# Normal Modal Preferential Consequence

Katarina Britz, Thomas Meyer, and Ivan Varzinczak

Centre for Artificial Intelligence Research
CSIR Meraka Institute and UKZN, South Africa
{arina.britz,tommie.meyer,ivan.varzinczak}@meraka.org.za

**Abstract.** One of the most successful approaches to the formalization of commonsense reasoning is the work by Lehmann and colleagues, known as the KLM approach, in which defeasible consequence relations with a preferential semantics are studied. In spite of its success, KLM is limited to propositional logic. In recent work we provided the semantic foundation for extending defeasible consequence relations to modal logics and description logics. In this paper we continue that line of investigation by going beyond the basic (propositional) KLM postulates, thereby making use of the additional expressivity provided by modal logic. In particular, we show that the additional constraints we impose on the preferential semantics ensure that the rule of necessitation holds for the corresponding consequence relations, as one would expect it to. We present a representation result for this tightened framework, and investigate appropriate notions of entailment in this context — normal entailment, and a rational version thereof.

**Keywords:** Non-monotonic reasoning; preferential consequence; modal logic

## 1   Introduction and Motivation

The formalization of commonsense reasoning, as usually studied in the AI tradition, depends crucially on the eschewal of the monotonicity property of classical logic, or, at the very least, on a careful neutralization thereof. This issue has been dealt with in a variety of ways in the non-monotonic literature. One particular approach that has been quite successful is the one by Lehmann and colleagues. In their seminal papers [12, 14], the authors consolidated what became known as the *KLM approach*, in which (propositional) defeasible consequence relations $\mid\!\sim$ with a preferential semantics are studied. In this setting, $\alpha \mid\!\sim \beta$ is given the meaning that "all normal (i.e., most preferred) $\alpha$-worlds are $\beta$-worlds", leaving it open for $\alpha$-worlds that are exceptional (or less preferred) not to satisfy $\beta$. The theory that has been developed around this notion allows us to cope with exceptionality when performing reasoning. Besides its simplicity and elegance, the type of consequence relations studied by Lehmann and colleagues has also played an important role in the formalization of commonsense reasoning in providing the foundation for the important notion of rational closure [14].

Notwithstanding its fruitfulness, the KLM approach is limited to *propositional* logic and so it remained until recently despite some attempts to recast it

in more expressive formalisms [5, 9, 11, 13, 15, 16]. Indeed, many scenarios that are interesting from the standpoint of modern AI cannot be satisfactorily formalized in a propositional language. Extensions of the KLM approach have therefore been driven by either extending the syntax [9, 11, 16] or the underlying preferential semantics [5, 13] to logics with more expressivity. A unifying semantics, with a corresponding representation result, was nevertheless still missing until a recent work by the present authors [6, 7] provided the semantic foundation for extending defeasible consequence relations to modal logics [2] and description logics (DLs) [1]. In the referred papers we lifted the notion of rational closure as defined by Lehmann and Magidor in the propositional case [14] to modal and description logics, thereby providing a preliminary account of this construction in logics with more structure than the propositional one.

It turns out that the aforementioned approach, although counting as a true extension of the KLM framework to non-propositional languages, is still limited in the sense that it does not make use of the additional expressivity of e.g. modal languages. To make this more precise, one can state defeasible statements of the form $\alpha \mathrel{|\sim} \beta$, where $\alpha$ and $\beta$ now can be *any* modal sentence; however the syntactic characterization of defeasible consequence (i.e., the set of Gentzen-style properties specifying the expected behavior of $\mathrel{|\sim}$) is confined to the original Boolean postulates proposed by Kraus et al. In other words, the additional expressivity of modal logic is not reflected in terms of new properties, which means that modal sentences are basically opaque to the postulates. Moreover, despite the underlying modal formalism, some inference rules that are seen as important in a modal context such as the necessitation rule behave in an unexpected way.

In this paper we analyze these issues and address them by proposing additional properties that a truly modal-based defeasible consequence relation $\mathrel{|\sim}$ ought to satisfy. In particular, we study what semantic constraints should be added to the original preferential semantics for the new properties to hold.

The remainder of the present paper is organized as follows: After some logical preliminaries (Section 2), we recap our preferential semantics for defeasible modal logic (Section 3). We then motivate the need for KLM-style properties reflecting the additional expressivity of modal languages (Section 4). In particular, we define appropriate semantic constraints warranting the new postulates and establish the corresponding representation result. In Section 5 we define entailment from defeasible knowledge bases and motivate the need to move beyond rational closure. We conclude with a summary of our contributions and directions for future research.

## 2   Modal Logic

We work in a (finite) set of *atomic propositions* $\mathcal{P}$, using the logical connectives $\wedge$ (conjunction), $\neg$ (negation), and a set of modal operators $\Box_i$, $1 \leq i \leq n$. Propositions are denoted by $p, q, \ldots$, and formulas by $\alpha, \beta, \ldots$, built up in the usual way according to the rule: $\alpha \ ::= p \mid \neg\alpha \mid \alpha \wedge \alpha \mid \Box_i\alpha$. All the other truth functional connectives ($\vee, \rightarrow, \leftrightarrow, \ldots$) are defined in terms of $\neg$ and $\wedge$ in

the usual way. Given $\Box_i$, with $\Diamond_i$ we denote its *dual* operator, i.e., for any $\alpha$, $\Diamond_i\alpha \equiv_{\mathrm{def}} \neg\Box_i\neg\alpha$. We use $\top$ as an abbreviation for $p \vee \neg p$ and $\bot$ for $p \wedge \neg p$, for some $p \in \mathcal{P}$. With $\mathcal{L}$ we denote the set of all formulas of the modal language.

As for the semantics, we assume the standard possible-worlds one:

**Definition 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 $W$, $1 \le i \le n$, and $V : W \times \mathcal{P} \longrightarrow \{0, 1\}$ is a valuation function.*

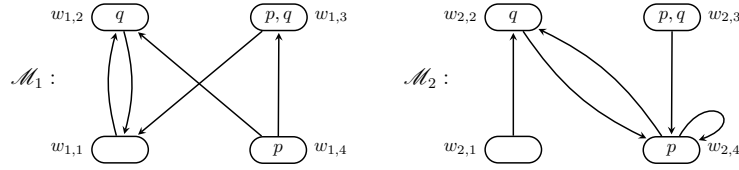Figure 1 depicts two examples of Kripke models for $\mathcal{P} = \{p, q\}$.



**Fig. 1.** Examples of Kripke models.

Sometimes it is convenient to talk about possible worlds in the context of their respective Kripke models. Given $\mathscr{M} = \langle W, R, V \rangle$ and $w \in W$, a pair $(\mathscr{M}, w)$ is a *pointed Kripke model.* Pointed Kripke models are not to be viewed as objects, as variables are commonly regarded in first-order contexts. A set of pointed Kripke models describes the *intention* of a modal statement — cf. Definition 2 below.

Formulas of our modal language are true or false relative to a possible world in a Kripke model. This is formalized by the following truth conditions:

**Definition 2.** *Given $\mathscr{M} = \langle W, R, V \rangle$ and $w \in W$:*

- *$\mathscr{M}, w \Vdash p$ if and only if $V(w, p) = 1$;*
- *$\mathscr{M}, w \Vdash \neg\alpha$ if and only if $\mathscr{M}, w \nVdash \alpha$;*
- *$\mathscr{M}, w \Vdash \alpha \wedge \beta$ if and only if $\mathscr{M}, w \Vdash \alpha$ and $\mathscr{M}, w \Vdash \beta$;*
- *$\mathscr{M}, w \Vdash \Box_i\alpha$ 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 $\mathscr{M} = \langle W, R, V \rangle$, $\mathscr{M}$ *satisfies* $\alpha$ if there is $w \in W$ such that $\mathscr{M}, w \Vdash \alpha$. We say that $\alpha$ is *true* in $\mathscr{M}$ (alias $\mathscr{M}$ is a *model* of $\alpha$) if $\mathscr{M}, w \Vdash \alpha$ for every $w \in W$. For a given system of modal logic, we say that $\alpha$ is *valid* (denoted $\models \alpha$) if $\alpha$ is true in every model of the underlying system. Here we shall assume the system of normal modal logic $\mathsf{K}$, of which all the other normal modal logics are extensions. Semantically, $\mathsf{K}$ is characterized by the class of all Kripke models (Definition 1). We say that $\alpha$ *locally entails* $\beta$ in the system $\mathsf{K}$ (denoted $\alpha \models \beta$) if for every model $\mathscr{M}$ and every $w$ in $\mathscr{M}$, $\mathscr{M}, w \Vdash \alpha$ implies $\mathscr{M}, w, \Vdash \beta$.

Syntactically, $\mathsf{K}$ corresponds to the smallest set of sentences containing all propositional tautologies, all instances of the axiom schema $K : \Box_i(\alpha \to \beta) \to (\Box_i\alpha \to \Box_i\beta)$, $1 \le i \le n$, and closed under the *rule of necessitation* below:

$$(\mathrm{RN}) \quad \frac{\alpha}{\Box_i\alpha} \tag{1}$$

The following are derived rules in the system $\mathsf{K}$:

$$(\text{RK}) \ \frac{\alpha_1 \wedge \ldots \wedge \alpha_k \rightarrow \beta}{\Box_i \alpha_1 \wedge \ldots \wedge \Box_i \alpha_k \rightarrow \Box_i \beta} \qquad (\text{Mon}) \ \frac{\alpha \rightarrow \beta}{\Box_i \alpha \rightarrow \Box_i \beta} \qquad (\text{Cgr}) \ \frac{\alpha \leftrightarrow \beta}{\Box_i \alpha \leftrightarrow \Box_i \beta}$$

Given a system of modal logic, from a knowledge representation perspective it is convenient to be able to work within a class of models $\mathcal{M}$ of the corresponding system, representing e.g. some background knowledge of relevance for a given application domain.

## 3    Modal Defeasible Consequence

A modal defeasible consequence relation $\mathrel{|\!\sim}$ is defined as a binary relation on formulas of our underlying modal logic, i.e., $\mathrel{|\!\sim} \subseteq \mathcal{L} \times \mathcal{L}$. We say that $\mathrel{|\!\sim}$ is a *preferential* consequence relation [6] if it satisfies the following set of properties (alias postulates or Gentzen-style rules, as they are sometimes referred to):

$$(\text{Ref}) \ \alpha \mathrel{|\!\sim} \alpha \qquad\qquad (\text{LLE}) \ \frac{\models \alpha \leftrightarrow \beta, \ \ \alpha \mathrel{|\!\sim} \gamma}{\beta \mathrel{|\!\sim} \gamma} \qquad (\text{And}) \ \frac{\alpha \mathrel{|\!\sim} \beta, \ \ \alpha \mathrel{|\!\sim} \gamma}{\alpha \mathrel{|\!\sim} \beta \wedge \gamma}$$

$$(\text{Or}) \ \frac{\alpha \mathrel{|\!\sim} \gamma, \ \ \beta \mathrel{|\!\sim} \gamma}{\alpha \vee \beta \mathrel{|\!\sim} \gamma} \qquad (\text{RW}) \ \frac{\alpha \mathrel{|\!\sim} \beta, \ \ \models \beta \rightarrow \gamma}{\alpha \mathrel{|\!\sim} \gamma} \qquad (\text{CM}) \ \frac{\alpha \mathrel{|\!\sim} \beta, \ \ \alpha \mathrel{|\!\sim} \gamma}{\alpha \wedge \beta \mathrel{|\!\sim} \gamma}$$

The semantics of preferential consequence relations is in terms of *modal preferential models*; these are partially ordered structures with states labeled by pointed Kripke models (cf. Section 2):

**Definition 3.** *Let $\mathcal{M}$ be a class of Kripke models. $\mathscr{U}_{\mathcal{M}} := \{(\mathcal{M}, w) \mid \mathcal{M} = \langle W, R, V \rangle \in \mathcal{M} \text{ and } w \in W\}$.*

Let $S$ be a set and $\prec \subseteq S \times S$ be a strict partial order on $S$, i.e., $\prec$ is *irreflexive* and *transitive*. Given $S' \subseteq S$, we say that $s \in S'$ is *minimal* in $S'$ if there is no $s' \in S'$ such that $s' \prec s$. With $\min_{\prec} S'$ we denote the minimal elements of $S' \subseteq S$. We say that $S' \subseteq S$ is *smooth* [12] if for every $s \in S'$ either $s$ is minimal in $S'$ or there is $s' \in S'$ such that $s'$ is minimal in $S'$ and $s' \prec s$.

**Definition 4 (Preferential Model).** *A preferential model is a tuple $\mathscr{P} = \langle S, \ell, \prec \rangle$ where $S$ is a set of states; $\ell : S \longrightarrow \mathscr{U}_{\mathcal{M}}$ is a labeling function; $\prec \subseteq S \times S$ is a strict partial order on $S$ satisfying the smoothness condition.*[1]

Given a preferential model $\mathscr{P} = \langle S, \ell, \prec \rangle$ and $\alpha \in \mathcal{L}$, with $\llbracket \alpha \rrbracket$ we denote the set of states satisfying $\alpha$ ($\alpha$-*states* for short) according to the following definition:

**Definition 5.** *Let $\mathscr{P} = \langle S, \ell, \prec \rangle$ and let $\alpha \in \mathcal{L}$. Then $\llbracket \alpha \rrbracket := \{s \in S \mid \ell(s) \Vdash \alpha\}$.*

---

[1] That is, for every $\alpha \in \mathcal{L}$, the set $\llbracket \alpha \rrbracket$ (cf. Definition 5) is smooth.

States lower down in the order are *more preferred* (or *more normal*) than those higher up. As an example, let $\mathcal{M}$ be the class of K-models depicted in Figure 1. Then $\mathscr{U}_{\mathcal{M}} = \{(\mathscr{M}_i, w_{i,j}) \mid i = 1, 2 \text{ and } 1 \leq j \leq 4\}$. Figure 2 below depicts the preferential model $\mathscr{P} = \langle S, \ell, \prec \rangle$ where $S = \{s_i \mid 1 \leq i \leq 8\}$, $\ell$ is such that $\ell(s_1) = (\mathscr{M}_1, w_{1,1})$, $\ell(s_2) = (\mathscr{M}_2, w_{2,1})$, $\ell(s_3) = (\mathscr{M}_1, w_{1,2})$, $\ell(s_4) = (\mathscr{M}_1, w_{1,3})$, $\ell(s_5) = (\mathscr{M}_2, w_{2,3})$, $\ell(s_6) = (\mathscr{M}_2, w_{2,2})$, $\ell(s_7) = (\mathscr{M}_1, w_{1,4})$, and $\ell(s_8) = (\mathscr{M}_2, w_{2,4})$, and $\prec$ is the transitive closure of $\{(s_1, s_3), (s_2, s_3), (s_3, s_4),$ $(s_3, s_5), (s_4, s_6), (s_5, s_6), (s_6, s_7), (s_6, s_8)\}$.
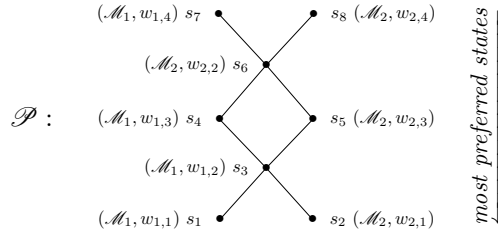


**Fig. 2.** A preferential model for $\mathcal{M} = \{\mathscr{M}_1, \mathscr{M}_2\}$, with $\mathscr{M}_1$ and $\mathscr{M}_2$ as in Figure 1.

Given $\mathscr{P} = \langle S, \ell, \prec \rangle$ and $\alpha \in \mathcal{L}$, $\alpha$ is *satisfiable* in $\mathscr{P}$ if $[\![\alpha]\!] \neq \emptyset$, otherwise $\alpha$ is *unsatisfiable* in $\mathscr{P}$. We say that $\alpha$ is *true* in $\mathscr{P}$ (denoted $\mathscr{P} \Vdash \alpha$) if $[\![\alpha]\!] = S$.

From the definition of a preferential model one can see that a class $\mathcal{M}$ of Kripke models determines a class of preferential models. We denote the class of preferential models based on $\mathcal{M}$ with $\mathcal{M}^{\mathscr{P}}$. We say that $\alpha$ is *valid* in $\mathcal{M}^{\mathscr{P}}$ if $\alpha$ is true in every preferential model $\mathscr{P}$ of $\mathcal{M}^{\mathscr{P}}$, i.e., $\mathscr{P} \Vdash \alpha$ for every $\mathscr{P} \in \mathcal{M}^{\mathscr{P}}$.

Given $\mathscr{P} = \langle S, \ell, \prec \rangle$, the defeasible statement $\alpha \mathrel{|\!\sim}_{\mathscr{P}} \beta$ holds in $\mathscr{P}$ if and only if $\min_{\prec} [\![\alpha]\!] \subseteq [\![\beta]\!]$, i.e., every $\prec$-minimal $\alpha$-state is a $\beta$-state. As an example, in the model $\mathscr{P}$ of Figure 2, we have $\neg q \mathrel{|\!\sim}_{\mathscr{P}} \Box \neg p$ and also $q \mathrel{|\!\sim}_{\mathscr{P}} \Diamond(\neg p \wedge \neg q)$.

The representation theorem for preferential consequence relations then states:

**Theorem 1 (Britz et al. [6]).** *A modal defeasible consequence relation is a preferential consequence relation if and only if it is defined by some preferential model, i.e., $\mathrel{|\!\sim}$ is preferential if and only if there exists $\mathscr{P}$ such that $\mathrel{|\!\sim} = \mathrel{|\!\sim}_{\mathscr{P}}$.*

If, in addition to the preferential properties, the defeasible consequence relation $\mathrel{|\!\sim}$ also satisfies the following Rational Monotonicity property [14], it is said to be a *rational* consequence relation:

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

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

**Definition 6 (Modular Order).** *Given a set $S$, $\prec \subseteq S \times S$ is modular if and only if 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 $rk(s) < rk(s')$.*

**Definition 7 (Ranked Model).** *A ranked model $\mathscr{R} = \langle S, \ell, \prec \rangle$ is a preferential model such that $\prec$ is modular.*

The preferential model in Figure 2 is also an example of a ranked model.

**Theorem 2 (Britz et al. [6]).** *A modal defeasible consequence relation is a rational consequence relation if and only if it is defined by some ranked model, i.e., $\mid\!\sim$ is rational if and only if there exists $\mathscr{R}$ such that $\mid\!\sim \;=\; \mid\!\sim_{\mathscr{R}}$.*

## 4    Beyond the KLM Postulates

Britz et al.'s constructions and representation results are with respect to the same set of properties used to characterize propositional rational consequence. This has the advantage that methods employed in a propositional non-monotonic setting translate seamlessly to a modal context. This includes reasoning tasks such as computing the rational closure of defeasible knowledge bases [6, 7]. In that respect, the definitions in Section 3 provide a good starting point for investigating more elaborated versions of modal rational consequence. Here we are interested in doing precisely this and we start by making an important observation:

**Proposition 1.** *Let $\alpha \in \mathcal{L}$ and let $\mathcal{M}$ be a class of Kripke models. Then $\alpha$ is valid in $\mathcal{M}$ if and only if $\alpha$ is valid in $\mathcal{M}^{\mathscr{P}}$.*

That is, all the validities of the underlying system of modal logic (or of the specific class of models we are working with) remain valid with respect to our preferential semantics. An immediate consequence of this is the following result:

**Corollary 1.** *All inference rules of the underlying modal logic are sound with respect to the preferential semantics.*

It is easy to see why: given a rule $\rho$ which is sound in the respective system of modal logic and of which the premise $\alpha$ is preferentially valid, from Proposition 1 follows that $\alpha$ is (modally) valid, from which follows the validity of $\rho$'s consequent $\beta$, which, by Proposition 1 again, must be preferentially valid.

In spite of preserving all modal validities and rules of inference, the current preferential semantics gives rise to a rather odd phenomenon: Contrary to the classical possible worlds semantics, in the preferential semantics some inference rules need not be satisfied by *individual* (preferential or ranked) models. To make this more precise, one can devise models in which the rule of necessitation (RN) does *not* hold. Indeed, the current preferential semantics is too liberal in the sense that it allows for legitimate models in which $\alpha$ is true (i.e., $[\![\alpha]\!] = S$) but in which $\Box_i\alpha$ fails to hold (i.e., $[\![\Box_i\alpha]\!] \neq S$), for some $1 \leq i \leq n$. To witness, let $\mathscr{M}_1$ be as in Figure 1, and let $S = \{s_1\}$, $\ell(s_1) = (\mathscr{M}_1, w_{1,4})$, and $\prec = \emptyset$. It is easy to check that $\mathscr{P} = \langle S, \ell, \prec \rangle$ is a preferential model and that $\mathscr{P} \Vdash p$ but $\mathscr{P} \not\Vdash \Box p$. Hence the preferential semantics for modal consequence is not, strictly speaking, truly normal in the modal sense.

It should not be that hard to see that this is not an inoffensive feature: In a given $\mathscr{P}$ one is told that every state is an $\alpha$-state, but at the same time it is

possible to refer to a $\neg\alpha$-world that is part of the structure of $\mathscr{P}$. It sounds as though $\mathscr{P}$ is not particularly accurate in the knowledge it conveys. The consequences of this become apparent when considering specific application domains. For instance, in an action context, this could mean that there is a possible execution of tossing a coin having as outcome "not-heads and not-tails", even though there is no configuration of states other than "heads or tails".

With a similar argument we can also show that the rule of congruence (Cgr) in general is not satisfied 'locally', i.e., in an arbitrary model: Let $\mathscr{M}_1$ be as in Figure 1, and let $S = \{s_1\}$, $\ell(s_1) = (\mathscr{M}_1, w_{1,1})$, and $\prec = \emptyset$. Then $\mathscr{P} = \langle S, \ell, \prec \rangle$ is a preferential model and $\mathscr{P} \Vdash \neg p \leftrightarrow \neg q$ but $\mathscr{P} \nVdash \Box\neg p \leftrightarrow \Box\neg q$. (That this is not particularly desirable should not be that hard to see.) The same counter-example applies to Rules RK and Mon.

Surely all of this has to do with the extra richness of modal structures when compared to the propositional ones, but also, as alluded to above, to the rather liberal character of our original preferential semantics and its interplay with the corresponding syntactic characterization.

The obvious direction to follow in tackling the above issues is through the requirement of additional restrictions in the semantics with appropriate postulates characterizing them. Indeed, the rationality properties from Section 3 seem too weak in a modal context, as they do not really make use of the full expressiveness of modal logic. Without properties referring directly to the extra operator $\Box$, modal formulas are treated in a completely opaque way by the remaining Boolean postulates. Hence we shall investigate extra KLM-style properties that do make use of the non-Boolean connectives of the underlying language.

The requirement that every preferential model also satisfy the rule of necessitation provides us with insights towards our stated aim. In what follows we shall have a closer look at it.

**Proposition 2.** *Let $\alpha \in \mathcal{L}$ and $\mathscr{P}$ be a preferential model. Then $\mathscr{P} \Vdash \alpha$ if and only if $\neg\alpha \hspace{1pt}\mid\hspace{-5pt}\sim_{\mathscr{P}} \bot$ holds in $\mathscr{P}$.*

'Local' (i.e., model-wise) satisfaction of the rule of necessitation amounts to having $\Box\alpha$ true in a model whenever $\alpha$ is true in the same model. That is to say, RN holds in $\mathscr{P}$ if $\mathscr{P} \Vdash \alpha$ implies $\mathscr{P} \Vdash \Box_i\alpha$, for every $1 \leq i \leq n$. Given this and Proposition 2, we obtain the following KLM-style version of RN:

$$\text{(Norm)} \quad \frac{\alpha \hspace{1pt}\mid\hspace{-5pt}\sim \bot}{\Diamond_i\alpha \hspace{1pt}\mid\hspace{-5pt}\sim \bot}, \text{ for } 1 \leq i \leq n \tag{2}$$

We call Property (2) *Normality*, as it is the KLM version of a fundamental property of normal modal logics, namely the rule of necessitation. Intuitively it says that what is inconsistent should not be possible.

The result in Proposition 2 also allows us to derive a KLM-style version of the rule of congruence. Cgr holds in $\mathscr{P}$ if $\mathscr{P} \Vdash \alpha \leftrightarrow \beta$ implies $\mathscr{P} \Vdash \Box_i\alpha \leftrightarrow \Box_i\beta$, for every $1 \leq i \leq n$, that is, if the following rule holds in $\mathscr{P}$:

$$\text{(Equiv)} \quad \frac{\neg(\alpha \leftrightarrow \beta) \hspace{1pt}\mid\hspace{-5pt}\sim \bot}{\neg(\Box_i\alpha \leftrightarrow \Box_i\beta) \hspace{1pt}\mid\hspace{-5pt}\sim \bot}, \text{ for } 1 \leq i \leq n \tag{3}$$

With an analogous argument one can derive the following properties ensuring, respectively, Mon and RK:

$$\text{(Imp)} \quad \frac{\neg(\alpha \to \beta) \mathrel{\mid\!\sim} \bot}{\neg(\Box_i\alpha \to \Box_i\beta) \mathrel{\mid\!\sim} \bot}, \text{ for } 1 \le i \le n \tag{4}$$

$$(\text{RK}^{\mid\!\sim}) \quad \frac{\neg(\alpha_1 \wedge \ldots \wedge \alpha_k \to \beta) \mathrel{\mid\!\sim} \bot}{\neg(\Box_i\alpha_1 \wedge \ldots \wedge \Box_i\alpha_k \to \Box_i\beta) \mathrel{\mid\!\sim} \bot}, \text{ for } 1 \le i \le n \tag{5}$$

Being derived from inference rules that hold in every system of classical normal modal logics, the postulates in (2)–(5) stand as reasonable properties to have in modal preferential reasoning. As we have seen, *none* of them hold in the standard preferential semantics.

New postulates are usually captured in the semantics by means of additional restrictions on the preferential models. For instance, in the evolution from preferential consequence relations to rational ones, the rational monotonicity property became warranted by requiring the partial order $\prec$ to be a modular ordering [14]. As we shall see, it turns out that one can also force extra properties (as the ones we stated above) by imposing additional restrictions on the set of *states* and on the *labeling* function, a route that seems not to have been explored so far.

Looking back at the counter-examples to local satisfaction of the inference rules, a common pattern emerges: a given rule is violated because of a pointed Kripke model acting "behind the curtain", i.e., a pointed model labeling no state whatsoever but which (implicitly) still interferes with states that are labeled with pointed models it relates to. We call these pointed models occurring implicitly in a preferential model *spurious models*.[2]

We claim that a modal preferential semantics should not allow for spurious models. The set of states and the labeling function must be disciplined in such a way as to prevent a pointed model from determining the truth of formulas without being itself associated with any state. We make this more precise now.

**Definition 8 (Non-Spuriousness).** *A pair $(S, \ell)$ is non-spurious if and only if, for all $s \in S$ with $\ell(s) = (\mathscr{M}, w)$ for some $\mathscr{M} = \langle W, R, V \rangle$, and for all $w' \in W$ such that $(w, w') \in R_i$ for some $i$, there exists $s' \in S$ with $\ell(s') = (\mathscr{M}, w')$.*

The non-spuriousness condition requires that whenever a possible world can be referred to indirectly, then it is indeed a world that can be accessed directly.

**Definition 9 (Non-Spurious Model).** *A non-spurious model $\mathscr{N} = \langle S, \ell, \prec \rangle$ is a preferential model such that $(S, \ell)$ is non-spurious.*

Given a non-spurious model $\mathscr{N} = \langle S, \ell, \prec \rangle$ and a formula $\alpha \in \mathcal{L}$, as before with $[\![\alpha]\!]$ we denote the set of elements of $S$ satisfying $\alpha$ (cf. Definition 5). The model depicted in Figure 2 is also an example of a non-spurious (ranked) model.

The definition of non-spurious model puts us in a position to state the first of the results leading us to the realization of our stated aims:

---

[2] Note that requiring the labeling function to be surjective would be too strong as it would require the cardinality of $S$ to be at least that of $\mathscr{U}_{\mathcal{M}}$.

**Lemma 1 (Soundness).** *Let $\mathcal{N} = \langle S, \ell, \prec \rangle$ be a non-spurious model and let $\vdash_{\mathcal{N}}$ be the defeasible consequence relation it defines. Then $\vdash_{\mathcal{N}}$ satisfies the preferential properties (Ref), (LLE), (And), (Or), (RW) and (CM), as well as the properties (Norm), (Equiv), (Imp) and ($RK^{\vdash}$). If $\mathcal{N}$ is moreover a ranked model, then $\vdash_{\mathcal{N}}$ also satisfies (RM).*

Before we address completeness, it is worth making an observation:

**Proposition 3.** *Let $\vdash$ be a preferential consequence relation. If $\vdash$ satisfies (Norm), then it satisfies (Equiv), (Imp) and ($RK^{\vdash}$).*

The proof of Proposition 3 relies on Propositions 1 and 2 and on the fact that Rules RK, Mon and Cgr can be derived from RN and the modal validities in the classical case. From the result above we conclude that it is enough to restrict our attention to those consequence relations satisfying (Norm).

**Definition 10 (Normal Consequence).** *A modal consequence relation $\vdash$ is a normal consequence relation if it satisfies all the preferential properties from Section 3 together with (Norm).*

**Lemma 2 (Completeness).** *Let $\vdash \subseteq \mathcal{L} \times \mathcal{L}$ be a normal consequence relation. Then there exists a non-spurious ranked model $\mathcal{N}$ such that $\vdash_{\mathcal{N}} = \vdash$.*

We are now ready to state one of the main results of the present paper:

**Theorem 3.** *A defeasible consequence relation is a normal consequence relation if and only if it is defined by some non-spurious model.*

If $\vdash$ is a normal consequence relation also satisfying RM, we call $\vdash$ a *rational normal* consequence relation. This leads us to our second representation result:

**Theorem 4.** *A defeasible consequence relation is a rational normal consequence relation if and only if it is defined by some non-spurious ranked model.*

## 5   Normal Entailment

So far we have assessed $\vdash$ from the perspective of consequence relations. Following Lehmann and Magidor [14], one can also view $\vdash$ as a connective in an enriched modal language, which allows us to write down defeasible statements (or 'conditionals', as they are also referred to). Given a set of defeasible statements of the form $\alpha \vdash \beta$, from a knowledge representation and reasoning perspective it becomes important to address the question of what it means for a defeasible statement to be *entailed* by others [14].

A *defeasible knowledge base* $\mathcal{K}^{\vdash}$ is a finite set of statements $\alpha \vdash \beta$, where $\alpha, \beta \in \mathcal{L}$ [6]. Given a non-spurious model $\mathcal{N}$, we extend the notion of satisfaction to knowledge bases in the obvious way: $\mathcal{N} \Vdash \mathcal{K}^{\vdash}$ if $\alpha \vdash_{\mathcal{N}} \beta$ for every $\alpha \vdash \beta \in \mathcal{K}^{\vdash}$. This leads us to an obvious definition of entailment:

**Definition 11 (Normal Entailment).** *$\mathcal{K}^{\vdash}$ normally entails $\alpha \vdash \beta$ if and only if for every non-spurious model $\mathcal{N}$, if $\mathcal{N} \Vdash \mathcal{K}^{\vdash}$, then $\alpha \vdash_{\mathcal{N}} \beta$.*

On a related note, the *normal* closure of $\mathcal{K}^{\vdash}$ is defined as the intersection of all the normal consequence relations containing $\mathcal{K}^{\vdash}$.

**Theorem 5.** *Let $\mathcal{K}^{\vdash}$ be a defeasible knowledge base. Then (i) the set of all sentences normally entailed by $\mathcal{K}^{\vdash}$ is a Tarskian consequence relation; (ii) it is a normal consequence relation; (iii) it coincides with the normal closure of $\mathcal{K}^{\vdash}$.*

In the context of normal consequence, normal entailment is therefore the appropriate notion of entailment for defeasible knowledge bases. However, if we shift our focus to the class of *rational normal* consequence relations, the obvious definition of rational normal entailment does not provide a desirable result.

**Definition 12 (Rational Normal Entailment).** *$\mathcal{K}^{\vdash}$ rationally normally entails $\alpha \mathrel{|\!\sim} \beta$ if and only if for every non-spurious ranked model $\mathcal{N}$, if $\mathcal{N} \Vdash \mathcal{K}^{\vdash}$, then $\alpha \mathrel{|\!\sim}_{\mathcal{N}} \beta$.*

**Theorem 6.** *Given a defeasible knowledge base $\mathcal{K}^{\vdash}$, the set of defeasible statements rationally normally entailed by $\mathcal{K}^{\vdash}$ is exactly the normal closure of $\mathcal{K}^{\vdash}$.*

So from Theorem 6 it follows that rational normal entailment generates a consequence relation that is normal, but is not always rational. This is similar to a result obtained for rational consequence relations [6]. The following proposal to define and construct a viable notion of rational normal closure is analogous to that proposed by Britz et al. [6] which, in turn, is based on the proposal by Lehmann and Magidor [14].

**Definition 13.** *Let $\mathcal{K}^{\vdash}$ be a defeasible knowledge base. The preference order $\ll$ generated by $\mathcal{K}^{\vdash}$ is a binary relation on the set of rational normal consequence relations containing $\mathcal{K}^{\vdash}$, 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 an $\alpha \mathrel{|\!\sim} \beta \in \mathrel{|\!\sim}_1 \setminus \mathrel{|\!\sim}_0$ such that for all $\gamma$ such that $\gamma \vee \alpha \mathrel{|\!\sim}_0 \neg\alpha$ and for all $\delta$ 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 $\rho \vee \gamma \mathrel{|\!\sim}_1 \neg\gamma$.*

The idea is to define rational normal closure as the most preferred (with respect to $\ll$) of all the rational normal consequence relations containing $\mathcal{K}^{\vdash}$.

**Definition 14.** *Let $\mathcal{K}^{\vdash}$ be a defeasible knowledge base, let $\mathcal{K}^R$ be the class of rational normal consequence relations containing $\mathcal{K}^{\vdash}$, and let $\ll$ be the preference ordering on $\mathcal{K}^R$ generated by $\mathcal{K}^{\vdash}$. If $\ll$ has a (unique) minimum element $\mathrel{|\!\sim}$, then the rational normal closure of $\mathcal{K}^{\vdash}$ is defined as $\mathrel{|\!\sim}$.*

In order to provide the conditions for the existence of rational normal closure, we first need to define a *ranking* of formulas with respect to $\mathcal{K}^{\vdash}$ which, in turn, is based on a notion of *exceptionality*. A formula $\alpha$ is said to be *exceptional* for a defeasible knowledge base $\mathcal{K}^{\vdash}$ if and only if $\mathcal{K}^{\vdash}$ *normally* entails $\top \mathrel{|\!\sim} \neg\alpha$. A defeasible statement $\alpha \mathrel{|\!\sim} \beta$ is exceptional for $\mathcal{K}^{\vdash}$ if and only if its antecedent $\alpha$ is exceptional for $\mathcal{K}^{\vdash}$.

Let $E(\mathcal{K}^{\vdash})$ denote the subset of $\mathcal{K}^{\vdash}$ containing statements that are exceptional for $\mathcal{K}^{\vdash}$. We define a non-increasing sequence of subsets of $\mathcal{K}^{\vdash}$ as follows: $\mathcal{E}_0 = \mathcal{K}^{\vdash}$, 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 $\mathcal{K}^{\vdash}$ as follows:[3] $r_{\mathcal{K}^{\vdash}}(\alpha)$ is the smallest integer $i$ such that $\alpha$ is not exceptional for $\mathcal{E}_i$. If $\alpha$ is exceptional for $\mathcal{E}_k$ (and therefore exceptional for all $\mathcal{E}$s), then $\alpha$ does not have a rank (denoted as $r_{\mathcal{K}^{\vdash}}(\alpha) = \infty$). Intuitively, the higher the rank of a formula, the more exceptional it is with respect to $\mathcal{K}^{\vdash}$.

**Theorem 7.** *Let $\mathcal{K}^{\vdash}$ be a defeasible knowledge base. The rational normal closure of $\mathcal{K}^{\vdash}$ exists and is the set $\mathcal{R}^{\vdash}$ of defeasible statements $\alpha \hspace{1pt}\mid\hspace{-3pt}\sim \beta$ such that either $r_{\mathcal{K}^{\vdash}}(\alpha) < r_{\mathcal{K}^{\vdash}}(\alpha \wedge \neg\beta)$, or $r_{\mathcal{K}^{\vdash}}(\alpha) = \infty$ (in which case $r_{\mathcal{K}^{\vdash}}(\alpha \wedge \neg\beta) = \infty$).*

We conclude this section by observing that exceptionality checking for normal entailment *cannot* be reduced to (local) classical entailment, as is the case for preferential entailment. More precisely, given a defeasible knowledge base $\mathcal{K}^{\vdash}$, let $\mathcal{K}^{\rightarrow}$ be its classical counterpart in which every defeasible statement of the form $\gamma \hspace{1pt}\mid\hspace{-3pt}\sim \delta$ in $\mathcal{K}^{\vdash}$ is replaced by $\gamma \rightarrow \delta$. It can be shown that $\mathcal{K}^{\vdash}$ preferentially entails $\neg\alpha$ if and only if $\neg\alpha$ is (locally) entailed by $\mathcal{K}^{\rightarrow}$ [6]. And while it is easy to show that $\neg\alpha$ being (locally) entailed by $\mathcal{K}^{\rightarrow}$ implies that $\mathcal{K}^{\vdash}$ normally entails $\neg\alpha$ (i.e., $\alpha$ is exceptional for $\mathcal{K}^{\vdash}$), it is just as easy to construct a counterexample which shows that the converse does not always hold.

On the one hand the result above is a negative one as it rules out a reduction to classical entailment for computing rational normal closure. On the other hand it is of theoretical importance since it is a concrete indication that normal rational consequence is a *true* extension of propositional defeasible consequence.

## 6   Concluding Remarks

Recapitulating the main contributions of this paper, they can be summarized as follows: (*i*) We have provided concrete evidence that the move from propositional to modal-based defeasible consequence relations bring about gaps that the original KLM postulates are not able to cope with; (*ii*) We have tightened our preferential semantics for modal logic by motivating and defining additional constraints on preferential models; (*iii*) We have motivated extra KLM-style postulates that do make use of the additional expressiveness of modal logic; (*iv*) We have proved new representation theorems establishing the link between the semantic constraints and the new set of postulates, and (*v*) We have extended the notion of rational closure for the case of normal consequence relations.

Crocco and Lamarre [10] as well as Boutilier [4] have also investigated defeasible consequence in a modal context. In particular, Boutilier showed that (propositional) nonmonotonic consequence can be embedded in conditional logics via a binary modality $\Rightarrow$. The links between our richer framework and the conditional $\Rightarrow$ remain to be explored in more detail, though.

---

[3] Observe that our terminology differs from that of Britz et al. [6], but is consistent with that of Lehmann and Magidor [14].

Our representation result paves the way for both the investigation into further modal properties and the definition of effective decision procedures for modal preferential reasoning. Another avenue for future research is the integration of the refined approach here presented with notions of *typicality* [3] and *defeasible modalities* [8], thereby establishing the foundation of a general framework for modal defeasible reasoning.

## Acknowledgements

## References

1. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P. (eds.): The Description Logic Handbooks. Cambridge University Press (2007)
2. Blackburn, P., van Benthem, J., Wolter, F.: Handbook of Modal Logic. Elsevier North-Holland (2006)
3. Booth, R., Meyer, T., Varzinczak, I.: PTL: A propositional typicality logic. In: Proc. 13th European Conference on Logics in Artificial Intelligence (JELIA) (2012)
4. Boutilier, C.: Conditional logics of normality: A modal approach. Artificial Intelligence 68(1), 87–154 (1994)
5. Britz, K., Heidema, J., Meyer, T.: Semantic preferential subsumption. In: Proc. of KR. pp. 476–484 (2008)
6. Britz, K., Meyer, T., Varzinczak, I.: Preferential reasoning for modal logics. Electronic Notes in Theoretical Computer Science 278, 55–69 (2011)
7. Britz, K., Meyer, T., Varzinczak, I.: Semantic foundation for preferential description logics. In: Proc. of Australasian AI Conference. pp. 491–500 (2011)
8. Britz, K., Varzinczak, I.: Defeasible modes of inference: A preferential perspective. In: 14th International Workshop on Nonmonotonic Reasoning (NMR) (2012)
9. Casini, G., Straccia, U.: Rational closure for defeasible description logics. In: Proc. JELIA. pp. 77–90 (2010)
10. Crocco, G., Lamarre, P.: On the connections between nonmonotonic inference systems and conditional logics. In: Proc. of KR. pp. 565–571 (1992)
11. Giordano, L., Olivetti, N., Gliozzi, V., Pozzato, G.: $\mathcal{ALC}+T$: a preferential extension of description logics. Fundamenta Informaticae 96(3), 341–372 (2009)
12. Kraus, S., Lehmann, D., Magidor, M.: Nonmonotonic reasoning, preferential models and cumulative logics. Artificial Intelligence 44, 167–207 (1990)
13. Lehmann, D., Magidor, M.: Preferential logics: the predicate calculus case. In: Proceedings of TARK. pp. 57–72 (1990)
14. Lehmann, D., Magidor, M.: What does a conditional knowledge base entail? Artificial Intelligence 55, 1–60 (1992)
15. Moodley, K., Meyer, T., Varzinczak, I.: A protégé plug-in for defeasible reasoning. In: Proceedings of the 25th International Workshop on Description Logics (2012)
16. Quantz, J.: A preference semantics for defaults in terminological logics. In: Proc. of KR. pp. 294–305 (1992)