for E. Let E' be any extension of Δ_G such that $E \neq E'$. Let δ be the $\mathcal{T}$-least default such that def(δ, E', E). Because $\mathcal{T}$ is a $\Delta_G, \mathcal{P}_G$-characteristic$_d$ ordering for E, it must be the case that def(δ', E) iff def(δ', E') for all $\delta' \mathcal{T} \delta$. Because the prerequisite of δ is in E' and def(δ', E) iff def(δ', E') for all $\delta' \in B_\delta$, the prerequisite of δ is in E as well by Lemma 7.2.8. Hence appl(δ, E, E').

Assume appl(δ', E, E') or appl(δ', E', E) for some $\delta' \mathcal{T} \delta$. By earlier assumptions def(δ'', E) iff def(δ'', E') for all $\delta'' \in B_{\delta'}$. Now because the prerequisite of δ' is in one of E and E', by Lemma 7.2.8 it must be in both E and E'. But as def(δ', E) iff def(δ', E') by assumption, we get a contradiction. Hence appl(δ', E) iff appl(δ', E') for all $\delta' \mathcal{T} \delta$. Now $\mathcal{T}$ is a $\Delta_G, \mathcal{P}_G$-characteristic ordering for E because appl(δ, E, E') and appl(δ', E', E) for no $\delta' \mathcal{T} \delta$. Therefore E is a $\mathcal{P}_G$-preferred extension of Δ_G. The proof that $\mathcal{P}_G$-preferredness implies $\mathcal{P}_G$-preferredness$_d$ is similar. □

In Section 7.3.1 we show that the translation given in this section coincides with a translational theory of inheritance presented by Gelfond and Przymusinska [1990]. Examples on how these translations work with respect to benchmark inheritance networks can be found in Section 7.2.4.

As inheritance networks are a very restricted subclass of all default theories, the question of extending the definitions of this section to a larger class of default theories. In Section 7.4 we show that removal of the acyclicity condition, and more surprisingly, even the introduction of disjunctive information to acyclic inheritance networks causes problems that suggest the unsuitability of prioritized default logic for more general forms of default reasoning.

7.2.3 Inheritance reasoning with prerequisite-free defaults

In this section we show how inheritance networks can be represented as prerequisite-free normal default theories, and hence – by results in Section 6.2 – in the preferred subtheories framework of Brewka [1989]. The translation we present works otherwise like the translation given in Section 7.2.2 except that the links in inheritance networks have the contraposition property. This is a consequence of representing them as defaults $:p \rightarrow \beta/p \rightarrow \beta$. The translation is a direct derivative of the one given in Section 7.2.2, and the results on the correspondence between the translations make only relatively weak assumptions about the priorities used. We proceed by establishing a connection between prioritized default theories with defaults of the form $p{:}q/q$ and $p{:}\neg q/\neg q$, and prioritized default theories with defaults of the form $:p \rightarrow q/p \rightarrow q$ and $:p \rightarrow \neg q/p \rightarrow \neg q$.

The translations of normal defaults with prerequisites to prerequisite-free normal defaults use the function $I : 2^{\mathcal{D}} \rightarrow 2^{\mathcal{D}}$ that is defined as $I(D) = \{:p \rightarrow \beta/p \rightarrow \beta | p{:}\beta/\beta \in D\}$. The inverse of I is $I^{-1}(D) = \{p{:}\beta/\beta | :p \rightarrow \beta/p \rightarrow \beta \in D\}$. For translation of strict links define $I_s(D) = \{p \rightarrow \beta | p{:}/\beta \in D\}$ and its inverse $I_s^{-1}(W) = \{p{:}/\beta | p \rightarrow \beta \in W\}$. The following lemmata show that for each extension of $\langle D, W\rangle$, there is an extension of $\langle I(D), W \cup I_s(D)\rangle$ that defeats the corresponding defaults, and vice versa. Like in the previous section, we assume that each inheritance network has exactly one individual x.

Lemma 7.2.11 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let $\Delta' = \langle I(D), W \cup I_s(D)\rangle$. Let E be an extension of Δ_G and $F = \{\delta \in D | not\ def(\delta, E)\}$. Then there is an extension E' of Δ' such that $I(F) \subseteq \{\delta \in I(D) | not\ def(\delta, E')\}$ and for all propositional variables p, $p \in E$ if and only if $p \in E'$.*

Proof: Let $S = \{\phi | :\phi/\phi \in I(F)\} \cup W \cup I_s(D)$.

Lemma A For all literals π, $\pi \in E$ implies $S \models \pi$.
Proof: By Theorem 2.1.3 and Lemma A.1.5 $E = Cn_W(F)$. Let $C_0, C_1, \ldots$ be the sets for $Cn_W(F)$ in the definition of $Cn_W(F)$ in Theorem 2.1.3. *Induction*

hypothesis: For all literals π, $C_i \models \pi$ implies $S \models \pi$. *Base case* $i = 0$: Because $C_0 = W$ and $W \subseteq S$, $C_0 \models \pi$ implies $S \models \pi$ for all literals π. *Inductive case* $i \geq 1$: Assume $C_i \models \pi$. If $\pi \in Cn(C_{i-1}) \subseteq Cn(C_i)$, then by monotonicity of $\models$ and by induction hypothesis $S \models \pi$. If $\pi \in Cn(C_i) \backslash Cn(C_{i-1})$, then because E is consistent by Lemma 7.2.5, there is $q{:}\sigma/\pi \in F$ such that $C_{i-1} \models q$. By induction hypothesis $S \models q$. Because $q \rightarrow \pi \in \{\phi | {:}\phi/\phi \in I(F)\} \cup I_s(D) \subseteq S$, finally $S \models \pi$. □

Let M be a propositional model such that for all propositional variables p, $M \models p$ iff $p \in E$.

Lemma B $M \models S$.
Proof: Because G is a good inheritance network, E is consistent by Lemma 7.2.5. If $\neg p \in E$, then $p \notin E$ because E is consistent, and hence $M \models \neg p$. Therefore $M \models \pi$ for all literals $\pi \in E$. Specifically $M \models W$ because $W \subseteq E$ and W is a set of literals. Assume $p \rightarrow \beta \in \{\phi | {:}\phi/\phi \in I(F)\} \cup I_s(D) = S \backslash W$ and $M \models p$. Now either by definition of F not def($p{:}\beta/\beta, E$), or because defaults without justifications cannot be defeated, not def($p{:}/\beta, E$). Because $M \models p$, $p \in E$. Because not def($p{:}\sigma/\beta, E$), $\beta \in E$. Hence $M \models \beta$ and $M \models p \rightarrow \beta$. Therefore $M \models S$. □

Lemma C For all propositional variables p, $S \models p$ if and only if $M \models p$.
Proof: Assume $S \models p$. Because $M \models S$, $M \models p$. Assume $S \not\models p$. By Lemma **A** $p \notin E$. By definition of M, $M \not\models p$. □

Lemma D $S \cup \{p \rightarrow \beta\}$ is inconsistent for all ${:}p \rightarrow \beta / p \rightarrow \beta \in I(D) \backslash I(F)$.
Proof: Assume the opposite. Hence $S \not\models p$ or $S \not\models \neg\beta$ for some ${:}p \rightarrow \beta / p \rightarrow \beta \in I(D) \backslash I(F)$. By Lemma **A** $p \notin E$ or $\neg\beta \notin E$, that is, not def($p{:}\beta/\beta, E$), which contradicts the definition of F. □

By Lemmata **B** and **D** S is a maximal consistent subset of $\{\phi | {:}\phi/\phi \in I(D)\} \cup W \cup I_s(D)$, and by Lemma 6.2.8 $E' = Cn(S)$ is an extension of $\langle I(D), W \cup I_s(D)\rangle$. Clearly, no member of $I(F)$ is defeated in $Cn(S)$. By Lemma **C** $p \in Cn(S)$ if and only if $M \models p$, and $M \models p$ if and only if $p \in E$ by definition of M. □

The proof in the opposite direction requires priorities.

Lemma 7.2.12 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let $\Delta' = \langle I(D), W \cup I_s(D)\rangle$. Let $\mathcal{P}$ be a strict partial order on $I(D)$ such that ${:}p \rightarrow \beta/p \rightarrow \beta \mathcal{P} {:}p' \rightarrow \beta'/p' \rightarrow \beta'$ whenever $\beta = p'$. Let E' be a $\mathcal{P}$-preferred extension of Δ' and $F' = \{\delta \in I(D) | not\ def(\delta, E')\}$. Then there is an extension $E \subseteq E'$ of Δ_G such that $I^{-1}(F') \cup \{p{:}/\beta \in D|\} \subseteq \{\delta \in D | not\ def(\delta, E)\}$ and for all propositional variables p, $p \in E$ if and only if $p \in E'$.*

Proof: Define

$$Tr(D,W,E) = Cn(W \cup \left\{ p \wedge \beta \in E \,\middle|\, \frac{p:\sigma}{\beta} \in D \right\})$$

for sets of defaults D and sets of formulae $W \subseteq \mathcal{L}$ and $E \subseteq \mathcal{L}$. Let $E = Tr(D, W, E')$. For $p{:}/\beta \in D$ not def$(p{:}/\beta, E)$ is immediate because defaults without justifications are defeated in no extension. Assume $p{:}\beta/\beta \in I^{-1}(F')$. Hence $p \wedge \neg\beta \notin E'$. Because $E \subseteq E'$, $p \wedge \neg\beta \notin E$. Hence not def$(p{:}\beta/\beta, E)$. Hence no member of $I^{-1}(F') \cup \{p{:}/\beta \in D|\}$ is defeated in E.

Next we show that E is an extension of Δ_G. Because E' is a $\mathcal{P}$-preferred extension of Δ' and defaults in Δ' are prerequisite-free normal, by Lemma 6.2.12 $E' = Cn(S)$ for some preferred subtheory S of $\langle\{\phi|{:}\phi/\phi \in I(D)\}, W \cup I_s(D), \mathcal{P}'\rangle$ where $\mathcal{P}' = \{\langle\phi, \phi'\rangle|{:}\phi/\phi\mathcal{P}{:}\phi'/\phi'\}$. Let $n = |I(D)|$. Let $\mathcal{T}$ be the strict total order and $S_0, \ldots, S_n$ the sets in Definition 6.2.2 with the properties $\mathcal{P}' \subseteq \mathcal{T}$ and $S = S_n$. Let $p_1 \to \beta_1, \ldots, p_n \to \beta_n$ be the ordering $\mathcal{T}$. Define $\delta_i = p_i{:}\beta_i/\beta_i$ for all $i \in \{1, \ldots, n\}$. Define $D_i = \{\delta_1, \ldots, \delta_i\} \cup \{p{:}/\beta \in D|\}$ for all $i \in \{0, \ldots, n\}$.

Induction hypothesis: $E_i = Tr(D_i, W, Cn(S_i))$ is an extension of $\langle D_i, W\rangle$.

Base case $i = 0$: Let $p'_1{:}/\beta'_1, \ldots, p'_m{:}/\beta'_m$ be a strict total order on D_0 so that $j < k$ whenever $\beta'_j = p'_k$. There is such an ordering because G is good and therefore acyclic. Define $D'_j = \{p'_1{:}/\beta'_1, \ldots, p'_j{:}/\beta'_j\}$ for all $j \in \{0, \ldots, m\}$. Define $S'_j = W \cup \{p \to \beta | p{:}/\beta \in D'_j\}$. Clearly, S'_j is the unique extension of $\langle\emptyset, W \cup I_s(D'_j)\rangle$ for all $j \in \{0, \ldots, m\}$. *Induction hypothesis*: $E'_j = Tr(D'_j, W, Cn(W \cup I_s(D'_j)))$ is an extension of $\langle D'_j, W\rangle$. *Base case* $j = 0$: The result is immediate, because $Tr(D'_0, W, Cn(W \cup I_s(D'_0))) = Cn(W)$ and $Cn(W)$ is an extension of $\langle\emptyset, W\rangle$. *Inductive case* $j \geq 1$: Assume $p \wedge \beta \in Cn(S'_j)\backslash Cn(S'_{j-1})$ and $p \to \beta \in W \cup I_s(D'_j)$. By Lemma A.2.3 the set of propositional variables in $Cn(S'_j)\backslash Cn(S'_{j-1})$ is a subset of $\{\beta'_j\}$. Because by assumption β'_j is not the antecedent of any implication in S'_j, $p \neq \beta'_j$ and $p \in Cn(S'_{j-1})$. Hence $\beta \in Cn(S'_j)\backslash Cn(S'_{j-1})$. Hence $p \to \beta \notin S'_{j-1}$, and therefore $p \to \beta = p'_j \to \beta'_j$ and $E'_j \in \{E'_{j-1}, Cn(E'_{j-1} \cup \{p'_j \wedge \beta'_j\})\}$. Assume $S'_{j-1} \models p'_j$. Hence $S'_j \models p'_j \wedge \beta'_j$, and $p'_j \wedge \beta'_j \in E'_j$ and $E'_j = Cn(E'_{j-1} \cup \{\beta'_j\})$. Because $S'_{j-1} \models p'_j$, by Lemma A.2.2 either $p'_j \in W$ or $S'_{j-1} \models p$ for some $p \to p'_j \in S'_{j-1}$. Hence $p'_j \in E'_{j-1}$. Therefore by Lemma A.1.14 E'_j is an extension of $\langle D'_j, W\rangle$. Assume $S'_{j-1} \not\models p'_j$. By Lemma A.2.3 $p'_j \notin Cn(S'_j)$. Then $p'_j \wedge \beta'_j \notin Cn(S'_j)$ and $p'_j \wedge \beta'_j \notin E'_i$. Therefore $E'_j = E'_{j-1}$. By induction hypothesis and Lemma A.1.11 E'_j is an extension of $\langle D'_j, W\rangle$.

Inductive case $i \geq 1$: Assume $S_{i-1} \cup \{p_i \to \beta_i\}$ is inconsistent. Then $S_i = S_{i-1}$. Now $p \wedge \neg\beta_i \in Cn(S_{i-1})$ and because S_{i-1} is consistent, $p_i \wedge \beta_i \notin Cn(S_{i-1})$, and hence $E_i = E_{i-1}$. Because $S_{i-1} \models \overline{\beta_i}$ and by assumption there are no formulae $\beta_i \to \beta'$ in S_{i-1}, by Lemma A.2.2 either $\overline{\beta_i} \in W \subseteq S_{i-1}$ or

$\beta' \rightarrow \overline{\beta_i} \in \{\phi | :\phi/\phi \in I(D)\} \cup I_s(D) \subseteq S_{i-1}$ and $S_{i-1} \models \beta'$. In both cases $\overline{\beta_i} \in E_{i-1}$. Hence by Lemma A.1.9 E_i is an extension of $\langle D_i, W\rangle$.

Assume $S_{i-1} \cup \{p_i \rightarrow \beta_i\}$ is consistent. Then $S_i = S_{i-1} \cup \{p_i \rightarrow \beta_i\}$. Assume $p \wedge \beta \in Cn(S_i) \backslash Cn(S_{i-1})$ and $p \rightarrow \beta \in \{\phi | :\phi/\phi \in I(D_{i-1})\}$. Because by Lemma A.2.3 the set of propositional variables in $Cn(S_i) \backslash Cn(S_{i-1})$ is a subset of $\{\beta_i\}$, and by assumption β_i is not the antecedent of any implication in S_i, $p \neq \beta_i$ and $p \in Cn(S_{i-1})$. Hence $\beta \in Cn(S_i) \backslash Cn(S_{i-1})$ and $p \rightarrow \beta \notin S_{i-1}$. Because S_{i-1} is a maximal consistent subset of $W \cup I_s(D) \cup \{\delta_1, \ldots, \delta_{i-1}\}$, $S_{i-1} \models p \wedge \neg\beta$ or $p \rightarrow \beta = p_i \rightarrow \beta_i$. Because $S_i \models p \wedge \beta$ and S_i is consistent, it must be the case that $p \rightarrow \beta = p_i \rightarrow \beta_i$. Therefore $E_i \in \{E_{i-1}, Cn(E_{i-1} \cup \{p_i \wedge \beta_i\})\}$. Assume $S_{i-1} \models p_i$. Hence $S_i \models p_i \wedge \beta_i$, and $p_i \wedge \beta_i \in E_i$ and $E_i = Cn(E_{i-1} \cup \{\beta_i\})$. Because $S_{i-1} \models p_i$, by Lemma A.2.2 either $p_i \in W$ or $S_{i-1} \models p$ for some $p \rightarrow p_i \in S_{i-1}$. Hence $p_i \in E_{i-1}$. Therefore by Lemma A.1.14 E_i is an extension of $\langle D_i, W\rangle$. Assume $S_{i-1} \not\models p_i$. By Lemma A.2.3 $p_i \notin Cn(S_i)$. Then $p_i \wedge \beta_i \notin Cn(S_i)$ and $p_i \wedge \beta_i \notin E_i$. Therefore $E_i = E_{i-1}$. By induction hypothesis and Lemma A.1.11 E_i is an extension of $\langle D_i, W\rangle$. This finishes the induction proof, and the case $i = n$ shows that $E = Tr(D_n, W, Cn(S_n))$ is an extension of $\langle D_n, W\rangle = \langle D, W\rangle$.

Because $E \subseteq E'$, for all propositional variables p, $p \in E$ implies $p \in E'$. Assume $p \in E' = Cn(S_n)$. By Lemma A.2.2 $p \in W$ or $q \rightarrow p \in S_n$ and $S_n \models q$. Hence either $p \in W \subseteq E$ or by definition $q \wedge p \in Tr(D, W, Cn(S_n))$. Hence for all propositional variables p, $p \in E$ if and only if $p \in E'$. □

Lemma 7.2.13 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let $\mathcal{P}$ be a strict partial order on D such that $p{:}\beta/\beta \mathcal{P} p'{:}\beta'/\beta'$ whenever $\beta = p'$. Let $\Delta' = \langle I(D), W \cup I_s(D)\rangle$. Let $\mathcal{P}' = \{\langle :p \rightarrow \beta/p \rightarrow \beta, :p' \rightarrow \beta'/p' \rightarrow \beta'\rangle | p{:}\beta/\beta \mathcal{P} p'{:}\beta'/\beta'\}$. 1. If E' is a $\mathcal{P}'$-preferred extension of Δ', then there is a $\mathcal{P}$-preferred extension of Δ_G such that for all propositional variables p, $p \in E$ if and only if $p \in E'$. 2. If E is a $\mathcal{P}$-preferred extension of Δ_G, then there is a $\mathcal{P}'$-preferred extension E' of Δ' such that for all propositional variables p, $p \in E$ if and only if $p \in E'$.*

Proof: 1. Let E' be a $\mathcal{P}'$-preferred extension of Δ'. Let $F' = \{\delta \in I(D) | \text{not def}(\delta, E')\}$. By Lemma 7.2.12 there is an extension E of Δ_G such that $I^{-1}(F') \cup \{p{:}/\beta \in D|\} \subseteq \{\delta \in D | \text{not def}(\delta, E)\} = F$. Assume E is not $\mathcal{P}$-preferred. Let $\mathcal{T}'$ be a $\Delta, \mathcal{P}'$-characteristic ordering for E'. Let $\mathcal{T}$ be a strict total order on D such that $\mathcal{P} \subseteq \mathcal{T}$ and $\mathcal{T}' = \{\langle :p \rightarrow \beta/p \rightarrow \beta, :p' \rightarrow \beta'/p' \rightarrow \beta'\rangle | \langle p{:}\beta/\beta, p'{:}\beta'/\beta'\rangle \in \mathcal{T}\}$. By Lemma 3.1.18 there is a $\mathcal{T}$-preferred extension E_2 of Δ_G. Clearly $\mathcal{T}$ is a $\Delta_G, \mathcal{T}$-characteristic ordering for E_2. Let $F_2 = \{\delta \in D | \text{not def}(\delta, E_2)\}$. By Lemma 7.2.11 there is an extension E'' of Δ' such that $I(F_2) \subseteq \{\delta \in I(D) | \text{not def}(\delta, E'')\}$. Let $F'' = \{\delta \in I(D) | \text{not def}(\delta, E'')\}$. Assume $:p \rightarrow \beta/p \rightarrow \beta \in I(D)$ and $\text{def}(:p \rightarrow \beta/p \rightarrow \beta, E'', E')$. By Lemma

A below, def($p{:}\beta/\beta, E_2, E$). Because $\mathcal{T}$ is a $\Delta_G, \mathcal{P}$-characteristic ordering for E_2, there is $p'{:}\beta'/\beta'\mathcal{T}p{:}\beta/\beta$ such that def($p'{:}\beta'/\beta', E, E_2$). By Lemma **A**, def($:p' \to \beta'/p' \to \beta', E', E''$) and $:p' \to \beta'/p' \to \beta'\mathcal{T}'{:}p \to \beta/p \to \beta$. This contradicts the fact that $\mathcal{T}'$ is a $\Delta', \mathcal{P}'$-characteristic ordering for E'. Hence the assumption that E is not $\mathcal{P}$-preferred is false. By Lemma 7.2.12, for all propositional variables p, $E \models_{\mathcal{P}_G} p$ if and only if $E' \models_{\mathcal{P}'} p$.

Lemma A If $I(F_2) \subseteq F''$ and $I^{-1}(F') \cup \{p{:}/\beta \in D|\} \subseteq F$, then def($:p \to \beta/p \to \beta, E''$) implies def($p{:}\beta/\beta, E_2$), and then def($p{:}\beta/\beta, E$) implies def($:p \to \beta/p \to \beta, E'$).
Proof: Assume def($:p \to \beta/p \to \beta, E''$). Hence $:p \to \beta/p \to \beta \in I(D)\backslash F''$. Because $I(F_2) \subseteq F''$, $:p \to \beta/p \to \beta \in I(D)\backslash I(F_2)$. Now

$$\frac{:p \to \beta}{p \to \beta} \in I(D)\backslash I(F_2)$$

$$\begin{array}{ll}
\text{iff} & \dfrac{:p \to \beta}{p \to \beta} \in I(D) \text{ and } \dfrac{:p \to \beta}{p \to \beta} \notin \left\{ \dfrac{:p \to \beta}{p \to \beta} \in I(D) \,\middle|\, \text{not def}(\dfrac{p:\beta}{\beta}, E_2) \right\} \\
\text{iff} & \dfrac{:p \to \beta}{p \to \beta} \in I(D) \text{ and } \left(\dfrac{:p \to \beta}{p \to \beta} \notin I(D) \text{ or def}(\dfrac{p:\beta}{\beta}, E_2) \right) \\
\text{iff} & \dfrac{:p \to \beta}{p \to \beta} \in I(D) \text{ and def}(\dfrac{p:\beta}{\beta}, E_2).
\end{array}$$

Assume def($p{:}\beta/\beta, E$). Hence $p{:}\beta/\beta \in D\backslash F$. Because $I^{-1}(F') \cup \{p{:}/\beta \in D|\} \subseteq F$, $p{:}\beta/\beta \in D\backslash(I^{-1}(F') \cup \{p{:}/\beta \in D|\})$ and

$$\begin{array}{lll}
\dfrac{p:\beta}{\beta} \in D\backslash I^{-1}(F') & \text{iff} & \dfrac{p:\beta}{\beta} \in D \text{ and } \dfrac{:p \to \beta}{p \to \beta} \notin F' \\
 & \text{iff} & \dfrac{p:\beta}{\beta} \in D \text{ and } \dfrac{:p \to \beta}{p \to \beta} \notin \{\delta \in I(D) | \text{not def}(\delta, E')\} \\
 & \text{iff} & \dfrac{p:\beta}{\beta} \in D \text{ and } \left(\dfrac{:p \to \beta}{p \to \beta} \notin I(D) \text{ or def}(\delta, E') \right) \\
 & \text{iff} & \dfrac{p:\beta}{\beta} \in D \text{ and def}(\dfrac{:p \to \beta}{p \to \beta}, E').
\end{array}$$

□

2. Let E be a $\mathcal{P}$-preferred extension of Δ. Let $F = \{\delta \in D | \text{not def}(\delta, E)\}$. By Lemma 7.2.11 there is an extension E' of Δ' such that $I(F) \subseteq F' = \{\delta \in I(D) | \text{not def}(\delta, E')\}$. Assume E' is not $\mathcal{P}'$-preferred. Let $\mathcal{T}$ be a $\Delta_G, \mathcal{P}$-characteristic ordering for E. Let $\mathcal{T}' = \{\langle :p \to \beta/p \to \beta, :p' \to \beta'/p' \to \beta'\rangle | \langle p{:}\beta/\beta, p'{:}\beta'/\beta'\rangle \in \mathcal{T}\}$. By Lemma 3.1.18 there is a $\mathcal{T}'$-preferred extension E'' of Δ'. Clearly $\mathcal{T}'$ is a $\Delta', \mathcal{T}'$-characteristic ordering for E'' and by Lemma 3.1.17 E'' is $\mathcal{P}'$-preferred. Let $F'' = \{\delta \in I(D) | \text{not def}(\delta, E'')\}$. By Lemma 7.2.12 there is an extension E_2 of Δ_G such that $I^{-1}(F'') \cup \{p{:}/\beta \in D|\} \subseteq F_2 = \{\delta \in D | \text{not def}(\delta, E_2)\}$. Assume $p{:}\beta/\beta \in D$ and def($p{:}\beta/\beta, E_2, E$). By

Lemma **A**, def($:p \rightarrow \beta/p \rightarrow \beta, E'', E'$). Because $\mathcal{T}'$ is a $\Delta', \mathcal{P}'$-characteristic ordering for E'', there is $:p' \rightarrow \beta'/p' \rightarrow \beta' \mathcal{T}' :p \rightarrow \beta/p \rightarrow \beta$ such that def($:p' \rightarrow \beta'/p' \rightarrow \beta', E', E''$). By Lemma **A**, def($p':\beta'/\beta', E, E_2$) and $p':\beta'/\beta' \mathcal{T} p:\beta/\beta$. This contradicts the fact that $\mathcal{T}$ is a $\Delta_G, \mathcal{P}$-characteristic ordering for E. Hence the assumption that E' is not $\mathcal{P}'$-preferred is false. By Lemma 7.2.11, for all propositional variables p, $E \models_{\mathcal{P}_G} p$ if and only if $E' \models_{\mathcal{P}'} p$. □

Finally, we establish the main result of this section that shows that translations of inheritance networks into prerequisite-free normal default theories differ from the translations into normal default theories only with respect to the fact that prerequisite-free normal defaults $:p \rightarrow \beta/p \rightarrow \beta$ sanction inferences by contraposition from $\neg\beta$ to $\neg p$.

Theorem 7.2.14 *Let G be a good inheritance network and Δ_G the translation of G as given in Definition 7.2.1. Let $\mathcal{P}_G$ be the strict total order for Δ_G given in Definition 7.2.6. Let $\mathcal{P} = \{\langle :p \rightarrow \beta/p \rightarrow \beta, :p' \rightarrow \beta'/p' \rightarrow \beta' \rangle | p:\beta/\beta \mathcal{P}_G p':\beta'/\beta'\}$. Then for all propositional variables p, $\langle D, W\rangle \models_{\mathcal{P}_G} p$ if and only if $\langle I(D), W \cup I_s(D)\rangle \models_{\mathcal{P}} p$.*

Proof: The strict partial order $\mathcal{P}$ fulfills the property that $:p \rightarrow \beta/p \rightarrow \beta \mathcal{P} :p' \rightarrow \beta'/p' \rightarrow \beta'$ whenever $\beta = p'$. The existence of a $\mathcal{P}_G$-preferred extension E of $\langle D, W\rangle$ such that the propositional variable p is not in E is by Lemma 7.2.13 equivalent to the existence of a $\mathcal{P}$-preferred extension E' of $\langle I(D), W \cup I_s(D)\rangle$ such that the propositional variable p is not in E. Hence for all propositional variables p, p is in all $\mathcal{P}_G$-preferred extensions of $\langle D, W\rangle$ if and only if p is in all $\mathcal{P}$-preferred extensions of $\langle I(D), W \cup I_s(D)\rangle$. □

7.2.4 Examples

In this section we discuss a number of benchmark inheritance networks for assessing the behavior of our inheritance theories.

Example 7.2.1 The inheritance network in Figure 7.4 motivates Definition 7.2.2 that defines *good* inheritance networks. In good inheritance networks every conflict is between a defeasible link and a fact or a strict link. According to the translation given in Section 7.2.2, the set of defaults that corresponds to the inheritance network is $D = \{b:/a, c:/\neg a, d:b/b, d:c/c\}$. The default theory $\langle D, \{d\}\rangle$ has no extensions because the normal defaults in D cannot be defeated, and if they are applied, a contradiction is obtained. The contradictory extension $\mathcal{L}$ is not an extension of $\langle D, \{d\}\rangle$. This inheritance network is problem if inference by contraposition cannot be applied to strict links, and the notion of defeat for defeasible links is the one supported by default logic. ■

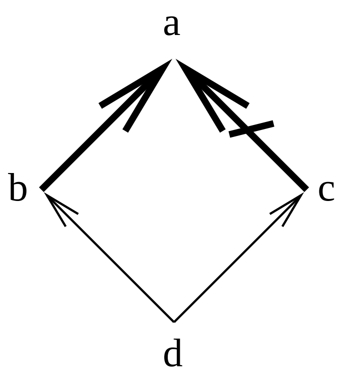


Figure 7.4: An inheritance network that is not good

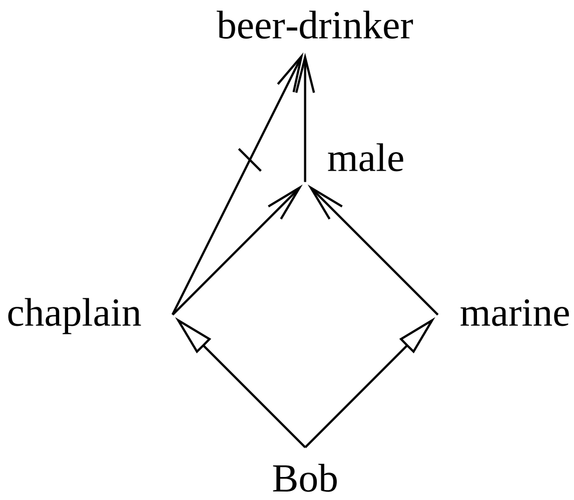


Figure 7.5: The marine chaplain

The inheritance network in Figure 7.5 illustrates the difference between on-path and off-path preemption [Horty, 1994]. With on-path preemption, the path Bob→marine→male→beer-drinker could be preempted only by a negative link that starts from a node that is on the path Bob→marine→male. Hence with on-path preemption the conflict between the links ending in the node beer-drinker is unresolved. Off-path preemption is possible also for negative links that are on some path between *Bob* and *male*, not necessarily on the path path Bob→marine→male. With off-path preemption the path Bob→marine→male→beer-drinker is preempted by the link from chaplain to beer-drinker, and hence the conclusion not beer-drinker is obtained. Both on-path and off-path preemption have their proponents [Touretzky *et al.*, 1987].

The inheritance theories in the framework of default logic and priorities do not define a preemption relation between paths, and therefore it is not possible to classify them directly as on-path or off-path. Because theories presented in Sections 7.2.2 and 7.2.3 assign the default $a{:}p/p$ a priority higher than

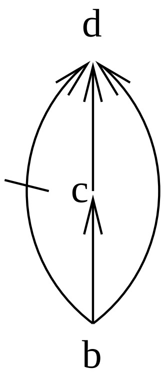


Figure 7.6: A redundant link

$b{:}\neg p/\neg p$ whenever a is lower than b in the inheritance network irrespective of whether b can be derived by using a sequence of links involving a, they somewhat resemble path-based theories with off-path preemption.

Example 7.2.2 The translation of the inheritance network in Figure 7.5 to default logic is as follows. Let $\Delta_G = \langle D, W \rangle$ where

$$\begin{aligned} W &= \{chaplain, marine\} \\ D &= \left\{ \frac{chaplain : man}{man}, \frac{marine : man}{man}, \right. \\ &\quad \left. \frac{chaplain : \neg beer\text{-}drinker}{\neg beer\text{-}drinker}, \frac{marine : \neg beer\text{-}drinker}{\neg beer\text{-}drinker} \right\}. \end{aligned}$$

The priorities for Δ_G are

$$\begin{aligned} \mathcal{P}_G &= \left\{ \left\langle \frac{chaplain : \neg beer\text{-}drinker}{\neg beer\text{-}drinker}, \frac{man : beer\text{-}drinker}{beer\text{-}drinker} \right\rangle, \right. \\ &\quad \left\langle \frac{chaplain : man}{man}, \frac{man : beer\text{-}drinker}{beer\text{-}drinker} \right\rangle, \\ &\quad \left. \left\langle \frac{marine : man}{man}, \frac{man : beer\text{-}drinker}{beer\text{-}drinker} \right\rangle \right\}. \end{aligned}$$

The extension $E = Cn(\{chaplain, marine, man, \neg beer\text{-}drinker\})$ is the only $\mathcal{P}_G$-preferred extension of Δ_G. ■

Example 7.2.3 A link in an inheritance network is *redundant* if it is obtained as a combination of several consecutive links the last of which is positive or negative and the others are positive. It is sometimes argued that adding or removing redundant links should not affect the conclusions sanctioned. The inheritance network in Figure 7.6 illustrates how adding a redundant link (or equivalently, splitting a link to two consecutive links) does not seem to affect the conclusions. The translation of the inheritance network is the set of defaults $D = \{b{:}\neg d/\neg d, b{:}c/c, c{:}d/d, b{:}d/d\}$ and the

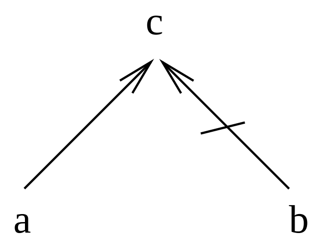


Figure 7.7: A simple inheritance network

strict partial order $\mathcal{P}_G = \{\langle b{:}\neg d/\neg d, c{:}d/d\rangle\}$. There are two extensions of $\langle D, \{b\}\rangle$, $E_1 = Cn(\{b, c, \neg d\})$ and $E_2 = Cn(\{b, c, d\})$. Both extensions are $\mathcal{P}_G$-preferred, and hence $\langle D, \{b\}\rangle \not\models_{\mathcal{P}_G} d$ nor $\langle D, \{b\}\rangle \not\models_{\mathcal{P}_G} \neg d$. The default theory $\langle D\backslash\{b{:}d/d\}, \{b\}\rangle$ for the inheritance network without the redundant link also has the extensions E_1 and E_2, and as the priorities are the same, still no conclusions concerning d can be obtained. ■

In Section 7.2.3 we gave an inheritance theory based on prerequisite-free normal defaults. Theorem 7.2.14 shows that the theory that uses prerequisite-free normal defaults and the one given in Section 7.2.2 that uses normal defaults with prerequisites have the same positive literals as their consequences. These theories differ with respect to negative literals.

Example 7.2.4 Figure 7.7 depicts an inheritance network with two links that correspond to the defaults $D = \{a{:}c/c, b{:}\neg c/\neg c\}$ (for the theory in Section 7.2.2) and $D' = \{{:}a \rightarrow c/a \rightarrow c, {:}b \rightarrow \neg c/b \rightarrow \neg c\}$ (for the theory in Section 7.2.3.) The priority relations associated with these inheritance networks are empty. The default theory $\langle D, \{b\}\rangle$ has the extension $E = Cn(\{b, \neg c\})$, and the default theory $\langle D', \{b\}\rangle$ has the extension $E' = Cn(\{b, b \rightarrow \neg c, a \rightarrow c\})$. The difference concerning negative literals is that $\neg a \notin E$ and $\neg a \in E'$. The theory in Section 7.2.3 sanctions contraposition while the theory in Section 7.2.2 does not. ■

7.3 Related work

In this section we show that the formalizations of inheritance networks by Gelfond and Przymusinska [1990] and Brewka [1987] can be translated to our formalization of inheritance networks presented in Section 7.2 (Sections 7.3.1 and 7.3.2), and how conditional entailment of Geffner and Pearl [1992] can be embedded in prioritized default logic (Section 7.3.3.) Some of the approaches proposed for inheritance reasoning in nonmonotonic logics seem to be incompatible with the kind of priorities investigated in this work. We briefly discuss approaches by Brewka (Section 7.3.4) and Lifschitz (Section

7.3.6), and the use of semi-normal defaults (Section 7.3.5.)

7.3.1 An inheritance theory by Gelfond and Przymusinska

Gelfond and Przymusinska [1990] present translations of inheritance networks to autoepistemic logic [Moore, 1985]. In Section 4b of their paper, they give two formalizations of inheritance reasoning and prove that they are equivalent. The first of these formalizations, defined for the class of good inheritance networks as defined in Definition 7.2.2, uses a form of priorities that induce an ordering on the stable expansions of autoepistemic theories. It turns out that the priorities and the ordering can be seen as a form of lexicographic ordering.

Definition 7.3.1 (Autoepistemic logic) *The language $\mathcal{L}_{ae}$ of autoepistemic logic includes the formulae in $\mathcal{L}$ and is closed under the unary operator L. The consequence relation $\models_{ae}$ for autoepistemic formulae extends the logical consequence relation of classical propositional logic and treats formulae $L\phi$ like propositional variables. The stable expansions of a set $\Sigma \subseteq \mathcal{L}_{ae}$ are sets E that fulfill the equation*

$$E = \{\phi \in \mathcal{L}_{ae} | \Sigma \cup \{L\phi \in \mathcal{L}_{ae} | \phi \in E\} \cup \{\neg L\phi \in \mathcal{L}_{ae} | \phi \notin E\} \models_{ae} \phi\}.$$

Definition 7.3.2 (Translation) *Let $G = \langle V, E, I, P, N\rangle$ be an acyclic inheritance network where V is finite and $I = \{x\}$. The translation $Th(G)$ of G to autoepistemic logic is the following set of formulae.*

$$\begin{array}{ll}
\{H(x,c) | \langle x,c\rangle \in P\} \cup & \{\neg H(x,c) | \langle x,c\rangle \in N\} \cup \\
\{PD(x,c) | \langle x,c,0\rangle \in E\} \cup & \{ND(x,c) | \langle x,c,1\rangle \in E\} \cup \\
\{PS(x,c) | \langle x,c,2\rangle \in E\} \cup & \{NS(x,c) | \langle x,c,3\rangle \in E\} \cup
\end{array}$$

$$\begin{array}{ll}
\{LPD(c,p) \wedge LH(x,c) \wedge \neg Lab(x,c,p) \rightarrow H(x,p) & | c \in V, p \in V, x \in I\} \cup \\
\{LND(c,p) \wedge LH(x,c) \wedge \neg Lab(x,c,p) \rightarrow \neg H(x,p) & | c \in V, p \in V, x \in I\} \cup \\
\{LPS(c,p) \wedge LH(x,c) \rightarrow H(x,p) & | c \in V, p \in V, x \in I\} \cup \\
\{LNS(c,p) \wedge LH(x,c) \rightarrow \neg H(x,p) & | c \in V, p \in V, x \in I\}
\end{array}$$

Definition 7.3.3 (Explanation) *Let $Th(G)$ be the translation of a finite and acylic inheritance network G. A set X is an* explanation, *if X consists of formulae of the form $ab(x, c_1, c_2)$ and $Th(G) \cup X$ has a consistent stable expansion.*

The purpose of an abnormality assumption $\mathrm{ab}(x, c_1, c_2)$ is to block inheritance of properties for x from c_2 to c_1. For an explanation X, the set of abnormality assumptions $\mathrm{ab}(x, c_1, c_2)$ with $c_2 = a$ is denoted by $X(a)$. The notion of explanation is motivated by the way inheritance networks G are translated to sets of formulae $\mathrm{Th}(G)$: conflicts between defeasible links are not resolved by formulae in Th(G), and in general Th(G) does not have stable expansions or the stable expansions are inconsistent.

An ordering on explanations of Th(G) is determined by the abnormality assumptions they contain. Gelfond and Przymusinska define the relation *better than* on sets $X(a)$ and the relation *preferable* on explanations[1].

Definition 7.3.4 (Better than) *Let X and X' be explanations of an inheritance network G and a a node in G. Then $X(a)$ is* better than *$X'(a)$, if $X(a) \subset X'(a)$, or $X'(a)$ is non-empty[2] and for every $ab(x,c,a) \in X(a)$ there is $ab(x,c',a) \in X'(a)$ and a non-empty positive path from c' to c.*

Lemma 7.3.5 *The relation* better than *is a strict partial order.*

Proof: We show asymmetry first. Let G be an inheritance network. Assume that for some explanations X and X' of G and some node a, $X(a)$ is better than $X'(a)$ and $X'(a)$ is better than $X(a)$.

Assume $X(a) \not\subset X'(a)$ and $X'(a) \not\subset X(a)$. Because $X'(a)$ is better than $X(a)$, $X(a)$ is non-empty and for each $\mathrm{ab}(x,c',a) \in X'(a)$ there is $\mathrm{ab}(x,c,a) \in X(a)$ such that there is a non-empty positive path from c to c'. Because $X(a)$ is better than $X'(a)$, $X'(a)$ is non-empty and for each $\mathrm{ab}(x,c,a) \in X(a)$ there is $\mathrm{ab}(x,c',a) \in X'(a)$ such that there is a non-empty positive path from c' to c. Take any $\mathrm{ab}(x,c_1,a) \in X(a)$. Starting from c_1 we now obtain an infinite chain of nodes $c_1, c'_1, c_2, c'_2, \ldots$ such that $\mathrm{ab}(x,c_i,a) \in X(a)$ for all $i \geq 1$ and $\mathrm{ab}(x,c'_i,a) \in X'(a)$ for all $i \geq 1$, and there is a non-empty positive path from c_i to c'_{i+1} and from c'_i to c_{i+1} for all $i \geq 1$. However, this contradicts the acyclicity and finiteness of G. Hence it is not possible that $X(a) \not\subset X'(a)$ and $X'(a) \not\subset X(a)$ and $X(a)$ is better than $X'(a)$ and $X'(a)$ is better than $X(a)$.

Assume $X(a) \subset X'(a)$. Let $\mathrm{ab}(x,c,a)$ be an element of $X'(a)$ such that there is no $\mathrm{ab}(x,c',a) \in X'(a)$ with a non-empty path from c to c'. Because of the finiteness and acyclicity of G, such an abnormality assumption exists. Because $X'(a)$ is better than $X(a)$, it should be the case that there is $\mathrm{ab}(x,c'',a) \in X(a)$ with a non-empty path from c'' to c. However, because of the above and $X(a) \subset X'(a)$, there are no such elements in $X(a)$. Hence it is not possible that $X(a) \subset X'(a)$ and $X'(a)$ is better than $X(a)$, and by symmetry it is not possible that $X'(a) \subset X(a)$ and $X(a)$ is better than $X'(a)$.

---

[1] In [Gelfond and Przymusinska, 1990, Proposition 1], the relation *preferable* is claimed to be a preorder. In the proof it is stated that by preorder an asymmetric and transitive relation is meant. However, the proposition and the proof are in error: neither the relation *better than* on which *preferable* is based nor *preferable*, as defined on Page 412 in their paper, are asymmetric. Our results that follow use a corrected definition of *better than.*

[2] This non-emptyness requirement is missing in the Gelfond and Przymusinska definition.

We have shown that the assumption that $X(a)$ is better than $X'(a)$ and vice versa is contradictory. Therefore the relation *better than* is asymmetric. Next we show that the relation *better than* is transitive, that is, if $X(a)$ is better than $X'(a)$ and $X'(a)$ is better than $X''(a)$, then $X(a)$ is better than $X''(a)$.

Clearly, if $X(a) \subset X'(a)$ and $X'(a) \subset X''(a)$, then $X(a) \subset X''(a)$ by the transitivity of $\subset$.

Assume $X(a) \subset X'(a)$ and $X'(a) \not\subset X''(a)$. Because $X'(a)$ is better than $X''(a)$ and $X'(a) \not\subset X''(a)$, $X''(a)$ is non-empty. Let $\mathrm{ab}(x, c, a)$ be any element of $X(a)$. Because $X(a) \subset X'(a)$, also $\mathrm{ab}(x, c, a) \in X'(a)$. Because $X'(a)$ is better than $X''(a)$ and $X'(a) \not\subset X''(a)$, there is $\mathrm{ab}(x, c', a) \in X''(a)$ such that there is a non-empty positive path from c' to c. Because such an element of $X''(a)$ exists for every element of $X(a)$, $X(a)$ is better than $X''(a)$.

Assume $X(a) \not\subset X'(a)$ and $X'(a) \subset X''(a)$. Because $X'(a) \subset X''(a)$, $X''(a)$ is non-empty. Let $\mathrm{ab}(x, c, a)$ be any element of $X(a)$. Because $X(a)$ is better than $X'(a)$ and $X(a) \not\subset X'(a)$, there is $\mathrm{ab}(x, c', a) \in X'(a)$ such that there is a non-empty path from c' to c. Because $X'(a) \subset X''(a)$, also $\mathrm{ab}(x, c, a) \in X''(a)$. Because such an element exists for every element of $X(a)$, $X(a)$ is better than $X''(a)$.

Assume $X(a) \not\subset X'(a)$ and $X'(a) \not\subset X''(a)$. Because $X'(a)$ is better than $X''(a)$ and $X'(a) \not\subset X''(a)$, $X''(a)$ is non-empty. Let $\mathrm{ab}(x, c, a)$ be any element of $X(a)$. Because $X(a)$ is better than $X'(a)$ and $X(a) \not\subset X'(a)$, there is $\mathrm{ab}(x, c', a) \in X'(a)$ such that there is a non-empty path from c' to c. Because $X'(a)$ is better than $X''(a)$ and $X'(a) \not\subset X''(a)$, there is $\mathrm{ab}(x, c'', a) \in X''(a)$ such that there is a non-empty path from c'' to c'. Therefore by the definition of path, there is a path from c'' to c. Hence for every $\mathrm{ab}(x, c, a) \in X(a)$ there is $\mathrm{ab}(x, c'', a) \in X''(a)$ such that there is a non-empty path from c'' to c. Therefore $X(a)$ is better than $X''(a)$. □

The preferability relation on explanations is defined in terms of the *better than* relation as follows.

Definition 7.3.6 (Preferable) *An explanation X is* preferable *to X', if for some $m \geq 0$,*

1. *$X(a) = X'(a)$ for all a such that $r(a) < m$,*
2. *for all a such that $r(a) = m$, $X(a) = X'(a)$ or $X(a)$ is better than $X'(a)$, and*
3. *for at least one a such that $r(a) = m$, $X(a)$ is better than $X'(a)$.*

An explanation X is a best *explanation, if no explanation is preferable to X.*

Lemma 7.3.7 ([Gelfond and Przymusinska, 1990]) *The preferability relation is a strict partial order.*

Proof: The proof assumes that *better than* is a strict partial order. □

We show that Gelfond's and Przymusinska's inheritance theory is equivalent to our translation of inheritance networks to prioritized default logic. First we show that without considering priorities and the preferability relation, there is a correspondence between the extensions of our translation of inheritance networks to default logic and the stable expansions of Gelfond and Przymusinska's translation. The correspondence is such that for every extension there is a stable expansion that contains the same literals indicating the membership and non-membership of individuals in classes, and vice versa. Then we show that Gelfond's and Przymusinska's preferability relation has a natural counterpart in default logic with lexicographic priorities and classification of defaults to defeated and non-defeated. Finally, we show that for a given inheritance network, the maximal extensions in the preferability relation are exactly the preferred extensions.

Our translation is simpler than Gelfond's and Przymusinska's. First, our translation avoids the reification of inheritance networks, that is, each link of an inheritance network can be represented as a default in Reiter's default logic, whereas Gelfond and Przymusinska have formulae of autoepistemic logic that describe the properties of links and an atomic formula for each link. Second, the multiple extensions for inheritance networks is simply due to multiple extensions of the corresponding default theories. Gelfond and Przymusinska resolve inheritance conflicts by introducing the notion of explanation that is not present in the definition of autoepistemic logic. Multiple extensions in inheritance networks correspond to the existence of multiple explanations. The explanations correspond to sets of defeated defaults in our translation. Third, by our results Gelfond's and Przymusinska's preferability relation on explanations can be seen as an ordering on extensions that is based on a lexicographic prioritization of defaults.

First we establish a correspondence with the class of default theories given in Definition 7.2.1 and the translations of inheritance networks to autoepistemic logic and explanations by Gelfond and Przymusinska.

Lemma 7.3.8 *Let Σ be a set of formulae $\{\chi_1, \ldots, \chi_n\}$ where $\chi_i = L\phi_{i,1} \wedge \cdots \wedge L\phi_{i,m_i} \wedge \neg L\psi_{i,1} \wedge \cdots \wedge \neg L\psi_{i,l_i} \rightarrow \pi_i$ and $\{\phi_{i,1}, \ldots, \phi_{i,m_i}, \psi_{i,1}, \ldots, \psi_{i,l_i}, \pi_i\} \subseteq \mathcal{L}$ and $m_i \geq 0$ and $l_i \geq 0$ for all $i \in \{1, \ldots, n\}$. Let I be a subset of $\{1, \ldots, n\}$ so that $B = \{\pi_i | i \in I\}$ is consistent. Then $Cn(B) = E \cap \mathcal{L}$ for a stable expansion E of Σ, if for all $i \in I$, $\{\phi_{i,1}, \ldots, \phi_{i,m_i}\} \subseteq Cn(B)$ and $\{\psi_{i,1}, \ldots, \psi_{i,l_i}\} \cap Cn(B) = \emptyset$, and for all $i \in \{1, \ldots, n\} \backslash I$, $\{\phi_{i,1}, \ldots, \phi_{i,m_i}\} \not\subseteq Cn(B)$ or $\{\psi_{i,1}, \ldots, \psi_{i,l_i}\} \cap Cn(B) \neq \emptyset$.*

If E is a stable expansion of Σ, then there is $I \subseteq \{1, \ldots, n\}$ such that $E \cap \mathcal{L} = Cn(\{\pi_i | i \in I\})$.

Proof: By [Marek and Truszczyński, 1991, Theorem 4.1c]. □

We say that a formula $L\phi_1 \wedge \cdots \wedge L\phi_n \wedge \neg L\psi_1 \wedge \cdots \wedge \neg L\psi_m \rightarrow \pi$ *has support in B* if $\{\phi_1, \ldots, \phi_n\} \subseteq Cn(B)$ and $\{\psi_1, \ldots, \psi_m\} \cap Cn(B) = \emptyset$.

Like in Sections 7.2.2 and 7.2.3, from now on we assume for simplicity that each inheritance network has only one individual x.

Lemma 7.3.9 *Let G be a good inheritance network and Δ_G the translation of G as given in Definition 7.2.1. Let X be a best explanation of $Th(G)$. Let E be a stable expansion of $X \cup Th(G)$. Then for all $ab(x, q', q) \in X$, $H(x, q') \in E$ and either $\neg H(x, q) \in E$ and $PD(q', q) \in Th(G)$ or $H(x, q) \in E$ and $ND(q', q) \in Th(G)$.*

Proof: Assume $\mathrm{ab}(x, q', q) \in X$. By Lemma 7.3.8 $E \cap \mathcal{L} = Cn(B)$ where $B = \{\pi | \pi \in \mathrm{Th}(G) \cup X$ or $\phi \rightarrow \pi \in \mathrm{Th}(G) \cup X$ has support in $B\}$.

First we show that $H(x, q') \in E$ and $\{\mathrm{PD}(q', q), \mathrm{ND}(q', q)\} \cap E \neq \emptyset$. So assume the opposite, that is, $H(x, q') \notin E$ or $\{\mathrm{PD}(q', q), \mathrm{ND}(q', q)\} \cap E = \emptyset$. Let $X' = X \backslash \{\mathrm{ab}(x, q', q)\}$. Let $B' = B \backslash \{\mathrm{ab}(x, q', q)\}$. We show that $E' \cap \mathcal{L} = Cn(B')$ for a stable expansion E' of $\mathrm{Th}(G) \cup X'$. By Lemma 7.3.8 it suffices to show that *for every formula $\pi \in B'$, either $\pi \in Th(G) \cup X'$ or there is $\phi \rightarrow \pi \in Th(G) \cup X'$ that has support in B', and no $\phi \rightarrow \pi \in Th(G) \cup X'$ such that $\pi \notin B'$ has support in B'*. It suffices to consider formulae in $\mathrm{Th}(G) \cup X'$ with occurrences of $\mathrm{ab}(x, q', q)$ because B and B' are consistent sets of literals and hence a formula with no occurrences of $\mathrm{ab}(x, q', q)$ has support in B if and only if it has support in B'. Assume $H(x, q) \in B'$ (proof for $\neg H(x, q)$ is similar.) Clearly $H(x, q) \in B$. Therefore $H(x, q) \in \mathrm{Th}(G)$ or there is $\phi \rightarrow H(x, q) \in \mathrm{Th}(G)$ that has support in B. In the first case $H(x, q) \in B'$ by definition of B'. It is not possible that $\phi = L\mathrm{PD}(q', q) \wedge LH(x, q') \wedge \neg L\mathrm{ab}(x, q', q)$, because by assumption $H(x, q') \notin E$ or $\{\mathrm{PD}(q', q), \mathrm{ND}(q', q)\} \cap E = \emptyset$. Hence ϕ does not have occurrences of $\mathrm{ab}(x, q', q)$, and therefore $\phi \rightarrow H(x, q)$ has support in B'. Assume $H(x, q) \notin B'$ (proof for $\neg H(x, q)$ is similar.) Clearly $H(x, q) \notin B$. Therefore no $\phi \rightarrow H(x, q) \in \mathrm{Th}(G)$ has support in B. Consider $\phi = L\mathrm{PD}(q', q) \wedge LH(x, q') \wedge \neg L\mathrm{ab}(x, q', q)$. By assumption $H(x, q') \notin E$ or $\{\mathrm{PD}(q', q), \mathrm{ND}(q', q)\} \cap E = \emptyset$ and hence $H(x, q') \notin B'$ or $\{\mathrm{PD}(q', q), \mathrm{ND}(q', q)\} \cap B' = \emptyset$. Therefore $\phi \rightarrow H(x, q)$ has no support in B'. Other ϕ have no occurrences of $\mathrm{ab}(x, q', q)$ and therefore no $\phi \rightarrow H(x, q) \in \mathrm{Th}(G)$ has support in B'. Now by Lemma 7.3.8 there is an extension E' of $\mathrm{Th}(G) \cup X'$ such that $E' \cap \mathcal{L} = Cn(B')$. Hence X' is an explanation of $\mathrm{Th}(G)$. Finally, $X(a) = X'(a)$ for all a such that $r(a) < r(q)$, and $X'(q)$

is better than $X(q)$ because $X'(q) \subset X(q)$, and $X(a) = X'(a)$ for all a such that $r(a) = r(q)$ and $a \neq q$. Therefore X' is preferable to X, which contradicts the fact that X is a best explanation. Therefore the assumption $H(x, q') \notin E$ or $\{\mathrm{PD}(q', q), \mathrm{ND}(q', q)\} \cap E = \emptyset$ is false, and finally $H(x, q') \in E$ and $\{\mathrm{PD}(q', q), \mathrm{ND}(q', q)\} \cap E \neq \emptyset$.

Because E is consistent, either $\mathrm{PD}(q', q) \in \mathrm{Th}(G)$ or $\mathrm{ND}(q', q) \in \mathrm{Th}(G)$. Because G is good, not both $\langle q', q, 0\rangle$ and $\langle q', q, 1\rangle$ are in E, and hence not both $\mathrm{PD}(q', q) \in \mathrm{Th}(G)$ and $\mathrm{ND}(q', q) \in \mathrm{Th}(G)$. Therefore exactly one of them is in $\mathrm{Th}(G)$.

Then we show that either $\neg H(x, q) \in E$ and $\mathrm{PD}(q', q) \in \mathrm{Th}(G)$ or $H(x, q) \in E$ and $\mathrm{ND}(q', q) \in \mathrm{Th}(G)$. So assume that $\mathrm{PD}(q', q) \in \mathrm{Th}(G)$ (the proof for $\mathrm{ND}(q', q) \in \mathrm{Th}(G)$ is similar.) And assume that $\neg H(x, q) \notin E$. Let $X' = (X \backslash \{\mathrm{ab}(x, q', q)\}) \cup \{\mathrm{ab}(x, q, t) | \langle q, t, n\rangle \text{ is in } G, n \in \{0, 1\}\}$. We show that X' is an explanation for $\mathrm{Th}(G)$ and X' is preferable to X. This contradicts the fact that X is a best explanation. Let $B' = B \backslash \{\mathrm{ab}(x, q', q)\} \cup \{\mathrm{ab}(x, q, t) | \langle q, t, n\rangle \text{ in } G, n \in \{0, 1\}\} \cup \{\chi\}$, where $\chi = H(x, q)$ if $\langle 'q, q, 0\rangle$ is in G ($\chi = \neg H(x, q)$ for the case in which $\langle q', q, 1\rangle$ is in G.) To prove that there is a stable expansion E' of $\mathrm{Th}(G) \cup X'$ such that $E' \cap \mathcal{L} = Cn(B')$, by Lemma 7.3.8 it suffices to show that *for every formula* $\pi \in B'$, *either* $\pi \in Th(G) \cup X'$ *or there is* $\phi \rightarrow \pi \in Th(G) \cup X'$ *that has support in* B', *and no* $\phi \rightarrow \pi \in Th(G) \cup X'$ *such that* $\pi \notin B'$ *has support in* B'. By assumption $\mathrm{PD}(q', q) \in E$ and hence $\langle q', q, 0\rangle$ is in G.

Consider $H(x, q) \in B'$. By assumpion $\mathrm{PD}(q', q) \in \mathrm{Th}(G)$. By definition $L\mathrm{PD}(q', q) \wedge LH(x, q') \wedge \neg L\mathrm{ab}(x, q', q) \rightarrow H(x, q) \in \mathrm{Th}(G)$. By the first part of this lemma, $H(x, q') \in B$ and therefore $H(x, q') \in B'$. By definition $\mathrm{ab}(x, q', q) \notin B'$. Hence $L\mathrm{PD}(q', q) \wedge LH(x, q') \wedge \neg L\mathrm{ab}(x, q', q) \rightarrow H(x, q) \in \mathrm{Th}(G)$ has support in B'.

Consider $H(x, p) \in B'$ such that $p \neq q$. Now $H(x, p) \in B$. Because $Cn(B) = E \cap \mathcal{L}$ for a stable expansion E of $\mathrm{Th}(G) \cup X$, by Lemma 7.3.8 there is $\phi \rightarrow H(x, p) \in \mathrm{Th}(G)$ that has support in B. Assume $\phi = L\mathrm{PD}(p', p) \wedge LH(x, p') \wedge \neg L\mathrm{ab}(x, p', p)$. Hence $\mathrm{ab}(x, p', p) \notin B$. Because $H(x, q) \notin B$, $p \neq q$. Therefore $\mathrm{ab}(x, p', p) \notin X'$. Also $\mathrm{PD}(p', p) \in B'$ and $H(x, p') \in B'$ because $\mathrm{PD}(p', p) \in B$ and $H(x, p') \in B$. Therefore $\phi \rightarrow H(x, p)$ has support in B'. Assume $\phi = L\mathrm{PS}(p', p) \wedge LH(x, p')$. Similarly $\mathrm{PS}(p', p) \in B'$ and $H(x, p') \in B'$ because $\mathrm{PS}(p', p) \in B$ and $H(x, p') \in B$. Therefore $\phi \rightarrow H(x, p)$ has support in B'. Proof for negative literals $\neg H(x, p) \in B'$ is similar.

Consider $\phi \rightarrow H(x, p) \in \mathrm{Th}(G)$ such that $H(x, p) \notin B'$. We show that it has no support in B'. Because $H(x, q) \in B'$, $p \neq q$. There are no $H(x, p) \in B \backslash B'$, and hence $H(x, p) \notin B$. By Lemma 7.3.8 $\phi \rightarrow H(x, p)$ has no support in B. Assume $\phi = L\mathrm{PD}(p', p) \wedge LH(x, p') \wedge \neg L\mathrm{ab}(x, p', p)$.

Hence $\{\mathrm{PD}(p',p), H(x,p')\} \not\subseteq B$ or $\mathrm{ab}(x,p',p) \in X$. If $\mathrm{ab}(x,p',p) \in X$, then $\mathrm{ab}(x,p',p) \in X'$ because $p \neq q$. Otherwise $\mathrm{ab}(x,p',p) \notin X$. If $\mathrm{PD}(p',p) \notin B$, then $\mathrm{PD}(p',p) \notin B'$ by definition. The remaining case is that $H(x,p') \notin B$. If $p' \neq q$, then clearly $H(x,p') \notin B'$. It is not possible that $p' = q$, because by definition $\mathrm{ab}(x,q',q) \in X'$ would hold, and by assumption this is not the case. Therefore $H(x,p') \notin B'$. Assume $\phi = L\mathrm{PS}(p',p) \wedge LH(x,p')$. Hence $\{\mathrm{PS}(p',p), H(x,p')\} \not\subseteq B$. If $\mathrm{PS}(p',p) \notin B$, then $\mathrm{PS}(p',p) \notin B'$ by definition. Otherwise assume that $\mathrm{PS}(p',p) \in B$. Hence $H(x,p') \notin B$. It is not possible that $p' = q$ because G is good and $\langle q',q,0\rangle$ is in G and by the fact that $\mathrm{PS}(p',p)$ is in G, $\langle p',p,2\rangle$ is in G. Therefore $H(x,p') \notin B'$.

Hence in all cases $\phi \rightarrow H(x,p)$ has no support in B'.

Therefore by Lemma 7.3.8 $E' \cap \mathcal{L} = Cn(B')$ for a stable expansion E' of $\mathrm{Th}(G) \cup X'$. Hence $X(a) = X'(a)$ for all a such that $r(a) < r(q)$ or $r(a) = r(q)$ and $a \neq q$, and $X'(q)$ is better than $X(q)$ because $X'(q) \subset X(q)$. Therefore X' is preferable to X, which contradicts the assumption that X is a best explanation. Therefore the assumption that $H(x,q) \notin E$ and $\neg H(x,q) \notin E$ is false. This establishes the result that either $H(x,q) \in E$ and $\mathrm{ND}(q',q) \in \mathrm{Th}(G)$ or $\neg H(x,q) \in E$ and $\mathrm{PD}(q',q) \in \mathrm{Th}(G)$. □

Lemma 7.3.10 *Let $G = \langle V, E_d, I, P, N\rangle$ be a good inheritance network and Δ_G the translation of G as given in Definition 7.2.1. Then for each best explanation X of Th(G) there is an extension E of Δ_G such that for all defeasible links from c_1 to c_2 in G, def(δ, E) if and only if ab(x, c_1, c_2) $\in X$, where δ is $c_1{:}c_2/c_2$ or $c_1{:}\neg c_2/\neg c_2$.*

Proof: Let X be a best explanation of $\mathrm{Th}(G)$. By definition of explanations, $\mathrm{Th}(G) \cup X$ has a stable expansion E that is consistent. Let $B' = \{p | H(x,p) \in E\} \cup \{\neg p | \neg H(x,p) \in E\}$. Let $E' = Cn(B')$. Because E is consistent, also B' and E' are consistent.

Lemma A For all $p{:}\pi/\pi \in D$, $\mathrm{def}(p{:}\pi/\pi, E')$ if and only if $\mathrm{ab}(x,p,q) \in X$, where $\pi \in \{q, \neg q\}$.
Proof: ($\Rightarrow$) Assume that $p{:}q/q \in D$ and $\mathrm{def}(p{:}q/q, E')$. Hence $p \in E'$ and $\neg q \in E'$. Because $E' = Cn(B')$ and B' is a consistent set of literals, $p \in B'$ and $\neg q \in B'$. By the definition of B', $H(x,p) \in E$ and $\neg H(x,q) \in E$. Because $p{:}q/q \in D$, there is a link $\langle p,q,0\rangle$ in G. By the definition of $\mathrm{Th}(G)$, $\mathrm{PD}(p,q) \in \mathrm{Th}(G) \subseteq E$. Because $L\mathrm{PD}(p,q) \wedge LH(x,p) \wedge \neg L\mathrm{ab}(x,p,q) \rightarrow H(x,q) \in \mathrm{Th}(G) \subseteq E$ and $\{\mathrm{PD}(p,q), H(x,p), \neg H(x,q)\} \subseteq E$, by Lemma 7.3.8 and the consistency of E, $\mathrm{ab}(x,p,q) \notin E$. Hence $\mathrm{ab}(x,p,q) \in X$. The proof of the implication from $\mathrm{def}(p{:}\neg q/\neg q, E')$ to $\mathrm{ab}(x,p,q) \in X$ is almost identical.

($\Leftarrow$) Assume that $p{:}q/q \in D$ and not def$(p{:}q/q, E')$. Hence $p \notin E'$ or $\neg q \notin E'$. Because $B' \subseteq E'$, $p \notin B'$ or $\neg q \notin B'$. By definition of B', $H(x,p) \notin E$ or $\neg H(x,q) \notin E$. Because $p{:}q/q \in D$, $\mathrm{PD}(p,q) \in \mathrm{Th}(G)$. Because X is a best explanation, by Lemma 7.3.9 $\mathrm{ab}(x,p,q) \notin X$. The proof of the implication from not def$(p{:}\neg q/\neg q, E')$ to $\mathrm{ab}(x,p,q) \notin X$ is almost identical. $\square$

We show that E' is an extension of $\Delta_G = \langle D, W\rangle$. Define $D_{E'} = \{p{:}\pi/\pi \in D | \mathrm{ab}(x,p,q) \notin X, \pi \in \{q, \neg q\}\} \cup \{p{:}/\pi \in D|\}$. By Lemma **A** members of $D_{E'}$ are not defeated in E'. By Theorem 2.1.3 and Lemma A.1.5 $E' = Cn_W(D_{E'})$ implies that E' is an extension of Δ_G. We derive the antecedent of this implication. Let $C_0, C_1, \ldots$ be the sets in the definition of $Cn_W(D_{E'})$.

Induction hypothesis: $\{l \in E' | l \text{ is a literal}, r(l) \leq i\} \subseteq C_i \subseteq E'$.

Base case $i = 0$: We give the proof for positive literals; the proof for negative literals is similar. Assume $p \in E'$. Because $r(p) = 0$, p is not the consequent of any link, $H(x,p) \in \mathrm{Th}(G)$ and $\langle x, p\rangle \in P$. By definition of Δ_G, $p \in W = C_0$. Assume $p \in C_0$. Because $C_0 = W$, $\langle x, p\rangle \in P$ by definition of Δ_G. By definition of $\mathrm{Th}(G)$, $H(x,p) \in \mathrm{Th}(G) \subseteq E$. By definition $p \in E'$.

Inductive case $i \geq 1$: We give the proof for positive literals; the proof for negative literals is similar. Assume $p \in E'$. Hence by the definition of E', $H(x,p) \in E$. Therefore by Lemma 7.3.8 either $H(x,p) \in \mathrm{Th}(G)$, or $L\mathrm{PD}(q,p) \wedge LH(x,q) \wedge \neg L\mathrm{ab}(x,q,p) \rightarrow H(x,p) \in \mathrm{Th}(G)$ for some $q \in \mathcal{L}$ and $\{\mathrm{PD}(q,p), H(x,q)\} \subseteq E$ and $\mathrm{ab}(x,q,p) \notin E$, or $L\mathrm{PS}(q,p) \wedge LH(x,q) \rightarrow H(x,p) \in \mathrm{Th}(G)$ for some $q \in \mathcal{L}$ and $\{\mathrm{PS}(q,p), H(x,q)\} \subseteq E$. In the first case, $\langle x,p\rangle \in P$ and by definition of Δ_G, $p \in W$. Hence $p \in C_i$ because $W \subseteq C_i$. In the second case $\langle q,p,0\rangle$ is in G and $q{:}p/p \in D$ by the definition of Δ_G. By the definition of E', $q \in E'$. Because $\mathrm{ab}(x,q,p) \notin E$, $q{:}p/p \in D_{E'}$. Because $r(q) \leq r(p) - 1$, by induction hypothesis $q \in C_{i-1}$. Hence $p \in C_i$. In the third case $\langle q,p,2\rangle$ is in G and by definition of Δ_G, $q{:}/p \in D$. By the definition of E', $q \in E'$. Because $r(q) \leq r(p) - 1$, by induction hypothesis $q \in C_{i-1}$. Because $q{:}/p \in D_{E'}$, $p \in C_i$ by definition.

Assume $p \in C_i$ (the proof for negative literals is similar.) If $p \in C_{i-1}$, then by induction hypothesis $p \in E'$. If $p \in C_i \backslash C_{i-1}$, then $q{:}p/p \in D_{E'}$ or $q{:}/p \in D_{E'}$ for some $q \in C_{i-1}$. By induction hypothesis $q \in E'$ and hence $H(x,q) \in E$. Case 1. If $q{:}p/p \in D_{E'}$, then $\langle p,q,0\rangle$ is in G and $\mathrm{ab}(x,q,p) \notin E$ and $\mathrm{PD}(q,p) \in \mathrm{Th}(G)$. By definition $L\mathrm{PD}(q,p) \wedge LH(x,q) \wedge \neg L\mathrm{ab}(x,p,q) \rightarrow H(x,p) \in Th(G) \subseteq E$. By Lemma 7.3.8, $H(x,p) \in E$. By definition $p \in E'$. Case 2. If $q{:}/p \in D_{E'}$, then $\langle p,q,2\rangle$ is in G and $\mathrm{PS}(q,p) \in \mathrm{Th}(G) \subseteq E$. By definition $L\mathrm{PS}(q,p) \wedge LH(x,q) \rightarrow H(x,p) \in Th(G) \subseteq E$. By Lemma 7.3.8 $H(x,p) \in E$. By definition $p \in E'$. This finished the induction proof.

Hence $\{l \in E' | l \text{ is a literal }, r(l) \geq 0\} \subseteq \bigcup_{i\geq 0} C_i$ and $\bigcup_{i \geq 0} C_i \subseteq E'$. Therefore $E' = \bigcup_{i\geq 0} C_i = Cn_W(D_{E'})$ and E' is an extension of Δ_G. $\square$

Lemma 7.3.11 *Let $G = \langle V, E_d, I, P, N\rangle$ be a good inheritance network and Δ_G the translation of G as given in Definition 7.2.1. Then for each extension E of Δ_G there is an explanation X of $Th(G)$ such that for all $\delta \in D$, def(δ, E) if and only if $ab(x, c_1, c_2) \in X$, where δ is $c_1{:}c_2/c_2$ or $c_1{:}\neg c_2/\neg c_2$.*

Proof: Let E be an extension of $\Delta_G = \langle D, W\rangle$. Let $X = \{\mathrm{ab}(x,p,q)|p{:}\pi/\pi \in D, \mathrm{def}(p{:}\pi/\pi, E), \pi \in \{q, \neg q\}\}$. We claim that X is an explanation for $\mathrm{Th}(G)$, that is, there is a consistent stable expansion E' of $\mathrm{Th}(G) \cup X$. Let $B = X \cup \{H(x,c)|c \in E\} \cup \{\neg H(x,c)|\neg c \in E\} \cup \{\mathrm{PD}(c_1,c_2)|c_1{:}c_2/c_2 \in D\} \cup \{\mathrm{ND}(c_1,c_2)|c_1{:}\neg c_2/\neg c_2 \in D\} \cup \{\mathrm{PS}(c_1,c_2)|c_1{:}/c_2 \in D\} \cup \{\mathrm{NS}(c_1,c_2)|c_1{:}/\neg c_2 \in D\}$. We show that $E' \cap \mathcal{L} = Cn(B)$ for a stable expansion E' of $\mathrm{Th}(G) \cup X$. By Lemma 7.3.8 it suffices to show that for every $\pi \in B$ either $\pi \in \mathrm{Th}(G) \cup X$ or there is $\phi \to \pi \in \mathrm{Th}(G) \cup X$ that has support in B, and no $\phi \to \pi \in \mathrm{Th}(G) \cup X$ such that $\pi \notin B$ has support in B. The set B is consistent because E is consistent by Lemma 7.2.5.

Assume $\mathrm{PD}(c,d) \in B$ (proofs for $\mathrm{ND}(c,d), \mathrm{PS}(c,d), \mathrm{NS}(c,d)$ are almost identical.) Now $\mathrm{PD}(c,d) \in B$ implies $c{:}d/d \in D$ implies $\langle c, d, 0\rangle \in E_d$ implies $\mathrm{PD}(c,d) \in \mathrm{Th}(G)$.

Assume $H(x,c) \in B$. If $\langle x, c\rangle \in P$, then $H(x,c) \in \mathrm{Th}(G)$ by definition. If $\langle x, c\rangle \notin P$, then c is the conclusion of a generating default $p{:}\sigma/c$ of E. Hence $p \in E$ and by definition $H(x,p) \in B$. Case 1. If σ is the empty sequence, then $\mathrm{PS}(p,c) \in \mathrm{Th}(G)$. Now $L\mathrm{PS}(p,c) \wedge LH(x,p) \to H(x,c) \in \mathrm{Th}(G)$ has support in B. Case 2. If σ consists of c, then $\mathrm{PD}(p,c) \in \mathrm{Th}(G)$. Because not $\mathrm{def}(p{:}c/c, E)$, $\mathrm{ab}(x,p,c) \notin X$ and hence $\mathrm{ab}(x,p,c) \notin E'$. Now $L\mathrm{PD}(p,c) \wedge LH(x,p) \wedge \neg L\mathrm{ab}(x,p,c) \to H(x,c) \in \mathrm{Th}(G)$ has support in B. The proof for $\neg H(x,c) \in B$ is almost identical.

Assume $H(x,c) \notin B$. Hence $c \notin E$. Now for all p such that $p{:}\sigma/c \in D$, either $p \notin E$ or $p \wedge \neg c \in E$. In the first case $H(x,p) \notin B$ and $H(x,p) \notin E'$. In the second case $\mathrm{def}(p{:}c/c, E)$, and hence $\mathrm{ab}(x,p,c) \in X$. Because E' is consistent, no $L\mathrm{PD}(p,c) \wedge LH(x,p) \wedge \neg L\mathrm{ab}(x,p,c) \to H(x,c) \in \mathrm{Th}(G)$ has support in B and no $L\mathrm{PS}(p,c) \wedge LH(x,p) \to H(x,c) \in \mathrm{Th}(G)$ has support in B. The proof for $\neg H(x,c) \notin B$ is almost identical. □

Next we give the definition of a counterpart of the relation *better than* in default logic. Here we use the correspondence between abnormality assumptions in explanations and defaults that are defeated in an extension. This definition is an intermediate stage in establishing the main result of this section. When E is an extension of a default theory $\langle D, W\rangle$, by $E(a)$ we denote the set of defaults $\delta \in D$ such that $\mathrm{def}(\delta, E)$ and the conclusion of δ is a or $\neg a$.

Definition 7.3.12 *Let $G = \langle V, E_d, I, P, N\rangle$ be an inheritance network and Δ_G the translation of G as given in Definition 7.2.1. Let E and E' be exten-*

sions of Δ_G, and $a \in V$. Then $E(a)$ is better than *$E'(a)$, if $E(a) \subset E'(a)$, or for each $c{:}\alpha/\alpha \in E(a)$ there is $c'{:}\alpha'/\alpha' \in E'(a)$ such that there is a non-empty positive path from c' to c in G and $E'(a)$ is non-empty.*

Definition 7.3.13 *Let $G = \langle V, E_d, I, P, N\rangle$ be an inheritance network and Δ_G the translation of G as given in Definition 7.2.1. Let E and E' be extensions of Δ_G. Then E is* preferable to *E', if for some $m \geq 0$,*

1. *$E(a) = E'(a)$ for all $a \in V$ such that $r(a) < m$,*
2. *for all $a \in V$ such that $r(a) = m$, $E(a) = E'(a)$ or $E(a)$ is better than $E'(a)$, and*
3. *for at least one $a \in V$ such that $r(a) = m$, $E(a)$ is better than $E'(a)$.*

An extension is a best *extension, if no extension is preferable to it.*

Lemma 7.3.14 *Let G be an inheritance network and Δ_G the translation of G as given in Definition 7.2.1. Let $Th(G)$ be the translation of G as given in Definition 7.3.2.*

If E is a best extension of Δ_G, then there is a best explanation X of $Th(G)$ such that for all links from c_1 to c_2 in G and defaults $\delta \in \{c_1{:}c_2/c_2, c_1{:}\neg c_2/\neg c_2\}$, $def(\delta, E)$ if and only if $ab(x, c_1, c_2) \in X$.

If X is a best explanation of $Th(G)$, then there is a best extension E of Δ_G such that for all links from c_1 to c_2 in G and defaults $\delta = c_1{:}c_2/c_2$ or $\delta = c_1{:}\neg c_2/\neg c_2$, $def(\delta, E)$ if and only if $ab(x, c_1, c_2) \in X$.

Proof: By Lemmata 7.3.10 and 7.3.11, for each extension there is a corresponding explanation and vice versa. The correspondence between best extensions and best explanations is due to the similarity of definitions of best. □

The following lemma shows that any extension of a default theory that is the translation of an inheritance network has as its subset an extension of a default theory that is the translation of a subset of the inheritance network that meets certain conditions.

Lemma 7.3.15 *Let G be a good inheritance network and $\Delta_G = \langle D \cup D', W\rangle$ the translation of G as given in Definition 7.2.1. Assume that the propositional variables in conclusions of D' do not occur in any prerequisite of a member of D and the defaults in D' are normal. Let E' be an extension of $\langle D \cup D', W\rangle$. Then there is an extension E of $\langle D, W\rangle$ such that for all $\delta \in D$, if $def(\delta, E)$ then $def(\delta, E')$, and if $appl(\delta, E')$ then $appl(\delta, E)$.*

Proof: $E' = Cn_W(\{\delta \in D \cup D' | \text{not def}(\delta, E')\})$ by Theorem 2.1.3 and Lemma A.1.5. Let $D_{nd} = \{\delta \in D | \text{not def}(\delta, E')\}$. By Lemma A.1.3 $E_0 = Cn_W(D_{nd}) \subseteq E'$. Let $D_a = \{\alpha{:}\sigma/\gamma \in D_{nd} | \alpha \in E_0\}$. Now by Lemma A.1.7, E_0 is an extension of $\langle D_a, W\rangle$ and $D_a = \text{GD}(E_0, \langle D_a, W\rangle)$. By repeated application of Lemma A.1.11 E_0 is an extension of $\langle D_{nd}, W\rangle$. Because $E_0 \subseteq E'$, for all $\delta \in D$, if def(δ, E_0), then def(δ, E'). By repeated application of Lemma A.1.9 E_0 is an extension of $\langle D_{nd} \cup D_d, W\rangle$ where $D_d = \{\delta \in D | \text{def}(\delta, E_0)\}$.

Lemma A For all $\alpha{:}\sigma/\gamma \in D_{nd} \cup D_d$, $\alpha \in E'$ implies $\alpha \in E_0$.
Proof: The proof is by induction. Let $C'_0, C'_1, \ldots$ be the sets in the definition of $Cn_W(\{\delta \in D \cup D' | \text{not def}(\delta, E')\})$ and $C_0, C_1, \ldots$ be the sets for $Cn_W(\{\delta \in D_{nd} \cup D_d | \text{not def}(\delta, E')\})$. *Induction hypothesis*: If $\alpha \in C'_i$ for $\alpha{:}\sigma/\gamma \in D$, then $\alpha \in C_i$. *Base case* $i = 0$: By definition $C_0 = W = C'_0$. *Inductive case* $i \geq 1$: Assume $\alpha \in C'_i$ for some $\alpha{:}\sigma/\gamma \in D_{nd} \cup D_d$. If $\alpha \in C'_{i-1}$ then by induction hypothesis $\alpha \in C_{i-1} \subseteq C_i$. If $\alpha \in C'_i \backslash C'_{i-1}$, then $C'_{i-1} \models \alpha'$ for some $\alpha'{:}\sigma'/\alpha \in \{\delta \in D \cup D' | \text{not def}(\delta, E')\}$. Because by assumption propositional variables in conclusions of defaults in D' do not occur in prerequisites of defaults in D, $\alpha'{:}\sigma'/\alpha \in D$. Because not def$(\alpha'{:}\sigma'/\alpha, E')$, $\alpha'{:}\sigma'/\alpha \in D_{nd}$. Because G is good, by Lemma 7.2.5 E' is consistent and its subset C'_{i-1} is consistent. Because C'_{i-1} is a consistent set of literals and $C'_{i-1} \models \alpha'$, $\alpha' \in C'_{i-1}$. Hence by induction hypothesis $\alpha' \in C_{i-1}$. Because $\alpha'{:}\sigma'/\alpha \in D_{nd}$, $\alpha \in C_i$. □

Lemma B $\alpha \in E_0$ for all $\alpha{:}\sigma/\gamma \in D \backslash (D_{nd} \cup D_d)$.
Proof: This is because def$(\alpha{:}\sigma/\gamma, E', E_0)$ by definition of D_{nd} and D_d, and by Lemma **A** $\alpha \in E_0$ iff $\alpha \in E'$. □

Lemma C Defaults in $D \backslash (D_{nd} \cup D_d)$ are normal.
Proof: Assume $\delta \in D \backslash (D_{nd} \cup D_d)$. Because $D \backslash (D_{nd} \cup D_d) \subseteq D \backslash D_{nd} = D \backslash \{\delta \in D | \text{not def}(\delta, E')\} = \{\delta \in D | \text{def}(\delta, E')\}$, def$(\delta, E')$. Therefore all $\delta \in D$ have at least one justification, and are normal. □

Lemma D The propositional variables in conclusions of $D \backslash (D_{nd} \cup D_d)$ do not occur in any prerequisite of a default in $D_{nd} \cup D_d$.
Proof: We show that for all $a{:}\pi/\pi \in D \backslash (D_{nd} \cup D_d)$ there is $b{:}\pi/\pi \in D' \cap \text{GD}(E', \Delta_G)$. Assume $a{:}\pi/\pi \in D \backslash (D_{nd} \cup D_d)$. Because $a{:}\pi/\pi \in D \backslash D_{nd}$, def$(a{:}\pi/\pi, E')$. Hence $a \wedge \overline{\pi} \in E'$. Hence $\overline{\pi} \in W$ or there is $b{:}\overline{\pi}/\overline{\pi} \in \text{GD}(E', \Delta_G)$. Because $a \in E'$, by Lemma **A** $a \in E_0$. Because $a{:}\pi/\pi \in D \backslash D_d$, not def$(a{:}\pi/\pi, E_0)$, and hence $a \wedge \overline{\pi} \notin E_0$. Because by Lemma **A** $a \in E'$, $\overline{\pi} \notin E_0$. Because $W \subseteq E_0$, $\overline{\pi} \notin W$. Hence there is $b{:}\overline{\pi}/\overline{\pi} \in \text{GD}(E', \Delta_G)$.

Assume $b{:}\overline{\pi}/\overline{\pi} \in D$. Because not def$(b{:}\overline{\pi}/\overline{\pi}, E')$, $b{:}\overline{\pi}/\overline{\pi} \in D_{nd}$. Hence by Lemma **A** $b \in E_0$. Because not def$(b{:}\overline{\pi}/\overline{\pi}, E_0)$, $\overline{\pi} \in E_0$. Because $a \in E_0$, def$(a{:}\pi/\pi, E_0)$, which contradicts the assumption that not def$(a{:}\pi/\pi, E_0)$. Hence $b{:}\overline{\pi}/\overline{\pi} \in D'$.

Hence if $a{:}\pi/\pi \in D\backslash(D_{nd}\cup D_d)$, then there is $b{:}\overline{\pi}/\overline{\pi} \in D'$. By the assumptions of the lemma the propositional variable in π does not occur in the prerequisite of any default in D. □

By Lemmata **C** and **D** and Lemma 7.2.4 there is an extension E of $\langle D, W\rangle$ such that for all $\delta \in D$, if def(δ, E) then def(δ, E'). It remains to show that for all $\delta \in D$, if appl(δ, E') then appl(δ, E). Because all $\delta \in D$ such that appl(δ, E') are in D_{nd}, and if appl(δ, E_0) then appl(δ, E) because def(δ, E) implies def(δ, E'), it suffices to show that $\alpha \in E_0$ for all $\alpha{:}\sigma/\gamma \in D_{nd}$ such that appl$(\alpha{:}\sigma/\gamma, E')$.

Induction hypothesis: for all $\alpha{:}\sigma/\gamma \in D_{nd}$ such that $r(\gamma) \leq i$, if $\alpha \in E'$ then $\alpha \in E_0$.

Base case $i = 0$: By definition of rank there are no $\alpha{:}\sigma/\gamma \in D_{nd}$ such that $r(\gamma) = 0$.

Inductive case $i \geq 1$: Let $\alpha{:}\sigma/\gamma$ be a member of D_{nd} such that $r(\gamma) = i$. Assume $\alpha \in E'$. Because E' is consistent by Lemma 7.2.5 and the conclusions of generating defaults of E' are literals, either $\alpha \in W$ or there is $\alpha'{:}\sigma'/\alpha \in \mathrm{GD}(E', \Delta_G)$. In the first case obviously $\alpha \in E_0$. In the second case by assumption $\alpha'{:}\sigma'/\alpha \in D$, and as not def$(\alpha'{:}\sigma'/\alpha, E')$, $\alpha'{:}\sigma'/\alpha \in D_{nd}$. By induction hypothesis $\alpha' \in E_0$. Because not def$(\alpha'{:}\sigma'/\alpha, E')$ and $E_0 \subseteq E'$, not def$(\alpha'{:}\sigma'/\alpha, E_0)$. As $\alpha' \in E_0$, by definition of extensions $\alpha \in E_0$. □

The following lemma shows that whenever all defaults with a rank less than some number are defeated in both of two extensions or in neither of them, then the conclusions of those defaults are in both of the extensions or in neither of them.

Lemma 7.3.16 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let E and E' be extensions of G such that $E(p) = E'(p)$ for all p such that $r(p) \leq n$. Then for all p such that $r(p) \leq n$, $p \in E$ if and only if $p \in E'$, and $\neg p \in E$ if and only if $\neg p \in E'$.*

Proof: Let $D' = \{p{:}\sigma/\beta \in D | r(\beta) \leq n\}$. Now by the definition of $r(\beta)$, $p'{:}\sigma'/\beta \in D'$ whenever $p{:}\sigma/\beta \in D'$ and $p'{:}\sigma'/\beta \in D$, and $p'{:}\sigma'/p \in D'$ whenever $p{:}\sigma/\beta \in D'$ and $p'{:}\sigma'/p \in D$. Hence by Lemma 7.2.8 $p \in E$ iff $p \in E'$ for all p such that $r(p) \leq n$. □

Lemma 7.3.17 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let E and E' be extensions of G such that $E(p) \neq E'(p)$ and $E(q) = E'(q)$ for all q such that $r(q) < r(p)$. If $\pi \in E$, where $\pi \in \{p, \neg p\}$, then $\overline{\pi} \in E'$.*

Proof: Assume $\pi \in E$. Because $E(p) \neq E'(p)$ there is $\delta = q{:}\beta/\beta \in E(p) \vartriangle E'(p)$ and $\beta \in \{p, \neg p\}$. Hence $q \wedge \overline{\beta} \in E$ or $q \wedge \overline{\beta} \in E'$. By Lemma 7.3.16 $q \in E$ and $q \in E'$. If def(δ, E, E'), $\beta = \overline{\pi}$ because $\pi \in E$ and E is consistent by Lemma 7.2.5. Because not def(δ, E'), $\pi \notin E'$ and by definition of extensions $\beta = \overline{\pi} \in E'$. If def$(\delta, E', E)$, $\overline{\beta} \in E'$. Because $q \in E$ and not def(δ, E), $\beta = \pi$. Hence $\overline{\pi} \in E'$. □

Whenever an extension is preferable to another extension, the latter can be improved at those lowest rank nodes where the former is better than the latter.

Lemma 7.3.18 *Let $G = \langle V, E_d, I, P, N\rangle$ be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let E and E' be extensions of Δ_G. Let $n \geq 0$ be such that $E(c) = E'(c)$ for all nodes c such that $r(c) < n$. Let $B = \{b \in V | E'(b)$ is better than $E(b), r(b) = n\}$. Let $\mathcal{P}_G$ be the strict partial order on defaults in Δ_G given in Definition 7.2.6. Then there is an extension E'' of Δ_G such that $E''(b) = E'(b)$ for all $b \in B$, and def(δ, E'') implies def(δ, E) for all $\delta \in D_L$, where $D_L = \{\delta \in D \backslash (D_B \cup D_{B'}) | \delta' \mathcal{P}_G \delta$ for no $\delta' \in D_B \cup D_{B'}\}$, $D_B = \{\alpha{:}\gamma/\gamma \in D | b \in B, \gamma \in \{b, \neg b\}\}$, and $D_{B'} = \{\alpha{:}/\gamma \in D | b \in B, \gamma \in \{b, \neg b\}\}$.*

Proof: Define $D_H = \{\delta \in D \backslash (D_B \cup D_{B'}) | \delta' \mathcal{P}_G \delta$ for some $\delta' \in D_B\}$.

Lemma A If $\alpha{:}/\beta \in D$ and $\beta \in \{a, \neg a\}$ for some $a \in V$ such that $r(a) = n$ and $E(a) \neq E'(a)$, then $\alpha \notin E \cup E'$.
Proof: Because $E(a) \neq E'(a)$, there is $\alpha{:}\pi/\pi \in D$ such that $\pi \in \{a, \neg a\}$ and def$(\alpha{:}\pi/\pi, E, E')$. Therefore $\overline{\pi} \in E$. By Lemma 7.3.17 $\pi \in E'$. Because G is good, by Lemma 7.2.5 $\pi \notin E$ and $\overline{\pi} \notin E'$. Because $\alpha{:}/\beta \in D$ and $\beta \in \{a, \neg a\}$, by definition of extensions either $\alpha \notin E$ or $\alpha \notin E'$. By Lemma 7.3.16 $\alpha \notin E \cup E'$. □

Lemma B For all $b \in B$, either $b \in E$ or $\neg b \in E$.
Proof: Let b be any member of B. Hence $E(b) \neq E'(b)$. Hence there is $\delta = p{:}\pi/\pi \in D$ such that def(δ, E, E') or def(δ, E', E). Hence either $\{p, \overline{\pi}\} \subseteq E$ or $\{p, \overline{\pi}\} \subseteq E'$. If $\{p, \overline{\pi}\} \in E$, then we are done. Otherwise $\{p, \overline{\pi}\} \subseteq E'$. By Lemma 7.3.16 $p \in E$. Because E is an extension of $\langle D, W\rangle$ and $p \in E$ and not def$(p{:}\pi/\pi, E)$, $\overline{\pi} \notin E$. By the definition of extensions $\pi \in E$. □

Lemma C Defaults $\delta \in D_H$ are normal.
Proof: Let δ be a member of D_H. By definition there is $\delta' \in D_B \cup D_{B'}$ such that $\delta' \mathcal{P}_G \delta$, that is, there is a path from the propositional variable in the conclusion of δ' to the prerequisite of δ. Because $\delta' \in D_B \cup D_{B'}$ there is $\delta'' \in D_B \cup D_{B'}$ such that δ' and δ'' have the same propositional variable in their conclusions and $\delta'' \in E(p) \vartriangle E'(p)$. Now δ'' is normal because it is defeated in an extension. Because there is a path from the propositional

variable in the conclusion of δ'', by definition, $\delta''\mathcal{P}_G\delta$. Because G is good, δ is normal. $\square$

By repeated applications of Lemma A.1.11 (justified by Lemma **A**) E is an extension of $\langle D_B \cup D_H \cup D_L, W\rangle$. Because defaults in D_B are normal by definition and defaults in D_H are normal by Lemma **C** and the propositional variables in conclusions of defaults in $D_B \cup D_H$ do not occur in the prerequisites of D_L (that would contradict the definitions of $\mathcal{P}_G$ and D_H and D_L), by Lemma 7.3.15 there is an extension E_1 of $\langle D_L, W\rangle$ such that for all $\delta \in D_L$, def(δ, E_1) implies def(δ, E), and appl(δ, E) implies appl(δ, E_1). Let $E_2 = Cn(E_1 \cup I)$ where $I = E' \cap (B \cup \{\neg p | p \in B\})$. Because members of B do not occur in D_L, def(δ, E_2) implies def(δ, E_1) for all $\delta \in D_L$.

We show that E_2 is an extension of $\langle D_L \cup D_B, W\rangle$. Let $\{\delta_1, \ldots, \delta_n\}$ be the defaults $p_i{:}\gamma_i/\gamma_i \in D_B$ such that $p_i \in E_1$. Define for all $i \in \{1, \ldots, n\}$,

$$\begin{array}{rcll} E'_0 & = & E_1 & \\ E'_i & = & Cn(E'_{i-1} \cup \{\gamma_i\}), & \text{if } \gamma_i \in I \\ E'_i & = & E'_{i-1}, & \text{if } \gamma_i \notin I. \end{array}$$

Induction hypothesis: E'_i is an extension of $\langle D_L \cup \{\delta_1, \ldots, \delta_i\}, W\rangle$.

Base case $i = 0$: Immediate, as $E'_0 = E_1$ is an extension of $\langle D_L, W\rangle$.

Inductive case $i \geq 1$: Assume that $\gamma_i \notin I$. By Lemma **B** $\overline{\gamma_i} \in I$. By Lemma A.1.9 E'_i is an extension of $\langle D_L \cup \{\delta_1, \ldots, \delta_i\}, W\rangle$. Assume $\gamma_i \in I$. By assumption $p_i \in E_1$, and because there are no defaults in $D_L \cup D_B$ with a prerequisite that is in B, $E'_i = Cn(E_{i-1} \cup \{\gamma\})$ is an extension of $\langle D_L \cup \{\delta_1, \ldots, \delta_i\}, W\rangle$ by Lemma A.1.14. This finishes the induction proof.

Hence E'_n is an extension of $\langle D_L \cup \{\delta_1, \ldots, \delta_n\}, W\rangle$. We have to show that $E_2 = E'_n$. Clearly $E'_n \subseteq E_2$. We show that for each $\beta \in I$ there is $\alpha{:}\beta/\beta \in D_B$ such that $\alpha \in E_1$, and hence $\beta \in E'_i \subseteq E'_n$ for some $i \in \{1, \ldots, n\}$. So assume $\beta \in I$. Because E' is consistent by Lemma 7.2.5 and conclusions of all generating defaults of E' are literals, either $\beta \in W$ or there is default $\alpha{:}\sigma/\beta \in \mathrm{GD}(E', \Delta_G)$. By Lemma 7.3.17 one of E or E' would be inconsistent if $\beta \in W$, and hence there is default $\alpha{:}\sigma/\beta \in \mathrm{GD}(E', \Delta_G)$. By construction of E_1 by Lemma 7.3.15, appl$(\alpha{:}\sigma/\beta, E_1)$. Hence $\alpha \in E_1 \subseteq E_2$.

By repeated applications of Lemma A.1.11 (justified by Lemma **A**) E_2 is an extension of $\langle D_B \cup D_{B'} \cup D_L, W\rangle$.

Because defaults in D_H are normal by Lemma **C**, by Lemma 7.2.4 there is an extension E_3 of $\langle D, W\rangle$ such that $E_2 \subseteq E_3$ and def(δ, E_3) implies def(δ, E_2) for all $\delta \in D_L$. Because def(δ, E_3) implies def(δ, E_2) implies def(δ, E_1) implies def(δ, E) for all $\delta \in D_L$, def(δ, E_3) implies def(δ, E) for all $\delta \in D_L$. $\square$

Lemma 7.3.19 *Let G be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let E and E' be extensions of Δ_G. Let $\mathcal{P}_G$ be the strict partial order given in Definition 7.2.6. Let a be a node in G such that $E(a)$ is better than $E'(a)$. Then def(δ, E', E) for all $\mathcal{P}_G$-minimal $\delta \in E(a) \vartriangle E'(a)$.*

Proof: Let δ be a member of $E(a) \vartriangle E'(a)$ such that def(δ, E, E'). Because $E(a)$ is better than $E'(a)$ and $E(a) \not\subset E'(a)$, there is default δ' in $E'(a)$ and a non-empty positive path from the prerequisite c' of δ' to the prerequisite c of δ. Choose δ' so that the positive path from c' to c is longest possible (that is, the set of links in it is set-inclusion maximal.) By definition of $\mathcal{P}_G$, $\delta'\mathcal{P}_G\delta$. If δ' were in $E(a)$, there would be $\delta'' \in E'(a)$ and a positive path from the prerequisite of δ'' to c, and this would contradict the maximality of the path from c' to c. Therefore $\delta' \notin E(a)$, and hence def(δ', E', E). Hence δ is not $\mathcal{P}_G$-minimal, and def(δ, E', E) for all $\mathcal{P}_G$-minimal members of $E(a) \vartriangle E'(a)$. □

Lemma 7.3.20 *Let $\mathcal{P}$ be a strict partial order on a finite set U. Let S be a subset of U such that $\mathcal{P} \cap (S \times S) = \emptyset$. Let P_1 and P_2 be sets such that $P_1 \cup P_2 = S$ and $P_1 \cap P_2 = \emptyset$. Let $\mathcal{P}'$ be the transitive closure of $\mathcal{P} \cup (P_1 \times P_2)$. Then $\mathcal{P}'$ is asymmetric, and therefore a strict partial order.*

Proof: We give a proof for the following statement, from which the claim can be obtained by selecting $\mathcal{R} = P_1 \times P_2$. If $\mathcal{P}$ is a strict partial order on a finite set U such that $\mathcal{P} \cap (S \times S) = \emptyset$ for a subset S of U, and $\mathcal{R}$ is a strict partial order on S, then the transitive closure of $\mathcal{P} \cup \mathcal{R}$ is a strict partial order.

Because every strict partial order has a strict total order that extends it, there is an injection f' from S to $\{1, \ldots, |S|\}$ such that $f(a) < f(b)$ if $a\mathcal{R}b$. Similarly, there is an injection f_1 from $\{c \in U\backslash S | c\mathcal{P}d \text{ for some } d \in S\}$ to $\{\ldots, -2, -1, 0\}$ such that $f(a) < f(b)$ if $a\mathcal{P}b$, and an injection f_2 from $\{c \in U\backslash S | c\mathcal{P}d \text{ for no } d \in S\}$ to $\{|S| + 1, |S| + 2, \ldots\}$ such that $f(a) < f(b)$ if $a\mathcal{P}b$. Define $f = f' \cup f_1 \cup f_2$ where functions are identified with their graphs. Clearly, f is an injection that defines a strict total order $\mathcal{T}$ on S as $a\mathcal{T}b$ if $f(a) < f(b)$. Because $\mathcal{P} \cup \mathcal{R} \subseteq \mathcal{T}$ and $\mathcal{T}$ is transitive and the transitive closure of $\mathcal{P} \cup \mathcal{R}$ is the smallest transitive relation that includes $\mathcal{P} \cup \mathcal{R}$, the transitive closure of $\mathcal{P} \cup \mathcal{R}$ is a subset of $\mathcal{T}$. Because $\mathcal{T}$ is asymmetric, every subset of it, including the transitive closure of $\mathcal{P} \cup \mathcal{R}$, is asymmetric. □

Theorem 7.3.21 *Let $G = \langle V, E_d, I, P, N\rangle$ be a good inheritance network and $\Delta_G = \langle D, W\rangle$ the translation of G as given in Definition 7.2.1. Let $\mathcal{P}_G$ be the strict partial order on D given in Definition 7.2.6. Then for every extension E of Δ_G, E is $\mathcal{P}_G$-preferred$_d$ ($\mathcal{P}_G$-preferred$_{2d}$) if and only if E is best.*

Proof: Let E be a $\mathcal{P}$-preferred$_{2d}$ extension of Δ_G and assume that there is an extension E' of Δ_G that is preferable to E. By Definition 7.3.13 there is n such that $E(c) = E'(c)$ for all nodes c such that $r(c) < n$, and $E'(c)$ is better than $E(c)$ or $E'(c) = E(c)$ for all c such that $r(c) = n$, and for some a such that $r(a) = n$, $E'(a)$ is better than $E(a)$. Let $B = \{b \in V | E'(b) \text{ is better than } E(b), r(b) = n\}$. This set is non-empty because $a \in B$. Let $D_B = \{\alpha{:}\sigma/\gamma \in D | b \in B, \gamma \in \{b, \neg b\}\}$. Let $D' = \{\delta \in D \backslash D_B | \delta' \mathcal{P}_G \delta \text{ for no } \delta' \in D_B\}$. By Lemma 7.3.18 there is an extension E'' of $\langle D, W\rangle$ such that $E'(b) = E''(b)$ for all $b \in B$, and def(δ, E'') implies def(δ, E) for all $\delta \in D'$.

Let $\mathcal{T}$ be any strict total order on D such that $\mathcal{P}_G \subseteq \mathcal{T}$. Let δ be the $\mathcal{T}$-least element of D such that def(δ, E, E'') or def(δ, E'', E). If $\delta \in D'$, then def(δ, E, E''). If $\delta \in E(b) \vartriangle E''(b)$ for some $b \in B$, then by Lemma 7.3.19 def(δ, E, E''). This exhausts all possible $\mathcal{T}$-least $\delta \in D$ such that def(δ, E, E'') or def(δ, E'', E), because for all $\delta' \in D \backslash (D' \cup D_B)$ there is $\delta'' \in D_B$ such that $\delta'' \mathcal{P}_G \delta'$, and by definition $E''(b) = E'(b)$ is better than $E(b)$ and hence there is $\delta''' \in D$ such that def(δ''', E, E''), and by definition of $\mathcal{P}_G$, whenever $\delta'' \mathcal{P}_G \delta'$ and δ''' and δ'' have the same propositional variable in the conclusion, $\delta''' \mathcal{P}_G \delta'$.

The conclusion that def(δ, E, E'') and def(δ', E) iff def(δ', E'') for all $\delta' \mathcal{T} \delta$ contradicts the fact that E is $\mathcal{P}_G$-preferred$_{2d}$. Hence the assumption that there is an extension that is preferable to E is false, and all $\mathcal{P}_G$-preferred$_{2d}$ extensions are best extensions. By Lemma 3.1.7 all $\mathcal{P}_G$-preferred$_d$ extensions are $\mathcal{P}_G$-preferred$_{2d}$ extensions and therefore best extensions.

The second direction, that is, each best extension is a $\mathcal{P}_G$-preferred$_d$ extension, is as follows.

Lemma A Let p be a node in G, and E and E' and E'' consistent extensions of Δ_G. If $E(p) \neq E'(p)$ and $E(p) \neq E''(p)$ and $E(q) = E'(q) = E''(q)$ for all q such that $r(q) < r(p)$, then $E'(p) = E''(p)$.
Proof: Assume $a{:}\pi/\pi \in E''(p)$. Hence $a \wedge \overline{\pi} \in E''$. Hence by Lemma 7.3.17 $\pi \in E$. Hence by Lemma 7.3.17 $\overline{\pi} \in E'$. By Lemma 7.3.16 $a \in E'$. Hence $a{:}\pi/\pi \in E'(p)$. Therefore $E''(p) \subseteq E'(p)$. A symmetric proof shows that $E'(p) \subseteq E''(p)$. Hence $E'(p) = E''(p)$. $\square$

Assume E is a best extension of Δ_G. A node p is a *d-node* (with respect to E'), if for an extension E', $E(q) = E'(q)$ for all q such that $r(q) < r(p)$ and $E(p) \neq E'(p)$ and $E'(p)$ is not better than $E(p)$. We construct a strict partial order $\mathcal{P}$ such that $\mathcal{P}_G \subseteq \mathcal{P}$ and $\alpha{:}\beta/\beta \mathcal{P} \alpha'{:}\beta'/\beta'$, where $\beta \in \{c, \neg c\}$ and $\beta' \in \{c', \neg c'\}$, whenever

1. $r(c) < r(c')$,

2. $r(c) = r(c')$ and c is a d-node and c' is not, or
3. $\alpha{:}\beta/\beta$ and $\alpha'{:}\beta'/\beta'$ are $\mathcal{P}_G$-minimal in $\{\alpha{:}\beta/\beta \in D | \alpha \in E, \beta \in \{c, \neg c\}\}$ and $\mathrm{appl}(\alpha{:}\beta/\beta, E)$ and $\mathrm{def}(\alpha'{:}\beta'/\beta', E)$.

Let $\mathcal{P}_1$ be the transitive closure of $\mathcal{P}_G \cup \{\langle \delta, \delta' \rangle | r(\delta) < r(\delta')\}$. The relation $\mathcal{P}_1$ is asymmetric because $\delta' \mathcal{P}_G \delta$ for no defaults such that $r(\delta) < r(\delta')$. Let $\mathcal{P}_2$ be the transitive closure of $\mathcal{P}_1 \cup \{\langle \alpha{:}\beta/\beta, \alpha'{:}\beta'/\beta' \rangle | r(\delta) = r(\delta'), \beta \in \{b, \neg b\}, \beta' \in \{b', \neg b'\}, b$ is a d-node , b' is not a d-node$\}$. The relation $\mathcal{P}_2$ is asymmetric because $\mathcal{P}_1$ does not relate defaults of the same rank that have different propositional variables in their conclusions. Define $\Lambda_c = \{\alpha{:}\beta/\beta \in D | \alpha \in E, \beta \in \{c, \neg c\}\}$ for every node $c \in V$. Let $\mathcal{P}$ be the transitive closure of $\mathcal{P}_2 \cup \{\langle \delta, \delta' \rangle \subseteq D \times D | \delta$ and δ' are $\mathcal{P}_G$-minimal in $\Lambda_c, c \in V, \mathrm{appl}(\delta, E), \mathrm{def}(\delta', E)\}$. Because $\mathcal{P}_2$ does not relate members of Λ_c, by Lemma 7.3.20 $\mathcal{P}$ is asymmetric.

Lemma R Let p be a d-node with respect to E'. Then for all $\delta \in E(p) \backslash E'(p)$ there is $\delta' \in E'(p) \backslash E(p)$ such that $\delta' \mathcal{P} \delta$.
Proof: Because $E'(p)$ is not better than $E(p)$, there is $\delta_0 \in E'(p) \backslash E(p)$ such that there is no $\delta_1 \in E(p) \backslash E'(p)$ with a non-empty positive path from the prerequisite of δ_1 to the prerequisite of δ_0. Because $\mathrm{def}(\delta_0, E')$, $a_0 \in E'$ where $\delta_0 = a_0{:}\pi/\pi$. By Lemma 7.3.16 $a_0 \in E$. If there is $\delta_2 \in E'(p) \backslash E(p)$ and a non-empty positive path from the prerequisite of δ_2 to a_0, then let $\delta_3 = \delta_2$, otherwise let $\delta_3 = \delta_0$. Clearly δ_3 is $\mathcal{P}_G$-minimal in $\{\alpha{:}\beta/\beta \in D | \alpha \in E, \beta \in \{p, \neg p\}\}$. Hence $\delta_3 \mathcal{P} \delta_0$ and $\delta_3 \in E'(p) \backslash E(p)$. □

Lemma S There is $\delta^* \in D$ such that $\mathrm{def}(\delta^*, E', E)$ and π^* is a d-node and $\mathrm{def}(\delta', E)$ if and only if $\mathrm{def}(\delta', E')$ for all $\delta' \in D$ such that $r(\pi) < r(\pi^*)$, where $\delta' = a{:}\pi/\pi$ and $\delta^* = a^*{:}\pi^*/\pi^*$.
Proof: Let n be the least number such that $E(p) = E'(p)$ for all p such that $r(p) < n$. If E' were preferable to E, $E(p) = E'(p)$ or $E'(p)$ is better than $E(p)$ for all p such that $r(p) = n$, and $E'(q)$ is better than $E(q)$ for some q such that $r(q) = n$. Because this is not the case, $E(p) \neq E'(p)$ and $E'(p)$ is not better than $E(p)$ for some p such that $r(p) = n$. Now p is a d-node. By the definition of *better than* there is $\delta \in E'(p) \backslash E(p)$. Hence $\mathrm{def}(\delta, E', E)$. □

By Lemma **S** there is $\delta^* = a^*{:}\pi^*/\pi^*$ with the given properties. Assume $\mathrm{def}(\delta, E, E')$ where $\delta = a{:}\pi/\pi$ and $\pi \in \{p, \neg p\}$. If $r(\pi) > r(\pi^*)$, by definition $\delta^* \mathcal{P} \delta$. Otherwise $r(\pi) = n$. If p is not a d-node, then by definition $\delta^* \mathcal{P} \delta$. Hence assume that p is a d-node. By Lemma **A** p is a d-node with respect to E'. By Lemma **R** there is $\delta' \mathcal{P} \delta$ such that $\delta' \in E'(p) \backslash E(p)$. Therefore for all δ such that $\mathrm{def}(\delta, E, E')$ there is $\delta' \mathcal{P} \delta$ such that $\mathrm{def}(\delta', E', E)$. Because $\mathcal{P} \subseteq \mathcal{T}$, also $\delta' \mathcal{T} \delta$.

As this holds for any E', $\mathcal{T}$ is a $\Delta_G, \mathcal{P}_G$-characteristic$_d$ ordering for E. Because E is a $\mathcal{P}_G$-preferred$_d$ extension of Δ_G and every $\mathcal{P}$-preferred$_d$ extension is also a $\mathcal{P}$-preferred$_{2d}$ extension by Lemma 3.1.7, E is a $\mathcal{P}_G$-preferred$_{2d}$ extension of Δ_G. □

The theorem shows that Gelfond and Przymusinska's inheritance theory coincides with the inheritance theory presented in Section 7.2 both in the case with $\mathcal{P}_G$-preferred and $\mathcal{P}_G$-preferred$_2$ extensions. By Lemma 7.2.10 $\mathcal{P}_G$-preferred extensions coincide with $\mathcal{P}_G$-preferred$_d$ extensions. We do not have an equivalence result for $\mathcal{P}_G$-preferred$_{2d}$ extensions.

Theorem 3.1.15 shows that preferred and preferred$_2$ extensions always coincide if priorities are layered and do not always coincide if they are not layered. In Section 7.4.1 it is shown that there are inheritance networks for which non-layered priorities are necessary. Hence the above theorem shows that for inheritance networks, even though the priorities in general are not layered, preferred and preferred$_2$ extensions still coincide.

7.3.2 An inheritance theory by Brewka

Brewka [1987] presents a translational inheritance theory by using predicate logic and circumscription. The theory is about the class of inheritance networks with the following properties. Defeasible links may not be followed by any links and a defeasible link and a strict link may not end in the same node. All strict links are positive. These restrictions correspond to a division of the nodes to classes and properties. All links between classes are strict, and all links between classes and properties are defeasible. No link starts from a property node. This class of inheritance networks has been investigated also by Gelfond and Przymusinska [1990].

Definition 7.3.22 (Class-property network) *Let $G = \langle V, E, I, P, N\rangle$ be an inheritance network. Then G is a* class-property network, *if there is a set $C \subseteq V$ such that $E \subseteq (C \times C \times \{2\}) \cup (C \times (V \backslash C) \times \{0, 1\})$. The members of C are the* classes *and the members of $V \backslash C$ are the* properties *in G.*

Brewka allows non-Boolean properties, that is, instead of only two kinds of links – positive and negative – there may be arbitrarily many different kinds of links from classes to properties. We consider only the Boolean case, and assume that inheritance networks have only one individual U. We give a translation from our definition of inheritance networks to Brewka's translation. Brewka's representation of inheritance networks uses a Lisp-like notation that is not presented here.

Let $G = \langle V, E, I, P, N\rangle$ be a class-property network where $I = \{U\}$ and U is a symbol. Let C be the classes in G. Links in E are translated to the following formulae of the predicate logic [Brewka, 1987].

$$\begin{aligned}\Phi_c \;\; = \;\; & \{\textsc{Specializes}(x,y) | \langle x,y,2\rangle \in E\} \\ & \cup \{\textsc{Has-slot}(x,y,T) | \langle x,y,0\rangle \in E\} \\ & \cup \{\textsc{Has-slot}(x,y,F) | \langle x,y,1\rangle \in E\}\end{aligned}$$

An individual may be a member of a class, and an individual may have or not have a property.

$$\begin{aligned}\Phi_i \;\; = \;\; & \{\textsc{Is}(U,c) | \langle U,c\rangle \in P, c \in C\} \\ & \cup \{\textsc{Holds}(p,U,T) | \langle U,p\rangle \in P, p \notin C\} \\ & \cup \{\textsc{Holds}(p,U,F) | \langle U,p\rangle \in N, p \notin C\}\end{aligned}$$

No two symbols refer to the same class or property.

$$\Phi_u = \{s \neq s' | \{s,s'\} \subseteq V \cup \{U,T,F\}, s \neq s'\}$$

In addition to the above formulae that depend on the inheritance networks, the following set Φ_b of formulae is needed. These formulae state that the following. The class-inclusion relation Specializes propagates class-membership. Class-inclusion is transitive. If an individual that is a member of a class is not exceptional with respect to a property associated with the class then the individual has the default value of the property. An individual is exceptional with respect to a property associated with a class if there is a subclass of the class with a different value for the same property and the individual belongs to that subclass. An individual may not have two different values for one property.

$$\begin{aligned}&\forall x \forall y(\textsc{Specializes}(x,y) \rightarrow (\textsc{Is}(U,x) \rightarrow \textsc{Is}(U,y)))\\&\forall x \forall y \forall z(\textsc{Specializes}(x,y) \wedge \textsc{Specializes}(y,z) \rightarrow \textsc{Specializes}(x,z))\\&\forall x \forall y \forall v(\textsc{Has-slot}(x,y,v) \wedge \textsc{Is}(U,x) \wedge \neg\textsc{Exceptional}(U,x,y) \rightarrow \textsc{Holds}(y,U,v))\\&\forall x_1 \forall x_2 \forall v_1 \forall v_2 \forall z(\textsc{Is}(U,x_1) \wedge \textsc{Has-slot}(x_1,z,v_1) \wedge \textsc{Specializes}(x_1,x_2) \wedge\\&\qquad\qquad \textsc{Has-slot}(x_2,z,v_2) \wedge v_1 \neq v_2 \rightarrow \textsc{Exceptional}(U,x_2,z))\\&\forall x \forall v_1 \forall v_2(\textsc{Holds}(x,U,v_1) \wedge \textsc{Holds}(x,U,v_2) \rightarrow v_1 = v_2)\end{aligned}$$

The use of the constant symbol U instead of a universally quantified variable u reflects our assumption that there is only one individual in each inheritance network.

Brewka says that Exceptional is circumscribed in $\Phi = \Phi_c \cup \Phi_i \cup \Phi_u \cup \Phi_b$ and all other predicates in Φ are varied. The definition of circumscription is given in 6.4.

We show the correspondence of Brewka's translation to our translation. The notation $M(S)$ denotes the interpretation of the predicate symbol S in the model M. Models of predicate logic are described for example in [Enderton, 1972].

Lemma 7.3.23 *Let $G = \langle V, E, I, P, N\rangle$ be a class-property network where $I = \{U\}$. Let Φ be the translation of G as defined by Brewka. Let M be a model such that $M \models \Phi$. Then there is a model M' such that $M' \models \Phi$ and $M'(\textsc{Exceptional}) = M(\textsc{Exceptional})$ and $M'(\textsc{Holds}) = M(\textsc{Holds})$, and for all x and y and z, $M' \models \textsc{Specializes}(x, y, z)$ if and only if $\Phi \models \textsc{Specializes}(x, y, z)$, and for all x and y, $M' \models \textsc{Is}(x, y)$ if and only if $\Phi \models \textsc{Is}(x, y)$, and for all x and y and z, $M' \models \textsc{Has-slot}(x, y, z)$ if and only if $\Phi \models \textsc{Has-slot}(x, y, z)$,*

Proof: For simplicity and without loss of generality (because of the formulae Φ_u in Φ, no two constant symbols refer to the same individual) we assume that constant symbols are interpreted as themselves in M.

Let M' be a model with the same universe as M and where constant symbols are interpreted as themselves and

$$\begin{aligned} M'(\textsc{Exceptional}) &= M(\textsc{Exceptional}), \\ M'(\textsc{Is}) &= \{\langle U, p\rangle | \Phi \models \textsc{Is}(U, p)\}, \\ M'(\textsc{Specializes}) &= \{\langle x, y\rangle | \Phi \models \textsc{Specializes}(x, y)\}, \\ M'(\textsc{Has-slot}) &= \{\langle x, y, v\rangle | \Phi \models \textsc{Has-slot}(x, y, v)\}, \text{and} \\ M'(\textsc{Holds}) &= M(\textsc{Holds}). \end{aligned}$$

We show that $M' \models \Phi$ by considering each member of Φ in turn. Assume that $\textsc{Specializes}(x, y) \in \Phi$. Then $\Phi \models \textsc{Specializes}(x, y)$ and by definition $\langle x, y\rangle \in M'(\textsc{Specializes})$. Hence $M' \models \textsc{Specializes}(x, y)$. Assume that $\textsc{Is}(x, y) \in \Phi$. Then $\Phi \models \textsc{Is}(x, y)$ and by definition $\langle x, y\rangle \in M'(\textsc{Is})$. Hence $M' \models \textsc{Is}(x, y)$. Assume that $\textsc{Has-slot}(x, y, v) \in \Phi$. Then $\Phi \models \textsc{Has-slot}(x, y, v)$ and by definition $\langle x, y, v\rangle \in M'(\textsc{Has-slot})$. Hence $M' \models \textsc{Has-slot}(x, y, v)$. Assume that $\textsc{Holds}(x, y, v) \in \Phi$. Because $M \models \Phi$ and $M(\textsc{Holds}) = M'(\textsc{Holds})$, $M' \models \textsc{Holds}(x, y, v)$.

Consider $\forall x \forall y(\textsc{Specializes}(x, y) \rightarrow (\textsc{Is}(U, x) \rightarrow \textsc{Is}(U, y))) \in \Phi$. Assume that $M' \models \textsc{Specializes}(x, y) \wedge \textsc{Is}(U, x)$ for some x and y. Hence by definition $\Phi \models \textsc{Specializes}(x, y) \wedge \textsc{Is}(U, x)$. Clearly $\Phi \models \textsc{Is}(U, y)$. Hence by definition $M' \models \textsc{Is}(U, y)$.

Consider $\forall x \forall y \forall z(\textsc{Specializes}(x, y) \wedge \textsc{Specializes}(y, z) \rightarrow \textsc{Specializes}(x, z)) \in \Phi$. Assume that $M' \models \textsc{Specializes}(x, y) \wedge \textsc{Specializes}(y, z)$ for some x, y and z. Hence by definition $\Phi \models \textsc{Specializes}(x, y) \wedge \textsc{Specializes}(y, z)$. Clearly $\Phi \models \textsc{Specializes}(x, z)$. Hence by definition $M' \models \textsc{Specializes}(x, z)$.

Consider $\forall x \forall y \forall v(\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \rightarrow \textsc{Holds}(y, U, v)) \in \Phi$. Assume that for some x, y and v, $M' \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y)$. Hence $\Phi \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x)$ and therefore $M \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x)$. Because $M'(\textsc{Exceptional}) = M(\textsc{Exceptional})$, $M \models \neg\textsc{Exceptional}(U, x, y)$. Because $M \models \phi$, $M \models$

$\textsc{Holds}(y, U, v)$. Because $M'(\textsc{Holds}) = M(\textsc{Holds})$, $M' \models \textsc{Holds}(y, U, v)$.

Consider $\phi = \forall x_1 \forall x_2 \forall v_1 \forall v_2 \forall z(\textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2 \rightarrow \textsc{Exceptional}(U, x_2, z)) \in \Phi$. Assume $M' \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$ for some x_1, x_2, v_1, v_2, z. Hence $\Phi \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$ and $M \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$. Because $M \models \phi$, $M \models \textsc{Exceptional}(U, x_2, z)$. Because $M'(\textsc{Exceptional}) = M(\textsc{Exceptional})$, $M' \models \textsc{Exceptional}(U, x_2, z)$.

Consider $\phi = \forall x \forall v_1 \forall v_2 (\textsc{Holds}(x, U, v_1) \wedge \textsc{Holds}(x, U, v_2) \rightarrow v_1 = v_2) \in \Phi$. Because $M \models \phi$ and $M(\textsc{Holds}) = M'(\textsc{Holds})$, $M' \models \phi$. □

Lemma 7.3.24 *Let $G = \langle V, E, I, P, N\rangle$ be a class-property network with $I = \{U\}$ and classes $C \subseteq V$. Let Φ be the translation of G as defined by Brewka. Let $\Delta_G = \langle D, W\rangle$ be the translation of G as given in Definition 7.2.1. Then for all $p \in C$, $W \cup I_s(D) \models p$ if and only if $\Phi \models \textsc{Is}(U, p)$ (I and I_s are defined in Section 7.2.3.)*

Proof: Assume $p \in C \backslash Cn(W \cup I_s(D))$. Hence there is a propositional model M such that $M \models W \cup I_s(D)$ and $M \not\models p$. Define the model M' as follows. The universe of M' is $V \cup \{U, T, F\}$, the constant symbols $V \cup \{U, T, F\}$ are interpreted as themselves, and the predicate symbols are interpreted as follows.

$$\begin{aligned}
M'(\textsc{Specializes}) &= \text{the transitive closure of } \{\langle x, y\rangle | \langle x, y, 2\rangle \in E\} \\
M'(\textsc{Has-slot}) &= \{\langle x, y, T\rangle | \langle x, y, 0\rangle \in E\} \\
&\quad \cup \{\langle x, y, F\rangle | \langle x, y, 1\rangle \in E\} \\
M'(\textsc{Exceptional}) &= (V \cup \{U, T, F\})^3 \\
M'(\textsc{Is}) &= \{U\} \times \{q \in C | M \models q\} \\
M'(\textsc{Holds}) &= \{\langle x, U, T\rangle | \langle U, x\rangle \in P\} \cup \{\langle x, U, F\rangle | \langle U, x\rangle \in N\}
\end{aligned}$$

Clearly $M' \not\models \textsc{Is}(U, p)$. We show that $M' \models \Phi_c \cup \Phi_i \cup \Phi_b \cup \Phi_u$. Assume $\textsc{Specializes}(x, y) \in \Phi_c$. By definition $\langle x, y, 2\rangle \in E$. Hence by definition $\langle x, y\rangle \in M'(\textsc{Specializes})$ and $M' \models \textsc{Specializes}(x, y)$. Assume $\textsc{Has-slot}(x, y, T) \in \Phi_c$. By definition $\langle x, y, 0\rangle \in E$. Hence by definition $\langle x, y, T\rangle \in M'(\textsc{Has-slot})$ and $M' \models \textsc{Has-slot}(x, y, T)$. Assume $\textsc{Has-slot}(x, y, F) \in \Phi_c$. By definition $\langle x, y, 1\rangle \in E$. Hence by definition $\langle x, y, F\rangle \in M'(\textsc{Has-slot})$ and $M' \models \textsc{Has-slot}(x, y, F)$. Assume $\textsc{Is}(U, c) \in \Phi_i$. By definition $\langle U, c\rangle \in P$. Hence by definition $c \in W$. Because $M \models W$, $M \models c$. By definition $\langle U, c\rangle \in M'(\textsc{Is})$. Hence $M' \models \textsc{Is}(U, c)$. Assume $\textsc{Holds}(p, U, T) \in \Phi_i$. By definition $\langle U, p\rangle \in P$. Hence by definition $\langle p, U, T\rangle \in M'(\textsc{Holds})$. Hence

$M' \models \textsc{Holds}(p, U, T)$. Assume $\textsc{Holds}(p, U, F) \in \Phi_i$. By definition $\langle U, p\rangle \in N$. Hence by definition $\langle p, U, F\rangle \in M'(\textsc{Holds})$. Hence $M' \models \textsc{Holds}(p, U, F)$.

Consider $\forall x \forall y(\textsc{Specializes}(x, y) \rightarrow (\textsc{Is}(U, x) \rightarrow \textsc{Is}(U, y))) \in \Phi_b$. Assume that for some x and y, $M' \models \textsc{Specializes}(x, y)$ and $M' \models \textsc{Is}(U, x)$. By definition there is a chain of links $\{\langle x, x_1, 2\rangle, \langle x_1, x_2, 2\rangle, \ldots, \langle x_n, y, 2\rangle\} \subseteq E$ and $M \models x$ for some $n \geq 0$. By definition $\{x{:}/x_1, x_1{:}/x_2, \ldots, x_n{:}/y\} \subseteq D$ and hence by definition $\{x \rightarrow x_1, x_1 \rightarrow x_2, \ldots, x_n \rightarrow y\} \subseteq I_s(D)$. Because $M \models I_s(D) \cup \{x\}$, $M \models y$. Hence by definition $\langle U, y\rangle \in M'(\textsc{Is})$ and $M' \models \textsc{Is}(U, y)$.

Consider $\forall x \forall y \forall z(\textsc{Specializes}(x, y) \wedge \textsc{Specializes}(y, z) \rightarrow \textsc{Specializes}(x, z)) \in \Phi_b$. This formula is true in M' by the transitivity of $M'(\textsc{Specializes})$.

Consider $\forall x \forall y \forall v(\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \rightarrow \textsc{Holds}(y, U, v)) \in \Phi_b$ and $\forall x_1 \forall x_2 \forall v_1 \forall v_2 \forall z(\textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2 \rightarrow \textsc{Exceptional}(U, x_2, z)) \in \Phi_b$. These formulae are true in M' because $M'(\textsc{Exceptional}) = (V \cup \{U, T, F\})^3$.

Consider $\forall x \forall v_1 \forall v_2(\textsc{Holds}(x, U, v_1) \wedge \textsc{Holds}(x, U, v_2) \rightarrow v_1 = v_2) \in \Phi_b$. Assume $M' \models \textsc{Holds}(x, U, v_1) \wedge \textsc{Holds}(x, U, v_2)$. Hence $\langle U, x\rangle \in P$ if $v_1 = T$ or $v_2 = T$ and $\langle U, x\rangle \in N$ if $v_1 = F$ or $v_2 = F$. If $v_1 \neq v_2$, then $\langle U, x\rangle \in P \cap N$ and by definition $\{x, \neg x\} \in W$. But W is consistent because $M \models W$. Therefore $v_1 = v_2$. Now because $M' \models \Phi$ and $M' \not\models \textsc{Is}(U, p)$, $\Phi \not\models p$. This completes the first direction of the proof of the equivalence.

Assume $\Phi \not\models \textsc{Is}(U, p)$ for $p \in C$, that is, there is model M such that $M \models \Phi$ and $M \not\models \textsc{Is}(U, p)$. Let M' be a propositional model such that $M' \models q$ if and only if $q \in C$ and $M \models \textsc{Is}(U, q)$ or $q \in V \backslash C$ and $M \models \textsc{Holds}(q, U, T)$. Clearly $M' \not\models p$. We show that $M' \models W \cup I_s(D)$. Assume $q \in W$. Then $\langle U, q\rangle \in P$. If $q \in C$, then $\textsc{Is}(U, q) \in \Phi$ and $M \models \textsc{Is}(U, q)$ and $M' \models q$ by definition. If $q \in V \backslash C$, then $\textsc{Holds}(q, U, T) \in \Phi$ and $M' \models q$ by definition. Assume $\neg q \in W$. Then $\langle U, q\rangle \in N$. Now $q \in V \backslash C$ and $\textsc{Holds}(q, U, F) \in \Phi$. Because $M \models \forall x \forall v_1 \forall v_2(\textsc{Holds}(x, U, v_1) \wedge \textsc{Holds}(x, U, v_2) \rightarrow v_1 = v_2)$ and $M \models T \neq F$ and $M \models \Phi$, $M \not\models \textsc{Holds}(q, U, T)$. Hence $M' \not\models q$ by definition and $M' \models \neg q$. Assume $q \rightarrow q' \in I_s(D)$. Hence $\langle q, q', 2\rangle \in E$, $\textsc{Specializes}(q, q') \in \Phi$ and $M \models \textsc{Specializes}(q, q')$. Assume $M' \models q$. Hence $M \models \textsc{Is}(U, q)$. Because $M \models \forall x \forall y(\textsc{Specializes}(x, y) \rightarrow (\textsc{Is}(U, x) \rightarrow \textsc{Is}(U, y)))$, $M \models \textsc{Is}(U, q')$. Hence by definition $M' \models q'$. Therefore $M' \models q \rightarrow q'$. □

Lemma 7.3.25 *Let $G = \langle V, E_d, I, P, N\rangle$ be a good class-property network with $I = \{U\}$ and classes $C \subseteq V$. Let Φ be the translation of G as defined by Brewka. Let Φ_{circ} be the result of circumscribing* Exceptional *in Φ with other predicates varied. Let $\Delta_G = \langle D, W\rangle$ be the translation of G as given in Definition 7.2.1 and let $\mathcal{P}_G$ be the priorities for Δ_G as given in Definition 7.2.6. Let $\Delta = \langle I(D), W \cup I_s(D)\rangle$ and $\mathcal{P} = \{\langle{:}p \rightarrow \beta/p \rightarrow \beta, {:}p' \rightarrow \beta'/p' \rightarrow \beta'\rangle | p{:}\beta/\beta \mathcal{P}_G p'{:}\beta'/\beta'\}$ (I and I_s are defined in Section 7.2.3.) Let M be a* *model of Φ_{circ}. Then there is a $\mathcal{P}$-preferred extension E of Δ such that for all $p \in V \backslash C$, if $p \in E$ then $M \models$ $\textsc{Holds}(p, U, T)$, and if $\neg p \in E$ then $M \models$ $\textsc{Holds}(p, U, F)$.*

Proof: Because G is good and therefore acyclic, by Lemma 7.2.7 $\mathcal{P}_G$ is a strict partial order. Assume that there is model M such that $M \models \Phi_{circ}$. For simplicity and without loss of generality we assume that constant symbols are interpreted as themselves in M. In the following we denote the propositional variable occurring in a literal β by b. Let M' be the model given in Lemma 7.3.23. Define $\mathcal{T}_1$ as the transitive closure of the relation $\mathcal{T}_0$ defined as $:p \rightarrow \beta/p \rightarrow \beta \mathcal{T}_0 :p' \rightarrow \beta'/p' \rightarrow \beta'$ whenever $:p \rightarrow \beta/p \rightarrow \beta \mathcal{P}_G :p' \rightarrow \beta'/p' \rightarrow \beta'$ or $\beta = \overline{\beta'}$ and $:p \rightarrow \beta/p \rightarrow \beta$ and $:p' \rightarrow \beta'/p' \rightarrow \beta'$ are $\mathcal{P}_G$-minimal elements in $\Lambda = \{:p \rightarrow \beta''/p \rightarrow \beta'' \in I(D) | \Phi \models \textsc{Is}(U, p)\}$ and $M \not\models \textsc{Exceptional}(U, p, b)$ and $M \models \textsc{Exceptional}(U, p', b')$. By Lemma 7.3.20, $\mathcal{T}_1$ is asymmetric and therefore a strict partial order.

Next we construct a preferred subtheory S of $\langle \{\phi | :\phi/\phi \in I(D)\}, W \cup I_s(D), \{\langle \phi, \psi \rangle | :\phi/\phi \mathcal{P} :\psi/\psi\} \rangle$. Let $\mathcal{T}$ be any strict total order on $I(D)$ such that $\mathcal{T}_1 \subseteq \mathcal{T}$. Let $:\phi_1/\phi_1, \ldots, :\phi_n/\phi_n$ be the ordering $\mathcal{T}$ on $I(D)$. Let $\phi_1, \ldots, \phi_n$ be the strict total order used in constructing $S_0, \ldots, S_n$ in Definition 6.2.2.

Induction hypothesis:

1. For all $p \in C$, $p \in Cn(S_i)$ if and only if $\Phi \models \textsc{Is}(U, p)$.
2. For all $j \in \{1, \ldots, i\}$, if $:p_j \rightarrow \beta_j/p_j \rightarrow \beta_j$ is $\mathcal{P}$-minimal in Λ, then $p_j \rightarrow \beta_j \in S_i$ if and only if $M \not\models \textsc{Exceptional}(U, p_j, b_j)$.
3. If $S_i \models p$ and $p \in V \backslash C$, then $M \models \textsc{Holds}(p, U, T)$. If $S_i \models \neg p$ and $p \in V \backslash C$, then $M \models \textsc{Holds}(p, U, F)$.

Base case $i = 0$:

1. By Lemma 7.3.24 for all $p \in C$, $p \in Cn(S_0)$ if and only if $\Phi \models \textsc{Is}(U, p)$.
2. Immediate because there are no $j \in \{1, \ldots, 0\}$.
3. Because G is good, the union of W and the consequents of the implications in $I_s(D)$ are consistent by Lemma 7.2.3. Therefore $S_0 = W \cup I_s(D)$ is consistent. Assume $S_0 \models p$ and $p \in V \backslash C$. If $p \in W$, then by definition $\langle U, p \rangle \in P$. By definition $\textsc{Holds}(p, U, T) \in \Phi$. Hence $M \models \textsc{Holds}(p, U, T)$. Because there are no occurrences of $q \in V \backslash C$ in $I_s(D)$ and S_0 is consistent, there is no q such that $q \notin W$ and $S_0 \models q$. Assume $S_0 \models \neg p$ and $p \in V \backslash C$. If $\neg p \in W$, then by definition $\langle U, p \rangle \in N$. By definition $\textsc{Holds}(p, U, F) \in \Phi$. Hence $M \models \textsc{Holds}(p, U, F)$. Because

there are no occurrences of $q \in V \backslash C$ in $I_s(D)$ and S_0 is consistent, there is no q such that $\neg q \notin W$ and $S_0 \models \neg q$.

Inductive case $i \geq 1$:

1. Assume $p \in Cn(S_i) \cap C$. If $S_{i-1} \cup \{p_i \rightarrow \beta_i\}$ is inconsistent and hence $S_i = S_{i-1}$, then the result is immediate by induction hypothesis. If $S_i = S_{i-1} \cup \{p_i \rightarrow \beta_i\}$, then by Lemma A.2.3 β_i is the only literal that may be in $Cn(S_i) \backslash Cn(S_{i-1})$. Because $b_i \in V \backslash C$, $p \neq b_i$. Hence $p \in Cn(S_{i-1})$. Hence by induction hypothesis $\Phi \models \textsc{Is}(U, p)$.

   Assume $\Phi \models \textsc{Is}(U, p)$. By induction hypothesis $p \in Cn(S_{i-1})$. Because $S_{i-1} \subseteq S_i$, $p \in Cn(S_i)$ by the monotonicity of Cn.

2. For $j \in \{1, \ldots, i-1\}$ the result is immediate by induction hypothesis. So let $:p_i \rightarrow \beta_i / p_i \rightarrow \beta_i$ be $\mathcal{P}$-minimal in Λ. Assume $p_i \rightarrow \beta_i \notin S_i$. Hence $S_i = S_{i-1}$ and $\{p_i \rightarrow \beta_i\} \cup S_{i-1}$ is inconsistent and $S_{i-1} \models \neg(p_i \rightarrow \beta_i)$ and $S_{i-1} \models p_i \wedge \neg\beta_i$. Hence by items 1 and 3 of the induction hypothesis $M \models \textsc{Is}(U, p_i)$ and $M \models \textsc{Holds}(b_i, U, T)$ if $\beta_i = \neg b_i$ and $M \models \textsc{Holds}(b_i, U, F)$ if $\beta_i = b_i$. Because $:p_i \rightarrow \beta_i / p_i \rightarrow \beta_i \in I(D)$, $\langle p_i, b_i, 0\rangle \in E_d$ if $\beta_i = b_i$ and $\langle p_i, b_i, 1\rangle \in E_d$ if $\beta_i = \neg b_i$. Therefore $\textsc{Has-slot}(p_i, b_i, T) \in \Phi$ if $\beta_i = b_i$ and $\textsc{Has-slot}(p_i, b_i, F) \in \Phi$ if $\beta_i = \neg b_i$. Because $\forall x \forall y \forall v(\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \rightarrow \textsc{Holds}(y, U, v)) \in \Phi$ and $M \models \Phi$, finally $M \models \textsc{Exceptional}(U, p_i, b_i)$.

   Assume that $M \models \textsc{Exceptional}(U, p_i, b_i)$. In the following we assume that $\beta_i = b_i$; the proof for the case $\beta_i = \neg b_i$ is almost identical.

   Lemma A $\Phi \models \textsc{Is}(U, p_i) \wedge \textsc{Has-slot}(p_i, b_i)$.
   Proof: Assume the opposite, that is, $\Phi \not\models \textsc{Is}(U, p_i) \wedge \textsc{Has-slot}(p_i, b_i, T)$. We show that $M''(\textsc{Exceptional}) \subset M'(\textsc{Exceptional})$ for a model M'' such that $M'' \models \Phi$. Let M'' be a model such that $M''(\textsc{Exceptional}) = M'(\textsc{Exceptional}) \backslash \{\langle U, p_i, b_i\rangle\}$ and $M''(P) = M'(P)$ for all other predicate symbols P. There are two formulae in Φ with occurrences of $\textsc{Exceptional}$.

   Consider $\phi = \forall x \forall y \forall v(\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \rightarrow \textsc{Holds}(y, U, v)) \in \Phi_b$. Assume $M'' \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y)$ for some x, y and v. By definition of M'' and M', $\Phi \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x)$. It is not possible that $x = p_i$ and $y = b_i$ because by assumption $\Phi \not\models \textsc{Has-slot}(p_i, b_i, T) \wedge \textsc{Is}(U, p_i)$. Therefore $M \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y)$ and because $M \models \phi$, $M \models \textsc{Holds}(y, U, v)$. By definition $M''(\textsc{Holds}) = M(\textsc{Holds})$ and hence $M'' \models \textsc{Holds}(y, U, v)$.

   Consider $\phi = \forall x_1 \forall x_2 \forall v_1 \forall v_2 \forall z(\textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2 \rightarrow \textsc{Exceptional}(U, x_2, z)) \in \Phi_b$.

Assume $M'' \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$ for some x_1, x_2, v_1, v_2, z. By definition of M'' and M' therefore $\Phi \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$ and $M \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$. Because $\forall x \forall y(\textsc{Specializes}(x, y) \rightarrow (\textsc{Is}(U, x) \rightarrow \textsc{Is}(U, y))) \in \Phi$, $\Phi \models \textsc{Is}(U, x_2)$. Because by assumption $\Phi \not\models \textsc{Is}(U, p_i) \wedge \textsc{Has-slot}(p_i, b_i, T)$, $x_2 \neq p_i$ or $z \neq b_i$. Because $\Phi \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$ and $M \models \Phi$, $M \models \textsc{Exceptional}(U, x_2, z)$ and $\langle U, x_2, z\rangle \in M(\textsc{Exceptional})$. Because $x_2 \neq p_i$ or $z \neq b_i$, $\langle U, x_2, z\rangle \in M''(\textsc{Exceptional})$, and therefore $M'' \models \phi$. This contradicts the minimality of M' with respect to $\textsc{Exceptional}$, and hence $\Phi \models \textsc{Is}(U, p_i) \wedge \textsc{Has-slot}(p_i, b_i, T)$. □

We analyze by cases.

(a) Consider the case where for some $q \in V$, $:q \rightarrow \overline{\beta_i}/q \rightarrow \overline{\beta_i}$ is $\mathcal{P}$-minimal in Λ and $M \not\models \textsc{Exceptional}(U, q, b)$. By Lemma **A** and induction hypothesis $p_i \in Cn(S_{i-1})$ and $q \in Cn(S_{i-1})$. For members $:p \rightarrow \beta'/p \rightarrow \beta'$ of Λ such that $M \models \neg\textsc{Exceptional}(U, p, b')$ clearly $M \models \textsc{Holds}(b', U, T)$ if $\beta' = b'$ and $M \models \textsc{Holds}(b', U, F)$ if $\beta' = \neg b'$. Hence $M \models \textsc{Holds}(b_i, U, F)$. If $:p_i \rightarrow \beta_i/p_i \rightarrow \beta_i$ is $\mathcal{P}$-minimal in Λ, then by definition $:q \rightarrow \overline{\beta_i}/q \rightarrow \overline{\beta_i}\mathcal{P}:p_i \rightarrow \beta_i/p_i \rightarrow \beta_i$. Hence $q \rightarrow \overline{\beta_i} = p_j \rightarrow \beta_j$ for some $j < i$. Because $M \not\models \textsc{Exceptional}(U, q, b)$ by induction hypothesis, $q \rightarrow \overline{\beta_i} \in S_{i-1}$. Because $S_{i-1} \models q$, $S_{i-1} \models \overline{\beta_i}$. Therefore $S_{i-1} \cup \{p_i \rightarrow \beta_i\}$ is inconsistent and $p_i \rightarrow \beta_i \notin S_i$.

(b) Consider the case where $\Phi \models \textsc{Holds}(b_i, U, F)$. Now $\langle U, b_i\rangle \in N$, and $\neg b_i \in W \subseteq S_{i-1}$. By induction hypothesis $p_i \in Cn(S_{i-1})$. Therefore $S_{i-1} \cup \{p_i \rightarrow \beta_i\}$ is inconsistent and $p_i \rightarrow \beta_i \notin S_i$.

(c) We show that the preconditions of (a) and (b) cannot be false. This is by deriving a contradiction with the fact $M' \models \Phi_{circ}$. So assume that the preconditions of (a) and (b) are false. We construct a model M'' such that $M'' \models \Phi$ and $M''(\textsc{Exceptional}) \subset M'(\textsc{Exceptional})$. Let the universe of M'' be $V \cup \{U, T, F\}$ and constant symbols are interpreted as themselves. Let

$$\begin{aligned} M''(\textsc{Exceptional}) &= M'(\textsc{Exceptional}) \backslash \{\langle U, p_i, b_i\rangle\}, \\ M''(\textsc{Holds}) &= M'(\textsc{Holds}) \backslash \{\langle b_i, U, F\rangle\} \cup \{\langle b_i, U, T\rangle\}, \text{ and} \\ M''(P) &= M'(P) \text{ for all other predicate symbols } P. \end{aligned}$$

We show that $M'' \models \Phi$. Formulae in Φ with no occurrences of $\textsc{Exceptional}$ or $\textsc{Holds}$ are true in M'' because $M' \models \Phi$ and M' and M'' differ only with respect to $\textsc{Exceptional}$ and $\textsc{Holds}$.

Assume $\textsc{Holds}(x, U, v) \in \Phi$. Because by assumption $\Phi \not\models \textsc{Holds}(b_i, U, F)$, it is not possible that $x = b_i$ and $v = F$. If $x \neq b_i$ or $v \neq F$, then $M'' \models \textsc{Holds}(x, U, v)$ because $M' \models \textsc{Holds}(x, U, v)$ and $M'(\textsc{Holds})\backslash\{\langle b_i, U, F\rangle\} \subseteq M''(\textsc{Holds})\backslash\{\langle b_i, U, F\rangle\}$.

Consider $\forall x \forall y \forall v(\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \rightarrow \textsc{Holds}(y, U, v)) \in \Phi_b$. Assume $M'' \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y)$ for some x, y and v. By definition $\Phi \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x)$. If $y = b_i$ and $v = T$, then by definition $M'' \models \textsc{Holds}(b_i, U, T)$. If $y \neq b_i$, then $M' \models \neg\textsc{Exceptional}(U, x, y)$ and because $M' \models \Phi$, $M' \models \textsc{Holds}(y, U, v)$. Because $\langle y, U, v\rangle \in M'(\textsc{Holds})$ implies $\langle y, U, v\rangle \in M''$ for $y \neq b_i$, $M'' \models \textsc{Holds}(y, U, v)$. The remaining case is $y = b_i$ and $v \neq T$. Now $v = F$. Because $\Phi \models \textsc{Has-slot}(q, b_i, F)$ for some $x = q \in C$, $\langle q, b_i, 1\rangle \in E_d$ and $:q \rightarrow \overline{\beta_i}/q \rightarrow \overline{\beta_i} \in I(D)$. By assumption $:q \rightarrow \overline{\beta_i}/q \rightarrow \overline{\beta_i}$ is not $\mathcal{P}$-minimal in Λ. Because $:q \rightarrow \overline{\beta_i}/q \rightarrow \overline{\beta_i} \in \Lambda$, there is $\delta \in I(D)$ that is $\mathcal{P}$-minimal in Λ. By assumption $\delta = :p \rightarrow \beta_i/p \rightarrow \beta_i$ for some $p \in V$. Now $\Phi \models \textsc{Is}(U, p) \wedge \textsc{Has-slot}(p, b_i, T) \wedge \textsc{Specializes}(p, q) \wedge \textsc{Has-slot}(q, b_i, F) \wedge T \neq F$ and $M'' \models \neg\textsc{Exceptional}(U, q, b_i)$. This contradicts $M' \models \forall x_1 \forall x_2 \forall v_1 \forall v_2 \forall z(\textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2 \rightarrow \textsc{Exceptional}(U, x_2, z))$. Hence it is not possible that $y = b_i$ and $v \neq T$.

Consider $\forall x_1 \forall x_2 \forall v_1 \forall v_2 \forall z(\textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2 \rightarrow \textsc{Exceptional}(U, x_2, z)) \in \Phi_b$. Assume $M'' \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2 \rightarrow \textsc{Exceptional}(U, x_2, z)$ for some x_1, x_2, v_1, v_2 and z. By definition $\Phi \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$. Hence $M \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$ and because $M \models \Phi_b$, $M \models \textsc{Exceptional}(U, x_2, z)$. If $x_2 \neq p_i$ or $z \neq b_i$ then $M'' \models \textsc{Exceptional}(U, x_2, z)$ because $M''(\textsc{Exceptional}) = M(\textsc{Exceptional}) \backslash\{\langle U, p_i, b_i\rangle\}$. So assume that $x_2 = p_i$ and $z = b_i$. Because $\Phi \models \textsc{Specializes}(x_1, p_i)$, there is a path from x_1 to p_i. This contradicts the assumption that there is no path from node q to p_i if $\textsc{Has-slot}(q, b_i, F) \in \Phi$ and $\Phi \models \textsc{Is}(U, q)$. Hence it is not possible that $x_2 = p_i$ and $z = b_i$.

3. Assume $S_{i-1} \cup \{p_i \rightarrow \beta_i\}$ is inconsistent. Hence $S_i = S_{i-1}$, and the result is immediate by induction hypothesis. Assume $S_{i-1} \cup \{p_i \rightarrow \beta_i\}$ is consistent. Hence $S_i = S_{i-1} \cup \{p_i \rightarrow \beta_i\}$. By Lemma A.2.3 $(Cn(S_i)\backslash Cn(S_{i-1})) \cap (V\backslash C) \subseteq \{\beta_i\}$. If $\beta_i \in Cn(S_i)\backslash Cn(S_{i-1})$, then $p_i \in Cn(S_{i-1})$ by Lemma A.2.3. By induction hypothesis $M \models \textsc{Is}(U, p_i)$.

Because $p_i \rightarrow \beta_i \in I(D)$, $\langle p_i, b_i, 0\rangle \in E_d$. Hence $\textsc{Has-slot}(p_i, b_i, T) \in \Phi$. If $:p_i \rightarrow \beta_i/p_i \rightarrow \beta_i$ is $\mathcal{P}$-minimal in Λ, then by item 2 above $M \not\models \textsc{Exceptional}(U, p_i, b_i)$. Because $M \models \Phi$ and $M \models \textsc{Is}(U, p_i)$ and $M \models \forall x \forall y \forall v(\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \rightarrow \textsc{Holds}(y, U, v))$, $M \models \textsc{Holds}(b_i, U, T)$. If $:p_i \rightarrow \beta_i/p_i \rightarrow \beta_i$ is not $\mathcal{P}$-minimal in Λ then there is $:p_j \rightarrow \beta_j/p_j \rightarrow \beta_j\mathcal{P}{:}p_i \rightarrow \beta_i/p_i \rightarrow \beta_i$ in Λ and by induction hypothesis $p_j \rightarrow \beta_j \in S_i$ if and only if $M \not\models \textsc{Exceptional}(U, p_j, b_j)$. By induction hypothesis $p_j \in Cn(S_i)$. If $\beta_j = \overline{\beta_i}$, then $p_j \rightarrow \beta_j \notin S_i$. Hence $S_{i-1} \models \beta_i$. By induction hypothesis $M' \models \textsc{Holds}(b_i, U, T)$. If $\beta_j = \beta_i$, then $p_j \rightarrow \beta_j \in S_i$ and $S_j \models \beta_j$. By induction hypothesis $M' \models \textsc{Holds}(b_i, U, T)$.

By definition $S = S_n$ is a preferred subtheory of $\langle\{\phi | {:}\phi/\phi \in I(D)\}, W \cup I_s(D), \{\langle\phi, \psi\rangle | {:}\phi/\phi\mathcal{P}{:}\psi/\psi\}\rangle$. By Lemma 6.2.11 $E = Cn(S)$ is a $\mathcal{P}$-preferred extension of $\Delta = \langle I(D), W \cup I_s(D)\rangle$. By item 3 of the induction hypothesis, for all $p \in V \backslash C$, if $p \notin E$ then $M \not\models \textsc{Holds}(p, U, T)$, and if $\neg p \notin E$ then $M \not\models \textsc{Holds}(p, U, F)$. □

Lemma 7.3.26 *Let $G = \langle V, E_d, I, P, N\rangle$ be a good class-property network with $I = \{U\}$ and classes $C \subseteq V$. Let Φ be the translation of G as defined by Brewka. Let Φ_{circ} be the result of circumscribing* $\textsc{Exceptional}$ *in Φ with other predicates varied. Let $\Delta_G = \langle D, W\rangle$ be the translation of G as given in Definition 7.2.1 and let $\mathcal{P}_G$ be the priorities for Δ_G as given in Definition 7.2.6. Let $\Delta = \langle I(D), W \cup I_s(D)\rangle$ and $\mathcal{P} = \{\langle{:}p \rightarrow \beta/p \rightarrow \beta, {:}p' \rightarrow \beta'/p' \rightarrow \beta'\rangle | p{:}\beta/\beta\mathcal{P}_G p'{:}\beta'/\beta'\}$. If E is a $\mathcal{P}$-preferred extension of Δ, then there is a model M such that $M \models \Phi_{circ}$ and for all $p \in V \backslash C$, if $p \notin E$ then $M \not\models$* $\textsc{Holds}(p, U, T)$.

Proof: Because G is acyclic, by Lemma 7.2.7 $\mathcal{P}_G$ is a strict partial order. In the following we denote the propositional variable occurring in a literal β by b. Assume there is a $\mathcal{P}$-preferred extension E of Δ. Because G is good, the union of W and the consequents of the implications in $I_s(D)$ are consistent by Lemma 7.2.3, and hence $W \cup I_s(D)$ is consistent. By Theorem A.1.1 E is consistent. We construct a model M such that $M \models \Phi_{circ}$. The universe of M is $V \cup \{U, T, F\}$, the constant symbols $V \cup \{U, T, F\}$ are interpreted as themselves, and the predicate symbols are interpreted as follows.

$$
\begin{aligned}
M(\textsc{Specializes}) &= \{\langle x, y\rangle | \Phi \models \textsc{Specializes}(x, y)\}\\
M(\textsc{Has-slot}) &= \{\langle x, y, v\rangle | \Phi \models \textsc{Has-slot}(x, y, v)\}\\
M(\textsc{Exceptional}) &= \left\{\langle U, p, b\rangle \,\middle|\, \frac{:p \rightarrow \beta}{p \rightarrow \beta} \in I(D), \mathrm{def}(\frac{:p \rightarrow \beta}{p \rightarrow \beta}, E), \beta \in \{b, \neg b\}\right\}\\
&\quad \cup \{\langle U, p, b\rangle | \Phi \models \textsc{Exceptional}(U, p, b)\}\\
M(\textsc{Is}) &= \{\langle U, x\rangle | \Phi \models \textsc{Is}(U, x)\}
\end{aligned}
$$

$$M(\textsc{Holds}) \quad = \quad \{\langle x, U, T\rangle | x \in E\} \cup \{\langle x, U, F\rangle | \neg x \in E\}$$

We show that $M \models \Phi_c \cup \Phi_i \cup \Phi_b \cup \Phi_u$. Assume that $\textsc{Specializes}(x, y) \in \Phi$. Then $\Phi \models \textsc{Specializes}(x, y)$ and by definition $\langle x, y\rangle \in M(\textsc{Specializes})$. Hence $M \models \textsc{Specializes}(x, y)$. Assume that $\textsc{Is}(x, y) \in \Phi$. Then $\Phi \models \textsc{Is}(x, y)$ and by definition $\langle x, y\rangle \in M(\textsc{Is})$. Hence $M \models \textsc{Is}(x, y)$. Assume that $\textsc{Has-slot}(x, y, v) \in \Phi$. Then $\Phi \models \textsc{Has-slot}(x, y, v)$ and by definition $\langle x, y, v\rangle \in M(\textsc{Has-slot})$. Hence $M \models \textsc{Has-slot}(x, y, v)$. Assume $\textsc{Holds}(p, U, T) \in \Phi_i$. By definition $\langle p, U, T\rangle \in P$. Hence by definition $p \in W$. Because $W \subseteq E$, $p \in E$. By definition $\langle p, U, T\rangle \in M(\textsc{Holds})$. Hence $M \models \textsc{Holds}(p, U, T)$. Assume $\textsc{Holds}(p, U, F) \in \Phi_i$. By definition $\langle p, U, F\rangle \in N$. Hence by definition $\neg p \in W$. Because $W \subseteq E$, $\neg p \in E$. By definition $\langle p, U, F\rangle \in M(\textsc{Holds})$. Hence $M \models \textsc{Holds}(p, U, F)$.

Consider $\forall x \forall y(\textsc{Specializes}(x, y) \to (\textsc{Is}(U, x) \to \textsc{Is}(U, y))) \in \Phi$. Assume that $M \models \textsc{Specializes}(x, y) \wedge \textsc{Is}(U, x)$ for some x and y. Hence by definition $\Phi \models \textsc{Specializes}(x, y) \wedge \textsc{Is}(U, x)$. Clearly $\Phi \models \textsc{Is}(U, y)$. Hence by definition $M \models \textsc{Is}(U, y)$.

Consider $\forall x \forall y \forall z(\textsc{Specializes}(x, y) \wedge \textsc{Specializes}(y, z) \to \textsc{Specializes}(x, z)) \in \Phi$. Assume that $M \models \textsc{Specializes}(x, y) \wedge \textsc{Specializes}(y, z)$ for some x, y and z. Hence by definition $\Phi \models \textsc{Specializes}(x, y) \wedge \textsc{Specializes}(y, z)$. Clearly $\Phi \models \textsc{Specializes}(x, z)$. Hence by definition $M \models \textsc{Specializes}(x, z)$.

Consider $\forall x \forall y \forall v(\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \to \textsc{Holds}(y, U, v)) \in \Phi_b$. Assume $M \models \textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y)$ for some symbols x, y and $v \in \{T, F\}$. Hence $\langle x, y, p\rangle \in E_d$, where $p = 0$ if $v = T$ and $p = 1$ otherwise, and by definition $\Phi \models \textsc{Is}(U, x)$ and hence by Lemma 7.3.24 $x \in E$, and not def$(:x \to y/x \to y, E)$ and not def$(:x \to \neg y/x \to \neg y, E)$. If $v = T$ then by definition $:x \to y/x \to y \in I(D)$ and hence $y \in E$ because not def$(:x \to y/x \to y, E)$. Then by definition $\langle y, U, T\rangle \in M(\textsc{Holds})$ and $M \models \textsc{Holds}(y, U, T)$. If $v = F$ then by definition $:x \to \neg y/x \to \neg y \in I(D)$ and hence $\neg y \in E$ because not def$(:x \to \neg y/x \to \neg y, E)$. Then by definition $\langle y, U, F\rangle \in M(\textsc{Holds})$ and $M \models \textsc{Holds}(y, U, F)$.

Consider $\forall x_1 \forall x_2 \forall v_1 \forall v_2 \forall z(\textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2 \to \textsc{Exceptional}(U, x_2, z)) \in \Phi_b$. Assume $M \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$ for some symbols x_1, x_2, z, v_1 and v_2. Hence $\Phi \models \textsc{Is}(U, x_1) \wedge \textsc{Has-slot}(x_1, z, v_1) \wedge \textsc{Specializes}(x_1, x_2) \wedge \textsc{Has-slot}(x_2, z, v_2) \wedge v_1 \neq v_2$. Hence by definition $M \models \textsc{Exceptional}(U, x_2, z)$.

Consider $\forall x \forall v_1 \forall v_2(\textsc{Holds}(x, U, v_1) \wedge \textsc{Holds}(x, U, v_2) \to v_1 = v_2) \in \Phi_b$. Assume $M \models \textsc{Holds}(x, U, v_1) \wedge \textsc{Holds}(x, U, v_2)$ for some symbols x, v_1 and v_2. By definition $\{v_1, v_2\} \subseteq \{T, F\}$. Assume $v_1 \neq v_2$. Then by definition $x \in E$

and $\neg x \in E$, which contradicts the consistency of E. Hence $v_1 = v_2$ and $M \models v_1 = v_2$.

It remains to show that the second conjunct of Φ_{circ} is true in M. This formula says that there are no models M' such that $M' \models \Phi$ and $M'(\textsc{Exceptional}) \subset M(\textsc{Exceptional})$. Assume there is such a model M'. Hence there is $\langle U, p, b\rangle \in M(\textsc{Exceptional}) \backslash M'(\textsc{Exceptional})$. Now $C = W \cup I_s(D) \cup A$, where $A = \{p \to \beta\} \cup \{p' \to \beta' | {:}p' \to \beta'/p' \to \beta' \in I(D), \langle U, p', b'\rangle \notin M'(\textsc{Exceptional})\}$, is consistent because of the following. Construct model M'' such that $M'' \models p$ if and only if $p \in C$ and $M' \models \textsc{Is}(U,p)$ or $p \in V \backslash C$ and $M' \models \textsc{Holds}(p, U, T)$. Assume $\neg p \in W$. Hence $\langle U, p\rangle \in N$, $\textsc{Holds}(p, U, F) \in \Phi$, $M' \models \textsc{Holds}(p, U, F)$, $M' \not\models \textsc{Holds}(p, U, T)$, $M'' \not\models p$. Assume $p \in W$. Hence $\langle U, p\rangle \in P$, $\textsc{Holds}(p, U, T) \in \Phi$ if $p \in V \backslash C$ and $\textsc{Is}(U, p) \in \Phi$ if $p \in C$, $M' \models \textsc{Holds}(p, U, T)$ if $p \in V \backslash C$ and $M' \models \textsc{Is}(U, p)$ if $p \in C$, $M'' \models p$. Assume $q \to q' \in I_s(D)$. Hence $\langle q, q', 2\rangle \in E_d$ and $\textsc{Specializes}(q, q') \in \Phi$. Assume $M'' \models q$. Hence $M' \models \textsc{Is}(U, q)$. Because $M' \models \textsc{Specializes}(q, q')$ and $M' \models \forall x \forall y (\textsc{Specializes}(x, y) \to (\textsc{Is}(U, x) \to \textsc{Is}(U, y)))$, $M' \models \textsc{Is}(U, q')$, and hence $M'' \models q'$. Assume $p \to \beta \in A$. Hence $M' \not\models \textsc{Exceptional}(U, p, b)$ and $M' \models \textsc{Has-slot}(p, b, T)$. Assume $M'' \models p$. Hence $M' \models \textsc{Is}(U, p)$. Because $M' \models \forall x \forall y \forall v (\textsc{Has-slot}(x, y, v) \wedge \textsc{Is}(U, x) \wedge \neg\textsc{Exceptional}(U, x, y) \to \textsc{Holds}(y, U, v))$, $M' \models \textsc{Holds}(p, b, T)$ and $M'' \models \beta$.

Let $\Lambda = \{{:}p' \to \beta'/p' \to \beta' \in I(D) | p' \in E\}$ and $\Gamma = \{\phi | {:}\phi/\phi \in \Lambda\}$.

Lemma A For all $\mathcal{P}$-minimal elements ${:}p' \to \beta'/p' \to \beta' \in \Lambda$, $\langle U, p', b'\rangle \in M(\textsc{Exceptional})$ if and only if $p' \to \beta' \notin E$.
Proof: The *if* part is by definition of M. Assume $\langle U, p', b'\rangle \in M(\textsc{Exceptional})$. If $\Phi \not\models \textsc{Exceptional}(U, p', b')$, then by definition $p' \to \beta' \notin E$. Otherwise $\Phi \models \textsc{Exceptional}(U, p', b')$. Because ${:}p' \to \beta'/p' \to \beta'$ is $\mathcal{P}$-minimal in Λ, there is no ${:}p'' \to \beta''/p'' \to \beta'' \in I(D)$ such that $p'' \in E$ and there is a path from p'' to p'. By Lemma 7.3.24 $W \cup I_s(D) \models p''$ iff $\Phi \models \textsc{Is}(U, p'')$ for all propositional variables p'', and by Lemma A.2.2 $E \models p''$ iff $\Phi \models \textsc{Is}(U, p'')$. Hence there is no ${:}p'' \to \beta''/p'' \to \beta'' \in I(D)$ such that $\Phi \models \textsc{Is}(U, p'') \wedge \textsc{Specializes}(p'', p)$. Because $\Phi \models \textsc{Exceptional}(U, p', b')$, it must be the case that $\Phi \models \textsc{Has-slot}(p', b', T) \wedge \textsc{Is}(U, p') \wedge \textsc{Holds}(b', U, F)$. Now $\langle U, p'\rangle \in P$, and $\langle U, b'\rangle \in N$ if $\beta' = b'$ and $\langle U, b'\rangle \in P$ if $\beta' = \neg b'$. Hence $\{p', \overline{\beta'}\} \subseteq W$. Because $\neg(p' \to \beta') \in Cn(W) \subseteq E$ and E is consistent, $p' \to \beta' \notin E$. $\square$

Because E is a $\mathcal{P}$-preferred extension of a prerequisite-free normal default theory $\langle I(D), W \cup I_s(D)\rangle$, by Lemma 6.2.12 there is a preferred subtheory S of $\langle I(D), W \cup I_s(D), \mathcal{B}\rangle$ such that $Cn(S) = E$ and $\mathcal{B} = \{\langle \phi, \phi'\rangle | \langle {:}\phi/\phi, {:}\phi'/\phi'\rangle \in \mathcal{P}\}$.

Lemma B Let $R = \{p' \to \beta' | \beta' \in \{\beta, \overline{\beta}\}, {:}p' \to \beta'/p' \to \beta'$ is $\mathcal{P}$-minimal in $\Lambda\}$. Then $W \cup I_s(D) \cup (R \cap E) \cup \{p \to \beta\}$ is inconsistent.

Proof: Let $\mathcal{T}$ be a total order such that $\mathcal{B} \subseteq \mathcal{T}$ and by using which S is obtained. Let $p_i \rightarrow \beta_i$ be the $\mathcal{T}$-least element in Γ. Because $p_i \rightarrow \beta_i \in \Gamma$, $p_i \in S_0 \subseteq S_i$. If $S_{i-1} \models \neg\beta_i$, then $\beta_i = \beta$. If $S_{i-1} \not\models \neg\beta_i$, then $S_i = S_{i-1} \cup \{p_i \rightarrow \beta_i\}$ and $\beta_i = \overline{\beta}$. In both cases $S_i \cup \{p \rightarrow \beta\}$ is inconsistent. Because $S_i \subseteq W \cup I_s(D) \cup (R \cap E) \cup \{p \rightarrow \beta\}$, $W \cup I_s(D) \cup (R \cap E) \cup \{p \rightarrow \beta\}$ is inconsistent. □

By Lemma **A** $N = W \cup I_s(D) \cup \{p' \rightarrow \beta' | {:}p' \rightarrow \beta'/p' \rightarrow \beta'$ is $\mathcal{P}$-minimal in $\Lambda, \beta' \in \{\beta, \overline{\beta}\}, \langle U, p', b'\rangle \notin M(\textsc{Exceptional})\}$ is a subset of E. The set $\{p \rightarrow \beta\} \cup N$ is consistent because it is a subset of the consistent set C. Because $M' \models \Phi$ and $M' \not\models \textsc{Exceptional}(U, p, b)$, it is not the case that $\Phi \models \textsc{Exceptional}(U, p, b)$. Hence def$({:}p \rightarrow \beta/p \rightarrow \beta, E)$ by definition of M, and $S \models \neg(p \rightarrow \beta)$. Now by Lemma **B** $\{p \rightarrow \beta\} \cup N$ is inconsistent. Therefore the assumption that there is a model M' of Φ such that $M'(\textsc{Exceptional}) \subset M(\textsc{Exceptional})$ is false, and $M \models \Phi_{circ}$. □

Theorem 7.3.27 *Let $G = \langle V, E_d, I, P, N\rangle$ be a good class-property network with $I = \{U\}$ and classes $C \subseteq V$. Let Φ be the translation of G as defined by Brewka. Let Φ_{circ} be the result of circumscribing* Exceptional *in Φ with other predicates varied. Let $\Delta_G = \langle D, W\rangle$ be the translation of G as given in Definition 7.2.1 and let $\mathcal{P}_G$ be the priorities for Δ_G as given in Definition 7.2.6. Let $\Delta = \langle I(D), W \cup I_s(D)\rangle$ and $\mathcal{P} = \{\langle {:}p \rightarrow \beta/p \rightarrow \beta, {:}p' \rightarrow \beta'/p' \rightarrow \beta'\rangle | p{:}\beta/\beta \mathcal{P}_G p'{:}\beta'/\beta'\}$. Then for all $p \in V \backslash C$, $\Phi_{circ} \models \textsc{Holds}(p, U, T)$ if and only if $\Delta \models_{\mathcal{P}} p$.*

Proof: Because G is good and hence acyclic, Lemmata 7.3.25 and 7.3.26 can be applied and by Lemma 7.2.7 $\mathcal{P}_G$ is a strict partial order. Assume $\Phi_{circ} \not\models \textsc{Holds}(p, U, T)$, that is, there is a model M such that $M \models \Phi_{circ}$ and $M \not\models \textsc{Holds}(p, U, T)$. By Lemma 7.3.25 there is a $\mathcal{P}$-preferred extension E of Δ such that $p \notin E$. Hence $\Delta \not\models_{\mathcal{P}} p$. Assume $\Delta \not\models_{\mathcal{P}} p$, that is, there is a $\mathcal{P}$-preferred extension E of Δ such that $p \notin E$. By Lemma 7.3.26 there is a model M such that $M \models \Phi_{circ}$ and $M \not\models \textsc{Holds}(p, U, T)$. Hence $\Phi_{circ} \not\models \textsc{Holds}(p, U, T)$. □

7.3.3 Conditional entailment

Geffner and Pearl formalize default theories T as pairs $\langle K, V\rangle$ where K is a *background context* and V is an *evidence set*. The context K is a pair $\langle L, D\rangle$ where L is a set of sentences and D is a set of defaults. We use the notation $\alpha \rightharpoonup \beta$ for defaults. The background context formalizes general rules concerning the domain of the knowledge base. The starting point for conditional entailment is p-entailment [Adams, 1975] or ϵ-entailment [Pearl, 1988].

These entailment relations can be justified probabilistically: they formalize probabilistically valid inference with infinitesimal probabilities [Adams, 1975; Pearl, 1988]. Alternatively, these entailment relations can be given semantics by using model preference [Kraus *et al.*, 1990]. These entailment relations are induced by the following rules.

$$\frac{p \rightharpoonup q \in D}{p \mid\!\sim_K q} \qquad \frac{V \mid\!\sim_K p \text{ and } V \mid\!\sim_K q}{V, p \mid\!\sim_K q} \qquad \frac{V, L \models p}{V \mid\!\sim_K p}$$

$$\frac{V \mid\!\sim_K p \text{ and } V, p \mid\!\sim_K q}{V \mid\!\sim_K q} \qquad \frac{V, p \mid\!\sim_K r \text{ and } V, q \mid\!\sim_K r}{V, p \vee q \mid\!\sim_K r}$$

Geffner and Pearl note that the entailment relations induced by the above rules are in general reasonable, but there is the unacceptable flaw that they are sensitive to irrelevant premises, that is, if $a \mid\!\sim b$ is derivable, $r, a \mid\!\sim b$ typically is not, even in cases where r is irrelevant to both a and b. The irrelevance problem is not present in entailment relations that are based on model minimization as in circumscription, or consistency-based approaches to nonmonotonic reasoning like autoepistemic logic or default logic. Geffner and Pearl set out to modify a minimal model semantics for defaults so that the properties expressed by the above rules are fulfilled. Viewing defaults $p \rightharpoonup q$ as implications $p \rightarrow q$ and by using minimal model semantics that prefers a model to another if it satisfies more defaults in the set-inclusion sense, Geffner and Pearl note that all the properties are satisfied except the one expressed by the first rule. They set the goal to define priorities on the defaults so that the first one would be satisfied as well.

Default theories used by Geffner and Pearl were presented in Section 6.5 along with a theorem connecting the minimal models semantics used in conditional entailment to prioritized default logic. We recall that default theories are of the form $T = \langle K, V\rangle$ with background $K = \langle L, D\rangle$ where $L \subseteq \mathcal{L}$ is a set of formulae and D is a set of defaults, and evidence $V \subseteq \mathcal{L}$. The formulae in the set L may have occurrences of assumptions in $\Delta_{\mathcal{L}}$.

Definition 7.3.28 ([Geffner and Pearl, 1992]) *A set of assumptions Δ is* in conflict with a default *$p \rightharpoonup \delta \in K$, if $L \cup \{p\} \cup \Delta \models \neg\delta$. A priority order $\prec$ over $\Delta_{\mathcal{L}}$ is* admissible *relative to a background context K if and only if every set Δ of assumptions in conflict with a default $p \rightharpoonup \delta$ in K contains an assumption δ' such that $\delta' \prec \delta$.*

Definition 7.3.29 ([Geffner and Pearl, 1992]) *A prioritized structure $\langle \mathcal{I}_{\mathcal{L}}, <, \Delta_{\mathcal{L}}, \prec\rangle$ is admissible with a background $K = \langle L, D\rangle$ if and only if the priority ordering $\prec$ is admissible with K.*

Definition 7.3.30 ([Geffner and Pearl, 1992]) *A formula ϕ is* conditionally entailed *by a default theory $T = \langle K, V\rangle$ if and only if ϕ is true in all the* *preferred models of T in every preferential structure admissible with K.*

Theorem 7.3.31 *A formula ϕ is conditionally entailed by $T = \langle K, V\rangle$, where $K = \langle L, D\rangle$, if and only if $\langle \{:\delta/\delta | \delta \in \Delta_{\mathcal{L}}\}, L \cup V\rangle \models^2_{\mathcal{P}} \phi$ for all strict partial orders $\mathcal{P} = \{\langle :\delta/\delta, :\delta'/\delta'\rangle | \delta \prec \delta'\}$ such that $\prec$ is admissible with K.*

Proof: Immediate by the above definitions and Theorem 6.5.8. □

A counterpart of Condition 7.1 discussed in Section 7.1.2 is Geffner and Pearl's [1992] motivation in introducing the notion of admissible prioritized structures. All admissible prioritized structures satisfy this condition. The reverse, however, does not hold because of a technicality in the definition of admissibility. Consider the following defaults.

$$q \wedge p \wedge \delta_0 \rightarrow r, p \wedge q \wedge \delta_1 \rightarrow r, p \wedge \delta_2 \rightarrow \neg r$$

Any ordering $\prec$ such that $\delta_2 \prec \delta_0$ satisfies Condition 7.1, but $\prec$ is not admissible if $\delta_2 \not\prec \delta_1$.

Delgrande's preference-based entailment relation

Delgrande [1994] proposes a system of nonmonotonic reasoning with specificity that strengthens conditional entailment of Geffner and Pearl [1992]. His system consists of two components. The first component is a logic that is used in determining a priority relation between defaults. The second component is a definition of minimal models similar to the one used in conditional entailment [Geffner and Pearl, 1992]. Delgrande claims that his system avoids the weakness of conclusions of conditional entailment. Delgrande gives several examples that show that his system is stronger than conditional entailment.

The logic used in ordering defaults is closely related to the preferential logics studied by Kraus, Lehmann and Magidor [1990]. The model theories of these logics are based on orderings of worlds with a propositional model or a set of propositional models associated with each world. A model is a triple $M = \langle W, E, P\rangle$ where W is a set of worlds, E is a reflexive, transitive and forward connected relation on W, and P is a mapping from propositional variables and worlds to truth values. Truth in a world $\models^M_w \phi$ for models M, worlds $w \in W$ and formulae ϕ is defined in the standard way. For defaults, $\models^M_w \alpha \rightharpoonup \beta$ if and only if $\models^M_{w'} \alpha \wedge \beta$ for some w' such that wEw' and $\models^M_{w''} \alpha \rightarrow \beta$ for all w'' such that wEw'', or $\models^M_{w'} \alpha$ for no wEw'.

Finally, defaults are ordered according to the ordering of the prerequisites of the defaults by $\prec$. The ordering $\prec$ is defined as $\alpha \prec \beta$ if and only if for all worlds w such that $\models^M_w \beta$ there is no w' such that wEw' and $\models^M_w \alpha$. This ordering is partly motivated by a remark Delgrande makes that $a \rightharpoonup b$ has to be given a higher priority than $b \rightharpoonup c$ so that with a and $\neg c$ the conclusion b would be obtained. However, Delgrande does not seem to be more generally aware of the of the second rule for priorities given in Section 7.1.3 as it is not difficult to find examples that suggest that the consequence relation Delgrande formalizes is most likely weaker than what he intends.

Example 7.3.1 Consider the defaults $T = \{a \rightharpoonup b, b \wedge c \rightharpoonup d\}$. To obtain the conclusion b given premises a, c and $\neg d$, it is necessary to order the two defaults. We show that Delgrande's definition does not do that because a is not given a precedence over $b \wedge c$. Specifically, there is no ordering between the formulae.

Consider the models $M_1 = \langle\{0, \ldots, 15\}, \mathcal{R}_1, P\rangle$ and $M_2 = \langle\{0, \ldots, 15\}, \mathcal{R}_2, P\rangle$ for a language that consists of propositional variables a, b, c, and d. The relations $\mathcal{R}_1$ and $\mathcal{R}_2$ are defined as $0000\mathcal{R}_1 0001\mathcal{R}_1 0010\mathcal{R}_1 0011\mathcal{R}_1 0100\mathcal{R}_1 0101\mathcal{R}_1 1000\mathcal{R}_1 1001\mathcal{R}_1 1010\mathcal{R}_1 1011\mathcal{R}_1 1100\mathcal{R}_1 1101\mathcal{R}_1 1110\mathcal{R}_1 1111\mathcal{R}_1 0110\mathcal{R}_1 0111$ and $0000\mathcal{R}_2 0001\mathcal{R}_2 0010\mathcal{R}_2 0011\mathcal{R}_2 0100\mathcal{R}_2 0101\mathcal{R}_2 0110\mathcal{R}_2 0111\mathcal{R}_2 1110\mathcal{R}_2 1111\mathcal{R}_2 1000\mathcal{R}_2 1001\mathcal{R}_2 1010\mathcal{R}_2 1011\mathcal{R}_2 1100\mathcal{R}_2 1101$, where the worlds (integers) are expressed as binary numbers. The function $P(n, p)$ maps a world n and a propositional variable $p \in \{a, b, c, d\}$ to true, if $p = a$ and the digit corresponding to 2^3 in the binary notation of n is 1, and it maps them to false otherwise. Similarly b corresponds to 2^2, c to 2^1, and d to 2^0. For example, $P(13, c) =$ false, because 13 in binary is 1101.

Both M_1 and M_2 are models of T because the consequents of the defaults in T are true in the minimal worlds where the antecedents are true. In the first model, the minimal $b \wedge c$ worlds (there is only one, 7, which is 0111 in binary) are below the minimal a worlds. In the second model, the minimal a worlds (there is only one, 13, which is 1101 in binary) are below the minimal $b \wedge c$ worlds. Hence there is no ordering $\prec$ between the consequents of the defaults in T.

Consider the set $P = \{a, c, \neg d\}$. The question is whether b is entailed by P, that is, whether b is true in the lexicographically minimal models of P and T. A default $\alpha \rightharpoonup \beta$ *prefers* a propositional model m_1 over m_2 if $m_1 \models \alpha \rightarrow \beta$ and $m_2 \models \alpha \wedge \neg\beta$. There are 16 assignments of truth values to the propositional variables a, b, c, d, only two of which satisfy P, $m_1 = 1010$ and $m_2 = 1110$. Delgrande defines that $m < m'$ if and only if there is a default δ in T such that δ prefers m to m' and for all defaults $\alpha \rightharpoonup \beta \in T$ such that $\alpha \rightharpoonup \beta$ prefers m' to m there is a default $\alpha' \rightharpoonup \beta'$ such that $\alpha' \rightharpoonup \beta'$ prefers m to m' and $\alpha \prec \alpha'$. Now $a \rightharpoonup b$ prefers m_2 over m_1 and $b \wedge c \rightharpoonup d$ prefers m_1 over m_2, and in neither case there is a more important default that has the opposite preference. Hence there is no ordering between m_1 and m_2, and consequently both are minimal. Because $m_1 \not\models b$, P does not entail b. ■

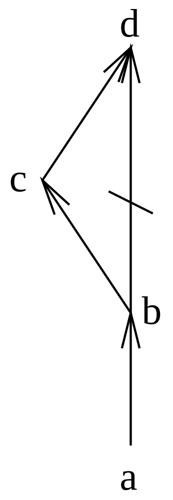


Figure 7.8: An extended penguin triangle

7.3.4 Default logic with specificity by Brewka

Brewka [1994] presents a version of prioritized default logic (see Section 6.1.2) and proposes a general method for computing priorities for a class of default theories that subsumes inheritance networks. Brewka's default theories are triples $\Delta = \langle D, T, C\rangle$ where D is a set of default rules, $T \subseteq \mathcal{L}$ is a set of background facts, and $C \subseteq \mathcal{L}$ is a set of contingent facts. The priorities for Δ are a function of D and T.

Definition 7.3.32 ([Brewka, 1994]) *Let $\Delta = \langle D, T, C\rangle$ be a default theory. A set of defaults $D' \subseteq D$ is* p-conflicting *(in Δ) if and only if for some $\pi{:}\gamma/\gamma \in D'$ we have $T \cup \{\pi\} \models_{D'} \bot$. A p-conflicting set D' is* resolvable *if it contains a default $\pi{:}\gamma/\gamma$ such that $T \cup \{\pi\} \not\models_{D'} \bot$.*

The consequence relation $\models_D$ can be defined by $\Sigma \models_D \phi$ whenever $\phi \in Cn_\Sigma(D)$, where $Cn_\Sigma(D)$ is defined in Theorem 2.1.3.

Definition 7.3.33 ([Brewka, 1994]) *Let $\Delta = \langle D, T, C\rangle$ be a default theory. $<_\Delta$ is* the specificity ordering *associated with Δ if and only if $<_\Delta$ is the smallest strict partial order on D such that $\delta_1 = \pi_1{:}\gamma_1/\gamma_1 <_\Delta \pi_2{:}\gamma_2/\gamma_2$ whenever δ_1 and δ_2 are in a minimal set $D' \subseteq D$ that is p-conflicting in Δ, and D' is not resolvable or $T \cup \{\pi_1\} \models_{D'} \bot$ and $T \cup \{\pi_2\} \not\models_{D'} \bot$.*

Brewka gives several examples of inheritance networks to demonstrate that his definition captures the intuitive idea of specificity. However, there are inheritance networks for which the definition may give unintuitive results.

Example 7.3.2 Consider the inheritance network in Figure 7.8. It corresponds to the set of defaults

$$D = \left\{ \frac{a : b}{b}, \frac{b : c}{c}, \frac{c : d}{d}, \frac{b : \neg d}{\neg d} \right\}.$$

When the set of background facts is empty, Brewka's definition gives the following ordering on the defaults.

$$<_D = \left\{ \left\langle \frac{b : \neg d}{\neg d}, \frac{c : d}{d} \right\rangle \right\}$$

We show that $E = Cn(\{a, b, c, d\})$ is a preferred extension of $\langle D, \emptyset, \{a, c\} \rangle$. One of the strict total orders $\ll$ that extend $<_D$ is $b{:}\neg d/\neg d$, $c{:}d/d$, $b{:}c/c$, $a{:}b/b$. Hence by Brewka's definitions we get the following sequence of sets E_i.

$$\begin{array}{rcl} E_0 & = & Cn(\{a, c\}) \\ E_1 & = & Cn(E_0 \cup \{d\}) \\ E_2 & = & Cn(E_1 \cup \{b\}) \\ E_3 & = & Cn(E_2) \\ E_i & = & E_3 = Cn(\{a, b, c, d\}) \text{ for all } i \geq 4 \end{array}$$

Therefore $E = \bigcup_{i \geq 0} E_i$ is the extension of $\langle D, \{a, c\}, <_D \rangle$ generated by $\ll$, and $\neg d$ is not a consequence of $\langle D, \emptyset, \{a, c\} \rangle$. ■

The problem identified in the above example could be avoided by imposing an ordering between defaults in the inheritance network that are not directly in conflict, as required by the second principle given in Section 7.1.3. The reason why these priorities are needed is not obvious, and it seems to be related to Example 6.1.2 in Section 6.1.2 that demonstrates how cumulativity [Makinson, 1994] is not preserved by the introduction of priorities in Brewka's default logic. Reasoning in the class of default theories that corresponds to acyclic inheritance networks is cumulative, but the introduction of priorities destroys this property. More priorities are needed to recover cumulativity. The motivation for these priorities is different in the context of Brewka's prioritized default logic than in the context of lexicographic priority mechanisms. It is easy to show that the introduction of the kind of lexicographic priority mechanism used in this work preserves cumulativity.

7.3.5 Non-normal defaults

One way of introducing priorities to sets of normal defaults is to modify the justifications of some of the defaults so that their application is prevented in case a conflicting higher priority default is applicable. This idea of using the class of seminormal defaults has been applied in translational theories of inheritance reasoning [Reiter and Criscuolo, 1981; Etherington and Reiter, 1983; Etherington, 1987; Etherington, 1988; Delgrande and Schaub, 1994].

Reiter and Criscuolo [1981] suggest that the application of the default $b{:}f/f$ can be prevented in the presence of the fact p by extending the justification of the default by the justification $\neg p$. If b, p and f are interpreted as in the penguin triangle this would simply mean that concluding that a bird flies is justified whenever there is no indication of the bird being a penguin. In general, the proposal by Reiter and Criscuolo says that $\alpha{:}\beta/\beta$ is given a higher priority than $\alpha'{:}\beta'/\beta'$ by replacing the latter with $\alpha'{:}\beta' \wedge \neg\alpha/\beta'$. Hence the lower priority default may not be applied in extensions that contain the prerequisite of the higher priority default.

Delgrande and Schaub [1994] propose a more sophisticated scheme for expressing priorities by using seminormal defaults. They suggest that the default $\alpha{:}\beta/\beta$ is given a higher priority than the default $\alpha'{:}\beta'/\beta'$ by adding to the latter default the new justification $\alpha \rightarrow \beta$. This strengthened justification guarantees that the latter default is not applied in extensions that contain α and $\neg\beta$. Delgrande and Schaub claim that their definition gives results that are more plausible than those obtained following Reiter and Criscuolo's proposal.

The next example demonstrates that the expression of priorities by using seminormal defaults fundamentally diverges from the use of lexicographic or other kinds of priority mechanisms that select a subset of the extensions that fulfill Reiter's [1980] definition.

Example 7.3.3 Consider the defaults $D = \{p{:}r \wedge s/r \wedge s, p{:}q/q, q{:}\neg r/\neg r\}$. To give the default $p{:}r \wedge s/r \wedge s$ a higher priority than $q{:}\neg r/\neg r$, the latter is replaced by the default $q{:}\neg r \wedge (p \rightarrow r \wedge s)/\neg r$. Let $D' = \{p{:}r \wedge s/r \wedge s, p{:}q/q, q{:}\neg r \wedge (p \rightarrow r \wedge s)/\neg r\}$. The default theory $\Delta' = \langle D', \{p, \neg s\}\rangle$ has only one extension $E'_1 = Cn(\{p, \neg s, q\})$. Hence $\Delta' \not\models_c r$ and $\Delta' \not\models_c \neg r$. The default theory $\langle D, \{p, \neg s\}\rangle$ has the extension $E_1 = Cn(\{p, \neg s, q, \neg r\})$, and the subset E'_1 of E_1 is not an extension of $\langle D, \{p, \neg s\}\rangle$. ■

If not all schemes of using seminormal defaults for expressing priorities, at least the ones presented by Delgrande and Schaub and by Reiter, Criscuolo and Etherington do not respect the following property. Define a priority scheme S on default theories as a function that maps a default theory Δ and an ordering relation $\mathcal{P}$ to a default theory Δ'. In the case of the schemes presented above, defaults in Δ are extended with justifications that make them seminormal. A priority scheme S is *faithful* if all extensions of Δ' are extensions of Δ[3]. As demonstrated in Example 7.3.3, the priority schemes

[3] There are presumably more general and reasonable definitions of *faithful* that allow a

of [Reiter and Criscuolo, 1981; Etherington and Reiter, 1983; Delgrande and Schaub, 1994] are not faithful. Faithfulness should be reflected against the idea that the extensions of Δ are the possible belief states of an agent (as suggested in Section 2.1), and priorities are a mechanism for selecting the most plausible belief states. A priority mechanism that is not faithful is not compatible with this idea.

7.3.6 Nested abnormality theories of Lifschitz

Lifschitz [1995] proposes a way of giving priorities in circumscription that is alternative to prioritized circumscription (see Section 6.4.) His *nested abnormality theories* are defined as trees with sets of formulae as nodes and with a designated root node. A subtree of a tree is called a *block*, the formulae in the block is the set of formulae in the root node of the subtree, and the blocks inside the block are the subtrees rooted at the children of the root node. There is a set of abnormality predicates associated with each block. Each block $B = [A_1, \ldots, A_j]\{\Phi; B_1, \ldots, B_m\}$ with abnormality predicates $\mathrm{ab}_1, \ldots, \mathrm{ab}_k$ is reduced to a set of formulae Ψ of the predicate logic as follows. First, each block B_i is recursively reduced to a set of formulae Ψ_i. Then the formulae $\Phi \cup \Psi_1 \cup \cdots \cup \Psi_m$ are circumscribed with respect to each abnormality predicate ab_i with the predicates $A_1, \ldots, A_j$ varied. The resulting sets of formulae are $\Phi_1, \ldots, \Phi_k$. Finally Ψ is defined as $\Phi_1 \cup \cdots \cup \Phi_k$.

A novelty in the definition of nested abnormality theories is that instead of abnormality predicates, existentially quantified predicate variables are used. This improves modularity because now there are no clashes between abnormality predicate symbols in different blocks. However, the ability to refer to abnormality in other blocks is lost.

Lifschitz proposes a way to construct nested abnormality theories that reflect the specificity relations between defaults. First, a more specific default is in a block that is inside the block of a less specific default. Second, defaults of equal or incomparable priority are in the same block and each has its own abnormality predicate.

Example 7.3.4 (Extended Nixon Diamond) The inheritance network is the same as in Figure 3.2 except that there are three additional links stating that republicans and quakers are Americans, and Americans are usually pacifists. The following nested abnormality theory is obtained by following

priority scheme for example to introduce auxiliary propositional variables to a default theory and still essentially choose a subset of the extensions of the original default theory.

Lifschitz's instructions [1995].

$$
\begin{array}{l}
\langle P : \forall x(A(x) \land \neg \mathrm{ab}_A(x) \rightarrow P(x)), \\
\quad \langle P : \forall x(R(x) \land \neg \mathrm{ab}_R(x) \rightarrow \neg P(x)), \\
\qquad \forall x(Q(x) \land \neg \mathrm{ab}_Q(x) \rightarrow P(x)) \\
\qquad \langle A : \forall x(R(x) \rightarrow A(x)), \\
\qquad\quad \forall x(Q(x) \rightarrow A(x)), \\
\qquad\quad R(n), Q(n) \rangle\rangle\rangle
\end{array}
$$

■

A restriction in using nested abnormality theories for inheritance reasoning the way Lifschitz suggests is that the expression of non-layered priorities by using the proposed principles is not possible, because a default that has an incomparable specificity should be in the same block as the two defaults of different specificity, and the block of one of the latter two defaults should be inside the block of the other.

The second restriction is more serious. Consider the extended Nixon diamond presented above. Circumscribing the abnormality predicates ab_R and ab_Q leaves the question of pacifism completely open (see [Lifschitz, 1994, Section 4.3].) At this point the nested abnormality theory has been reduced to the following formulae.

$$
\begin{array}{l}
\forall x(A(x) \land \neg \mathrm{ab}_A(x) \rightarrow P(x)), \\
\forall x(R(x) \land \neg Q(x) \rightarrow \neg P(x)) \\
\forall x(Q(x) \land \neg R(x) \rightarrow P(x)) \\
\forall x(R(x) \rightarrow A(x)) \\
\forall x(Q(x) \rightarrow A(x)) \\
R(n), Q(n)
\end{array}
$$

It remains to circumscribe the abnormality predicate variable ab_A in the outermost block. As the above formulae do not entail $P(c)$ or $\neg P(c)$ for any individual, the result of the circumscription is $\forall x \neg \mathrm{ab}_A(x)$, that is, we obtain the universally quantified implication $\forall x(A(x) \rightarrow P(x))$ that says that all Americans – including Nixon – are pacifists. The conflict between the two higher priority defaults makes it possible for the lower priority default to override the conflicting higher priority default. Only the highest priority applicable defaults should be used in determining the conclusions, and this is not the case in the above formalization of inheritance networks as nested abnormality theories.

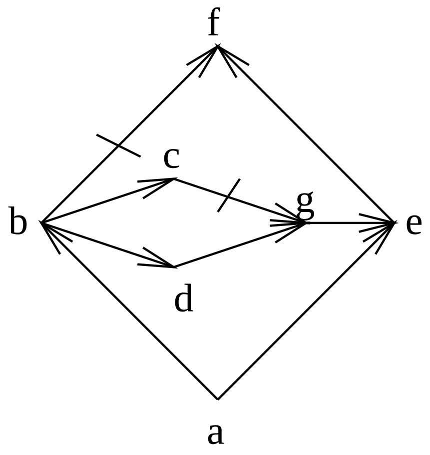


Figure 7.9: An inheritance network with varying priorities

7.4 Weakness of certain priority mechanisms

In this section we argue that lexicographic priority mechanism are not in general suitable for achieving default reasoning that properly takes into account the specificity of default rules. The argument proceeds by demonstrating the inexistence of sufficiently strong priorities that would yield the correct conclusions from certain kinds of inheritance networks. Because sufficiently strong priorities do not exist, the conclusions that can be obtained are too weak.

Examples that suggest the inappropriateness of fixed priorities in inheritance reasoning have been presented earlier [Horty, 1994]. Specifically, it has been argued that whenever each inheritance network is associated with fixed priorities that do not depend on the particular inferences sanctioned by the inheritance network, some of the conclusions that are obtained are unintuitive, and hence nonmonotonic logics with priorities produce conclusions that are too strong. Our arguments complement this view: for all members of some class of inheritance networks, any given priorities either produce too weak conclusions, that is, some of the intended conclusions are not obtained, or, they produce too strong conclusions, that is, some of the conclusions that are obtained are not intended. Therefore there is a mismatch between the capabilities of the priority mechanisms and what is required of intuitively correct inheritance reasoning. An inheritance network where fixed priorities seem to produce too strong conclusions is the following [Horty, 1994].

Example 7.4.1 Consider the inheritance network in Figure 7.9. Fulfillment of Condition 7.1 requires that the default $b{:}\neg f/\neg f$ is given a higher priority than the default $e{:}f/f$. Now consider extensions of $\Delta = \langle D, \{a\}\rangle$. Because $b{:}\neg f/\neg f$ has a higher priority than $e{:}f/f$, every extension of Δ contains $\neg f$ instead of f. However, researchers argue that $b{:}\neg f/\neg f$ is more specific than $e{:}f/f$ only in extensions where all of $b{:}d/d$, $d{:}g/g$ and $g{:}e/e$ are applied, that is, when $d{:}g/g$ is applied instead of $c{:}\neg g/\neg g$. This would allow an extension that contains f to be preferred. ■

Example 7.4.1 seems to indicate that Condition 7.1 and fixed priorities produces too strong conclusions in some cases. This example or the ones to be presented later cast doubt on the applicability of fixed priorities in solving problems of default reasoning. None of the examples given are conclusive in the sense that they would show the impossibility of useful and justifiable systems of nonmonotonic reasoning with priorities. Rather, we argue that known ways of using priorities are inherently constrained in ways that should be taken into account more seriously.

The following discussion relies on a number of assumptions on the nature and purpose of priorities. Priorities are viewed as an integral part of a knowledge base (an inheritance network.) The purpose of a knowledge base is to provide answers to queries. Queries do not affect the contents of the knowledge base, in particular, the priorities associated with the knowledge base are not affected by the queries, and the priorities may not depend on the particular queries made. Hence we reject the idea that a given knowledge base and a query together determine the priorities that are used in answering the query.

The knowledge bases or inheritance networks given in this section are quite ordinary, consisting of either two separate links or two separate penguin triangles. However, contrary to earlier research on inheritance reasoning, the queries may be arbitrary formulae containing for example material implications. This kind of non-atomic queries together with some basic assumptions on what qualifies as intuitively correct inheritance reasoning lead to problems.

Research on inheritance networks has usually considered only networks that are acyclic [Horty, 1994]. That lexicographic priorities have difficulties with cyclic inheritance networks is shown in the following example. The examples given later in this section are acyclic.

Example 7.4.2 (Cyclic inheritance networks) Consider the inheritance networks in Figure 7.10. The inheritance network on the right corresponds to the set of formulae $T = \{p_1 \rightarrow p_2, p_2 \rightarrow p_3, p_3 \rightarrow p_1\}$. Assume that $p_1 \rightarrow p_2$ does not have a higher priority than $p_2 \rightarrow p_3$. Now $T \cup \{p_1, \neg p_3\}$ has two preferred subtheories $S_1 = \{p_1 \rightarrow p_2, p_3 \rightarrow p_1, p_1, \neg p_3\}$ and $S_2 = \{p_2 \rightarrow p_3, p_3 \rightarrow p_1, p_1, \neg p_3\}$. Now $S_1 \models p_2$ and $S_1 \not\models p_2$. Hence by strong provability p_2 is not a consequence of $T \cup \{p_1, p_2\}$. Because of this, $p_1 \rightarrow p_2$ has to be given a higher priority than $p_2 \rightarrow p_3$, which in turn has to be given a higher priority than $p_3 \rightarrow p_1$ (by to similar reasoning), which in turn has to be given a higher priority than $p_1 \rightarrow p_2$. Clearly, no strict partial order satisfies these requirements.

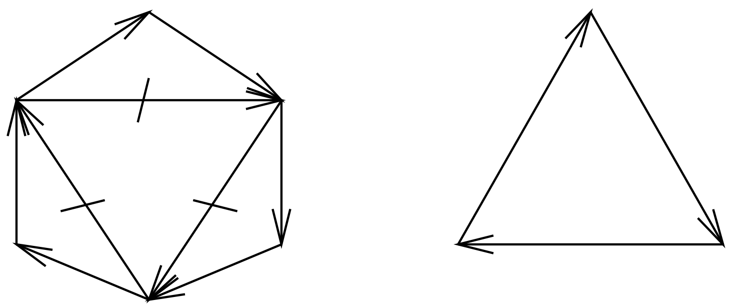


Figure 7.10: Two cyclic inheritance networks

The inheritance network on the left corresponds to the defaults

$$\begin{aligned} D \quad = \quad & \left\{ \frac{p_1 : \neg p_2}{\neg p_2}, \frac{p_1 : q_1}{q_1}, \frac{q_1 : p_2}{p_2}, \frac{p_2 : \neg p_3}{\neg p_3}, \frac{p_2 : q_2}{q_2}, \frac{q_2 : p_3}{p_3}, \right. \\ & \left. \frac{p_3 : \neg p_1}{\neg p_1}, \frac{p_1 : q_3}{q_3}, \frac{q_3 : p_1}{p_1} \right\}. \end{aligned}$$

Assume that $p_1{:}\neg p_2/\neg p_2$ is not given a higher priority than $p_2{:}\neg p_3/\neg p_3$. The default theory $\langle D, \{p_1\}\rangle$ has the extensions $E_1 = Cn(\{p_1, q_1, \neg p_2\})$, $E_2 = Cn(\{p_1, q_1, p_2, q_2, \neg p_3\})$, and $E_3 = Cn(\{p_1, q_1, p_2, q_2, p_3, q_3\})$. The extensions E_2 and E_3 do not contain the desired conclusion $\neg p_2$. To make E_1 preferred, $p_1{:}\neg p_2/\neg p_2$ should be given a higher priority than $p_2{:}\neg p_2/\neg p_2$ (examples can be given that demonstrate for example the insufficiency of prioritizing $p_1{:}\neg p_2/\neg p_2$ over $p_2{:}q_2/q_2$.) By similar reasoning, $p_2{:}\neg p_3/\neg p_3$ should be given a higher priority than $p_3{:}\neg p_1/\neg p_1$, and $p_3{:}\neg p_1/\neg p_1$ a higher priority than $p_1{:}\neg p_2/\neg p_2$. Clearly, no strict partial order satisfies these requirements. ∎

Conditional entailment and preferred subtheories

We give a simple example that suggests the unsuitability of the conditional entailment framework [Geffner and Pearl, 1992] and the preferred subtheories framework [Brewka, 1989] for achieving default reasoning that takes specificity into account. Model minimization with priorities as used in conditional entailment is equivalent to preferred subtheories for a large class of priorities, including layered priorities (by results in Section 6.2.) In our example there are only two defaults, and because all strict partial orders on two-element sets are layered, it therefore suffices to give the example for preferred subtheories only.

Let $T = \{a \rightarrow b, c \rightarrow d\}$ be a set of defaults. There are no defaults that contradict $a \rightarrow b$. Therefore, for the set of premises $P_1 = \{a, b \rightarrow c, \neg d\}$, we would like to have b as a conclusion (we think of $\{a, b \rightarrow c, \neg d\} \mid\!\sim b$ as a query made to the knowledge base $\{a \rightarrow b, c \rightarrow d\}$.) The previous sentence

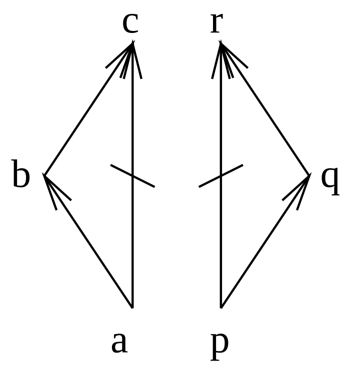


Figure 7.11: Two independent penguin triangles

relies on the idea of the default $a \rightarrow b$ as a directional rule. Assume $a \rightarrow b$ is not prioritized over $c \rightarrow d$. Then there is the total ordering $c \rightarrow d, a \rightarrow b$ of T. By Brewka's definitions, $S_0 = \{a, b \rightarrow c, \neg d\}$, $S_1 = S_0 \cup \{c \rightarrow d\}$, and $S_2 = S_1$. Consequently, $E = \{a, b \rightarrow c, \neg d, c \rightarrow d\}$ is a preferred subtheory of $T \cup P_1$. Clearly, $E \not\models b$, and b is not strongly provable from $T \cup P_1$. To make b strongly provable from $T \cup P_1$, $a \rightarrow b$ should be given a higher priority than $c \rightarrow d$. But, because of symmetry, $c \rightarrow d$ should be given a higher priority than $a \rightarrow b$ so that d were strongly provable from $T \cup P_2$ where $P_2 = \{c, d \rightarrow a, \neg b\}$. Clearly, there are no strict ordering relations on the defaults that fulfill both of these requirements. Therefore the intended inferences cannot be obtained in the framework of preferred subtheories or by using the prioritized model minimization used in conditional entailment.

Prioritized default logic

The above example suggesting that the expressivity of two well-known priority mechanisms is insufficient may seem to rely on the contraposition property of the defaults. However, essentially the same example can be reconstructed in the framework of default logic [Reiter, 1980] together with the lexicographic priority mechanism presented in Chapter 3. Default rules in default logic are not contraposable. The inheritance network is modified by replacing each default with an instance of the penguin triangle, as depicted in Figure 7.11.

We show that to obtain the intended conclusions the default $a{:}\neg c/\neg c$ has to be given a higher priority than $p{:}\neg r/\neg r$, and because of symmetry, $p{:}\neg r/\neg r$ has to be given a higher priority than $a{:}\neg c/\neg c$. This is impossible, and therefore there are no priorities that produce the intended conclusions. The argument is based on the following requirement.

$$\langle D, \{a, c \rightarrow p\}\rangle \models_{\mathcal{P}} \neg c$$

We want to derive $\neg c$ from $\{a\}$ and from any other set of premises that enables the application of defaults contradicting $a{:}\neg c/\neg c$ that are less specific than it.

Clearly, all defaults in the example that contradict $a{:}\neg c/\neg c$ are less specific than it. The default theory $\Delta = \langle D, \{a, c \rightarrow p\}\rangle$ has three extensions that apply the defaults shown below.

$$\begin{aligned}
\mathrm{GD}(E_1, \Delta) &= \left\{\frac{a:b}{b}, \frac{a:\neg c}{\neg c}\right\} \\
\mathrm{GD}(E_2, \Delta) &= \left\{\frac{a:b}{b}, \frac{b:c}{c}, \frac{p:q}{q}, \frac{q:r}{r}\right\} \\
\mathrm{GD}(E_3, \Delta) &= \left\{\frac{a:b}{b}, \frac{b:c}{c}, \frac{p:q}{q}, \frac{p:\neg r}{\neg r}\right\}
\end{aligned}$$

If E_2 or E_3 is a $\mathcal{P}$-preferred extension, the conclusion $\neg c$ cannot be obtained. Hence E_1 should be the only $\mathcal{P}$-preferred extension. Assume that not $a{:}\neg c/\neg c \mathcal{P} p{:}\neg r/\neg r$. Now there is the strict total order $\mathcal{T}$ on D such that $\mathcal{P} \subseteq \mathcal{T}$ and $p{:}\neg r/\neg r \mathcal{T} a{:}\neg c/\neg c$. We show that $\mathcal{T}$ is a $\Delta, \mathcal{P}$-characteristic ordering for E_3. Let δ be any default such that $\mathrm{appl}(\delta, E_1, E_3)$. Only $a{:}\neg c/\neg c$ is such default. Now $p{:}\neg r/\neg r \mathcal{T} a{:}\neg c/\neg c$ and $\mathrm{appl}(p{:}\neg r/\neg r, E_3, E_1)$. To solve the conflict between $p{:}\neg r/\neg r$ and $q{:}r/r$ given p, it has to be the case that $p{:}\neg r/\neg r \mathcal{P} q{:}r/r$. Let δ be any default such that $\mathrm{appl}(\delta, E_2, E_3)$. Only $q{:}r/r$ is such default. Now $p{:}\neg r/\neg r \mathcal{T} q{:}r/r$ and $\mathrm{appl}(p{:}\neg r/\neg r, E_3, E_2)$. Therefore E_3 is a $\mathcal{P}$-preferred extension of Δ, and because $\neg c \notin E_3$, finally $\langle D, \{a, c \rightarrow p\}\rangle \not\models_{\mathcal{P}} \neg c$.

Thus, as in the case with preferred subtheories and the model minimization used in conditional entailment, the example gives two requirements for the priorities that cannot be simultaneously satisfied. This allows us to conclude that no priorities satisfy the requirements. The same question can be raised in the context of preferred$_2$ extensions, and the answer is immediate by Lemma 3.1.8: the consequence relation $\models^2_{\mathcal{P}}$ is weaker than $\models_{\mathcal{P}}$ and hence does not sanction the intended conclusions.

Prioritized default logics by Baader and Hollunder, Brewka, and Marek and Truszczyński

Surprisingly, although the use of priorities in [Baader and Hollunder, 1993; Brewka, 1994; Marek and Truszczyński, 1993] is not based on lexicographic comparison, our example can be reconstructed in these systems of prioritized default logic.

Let $\langle D, W\rangle$ be a default theory where the defaults D are as in Figure 7.11 and $W = \emptyset$. Let $\mathcal{B}$ be a strict partial order on D. Assume that not $b{:}c/c \mathcal{B} q{:}r/r$. Now by using Definition 6.1.3 and taking $E = Cn(\{b, c \rightarrow p, q, c, r\})$,

$$\begin{aligned}
E_0 &= \{b, c \rightarrow p, q\}, \\
E_1 &= E_0 \cup \{c, r\}, \text{ and} \\
E_i &= E_1 \text{ for all } i \geq 2.
\end{aligned}$$

Therefore $E = Cn(\{b, c \rightarrow p, q, c, r\})$ is an S-extension of $\langle D, W\rangle$ contrary to what is intended. Therefore we need $\langle b{:}c/c, q{:}r/r\rangle$ in $\mathcal{B}$, and because of symmetry, also $\langle q{:}r/r, b{:}c/c\rangle \in \mathcal{B}$. And this is not possible because $\mathcal{B}$ should be asymmetric. Hence unintended extensions are obtained in prioritized default logics by Baader and Hollunder, Brewka, and Marek and Truszczyński.

We conclude that in the cases where lexicographic priorities are unsuitable for default reasoning with specificity, these three variants of prioritized default logic are unsuitable as well.

Conclusions

The examples we have given are fairly simple, and involve only sets of default that correspond to acyclic inheritance networks. The kind of pair of defaults that form the example concerning preferred subtheories and the model minimization used in conditional entailment can be found in almost all inheritance networks that would be used in practical applications. Similarly, the penguin triangles used in the example for prioritized default logic are the simplest instance of inheritance networks where specificity needs to be taken into account. The simplicity of the examples strongly suggests that there is nothing wrong with the sets of defaults per se (as far as we accept the possibility of conflicting defaults, the basic motivation for the research on nonmonotonic reasoning), and the problem we have identified is genuine. There seem to be several ways around the problem.

- Let the priorities depend on the query.

  It could be argued that the problem with the examples is that specificity is determined by the facts contained in the query, and therefore it is unnatural that priorities may not depend on the query. However, this would destroy the main advantage of using priorities: priorities could not be viewed as declarative data that are part of the knowledge base. If they are auxiliary data used during the evaluation of queries, it is not necessary nor useful to make them explicit.

- Do not allow disjunctivity in the queries.

  This is a subject that may be of worth pursuing. Without disjunctions the queries are essentially of the form $\Gamma \mathrel{|\!\sim} \phi$, where Γ is a set of literals and ϕ may be an arbitrary formula. Clearly, disjunctions are needed for expressing important forms of uncertainty, and this class of queries does not support applications for which such uncertainty is essential. Note that theories of inheritance presented by us in Section 7.2 and one of the systems presented by Gelfond and Przymusinska [1990] can be seen as an instance of the kind of restriction suggested here.

- Do not use lexicographic comparison.

  We are not aware of promising alternatives. Other ways of using priorities in default reasoning include best-first comparison [Benferhat *et al.*, 1993; Pearl, 1990], lexicographic comparison using the cardinality of sets of violated assumptions [Benferhat *et al.*, 1993], and the use of priorities in selecting the order in which default rules are applied [Baader and Hollunder, 1993; Marek and Truszczyński, 1993; Brewka, 1994]. Best-first comparison is weaker than lexicographic comparison, and exhibits the problem discussed in this section. The priorities that use cardinality of sets of violated assumptions has been proposed as a remedy to a problem present in best-first comparison, but does not address the problem presented in this section. As demonstrated above, prioritized default logics as presented by Baader and Hollunder, Brewka, and Marek and Truszczyński suffer from a similar problem.

7.4.1 Applicability of layered priorities in inheritance

Layered priorities (see Theorem 3.1.15) correspond to mappings of defaults to integers. In this section we point out some of the restrictions of layered priorities from the point of view of inheritance reasoning. Intuitively, the restriction to layered priorities precludes the possibility that a default has a priority that is incomparable to the priorities of two other defaults one of which has a higher priority than the other. There is a wide class of inheritance networks that can be represented by using layered priorities. In this section we give examples of inheritance networks that do not have natural such representations.

Definition 7.4.1 *Let $G = \langle V, E, I, P, N\rangle$ be an inheritance network. Let L be the set of links that end in a node. Then G is* non-layered*, if there is no assignment A of integers to L such that for all $n \in L$ and $m \in L$ such that the polarity of one is positive and the polarity of the other is negative, $A(n) < A(m)$ if and only if there is a positive path from the antecedent of n to the antecedent of m.*

Non-layered inheritance networks do not have a natural representation in systems of nonmonotonic reasoning with layered priorities, like in the less general version of Brewka's preferred subtheories [Brewka, 1989] in Section 6.2. For non-layered inheritance networks and given priorities that respect specificity it is possible to choose the members of P and N so that there is a conflict between two links neither of which is more specific than the other but nevertheless one of which has a higher priority than the other.

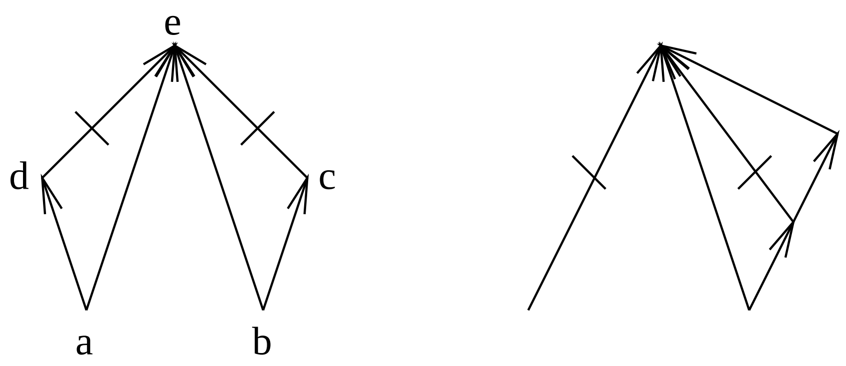


Figure 7.12: Two inheritance networks that require non-layered priorities

Example 7.4.3 Consider the leftmost inheritance network in Figure 7.12. Clearly $a{:}e/e$ should have a higher priority than $d{:}\neg e/\neg e$, and $b{:}e/e$ should have a higher priority than $c{:}\neg e/\neg e$. Neither of the defaults $a{:}e/e$ and $c{:}\neg e/\neg e$ should have a higher priority than the other, and neither of the defaults $b{:}e/e$ and $d{:}\neg e/\neg e$ should have a higher priority than the other. With layered priorities, the only way to not impose an ordering on two defaults is to have them on the same layer. Hence $a{:}e/e$ should be higher than $d{:}\neg e/\neg e$ which should be at the same layer as $b{:}e/e$ which should be higher than $c{:}\neg e/\neg e$. This however is impossible because $c{:}\neg e/\neg e$ should be at the same layer as $a{:}e/e$. Similarly for the second network, where the leftmost default should be incomparable in priority to the two conflicting defaults on the right, which in turn should have different priorities. This is an intuitive explanation why there are no satisfactory priorities for the inheritance network. This fact can be verified by an exhaustive case analysis of all possible priorities on the four defaults. ■

Note that from the existence of an counterexample (Section 6.4) that demonstrates that non-layered priorities are not in general expressible in prioritized circumscription it does not follow that some acyclic inheritance networks cannot be represented in prioritized circumscription. However, the priorities for the inheritance network in Example 7.4.3 are exactly the same that were used in the counterexample. Hence prioritized circumscription cannot represent this inheritance network in the obvious way.

Grosof [1991] argues for the use of arbitrary strict partial orders instead of layered strict partial orders. All his examples, however, only suggest that intuitively there is no reason to order certain defaults and therefore priorities should be non-layered. He does not argue for any of his examples that there are no layered priorities by using which exactly the same conclusions would be obtained.

Chapter 8

Conclusions

We have investigated the use of priorities in nonmonotonic logics, the computational properties of of nonmonotonic reasoning with priorities, and the applicability of nonmonotonic logics with priorities to the representation of knowledge. In particular, we have proposed a new way of extending Reiter's default logic [1980] with priorities that are based on lexicographic comparison. We have identified the location of the decision problems of the resulting system in the polynomial hierarchy, and identified restricted cases where the decision problems can be solved in polynomial time. We have compared our system of nonmonotonic reasoning with earlier approaches by Brewka [1989], Ryan [1992a], McCarthy [1986] and Lifschitz [1985], and Geffner and Pearl [1992]. We have presented a procedure for automated reasoning in prioritized default logic and proved it correct. Finally, we have established links between different translational theories of inheritance, for example that inheritance theories presented by Gelfond and Przymusinska [1990] and Brewka [1987] coincide with translational theories presented in this work.

References

Adams, E. W.: *The Logic of Conditionals: An Application of Probability to Deductive Logic*, volume 86 of *Synthese Library*. D. Reidel Publishing Company, Dordrecht, Holland, 1975.

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

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

Balcázar, J. L., I. Díaz, and J. Gabarró: *Structural Complexity I*. Springer-Verlag, Berlin, 1995.

Benferhat, S., C. Cayrol, D. Dubois, J. Lang, and H. Prade: Inconsistency management and prioritized syntax-based entailment. In R. Bajcsy, editor, *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, volume 1, pages 640–645, Chambery, France, 1993.

Brewka, G.: The logic of inheritance in frame systems. In J. McDermott, editor, *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, volume 1, pages 483–488, Milano, August 1987. Morgan Kaufmann Publishers.

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

Brewka, G.: *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge University Press, Cambridge, England, 1991.

Brewka, G.: Adding priorities and specificity to default logic. In *Logic in Artificial Intelligence: Proceedings of the 4th European Workshop JELIA'94*, number 838 in Lecture Notes in Artificial Intelligence, pages 247–260, York, England, 1994. Springer-Verlag.

Bylander, T., D. Allemang, M. C. Tanner, and J. R. Josephson: The computational complexity of abduction. *Artificial Intelligence*, 49:25–60, 1991.

Cook, S. A.: The complexity of theorem proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, pages 151–158, 1971.

Delgrande, J. P. and T. H. Schaub: A general approach to specificity in default reasoning. In *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*, pages 146–157, Bonn, May 1994. Morgan Kaufmann Publishers.

Delgrande, J. P.: A preference-based approach to default reasoning: Preliminary report. In *Proceedings of the 12th National Conference on Artificial Intelligence*, pages 902–908, Seattle, Washington, 1994.

Doyle, J.: A truth maintenance system. *Artificial Intelligence*, 12:231–272, 1979.

Eiter, T. and G. Gottlob: The complexity of logic-based abduction. *Journal of the ACM*, 42(1):3–42, 1995.

Eiter, T., G. Gottlob, and N. Leone: Semantics and complexity of abduction from default theories. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, volume 1, pages 870–876, Montréal, Canada, 1995. Morgan Kaufmann Publishers.

Britannica Online (version 1.31). Britannica Advanced Publishing, Inc., 1995.

Enderton, H. B.: *A Mathematical Introduction to Logic.* Academic Press, New York, 1972.

Etherington, D. and R. Reiter: On inheritance hierarchies with exceptions. In *Proceedings of the 2nd National Conference on Artificial Intelligence*, pages 104–108, Washington, DC, 1983.

Etherington, D. W.: Formalizing nonmonotonic reasoning systems. *Artificial Intelligence*, 31:41–85, 1987.

Etherington, D. W.: *Reasoning with Incomplete Information.* Pitman, London, 1988.

Fariñas del Cerro, L., A. Herzig, and J. Lang: From ordering-based nonmonotonic reasoning to conditional logics. *Artificial Intelligence*, 66:375–393, 1994.

Gärdenfors, P. and D. Makinson: Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65:197–245, 1994.

Gärdenfors, P.: *Knowledge in Flux.* The MIT Press, Cambridge, Massachusetts, 1988.

Garey, M. R. and D. S. Johnson: *Computers and Intractability.* W.H. Freeman and Company, San Francisco, 1979.

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

Geffner, H.: *Default Reasoning: Causal and Conditional Theories*. The MIT Press, Cambridge, Massachusetts, 1992.

Gelfond, M. and H. Przymusinska: Formalization of inheritance reasoning in autoepistemic logic. *Fundamenta Informaticae*, 13(4):403–443, 1990.

Gelfond, M., V. Lifschitz, H. Przymusinska, and M. Truszczyński: Disjunctive defaults. In J. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 230–237, Cambridge, Massachusetts, April 1991. Morgan Kaufmann Publishers.

Goldszmidt, M., P. Morris, and J. Pearl: A maximum entropy approach to nonmonotonic reasoning. In *Proceedings of the 8th National Conference on Artificial Intelligence*, pages 646–652, Boston, 1990.

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

Grosof, B. N.: Generalizing prioritization. In J. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 289–300, Cambridge, Massachusetts, April 1991. Morgan Kaufmann Publishers.

Hayes, P.: The logic of frames. In D. Metzing, editor, *Frame Conceptions and Text Understanding*, pages 46–61. Walter de Gruyter and Co., 1979. Reprinted as [Hayes, 1985].

Hayes, P.: The logic of frames. In R. Brachman and H. Levesque, editors, *Readings in Knowledge Representation*, pages 287–297. Morgan Kaufmann Publishers, 1985.

Horty, J. F.: Some direct theories of nonmonotonic inheritance. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming: Nonmonotonic Reasoning and Uncertain Reasoning*, volume 3, pages 111–187. Clarendon Press, Oxford, 1994.

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

Kanellakis, P. C.: Elements of relational database theory. In J. Van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 1073–1156. Elsevier Science Publishers, 1990.

Kautz, H. and B. Selman: Hard problems for simple default theories. *Artificial Intelligence*, 49:243–279, 1991.

Konolige, K.: Hierarchic autoepistemic theories for nonmonotonic reasoning: Preliminary report. In M. Reinfrank, J. de Kleer, M. L. Ginsberg, and E. Sandewall, editors, *Proceedings of the 2nd Workshop on Non-Monotonic Reasoning*, number 346 in Lecture Notes in Artificial Intelligence, pages 42–59, Grassau, Germany, June 1988. Springer-Verlag.

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

Krentel, M. W.: Generalizations of Opt P to the polynomial hierarchy. *Theoretical Computer Science*, 97:183–198, 1992.

Lehmann, D. and M. Magidor: What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.

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

Lifschitz, V.: Circumscriptive theories: a logic-based framework for knowledge representation. In *Proceedings of the 6th National Conference on Artificial Intelligence*, pages 364–368, Seattle, Washington, 1987.

Lifschitz, V.: Restricted monotonicity. In *Proceedings of the 11th National Conference on Artificial Intelligence*, pages 432–437, Washington, DC, 1993.

Lifschitz, V.: Circumscription. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming: Nonmonotonic Reasoning and Uncertain Reasoning*, volume 3, pages 297–352. Clarendon Press, Oxford, 1994.

Lifschitz, V.: Nested abnormality theories. *Artificial Intelligence*, 74:351–365, 1995.

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

Makinson, D.: General patterns in nonmonotonic reasoning. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial*

Intelligence and Logic Programming: Nonmonotonic Reasoning and Uncertain Reasoning, volume 3, pages 35–110. Clarendon Press, Oxford, 1994.

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

Marek, V. W. and M. Truszczyński: *Nonmonotonic Logic: Context-Dependent Reasoning*. Springer-Verlag, 1993.

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

McCarthy, J. and P. J. Hayes: Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, Edinburgh, 1969.

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

McCarthy, J.: Applications of circumscription to formalizing commonsense knowledge. *Artificial Intelligence*, 28:89–116, 1986.

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

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

Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, San Mateo, California, 1988.

Pearl, J.: Probabilistic semantics for nonmonotonic reasoning: a survey. In R. J. Brachman, H. J. Levesque, and R. Reiter, editors, *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, pages 505–516. Morgan Kaufmann Publishers, 1989. Reprinted in *Readings in Uncertain Reasoning*, G. Shafer and J. Pearl (eds.), Morgan Kaufmann Publishers, San Francisco, California, 1990, pp. 699–710.

Pearl, J.: System Z: a natural ordering of defaults with tractable applications to nonmonotonic logic. In *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 121–135, Pacific Grove, California, March 1990. Morgan Kaufmann Publishers.

Pollock, J. L.: Justification and defeat. *Artificial Intelligence*, 67:377–407, 1994.

Poole, D. L.: On the comparison of theories: preferring the most specific explanation. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, pages 144–147, Los Angeles, California, 1985.

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

Poole, D.: The effect of knowledge on belief: conditioning, specificity and the lottery paradox in default reasoning. *Artificial Intelligence*, 49(1–3):281–307, 1991.

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

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

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

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

Ryan, M. D.: *Ordered Presentations of Theories: Default Reasoning and Belief Revision*. PhD thesis, University of London, 1992.

Simari, G. R. and R. P. Loui: A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence*, 53:125–157, 1992.

Stillman, J.: It is not my default: the complexity of reasoning in default logic. In *Proceedings of the 8th National Conference on Artificial Intelligence*, pages 571–578, Boston, August 1990. Morgan Kaufmann Publishers.

Stillman, J.: The complexity of propositional default logics. In *Proceedings of the 10th National Conference on Artificial Intelligence*, pages 794–799, San Jose, California, 1992.

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

Appendix A

Auxiliary results

A.1 Default logic

Theorem A.1.1 *Let $\langle D, W\rangle$ be a default theory where all defaults in D have at least one justification. Let E be an extension of $\langle D, W\rangle$. Then E is consistent if and only if W is consistent.*

Proof: Assume that W is inconsistent. Because by Theorem 2.1.5 $E = Cn(W \cup \{\gamma | \alpha{:}\sigma/\gamma \in \mathrm{GD}(E, \Delta)\})$, E is inconsistent. Assume that E is inconsistent. By Theorem 2.1.3 $E = Cn_W(D^E)$. Because $\neg\beta \in E$ for all $\alpha{:}\sigma/\gamma \in D$ and β in σ and σ is not the empty sequence, $D^E = \emptyset$. Now $E = Cn_W(D^E) = Cn(W)$, and hence W is inconsistent. $\square$

Lemma A.1.2 *If an extension E of a default theory Δ is inconsistent, then E is the only extension of Δ.*

Proof: Assume E is an inconsistent extension of Δ, that is, $E = \mathcal{L}$, and assume that E' is an extension of Δ. Because $E' \subseteq E$, by [Reiter, 1980, Theorem 2.4], $E = E'$. Therefore E is the only extension of Δ. $\square$

Lemma A.1.6 shows that the defaults defeated in extensions of a default theory have the same property as the generating defaults of extensions: for two different extensions the defeated defaults in one are not a subset of the defeated defaults in the other. The proof of the lemma is based on Theorem 2.1.3 and the following lemmata.

Lemma A.1.3 *Let D and D' be sets of defaults such that $D' \subseteq D$. Then $Cn_W(D') \subseteq Cn_W(D)$.*

Proof: Let $C_0, C_1, \ldots$ be the sets for $Cn_W(D)$ and $C'_0, C'_1, \ldots$ the sets for $Cn_W(D')$ as defined in Theorem 2.1.3. *Induction hypothesis*: $C'_i \subseteq C_i$. *Base case* $i = 0$: Immediate as $C_0 = W = C'_0$. *Inductive case* $i \geq 1$: Assume that $\gamma \in C'_i$. If $\gamma \in C'_{i-1}$, then by induction hypothesis $\gamma \in C_{i-1}$ and hence $\gamma \in C_i$. If $\gamma \in C'_i \backslash C'_{i-1}$, then there is a default $\alpha{:}\sigma/\gamma \in D'$ and $C'_{i-1} \models \alpha$. By induction hypothesis $C'_{i-1} \subseteq C_{i-1}$, and hence $C_{i-1} \models \alpha$. Because $D' \subseteq D$, $\alpha{:}\sigma/\gamma \in D$, and hence $\gamma \in C_i$ by definition. Hence for all $i \geq 0$, $C'_i \subseteq C_i$ and therefore $Cn_W(D') = \bigcup_{i\geq 0} Cn(C'_i) \subseteq \bigcup_{i \geq 0} Cn(C_i) = Cn_W(D)$. □

Lemma A.1.4 *Let $E \subseteq \mathcal{L}$ be a set of formulae and D a set of defaults. Let $D^E = \{\alpha{:}\beta_1, \ldots, \beta_n/\gamma \in D | \{\neg\beta_1, \ldots, \neg\beta_n\} \cap E = \emptyset\}$ and $D_E = \{\delta \in D | not\ def(\delta, E)\}$. Then $D^E \subseteq D_E$.*

Proof: Assume that $\alpha{:}\beta_1, \ldots, \beta_n/\gamma \in D^E$. Hence by definition of D^E $\neg\beta_i \in E$ for no $i \in \{1, \ldots, n\}$. Therefore not def$(\alpha{:}\beta_1, \ldots, \beta_n/\gamma, E)$ and finally $\alpha{:}\beta_1, \ldots, \beta_n/\gamma \in D_E$. □

Lemma A.1.5 *For all $E \subseteq \mathcal{L}$, $Cn_W(D_E) = Cn_W(D^E)$, where $D_E = \{\delta \in D | not\ def(\delta, E)\}$.*

Proof: By Lemmata A.1.3 and A.1.4 $Cn_W(D^E) \subseteq Cn_W(D_E)$. It remains to show that $Cn_W(D_E) \subseteq Cn_W(D^E)$. Let $C_0, C_1, \ldots$ be the sets for $Cn_W(D^E)$ and $C'_0, C'_1, \ldots$ the sets for $Cn_W(D_E)$ as defined in Theorem 2.1.3. *Induction hypothesis*: $C'_i \subseteq C_i$. *Base case* $i = 0$: Immediate as $C'_0 = W = C_0$. *Inductive case* $i \geq 1$: Assume that $\gamma \in C'_i$. If $\gamma \in C'_{i-1}$, then by induction hypothesis $\gamma \in C_{i-1} \subseteq C_i$. If $\gamma \in C'_i \backslash C'_{i-1}$, then there is a default $\alpha{:}\sigma/\gamma \in D_E$ and $C'_{i-1} \models \alpha$. By induction hypothesis $C'_{i-1} \subseteq C_{i-1}$ and hence $C_{i-1} \models \alpha$. Because $\alpha{:}\sigma/\gamma \in D_E$, either $\alpha \notin E$ or $\neg\beta \in E$ for no β in σ. Because $C_{i-1} \models \alpha$ and $C_{i-1} \subseteq E$, $\neg\beta \in E$ for no β in σ, and hence $\alpha{:}\sigma/\beta \in D^E$. Hence $\gamma \in C_i$.

Hence $C'_i \subseteq C_i$ for all $i \geq 0$ and therefore $Cn_W(D_E) = \bigcup_{i \geq 0} Cn(C'_i) \subseteq \bigcup_{i\geq 0} Cn(C_i) = Cn_W(D^E)$. As $Cn_W(D^E) \subseteq Cn_W(D_E)$, finally $Cn_W(D^E) = Cn_W(D_E)$. □

Lemma A.1.6 *Let E and E' be extensions of a default theory $\Delta = \langle D, W\rangle$ and $E \neq E'$. Then $\{\delta \in D | not\ def(\delta, E)\} \not\subseteq \{\delta \in D | not\ def(\delta, E')\}$.*

Proof: Let $D_E = \{\delta \in D | \text{not def}(\delta, E)\}$ and $D_{E'} = \{\delta \in D | \text{not def}(\delta, E')\}$. Assume that $D_E \subseteq D_{E'}$. By Theorem 2.1.3 and Lemma A.1.5, $E = Cn_W(D_E)$ and $E' = Cn_W(D_{E'})$. As $D_E \subseteq D_{E'}$, by Lemma A.1.3 $E \subseteq E'$. This contradicts [Reiter, 1980, Theorem 2.4]. □

Lemma A.1.7 *Let D be a set of defaults and W a set of formulae such that $E = Cn_W(D)$ and $\neg\beta \notin E$ for all $\alpha{:}\sigma/\gamma \in D$ and β in σ. Let $D' = \{\alpha{:}\sigma/\gamma \in D | \alpha \in Cn_W(D)\}$. Then E is an extension of $\langle D', W\rangle$ and $D' = GD(E, \langle D', W\rangle)$.*

Proof: Because $\neg\beta \in E$ for no $\alpha{:}\sigma/\gamma \in D'$ and β in σ, ${D'}^E = D'$. Because $D' \subseteq D$, by Lemma A.1.3 $Cn_W(D') \subseteq Cn_W(D)$. We show by induction that $Cn_W(D) \subseteq Cn_W(D')$. Let $C_0, C_1, \ldots$ be the sets defined in Theorem 2.1.3 for $Cn_W(D)$, and $C'_0, C'_1, \ldots$ for $Cn_W(D')$. *Induction hypothesis*: $C_i \subseteq C'_i$. *Base case* $i = 0$: $C'_0 = W = C_0$ by definition. *Inductive case* $i \geq 1$: Assume that $\gamma \in C_i$. If $\gamma \in C_{i-1}$ then by induction hypothesis $\gamma \in C'_{i-1} \subseteq C_{i-1} \subseteq C_i$. If $\gamma \in C_i \backslash C_{i-1}$, then $C_{i-1} \models \alpha$ for some $\alpha{:}\sigma/\gamma \in D$. Because $C_{i-1} \subseteq E$ and E is closed under logical consequence, $\alpha \in E$ and hence $\alpha{:}\sigma/\gamma \in D'$. By induction hypothesis $C_{i-1} \subseteq C'_{i-1}$ and hence $C'_{i-1} \models \alpha$. Therefore $\gamma \in C'_i$.

Hence $Cn_W(D) = Cn_W(D')$ and $Cn_W(D')$ is an extension of $\langle D', W\rangle$. Because for all $\alpha{:}\sigma/\gamma \in D'$ $\alpha \in E$ and $\neg\beta \in E$ for no $\beta \in \sigma$, $D' = \mathrm{GD}(E, \langle D', W\rangle)$ by definition. $\Box$

The extensions of two default theories that differ only slightly are in many cases closely related. The following lemmata describe several connections of this kind.

Lemma A.1.8 *Let α be a formula and γ a consistent formula, σ a finite sequence of formulae, and E an extension of a default theory $\langle D, W\rangle$ such that $\alpha \in E$ and $\neg\beta \notin Cn(E \cup \{\gamma\})$ for all β in σ, and propositional variables in γ do not occur anywhere in W and occur only in defaults $\alpha'{:}\beta'_1, \ldots, \beta'_n/\gamma' \in D$ such that $\{\neg\beta'_1, \ldots, \neg\beta'_n\} \cap E \neq \emptyset$. Then $Cn(E \cup \{\gamma\})$ is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$.*

Proof: Let $E' = Cn(E \cup \{\gamma\})$. Assume that $\alpha'{:}\sigma'/\gamma' \in D^E$. Hence for all β' in σ', $\beta' \notin E$ and there is a propositional model M such that $M \models E$ and $M \not\models \neg\beta'$. Because the propositional variables in W and D^E do not occur in γ, M can be modified to M' such that $M' \models E$ and $M' \not\models \neg\beta'$ and $M' \models \gamma$. Hence $Cn(E \cup \{\gamma\}) \not\models \neg\beta'$. Therefore $\alpha'{:}\sigma'/\gamma' \in D^{Cn(E\cup\{\gamma\})}$. Because $E \subseteq E \cup \{\gamma\}$, $D^{Cn(E\cup\{\gamma\})} \subseteq D^E$, and hence $D^E = D^{Cn(E\cup\{\gamma\})}$. Because $\neg\beta \in Cn(E \cup \{\gamma\})$ for no β in σ, $(D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})} = D^{Cn(E\cup\{\gamma\})} \cup \{\alpha{:}\sigma/\gamma\}$. We show that $Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})}) = Cn(E \cup \{\gamma\})$ and hence by Theorem 2.1.3 $Cn(E \cup \{\gamma\})$ is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$. Because E is an extension of $\langle D, W\rangle$, $E = Cn_W(D^E)$. Let $C_0, C_1, \ldots$ be the sets in the definition of $Cn_W(D^E)$, and $C'_0, C'_1, \ldots$ the sets in the definition of $Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})})$. *Induction hypothesis*: $C_i \subseteq C'_i \subseteq C_i \cup \{\gamma\}$. *Base case* $i = 0$: By definition $C'_0 = W = C_0$ and hence $C_0 \subseteq C'_0 \subseteq C_0 \cup \{\gamma\}$. *Inductive case* $i \geq 1$: Assume that $\gamma' \in C_i$. If $\gamma' \in C_{i-1}$, then by induction hypothesis $\gamma' \in C'_{i-1} \subseteq C'_i$. If $\gamma' \in C_i \backslash C_{i-1}$, then there is $\alpha'{:}\sigma'/\gamma' \in D^E$ such that $C_{i-1} \models \alpha'$. By induction hypothesis $C_{i-1} \subseteq C'_{i-1}$ and hence $C'_{i-1} \models \alpha'$. Because $D^E \subseteq D^{Cn(E\cup\{\gamma\})} \cup \{\alpha{:}\sigma/\gamma\}$, $\alpha'{:}\sigma'/\gamma' \in (D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})}$. Therefore $\gamma' \in C'_i$. Next we show the second inclusion. Assume that $\gamma' \in C'_i$. If $\gamma' \in C'_{i-1}$, then by induction hypothesis $\gamma' \in C_{i-1} \cup \{\gamma\} \subseteq C_i \cup \{\gamma\}$. If $\gamma' = \gamma$, then obviously $\gamma' \in C_i \cup \{\gamma\}$. If $\gamma' \in C'_i \backslash C'_{i-1}$ and $\gamma \neq \gamma'$, then there is $\alpha'{:}\sigma'/\gamma' \in D^{Cn(E\cup\{\gamma\})}$ such that $C'_{i-1} \models \alpha'$. Assume that $C_{i-1} \not\models \alpha'$, that is, there is a model M such that $M \models C_{i-1}$ and $M \not\models \alpha'$. Because propositional variables in γ do not occur in $D^{Cn(E\cup\{\gamma\})}$ or W, we can construct a model M' such that $M' \models C_{i-1}$ and $M' \models \gamma$ and $M' \not\models \alpha'$. Hence $C_{i-1} \cup \{\gamma\} \not\models \alpha'$. Because $C'_{i-1} \subseteq C_{i-1} \cup \{\gamma\}$, this contradicts the fact that $C'_{i-1} \models \alpha'$. Hence $C_{i-1} \models \alpha'$ and by definition $\gamma' \in C_i$. This finishes the induction proof.

Hence $\alpha \in \bigcup_{i\geq 0} Cn(C_i) \subseteq \bigcup_{i\geq 0} Cn(C'_i) = Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})})$. Hence by definition $\gamma \in Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})})$, and $Cn(E \cup \{\gamma\}) \subseteq Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})})$. Because $Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})}) \subseteq E\cup\{\gamma\}$, $Cn(E\cup\{\gamma\}) = Cn_W((D\cup\{\alpha{:}\sigma/\gamma\})^{Cn(E\cup\{\gamma\})})$, and finally $Cn(E\cup\{\gamma\})$ is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$. □

Lemma A.1.9 *Let α and γ be formulae and σ a sequence of formulae tha contains β. Let $E \subseteq \mathcal{L}$ be a set of formulae such that $\neg\beta \in E$. Then E is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$ if and only if E is an extension of $\langle D, W\rangle$.*

Proof: Let $D' = D \cup \{\alpha{:}\sigma/\gamma\}$. Because $\neg\beta \in E$ for some β in σ, $D^E = D'^E$. Hence $Cn_W(D^E) = Cn_W(D'^E)$, and by Theorem 2.1.3 E is an extension of $\langle D, W\rangle$ if and only if E is an extension of $\langle D', W\rangle$. □

Lemma A.1.10 *Let $\langle D, W\rangle$ be a default theory and E an extension of $\langle D, W\rangle$. Let ϕ be a consistent formula such that propositional variables in ϕ do not occur anywhere in D or W. Then $Cn(E\cup\{\phi\})$ is an extension of $\langle D, W\cup\{\phi\}\rangle$.*

Proof: First we show that $D^E = D^{Cn(E\cup\{\phi\})}$. Assume that $\alpha{:}\sigma/\gamma \in D^{Cn(E\cup\{\phi\})}$. Hence $\neg\beta \in Cn(E \cup \{\phi\})$ for no β in σ. Because $E \subseteq Cn(E \cup \{\phi\})$, $\neg\beta \in E$ for no β in σ and $\alpha{:}\sigma/\gamma \in D^E$. Therefore $D^{Cn(E\cup\{\phi\})} \subseteq D^E$. Assume that $\alpha{:}\sigma/\gamma \in D^E$. Hence $\neg\beta \in E$ for no β in σ. Hence for every β in σ there is a propositional model M such that $M \models E$ and $M \not\models \neg\beta$. Because $E = Cn(W \cup G)$ for a set G of conclusions of defaults in D and ϕ is consistent and propositional variables in ϕ do not occur anywhere in D or W, a model M' can be constructed so that $M' \models E$ and $M' \not\models \neg\beta$ and $M' \models \phi$. Hence $\neg\beta \in Cn(E \cup \{\phi\})$ for no β in σ. Therefore $\alpha{:}\sigma/\gamma \in D^{Cn(E\cup\{\phi\})}$ and $D^E \subseteq D^{Cn(E\cup\{\phi\})}$.

By Theorem 2.1.3 $E = Cn_W(D^E)$. Let $C_0, C_1, \ldots$ be the sets in the definition of $Cn_W(D^E)$. Let $C'_0, C'_1, \ldots$ be the sets in the definition of $Cn_{W\cup\{\phi\}}(D^{Cn(E\cup\{\phi\})}) =$

$Cn_{W\cup\{\phi\}}(D^E)$. *Induction hypothesis*: $C'_i = C_i \cup \{\phi\}$. *Base case* $i = 0$: By definition $C'_0 = W \cup \{\phi\} = C_0 \cup \{\phi\}$. *Inductive case* $i \geq 1$: Assume that $\gamma \in C_i \cup \{\phi\}$. If $\gamma = \phi$, then by induction hypothesis $\phi \in C'_{i-1} \subseteq C'_i$. If $\gamma \in C_{i-1}$, then by induction hypothesis $\gamma \in C'_{i-1} \subseteq C'_i$. If $\gamma \in C_i \backslash C_{i-1}$, then $C_{i-1} \models \alpha$ for some $\alpha{:}\sigma/\gamma \in D^E$. Because $C_{i-1} \subseteq C'_{i-1}$, $C'_{i-1} \models \alpha$ and consequently $\gamma \in C'_i$. Hence $C_i \cup \{\phi\} \subseteq C'_i$. Assume that $\gamma \in C'_i$. If $\gamma \in C'_{i-1}$, then by induction hypothesis $\gamma \in C_{i-1} \cup \{\phi\} \subseteq C_i \cup \{\phi\}$. If $\gamma \in C'_i \backslash C'_{i-1}$, then there is $\alpha{:}\sigma/\gamma \in D^E$ such that $C'_{i-1} \models \alpha$. Assume that $C_{i-1} \not\models \alpha$. Hence there is a model M such that $M \models C_{i-1}$ and $M \not\models \alpha$. Because propositional variables in ϕ do not occur in D or W, M can be modified to M' such that $M' \models C_{i-1}$ and $M' \models \phi$ and $M' \not\models \alpha$. Hence $C_{i-1} \cup \{\phi\} \not\models \alpha$. This contradicts the fact that $C'_{i-1} \models \alpha$ and by induction hypothesis $C'_{i-1} = C_{i-1} \cup \{\phi\}$. Hence $C_{i-1} \models \alpha$, and because $\alpha{:}\sigma/\gamma \in D^E$, $\gamma \in C_i \cup \{\phi\}$.

Finally, $Cn(E\cup\{\phi\})$ is an extension of $\langle D, W\cup\{\phi\}\rangle$, as $Cn_{W\cup\{\phi\}}(D^{Cn(E\cup\{\phi\})}) = \bigcup_{i\geq 0} Cn(C'_i) = \bigcup_{i\geq 0} Cn(C_i \cup \{\phi\}) = Cn(Cn_W(D^E) \cup \{\phi\}) = Cn(E \cup \{\phi\})$. □

Lemma A.1.11 *Let $E \subseteq \mathcal{L}$ be a set of formulae such that $\alpha \notin E$. Then E is an extension of $\langle D, W\rangle$ if and only if E is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$.*

Proof: Let $D' = D \cup \{\alpha{:}\sigma/\gamma\}$. Consider any $E \subseteq \mathcal{L}$. Assume that $\neg\beta \in E$ for some β in σ. Hence by Lemma A.1.9 E is an extension of $\langle D, W\rangle$ if and only if it is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$. Otherwise $D'^E = D^E \cup \{\alpha{:}\sigma/\gamma\}$. We show that $Cn_W(D'^E) = Cn_W(D^E)$. Let $C_0, C_1, \ldots,$ be the sets in the definition of $Cn_W(D^E)$ as given in Theorem 2.1.3, and $C'_0, C'_1, \ldots,$ the respective sets for $Cn_W(D'^E)$. *Induction hypothesis*: $C_i = C'_i$. *Base case* $i = 0$: $C_0 = W = C'_0$. *Inductive case* $i \geq 1$: Because by induction hypothesis $C_{i-1} = C'_{i-1}$, for all $\alpha'{:}\sigma'/\gamma' \in D^E$, $\gamma' \in C'_i$ iff $\gamma' \in C_i$ by definition. Because $\alpha \notin C'_{i-1} \subseteq E$, $\gamma \in C_i$ iff $\gamma \in C'_i$. Hence $C_i = C'_i$. Therefore $Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^E) = Cn_W(D^E)$ for any E such that $\alpha \notin E$ and $\neg\beta \in E$ for no β in σ. Hence by Theorem 2.1.3, for all $E \subseteq \mathcal{L}$, E is an extension of $\langle D, W\rangle$ if and only if E is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$. □

Lemma A.1.12 *Let E' be a consistent extension of the default theory $\Delta' = \langle D \cup D', W \cup W'\rangle$ where the propositional variables in D and W are distinct from the propositional variables in W' and in conclusions of defaults in D'. Then $E = Cn(\{\gamma|\alpha{:}\sigma/\gamma \in D \cap GD(E', \Delta')\} \cup W)$ is an extension of $\Delta = \langle D, W\rangle$.*

Proof: We need two lemmata for the proof. **Lemma A** Let Φ and Φ' be consistent sets of formulae such that the propositional variables in one do not occur in the other, and propositional variables in the formula ϕ do not occur in Φ'. Then $\Phi \models \phi$ if and only if $\Phi \cup \Phi' \models \phi$.

Proof: The "only if" direction is directly by the monotonicity of $\models$. Assume that $\Phi \not\models \phi$, that is, there is a model M such that $M \models \Phi$ and $M \not\models \phi$. Because propositional variables in Φ' are distinct from the propositional variables in Φ and ϕ and Φ' is consistent, a model M' can be constructed so that $M' \models \Phi'$ and the truth-values of propositional variables not in Φ' are the same in M and M'. Therefore $M' \models \Phi \cup \Phi'$ and $M' \not\models \phi$, and finally $\Phi \cup \Phi' \not\models \phi$. □

Let $G' = \{\gamma | \alpha{:}\sigma/\gamma \in \mathrm{GD}(E', \Delta')\}$. Let $G = \{\gamma | \alpha{:}\sigma/\gamma \in D \cap \mathrm{GD}(E', \Delta')\}$. Now $E = Cn(G \cup W)$ and by Theorem 2.1.5 $E' = Cn(G' \cup W \cup W')$.

Lemma B $D^E = D^{E'}$.
Proof: Because $E \subseteq E'$, clearly $D^{E'} \subseteq D^E$. Assume that $\alpha{:}\sigma/\gamma \in D \backslash D^{E'}$, that is, $\neg\beta \in E'$ for some β in σ. Hence $G' \cup W \cup W' \models \neg\beta$. By Lemma **A** $G \cup W \models \neg\beta$ because by assumption the sets of propositional variables occurring in $G \cup W$ and in $(G' \backslash G) \cup W'$ are disjoint. Hence $\alpha{:}\sigma/\gamma \in D \backslash D^E$. Hence $D^E \subseteq D^{E'}$ and $D^E = D^{E'}$. □

Let $R = \{\gamma | \alpha{:}\sigma/\gamma \in D\} \cup W$. Let $C_0, C_1, \ldots$ be the sets in the definition of $Cn_W(D^E)$, and $C'_0, C'_1, \ldots$ the sets in the definition of $Cn_{W \cup W'}((D \cup D')^{E'})$. *Induction hypothesis*: $C_i = C'_i \cap R$. *Base case* $i = 0$: $C_0 = W = (W \cup W') \cap R = C'_0 \cap R$. *Inductive case* $i \geq 1$: ($\subseteq$) Assume that $\gamma \in C_i$. If $\gamma \in C_{i-1}$ then by induction hypothesis $\gamma \in C'_{i-1} \cap R \subseteq C_i \cap R$ If $\gamma \in C_i \backslash C_{i-1}$, then there is $\alpha{:}\sigma/\gamma \in D^E$ such that $C_{i-1} \models \alpha$. Because $C_{i-1} \subseteq C'_{i-1}$, also $C'_{i-1} \models \alpha$. Because $D^E \subseteq (D \cup D')^{E'}$, $\gamma \in C'_i$. Because $\gamma \in R$, $\gamma \in C'_i \cap R$. ($\supseteq$) Assume that $\gamma \in C'_i \cap R$. If $\gamma \in C'_{i-1} \cap R$ then by induction hypothesis $\gamma \in C_{i-1} \subseteq C_i$. If $\gamma \in (C'_i \cap R) \backslash (C'_{i-1} \cap R)$, then there is $\alpha{:}\sigma/\gamma \in D^{E'}$ such that $C'_{i-1} \models \alpha$. As $D^E = D^{E'}$ by Lemma **B**, $\alpha{:}\sigma/\gamma \in D^E$. The set $C'_{i-1} \backslash C_{i-1}$ consists of members of $(G' \backslash G) \cup W'$, and propositional variables occurring in W' do not occur in C_{i-1} or α. Hence by Lemma **A** $C_{i-1} \models \alpha$, and $\gamma \in C_i$ by definition. This finishes the inductive step that establishes $C_i = C'_i \cap R$.

Hence $Cn_W(D^{Cn(G \cup W)}) = \bigcup_{i \geq 0} Cn(C_i) = \bigcup_{i \geq 0} Cn(C'_i \cap R) = Cn((G' \cup W \cup W) \cap R) = Cn(G \cup W)$, and therefore $E = Cn(G \cup W)$ is an extension of Δ.
□

Lemma A.1.13 *Let E be an extension of $\langle D, W \rangle$ such that $\gamma \in E$. Then E is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W \rangle$.*

Proof: **Lemma A** Let D' be a set of defaults and W' a set of formulae. If $E' = Cn_{W'}(D')$ and $\alpha' \in E'$ and $\alpha'{:}\sigma'/\gamma' \in D'$, then $\gamma' \in E'$.
Proof: Because $\alpha' \in E' = \bigcup_{i \geq 0} C'_i$ as defined in Theorem 2.1.3, $\alpha' \in C'_j$ for some $j \geq 0$. By definition $\gamma' \in C'_{j+1} \subseteq E'$. □

Let $C_0, C_1, \ldots$ be the sets for $Cn_W(D^E)$ as given in Theorem 2.1.3 and $C'_0, C'_1, \ldots$ the sets for $Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^E)$. Clearly either $(D \cup \{\alpha{:}\sigma/\gamma\})^E = D^E$ or

$(D \cup \{\alpha{:}\sigma/\gamma\})^E = D^E \cup \{\alpha{:}\sigma/\gamma\}$. *Induction hypothesis*: $C_i \subseteq C'_i \subseteq E$. *Base case* $i = 0$: $C_0 = C'_0 = W \subseteq E$. *Inductive case* $i \geq 1$: Assume that $\gamma' \in C_i$. If $\gamma' \in C_{i-1}$, then by induction hypothesis $\gamma' \in C'_{i-1} \subseteq C'_i$. If $\gamma' \in C_i \backslash C_{i-1}$, then $\alpha' \in C_{i-1}$ for some $\alpha'{:}\sigma'/\gamma' \in D^E$. By induction hypothesis $C_{i-1} \subseteq C'_{i-1}$ and hence $\alpha' \in C'_{i-1}$. Because $\alpha'{:}\sigma'/\gamma' \in (D \cup \{\alpha{:}\sigma/\gamma\})^E$, hence $\gamma' \in C'_i$. Then we show the second inclusion. Assume that $\gamma' \in C'_i$. If $\gamma' \in C'_{i-1}$, then by induction hypothesis $\gamma' \in E$. If $\gamma' \in C'_i \backslash C'_{i-1}$, then $\alpha' \in C'_{i-1}$ for some $\alpha'{:}\sigma'/\gamma' \in (D \cup \{\alpha{:}\sigma/\gamma\})^E$. If $\alpha'{:}\sigma'/\gamma' \neq \alpha{:}\sigma/\gamma$, then by Lemma **A** and because by induction hypothesis $C'_{i-1} \subseteq E$, $\gamma' \in E$. Otherwise $\alpha'{:}\sigma'/\gamma' = \alpha{:}\sigma/\gamma$ and by assumption $\gamma \in E$.

Hence $E = Cn_W(D^E) = Cn_W((D \cup \{\alpha{:}\sigma/\gamma\})^E)$ and E is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$. □

Lemma A.1.14 *Let $\langle D, W\rangle$ be a default theory where prerequisites, justifications and conclusions in D are literals, and every default in D is either normal or has no justifications, and W is a consistent set of literals. Let E be an extension of $\langle D, W\rangle$. Let $\alpha{:}\sigma/\gamma$ be a default such that $\alpha \in E$ and $\neg\gamma \notin E$ and σ is the empty sequence or consists of γ, and $\gamma{:}\sigma'/\gamma' \in D$ for no σ' and γ'. Then $E' = Cn(E \cup \{\gamma\})$ is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$.*

Proof: Define $D_E = \{\delta \in D | \text{not def}(\delta, E)\}$ and $D'_{E'} = \{\delta \in D' | \text{not def}(\delta, E')\}$ where $D' = D \cup \{\alpha{:}\sigma/\gamma\}$. Assume that $\alpha'{:}\sigma'/\gamma' \in D_E$. Hence either $\alpha' \notin E$, or $\alpha' \in E$ and $\neg\beta' \in E$ for no β' in σ'. In the first case, because γ is not the prerequisite of any default in D and E' is the closure of a set of literals under logical consequence, $\alpha' \notin E'$. In the second case, $\gamma' \in E$, and hence $\gamma' \neq \overline{\gamma}$. Because E' is the closure of a set of literals under logical consequence and σ' is the empty sequence or it consists of γ', $\neg\beta' \in E'$ for no β' in σ'. Hence $\alpha'{:}\sigma'/\gamma' \in D'_{E'}$ and $D_E \subseteq D'_E$. That $D'_{E'} \subseteq D_E \cup \{\alpha{:}\sigma/\gamma\}$ is immediate because $E \subseteq E'$. Because $\neg\gamma \notin E'$, $\alpha{:}\sigma/\gamma \in D'_{E'}$. Hence $D'_{E'} = D_E \cup \{\alpha{:}\sigma/\gamma\}$.

By Theorem 2.1.3 and Lemma A.1.5 $E = Cn_W(D_E)$. We show that $E' = Cn_W(D'_{E'})$. Let $C_0, C_1, \ldots$ be the sets in the definition of $Cn_W(D^E)$ in Theorem 2.1.3, and $C'_0, C'_1, \ldots$ the sets for $Cn_W(D'_{E'})$.

Induction hypothesis: $C_i \subseteq C'_i \subseteq C_i \cup \{\gamma\}$. *Base case* $i = 0$: $C_0 = W = C'_0$. *Inductive case* $i \geq 1$: Assume that $\gamma' \in C_i$. If $\gamma' \in C_{i-1}$, then by induction hypothesis $\gamma' \in C'_{i-1} \subseteq C'_i$. If $\gamma' \in C_i \backslash C_{i-1}$, then there is $\alpha'{:}\sigma'/\gamma' \in D_E$ such that $C_{i-1} \models \alpha'$. By induction hypothesis $C_{i-1} \subseteq C'_{i-1}$, and because $D_E \subseteq D'_{E'}$, $\gamma' \in C'_i$ by definition. Then we show the second inclusion. Assume that $\gamma' \in C'_i$. If $\gamma' \in C'_{i-1}$, then by induction hypothesis $\gamma' \in C_{i-1} \cup \{\gamma\} \subseteq C_i \cup \{\gamma\}$. If $\gamma' \in C'_i \backslash C'_{i-1}$, then $C'_{i-1} \models \alpha'$ for some $\alpha'{:}\sigma'/\gamma' \in D'_{E'}$. If $\gamma' = \gamma$, then clearly $\gamma \in C_i \cup \{\gamma\}$. Otherwise $\alpha'{:}\sigma'/\gamma' \in D_E$. Because by induction hypothesis $C'_{i-1} = C_{i-1} \cup \{\gamma\}$, and C'_{i-1} is a set of literals and $C'_{i-1} \models \alpha'$,

$\alpha' \in C'_{i-1}$. Because γ is not the prerequisite of any default in D, $\alpha' \in C_{i-1}$. Hence $C_{i-1} \models \alpha'$ and $\gamma' \in C_i$ by definition.

Because $\alpha \in E$, $\alpha \in C_j \subseteq C'_j$ for some $j \geq 0$. Therefore by definition $\gamma \in C'_{j+1} \subseteq \bigcup_{i\geq 0} Cn(C'_i)$. Therefore $E' = Cn_W(D'_{E'})$ and $Cn(E \cup \{\gamma\})$ is an extension of $\langle D \cup \{\alpha{:}\sigma/\gamma\}, W\rangle$. □

A.2 Horn clauses

The following lemmata on propositional Horn clauses are used for example in Section 7.2.3. We define Horn clauses as formulae $p_1 \wedge \cdots \wedge p_n \rightarrow l$ where p_i are propositional variables and l is a literal. The first lemma states an elementary property of models of sets of Horn clauses.

Lemma A.2.1 *Let S be a consistent set of Horn clauses. Let M be a model such that for all propositional variables p, $M \models p$ if and only if $M' \models p$ for all models M' such that $M' \models S$. Then $M \models S$ and for all propositional variables p, $M \models p$ if and only if $S \models p$.*

Proof: First we show that $M \models p$ if and only if $S \models p$. Let p be any propositional variable. If $S \models p$, then p is true in all models M' such that $M' \models S$ by definition of logical consequence. Hence $M \models p$. Assume that $S \not\models p$. Then there is a model M' such that $M' \models S$ and $M' \not\models p$. Hence $M \not\models p$.

Then we show that $M \models S$. Let $\phi = p_1 \wedge \cdots \wedge p_n \rightarrow p$ be a member of S. Assume that $p_1, \ldots, p_n$ are true in all models of S. Hence p is true in all models of S. By definition $M \models p$ and hence $M \models \phi$. Otherwise one of $p_1, \ldots, p_n$ is false in some model of S. By definition $M \not\models p_i$ for some $i \in \{1, \ldots, n\}$ and hence $M \models \phi$. Let $\phi = p_1 \wedge \cdots \wedge p_{n-1} \rightarrow \neg p_n$ be a member of S. Hence in every model of S at least one of $p_1, \ldots, p_n$ is false. By definition at least one of them is false in M, and consequently $M \models \phi$. □

Lemma A.2.2 *Let β be a literal and S a consistent set of literals β' and implications $q \rightarrow \beta'$ such that q is a propositional variable and β' is a literal and $\beta \neq \neg q$. If $S \models \beta$, then $\beta \in S$ or $S \models p$ for some $p \rightarrow \beta \in S$.*

Proof: Assume that $S \models \beta$ and $\beta \notin S$. Assume that for all p such that $p \rightarrow \beta \in S$, $S \not\models p$, that is, there is a model M such that $M \models S$ and $M \not\models p$. Because S is a set of Horn clauses, by Lemma A.2.1 there is a model M' such that $M' \models S$ and $M' \not\models p$ for all $p \rightarrow \beta \in S$. Now modify M' to M'' such that $M'' \not\models \beta$ and $M'' \models q$ iff $M' \models q$ for all propositional variables q that do not occur in β. We show that $M'' \models S$. By assumption $\beta \notin S$. If $\neg\beta \in S$, then $M'' \models \neg\beta$ by definition of M''. If β does not occur in $\phi \in S$, then clearly $M'' \models \phi$. If $p \rightarrow \overline{\beta} \in S$ for some p, then $M'' \models p \rightarrow \overline{\beta}$ because $M'' \models \overline{\beta}$. By assumption $M' \not\models p$ for all $p \rightarrow \beta \in S$, and hence $M'' \models p \rightarrow \beta$. If $\beta \rightarrow \beta' \in S$, then $M'' \models \beta \rightarrow \beta'$ because $M'' \not\models \beta$. This exhausts all formulae in S, and therefore $M'' \models S$. This contradicts the assumption $S \models \beta$ because $M'' \not\models \beta$, and hence the assumption that $S \not\models p$ for all p such that $p \rightarrow \beta \in S$ is false. Therefore $\beta \in S$ or $S \models p$ for some $p \rightarrow \beta \in S$, if $S \models \beta$. □

Lemma A.2.3 *Let $S \cup \{p \rightarrow \beta\}$ be a consistent set of literals and formulae $p' \rightarrow \beta'$ such that p' is a propositional variable and β' is a literal and $p' \neq \beta$. Then if $S \not\models q$ and $S \cup \{p \rightarrow \beta\} \models q$ for a propositional variable q, then $q = \beta$.*

Proof: Because S is a consistent set of Horn clauses, by Lemma A.2.1 there is a model M such that for propositional variables q, $M \models q$ if and only if $M' \models q$ for all models M' of S, and $S \models q$ if and only if $M \models q$. If $M \models p \rightarrow \beta$, there is no propositional variable q such that $S \not\models q$ and $S \cup \{p \rightarrow \beta\} \models q$. Otherwise $M \models p$ and $M \not\models \beta$. Because $M \models p$, $S \models p$. Similarly, β has to be a propositional variable because otherwise $M \models b$ and $\beta = \neg b$ and $S \models p \wedge \neg\beta$ and $S \cup \{p \rightarrow \beta\}$ would be inconsistent contrary to our assumptions. Let M' be a model such that $M' \models \beta$ and M' interprets other propositionla variables like M does. We show that $M' \models S \cup \{p \rightarrow \beta\}$. Because $M' \models \beta$, $M' \models q \rightarrow \beta$ for any $q \rightarrow \beta \in S \cup \{p \rightarrow \beta\}$ and propositional variable q. If $\phi \in S$ and there are no occurrences of β in ϕ, then clearly $M' \models \phi$. It is not possible that $\beta \in S$ because $M \not\models \beta$ and $M \models S$. It is not possible that $\neg\beta \in S$, because $S \models p$ and $S \cup \{p \rightarrow \beta\}$ is consistent. If $p' \rightarrow \neg\beta \in S$, then $M \not\models p'$ because $S \models p$ and $S \cup \{p \rightarrow \beta\}$ is consistent. There are no formulae $\beta \rightarrow \beta' \in S$ because by assumption for all $p' \rightarrow \beta' \in S$, $p' \neq \beta$. This exhausts all members of S, and therefore $M' \models S \cup \{p \rightarrow \beta\}$. Therefore $q = \beta$ is the only propositional variable such that $S \not\models q$ and $S \cup \{p \rightarrow \beta\} \models q$. □

Index