Preferential Reasoning for Modal Logics

Katarina Britz   Thomas Meyer   Ivan Varzinczak

Centre for Artificial Intelligence Research
CSIR Meraka Institute and University of KwaZulu-Natal, South Africa

Abstract

Modal logic is the foundation for a versatile and well-established class of knowledge representation formalisms in artificial intelligence. Enriching modal logics with non-monotonic reasoning capabilities such as preferential reasoning as developed by Lehmann and colleagues would therefore constitute a natural extension of such KR formalisms. Nevertheless, there is at present no generally accepted semantics, with corresponding syntactic characterization, for preferential consequence in modal logics. In this paper we fill this gap by providing a natural and intuitive semantics for preferential and rational modal consequence. We do so by placing a preference order on possible worlds indexed by Kripke models they belong to. We also prove representation results for both preferential and rational consequence, which paves the way for effective decision procedures for modal preferential reasoning. We then illustrate applications of our constructions to modal logics widely used in AI, notably in the contexts of reasoning about actions, knowledge and beliefs. We argue that our semantics constitutes the foundation on which to explore preferential reasoning in modal logics in general.

Keywords: non-monotonic reasoning, preference order, rational closure, actions, beliefs

1   Introduction and Motivation

Consider the following scenario depicting a nuclear power station [3]: In a particular power plant there is an atomic pile and a cooling system, both of which can be either on or off. An agent is in charge of detecting hazardous situations and preventing the plant from malfunctioning (Figure 1).
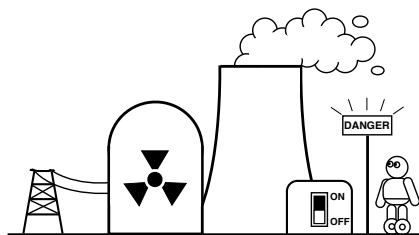

Fig. 1. The nuclear power station and its controlling agent.

It is common knowledge that situations in which the pile is on are usually not hazardous. However, in more specific contexts, say when the pile is on but the cooling system is down, one would expect it to be a hazardous situation. One

may then also want to draw conclusions like "if a situation is hazardous, then it is usually the case that the *effect* of switching the pile off brings about a non-hazardous situation"; or "if the pile is on and the cooling system is off, then usually the surveillance agent *knows* that a malfunction is imminent"; or "situations in which the pile is on usually *ought* to be non-hazardous"; or even the more complex "if the agent *believes* that there is danger, then usually he *must* perform the *action* of switching the pile off".

All of these are examples of defeasible inferences. The first two examples are instances of propositional defeasible consequence and are adequately dealt with within the framework for preferential reasoning developed by Lehmann and colleagues in the 90's [17,20]. For the last four examples, however, it is a different story: Firstly, for their specification we need a logical language that is richer than the propositional one: They require, respectively, the ability to express the effects and preconditions of *actions*; an agent's *knowledge* and *beliefs*; *regulations* or *obligations*; and combinations thereof. Well established formalisms for dealing with these notions in the AI literature are mostly variants of modal logic: The examples above illustrate applications of dynamic logic [14], epistemic logic [11], and deontic logic [22].

Secondly, it turns out that research on preferential reasoning has really only reached maturity in a propositional context, whereas many logics of interest, like the ones mentioned above, have more structure in both syntax and semantics. If one wants to be able to capture the forms of reasoning exemplified above, then one has to move beyond propositional preferential consequence.

There has by now been quite a substantial number of attempts to incorporate defeasible reasoning in logics other than propositional logic. After a first tentative exploration of preferential predicate logics by Lehmann and Magidor [19], some more recent investigations have attempted to define notions of defeasibility in deontic logics [23], and of defeasible subsumption for description logics [4,13,6]. Nevertheless, a generally accepted semantics for preferential reasoning in modal logics, with a corresponding syntactic characterization, does not yet exist.

In the present paper we aim at filling this gap by providing a generalization of classical preferential consequence to an important family of modal logics — we present the semantic foundation, prove the required representation results, point out computational consequences and benefits, and suggest applications of modal preferential reasoning. The good balance between expressivity and computational properties of modal logics makes them good candidates for the type of extension of preferential reasoning that we have in mind — with (propositional) modal logic one can express more than with classical propositional logic without being hampered by the undecidability of many first-order based languages.

The remainder of this paper is organized as follows: In Section 2 we briefly recap the seminal work by Lehmann and colleagues on propositional preferential and rational consequence, and in Section 3 we provide the basic modal logic terminology. We then present an account of modal preferential reasoning in Section 4. Importantly, our representation results are with respect to the corresponding *propositional* properties or rules, and methods employed in a propositional non-monotonic setting therefore translate seamlessly to a modal context. This includes reasoning tasks such as computing the preferential or rational closure of a defeasible knowledge base [20].

Our modal semantics therefore forms the foundation of preferential consequence for a whole class of modal-based formalisms, which we illustrate with some case studies, namely modal logics of action (Section 5) and knowledge (Section 6). In Section 7 we propose one avenue for future work, viz. defeasible accessibility. After discussing some related work (Section 8), we conclude in Section 9.

2 Propositional Preferential Reasoning

Here we give a brief outline of propositional preferential and rational consequence, as initially defined by Kraus et al. [17]. A propositional defeasible consequence relation $\mathrel{|\!\sim}$ is defined as a binary relation on formulas of an underlying (possibly infinitely generated) propositional logic. $\mathrel{|\!\sim}$ is said to be *preferential* if it satisfies the following set of properties (below $\alpha, \beta, \ldots$ denote propositional formulas, and $\models$ and $\equiv$ denote, respectively, propositional entailment and logical equivalence):

(Ref) $\alpha \mathrel{|\!\sim} \alpha$
$\qquad$ (And) $\dfrac{\alpha \mathrel{|\!\sim} \beta, \ \ \alpha \mathrel{|\!\sim} \gamma}{\alpha \mathrel{|\!\sim} \beta \wedge \gamma}$
$\qquad$ (Or) $\dfrac{\alpha \mathrel{|\!\sim} \gamma, \ \ \beta \mathrel{|\!\sim} \gamma}{\alpha \vee \beta \mathrel{|\!\sim} \gamma}$

(LLE) $\dfrac{\alpha \equiv \beta, \ \ \alpha \mathrel{|\!\sim} \gamma}{\beta \mathrel{|\!\sim} \gamma}$
$\qquad$ (RW) $\dfrac{\alpha \mathrel{|\!\sim} \beta, \ \ \beta \models \gamma}{\alpha \mathrel{|\!\sim} \gamma}$
$\qquad$ (CM) $\dfrac{\alpha \mathrel{|\!\sim} \beta, \ \ \alpha \mathrel{|\!\sim} \gamma}{\alpha \wedge \beta \mathrel{|\!\sim} \gamma}$

The semantics of (propositional) preferential consequence relations is in terms of *preferential models*; these are partially ordered structures with states labeled by propositional valuations. We shall make this terminology more precise in the upcoming section on modal preferential consequence, but it essentially allows for a partial order on states, with states lower down in the order being more preferred than those higher up. Given a preferential model $\mathscr{P}$, a pair $\alpha \mathrel{|\!\sim} \beta$ is in the consequence relation defined by $\mathscr{P}$ if and only if the most preferred α-states are also β-states. The representation theorem for preferential consequence relations then states:

Theorem 2.1 (Kraus et al. [17]) *A defeasible consequence relation is a preferential consequence relation if and only if it is defined by some preferential model.*

If, in addition to the properties of preferential consequence, $\mathrel{|\!\sim}$ also satisfies the following Rational Monotony property, it is said to be a *rational* consequence relation:

(RM) $\dfrac{\alpha \mathrel{|\!\sim} \beta, \ \ \alpha \mathrel{|\!\not\sim} \neg\gamma}{\alpha \wedge \gamma \mathrel{|\!\sim} \beta}$

The semantics of rational consequence relations is in terms of *ranked* preferential models, i.e., preferential models in which the preference order is *modular*:

Definition 2.2 Given a set S, $\prec \subseteq S \times S$ is *modular* if and only if $\prec$ is a partial order and there is a ranking function $rk : S \longrightarrow \mathbb{N}$ such that for every $s, s' \in S$, $s \prec s'$ if and only if $rk(s) < rk(s')$.

The representation theorem for rational consequence relations then states:

Theorem 2.3 (Lehmann & Magidor [20]) *A defeasible consequence relation is a rational consequence relation if and only if it is defined by some ranked model.*

3 Modal Logic

We work in a set of *atomic propositions* $\mathcal{P}$ (together with the distinguished atom $\bot$), using the logical connectives $\wedge$, $\neg$, and a set of modal operators $\Box_i$, $1 \leq i \leq n$. Propositions are denoted by $p, q, \ldots$, and formulas by $\alpha, \beta, \ldots$, constructed in the usual way. With $\mathcal{L}$ we denote the set of all formulas of the modal language.

The semantics is the standard possible-worlds one:

Definition 3.1 A *Kripke model* is a tuple $\mathscr{M} = \langle W, R, V \rangle$ where W is a set of *possible worlds*, $R = \langle R_1, \ldots, R_n \rangle$, where each $R_i \subseteq W \times W$ is an *accessibility relation* on W, $1 \leq i \leq n$, and $V \colon W \longrightarrow 2^{\mathcal{P}}$ is a *valuation function*.

Definition 3.2 Given $\mathscr{M} = \langle W, R, V \rangle$ and $w \in W$:

- $\mathscr{M}, w \nVdash \bot$;
- $\mathscr{M}, w \Vdash p$ if and only if $p \in V(w)$;
- $\mathscr{M}, w \Vdash \alpha \wedge \beta$ if and only if $\mathscr{M}, w \Vdash \alpha$ and $\mathscr{M}, w \Vdash \beta$;
- $\mathscr{M}, w \Vdash \neg\alpha$ if and only if $\mathscr{M}, w \nVdash \alpha$;
- $\mathscr{M}, w \Vdash \Box_i\alpha$ if and only if $\mathscr{M}, w' \Vdash \alpha$ for all w' such that $(w, w') \in R_i$.

Given $\alpha \in \mathcal{L}$ and $\mathscr{M} = \langle W, R, V \rangle$, we say that $\mathscr{M}$ is a *model of* α if and only if $\mathscr{M}, w \Vdash \alpha$ for every $w \in W$. We say that $\mathscr{M}$ *satisfies* α if there is at least one $w \in W$ such that $\mathscr{M}, w \Vdash \alpha$. Given a set of formulas $\mathcal{K}$, $\mathscr{M}$ is a model of $\mathcal{K}$ if and only if $\mathscr{M}$ is a model of every formula in $\mathcal{K}$. A formula α is (globally) *entailed* by $\mathcal{K}$, denoted $\mathcal{K} \models \alpha$, if and only if every model of $\mathcal{K}$ is also a model of α.

For more details on modal logic, we refer the reader to the handbook by Blackburn, van Benthem and Wolter [2].

4 Modal Preferential Consequence

In this section we present the semantic foundation of preferential reasoning for modal and multi-modal logics. We interpret the defeasible consequence relation $\mathrel|\!\sim$ as a relation on formulas of a given underlying modal language.

Definition 4.1 A defeasible consequence relation $\mathrel|\!\sim \; \subseteq \mathcal{L} \times \mathcal{L}$ is a *preferential consequence relation* if and only if it satisfies the properties (Ref), (LLE), (And), (RW), (Or), and (CM), with propositional entailment replaced by modal entailment in the corresponding modal system. $\mathrel|\!\sim$ is *rational* if and only if in addition to being preferential it also satisfies the property (RM).

Up until now it has not been clear how to best generalize the propositional semantics for preferential reasoning to the more expressive cases. We propose replacing the propositional valuations in preferential models with model-world pairs, also sometimes referred to as pointed Kripke models.

Given a particular class (set) of Kripke models $\mathcal{M}$ as in Definition 3.1, let $\mathscr{U}_{\mathcal{M}}$ denote the set of all pairs $(\mathscr{M}, w)$, where $\mathscr{M} = \langle W, R, V \rangle$ is a Kripke model in $\mathcal{M}$ and $w \in W$. As alluded to above, our semantics replaces the propositional valuations used in the preferential models of Lehmann et al. with structures over the set $\mathscr{U}_{\mathcal{M}}$. We make this more precise now. (We shall follow the notation and terminology

defined by Lehmann and colleagues. For more details, the reader is referred to the papers by Kraus et al. [17] and Lehmann and Magidor [20].)

Let S be a set, the elements of which are called *states*. Let $\mathcal{M}$ be a given class of Kripke models, and $\ell : S \longrightarrow \mathscr{U}_{\mathcal{M}}$ be a *labeling function* mapping every state to a pair $(\mathscr{M}, w)$ where $\mathscr{M} = \langle W, R, V \rangle$ is a Kripke model such that $w \in W$. Let $\prec \subseteq S \times S$. Given $\alpha \in \mathcal{L}$, we say that $s \in S$ *satisfies* α if and only if $\ell(s) \Vdash \alpha$, i.e., $\mathscr{M}, w \Vdash \alpha$, where $\ell(s) = (\mathscr{M}, w)$. We define $\widehat{\alpha} = \{ s \in S \mid \ell(s) \Vdash \alpha \}$. We say that $\widehat{\alpha}$ is *smooth* if and only if each $s \in \widehat{\alpha}$ is either $\prec$-minimal in $\widehat{\alpha}$, or there is $s' \in \widehat{\alpha}$ such that $s' \prec s$ and s' is $\prec$-minimal in $\widehat{\alpha}$. We say that S satisfies the smoothness condition if and only if for every $\alpha \in \mathcal{L}$, $\widehat{\alpha}$ is smooth. We can now define modal preferential models:

Definition 4.2 Let $\mathcal{M}$ be a class of Kripke models. A *preferential model* is a triple $\mathscr{P} = \langle S, \ell, \prec \rangle$ where S is a set of states satisfying the smoothness condition, ℓ is a labeling function mapping states to elements of $\mathscr{U}_{\mathcal{M}}$, and $\prec$ is a strict partial order on S.

These formal constructions closely resemble those of Kraus et al. [17] and of Lehmann and Magidor [20], the difference being that propositional valuations are replaced with elements of the set $\mathscr{U}_{\mathcal{M}}$. In Section 5 we provide an extended example including an instance of a modal preferential model.

Definition 4.3 Given $\alpha, \beta \in \mathcal{L}$ and a preferential model $\mathscr{P} = \langle S, \ell, \prec \rangle$, the consequence relation defined by $\mathscr{P}$ is denoted by $\mathrel|\!\!\sim_{\mathscr{P}}$ and is defined by: $\alpha \mathrel|\!\!\sim_{\mathscr{P}} \beta$ if and only if every $\prec$-minimal state $s \in \widehat{\alpha}$ is such that $s \in \widehat{\beta}$.

We say that a modal preferential model $\mathscr{P}$ *satisfies* the defeasible statement $\alpha \mathrel|\!\!\sim \beta$ if and only if $\alpha \mathrel|\!\!\sim_{\mathscr{P}} \beta$ holds.

We are now in a position to state one of our central results:

Theorem 4.4 *A modal defeasible consequence relation is a preferential consequence relation if and only if it is defined by some preferential model.*

Proof. The proof is outlined in Appendix A. $\qquad\qquad\square$

The significance of this is that the representation result is proved with respect to the same set of properties used to characterize propositional preferential consequence. We therefore argue that our definition of preferential models provide the foundation for a semantics for preferential (and rational) consequence for a whole class of multi-modal logics. We do not claim that this is *the* appropriate notion of preferential consequence for all modal logics, but rather that it describes the basic framework within which to investigate such notions.

In order to obtain a similar result for rational consequence, we restrict ourselves to those preferential models in which $\prec$ is a modular order on states (cf. Definition 2.2):

Definition 4.5 A *ranked model* $\mathscr{P}_r$ is a preferential model $\langle S, \ell, \prec \rangle$ in which the strict partial order $\prec$ on S is modular.

Since ranked models are preferential models, the notion of rational consequence is as in Definition 4.3. We can then state the following result:

Theorem 4.6 *A defeasible consequence relation is a rational consequence relation if and only if it is defined by some ranked model.*

Proof. The proof is outlined in Appendix B. □

One of the primary reasons for defining non-monotonic consequence relations of the kind we have presented above is to obtain a semantic notion, with corresponding proof-theoretic decision procedure, of *defeasible entailment* [20]: Given a set of defeasible statements of the form $\alpha \mathrel{|\sim} \beta$, where $\mathrel{|\sim}$ is now viewed as a connective in an enriched modal language, which other defeasible statements should one be able to derive from this?

Definition 4.7 A modal *defeasible knowledge base* $\mathcal{K}^{|\sim}$ is a finite set of statements $\alpha \mathrel{|\sim} \beta$, where $\alpha, \beta \in \mathcal{L}$.

A preferential model $\mathscr{P}$ satisfies a defeasible knowledge base $\mathcal{K}^{|\sim}$ if and only if $\mathscr{P}$ satisfies every statement in $\mathcal{K}^{|\sim}$, i.e., for every $\alpha \mathrel{|\sim} \beta \in \mathcal{K}^{|\sim}$, $\alpha \mathrel{|\sim_{\mathscr{P}}} \beta$ holds.

We define modal preferential entailment in the obvious way:

Definition 4.8 $\alpha \mathrel{|\sim} \beta$ is *preferentially entailed* by $\mathcal{K}^{|\sim}$ if and only if $\alpha \mathrel{|\sim_{\mathscr{P}}} \beta$ for every preferential model $\mathscr{P}$ satisfying $\mathcal{K}^{|\sim}$. The set of all defeasible statements preferentially entailed by $\mathcal{K}^{|\sim}$ is called the modal *preferential closure* of $\mathcal{K}^{|\sim}$.

The modal preferential closure of $\mathcal{K}^{|\sim}$ coincides with the intersection of all preferential consequence relations containing $\mathcal{K}^{|\sim}$, and is a preferential consequence relation. Modal rational entailment can be defined in a similar way:

Definition 4.9 $\alpha \mathrel{|\sim} \beta$ is *rationally entailed* by $\mathcal{K}^{|\sim}$ if and only if $\alpha \mathrel{|\sim_{\mathscr{P}_r}} \beta$ for every ranked model $\mathscr{P}_r$ satisfying $\mathcal{K}^{|\sim}$.

However, as in the propositional case [20], the obvious definition of modal rational closure based on modal rational entailment does not produce an appropriate result.

Theorem 4.10 *Given a defeasible knowledge base $\mathcal{K}^{|\sim}$, the set of defeasible statements rationally entailed by $\mathcal{K}^{|\sim}$ is exactly the preferential closure of $\mathcal{K}^{|\sim}$.*

Proof. The proof is similar to the propositional case [20, Theorem 4.2]. □

The main consequence of this result is that modal rational entailment does not, in general, produce a consequence relation which is rational.

The argument to define and construct a viable notion of rational closure is analogous to that given by Lehmann and Magidor [20] in the propositional case. First, we define a preference ordering on consequence relations, with relations further down in the ordering interpreted as more preferred.

Definition 4.11 Let $\mathcal{K}^{|\sim}$ be a defeasible knowledge base. The preference order $\ll$ generated by $\mathcal{K}^{|\sim}$ is a binary relation on the set of rational consequence relations containing $\mathcal{K}^{|\sim}$, defined as follows: $\mathrel{|\sim_0}$ is *preferable to* $\mathrel{|\sim_1}$ (written $\mathrel{|\sim_0} \ll \mathrel{|\sim_1}$) if and only if

- there is $(\alpha, \beta) \in \mathrel{|\sim_1} \setminus \mathrel{|\sim_0}$ such that for all γ such that $\gamma \vee \alpha \mathrel{|\sim_0} \neg\alpha$ and for all δ such that $\gamma \mathrel{|\sim_0} \delta$, we also have $\gamma \mathrel{|\sim_1} \delta$; and

- for every $\gamma, \delta \in \mathcal{L}$, if $\gamma \mathrel{|\!\sim} \delta$ is in $\mathrel{|\!\sim}_0 \setminus \mathrel{|\!\sim}_1$, then there is an assertion $\rho \mathrel{|\!\sim} \nu$ in $\mathrel{|\!\sim}_1 \setminus \mathrel{|\!\sim}_0$ such that $\rho \vee \gamma \mathrel{|\!\sim}_1 \neg\gamma$.

Given a defeasible knowledge base $\mathcal{K}^{|\!\sim}$, the idea is to define modal rational closure as the most preferred (with respect to $\ll$) of all the rational consequence relations containing $\mathcal{K}^{|\!\sim}$. This leads to the following important definition of modal rational closure for defeasible knowledge bases:

Definition 4.12 Let $\mathcal{K}^{|\!\sim}$ be a defeasible knowledge base, let $\mathcal{K}^R$ be the class of rational consequence relations containing $\mathcal{K}^{|\!\sim}$, and let $\ll$ be the preference ordering on $\mathcal{K}^R$ generated by $\mathcal{K}^{|\!\sim}$. If $\ll$ has a (unique) minimum element $\mathrel{|\!\sim}$, then the modal rational closure of $\mathcal{K}^{|\!\sim}$ is defined as $\mathrel{|\!\sim}$.

Clearly then, if the modal rational closure exists, it is a rational consequence relation. In order to provide the conditions for the existence of modal rational closure, we first need to define a *ranking* of formulas with respect to $\mathcal{K}^{|\!\sim}$ which, in turn, is based on a notion of *exceptionality*. The ranking of formulas can also be used to define an algorithm for computing the modal rational closure, when it exists (see below and the discussion in Section 8).

A formula α is said to be *exceptional* for a defeasible knowledge base $\mathcal{K}^{|\!\sim}$ if and only if $\mathcal{K}^{|\!\sim}$ *preferentially* entails $\top \mathrel{|\!\sim} \neg\alpha$. A defeasible statement $\alpha \mathrel{|\!\sim} \beta$ is exceptional for $\mathcal{K}^{|\!\sim}$ if and only if its antecedent α is exceptional for $\mathcal{K}^{|\!\sim}$.

It turns out that checking for exceptionality can be reduced to classical modal entailment checking.

Lemma 4.13 *Given a defeasible knowledge base $\mathcal{K}^{|\!\sim}$, let $\mathcal{K}^{\to}$ be its classical counterpart in which every defeasible statement of the form $\gamma \mathrel{|\!\sim} \delta$ in $\mathcal{K}^{|\!\sim}$ is replaced by $\gamma \to \delta$. The formula α is exceptional for $\mathcal{K}^{|\!\sim}$ if and only if $\neg\alpha$ is globally entailed by $\mathcal{K}^{\to}$: i.e., if and only if $\mathcal{K}^{\to} \models \neg\alpha$.*

Proof. The proof is similar to the propositional case [20, Corollary 5.22]. $\square$

Let $E(\mathcal{K}^{|\!\sim})$ denote the subset of $\mathcal{K}^{|\!\sim}$ containing statements that are exceptional for $\mathcal{K}^{|\!\sim}$. We define a non-increasing sequence of subsets of $\mathcal{K}^{|\!\sim}$ as follows: $\mathcal{E}_0 = \mathcal{K}^{|\!\sim}$, and for $i > 0$, $\mathcal{E}_i = E(\mathcal{E}_{i-1})$. Clearly there is a smallest integer k such that for all $j \geq k$, $\mathcal{E}_j = \mathcal{E}_{j+1}$. From this we define the *rank* of a formula with respect to $\mathcal{K}^{|\!\sim}$ as follows: $r_{\mathcal{K}^{|\!\sim}}(\alpha) = k - i$, where i is the smallest integer such that α is not exceptional for $\mathcal{E}_i$. If α is exceptional for $\mathcal{E}_k$ (and therefore exceptional for all $\mathcal{E}$s), then $r_{\mathcal{K}^{|\!\sim}}(\alpha) = 0$. Intuitively, the lower the rank of a formula, the more exceptional it is with respect to the defeasible knowledge base $\mathcal{K}^{|\!\sim}$.

We are now in a position to provide the conditions for the existence of modal rational closure.

Theorem 4.14 *Let $\mathcal{K}^{|\!\sim}$ be a finite defeasible knowledge base. The rational closure of $\mathcal{K}^{|\!\sim}$ exists and is the set $\mathcal{R}^{|\!\sim}$ of defeasible statements $\alpha \mathrel{|\!\sim} \beta$ such that either $r_{\mathcal{K}^{|\!\sim}}(\alpha) > r_{\mathcal{K}^{|\!\sim}}(\alpha \wedge \neg\beta)$, or $r_{\mathcal{K}^{|\!\sim}}(\alpha) = 0$ (in which case $r_{\mathcal{K}^{|\!\sim}}(\alpha \wedge \neg\beta) = 0$ as well).*

Proof. The proof is outlined in Appendix C. $\square$

Observe from Lemma 4.13 that the determination of the rank of a formula can be reduced to classical (global) entailment checking for the modal logic under consid-

eration. It is therefore easy to construct a (naïve) decidable algorithm to determine whether a given defeasible statement is in the modal rational closure of a finite defeasible knowledge base $\mathcal{K}^{\mid\sim}$.

Recent work by Casini and Straccia [6] in the context of description logics makes it possible to improve on this algorithm, from an implementation point of view. Casini and Straccia proposed a syntactic operational characterization of rational closure for the description logic $\mathcal{ALC}$, but without providing a semantic counterpart. The algorithm draws on work by Lehmann and Magidor [20], Freund [12] and Poole [24], is based on performing a number of classical entailment checks (for $\mathcal{ALC}$), and is easily amenable to implementation. From their description it is clear that the number of classical entailment checks needed is quadratic in the size of the knowledge base under consideration.

The significance of their work for us is that their algorithm can readily be adjusted to determine the modal rational closure of a defeasible knowledge base $\mathcal{K}^{\mid\sim}$, with rational closure membership checking being reduced to a number of modal global entailment checks that is quadratic in the size of $\mathcal{K}^{\mid\sim}$. For the modal logics we consider here, determining membership of the modal rational closure is therefore no harder than global entailment checking.

In what follows we analyze applications of our constructions to some classes of modal logics commonly used in AI. A further application area, in the context of the description logic $\mathcal{ALC}$, was recently explored by the present authors [5].

5 Preferential Reasoning about Actions

There is in the AI literature a fair number of modal-based formalisms for reasoning about actions and change [9,8,27,10]. These are essentially variants of the multi-modal logic K.

In these modal logics of action, the modal operators are determined by a (finite) set of *actions* $\mathcal{A} = \{a_1, \ldots, a_n\}$: For each $a \in \mathcal{A}$, there is associated a modal operator $[a]$. Formulas are recursively defined in the usual way, and the semantics is as in Definition 3.2. Formulas of the form $[a]\alpha$ are used to specify the *effects* of actions and they read as "after *every* execution of action a, the formula α holds". The operator $\langle a \rangle$ is mostly used to specify the *executability* of actions: $\langle a \rangle\top$ reads as "there is a possible execution of action a".

In our nuclear power plant example, we have the atoms $\mathcal{P} = \{p, c, h\}$, where p stands for "the atomic pile is on", c for "the cooling system is on", and h for "hazardous situation". As for the actions, we have $\mathcal{A} = \{f, m\}$, where f stands for "flipping the pile switch", and m for "malfunction". One possible specification of such a scenario is given by the following knowledge base:

$$\mathcal{K} = \{(p \wedge \neg c) \leftrightarrow h,\ h \rightarrow \langle m \rangle\top,\ p \rightarrow [f]\neg p,\ \langle f \rangle\top\}$$

(We note that specifying a solution to the frame problem is beyond the scope of this paper. Instead, we refer the reader to the solutions provided by the above-mentioned frameworks which can, in principle, be integrated into the present formalism in a straightforward way.) $\mathcal{K}$ basically says that "a hazardous situation is one in which

the pile is on and the cooler off", "a hazardous situation may lead to a malfunction", "if the pile is on, then flipping switches it off", and "one can always flip the pile switch". We can then conclude $\mathcal{K} \models \mathsf{p} \rightarrow [\mathsf{f}]\neg\mathsf{h}$, $\mathcal{K} \models [\mathsf{m}]\bot \rightarrow (\neg\mathsf{p} \vee \mathsf{c})$, and $\mathcal{K} \not\models [\mathsf{m}]\langle\mathsf{f}\rangle\neg\mathsf{p} \vee [\mathsf{f}]\mathsf{h}$.

Recalling our discussion in the Introduction, with such a specification one cannot reason with *exceptionalities* given by statements of the form "a situation in which the pile is on is usually not hazardous" and "a situation in which the pile is on and the cooling system is off is usually hazardous". Thanks to our constructions from the previous section, we can move to a defeasible version of a modal logic of actions in which one can capture defeasible consequences.

As an example, assume we are given $\mathcal{M} = \{\mathscr{M}_1, \mathscr{M}_2\}$, where $\mathscr{M}_1$ and $\mathscr{M}_2$ are as depicted in Figure 2.
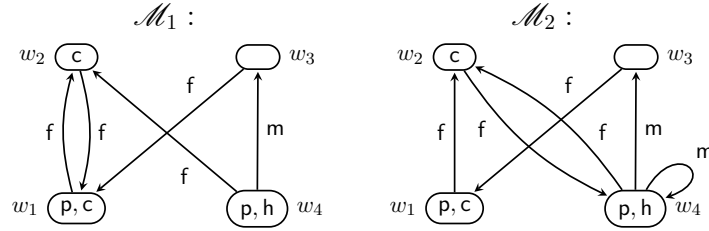

Fig. 2. Models depicting the behavior of actions in our nuclear power station scenario.

Hence $\mathscr{U}_{\mathcal{M}} = \{(\mathscr{M}_i, w_j) \mid i \in \{1, 2\}, j \in \{1, 2, 3, 4\}\}$. Assume $S = \{s_i \mid 1 \leq i \leq 8\}$, and let a labeling function ℓ be such that $\ell(s_1) = (\mathscr{M}_1, w_1)$, $\ell(s_2) = (\mathscr{M}_2, w_1)$, $\ell(s_3) = (\mathscr{M}_1, w_2)$, $\ell(s_4) = (\mathscr{M}_1, w_3)$, $\ell(s_5) = (\mathscr{M}_2, w_3)$, $\ell(s_6) = (\mathscr{M}_2, w_2)$, $\ell(s_7) = (\mathscr{M}_1, w_4)$, and $\ell(s_8) = (\mathscr{M}_2, w_4)$. The order $\prec$ is given by: $s_1 \prec s_3$, $s_2 \prec s_3$, $s_3 \prec s_4$, $s_3 \prec s_5$, $s_4 \prec s_6$, $s_5 \prec s_6$, $s_6 \prec s_7$, and $s_6 \prec s_8$. Figure 3 below depicts the preferential model $\mathscr{P} = \langle S, \ell, \prec \rangle$.
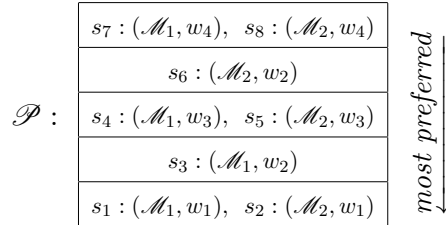

Fig. 3. Preferential model for the nuclear power plant scenario.

The rationale of this partial order is as follows: The utility company selling the electricity generated by the power plant tries as far as possible to keep both the pile and the cooling system on, ensuring that the pile can be easily switched off (states s_1 and s_2); sometimes the company has to switch the pile off for maintenance but then tries to keep the cooler running, preferably if turning the pile on again does not cause a fault in the cooling system (state s_3); more rarely the company needs to switch off both the pile and the cooler, e.g. when the latter needs maintenance (states s_4 and s_5); in an exceptional situation, turning the pile on may interfere with the cooler switching it off (state s_6); and, finally, only in very exceptional situations would the pile be on while the cooler is off, e.g. during a serious malfunction (states s_7 and s_8).

9

Let $\mathrel|\!\sim_{\mathscr{P}}$ be the defeasible consequence relation defined by the model $\mathscr{P}$ above (cf. Definition 4.3). Then we can check that $\mathsf{c} \wedge \langle\mathsf{f}\rangle\top \mathrel|\!\sim_{\mathscr{P}} [\mathsf{f}]\mathsf{c}$ (usually, flipping the switch does not shut down the cooler); $\neg\mathsf{p} \wedge \mathsf{c} \mathrel|\!\sim_{\mathscr{P}} [\mathsf{f}]\neg\mathsf{h}$ (turning the pile on, when the cooler is on, usually does not lead to a hazardous situation); $\top \mathrel|\!\sim_{\mathscr{P}} [\mathsf{m}]\bot$ (meltdowns are unlikely); $\mathsf{p}\wedge\neg\mathsf{c} \mathrel|\!\not\sim_{\mathscr{P}} [\mathsf{m}]\neg\mathsf{h}$ (it is not the case that once a malfunction has taken place in a hazardous situation, one is usually in a safe situation); and $\neg\mathsf{p} \wedge \mathsf{c} \mathrel|\!\not\sim_{\mathscr{P}} \langle\mathsf{f}\rangle\langle\mathsf{m}\rangle\top$ (it is not usually the case that turning the pile on may lead to a situation in which a malfunction is possible).

Moreover, thanks to our representation result (Theorem 4.4), we know that $\mathrel|\!\sim_{\mathscr{P}}$ is a preferential consequence relation (actually it turns out to be a *rational* consequence relation, since $\prec$ here is a modular partial order).

We started our example with a classical knowledge base, and then constructed a semantic rational consequence relation. This illustrated the semantics well, but in practice we may rather start with a defeasible knowledge base $\mathcal{K}^{\mathrel|\!\sim}$, from which we may draw defeasible conclusions. The semantics is then hidden from the user, as it forms part of the internal working of the reasoner. In the next section we analyze an example of reasoning with a defeasible knowledge base.

6 Epistemic Preferential Reasoning

Another family of modal logics that is of great interest from the standpoint of AI is that of epistemic logics, which allow for reasoning about knowledge [11].

The language of basic epistemic logic contains a (finite) set of *agents* $\mathcal{A} = \{A_1, \ldots, A_n\}$. For each agent $A \in \mathcal{A}$ there is a *knowledge operator* K_A. Formulas are recursively defined in the usual way with $\{K_{A_1}, \ldots, K_{A_n}\}$ as modal operators, and the semantics is as in Definition 3.2. A formula of the form $K_A\alpha$ is used to specify agent A's *knowledge* about the world and it is read as "agent A *knows* that α is the case". For instance, $K_{A_1}(\alpha \wedge K_{A_2}\alpha)$ formalizes the fact that "agent A_1 knows both α and that agent A_2 also knows α".

The core of epistemic logic is the normal multi-modal logic K_m. Hence, the following version of axiom schema K is valid: $K_A\alpha \wedge K_A(\alpha \to \beta) \to K_A\beta$, i.e., "if A knows both α and $\alpha \to \beta$, then she also knows β". Stronger epistemic logics are obtained by adding additional schemata, expressing specific desired properties of knowledge, to the basic system K. Since K is at the heart of these logics, we shall suffice with it in our exposition below.

In our running example, let us assume that we have two agents, say A and B. The set $\mathcal{P}$ is as in the previous section, with the propositions p, c and h keeping their previous intuition. We have the following defeasible knowledge base:

$$\mathcal{K}^{\mathrel|\!\sim} = \left\{ \begin{array}{c} \mathsf{h} \mathrel|\!\sim \mathsf{p} \wedge \neg\mathsf{c},\ \mathsf{p} \wedge \neg\mathsf{c} \wedge \neg\mathsf{h} \mathrel|\!\sim \bot,\ \mathsf{h} \mathrel|\!\sim K_\mathsf{A}(\mathsf{p} \wedge \neg\mathsf{c}),\ \mathsf{h} \wedge \mathsf{p} \wedge \neg\mathsf{c} \mathrel|\!\sim K_\mathsf{B}K_\mathsf{A}\neg\mathsf{c}, \\[2mm] K_\mathsf{B}\mathsf{p} \mathrel|\!\sim \mathsf{p} \wedge \mathsf{c},\ K_\mathsf{B}\mathsf{p} \wedge K_\mathsf{A}\neg\mathsf{p} \mathrel|\!\sim \neg\mathsf{p},\ K_\mathsf{B}\mathsf{p} \wedge K_\mathsf{B}\neg\mathsf{c} \mathrel|\!\sim \mathsf{p} \wedge \neg\mathsf{c} \end{array} \right\}$$

The intuition conveyed by $\mathcal{K}^{\mathrel|\!\sim}$ is as follows: Hazardous situations are usually ones in which the pile is on and the cooler is off. The statement $\mathsf{p} \wedge \neg\mathsf{c} \wedge \neg\mathsf{h} \mathrel|\!\sim \bot$ encodes the 'hard' constraint $\mathsf{p} \wedge \neg\mathsf{c} \to \mathsf{h}$ [20, p. 6]. Then we have that in hazardous

situations our agent A usually believes the pile is indeed on and the cooler is off. In hazardous situations where the pile is on and the cooler is off, agent B knows that agent A knows that the cooler is down. Then B is usually right about his belief that the pile is on (unless A believes the opposite, in which case it is off); moreover the cooler is also on (unless he knows otherwise, in which case it is off).

From the defeasible knowledge base $\mathcal{K}^{|\sim}$ we can preferentially derive $h \wedge p \wedge \neg c \mid\sim K_A \neg c \wedge K_B K_A \neg c$: in hazardous situations due to the pile being on while the cooler is off, usually A knows that the cooling is down and B knows that A knows this is the case. Since we know that the preferential closure is a preferential consequence relation (cf. comment after Definition 4.8), this can be derived by applying properties (RW), (CM) and (And) to the statements in $\mathcal{K}^{|\sim}$.

$\mathcal{K}^{|\sim}$ does *not* entail $K_B p \wedge K_B \neg c \wedge \neg K_A h \mid\sim h$ preferentially: situations in which B knows both that the pile is on and the cooler is off, and in which A does not know there is a hazard, are not usually hazardous. We (i.e., the underlying reasoner) can easily construct a modal preferential model satisfying $\mathcal{K}^{|\sim}$ that does not satisfy the statement above. Let $\mathscr{M}$ be the Kripke model in Figure 4.
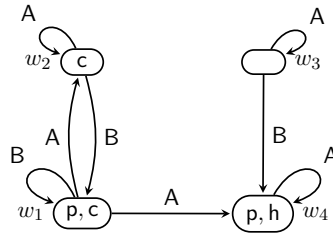

Fig. 4. A Kripke model depicting knowledge of two agents in our nuclear power station scenario.

From $\mathscr{M}$ we construct a modal preferential model $\mathscr{P} = \langle S, \ell, \prec \rangle$, where $S = \{s_1, s_2, s_3, s_4\}$, $\ell(s_i) = (\mathscr{M}, w_i)$, $1 \leq i \leq 4$, and $s_1 \prec s_2$, $s_2 \prec s_3$, and $s_3 \prec s_4$. It can be checked that $\mathscr{P}$ satisfies $\mathcal{K}^{|\sim}$. However, $K_B p \wedge K_B \neg c \wedge \neg K_A h \not\mid\sim_{\mathscr{P}} h$.

It can be checked that $K_B p \wedge K_B \neg c \wedge \neg K_A h \mid\sim h$ *is* in the modal rational closure of the defeasible knowledge base $\mathcal{K}^{|\sim}$.

7 Defeasible Accessibility Relations

Theorems 4.4 and 4.6 characterize the semantic notions of preferential and rational consequence by their respective properties. Adding more properties to a preferential consequence relation necessitates placing further restrictions on the corresponding semantic structures. For example, adding the property (RM) to a preferential consequence relation necessitates a ranked order on states in modal preferential models.

But the addition of further properties to a preferential consequence relation can also influence its corresponding preferential semantic structures in another way. Consider the additional property $\Box_i \alpha \mid\sim \alpha$, the semantics of which is that every minimal state satisfying $\Box_i \alpha$ also satisfies α. Now replace $\mid\sim$ with *local* classical modal entailment, and we get $\models \Box_i \alpha \rightarrow \alpha$, i.e., every world in a Kripke model satisfying $\Box_i \alpha$ also satisfies α. This is of course the axiom schema characterizing the class of modal frames with reflexive accessibility relations. Now, $\Box_i \alpha \mid\sim \alpha$ is quite different from the Reflexivity property (Ref) for preferential consequence,

i.e., $\alpha \mathrel{|\!\sim} \alpha$, which states that $\mathrel{|\!\sim}$ is reflexive. Instead, $\Box_i\alpha \mathrel{|\!\sim} \alpha$ states that any normal world in a Kripke model that satisfies $\Box_i\alpha$, also satisfies α. That is, R_i is *normally* reflexive, and it is only for some world-model pairs that are exceptional with respect to $\Box_i\alpha$, that we may have $\mathscr{M}, w \Vdash \Box_i\alpha$ but $\mathscr{M}, w \Vdash \neg\alpha$. Thus the property $\Box_i\alpha \mathrel{|\!\sim} \alpha$ captures the notion of a *locally* defeasible reflexive accessibility relation R_i. Similarly, we can capture the notion of a *globally* defeasible accessibility relation R_i by the property $\top \mathrel{|\!\sim} \Box_i\alpha \to \alpha$.

For each such defeasible property of the class of accessibility relations, we still need to characterize the corresponding class of modal preferential models, but for now we leave this for future investigation.

8 Related Work

This paper builds on current work on $\mathcal{ALC}$ [5], in which rational closure for $\mathcal{ALC}$ is studied in more detail.

Also in a DL setting, Britz et al. [4] and Giordano et al. [13] use typicality orderings on *objects* in first-order domains to define defeasible subsumption in $\mathcal{ALC}$. Both approaches define rational consequence relations, but without representation results. In contrast, we provide a general semantic framework which, unlike these proposals, is relevant to all logics with a possible worlds semantics. This is because our semantics for defeasible consequence yields a single order on relational structures at the meta level, rather than ad hoc relativized orders at the object level.

The notion of defeasibility in action theories has been dealt with in non-modal frameworks for reasoning about actions [1,28]. Contrary to ours, their work is not concerned with extending preferential reasoning to more expressive logics. However, defeasible statements of the kind we studied here can be used in reasoning about the *qualifications* of actions: In situations where α holds, the action a is usually executable; but in the more specific context α', a's execution fails. In that sense, our framework also stands as an approach to the qualification problem [21] and to the more general problem of *revising* action domain descriptions [25,26].

As mentioned in Section 4, Casini and Straccia [6] recently proposed an algorithm for determining rational closure in the context of description logics. Their algorithm can be adjusted to determine the modal rational closure as defined in Definition 4.12. From the description of this algorithm it follows that, for the modal logics under consideration here, determining membership of the modal rational closure is no harder than global entailment checking. Furthermore, the algorithm can readily be extended to implement refinements presented by *presumptive* reasoning [18], and inheritance networks [7]. We are currently investigating further optimizations of the algorithm via suitable definitions of modularity for defeasible knowledge bases [15,16].

9 Concluding Remarks

The contributions of this paper are as follows: (i) we filled an important gap in the non-monotonic reasoning community by providing a natural and intuitive semantics for preferential reasoning in modal logics; (ii) we gave to our semantics

a corresponding syntactic characterization via our representation results; (*iii*) we established the basis with which to 'lift' the propositional notions of defeasible consequence and closure to modal logics in general; (*iv*) we illustrated how our constructions can be applied in two important families of modal logic; and (*v*) we proposed defeasible versions of some classical properties of accessibility relations.

The simple modal logics that we have assessed here are the backbone of well established formalisms for reasoning about actions, knowledge, obligations, and combinations thereof. A proper understanding of preferential reasoning for basic modal logic is therefore important for finding specific definitions of defeasible reasoning in these logics. In that sense, we believe that the results we develop here pave the way for extending preferential and rational consequence, and hence also rational closure, to a whole class of logics for knowledge representation and reasoning.

References

[1] Baral, C. and J. Lobo, *Defeasible specifications in action theories*, in: *Proc. IJCAI*, 1997, pp. 1441–1446.

[2] Blackburn, P., J. van Benthem and F. Wolter, "Handbook of Modal Logic," Elsevier North-Holland, 2006.

[3] Britz, K., J. Heidema and W. Labuschagne, *Semantics for dual preferential entailment*, Journal of Philosophical Logic **38** (2009), pp. 433–446.

[4] Britz, K., J. Heidema and T. Meyer, *Semantic preferential subsumption*, in: *Proc. KR*, 2008, pp. 476–484.

[5] Britz, K., T. Meyer and I. Varzinczak, *Semantic foundation for preferential description logics*, in: *Proc. 24th Australasian Conference on AI*, 2011.

[6] Casini, G. and U. Straccia, *Rational closure for defeasible description logics*, in: *Proc. JELIA*, 2010, pp. 77–90.

[7] Casini, G. and U. Straccia, *Defeasible inheritance-based description logics*, in: *Proc. IJCAI*, 2011, pp. 813–818.

[8] Castilho, M., A. Herzig and I. Varzinczak, *It depends on the context! A decidable logic of actions and plans based on a ternary dependence relation*, in: *9th International Workshop on Nonmonotonic Reasoning (NMR)*, 2002.

[9] De Giacomo, G. and M. Lenzerini, *PDL-based framework for reasoning about actions*, in: *IA*AI*, 1995, pp. 103–114.

[10] Demolombe, R., A. Herzig and I. Varzinczak, *Regression in modal logic*, Journal of Applied Non-Classical Logics **13** (2003), pp. 165–185.

[11] Fagin, R., J. Halpern, Y. Moses and M. Vardi, "Reasoning about Knowledge," MIT Press, 1995.

[12] Freund, M., *Preferential reasoning in the perspective of Poole default logic*, Artificial Intelligence **98** (1998), pp. 209–235.

[13] Giordano, L., N. Olivetti, V. Gliozzi and G. Pozzato, $\mathcal{ALC} + T$: *a preferential extension of description logics*, Fundamenta Informatica **96** (2009), pp. 341–372.

[14] Harel, D., J. Tiuryn and D. Kozen, "Dynamic Logic," MIT Press, 2000.

[15] Herzig, A. and I. Varzinczak, *On the modularity of theories*, in: *Advances in Modal Logic* (2005), pp. 93–109.

[16] Herzig, A. and I. Varzinczak, *A modularity approach for a fragment of $\mathcal{ALC}$*, in: M. Fisher, W. van der Hoek, B. Konev and A. Lisitsa, editors, *Proceedings of the 10th European Conference on Logics in Artificial Intelligence (JELIA)*, number 4160 in LNAI (2006), pp. 216–228.

[17] Kraus, S., D. Lehmann and M. Magidor, *Nonmonotonic reasoning, preferential models and cumulative logics*, Artificial Intelligence **44** (1990), pp. 167–207.

[18] Lehmann, D., *Another perspective on default reasoning*, Annals of Mathematics and Artificial Intelligence **15** (1995), pp. 61–82.

[19] Lehmann, D. and M. Magidor, *Preferential logics: the predicate calculus case*, in: *Proc. TARK*, 1990, pp. 57–72.

[20] Lehmann, D. and M. Magidor, *What does a conditional knowledge base entail?*, Artificial Intelligence **55** (1992), pp. 1–60.

[21] McCarthy, J., *Epistemological problems of artificial intelligence*, in: *Proc. IJCAI*, 1977, pp. 1038–1044.

[22] McNamara, P., *Deontic logic*, in: *Handbook of the History of Logic, Vol. 7*, Elsevier N-Holland, 2006 pp. 197–288.

[23] Nute, D., editor, "Defeasible Deontic Logic," Synthese Library **263**, Kluwer Academic Publishers, 1997.

[24] Poole, D., *A logical framework for default reasoning*, Artificial Intelligence **36** (1988), pp. 27–47.

[25] Varzinczak, I., *Action theory contraction and minimal change*, in: *Proc. KR*, 2008, pp. 651–661.

[26] Varzinczak, I., *On action theory change*, Journal of Artificial Intelligence Research **37** (2010), pp. 189–246.

[27] Zhang, D. and N. Foo, *EPDL: A logic for causal reasoning*, in: *Proc. IJCAI*, 2001, pp. 131–138.

[28] Zhang, Y., *Handling defeasibilities in action domains*, Theory and Practice of Logic Programming **3** (2003), pp. 329–376.

A   Proof Sketch for Theorem 4.4

Lemma A.1 (Soundness) *For any preferential model $\mathscr{P}$, the consequence relation it defines is a preferential relation.*

Let $\mid\!\sim$ denote a preferential relation, i.e., $\mid\!\sim$ satisfies the properties for preferential reasoning from Section 2.

Definition A.2 A pair $(\mathscr{M}, w) \in \mathscr{U}$ is *normal* for $\alpha \in \mathcal{L}$ iff for every $\beta \in \mathcal{L}$ s.t. $\alpha \mid\!\sim \beta$, $\mathscr{M}, w \Vdash \beta$.

Lemma A.3 *Let $\mid\!\sim$ satisfy (Ref), (RW), and (And), and let $\alpha, \beta \in \mathcal{L}$. All normal $(\mathscr{M}, w)$ for α satisfy β iff $\alpha \mid\!\sim \beta$.*

Definition A.4 Let $\alpha, \beta \in \mathcal{L}$. $\alpha \leqslant \beta$ iff $\alpha \vee \beta \mid\!\sim \alpha$.

We construct a preferential model as in Definition 4.2. Let $\mathscr{P} = \langle S, \ell, \prec \rangle$, where

- $S = \{\langle \mathscr{M}, w, \alpha \rangle \mid (\mathscr{M}, w) \in \mathscr{U} \text{ is normal for } \alpha \in \mathcal{L}\}$;
- $\ell(\langle \mathscr{M}, w, \alpha \rangle) = (\mathscr{M}, w)$;
- $\langle \mathscr{M}, w, \alpha \rangle \prec \langle \mathscr{M}', w', \beta \rangle$ iff $\alpha \leqslant \beta$ and $\mathscr{M}, w \not\Vdash \beta$.

Lemma A.5 *$\prec$ is a strict partial order on S.*

Lemma A.6 *For any $\alpha \in \mathcal{L}$, $\widehat{\alpha} \subseteq S$ is smooth.*

Lemma A.7 *$\alpha \mid\!\sim \beta$ iff $\alpha \mid\!\sim_{\mathscr{P}} \beta$.*

Proof Sketch for Theorem 4.4. Soundness is given by Lemma A.1. For the only if part, let $\mid\!\sim$ be a preferential consequence relation, and let $\mathscr{P}$ be defined as above. We showed in Lemmas A.5 and A.6 that $\mathscr{P}$ is indeed a preferential model. Lemma A.7 shows that $\mathscr{P}$ defines a consequence relation that is exactly $\mid\!\sim$.   □

B   Proof Sketch for Theorem 4.6

Lemma B.1 (Soundness) *If $\mathscr{P}_r$ is a ranked model, then the consequence relation $\mathrel{\vdash\!\!\!\sim}_{\mathscr{P}_r}$ it defines is rational.*

Let $\mathrel{\vdash\!\!\!\sim}$ denote a rational relation, i.e., $\mathrel{\vdash\!\!\!\sim}$ satisfies the properties for preferential reasoning from Section 2 plus (RM).

Definition B.2 $\alpha \in \mathcal{L}$ is *consistent* w.r.t. $\mathrel{\vdash\!\!\!\sim}$ iff $\alpha \mathrel{\not\vdash\!\!\!\sim} \bot$. Given $\mathscr{P}_r = \langle S, \ell, \prec \rangle$, α is *consistent* w.r.t. $\mathrel{\vdash\!\!\!\sim}_{\mathscr{P}_r}$ iff $\alpha \mathrel{\not\vdash\!\!\!\sim}_{\mathscr{P}_r} \bot$, i.e., iff there is $s \in S$ s.t. $s \in \widehat{\alpha}$.

Let $\mathscr{C} = \{\alpha \in \mathcal{L} \mid \alpha \text{ is consistent w.r.t. } \mathrel{\vdash\!\!\!\sim}\}$.

Definition B.3 Given $\alpha, \beta \in \mathscr{C}$, α is *not more exceptional* than β, written $\alpha \mathcal{R} \beta$, iff $\alpha \vee \beta \mathrel{\not\vdash\!\!\!\sim} \neg\alpha$. α is *as exceptional as* β, written $\alpha \sim \beta$, iff $\alpha\mathcal{R}\beta$ and $\beta\mathcal{R}\alpha$.

That $\sim$ is an equivalence relation follows from the fact that $\mathcal{R}$ is reflexive and transitive. With $\overline{\alpha}$ we denote the equivalence class of α. The set of equivalence classes of elements of $\mathscr{C}$ under $\sim$ is denoted by $\overline{\mathscr{C}}$. We write $\overline{\alpha} \leq \overline{\beta}$ iff $\alpha\mathcal{R}\beta$, and $\overline{\alpha} < \overline{\beta}$ iff $\overline{\alpha} \leq \overline{\beta}$ and $\alpha \not\sim \beta$.

We construct a ranked model as in Definition 4.5. Let $\mathscr{P}_r = \langle S, \ell, \prec \rangle$, where:

- $S = \{\langle \mathscr{M}, w, \alpha \rangle \mid (\mathscr{M}, w) \text{ is normal for } \alpha \in \mathcal{L}\}$;
- $\ell(\langle \mathscr{M}, w, \alpha \rangle) = (\mathscr{M}, w)$;
- $\langle \mathscr{M}, w, \alpha \rangle \prec \langle \mathscr{M}', w', \beta \rangle$ iff $\overline{\alpha} < \overline{\beta}$.

Lemma B.4 $\prec$ *is a modular partial order.*

Lemma B.5 *Let $\alpha \in \mathcal{L}$ be consistent. $\widehat{\alpha} \subseteq S$ is smooth.*

Lemma B.6 $\alpha \mathrel{\vdash\!\!\!\sim} \beta$ *iff* $\alpha \mathrel{\vdash\!\!\!\sim}_{\mathscr{P}_r} \beta$.

Proof Sketch for Theorem 4.6. Soundness is given by Lemma B.1. For the only if part, let $\mathrel{\vdash\!\!\!\sim}$ be a rational consequence relation, and $\mathscr{P}_r$ be as defined above. By Lemmas B.4 and B.5, $\mathscr{P}_r$ is a ranked model. Lemma B.6 shows that $\mathscr{P}_r$ defines a consequence relation that is exactly $\mathrel{\vdash\!\!\!\sim}$. $\square$

C   Proof Sketch for Theorem 4.14

We first show that $\mathcal{R}^{\mathrel{\vdash\!\!\!\sim}}$ contains $\mathcal{K}^{\mathrel{\vdash\!\!\!\sim}}$ and then that $\mathcal{R}^{\mathrel{\vdash\!\!\!\sim}}$ is a rational consequence relation. The remainder of the proof is concerned with showing that if $\mathcal{R}_1^{\mathrel{\vdash\!\!\!\sim}}$ is a different rational consequence relation containing $\mathcal{K}^{\mathrel{\vdash\!\!\!\sim}}$, then $\mathcal{R}^{\mathrel{\vdash\!\!\!\sim}} \ll \mathcal{R}_1^{\mathrel{\vdash\!\!\!\sim}}$, which proves that $\mathcal{R}^{\mathrel{\vdash\!\!\!\sim}}$ is the modal rational closure of $\mathcal{K}^{\mathrel{\vdash\!\!\!\sim}}$. This part of the proof makes use of the ranks of formulas, but the crucial aspect is the use of three ranked models $\mathscr{P}_r$, $\mathscr{P}_r'$, and $\mathscr{P}_r''$. $\mathscr{P}_r$ is a ranked model characterizing $\mathcal{R}_1^{\mathrel{\vdash\!\!\!\sim}}$, while $\mathscr{P}_r''$ is constructed from $\mathscr{P}_r$, and $\mathscr{P}_r'$ is constructed from $\mathscr{P}_r''$. The models $\mathscr{P}_r''$ and $\mathscr{P}_r'$ are used to show that $\mathcal{R}^{\mathrel{\vdash\!\!\!\sim}} \ll \mathcal{R}_1^{\mathrel{\vdash\!\!\!\sim}}$. $\square$