

Model-Based Defeasible Reasoning Literature Review

Jaron Cohen

University of Cape Town
Cape Town, South Africa
CHNJAR003@myuct.ac.za

ABSTRACT

Defeasible reasoning is a form of non-classical reasoning that formalises common sense patterns of dealing with exceptions to explicit information. Reasoning systems developed for classical propositional logics have been studied extensively in the literature with many efficient implementations having been developed. On the other hand, comparatively fewer approaches to developing defeasible reasoning systems have been attempted and implemented. We therefore review the current approaches to computing defeasible entailment through the lens of one of the most prominent frameworks - the KLM framework. Our review elicits a gap in the literature for the design and feasibility of novel 'bottom-up' model-based methods for implementing defeasible reasoning.

CCS CONCEPTS

• **Theory of computation** → **Automated reasoning**; • **Computing methodologies** → **Nonmonotonic, default reasoning and belief revision**.

KEYWORDS

Artificial Intelligence, Knowledge Representation and Reasoning, Defeasible Reasoning, Propositional Logic, Rational Closure, Lexicographic Closure

1 INTRODUCTION AND MOTIVATION

Knowledge Representation and Reasoning (KRR) is a subfield of Artificial Intelligence (AI) that takes a formal and at times philosophical approach to the problem of simulating intelligence. At the heart of this subfield is the notion of a knowledge-based system, where information is syntactically encoded in the form of a symbolic structure or *knowledge base* that allows for the development of reasoning services or mechanical operations to derive conclusions from said structure.

Approaches to KRR can be both logic-based and non-logic based [1]. Our work will build on other logic-based approaches. A logic or logical system consists of two parts, its language consisting of both a syntax and a semantics, as well as a method of reasoning [11]. There are a number of logical systems that one can choose from that vary in terms of their expressivity and complexity. Our work utilises propositional logic as the foundation for both representation and reasoning. In terms of reasoning, we will focus on the aspects of how new information or knowledge is derived from existing knowledge - how implicit inferences are drawn from explicit information. This is often referred to simply as the notion of entailment or logical consequence.

Classical logics are monotonic and therefore lack the ability to explicitly express typicality or reason with uncertainty as humans do. Thus we turn to a class of logics referred to as nonmonotonic

logics that have the ability to retract certain conclusions upon the addition of new knowledge. We refer to their associated reasoning services as defeasible reasoning. Unlike classical entailment, there is no single answer for what defeasible entailment should look like or how it ought to behave. Thus there are many patterns of defeasible reasoning. Kraus, Lehmann and Magidor (KLM), proposed a set of properties as a thesis for how to define a 'sensible' or 'rational' notion of defeasible entailment. The focus of this review will therefore be to look at the framework proposed by KLM, its extensions and current KLM-style approaches to defeasible reasoning due to their desirable theoretical and computational properties [12].

2 PROPOSITIONAL LOGIC

2.1 Motivation

Propositional logic is sometimes referred to as zeroth-order logic, which adequately conveys its position as the foundation of many other logic systems. Hence, it often serves as a useful testing ground for ideas within KRR as results can be easily extended to more expressive logics such as first order logic, modal logics and description logics.

2.2 Syntax

The language of propositional logic consists of *atomic propositions* or just *atoms* [2], which are combined together using logical connectives to form what are referred to as *formulas*. Each atom at a fundamental level represent indivisible facts about the world or domain one is attempting to represent and can be either true or false. \mathcal{P} is used to denote a set, $\{p, q, r, \dots\}$, of meta-variables for such atomic propositions. For example, we may represent the atom, 'the sun is shining', with s .

There are a number of logical connectives, one unary, negation (\neg), and four binary: conjunction (\wedge), disjunction (\vee), material implication (\rightarrow) and bi-implication (\leftrightarrow). The logical connectives from highest to lowest precedence are: \neg , \wedge , \vee , \rightarrow and \leftrightarrow . Greek letters α, β, \dots are typically used to represent formulas, which can be recursively defined by the following grammar [2]:

$$\begin{aligned} \langle \text{formula} \rangle &::= p \in \mathcal{P} \mid \top \mid \perp \\ \langle \text{formula} \rangle &::= \neg \langle \text{formula} \rangle \\ \langle \text{formula} \rangle &::= \langle \text{formula} \rangle \langle \text{op} \rangle \langle \text{formula} \rangle \\ \langle \text{op} \rangle &::= \vee \mid \wedge \mid \rightarrow \mid \leftrightarrow \end{aligned}$$

We will use, \mathcal{L} , to refer to the set of all well-formed formulas. Note, we also include the constants \top and \perp in \mathcal{L} , which denote a formula that is always true and a formula that is always false, respectively. We refer to a finite set of formulas as a knowledge base, often denoted by \mathcal{K} .

Now consider a small example of how we can use propositional logic to represent knowledge about the world:

Example 2.1. We have the following meta-variables,

$$\mathcal{P} = \{m, l, p\},$$

that represent the atomic propositions “being a mammal”, “giving birth to live young” and “primates” respectively. We can then encode the knowledge:

- (1) Mammals give birth to live young ($m \rightarrow l$)
- (2) Primates are mammals ($p \rightarrow m$)

in a knowledge base $\mathcal{K} = \{m \rightarrow l, p \rightarrow m\}$.

2.3 Semantics

In much the same way that an arithmetic expression, such as $a^2 + b^2 = c^2$, is true or false depending on the values assigned to the variables a , b and c , the truth of a propositional formula is dependent on the values assigned to its constituent atoms [2]. An *interpretation* is a function $\mathcal{I} : \mathcal{P} \mapsto \{T, F\}$, which assigns the values of T (true) and F (false) to every atom in \mathcal{P} with \mathcal{U} being used to denote the set of all interpretations. We tend to represent interpretations as a sequence of atoms where an atom with a bar above (e.g. \bar{p}) is read as being false and true otherwise. For example, $p\bar{q}$ represents the interpretation where p is true and q is false.

We can then recursively evaluate the truth of a formula using the notion of *satisfaction*. An interpretation \mathcal{I} *satisfies* a formula α , denoted $\mathcal{I} \models \alpha$, if α recursively evaluates to true using the Boolean semantics of the operators it is composed of, in which case \mathcal{I} is said to be a *model* of α . Formally, $\mathcal{I} \models \alpha$ if and only if one of the following conditions holds:

- $\alpha \in \mathcal{P}$ and $\mathcal{I}(\alpha) = T$
- $\alpha = \neg\beta$ and \mathcal{I} does not satisfy β
- $\alpha = \beta \wedge \gamma$ and both $\mathcal{I} \models \beta$ and $\mathcal{I} \models \gamma$
- $\alpha = \beta \vee \gamma$ and at least one of $\mathcal{I} \models \beta$ or $\mathcal{I} \models \gamma$
- $\alpha = \beta \rightarrow \gamma$ and at least one of $\mathcal{I} \models \neg\beta$ or $\mathcal{I} \models \gamma$
- $\alpha = \beta \leftrightarrow \gamma$ and either both $\mathcal{I} \models \beta$ and $\mathcal{I} \models \gamma$ or \mathcal{I} satisfies neither β nor γ

We denote the set of models of a given formula α as $Mod(\alpha)$ or $\llbracket \alpha \rrbracket$. Similarly, an interpretation \mathcal{I} is a model of a knowledge base \mathcal{K} if it is a model of every formula in \mathcal{K} , that is $Mod(\mathcal{K}) = \cap \{Mod(\alpha) \mid \alpha \in \mathcal{K}\}$.

We illustrate these concepts with another example:

Example 2.2. Consider Example 2.1 from before. A possible interpretation for the meta-variables in \mathcal{P} is

$$m\bar{l}\bar{p} \in \mathcal{U},$$

and we can see that $m\bar{l}\bar{p} \models m \rightarrow l$, i.e. $m\bar{l}\bar{p} \in Mod(m \rightarrow l)$. Furthermore, we can see that $m\bar{l}\bar{p} \not\models \mathcal{K}$, i.e. $m\bar{l}\bar{p} \notin Mod(\mathcal{K})$.

We briefly note the distinction between what are referred to as the *object-language* and the *meta-language*. The object-language essentially refers to the formulas used to represent knowledge itself, i.e. \mathcal{L} , while the meta-language is a separate set of symbols with which we use to reason about the knowledge on the object-level. Satisfaction (\models) and entailment (\models) are therefore meta-level concepts, whilst material implication (\rightarrow) for example is an object-level connective.

2.4 Classical Reasoning

We are now ready give a model-theoretic definition of the meta-level notion of *entailment* (or *logical consequence*) for classical propositional logic. We say that a formula, α , entails a formula, β , denoted $\alpha \models \beta$, if and only if $Mod(\alpha) \subseteq Mod(\beta)$. Intuitively, this says that knowing that α is true is enough to conclude that β is true as well. For example, $\alpha \wedge \beta \models \alpha$.

By extension, for a knowledge base, \mathcal{K} , $\mathcal{K} \models \alpha$ if and only if $Mod(\mathcal{K}) \subseteq Mod(\alpha)$. In the case that \mathcal{K} does not entail α , that is $Mod(\mathcal{K}) \not\subseteq Mod(\alpha)$, we represent this as $\mathcal{K} \not\models \alpha$. Consider the following small example:

Example 2.3. Let $\mathcal{K} = \{p \rightarrow q, \neg p\}$. Therefore $\mathcal{K} \models \neg p$ since

$$Mod(\mathcal{K}) = \{\bar{p}\bar{q}\} \subseteq \{\bar{p}\bar{q}, \bar{p}q\} = Mod(\neg p).$$

In the context of Example 2.1, we can sensibly conclude using the provided definition of entailment, that primates give birth to live young ($p \rightarrow l$) since $\mathcal{K} = \{m \rightarrow l, p \rightarrow m\} \models p \rightarrow l$.

3 DEFEASIBLE REASONING

3.1 Motivation

In terms of the AI concept of an intelligent agent, and particularly a knowledge-based agent as described by Craik [7], we are concerned with its ability to make rational or sensible actions based on some internal representation of knowledge. Classical logics and the their reasoning services, based on Tarskian notions of consequence [16], obey the property of *monotonicity*, which simply states that adding information or new knowledge to a knowledge base cannot result in the retraction of any conclusions that could have been entailed before such information was added. As a result, classical logics are impractical for representing and reasoning with exceptional knowledge since they have no means of explicitly expressing the notion of *typicality*. We illustrate this with a similar mammalian example:

Example 3.1. We have the following meta-variables

$$\mathcal{P} = \{m, l, p\}$$

that represent the atomic propositions “being a mammal”, “giving birth to live young” and “being a platypus” respectively. We can encode the knowledge:

- (1) Mammals give birth to live young ($m \rightarrow l$)
- (2) Platypuses are mammals ($p \rightarrow m$)

in a knowledge base $\mathcal{K} = \{m \rightarrow l, p \rightarrow m\}$.

Using this knowledge, we can conclude using classical entailment that platypuses give birth to live young. However this is not an accurate reflection of reality since platypuses are exceptional mammals that lay eggs. Thus, we can add the formula

$$p \rightarrow \neg l$$

to \mathcal{K} to reflect the fact they do not give birth to live young. Unfortunately, classical reasoning will force us to conclude that platypuses cannot exist, since any model of \mathcal{K} will require the atom p to be false.

In theory, one can adjust the knowledge base to specifically address platypuses as being exceptional, but this quickly becomes impractical as each further exception, for example the echidna,

would warrant such a remodelling of the knowledge base. There is no convenient mechanism for an agent using classical logic to reconcile its ‘beliefs’ about the world when it is presented with new information inconsistent with its prior ‘beliefs’ [12].

Example 3.1 illustrates the inability to express the notion of typicality with classical logic alone. In essence, we wanted to express the notion that mammals *typically* give live birth and still have the ability to reason about the existence of platypuses.

Whereas classical logics have a unique and well-defined notion of entailment, there are many forms of defeasible entailment across many different nonmonotonic formalisms. An overview of some of these competing formalisms can be found in [12]. Defining a notion of defeasible entailment is often contentious in practice and there is likely no one-size-fits-all solution, however as we will describe in the next section, KLM introduced a set of properties that they felt sensible defeasible entailment relations should satisfy.

3.2 The KLM Framework

Shoham introduced *preferential semantics* for nonmonotonic reasoning in [18–20] with the idea being that an agent performing a form of nonmonotonic reasoning would have a preference for interpretations that assign more ‘typical’ values to the atoms in question. The seminal KLM paper [13] set out to characterise the preferential model-theoretic approach to nonmonotonic entailment taken by Shoham in proof-theoretic terms applied to *consequence relations*, inspired by the work of Gabbay in [9].

Consequence Relations (CRs) provide an abstract mathematical view of the notion of entailment or logical consequence and correspond to the implicit information an intelligent agent may have [15]. Hence, KLM initially introduced, \sim , in [13] as a binary relation on a language \mathcal{L} , i.e. $\sim \subseteq \mathcal{L} \times \mathcal{L}$, to consider nonmonotonic reasoning in terms of CRs on the meta-level atop classical logics. CRs are flexible by design in order to define many patterns of reasoning. Although there are many CRs for a given language, not every CR corresponds to a sensible form of reasoning.

It proves necessary to consider restricted subsets of CRs that have desirable properties. As such, KLM did not initially propose a specific form of nonmonotonic reasoning, but in fact explored five logical systems and their corresponding CRs that satisfy varying properties. Most notably is the preferential logical system **P**, for which they give an axiomatic description and refine the semantics initially proposed by Shoham in [19]. This was done to elucidate the various closure properties that \sim should satisfy, of which six properties, often referred to as the KLM postulates were presented for the logical system **P**. It is this very preferential approach that forms the basis of what is referred to as the KLM framework.

KLM originally proposed preferential semantics for \sim in the form of *preferential interpretations* and proved a representation theorem linking preferential interpretations to *preferential consequence relations* in [13]. These were general enough to capture the preferential ordering intuitions proposed by Shoham but were not restrictive enough to define all reasonable defeasible entailment relations [15].

Lehmann and Magidor later extended propositional logic with, \sim , as a now more familiar object-level connective in [15] and elicited a seventh property to incorporate the idea that such CRs should be

‘rational’ and obey the property of Rational Monotonicity. They also refined preferential interpretations to consider a restricted family of preferential interpretations referred to as *ranked interpretations* and proved a similar representation theorem linking them to *rational consequence relations* [15]. In this context, two entailment relations were initially explored, *preferential entailment* and *ranked entailment*. Both of which, were based on Tarskian notions of logical consequence and were shown to be equivalent and still monotonic - thus not sufficient for true defeasible reasoning [12].

The first true nonmonotonic entailment relation, *Rational Closure* (RC), was also proposed by Lehmann and Magidor in [15] as a foundational form of rational defeasible entailment alongside the thesis that any reasonable defeasible entailment relation should entail at least as many conclusions as RC. Hence, RC is consistent with a conservative form of reasoning known as *prototypical reasoning* where in linguistic terms, objects that are *typical* examples inherit properties by default.

Lehmann expands on this notion in [14] and proposed Lexicographic Closure (LC) partially to address the final thesis presented by Lehmann and Magidor in [15] but also to construct a defeasible entailment relation capturing the intuitive ideas from Reiter’s Default Logic in [17]. As such, LC is superset of RC and corresponds to a bolder form of reasoning referred to as *presumptive reasoning* that grants properties to an object as long as there is nothing to contradict such a conclusion [6, 14].

Consider Example 3.1, where we have that “mammals give birth to live young”, “platypuses are mammals” and “platypuses do not give birth to live young”. If we then added that “mammals are typically warm-blooded”, one would *not* prototypically conclude that platypuses are warm-blooded since they are an atypical mammal, whereas one could presumptively conclude that they are since there is no current information to suggest otherwise. We will demonstrate that this is indeed the case when we describe RC and LC.

3.3 KLM-style Syntax and Semantics

Casini et al. presented a systematic approach for extending the KLM framework for defeasible entailment in [6]. In this approach, Casini et al. enrich the language of propositional logic with an additional object-level connective, \sim , and propose *rational defeasible entailment* as the refinement needed to characterise sensible defeasible entailment relations that uphold rational closure as the nonmonotonic core of defeasible entailment [6].

3.3.1 Syntax. Given α and β from the language of propositional logic, \mathcal{L} , we now allow for formulas of the form $\alpha \sim \beta$ and refer to them as KLM-style *defeasible implications* (DIs) [6]. The connective, \sim , can be interpreted as the defeasible complement to material implication \rightarrow and read as saying “ α typically implies β ”. We clarify that nesting of the connective, \sim , is not allowed. More formally, we can define this extension of \mathcal{L} as $\mathcal{L}\mathcal{P} = \mathcal{L} \cup \{\alpha \sim \beta \mid \alpha, \beta \in \mathcal{L}\}$ [12].

We briefly note that ideas of increasing the expressivity of typicality on the object-level have been explored in the context of Propositional Typicality Logic (PTL) [3], which builds on the same KLM approach to defeasible reasoning.

Example 3.2. Consider Example 3.1 from earlier, we are now able to express that mammals typically give live birth ($m \sim l$), that

platypuses are mammals ($p \rightarrow m$) and that platypuses typically do not give birth to live young ($p \vdash \neg l$), on the object-level, in a knowledge base $\mathcal{K} = \{m \vdash l, p \rightarrow m, p \vdash \neg l\}$.

3.3.2 Semantics. We provide semantics for KLM-style DIs through structures referred to as *ranked interpretations*. A ranked interpretation, is a function $\mathcal{R} \mapsto \mathbb{N} \cup \{\infty\}$, such that $\mathcal{R}(u) = 0$ for some $u \in \mathcal{U}$, and satisfying the *convexity* property: for every $i \in \mathbb{N}$, if $\mathcal{R}(v) = i$, then, for every j s.t. $0 \leq j < i$, there is a $u \in \mathcal{U}$ for which $\mathcal{R}(u) = j$ [6]. We sometimes use $\mathcal{R}^{\mathcal{K}}$ to refer to all the ranked interpretations w.r.t. a knowledge base \mathcal{K} [6]. Intuitively, a ranked interpretation is an ordering of all the interpretations, \mathcal{U} , in order of their ‘typicality’, from rank 0 to rank n , with a reserved infinite rank, ∞ . Those interpretations assigned a ranking of 0 therefore correspond to the most typical interpretations in the eyes of an agent with interpretations of rank ∞ being deemed ‘impossible’. Hence, we sometimes use $\mathcal{U}^{\mathcal{R}} := \{u \in \mathcal{U} \mid \mathcal{R}(u) < \infty\}$ to represent the ‘possible’ interpretations w.r.t \mathcal{R} [6]. Given a DI, $\alpha \sim \beta$, we say that \mathcal{R} satisfies $\alpha \sim \beta$, denoted as expected $\mathcal{R} \Vdash \alpha \sim \beta$, if in the lowest rank in which α is satisfied, β holds in all such models of α within this rank. In which case, \mathcal{R} is said to be a model of $\alpha \sim \beta$ [6]. Intuitively, we have that $\mathcal{R} \Vdash \alpha \sim \beta$ iff $\alpha \rightarrow \beta$ holds in all the ‘best’ or most typical worlds of α .

Given some formula $\alpha \in \mathcal{L}$, we say that \mathcal{R} satisfies α , denoted $\mathcal{R} \Vdash \alpha$, if α is satisfied by *all* the *possible* interpretations w.r.t. \mathcal{R} , that is $\forall u \in \mathcal{U}^{\mathcal{R}}, u \Vdash \alpha$. Intuitively, this is reasonable since classical propositional formulas do not express any degree of typicality.

We note that, by virtue of these semantics, any $\alpha \in \mathcal{L}$ can be expressed as an equivalent DI $\neg\alpha \vdash \perp$ [5, 6]. As such, a knowledge base, \mathcal{K} , will now be said to contain a set of DIs. Since these are somewhat interchangeable syntactic forms, we will not enforce this constraint for the sake of readability in our examples.

Example 3.3. By the convexity property of the order on ranked interpretations, we can graphically represent them in the form of a table with no empty rows. For example, here is a ranked interpretation for \mathcal{K} in Example 3.2:

∞	$\overline{pml} \overline{pml}$
2	$pml \overline{pml}$
1	$\overline{pml} \overline{pml} \overline{pml}$
0	\overline{pml}

Figure 1: Ranked Interpretation for Example 3.3

In this particular ranked interpretation, we can see that \overline{pml} and \overline{pml} are impossible since $\mathcal{R}(\overline{pml}) = \mathcal{R}(pml) = \infty$, whilst $\mathcal{R}(pml) = 0$ and therefore \overline{pml} is most preferred and represents the most typical world.

3.4 KLM-style Defeasible Entailment

Although the original six KLM postulates proposed by KLM in [13], together with the seventh proposed by Lehmann and Magidor in [15], were used to characterise a restricted class of *rational* consequence relations using \sim on the meta-level, we now present

a reformulation of the seven postulates by Casini et al. in [6] to characterise defeasible entailment on the meta-level, denoted by \approx :

$$\begin{array}{ll}
 \text{(Ref)} & \mathcal{K} \approx \alpha \sim \alpha \\
 \text{(LLE)} & \frac{\mathcal{K} \approx \alpha \leftrightarrow \beta, \mathcal{K} \approx \alpha \sim \gamma}{\mathcal{K} \approx \beta \sim \gamma} \quad \text{Or} \quad \frac{\mathcal{K} \approx \alpha \sim \gamma, \mathcal{K} \approx \beta \sim \gamma}{\mathcal{K} \approx \alpha \wedge \beta \sim \gamma} \\
 \text{(RW)} & \frac{\mathcal{K} \approx \alpha \rightarrow \beta, \mathcal{K} \approx \gamma \sim \alpha}{\mathcal{K} \approx \gamma \sim \beta} \quad \text{(CM)} \quad \frac{\mathcal{K} \approx \alpha \sim \gamma, \mathcal{K} \approx \alpha \sim \beta}{\mathcal{K} \approx \alpha \wedge \beta \sim \gamma} \\
 \text{(RM)} & \frac{\mathcal{K} \approx \alpha \sim \gamma, \mathcal{K} \approx \alpha \sim \neg \beta}{\mathcal{K} \approx \alpha \wedge \beta \sim \gamma}
 \end{array}$$

We refer to a defeasible entailment relation, \approx , that satisfies these properties as being LM-rational. Rational Closure and Lexicographic Closure are both LM-rational defeasible entailment relations [6]. For completeness sake, Relevant Closure proposed by Casini et al. in [4] for description logics, but easily extended to propositional logic, is an example of a sensible form of defeasible entailment that is stronger than Rational Closure but not LM-rational [4, 6].

3.5 Rational Closure

We first give a semantic characterisation of RC by ranking the ranked interpretations themselves and showing that RC can be defined according to the ranked interpretation that is minimal according to this ordering. Thereafter, we give a syntactic characterisation of RC that assigns ranks to the DIs within a given knowledge base, which allows for the definition of an algorithm to compute defeasible entailment queries.

3.5.1 Semantic Definition - Minimal Ranked Entailment. Given a knowledge base, \mathcal{K} , Casini et al. [6] define a partial order, $\leq_{\mathcal{K}}$, over all its ranked models as follows: for $\mathcal{R}_1, \mathcal{R}_2 \in \text{Mod}(\mathcal{K})$, we say $\mathcal{R}_1 \leq_{\mathcal{K}} \mathcal{R}_2$ if and only if $\forall v \in \mathcal{U}, \mathcal{R}_1(v) \leq_{\mathcal{K}} \mathcal{R}_2(v)$. Intuitively, the ranked models themselves are partially ordered in terms of their overall typicality, with the unique minimal ranked model, i.e. the most conservative, being used to define RC. Giordano et al. [10] showed that such a unique $\leq_{\mathcal{K}}$ -minimal model, denoted $\mathcal{R}_{\mathcal{K}}^{RC}$, exists. We are then able to define the entailment relation for RC, denoted \approx_{RC} , as follows: $\mathcal{K} \approx_{RC} \alpha \sim \beta$ if and only if $\mathcal{R}_{\mathcal{K}}^{RC} \Vdash \alpha \sim \beta$ [6]. In other words, the rational closure of a knowledge base is defined as the set of formulas satisfied by a single ranked interpretation, $\mathcal{R}_{\mathcal{K}}^{RC}$. By the representation theorem for rational defeasible entailment relations and ranked interpretations proven by Lehmann and Magidor in [15], we can conclude that RC is LM-rational.

Example 3.4. We now consider a small example to illustrate how one is able to use $\mathcal{R}_{\mathcal{K}}^{RC}$ to answer a defeasible entailment query. Consider the knowledge base \mathcal{K} from Example 3.2, with the addition that mammals are typically warm-blooded ($m \sim w$): $\mathcal{K} = \{m \vdash l, p \rightarrow m, p \vdash \neg l, m \sim w\}$. It is then possible to find its, unique $\leq_{\mathcal{K}}$ -minimal model, $\mathcal{R}_{\mathcal{K}}^{RC}$, for which we include its graphical representation below (we leave out the infinite rank for compactness).

2	$pml\bar{w} pmlw$
1	$\bar{p}ml\bar{w} \bar{p}ml\bar{w} \textcircled{pml\bar{w}} \textcircled{pml\bar{w}} \bar{p}ml\bar{w}$
0	$\bar{p}ml\bar{w} \bar{p}ml\bar{w} \bar{p}ml\bar{w} \bar{p}ml\bar{w} \bar{p}ml\bar{w}$

Figure 2: Representation of $\mathcal{R}_{\mathcal{K}}^{RC}$ for Example 3.4

Now suppose we wanted to confirm whether platypuses are typically warm-blooded ($p \sim w$). Then we must consider all the best p worlds in $\mathcal{R}_{\mathcal{K}}^{RC}$ and check whether they are worlds in which w is true as well. We circle such worlds on Figure 2. After doing so, we can conclude that $\mathcal{R}_{\mathcal{K}}^{RC} \not\models p \sim w$, which means that $p \sim w$ is not in the rational closure of \mathcal{K} , or $\mathcal{K} \not\models_{RC} p \sim w$, and can be interpreted as saying that platypuses are not typically warm-blooded.

3.5.2 Syntactic Definition - Base Ranks. First we must define some preliminary terminology. Given a DI, $\alpha \sim \beta$, we refer to $\alpha \rightarrow \beta$ as its *material counterpart*. Hence, given a defeasible knowledge base, \mathcal{K} , its material counterpart or *materialisation*, $\vec{\mathcal{K}}$, can be defined as:

$$\vec{\mathcal{K}} := \{\alpha \rightarrow \beta \mid \alpha \sim \beta \in \mathcal{K}\}.$$

A propositional formula, $\alpha \in \mathcal{L}$, is said to be *exceptional* w.r.t. a defeasible knowledge base, \mathcal{K} , if and only if $\vec{\mathcal{K}} \models \neg\alpha$ [6, 15]. In other words, α is false in the most preferred or typical (rank 0) interpretations (but may be true in others) in every ranked model of \mathcal{K} [6, 12]. Hence we can define a function, ε , that maps a knowledge base to its subset of exceptional formulas:

$$\varepsilon(\mathcal{K}) := \{\alpha \sim \beta \in \mathcal{K} \mid \text{such that } \vec{\mathcal{K}} \models \neg\alpha\}.$$

This allows us to recursively define a non-increasing sequence of knowledge bases, $\mathcal{E}_0^{\mathcal{K}}, \dots, \mathcal{E}_n^{\mathcal{K}}$, by defining $\mathcal{E}_0^{\mathcal{K}} := \mathcal{K}$ and then setting $\mathcal{E}_i^{\mathcal{K}} = \varepsilon(\mathcal{E}_{i-1}^{\mathcal{K}})$ for $0 < i < n$, with $\mathcal{E}_\infty^{\mathcal{K}} := \mathcal{E}_n^{\mathcal{K}}$, where n is the smallest i such that $\mathcal{E}_i^{\mathcal{K}} = \mathcal{E}_{i+1}^{\mathcal{K}}$ [5].

The *base rank*, $br_{\mathcal{K}}(\alpha)$, of a propositional formula α or the *antecedent* of a DI $\alpha \sim \beta$, can then be defined as the minimal index i such that α is *not* exceptional w.r.t. the sequence $\mathcal{E}_i^{\mathcal{K}}$ [5]. Formally:

$$br^{\mathcal{K}}(\alpha) := \min\{i \mid \mathcal{E}_i^{\mathcal{K}} \not\models \neg\alpha\}$$

Giordano et al. proved in [10] an alternative definition for the Rational Closure in terms of base ranks: Given $\alpha \sim \beta$,

$$\mathcal{K} \models_{RC} \alpha \sim \beta \text{ iff } br^{\mathcal{K}}(\alpha) < br^{\mathcal{K}}(\alpha \wedge \neg\beta) \text{ or } br^{\mathcal{K}}(\alpha) = \infty.$$

3.5.3 Algorithm. We now are in a position to define an algorithm in two parts, BaseRank and RationalClosure, that is able to answer defeasible entailment queries of the form $\alpha \sim \beta$ as described by Casini et al. in [6].

BaseRank takes as input a knowledge base \mathcal{K} and defines a sequence of materialisations, $E_0, \dots, E_{n-1}, E_\infty$ that are used to partition $\vec{\mathcal{K}}$ into a sequence of levels $R_0, \dots, R_{n-1}, R_\infty$ such that R_i corresponds to all the material counterparts of the DIs in \mathcal{K} of base rank i . We can define the sequence of E_i s using our earlier sequence as follows: $\forall i, E_i := \vec{\mathcal{E}}_i^{\mathcal{K}}$. Using this sequence, we can inductively define the sequence of R_i s: $R_i := E_i \setminus E_{i+1}$ for $0 \leq i \leq n-1$ with $R_\infty := E_\infty$.

RationalClosure, then answers the given query, using the sequence of R_i s produced by BaseRank. The algorithm does this by first checking whether $R_0 \cup \dots \cup R_{n-1} \cup R_\infty \models \neg\alpha$ holds (i.e. whether the antecedent, and by extension the query itself, is exceptional w.r.t. the entire knowledge base). If this is indeed the case, the algorithm will repeatedly remove the lowest level, each time performing the same check with the remaining ranks, until an i is found such that $R_i \cup \dots \cup R_{n-1} \cup R_\infty \models \neg\alpha$ holds or until only R_∞ remains. At which point, the algorithm then checks whether the materialisation of the query, i.e. $\alpha \rightarrow \beta$, is entailed by these remaining ranks - returning **true** if it is and **false** if it is not. A result by Freund in [8], guarantees that RationalClosure returns *true* iff $KB \models_{RC} \alpha \sim \beta$.

We finally note that computing RationalClosure is not computationally harder than classical entailment, since its implementation will make use of classical entailment checkers a polynomial number of times in the size of \mathcal{K} [6].

Example 3.5. Suppose again we wanted to answer the query whether platypuses are typically warm-blooded ($p \sim w$) for the knowledge base $\mathcal{K} = \{m \sim l, p \rightarrow m, p \sim \neg l, m \sim w\}$ to corroborate our result from earlier.

Upon giving \mathcal{K} as input to the BaseRank algorithm, we will receive the following ranking of formulas according to their base ranks:

R_∞	$p \rightarrow m$
R_1	$p \rightarrow \neg l$
R_0	$m \rightarrow l, m \rightarrow w$

Figure 3: Partition of \mathcal{K} by Base Ranks

The algorithm RationalClosure will then check whether $R_0 \cup R_1 \cup R_\infty \models \neg p$. Since it does, which means that p is exceptional w.r.t. $\vec{\mathcal{K}}$, we throw away R_0 and check whether $R_1 \cup R_\infty \models \neg p$ holds, which it does not. Hence, we now answer the query by checking whether $R_1 \cup R_\infty \models p \rightarrow w$ holds, which it also does not, thus we reach the same conclusion as in Example 3.4 that $\mathcal{K} \not\models_{RC} p \sim w$.

3.6 Lexicographic Closure

Lexicographic Closure was initially described in terms of *normal defaults*. We will instead describe Lexicographic Closure using the more familiar notation we have used up until this point i.e. knowledge bases containing DIs. Lehmann defines Lexicographic in two ways in [14]: a semantic or model-theoretic definition in terms of a ranked interpretation and another using *bases*. The model-theoretic definition being more relevant to our work. Both aforementioned definitions utilise the idea of imposing a ‘seriousness’ ordering on subsets of the knowledge base. Lehmann proposes to order subsets of the knowledge base by two criteria that in theory give an indication of its ‘seriousness’: its size or cardinality, and the specificity of the elements it contains in terms of base ranks. Each criterion defines its own modular order, hence the two are combined *lexicographically* with the specificity criterion being used as the major criterion [14].

3.6.1 Semantic Definition. The seriousness ordering, once defined for subsets of the defeasible knowledge base, is used to define a

modular ordering of the propositional interpretations themselves, resulting in the ranked interpretation that corresponds to Lexicographic Closure. Intuitively, each subset is associated with its own ‘characteristic’ numeric tuple that is then lexicographically ordered against the ‘characteristic’ numeric tuples of other subsets.

We define the *order of a defeasible knowledge base* \mathcal{K} to be the maximum base rank assigned to a formula in the knowledge base excluding formulas of rank ∞ . Thus, given a knowledge base of order k , we define for each $D \subseteq \mathcal{K}$, a $k+1$ tuple of natural numbers denoted n_D , (n_0, \dots, n_k) , where

$$n_0 = |\{\alpha \sim \beta \in D \mid \text{br}_{\mathcal{K}}(\alpha) = \infty\}|, \text{ and}$$

$$n_i = |\{\alpha \sim \beta \in D \mid \text{br}_{\mathcal{K}}(\alpha) = k - i\}| \text{ for } 1 \leq i \leq k.$$

[12, 14]

Hence, we can now define the seriousness modular ordering, $<_S$, of subsets of \mathcal{K} as follows: given $D_1, D_2 \subseteq \mathcal{K}$, we say that $D_1 <_S D_2$ iff $n_{D_1} < n_{D_2}$ where $<$ denotes the natural lexicographic ordering of tuples of natural numbers.

It is now straightforward to define a modular order on the propositional interpretations themselves that corresponds to the ranked interpretation defining LC. Intuitively, we are defining a preference ordering of the propositional interpretations that favours interpretations that ‘violate’ less serious subsets of DIs [14]. Lehmann defines such a modular order, $<_{LC}$ in [14] as: given $m, n \in \mathcal{U}$, $m <_{LC} n$ iff $V(m) <_S V(n)$ where $V(m) \subseteq \mathcal{K}$ is the set of DIs violated by $m \in \mathcal{U}$. We note that the ranked interpretation defined by this modular order is closely related to the ranked interpretation that defines rational closure.

One can view the ranked interpretation corresponding to lexicographic closure as a refinement of the one that corresponds to rational closure in the sense that the ‘levels’ may be split into a number of ‘sub-levels’ [14] such that the overall order of levels is maintained. Casini et al. [6] take such an approach to constructing the ranked interpretation corresponding to Lexicographic Closure. Assuming one has the ranked interpretation corresponding to RC, $\mathcal{R}_{\mathcal{K}}^{RC}$, we define the order, $\leq_{LC}^{\mathcal{K}}$, on \mathcal{U} as follows: $u, v \in \mathcal{U}$, $u \leq_{LC}^{\mathcal{K}} v$ iff $\mathcal{R}_{\mathcal{K}}^{RC}(u) = \infty$, or $\mathcal{R}_{\mathcal{K}}^{RC}(v) < \mathcal{R}_{\mathcal{K}}^{RC}(u)$, or $\mathcal{R}_{\mathcal{K}}^{RC}(v) = \mathcal{R}_{\mathcal{K}}^{RC}(u)$ and $C^{\mathcal{K}}(v) \geq C^{\mathcal{K}}(u)$ where

$$C^{\mathcal{K}}(v) := |\{\alpha \sim \beta \in \mathcal{K} \mid v \Vdash \alpha \rightarrow \beta\}|.$$

The intuition behind $C^{\mathcal{K}}$ being that it allows one to compare propositional interpretations within levels in $\mathcal{R}_{\mathcal{K}}^{RC}$ by their typicality in terms of the number of DI material counterparts they satisfy with those satisfying more being placed on a lower sub-level.

Once the ranked interpretation, $\mathcal{R}_{LC}^{\mathcal{K}}$, has been constructed, either by means of the seriousness ordering or by means of the function $C^{\mathcal{K}}$, it is straightforward to define the corresponding entailment relation, \models_{LC} as follows: given $\alpha \sim \beta$, $\mathcal{K} \models_{LC} \alpha \sim \beta$ iff $\mathcal{R}_{LC}^{\mathcal{K}} \Vdash \alpha \sim \beta$.

Example 3.6. As a means of comparison with Rational Closure, we consider the same knowledge base from Example 3.4: $\mathcal{K} = \{m \sim l, p \rightarrow m, p \sim \neg l, m \sim w\}$.

Using the semantics provided above, we can construct the ranked model of \mathcal{K} corresponding to Lexicographic Closure, $\mathcal{R}_{LC}^{\mathcal{K}}$, for which we represent graphically below (again we leave off the infinite rank):

5	$pb\bar{f}\bar{w}$
4	$p\bar{b}f\bar{w}$
3	$p\bar{b}\bar{f}\bar{w} \ \bar{p}b\bar{f}\bar{w}$
2	$\bar{p}\bar{b}\bar{f}\bar{w} \ \bar{p}\bar{b}f\bar{w}$
1	$\bar{p}\bar{b}f\bar{w}$
0	$p\bar{b}f\bar{w} \ p\bar{b}\bar{f}\bar{w} \ \bar{p}\bar{b}\bar{f}\bar{w} \ \bar{p}\bar{b}f\bar{w} \ \bar{p}\bar{b}f\bar{w}$

Figure 4: Representation of $\mathcal{R}_{\mathcal{K}}^{LC}$ for Example 3.6

Let us now turn back to our original query whether platypuses are typically warm-blooded ($p \sim w$) from Example 3.4. Again, we must consider all the best p worlds in $\mathcal{R}_{\mathcal{K}}^{LC}$ and check whether they are worlds in which w is true as well. We circle the only such world on Figure 4. In the case of Lexicographic Closure, we find that $\mathcal{R}_{\mathcal{K}}^{LC} \Vdash p \sim w$, which means that $p \sim w$ is in the Lexicographic Closure of \mathcal{K} , or $\mathcal{K} \models_{LC} p \sim w$, and can be interpreted as saying that platypuses are typically warm-blooded.

3.6.2 Bases. As mentioned briefly before, there is another characterisation of Lexicographic Closure by what Lehmann terms *bases* [14]. We choose not to define the specifics of such characterisation as our work surrounds the model-theoretic definition of Lexicographic Closure described above. We note that a detailed description of this characterisation, using notation consistent with this review, can be found in [12].

3.6.3 Lexicographic Closure Algorithm. Casini et al. [6] provide a general algorithm capable of computing the defeasible entailment relation generated by what they refer to as a \mathcal{K} -faithful rank function. This general algorithm is essentially a modification of earlier algorithm described for RC, where RationalClosure has been replaced with DefeasibleEntailment that makes a call to a new Rank function that replaces BaseRank.

The notion of being \mathcal{K} -faithful was first introduced to characterise *basic defeasible entailment* relations. It was subsequently shown that basic defeasible entailment was too permissive of a concept since it did not enforce that such an entailment relation endorse all the entailments of RC [6]. Hence, *rational defeasible entailment* relations were defined as basic defeasible entailment relations that *did* satisfy the property of extending RC - in line with the final thesis in [15]. It was also shown that rational defeasible entailment relations could be characterised semantically by requiring that \mathcal{K} -faithful ranked models be *rank preserving*: a \mathcal{K} -faithful ranked model \mathcal{R} is said to be rank preserving iff $\forall u, v \in \mathcal{U}$, if $\mathcal{R}_{\mathcal{K}}^{RC}(v) < \mathcal{R}_{\mathcal{K}}^{RC}(u)$, then $\mathcal{R}(v) < \mathcal{R}(u)$ [6]. This now formalises the idea that one can define *any* rational defeasible entailment relation in terms of a ranked model that is a refinement of $\mathcal{R}_{\mathcal{K}}^{RC}$ in the sense that certain levels of $\mathcal{R}_{\mathcal{K}}^{RC}$ may be split up into a number of sub-levels such that the relative positions of the interpretations in the original levels is maintained [6, 14].

Finally, to utilise the general algorithm proposed by Casini et al. to compute Lexicographic Closure, one needs to define a \mathcal{K} -faithful rank preserving rank function, $r_{\mathcal{K}}^{LC}$, that corresponds to

Lexicographic Closure. Casini et al. define it as follows: $r_{\mathcal{K}}^{LC}(\alpha) := \min\{\mathcal{R}_{\mathcal{K}}^{LC}(v) \mid v \in \text{Mod}(\alpha)\}$ [6].

We will not delve into the specifics of how this is done and the mechanics of how the algorithm answers a given query in this review as the complexities can be abstracted from our earlier demonstration of RC and found in [5, 6]. However, we will note that computing this algorithm is computationally harder than computing classical entailment since its implementation may result in a number of calls to a classical entailment checker that is exponential in the size of \mathcal{K} [6].

4 DISCUSSION

Our examples for Rational Closure and Lexicographic Closure contrasted their differences in terms of their prototypical and presumptive reasoning styles. We clarify that this does not inherently imply that LC is necessarily a more valid form of defeasible entailment than RC since it ‘correctly’ entailed that platypuses are warm-blooded mammals whereas RC did not. We mentioned earlier than there is likely no correct-most form of defeasible entailment, but merely ones that are potentially better suited to certain domains than others. Ultimately, choosing certain forms of defeasible entailment begins to encroach upon the area of Ontology Engineering.

5 CONCLUSIONS

We started this review with a demonstration of how one can use formal logic to represent and reason about knowledge. Classical logics and their associated reasoning services, burdened by the property of monotonicity, are shown to be inadequate for modelling real-world situations as they cannot express any degree of uncertainty or typicality. On the other hand, reasoning with uncertainty and the ability to retract prior conclusions, whilst in theory is better suited, is not unique. Hence, we briefly motivated for and chronicled the development of the prominent KLM framework used to formally assess sensible forms of defeasible entailment.

In particular, we examined two principle defeasible entailment relations, Rational Closure and Lexicographic Closure, that fall under the umbrella of *rational defeasible entailment* relations in an extension of the KLM framework [6] proposed by Casini et al. It was shown how both could be characterised using ranked models as well as ranked formulas - with an emphasis on the former.

A key result for our work is that any form of rational defeasible entailment can be characterised by a ranked model that is a refinement of the ranked model corresponding to Rational Closure. This may suggest that a model-based algorithm devised for Rational Closure may be generalisable to other forms of rational defeasible entailment.

Whilst focusing on their model-based semantics, we noted that the current algorithm for computing Rational Closure is equivalent in computational complexity to classical entailment checking, however, the general algorithm available for computing rational defeasible entailment is less efficient and computationally harder. Furthermore, it is evident that there presently does not appear to exist an efficient way of constructing the minimal ranked model corresponding to Rational Closure or the ranked model corresponding to Lexicographic Closure other than by some ‘brute-force’ means of applying their semantic definitions.

These gaps in the literature point to the need for further investigation into the design of novel algorithms that construct and use ranked models for defeasible entailment checking along with the feasibility of their implementations compared to current formula-based ‘top-down’ approaches.

REFERENCES

- [1] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider (Eds.). 2003. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, USA.
- [2] Mordechai Ben-Ari. 2012. *Propositional Logic: Formulas, Models, Tableaux*. Springer London, London, 1, 7–47.
- [3] Richard Booth, Thomas Meyer, and Ivan Varzinczak. 2012. PTL: A Propositional Typicality Logic. In *Logics in Artificial Intelligence*, Luis Fariñas del Cerro, Andreas Herzig, and Jérôme Mengin (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 107–119.
- [4] Giovanni Casini, Thomas Meyer, Kodylan Moodley, and Riku Nortjé. 2014. Relevant Closure: A New Form of Defeasible Reasoning for Description Logics. In *Proceedings of the 14th European Conference on Logics in Artificial Intelligence - Volume 8761*. Springer-Verlag, Berlin, Heidelberg, 92–106. https://doi.org/10.1007/978-3-319-11558-0_7
- [5] Giovanni Casini, Thomas Meyer, and Ivan Varzinczak. 2018. Defeasible Entailment: from Rational Closure to Lexicographic Closure and Beyond. In *Proceedings of 17th INTERNATIONAL WORKSHOP ON NON-MONOTONIC REASONING*. 109–118.
- [6] Giovanni Casini, Thomas Meyer, and Ivan Varzinczak. 2019. Taking Defeasible Entailment Beyond Rational Closure. In *Logics in Artificial Intelligence*, Francesco Calimeri, Nicola Leone, and Marco Manna (Eds.). Springer International Publishing, Cham, 182–197.
- [7] Kenneth James Williams Craik. 1952. *The nature of explanation*. Vol. 445. CUP Archive.
- [8] Michael Freund. 1998. Preferential reasoning in the perspective of Poole default logic. *Artificial Intelligence* 98, 1 (1998), 209–235. [https://doi.org/10.1016/S0004-3702\(97\)00053-2](https://doi.org/10.1016/S0004-3702(97)00053-2)
- [9] D. M. Gabbay. 1985. Theoretical Foundations for Non-Monotonic Reasoning in Expert Systems. In *Logics and Models of Concurrent Systems*, Krzysztof R. Apt (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 439–457.
- [10] L. Giordano, V. Gliozzi, N. Olivetti, and G.L. Pozzato. 2015. Semantic characterization of rational closure: From propositional logic to description logics. *Artificial Intelligence* 226 (2015), 1–33. <https://doi.org/10.1016/j.artint.2015.05.001>
- [11] Crina Grosan and Ajith Abraham. 2011. *Knowledge Representation and Reasoning*. Springer Berlin Heidelberg, Berlin, Heidelberg, 131–147. https://doi.org/10.1007/978-3-642-21004-4_6
- [12] Adam Kaliski. 2020. *An Overview of KLM-Style Defeasible Entailment*. Master’s thesis. Faculty of Science, University of Cape Town, Rondebosch, Cape Town, 7700.
- [13] Sarit Kraus, Daniel Lehmann, and Menachem Magidor. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44, 1-7 (1990). [https://doi.org/10.1016/0004-3702\(90\)90101-5](https://doi.org/10.1016/0004-3702(90)90101-5)
- [14] Daniel Lehmann. 1999. Another perspective on Default Reasoning. *Annals of Mathematics and Artificial Intelligence* 15 (11 1999). <https://doi.org/10.1007/BF01535841>
- [15] Daniel Lehmann and Menachem Magidor. 1992. What does a conditional knowledge base entail? *Artificial Intelligence* 55, 1 (1992), 1–60. [https://doi.org/10.1016/0004-3702\(92\)90041-U](https://doi.org/10.1016/0004-3702(92)90041-U)
- [16] J. Donald Monk. 1986. The Contributions of Alfred Tarski to Algebraic Logic. *The Journal of Symbolic Logic* 51, 4 (1986), 899–906. <http://www.jstor.org/stable/2273903>
- [17] R. Reiter. 1980. A logic for default reasoning. *Artificial Intelligence* 13, 1 (1980), 81–132. [https://doi.org/10.1016/0004-3702\(80\)90014-4](https://doi.org/10.1016/0004-3702(80)90014-4) Special Issue on Non-Monotonic Logic.
- [18] Yoav Shoham. 1987. Nonmonotonic Logics: Meaning and Utility. In *IJCAI*.
- [19] Y. Shoham. 1987. A Semantical Approach to Nonmonotonic Logics. In *LICS 1987*.
- [20] Yoav Shoham. 1988. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*. MIT Press, Cambridge, MA, USA.