KLM-Style Defeasibility for Restricted First-Order Logic

Giovanni Casini[1,2,3][0000−0002−4267−4447], Thomas Meyer[1,2,3][0000−0003−2204−6969], Guy Paterson-Jones[2,3], and Ivan Varzinczak[4,3,5,1][0000−0002−0025−9632]

[1] ISTI–CNR, Italy
[2] University of Cape Town, South Africa
[3] CAIR, South Africa
[4] LIASD, Université Paris 8, France
[5] Stellenbosch University, South Africa

Abstract. In this paper, we extend the KLM approach to defeasible reasoning beyond the propositional setting. We do so by making it applicable to a restricted version of first-order logic. We describe defeasibility for this logic using a set of rationality postulates, provide a suitable and intuitive semantics for it, and present a representation result characterising the semantic description of defeasibility in terms of our postulates. An advantage of our semantics is that it is sufficiently general to be applicable to other restricted versions of first-order logic as well. Based on this theoretical core, we then propose a version of defeasible entailment that is inspired by the well-known notion of Rational Closure as it is defined for defeasible propositional logic and defeasible description logics. We show that this form of defeasible entailment is rational in the sense that it adheres to the full set of rationality postulates.

Keywords: Defeasible reasoning · First-order logic · Rationality

1 Introduction

The past 15 years have seen a flurry of activity to introduce defeasible-reasoning capabilities into languages that are more expressive than that of propositional logic [5,6,9,16,17,27]. Most of the focus has been on defeasibility for description logics (DLs), with much of it devoted to versions of the so-called KLM approach to defeasible reasoning initially advocated for propositional logic by Kraus et al. [22]. In DLs, knowledge is expressed as class inclusions of the form $C \sqsubseteq D$, with the intended meaning that every instance of C is also an instance of D. Defeasible DLs allow, in addition, for defeasible inclusions of the form $C \mathrel{\sqsubseteq\!\!\!\!\sim} D$ with the intended meaning that instances of C are *usually* instances of D. For example, $\mathsf{Student} \mathrel{\sqsubseteq\!\!\!\!\sim} \neg\exists \mathsf{pays}.\mathsf{Tax}$ (students usually don't pay tax) is a defeasible version of $\mathsf{Student} \sqsubseteq \neg\exists \mathsf{pays}.\mathsf{Tax}$ (students don't pay tax).

In this paper, we focus instead on a restricted version of first-order logic (RFOL), for which a semantics in terms of Herbrand interpretations suffices.

We provide the theoretical foundations for an extension of RFOL modelling defeasible reasoning (DRFOL). However, the availability of non-unary predicates means that the definition of an appropriate semantics for DRFOL is a non-trivial exercise. This is because the intuition underlying KLM-style defeasibility generally depends on the underlying language. For propositional logics the intuition dictates a notion of typicality over *possible worlds*. The statement "birds usually fly", formalised as $\mathsf{bird} \mid\!\sim \mathsf{fly}$, says that in the most typical worlds in which bird is true, fly is also true. In contrast, defeasibility in DLs invokes a form of typicality over *individuals*. Thus $\mathsf{Student} \mathrel{\sqsubseteq\!\!\!\!\sim} \neg\exists \mathsf{pays}.\mathsf{Tax}$ states that of all those individuals that are students, the most typical ones don't pay taxes. To see the problem in extending either of these intuitions to the case with non-unary predicates, consider the following version of an example by Delgrande [13].

Example 1. The following DRFOL knowledge base states that humans don't feed wild animals, that elephants are usually wild animals, that keepers are usually human, and that keepers usually feed elephants, but that Fred the keeper usually does not feed elephants (the connective $\rightsquigarrow$ refers to defeasible implication and variables are implicitly quantified).

$$\mathcal{K} = \left\{ \begin{array}{c} \mathsf{wild\_animal}(\mathsf{x}) \wedge \mathsf{human}(\mathsf{y}) \rightarrow \neg\mathsf{feeds}(\mathsf{y},\mathsf{x}), \\ \mathsf{elephant}(\mathsf{x}) \rightsquigarrow \mathsf{wild\_animal}(\mathsf{x}), \\ \mathsf{keeper}(\mathsf{x}) \rightsquigarrow \mathsf{human}(\mathsf{x}), \\ \mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{y}) \rightsquigarrow \mathsf{feeds}(\mathsf{y},\mathsf{x}), \\ \mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{fred}) \rightsquigarrow \neg\mathsf{feeds}(\mathsf{fred},\mathsf{x}) \end{array} \right\}$$

For any appropriate semantics, $\mathcal{K}$ above should be satisfiable (given a suitable notion of satisfiability). Then it soon becomes clear that the propositional approach cannot achieve this. To see why, note that applying the propositional intuition to the example would result in $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{y}) \rightsquigarrow \mathsf{feeds}(\mathsf{y},\mathsf{x})$, meaning that in the most typical worlds (Herbrand interpretations in this case) all keepers feed all elephants. This is in conflict with $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{fred}) \rightsquigarrow \neg\mathsf{feeds}(\mathsf{fred},\mathsf{x})$, which states that in the most typical Herbrand interpretations, keeper Fred does not feed any elephants. For any reasonable definition of satisfiability, this would render the knowledge base unsatisfiable.

The DL-based intuition of object typicality is also problematic. Under this intuition, the statement $\mathsf{elephant}(\mathsf{x}) \rightsquigarrow \mathsf{wild\_animal}(\mathsf{x})$ would mean that the most typical elephants are wild animals. Similarly, $\mathsf{keeper}(\mathsf{x}) \rightsquigarrow \mathsf{human}(\mathsf{x})$ would mean that the most typical keepers are human. Combined with the first statement in $\mathcal{K}$, it would then follow that the most typical keepers (being humans) do not feed the most typical elephants (being wild animals). On the other hand, the fourth statement in $\mathcal{K}$ explicitly states that the most typical keepers feed the most typical elephants, from which we obtain the counter-intuitive conclusion that typical elephants and typical keepers cannot exist simultaneously. Some reflection on this example should be sufficient to indicate that it represents a genuine limitation of the standard propositional and DL approaches to defeasibility when applied to FOL.

In this paper, we resolve this matter with a semantics that is in line with the propositional intuition of a typicality ordering over worlds, but also includes aspects of the DL intuition of typicality of individuals. We achieve the latter by enriching our semantics with *typicality objects*, which are used to represent *typical* individuals. Thus, $\mathsf{elephant}(\mathsf{x}) \land \mathsf{keeper}(\mathsf{y}) \rightsquigarrow \mathsf{feeds}(\mathsf{y},\mathsf{x})$ means that in the most typical enriched Herbrand interpretations, all typical keepers feed all typical elephants, with the understanding that there may be exceptional keepers that don't feed some elephants. Note that the term *typical* is used here in two different, but related, ways.

Our central theoretical result is a representation result (Theorems 1 and 2), showing that defeasible implication defined in this way can be characterised w.r.t. a set of KLM-style rationality postulates adapted for DRFOL. Another important consequence of our representation result is that it provides the theoretical foundation for the definition of various forms of defeasible entailment for DRFOL. We present one such form of defeasible entailment and show that it can be viewed as the DRFOL analogue of Rational Closure as originally defined for the propositional case [24].

In the rest of the paper, we start by providing a brief introduction to RFOL and to KLM-style defeasible reasoning (Section 2). In Section 3, we introduce DRFOL, describe an abstract notion of satisfaction w.r.t. a set of KLM-style postulates, provide a suitable semantics, and prove a representation result, showing that the KLM-style postulates characterise the semantic construction. In Section 4, we present a form of defeasible entailment for DRFOL that can be viewed as the DRFOL equivalent of the well-known notion of Rational Closure. Before concluding the paper, we discuss related work in Section 5.

2 Background

The language of RFOL builds on three disjoint sets of symbols: a finite set of constants $\textsc{const}$, a countably infinite set of variable symbols $\textsc{var}$, and a finite set of predicate symbols $\textsc{pred}$. It has no function symbols. A *term* is an element of $\textsc{const} \cup \textsc{var}$. Each predicate symbol $\alpha \in \textsc{pred}$ has an *arity*, denoted $\mathfrak{ar}(\alpha) \in \mathbb{N}$, representing the number of terms it takes as arguments. We assume the existence of predicate symbols $\top$ and $\bot$, which have arity 0. An *atom* is an expression of the form $\alpha(t_1, \ldots, t_{\mathfrak{ar}(\alpha)})$, where $\alpha \in \textsc{pred}$ and each t_i is a term. Observe that $\top$ and $\bot$ are atoms as well.

A *compound* is a Boolean combination of atoms (i.e., built from atoms and the logical connectives $\neg$, $\land$, and $\lor$). An *implication* has the form $A(\vec{x}) \to B(\vec{y})$, where $A(\vec{x})$ and $B(\vec{y})$ are compounds, and where the terms occurring in $\vec{x}$ and $\vec{y}$ may overlap. A compound (resp. implication) is *ground* if all the terms contained in it are constants; otherwise it is *open*. Ground atoms are also known as *facts*.

The only formulas we permit are compounds and implications and these are understood to be implicitly universally quantified. We shall also adopt the following conventions. Constant symbols and variables are written in lowercase, with early letters used for constants $(a, b, \ldots)$ and later letters for variables

$(x, y, \dots)$. Compounds are denoted by uppercase letters $(A, B, \dots)$. Tuples of variables or constants are written with overbars, such as $\vec{x}$ and $\vec{a}$ resp., and $A(\vec{x})$ and $B(\vec{a})$ are used as shorthand for compounds over their respective tuples of terms. We use lowercase early Greek letter $(\alpha, \beta, \dots)$ to denote RFOL formulas, sometimes with tuples of terms $(\alpha(\vec{x}))$. The set of all formulas (compounds and implications) is denoted by $\mathcal{L}$. A *knowledge base* $\mathcal{K}$ is a finite subset of $\mathcal{L}$.

The Herbrand universe $\mathbb{U}$ is the set CONST. The *Herbrand base* of $\mathbb{U}$, denoted $\mathbb{B}$, is the set of facts defined over $\mathbb{U}$. A *Herbrand interpretation* is a subset $\mathcal{H} \subseteq \mathbb{B}$. The set of Herbrand interpretations is denoted by $\mathscr{H}$. *Substitutions* are defined to be functions $\varphi : \text{VAR} \to \text{VAR} \cup \text{CONST}$ assigning a term to each variable symbol. *Variable substitutions* are substitutions that assign only variables, and *ground substitutions* are substitutions that assign only constants. The application of a substitution φ to a compound $A(\vec{x})$ is denoted $A(\varphi(\vec{x}))$.

RFOL knowledge bases are interpreted by Herbrand interpretations $\mathcal{H}$ as follows: (1) if $A(\vec{a})$ is a ground atom, then $\mathcal{H} \Vdash A(\vec{a})$ iff $A(\vec{a}) \in \mathcal{H}$; (2) if $A(\vec{a})$ and $B(\vec{b})$ are ground compounds (where $\vec{a}$ and $\vec{b}$ may overlap), then $\mathcal{H} \Vdash A(\vec{a})$ and $\mathcal{H} \Vdash A(\vec{a}) \to B(\vec{b})$ as usual for Boolean connectives; (3) if $A(\vec{x})$ is an open compound, then $\mathcal{H} \Vdash A(\vec{x})$ iff $\mathcal{H} \Vdash A(\varphi(\vec{x}))$ for every ground substitution φ; (4) if $A(\vec{x}) \to B(\vec{y})$ is an open implication (where $\vec{x}$ and $\vec{y}$ may overlap), then $\mathcal{H} \Vdash A(\vec{x}) \to B(\vec{y})$ iff $\mathcal{H} \Vdash A(\varphi(\vec{x})) \to B(\varphi(\vec{y}))$ for every ground substitution φ, and (5) if $\mathcal{K}$ is a knowledge base, then $\mathcal{H} \Vdash \mathcal{K}$ iff $\mathcal{H} \Vdash \alpha$ for every $\alpha \in \mathcal{K}$. A Herbrand interpretation satisfying a knowledge base $\mathcal{K}$ is a *Herbrand model* of $\mathcal{K}$.

Kraus et al. [22] originally defined $\mid\sim$ as a consequence relation over a propositional language, with statements of the form $\alpha \mid\sim \beta$ to be interpreted as the meta-statement "β is a defeasible consequence of α". Subsequently, Lehmann and Magidor [24] made a subtle shift in considering an object-level language containing statements of the form $\alpha \mid\sim \beta$, to be interpreted as the object-level statement "α defeasibly implies β", and with $\mid\sim$ viewed as an object-level connective. This view is captured by a set of *rationality postulates*, which have been widely discussed in the literature. We do not repeat these rationality postulates here, but note that Definition 3, our definition of rationality for DRFOL, the defeasible version of RFOL, relies heavily on versions of the KLM rationality postulates that are lifted to DRFOL (see Section 3).

A semantics for defeasible implications is provided by *ranked interpretations* $\mathscr{R}$, with $\mathscr{R}$ a function from U (the set of all valuations) to $\mathbb{N} \cup \{\infty\}$, satisfying the following *convexity property*: for every $i \in \mathbb{N}$, if $\mathscr{R}(u) = i$, then, for every $j < i$, there is a $u' \in U$ for which $\mathscr{R}(u') = j$. $\mathscr{R}(v)$ indicates the degree of *atypicality* of v. The valuations judged most typical are those with rank 0, while those with infinite rank are judged so atypical as to be impossible. A defeasible statement $\alpha \mid\sim \beta$ is *satisfied in* $\mathscr{R}$ $(\mathscr{R} \Vdash \alpha \mid\sim \beta)$ if the models of α with the smallest *finite* rank in $\mathscr{R}$ are all models of β. A classical statement α is satisfied in $\mathscr{R}$ $(\mathscr{R} \Vdash \alpha)$ if every valuation of finite rank satisfies α.

Note that $\mathscr{R} \Vdash \neg\alpha \mid\sim \bot$ iff all the models of $\neg\alpha$ have infinite rank, which is equivalent by definition to $\mathscr{R} \Vdash \alpha$.

3 Defeasible restricted first-order logic

Defeasible Restricted First-Order Logic (DRFOL) extends the logic RFOL presented above with *defeasible implications* of the form $A(\vec{x}) \leadsto B(\vec{y})$, where $A(\vec{x})$ and $B(\vec{y})$ are compounds, and where $\vec{x}$ and $\vec{y}$ may overlap. The set of defeasible implications is denoted $\mathcal{L}^{\leadsto}$, and a *DRFOL knowledge base* $\mathcal{K}$ is defined to be a subset of $\mathcal{L} \cup \mathcal{L}^{\leadsto}$. Note that DRFOL knowledge bases may include (classical) RFOL formulas.

As demonstrated in Example 1, defeasible implications are intended to model properties that *typically* hold, but which may have exceptions. In this example, for instance, $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \leadsto \neg\mathsf{feeds(fred,x)}$, is an exception to $\mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \leadsto \mathsf{feeds(y,x)}$. A DRFOL knowledge base containing these statements ought to be satisfiable (for an appropriate notion of satisfaction). The same goes for the DRFOL knowledge base $\{\mathsf{bird(x)} \leadsto \mathsf{fly(x)}, \mathsf{bird(tweety)}, \neg\mathsf{fly(tweety)}\}$. To formalise these intuitions we first describe the intended behaviour of the defeasible connective $\leadsto$ and its interaction with (classical) RFOL formulas in terms of a set of rationality postulates in the KLM style [22,24]. These postulates are expressed via an abstract notion of satisfaction:

Definition 1. *A* satisfaction set *is a subset* $\mathcal{S} \subseteq \mathcal{L} \cup \mathcal{L}^{\leadsto}$.

We denote the classical part of a satisfaction set by $\mathcal{S}_C = \mathcal{S} \cap \mathcal{L}$. The first postulate we consider ensures $\mathcal{S}$ respects the classical notion of satisfaction when restricted to classical formulas, where $\models$ refers to classical entailment:

$$(\textsc{Cla}) \quad \frac{\mathcal{S}_C \models \alpha}{\alpha \in \mathcal{S}}$$

Next, we consider the interaction between classical and defeasible implications:

$$(\textsc{Sup}) \quad \frac{A(\vec{x}) \in \mathcal{S}}{\neg A(\vec{x}) \leadsto \bot \in \mathcal{S}}$$

We now consider the core of the proposal for defining rational satisfaction sets, the KLM rationality postulates, lifted to DRFOL, and expressed in terms of satisfaction sets:

$$(\textsc{Refl}) \quad A(\vec{x}) \leadsto A(\vec{x}) \in \mathcal{S}$$

$$(\textsc{Rw}) \quad \frac{A(\vec{x}) \leadsto B(\vec{y}) \in \mathcal{S},\ \models B(\vec{y}) \to C(\vec{z})}{A(\vec{x}) \leadsto C(\vec{z}) \in \mathcal{S}}$$

$$(\textsc{Lle}) \quad \frac{A(\vec{x}) \leadsto C(\vec{z}) \in \mathcal{S},\ \models A(\vec{x}) \to B(\vec{y}),\ \models B(\vec{y}) \to A(\vec{x})}{B(\vec{y}) \leadsto C(\vec{z}) \in \mathcal{S}}$$

$$(\textsc{And}) \quad \frac{A(\vec{x}) \leadsto B(\vec{y}) \in \mathcal{S},\ A(\vec{x}) \leadsto C(\vec{z}) \in \mathcal{S}}{A(\vec{x}) \leadsto B(\vec{y}) \wedge C(\vec{z}) \in \mathcal{S}}$$

$$(\textsc{Or}) \quad \frac{A(\vec{x}) \leadsto C(\vec{z}) \in \mathcal{S},\ B(\vec{y}) \leadsto C(\vec{z}) \in \mathcal{S}}{A(\vec{x}) \vee B(\vec{y}) \leadsto C(\vec{z}) \in \mathcal{S}}$$

$$(\textsc{Rm}) \quad \frac{A(\vec{x}) \leadsto \neg B(\vec{y}) \notin \mathcal{S},\ A(\vec{x}) \wedge B(\vec{y}) \leadsto C(\vec{z}) \notin \mathcal{S}}{A(\vec{x}) \leadsto C(\vec{z}) \notin \mathcal{S}}$$

Next we consider *instantiations* of implications (applicable to all substitutions of the right type):

$$(\textsc{Dui})\quad \frac{A(\vec{x}) \rightsquigarrow B(\vec{y}) \in \mathcal{S}}{A(\varphi(\vec{x})) \rightsquigarrow B(\varphi(\vec{y})) \in \mathcal{S}}$$

To begin with, note that universal instantiation is *not* a desirable property for defeasible implications. To see why, consider a satisfaction set $\mathcal{S}$ containing $\mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \rightsquigarrow \mathsf{feeds(y,x)}$ and $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \neg\mathsf{feeds(fred,x)}$. From (Dui) we have $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \mathsf{feeds(fred,x)} \in \mathcal{S}$, and hence by (And) and (Rw) that $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \bot \in \mathcal{S}$ as well, which is in conflict with the intuition that exceptional cases (all elephants usually not being fed by keeper Fred) should be permitted to exist alongside the general case (all elephants usually being fed by all keepers).

Weaker forms of instantiation for defeasible implications are more reasonable. Consider $\mathsf{keeper(x)} \rightsquigarrow \mathsf{feeds(x,y)}$, which states that keepers typically feed everything. While we cannot conclude anything about instances of x, for the reasons discussed above, we should at least be able to conclude things about instances of y, since y only appears in the consequent of the implication. This motivates the following postulate (again, applicable to all substitutions of the right type), where ψ is a variable substitution and $\vec{x} \cap \vec{y} = \emptyset$:

$$(\textsc{Irr})\quad \frac{A(\vec{x}) \rightsquigarrow B(\vec{x},\vec{y}) \in \mathcal{S}}{A(\vec{x}) \rightsquigarrow B(\vec{x},\varphi(\vec{y})) \in \mathcal{S}}$$

There are some more subtle forms of defeasible instantiation that seem reasonable as well. Consider the following relation defined over $\mathcal{L}$:

Definition 2. *$A(\vec{x})$ is* at least as typical *as $B(\vec{y})$ w.r.t. $\mathcal{S}$, denoted $A(\vec{x}) \preccurlyeq_{\mathcal{S}} B(\vec{y})$, iff $A(\vec{x}) \vee B(\vec{y}) \rightsquigarrow \neg A(\vec{x}) \notin \mathcal{S}$.*

Intuitively, $A(\vec{x}) \preccurlyeq_{\mathcal{S}} B(\vec{y})$ states that typical instances of $A(\vec{x})$ are at least as typical as typical instances of $B(\vec{y})$. Note that for any variable substitution φ, a typical instance of $A(\varphi(\vec{x}))$ is always an instance of $A(\vec{x})$. Thus the following postulate should hold, where φ is any variable substitution:

$$(\textsc{Typ})\quad A(\vec{x}) \preccurlyeq_{\mathcal{S}} A(\varphi(\vec{x}))$$

The last postulate we consider has to do with defeasibly impossible formulas. Suppose $A(\varphi(\vec{x})) \rightsquigarrow \bot \in \mathcal{S}$ for all substitutions $\varphi : \textsc{var} \to \textsc{var} \cup \mathbb{U}$. This states that if *all* specialisations of $A(\vec{x})$ are defeasibly impossible, then we should expect that there are in fact no instances of $A(\vec{x})$ at all:

$$(\textsc{Imp})\quad \frac{A(\varphi(\vec{x})) \rightsquigarrow \bot \in \mathcal{S} \text{ for all } \varphi : \textsc{var} \to \textsc{var} \cup \mathbb{U}}{\neg A(\vec{x}) \in \mathcal{S}}$$

This puts us in a position to define the central construction of the paper, namely that of a *rational* satisfaction set.

Definition 3. *$\mathcal{S}$ is* rational *iff it satisfies* (Cla)*,* (Sup)*,* (Irr)*,* (Typ)*,* (Imp) *and* (Refl)-(Rm)*.*

Rational satisfaction sets satisfy the following form of label invariance for defeasible implications, where the variable substitution φ is a *permutation*:

$$(\textsc{Per}) \quad \frac{A(\vec{x}) \rightsquigarrow B(\vec{y}) \in \mathcal{S}}{A(\varphi(\vec{x})) \rightsquigarrow B(\varphi(\vec{y})) \in \mathcal{S}}$$

Proposition 1. *Let $\mathcal{S}$ be a rational satisfaction set. Then $\mathcal{S}$ satisfies* (Per)*.*

We define a semantics for defeasible implications by enriching the Herbrand universe with a set $\mathcal{T}$ of *typicality objects*. Typicality objects represent individuals that are not explicitly mentioned in a given knowledge base, and are used here to interpret defeasible implications in a ranking of (enriched) Herbrand interpretations.

Definition 4. *Given a set of typicality objects $\mathcal{T}$, the corresponding* enriched Herbrand universe *is defined to be the set $\mathbb{U}_{\mathcal{T}} = \mathbb{U} \cup \mathcal{T}$. For each possible partition of $\mathbb{U}$ into two sets $\mathbb{U}_t$ and $\mathbb{U}_e$ (both possibly empty), we have a* typicality set *$Typ = \mathbb{U}_t \cup \mathcal{T}$. An* enriched Herbrand interpretation *(or EHI) $\mathcal{E}$ is a Herbrand interpretation defined over an enriched Herbrand universe $\mathbb{U}_{\mathcal{T}}$, and associated with $Typ_{\mathcal{E}}$, one of the possible typicality sets in $\mathbb{U}_{\mathcal{T}}$.*

Using the typicality sets in enriched Herbrand interpretations we distinguish between typical and atypical objects. That is, we assume that, given an interpretation $\mathcal{E}$, all the objects in $Typ_{\mathcal{E}}$ are typical objects, while the set $\mathbb{U}_e = \mathbb{U}_{\mathcal{T}} \setminus Typ_{\mathcal{E}}$ represents the exceptional ones.

Every EHI $\mathcal{E}$ restricts to a unique Herbrand interpretation $\mathcal{H}^{\mathcal{E}}$ over $\mathbb{U}$, defined by $\mathcal{H}^{\mathcal{E}} = \mathcal{E} \cap \mathbb{B}$. The set of EHIs over $\mathcal{T}$ is denoted by $\mathscr{H}_{\mathcal{T}}$. To interpret defeasible implications we make use of preference rankings over $\mathscr{H}_{\mathcal{T}}$.

Definition 5. *A* ranked interpretation *is a function $rk : \mathscr{H}_{\mathcal{T}} \to \Omega \cup \{\infty\}$, for some linear poset Ω, satisfying the following properties, where we define $\mathscr{H}^{rk}_{\mathcal{T}} = \{\mathcal{E} \in \mathscr{H}_{\mathcal{T}} : rk(\mathcal{E}) \neq \infty\}$ to be the set of possible EHIs w.r.t. rk, and $\mathscr{H}^{rk}_{\mathcal{T}}(A(\vec{x})) = \{\mathcal{E} \in \mathscr{H}^{rk}_{\mathcal{T}} : \mathcal{E} \Vdash A(\varphi(\vec{x}))$ for some $\varphi : \textsc{var} \to Typ_{\mathcal{E}}\}$ to be the set of possible EHIs w.r.t. rk satisfying some typical instance of $A(\vec{x}) \in \mathcal{L}$:*

1. *if $rk(\mathcal{E}) = x < \infty$, then for every $y \leq x$ there is some $\mathcal{E}' \in \mathscr{H}_{\mathcal{T}}$ such that $rk(\mathcal{E}') = y$;*
2. *for all $A(\vec{x}) \in \mathcal{L}$, $\mathscr{H}^{rk}_{\mathcal{T}}(A(\vec{x}))$ is either empty or has an element that is an rk-minimal model of $A(\vec{x})$. This is* smoothness *[22].*

The set of ranked interpretations over $\mathcal{T}$ is denoted $\mathcal{R}_{\mathcal{T}}$.

Definition 6. *Let rk be a ranked interpretation. For all $A(\vec{x}), B(\vec{y}) \in \mathcal{L}$:*

1. *$rk \Vdash A(\vec{x})$ iff $\mathcal{E} \Vdash A(\vec{x})$ for all $\mathcal{E} \in \mathscr{H}^{rk}_{\mathcal{T}}$;*
2. *$rk \Vdash A(\vec{x}) \to B(\vec{y})$ iff $\mathcal{E} \Vdash A(\vec{x}) \to B(\vec{y})$ for all $\mathcal{E} \in \mathscr{H}^{rk}_{\mathcal{T}}$;*
3. *$rk \Vdash A(\vec{x}) \rightsquigarrow B(\vec{y})$ iff $\mathcal{E} \Vdash A(\varphi(\vec{x})) \to B(\varphi(\vec{y}))$ for all $\mathcal{E} \in \min_{rk} \mathscr{H}^{rk}_{\mathcal{T}}(A(\vec{x}))$ and all $\varphi : \textsc{var} \to Typ_{\mathcal{E}}$.*

Thus, compounds and classical implications are true in a ranked interpretation *rk* if they are true in all possible EHIs w.r.t. *rk*, while a defeasible implication is true in *rk* if its classical counterparts, with variables substituted by typicality objects, are true in all minimal EHIs (possible w.r.t. *rk*) in which the antecedent of the defeasible implication is true. A ranked interpretation in which a statement is true is a *ranked model* of the statement.

Example 2. This is a (slightly modified) example proposed by Delgrande [13]. Let $\textsc{const} = \{\mathsf{clyde}, \mathsf{fred}\}$, $\textsc{var} = \{\mathsf{x}, \mathsf{y}\}$, and $\textsc{pred} = \{\mathsf{elephant}, \mathsf{keeper}, \mathsf{likes}\}$. The following DRFOL knowledge base states that elephants and keepers are disjoint, that elephants usually like keepers, that elephants usually *don't* like keeper Fred, and that elephant Clyde usually *does* like Fred:

$$\mathcal{K} = \left\{ \begin{array}{c} \mathsf{elephant(x)} \rightarrow \neg\mathsf{keeper(x)}, \\ \mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \rightsquigarrow \mathsf{likes(x,y)}, \\ \mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \neg\mathsf{likes(x,fred)}, \\ \mathsf{elephant(clyde)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \mathsf{likes(clyde,fred)} \end{array} \right\}$$

Let $\mathcal{T} = \{\mathsf{t}_1, \ldots\}$ be the set of typicality objects. For readability we abbreviate $\mathsf{elephant}$ with e, keeper with k and likes with l.

Consider the EHIs $\mathcal{E}_1 = \{\mathsf{e(t_1)}, \mathsf{k(t_2)}, \mathsf{l(t_1,t_2)}, \mathsf{e(t_2)}, \mathsf{e(clyde)}, \mathsf{k(fred)}, \mathsf{l(clyde,fred)}\}$, $\mathcal{E}_2 = \{\mathsf{e(t_1)}, \mathsf{k(t_2)}, \mathsf{l(t_1,t_2)}, \mathsf{k(t_3)}, \mathsf{l(t_1,t_3)}, \mathsf{e(clyde)}, \mathsf{k(fred)}, \mathsf{l(clyde,fred)}\}$, and $\mathcal{E}_3 = \{\mathsf{e(t_1)}, \mathsf{k(t_2)}, \mathsf{e(t_2)}, \mathsf{e(clyde)}, \mathsf{k(fred)}, \mathsf{l(clyde,fred)}\}$. In all these EHIs let $\mathbb{U}_t = \emptyset$ and consequently $\mathit{Typ} = \mathcal{T}$. That is, in each of them the defeasible implications are evaluated only w.r.t. the typicality objects. Let $\mathit{rk}_1(\mathcal{E}_1) = \mathit{rk}_1(\mathcal{E}_2) = 0$, $\mathit{rk}_1(\mathcal{E}_3) = 1$, and $\mathit{rk}_1(\mathcal{E}) = \infty$ for all other EHIs. Then rk_1 is a ranked model of the knowledge base above. Let $\mathit{rk}_2(\mathcal{E}_1) = \mathit{rk}_2(\mathcal{E}_3) = 0$, $\mathit{rk}_2(\mathcal{E}_2) = 1$, and $\mathit{rk}_2(\mathcal{E}) = \infty$ for all other EHIs. Then rk_2 is not a ranked model of $\mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \rightsquigarrow \mathsf{likes(x,y)}$, but is a ranked model of $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \neg\mathsf{likes(x,fred)}$ and $\mathsf{elephant(clyde)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \mathsf{likes(clyde,fred)}$.

The main important technical result of the paper is a representation result, comprising a *soundness* result (Theorem 1) and a *completeness* result (Theorem 2), showing that ranked interpretations precisely characterise rational satisfaction sets:

Definition 7. *The satisfaction set $\mathcal{S}^{rk}$ corresponding to a ranked interpretation rk is: $\mathcal{S}^{rk} = \{\alpha \in \mathcal{L} \cup \mathcal{L}^{\rightsquigarrow} : rk \Vdash \alpha\}$.*

First we show that all ranked interpretations generate rational satisfaction sets as defined above:

Theorem 1. *For every ranked interpretation rk, $\mathcal{S}^{rk}$ is a rational satisfaction set.*

Then we show every rational set $\mathcal{S}$ can be realised as the satisfaction set corresponding to some ranked interpretation:

Theorem 2. *For every rational satisfaction set $\mathcal{S}$ there exists a ranked interpretation rk, over an infinite set of $\mathcal{T}$ of typicality objects, such that $\mathcal{S} = \mathcal{S}^{rk}$.*

4 Defeasible entailment

A central question that we have postponed until now is *entailment*. That is, given a DRFOL knowledge base $\mathcal{K}$, when are we justified in asserting that a DRFOL formula α follows defeasibly from $\mathcal{K}$? In this section, we provide one answer to this question by defining a semantic version of *Rational Closure* [24] for DRFOL. It is, by now, well-established that systems for defeasible reasoning are amenable to multiple forms of entailment, and the work we present in this section should therefore be viewed as the first step in a larger investigation into defeasible entailment.

In this section we consider the question of defeasible entailment for DRFOL and define a semantic version of *Rational Closure* [24] for DRFOL. Due to the so-called *drowning effect* [4], it is considered inferentially too weak for some application domains. Despite that, it is a semantic construction that can be extended to obtain other interesting entailment relations [23,12,10,15]. It has gained attention in the framework of DLs [11,9,17,6]. An equivalent semantic construction, System Z [26], has been considered for unary first-order logic [20,2,3]. Several equivalent definitions of Rational Closure can be found in the literature. Here we refer to the approach due to Booth and Paris [7] and Giordano et al. [17].

Let a knowledge base $\mathcal{K}$ be a set of propositional defeasible implications $\alpha \mid\sim \beta$. Booth and Paris provide a construction with the following two immediate consequences: (i) Given all the ranked models of $\mathcal{K}$, there is a model $\mathscr{R}^*$ of $\mathcal{K}$, that we can call the *minimal* one, which assigns to every propositional valuation v the *minimal* rank assigned to it by any of the ranked models of $\mathcal{K}$. (ii) Propositional Rational Closure can be characterised using $\mathscr{R}^*$. That is, $\alpha \mid\sim \beta$ is in the (propositional) Rational Closure of $\mathcal{K}$ iff $\mathscr{R}^* \Vdash \alpha \mid\sim \beta$. The intuition behind the use of the ranked model $\mathscr{R}^*$ for the definition of entailment is that it formalises the *presumption of typicality* [23]: assigning to each valuation the lowest possible rank, we model a reasoning pattern in which we assume that we are in one of the most typical situations that are compatible with our knowledge base.

We can define an analogous construction for DRFOL, but to do so we first need to address a technical restriction regarding typicality objects. More specifically, Theorem 2 requires an infinite set of typicality objects to be true in general. The next result shows that ranked interpretations can be restricted to finite sets of typicality objects, which is exactly what we need for our definition of defeasible entailment.

Proposition 2. *Let $\mathcal{K} \subseteq \mathcal{L} \cup \mathcal{L}^{\rightsquigarrow}$. Then $\mathcal{K}$ has a unique minimal ranked model iff it has a unique minimal ranked model over a* finite *set $\mathcal{T}'$ of typicality objects, with the size of $\mathcal{T}'$ referred to as the* order *of $\mathcal{K}$.*

The order of $\mathcal{K}$ depends on the number of formulas in $\mathcal{K}$ and the number of quantifier-bound variables in the formula, and is easy to calculate. The minimal ranked interpretation is defined in two stages, combining the two minimisation approaches used in propositional logic and DLs, respectively: first the rank $rk^*_{\mathcal{K}}$, a minimisation with respect to the rank of the EHIs, in line with the propositional approach [7,17]; then we refine it into the rank $rk_{\mathcal{K}}$, based on the minimisation of the position of the constants inside the EHIs, in line with the DL approach [17,9].

Definition 8. *Let $\mathcal{K} \subseteq \mathcal{L} \cup \mathcal{L}^{\leadsto}$ be of order n, and take $\mathcal{T}' \subset \mathcal{T}$ to be a finite set of typicality objects of cardinality n. The rank $rk^*_{\mathcal{K}} : \mathscr{H}_{\mathcal{T}'} \to \mathbb{N} \cup \{\infty\}$ is defined as follows:*

$$rk^*_{\mathcal{K}}(\mathcal{E}) = \min\{rk(\mathcal{E}) : rk \in \mathcal{R}_{\mathcal{T}'} \text{ and } rk \Vdash \mathcal{K}\}.$$

The minimal ranked model *of $\mathcal{K}$, which we denote by $rk_{\mathcal{K}} : \mathscr{H}_{\mathcal{T}'} \to (\mathbb{N} \times \mathbb{N}) \cup \{\infty\}$, is defined as:*

- *$rk_{\mathcal{K}}(\mathcal{E}) = \infty$, if $rk^*_{\mathcal{K}}(\mathcal{E}) = \infty$;*
- *$rk_{\mathcal{K}}(\mathcal{E}) = (i, j)$, if:*
  a) $rk^*_{\mathcal{K}}(\mathcal{E}) = i$ ($i \in \mathbb{N}$); and
  b) for every $k \geq j$, there is no $\mathcal{E}'$ s.t. $Typ'_{\mathcal{E}} \supset Typ_{\mathcal{E}}$ and $rk_{\mathcal{K}}(\mathcal{E}') = (i, k)$; and
  c) for every $l < j$, there is some $\mathcal{E}'$ s.t. $Typ'_{\mathcal{E}} \supset Typ_{\mathcal{E}}$ and $rk_{\mathcal{K}}(\mathcal{E}') = (i, l)$.

The order is defined lexicographically: $(i, j) \leq (k, l)$ iff $i < j$, or $i = j$ and $j \leq l$.

Given a consistent $\mathcal{K}$ and fixed a finite set of typicality constants, $rk_{\mathcal{K}}$ exists and is unique.

Proposition 3. *Let $\mathcal{K}$ be a knowledge base with a ranked model rk. Then, for a fixed finite enriched Herbrand universe $\mathbb{U}_{\mathcal{T}}$, $\mathcal{K}$ has exactly one minimal ranked model $rk_{\mathcal{K}}$.*

Note that by convention $\min \emptyset = \infty$, and $rk_{\mathcal{K}}$ is a ranked interpretation over $\mathcal{T}'$, since the lexicographic order defined in Definition 8 can easily be translated into an order defined over $\mathbb{N} \cup \infty$ satisfying the constraints from Definition 5. Hence $rk_{\mathcal{K}} \in \mathcal{R}_{\mathcal{T}'}$. Intuitively, $rk_{\mathcal{K}}$ is the result of first "pushing" every EHI rank as low as possible amongst the models of $\mathcal{K}$, similar to how it's done in the propositional approach, and then giving priority to the EHIs that have a bigger set of objects considered typical. That is, a bigger set *Typ*, in line with the DL approach. This minimal ranked model can be used to define a defeasible entailment relation for DRFOL:

Definition 9. *Let $\mathcal{K} \subseteq \mathcal{L} \cup \mathcal{L}^{\leadsto}$ and $\alpha \in \mathcal{L} \cup \mathcal{L}^{\leadsto}$. Then α is in the* Rational Closure *of $\mathcal{K}$, denoted $\mathcal{K} \mathrel{|\!\!\approx}_{rc} \alpha$, iff $rk_{\mathcal{K}} \Vdash \alpha$.*

The idea is that we give preference to the EHIs in which the set of typical individuals is maximal. That is, we assume that as many objects as possible behave according to our expectations.

Example 3. Assume $\mathcal{K}$ as in Example 2. The order of $\mathcal{K}$ is 2, so we build our minimal model $rk_{\mathcal{K}}$ using the set of EHIs $\mathscr{H}_{\mathcal{T}'}$, where the set of typical constants is $\mathcal{T}' = \{t_1, t_2\}$. Each EHI $\mathcal{E}$ satisfying $\mathcal{K}$ will be assigned rank $rk^*_{\mathcal{K}'}(\mathcal{E}) = 0$. That is, all the EHIs in which, given two constants $a, b \in Typ_{\mathcal{E}}$, if a is an elephant and b is a keeper, a likes b but, if fred is a keeper, a does not like fred. Also, if fred is a keeper and clyde is an elephant, clyde likes fred. All the other EHIs will be assigned rank 1, apart those in which keepers and elephants are not disjoint, that will have rank ∞. For example, the EHI $\mathcal{E}_1$ from Example 2 would have rank 0, while $\mathcal{E}_3$ would have rank 1, since it does not satisfy the formula $\mathsf{elephant(x)} \land \mathsf{keeper(y)} \leadsto \mathsf{likes(x,y)}$ ($\mathcal{E}_2$ is not considered in $rk_{\mathcal{K}}$, since it uses the constant t_3).

Now extend $\mathcal{K}$ into $\mathcal{K}'$ by adding the facts $\mathsf{elephant(dustin)}$ and $\mathsf{keeper(george)}$. Also, add the unary predicate $\mathsf{purple(x)}$ to PRED . The order of $\mathcal{K}'$ is still 2, so we build our minimal model $rk_{\mathcal{K}'}$ using again the set of EHIs $\mathscr{H}_{\mathcal{T}'}$. Again, each EHI $\mathcal{E}$ satisfying $\mathcal{K}'$ will be assigned rank $rk^*_{\mathcal{K}'}(\mathcal{E}) = 0$, while only the EHIs in which elephants and keepers are not disjoint, and either dustin is not an elephant or george is not a keeper, will have rank ∞.

We need to refine $rk^*_{\mathcal{K}}$ into $rk_{\mathcal{K}}$ looking at the relative sizes of the sets Typ associated to each EHI. Among the EHIs $\mathcal{E}$ s.t. $rk^*_{\mathcal{K}}(\mathcal{E}) = 0$, the ones in which $Typ_{\mathcal{E}}$ is bigger are those in which $Typ_{\mathcal{E}} = \mathcal{T} \cup \mathbb{U}$. In order to satisfy $\mathcal{K}'$, in such EHIs it is necessary that fred is not a keeper. Such EHIs will have rank $(0,0)$ in $rk_{\mathcal{K}'}$. Since we have no information forcing the exceptionality of dustin and george, such minimal models must satisfy $\mathsf{likes(dustin, george)}$, and we obtain the intuitive conclusion that $\mathcal{K}' \models\!\!\approx_{rc} \top \leadsto \mathsf{likes(dustin, george)}$.

Being a ranked interpretation, the desirable form of monotonicity (RM) holds. For example, note that all EHIs $\mathcal{E}$ at rank $(0,0)$ in the minimal model $rk_{\mathcal{K}'}$ would either satisfy $\mathsf{purple}(a)$ or not for any $a \in Typ_{\mathcal{E}}$, since it is irrelevant w.r.t. the satisfaction of $\mathcal{K}'$. The outcome would be that, while satisfying $\mathsf{elephant(x)} \land \mathsf{keeper(fred)} \leadsto \neg\mathsf{likes(x, fred)}$ (which is in $\mathcal{K}'$), $rk_{\mathcal{K}'}$ would not satisfy $\mathsf{elephant(x)} \land \mathsf{keeper(fred)} \leadsto \neg\mathsf{purple}(x)$, while it would satisfy $\mathsf{elephant(x)} \land \mathsf{purple}(x) \land \mathsf{keeper(fred)} \leadsto \neg\mathsf{likes(x, fred)}$.

More generally, Rational Closure, in the propositional and DL cases, satisfies a number of attractive properties:

$$\text{(INCL)} \quad \alpha \in \mathcal{K} \text{ implies } \mathcal{K} \models\!\!\approx_{rc} \alpha$$

$$\text{(SMP)} \quad \mathcal{S} = \{\alpha : \mathcal{K} \models\!\!\approx_{rc} \alpha\} \text{ is rational}$$

It is straightforward that these properties carry over to our definition of $\models\!\!\approx_{rc}$.

Theorem 3. $\models\!\!\approx_{rc}$ *satisfies* (INCL) *and* (SMP).

It is worthwhile delving a bit deeper into each of these properties. The first one, (INCL), also known as Inclusion, simply requires that statements in $\mathcal{K}$ also be defeasibly entailed by $\mathcal{K}$. It is a meta-version of the (REFL) rationality postulate for propositional logic (described in Section 2) and for DRFOL (described in Section 3). While the property itself might seem self-evident, it is instructive to view it in concert with the definition of $rk_{\mathcal{K}}$. From this it follows that $rk_{\mathcal{K}}$, which essentially defines Rational Closure, is the ranked interpretation in which EHIs are assigned a ranking that is truly as low (i.e., as typical) as possible, subject to the constraint that $rk_{\mathcal{K}}$ is a model of $\mathcal{K}$. This aligns with the intuition of propositional Rational Closure which requires of valuations in a ranked interpretation to be as typical as possible.

(Smp) requires the set of statements corresponding to the Rational Closure of $\mathcal{K}$ to be rational (cf. Definition 3). By virtue of Theorem 2, this requires defeasible entailment to be characterised by a *single* ranked interpretation, whence the fact the property is also referred to as Single Model Property.

5 Related work

Defeasible reasoning is part of a broader research programme on conditional reasoning [1], most of which was developed for propositional logic. This paper falls in the class of approaches aimed at moving beyond propositional expressivity. Besides the many extensions of defeasible reasoning to DLs in the recent literature [5,9,17], there have also been proposals to extend this approach to FOL. Most of these define a preference order on the domain [28,8,14], in line with some of the aforementioned DL proposals, and present rationality postulates, but they do not provide characterisations in terms of rationality postulates. Others [13,21] are formally closer to our work in that they use preference orders over interpretations.

Delgrande [13] proposes a semantics closer to the intuitions behind *circumscription* [25], giving preference to interpretations minimising counter-examples to defeasible conditionals. On the other hand, Kern-Isberner and Thimm [21] propose a technical solution much closer to the work we present here. Like ours, their semantics is based on Herbrand interpretations. They define *ordinal conditional functions* over the set of Herbrand interpretations, obtaining a structure that is very close to our ranked interpretations. They identify some individuals as *representatives* of a conditional. This is done to formalise the same intuition (or, at least, an intuition that is very similar) that underlies our decision to introduce typicality objects. Apart from other formal differences (e.g. the expressivity of their language is slightly different), their work focuses on the definition of a notion of entailment based on a specific semantic construction carried over from the propositional framework known as *c-representations* of a conditional knowledge base [18,19]. In contrast, our focus in this paper is on getting the theoretical foundations of defeasible reasoning for restricted FOL in place. Thus, our work here is centred around a representation result that provides a characterisation of the semantics in terms of structural properties. And while we present some results on defeasible entailment, we have left a more in-depth study of this important topic as future work. Indeed, it is our conjecture that the foundations we have put in place in this paper will allow for the definition of more than one form of defeasible entailment. At the same time, a more in-depth comparison with the proposal of Kern-Isberner and Thimm remains to be done.

Kern-Isberner and Beierle [20] and Beierle et al. [2,3] use the same semantic approach of Kern-Isberner and Thimm [21] to develop an extension of Pearl's System Z [26] for first-order logic, but they restrict their attention to unary predicates. System Z is a form of entailment that is very close to the approach we introduce here.

Brafman [8] suggests preference orders over the domain should result in forms of reasoning quite different from the use of preference orders on interpretations, comparable to the difference between statistical and subjective readings of probabilities. We leave an investigation of the differences between these two modelling solutions as future work.

We conclude this section with some remarks on the differences between DRFOL and the defeasible DL $\mathcal{DALC}$ [9]. When $\mathcal{DALC}$ is stripped of existential and value restrictions and confined to TBox statements, and when DRFOL is restricted to unary predicates and open implications (defeasible and classical), every concept C in $\mathcal{DALC}$ can be mapped to a compound $C(x)$ in DRFOL, and vice versa. It is then possible to obtain a result that is analogous to the propositional case, with one exception: a defeasible implication of the form $C(x) \rightsquigarrow \bot$ has a meaning that is different than $C \precsim \bot$, its $\mathcal{DALC}$ counterpart.

This marks an important distinction between DRFOL and both the propositional KLM framework and $\mathcal{DALC}$, in which classical statements are equivalent to certain defeasible implications. In the propositional case, α is equivalent to $\neg\alpha \mid\!\sim \bot$ ($\mathscr{R} \Vdash \alpha$ iff $\mathscr{R} \Vdash \neg\alpha \mid\!\sim \bot$ for all $\mathscr{R}$) while, for $\mathcal{DALC}$, $C \sqsubseteq \bot$ is equivalent to $C \precsim \bot$. But in DRFOL, defeasible implications *cannot* inform us about compounds or classical implications. Formally, rational satisfaction sets do *not* necessarily satisfy the following postulate:

$$(\textsc{Sub}) \quad \frac{A(\vec{x}) \rightsquigarrow \bot \in \mathcal{S}}{A(\vec{x}) \rightarrow \bot \in \mathcal{S}}$$

Note nevertheless that for a ground compound α (including those containing 0-ary predicates) it is indeed the case that $\alpha \rightsquigarrow \bot$ is equivalent to $\alpha \rightarrow \bot$. It is when α is an *open* compound that (Sub) need not hold. As result, DRFOL provides the domain modeler with greater flexibility in that it leaves open the possibility of there being only atypical objects, something that is not possible in the propositional and DL cases.

6 Conclusion and future work

In this paper, we have laid the theoretical groundwork for KLM-style defeasible RFOL. Our primary contribution is a set of rationality postulates describing the behaviour of DRFOL, a typicality semantics for interpreting defeasibility, and a representation result, proving that the proposed postulates characterise the semantic behaviour precisely.

With the theoretical core in place, we then proceeded to define a form of defeasible entailment for DRFOL that can be viewed as the DRFOL equivalent of the propositional form of defeasible entailment known as Rational Closure.

With a suitable definition of DRFOL defeasible entailment in place, the next step is to design algorithms for computing DRFOL defeasible entailment. Here we plan to draw inspiration from both the propositional and DL cases, where defeasible entailment can be reduced to a series of classical entailment checks, sometimes in polynomial time and with a polynomial number of classical entailment checks.

The theoretical framework presented in this paper also places us in a position to investigate extensions to other restricted versions of first-order logic.

Acknowledgments

This work was partially supported by the ANR Chaire IA BE4musIA: BElief change FOR better MUlti-Source Information Analysis, and by TAILOR, a project funded by EU Horizon 2020 research and innovation programme under GA No 952215.

References

1. Arlo-Costa, H.: The Logic of Conditionals. In: The Stanford Encyclopedia of Philosophy. Summer 2019 edn. (2019)
2. Beierle, C., Falke, T., Kutsch, S., Kern-Isberner, G.: Minimal Tolerance Pairs for System Z-Like Ranking Functions for First-Order Conditional Knowledge Bases. In: Proc. of FLAIRS 2016. pp. 626–631. AAAI Press (2016)
3. Beierle, C., Falke, T., Kutsch, S., Kern-Isberner, G.: System Z^{FO}: Default reasoning with system Z-like ranking functions for unary first-order conditional knowledge bases. Int. J. Approx. Reason. **90**, 120–143 (2017)
4. Benferhat, S., Cayrol, C., Dubois, D., Lang, J., Prade, H.: Inconsistency Management and Prioritized Syntax-based Entailment. In: Proc. of IJCAI-93. pp. 640–645. Morgan Kaufmann Publishers Inc. (1993)
5. Bonatti, P.A.: Rational closure for all description logics. Artificial Intelligence **274**, 197–223 (2019)
6. Bonatti, P.A., Faella, M., Petrova, I.M., Sauro, L.: A new semantics for overriding in description logics. Artificial Intelligence **222**, 1–48 (2015)
7. Booth, R., Paris, J.B.: A Note on the Rational Closure of Knowledge Bases with Both Positive and Negative Knowledge. J. Log. Lang. Inf. **7**(2), 165–190 (1998)
8. Brafman, R.I.: A First-order Conditional Logic with Qualitative Statistical Semantics. J. Log. Comput. **7**(6), 777–803 (1997)
9. Britz, K., Casini, G., Meyer, T., Moodley, K., Sattler, U., Varzinczak, I.: Principles of KLM-Style Defeasible Description Logics. ACM T. Comput. Log. **22**(1) (2021)
10. Casini, G., Meyer, T., Moodley, K., Nortjé, R.: Relevant closure: A new form of defeasible reasoning for description logics. In: Proc. of JELIA 2014, LNCS, vol. 8761, pp. 92–106. Springer (2014)
11. Casini, G., Straccia, U.: Rational Closure for Defeasible Description Logics. In: Proc. of JELIA 2010. LNCS, vol. 6341, pp. 77–90. Springer-Verlag (2010)
12. Casini, G., Straccia, U.: Defeasible Inheritance-Based Description Logics. Journal of Artificial Intelligence Research **48**, 415–473 (2013)
13. Delgrande, J.P.: On first-order conditional logics. Artificial Intelligence **105**(1), 105 – 137 (1998)
14. Delgrande, J.P., Rantsoudis, C.: A Preference-Based Approach for Representing Defaults in First-Order Logic. In: Proc. of NMR 2020. pp. 120–129 (2020)

15. Giordano, L., Gliozzi, V.: Strengthening the Rational Closure for Description Logics: An Overview. In: Proc. of CILC 2019. pp. 68–81. CEUR-WS.org (2019)
16. Giordano, L., Gliozzi, V., Olivetti, N., Pozzato, G.L.: A non-monotonic Description Logic for reasoning about typicality. Artificial Intelligence **195**, 165–202 (2013)
17. Giordano, L., Gliozzi, V., Olivetti, N., Pozzato, G.L.: Semantic characterization of rational closure: From propositional logic to description logics. Art. Int. **226**, 1–33 (2015)
18. Kern-Isberner, G.: Conditionals in Nonmonotonic Reasoning and Belief Revision - Considering Conditionals as Agents, LNCS, vol. 2087. Springer (2001)
19. Kern-Isberner, G.: A Thorough Axiomatization of a Principle of Conditional Preservation in Belief Revision. Ann. Math. Artif. Intell. **40**(1-2), 127–164 (2004)
20. Kern-Isberner, G., Beierle, C.: A System Z-like Approach for First-Order Default Reasoning. In: Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation. LNCS, vol. 9060, pp. 81–95. Springer (2015)
21. Kern-Isberner, G., Thimm, M.: A ranking semantics for first-order conditionals. In: Proc. of ECAI 2012. p. 456–461. IOS Press (2012)
22. Kraus, S., Lehmann, D., Magidor, M.: Nonmonotonic reasoning, preferential models and cumulative logics. Artificial Intelligence **44**, 167–207 (1990)
23. Lehmann, D.: Another perspective on default reasoning. Ann. Math. Artif. Intell. **15**(1), 61–82 (1995)
24. Lehmann, D., Magidor, M.: What does a conditional knowledge base entail? Art. Intell. **55**, 1–60 (1992)
25. McCarthy, J.: Circumscription, a form of nonmonotonic reasoning. Art. Intell. **13**(1-2), 27–39 (1980)
26. Pearl, J.: System Z: a natural ordering of defaults with tractable applications to nonmonotonic reasoning. In: Proc. of TARK 1990 (1990)
27. Pensel, M., Turhan, A.Y.: Reasoning in the Defeasible Description Logic $\mathcal{EL}_{\bot}$ - computing standard inferences under rational and relevant semantics. Int. J. Approx. Reason. **103**, 28 – 70 (2018)
28. Schlechta, K.: Defaults as Generalized Quantifiers. Journal of Logic and Computation **5**(4), 473–494 (1995)