KLM-Style Defeasibility for Restricted First-Order Logic

Giovanni Casini[1,2], **Thomas Meyer**[1,2], **Guy Paterson-Jones**[2]
[1] ISTI–CNR, Italy
[2] University of Cape Town and CAIR, South Africa
giovanni.casini@isti.cnr.it, tmeyer@cs.uct.ac.za, PTRGUY002@myuct.ac.za

Abstract

We extend the KLM approach to defeasible reasoning to be applicable to a restricted version of first-order logic. We describe defeasibility for this logic using a set of rationality postulates, provide an appropriate semantics for it, and present a representation result that characterises the semantic description of defeasibility in terms of the rationality postulates. Based on this theoretical core, we then propose a version of defeasible entailment that is inspired by 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 our rationality postulates. The work in this paper is the first step towards our ultimate goal of introducing KLM-style defeasible reasoning into the family of Datalog+/- ontology languages.

1 Introduction

The past 15 years have seen a flurry of activity to introduce defeasible-reasoning capabilities into languages that are more expressive than propositional logic (Casini and Straccia 2010, 2013; Casini et al. 2015; Giordano et al. 2013, 2015; Bonatti et al. 2015; Bonatti 2019; Pensel and Turhan 2018). 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, Lehmann, and Magidor (1990), and Lehmann and Magidor (1992). In DLs, knowledge is expressed as general concept inclusions of the form $C \sqsubseteq D$, where C and D are concepts, with the intended meaning that every instance of C is also an instance of D. Defeasible DLs allow, in addition, for defeasible concept inclusions of the form $C \mathrel{\sqsubseteq\!\!\!\sim} D$ with the intended meaning that instances of C are *usually* instances of D. For instance, $\mathsf{Student} \mathrel{\sqsubseteq\!\!\!\sim} \neg\exists\mathsf{pays}.\mathsf{Tax}$ (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).

Given the tight formal relationship between DLs and the family of Datalog+/- ontology languages (Calì et al. 2010; Calì, Gottlob, and Lukasiewicz 2012), it is surprising that this form of defeasibility has not yet found its way into Datalog+/-. In this paper we take the first steps to fill that gap by providing the theoretical foundations for defeasibility in a restricted version of first-order logic. We refer to the classical version of the logic as RFOL and the defeasible extension as DRFOL. It suffices to to use Herbrand interpretations for the semantics of RFOL. 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 type of language in which it is implemented. For propositional logics the intuition dictates a notion of typicality over possible worlds. The statement "birds usually fly", formalised as $\mathsf{bird} \mathrel{|\!\sim} \mathsf{fly}$, is intended to convey 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. The statement $\mathsf{Student} \mathrel{\sqsubseteq\!\!\!\sim} \neg\exists\mathsf{pays}.\mathsf{Tax}$ states that of all those individuals that are students, the most typical ones don't pay taxes. Consider, for instance, the following example of (Delgrande 1998):

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).

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

Note that all statements, except for the first one, are defeasible. For any appropriate semantics, the knowledge base in the example should be satisfiable (given an appropriate notion of satisfiability). With this in mind 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 $\mathsf{elephant}(\mathsf{x}) \sim\!\!\!\rightarrow \mathsf{wild\_animal}(\mathsf{x})$ would mean that the most typical elephants are wild animals. Similarly, $\mathsf{keeper}(\mathsf{x}) \sim\!\!\!\rightarrow \mathsf{human}(\mathsf{x})$ would mean that the most typical keepers are human. Combined with the first statement in the knowledge base, 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 the knowledge base 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.

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 the typicality of individuals. We achieve the latter by enriching our semantics with a set of *typicality objects*, the elements of which are used to represent *typical* individuals. Thus, $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{y}) \sim\!\!\!\rightarrow \mathsf{feeds}(\mathsf{y},\mathsf{x})$ 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.

The central theoretical result of the paper is a representation result (Theorems 2 and 3), showing that defeasible implication defined in this way can be characterised w.r.t. a set of KLM-style rationality postulates adapted for DRFOL. We show that DRFOL formally generalises propositional KLM-style defeasible reasoning in two ways. The cases where DRFOL, restricted to 0-ary predicates, or where n-ary predicates for any $n > 0$ are allowed, but with a restriction to variable-free statements, both reduce to propositional KLM-style defeasibility. A comparison with defeasible DLs is more complicated, but the semantics of defeasible DLs, for the most part, carries over to DRFOL. An important exception is that whereas a defeasible DL statement of the form $A \mathrel{\tilde{\sqsubseteq}} \bot$ is equivalent to its classical counterpart $A \sqsubseteq \bot$, it is possible to distinguish between the DRFOL version of the same statement, $A(x) \sim\!\!\!\rightarrow \bot$, and its classical counterpart $A(x) \rightarrow \bot$. In fact, the former is weaker than the latter.

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 in Section 5, and show that it can be viewed as the DRFOL analogue of Rational Closure, as originally defined for the propositional case (Kraus, Lehmann, and Magidor 1990; Lehmann and Magidor 1992).

The rest of the paper is structured as follows. Section 2 is a brief introduction to RFOL, as well as to KLM-style defeasible reasoning for propositional logics. Section 4 is the heart of the paper. It introduces DRFOL, describes an abstract notion of satisfaction w.r.t. a set of KLM-style postulates, provides a semantics, and proves a representation result, showing that the KLM-style postulates characterise the semantic construction. Section 5 presents a form of defeasible entailment for DRFOL that can be viewed as the DRFOL equivalent of the well-known propositional form of defeasible entailment known as Rational Closure. Section 6 compares defeasible reasoning in DRFOL with KLM-style defeasible reasoning in propositional logic and DLs. Section 7 provides an overview of related work, while Section 8 concludes and briefly discusses future work. The proofs can be found in an appendix: https://tinyurl.com/7472fn2a.

2 Background

We consider a restricted version of a first-order language, which we refer to as RFOL. The language of RFOL is defined by three disjoint sets of symbols: CONST, a finite set of constants; VAR, a countably infinite set of variable symbols; and PRED, a finite set of predicate symbols. It has no function symbols. Associated with each predicate symbol $\alpha \in$ PRED is an *arity*, denoted $\mathfrak{ar}(\alpha) \in \mathbb{N}$, which represents the number of terms it takes as arguments. We assume the existence of predicate symbols $\top$ and $\bot$, which we take to have arity 0. A *term* is an element of CONST $\cup$ VAR. An *atom* is an expression of the form $\alpha(t_1, \ldots, t_{\mathfrak{ar}(\alpha)})$ where $\alpha \in$ PRED and the t_i are terms. Observe that $\top$ and $\bot$ are atoms as well.

A *compound* is defined to be a boolean combination of atoms, i.e. an expression built out of atoms and the standard logical connectives $\neg$, $\wedge$, and $\vee$. An *implication* is defined to have the form $A(\vec{x}) \rightarrow B(\vec{y})$ 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 (respectively, implication) is said to be *ground* if all the terms contained in it are constants; otherwise it is *open*. In RFOL, the only formulas we permit are compounds and implications. When viewed as formulas, compounds and implications are understood to be implicitly universally quantified.

We adopt the following conventions for various kinds of formula. Constant symbols and variables will be written in lowercase English, with early letters used for constants $(a, b, \ldots)$ and later letters used for variables $(x, y, \ldots)$. Compounds will be written in uppercase English $(A, B, \ldots)$. A tuple of variables or constants will be written with overbars, such as $\vec{x}$ and $\vec{a}$ respectively, and $A(\vec{x})$ and $B(\vec{a})$ will be used as shorthand for compounds over their respective tuples of terms. We use lowercase greek $(\alpha, \beta, \ldots)$ to denote RFOL formulas.

We omit specifying the symbol sets under consideration, as they can be inferred from context. The set of all formulas (compounds and implications) is denoted by $\mathcal{L}$, and a *knowledge base* $\mathcal{K}$ is defined to be a finite subset of $\mathcal{L}$.

RFOL can be thought of as an extension of Datalog (Abiteboul, Hull, and Vianu 1995). In fact, we use *Herbrand interpretations* to specify the semantics of RFOL. The Herbrand universe $\mathbb{U}$ is the set of constant symbols 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 $\mathcal{H} \subseteq \mathbb{B}$.

Substitutions are defined to be functions φ : VAR $\rightarrow$ VAR $\cup$ 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 $\mathcal{H} \Vdash A(\vec{a}) \rightarrow B(\vec{b})$ according to the usual laws of 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 $A(\vec{x}) \rightarrow B(\vec{y})$ is an open implication (where $\vec{x}$ and $\vec{y}$ may overlap), then $\mathcal{H} \Vdash A(\vec{x}) \rightarrow B(\vec{y})$ iff $\mathcal{H} \Vdash A(\varphi(\vec{x})) \rightarrow B(\varphi(\vec{y}))$ for every ground substitution φ.
5. If $\mathcal{K}$ is a knowledge base, then $\mathcal{H} \Vdash \mathcal{K}$ iff $\mathcal{H} \Vdash \alpha$ for every $\alpha \in \mathcal{K}$.

The set of Herbrand interpretations is denoted by $\mathscr{H}$. A Herbrand interpretation that satisfies a knowledge base $\mathcal{K}$ is a *Herbrand model* of $\mathcal{K}$.

3 Propositional Defeasible Reasoning

Kraus, Lehmann, and Magidor (1990) originally define $\mid\!\sim$ 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". Lehmann and Magidor (1992) subsequently shift to interpreting $\alpha \mid\!\sim \beta$ as the object-level statement "α defeasibly implies β", with $\mid\!\sim$ viewed as an object-level connective. An abstract notion of satisfaction can then be defined in terms of *satisfaction sets*. A satisfaction set $\mathcal{S}$ of statements of the form $\alpha \mid\!\sim \beta$ is said to be *rational* if it satisfies the well-known KLM properties below (Lehmann and Magidor 1992). Lehmann and Magidor did not refer to satisfaction sets, but our formulation here is equivalent to theirs for the propositional case:

$$\text{(Refl)}\ \ \alpha \mid\!\sim \alpha \in \mathcal{S}$$

$$\text{(Rw)}\ \ \frac{\alpha \mid\!\sim \beta \in \mathcal{S},\ \models \beta \rightarrow \gamma}{\alpha \mid\!\sim \gamma \in \mathcal{S}}$$

$$\text{(Lle)}\ \ \frac{\models \alpha \leftrightarrow \beta,\ \alpha \mid\!\sim \gamma \in \mathcal{S}}{\beta \mid\!\sim \gamma \in \mathcal{S}}$$

$$\text{(And)}\ \ \frac{\alpha \mid\!\sim \beta \in \mathcal{S},\ \alpha \mid\!\sim \gamma \in \mathcal{S}}{\alpha \mid\!\sim \beta \wedge \gamma \in \mathcal{S}}$$

$$\text{(Or)}\ \ \frac{\alpha \mid\!\sim \gamma \in \mathcal{S},\ \beta \mid\!\sim \gamma \in \mathcal{S}}{\alpha \vee \beta \mid\!\sim \gamma \in \mathcal{S}}$$

$$\text{(Rm)}\ \ \frac{\alpha \mid\!\sim \beta \in \mathcal{S},\ \alpha \mid\!\sim \neg\gamma \notin \mathcal{S}}{\alpha \wedge \gamma \mid\!\sim \beta \in \mathcal{S}}$$

A semantics for defeasible implications is provided by *ranked interpretations* $\mathscr{R}$, which are defined to be total preorders over a subset $U_{\mathscr{R}} \subseteq U$ of valuations. Valuations that are lower in the ordering are considered to be more typical, whereas valuations that are not in $U_{\mathscr{R}}$ are impossibly atypical. A defeasible statement $\alpha \mid\!\sim \beta$ is *satisfied in* $\mathscr{R}$ ($\mathscr{R} \Vdash \alpha \mid\!\sim \beta$) iff the $\mathscr{R}$-minimal models of α are also models of β, which formalises the intuition that β holds in the most typical situations in which α is true. A classical statement α is satisfied by $\mathscr{R}$ iff every valuation in $U_{\mathscr{R}}$ satisfies α.

Lehmann and Magidor (1992) prove the following correspondence between rational satisfaction sets and ranked interpretations:

Theorem 1. *(Lehmann and Magidor 1992). A set $\mathcal{S}$ of statements of the form $\alpha \mid\!\sim \beta$ is a rational satisfaction set iff there is a ranked interpretation $\mathscr{R}$ such that $\alpha \mid\!\sim \beta \in \mathcal{S}$ iff $\mathscr{R} \Vdash \alpha \mid\!\sim \beta$.*

To conclude this section, observe that $\mathscr{R} \Vdash \neg\alpha \mid\!\sim \bot$ iff $\mathscr{R}$ contains no models of $\neg\alpha$ (which are therefore viewed as impossible). In other words, $\mathscr{R} \Vdash \neg\alpha \mid\!\sim \bot$ iff $\mathscr{R} \Vdash \alpha$. We return to this property of propositional defeasible reasoning in Section 6.

4 Defeasible Restricted First-Order Logic

Defeasible Restricted First-Order Logic (DRFOL for short) extends the logic RFOL that was presented in Section 2 with *defeasible implications* of the form $A(\vec{x}) \rightsquigarrow B(\vec{y})$, where $A(\vec{x})$ and $B(\vec{y})$ are compounds, and where $\vec{x}$ and $\vec{y}$ may overlap. Observe that $\rightsquigarrow$ is intended to be the defeasible analogue of classical implication. That is, $A(\vec{x}) \rightsquigarrow B(\vec{y})$ is the defeasible analogue of the RFOL formula $A(\vec{x}) \rightarrow B(\vec{y})$. The set of defeasible implications is denoted $\mathcal{L}^{\rightsquigarrow}$, and a *DRFOL knowledge base* $\mathcal{K}$ is defined to be a subset of $\mathcal{L} \cup \mathcal{L}^{\rightsquigarrow}$. 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, $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \neg\mathsf{feeds(x, fred)}$, is an exception to $\mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \rightsquigarrow \mathsf{feeds(x, y)}$. 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 $\{\mathsf{bird(x)} \rightsquigarrow \mathsf{fly(x)}, \mathsf{bird(tweety)}, \neg\mathsf{fly(tweety)}\}$. To formalise these intuitions, we describe the intended behaviour of the defeasible connective $\rightsquigarrow$, and its interaction with (classical) RFOL formulas, in terms of a set of rationality postulates in the KLM style (Kraus, Lehmann, and Magidor 1990; Lehmann and Magidor 1992). These postulates are expressed via an abstract notion of satisfaction:

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

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

$$\text{(ClassF)}\ \ \frac{\mathcal{S}_C \models A(\vec{x})}{A(\vec{x}) \in \mathcal{S}}$$

$$\text{(ClassR)}\ \ \frac{\mathcal{S}_C \models A(\vec{x}) \rightarrow B(\vec{y})}{A(\vec{x}) \rightarrow B(\vec{y}) \in \mathcal{S}}$$

Next, we consider the interaction between classical and defeasible implications. We expect the following supraclassicality postulate to hold:

$$\text{(SupR)}\ \ \frac{A(\vec{x}) \rightarrow B(\vec{y}) \in \mathcal{S}}{A(\vec{x}) \rightsquigarrow B(\vec{y}) \in \mathcal{S}}$$

A similar postulate for compounds then holds:

$$\text{(SupF)}\ \frac{A(\vec{x}) \in \mathcal{S}}{\neg A(\vec{x}) \rightsquigarrow \bot \in \mathcal{S}}$$

Proposition 1. (SupF) *follows from* (ClassR) *and* (SupR).

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

$$\text{(Refl)}\ A(\vec{x}) \rightsquigarrow A(\vec{x}) \in \mathcal{S}$$

$$\text{(Rw)}\ \frac{A(\vec{x}) \rightsquigarrow B(\vec{y}) \in \mathcal{S},\ \models B(\vec{y}) \rightarrow C(\vec{z})}{A(\vec{x}) \rightsquigarrow C(\vec{z}) \in \mathcal{S}}$$

$$\text{(Lle)}\ \frac{A(\vec{x}) \rightsquigarrow C(\vec{z}) \in \mathcal{S},\ \models A(\vec{x}) \rightarrow B(\vec{y}),\ \models B(\vec{y}) \rightarrow A(\vec{x})}{B(\vec{y}) \rightsquigarrow C(\vec{z}) \in \mathcal{S}}$$

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

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

$$\text{(Rm)}\ \frac{A(\vec{x}) \rightsquigarrow B(\vec{y}) \in \mathcal{S},\ A(\vec{x}) \rightsquigarrow \neg C(\vec{z}) \notin \mathcal{S}}{A(\vec{x}) \wedge C(\vec{z}) \rightsquigarrow B(\vec{y}) \in \mathcal{S}}$$

Next we consider *instantiations* of implications. To begin with, note that universal instantiation is *not* a desirable property for defeasible implications:

$$\text{(Duir)}\ \frac{A(\vec{x}) \rightsquigarrow B(\vec{y}) \in \mathcal{S}}{A(\varphi(\vec{x})) \rightsquigarrow B(\varphi(\vec{y})) \in \mathcal{S}}$$

To see why, consider a satisfaction set $\mathcal{S}$ containing $\mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \rightsquigarrow \mathsf{feeds(y,x)}$ and $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \neg\mathsf{feeds(y,fred)}$. From (Duir) we have $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \rightsquigarrow \mathsf{feeds(y,fred)} \in \mathcal{S}$, 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, where ψ is a variable substitution and $\vec{x} \cap \vec{y} = \emptyset$:

$$\text{(Irr)}\ \frac{A(\vec{x}) \rightsquigarrow B(\vec{x},\vec{y}) \in \mathcal{S}}{A(\vec{x}) \rightsquigarrow B(\vec{x},\psi(\vec{y})) \in \mathcal{S}}$$

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 $B(\vec{y})$ with respect to $\mathcal{S}$, 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 $\preccurlyeq_{\mathcal{S}}$ does *not* partially order $\mathcal{L}$ in general, but is rather a partial ordering of the subset of *consistent* formulas of $\mathcal{L}$, i.e. $A(\vec{x}) \in \mathcal{L}$ such that $A(\vec{x}) \rightsquigarrow \bot \notin \mathcal{S}$.

For any variable substitution ψ, a typical instance of $A(\psi(\vec{x}))$ is always an instance of $A(\vec{x})$. Thus we should expect the following postulate to hold, where ψ is any variable substitution:

$$\text{(Typ)}\ A(\vec{x}) \preccurlyeq_{\mathcal{S}} A(\psi(\vec{x}))$$

The last postulate we consider has to do with defeasibly impossible formulas. Suppose that $A(\varphi(\vec{x})) \rightsquigarrow \bot \in \mathcal{S}$ for all substitutions $\varphi : \text{VAR} \rightarrow \text{VAR} \cup \mathbb{U}$. This intuitively states that there are no typical instances of *any* specialisation of $A(\vec{x})$. Thus we should expect that there are in fact no instances of $A(\vec{x})$ at all:

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

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

Definition 3. A satisfaction set $\mathcal{S}$ is *rational* iff it satisfies (ClassF), (ClassR), (SupR), (Irr), (Typ), (Imp) and (Refl)-(Rm).

Note that rational satisfaction sets satisfy the following form of label invariance for defeasible implications, where the variable substitution ψ is a *permutation*:

$$\text{(Per)}\ \frac{A(\vec{x}) \rightsquigarrow B(\vec{y}) \in \mathcal{S}}{A(\psi(\vec{x})) \rightsquigarrow B(\psi(\vec{y})) \in \mathcal{S}}$$

Proposition 2. (Per) *follows from* (Refl)-(Rm), (Irr) *and* (Typ).

4.1 Semantics

We now proceed to define an appropriate semantics for defeasible implications. The first step is to enrich the Herbrand universe with a set $\mathcal{T}$ of *typicality objects*. Typicality objects represent the individuals that aren't explicitly mentioned in a given knowledge base, and are used to interpret defeasible implications in a ranking of (enriched) Herbrand interpretations.

Definition 4. The *enriched Herbrand universe* is defined to be the set $\mathbb{U}_{\mathcal{T}} = \mathbb{U} \cup \mathcal{T}$. An *enriched Herbrand interpretation* (or EHI) $\mathcal{E}$ is a Herbrand interpretation over the enriched Herbrand universe.

Observe that every enriched Herbrand interpretation $\mathcal{E}$ 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 $rk : \mathscr{H}_{\mathcal{T}} \rightarrow \Omega \cup \{\infty\}$, 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 $\varphi : \text{VAR} \rightarrow \mathcal{T}\}$ 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* (Kraus, Lehmann, and Magidor 1990).

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

Definition 6. For $A(\vec{x}), B(\vec{y}) \in \mathcal{L}$:

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* $\Vdash A(\vec{x}) \rightarrow B(\vec{y})$ iff $\mathcal{E} \Vdash A(\vec{x}) \rightarrow B(\vec{y})$ for all $\mathcal{E} \in \mathscr{H}^{rk}_{\mathcal{T}}$.
3. *rk* $\Vdash A(\vec{x}) \sim\!\!\!\rightarrow B(\vec{y})$ iff $\mathcal{E} \Vdash A(\varphi(\vec{x})) \rightarrow B(\varphi(\vec{y}))$ for all $\mathcal{E} \in \min_{rk} \mathscr{H}^{rk}_{\mathcal{T}}(A(\vec{x}))$ and all $\varphi : \text{VAR} \rightarrow \mathcal{T}$.

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 counterpart, with variables substituted for 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 an example proposed by Delgrande (1998). The following DRFOL knowledge base states that elephants usually like keepers, that elephants usually *don't* like the keeper Fred, and that the elephant Clyde usually *does* like Fred:

$$\mathcal{K} = \{\mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \sim\!\!\!\rightarrow \mathsf{likes(x,y)},\ \mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \sim\!\!\!\rightarrow \neg\mathsf{likes(x,fred)},\ \mathsf{elephant(clyde)} \wedge \mathsf{keeper(fred)} \sim\!\!\!\rightarrow \mathsf{likes(clyde,fred)}\}.$$

Let $\mathcal{T} = \{\mathsf{t}_1, \ldots\}$ be the set of typicality objects. For readability we abbreviate elephant by e, keeper by k and likes by 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 $\mathcal{E}_3 = \{\mathsf{e(t_1)}, \mathsf{k(t_2)}, \mathsf{e(t_2)}, \mathsf{e(clyde)}, \mathsf{k(fred)}, \mathsf{l(clyde,fred)}\}$.

Let $rk_1(\mathcal{E}_1) = rk_1(\mathcal{E}_2) = 0$, $rk_1(\mathcal{E}_3) = 0$, and $rk_1(\mathcal{E}) = \infty$ for all other EHIs. Then rk_1 is a ranked model of the knowledge base above. Let $rk_2(\mathcal{E}_1) = rk_2(\mathcal{E}_3) = 0$, $rk_2(\mathcal{E}_2) = 1$, and $rk_2(\mathcal{E}) = \infty$ for all other EHIs. Then rk_2 is not a ranked model of $\mathsf{elephant(x)} \wedge \mathsf{keeper(y)} \sim\!\!\!\rightarrow \mathsf{likes(x,y)}$, but is a ranked model of $\mathsf{elephant(x)} \wedge \mathsf{keeper(fred)} \sim\!\!\!\rightarrow \neg\mathsf{likes(x,fred)}$ and $\mathsf{elephant(clyde)} \wedge \mathsf{keeper(fred)} \sim\!\!\!\rightarrow \mathsf{likes(clyde,fred)}$.

4.2 A Representation Result

In this section we show that ranked interpretations precisely characterise rational satisfaction sets.

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

Our representation result is obtained by showing that all ranked interpretations generate rational satisfaction sets (Theorem 2), and that every rational satisfaction set $\mathcal{S}$ can be realised as the satisfaction set corresponding to some ranked interpretation (Theorem 3).

Theorem 2. *For every ranked interpretation rk, $\mathcal{S}^{rk}$ is a rational satisfaction set.*

To show the converse of Theorem 2, we adapt the representation proof of Lehmann and Magidor Lehmann and Magidor (1992) to the DRFOL setting. The main idea is to show that the defeasible implications in a given rational satisfaction set can be completely characterised by *normal EHIs*, which are EHIs that satisfy all of the defeasible consequences of some compound $A(\vec{x})$. By ranking these normal EHIs over an appropriate linear poset, we can capture the satisfaction set exactly.

Definition 8. For a rational satisfaction set $\mathcal{S}$, the compounds $A(\vec{x}), B(\vec{y})$ are *equally typical* w.r.t. $\mathcal{S}$ (denoted $A(\vec{x}) \equiv_{\mathcal{S}} B(\vec{y})$) iff $A(\vec{x}) \preccurlyeq_{\mathcal{S}} B(\vec{y})$ and $B(\vec{y}) \preccurlyeq_{\mathcal{S}} A(\vec{x})$.

We denote the equivalence class of a compound $A(\vec{x}) \in \mathcal{L}$ with respect to $\equiv_{\mathcal{S}}$ by $[A(\vec{x})]_{\mathcal{S}}$. As predicates can have arbitrarily high arity in general, it is necessary in what follows to assume that $\mathcal{T}$ is a countably infinite set of typicality objects.

Definition 9. Let $\mathcal{S}$ be a rational satisfaction set. Then $\mathcal{E} \in \mathscr{H}_{\mathcal{T}}$ is *normal* for a formula $A(\vec{x}) \in \mathcal{L}$ w.r.t. $\mathcal{S}$ iff the following properties hold:

1. $\mathcal{E} \Vdash \alpha$ for all $\alpha \in \mathcal{S}_C$.
2. $\mathcal{E} \Vdash A(\varphi(\vec{x}))$ for some $\varphi : \text{VAR} \rightarrow \mathcal{T}$.
3. for all $B(\vec{y}) \in [A(\vec{x})]_{\mathcal{S}}$ and $\varphi : \text{VAR} \rightarrow \mathcal{T}$, $B(\vec{y}) \sim\!\!\!\rightarrow C(\vec{z}) \in \mathcal{S}$ implies that $\mathcal{E} \Vdash B(\varphi(\vec{y})) \rightarrow C(\varphi(\vec{z}))$.

The set of normal EHIs for $A(\vec{x})$ is denoted $\text{norm}_{\mathcal{S}}(A(\vec{x}))$. For the rest of this section, we will consider a *fixed* rational satisfaction set $\mathcal{S}$, and sketch the construction of a ranked interpretation $rk : \mathscr{H}_{\mathcal{T}} \rightarrow \Omega \cup \{\infty\}$ such that $\mathcal{S} = \mathcal{S}^{rk}$. First, we show that normal EHIs completely characterise the defeasible implications in a given rational satisfaction set:

Lemma 1. *$A(\vec{x}) \sim\!\!\!\rightarrow B(\vec{y}) \in \mathcal{S}$ iff for every $\mathcal{E} \in \text{norm}_{\mathcal{S}}(A(\vec{x}))$ and substitution $\varphi : \text{VAR} \rightarrow \mathcal{T}$ we have $\mathcal{E} \Vdash A(\varphi(\vec{x})) \rightarrow B(\varphi(\vec{y}))$.*

Corollary 1. *$A(\vec{x})$ has a normal EHI iff $A(\vec{x})$ is consistent with respect to $\mathcal{S}$, i.e. $A(\vec{x}) \sim\!\!\!\rightarrow \bot \notin \mathcal{S}$.*

Let $\Omega^* = \{\langle A(\vec{x}), \mathcal{E}\rangle : A(\vec{x}) \in \mathcal{L},\ \mathcal{E} \in \text{norm}_{\mathcal{S}}(A(\vec{x}))\}$. We order elements of Ω^* using the relation $\preccurlyeq_{\mathcal{S}}$ as follows:

$$\langle A(\vec{x}), \mathcal{E}^A\rangle \leq \langle B(\vec{y}), \mathcal{E}^B\rangle \text{ iff } A(\vec{x}) \preccurlyeq_{\mathcal{S}} B(\vec{y})$$

Proposition 3. *$\leq$ is reflexive, transitive and total over Ω^*.*

Let $\Omega = \Omega^*/\sim$ be the quotient of Ω^* with respect to its equivalence classes, which we denote by $[\alpha]_{\leq}$ for $\alpha \in \Omega^*$. By Proposition 3, Ω is a linear poset, though in general it is not well-ordered. We now show that any given EHI is contained in at most one equivalence class:

Lemma 2. *For any $\mathcal{E} \in \mathscr{H}_{\mathcal{T}}$, the following set is either empty or contains a single element:*

$$\Omega(\mathcal{E}) = \{[\langle A(\vec{x}), \mathcal{E}\rangle]_{\leq} : \langle A(\vec{x}), \mathcal{E}\rangle \in \Omega^*\}.$$

This lets us construct a ranking function $rk : \mathscr{H}_{\mathcal{T}} \rightarrow \Omega \cup \{\infty\}$ as follows:

$$rk(\mathcal{E}) = \begin{cases} [\langle A(\vec{x}), \mathcal{E}\rangle]_{\leq} & \text{if } \Omega(\mathcal{E}) = \{[\langle A(\vec{x}), \mathcal{E}\rangle]_{\leq}\} \\ \infty & \text{if } \Omega(\mathcal{E}) = \emptyset \end{cases}$$

Proposition 4. *The ranking function $rk : \mathscr{H}_{\mathcal{T}} \to \Omega \cup \{\infty\}$ is a ranked interpretation.*

Finally, we have the following result relating normal EHIs to minimal elements in *rk*:

Lemma 3. *For any formula $A(\vec{x}) \in \mathcal{L}$, we have that $\min_{rk} \mathscr{H}^{rk}_{\mathcal{T}}(A(\vec{x})) = norm_{\mathcal{S}}(A(\vec{x}))$.*

Lemmas 1 and 3 prove the converse to Theorem 2.

Theorem 3. *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 $\mathcal{S} = \mathcal{S}^{rk}$.*

4.3 Finite Sets of Typicality Objects

Theorem 3 has some limitations in that it requires an infinite set of typicality objects to be true in general. In this section we detail some ways ranked interpretations can be restricted to *finite* sets of typicality objects, which will be useful for defining a basic notion of entailment for DRFOL knowledge bases.

First, consider a fixed finite set $\mathcal{T}' \subset \mathcal{T}$. Note that the set of EHIs over $\mathcal{T}'$ is finite, as there are only finitely many possible atoms over the extended Herbrand base $\mathbb{B}_{\mathcal{T}'}$. Furthermore, given any such $\mathcal{E} \in \mathscr{H}_{\mathcal{T}'}$, we can define a *characteristic compound* for $\mathcal{E}$ that parallels the notion of characteristic formula for a propositional valuation:

Definition 10. Let $\mathcal{E} \in \mathscr{H}_{\mathcal{T}'}$ be an EHI over $\mathcal{T}'$, and $\pi : \mathcal{T}' \to \text{VAR}$ any injective function. Then the *characteristic compound* for $\mathcal{E}$, denoted $\text{ch}_\pi(\mathcal{E})$, is defined as follows:

$$\text{ch}_\pi(\mathcal{E}) = \bigwedge_{A(\vec{c},\vec{t}) \in \mathbb{B}_{\mathcal{T}'}} \pm A(\vec{c}, \pi(\vec{t}))$$

Here, $\vec{c}$ is a tuple of constants, $\vec{t}$ is a tuple of typicality objects, and $\pm A(\vec{x}, \pi(\vec{t}))$ means $A(\vec{c}, \pi(\vec{t}))$ if $\mathcal{E} \Vdash A(\vec{c}, \vec{t})$, or $\neg A(\vec{c}, \pi(\vec{t}))$ otherwise.

Note that, while $\text{ch}_\pi(\mathcal{E})$ depends on π, the characteristic formula is nevertheless unique up to relabelling of variables and the order of clauses. For this reason we will omit defining π explicitly where we refer to it. The important fact about characteristic formulas is that they reflect satisfaction properties of the underlying EHI $\mathcal{E}$:

Lemma 4. *Let $\mathcal{E} \in \mathscr{H}_{\mathcal{T}}$ and $\mathcal{E}' \in \mathscr{H}_{\mathcal{T}'}$ be any two EHIs over $\mathcal{T}$ and $\mathcal{T}'$ respectively such that $\mathcal{E} \Vdash \varphi(ch_\pi(\mathcal{E}'))$ for some $\varphi : \text{VAR} \to \mathcal{T}$. Then for any compound $A(\vec{x})$ and substitution $\psi : \text{VAR} \to \mathcal{T}'$, $\mathcal{E}' \Vdash A(\psi(\vec{x}))$ iff $\mathcal{E} \Vdash A(\varphi \circ \pi \circ \psi(\vec{x}))$.*

The number of typicality objects required to model a defeasible formula depends on the number of variables in the formula. With this in mind, we define the *order* of a formula $A(\vec{x})$ to be the length of the tuple $\vec{x}$.

Definition 11. For any ranked interpretation $rk \in \mathcal{R}_{\mathcal{T}}$, the *restriction of rk to $\mathcal{E}'$*, denoted $rk^* \in \mathcal{R}_{\mathcal{T}'}$, is defined by $rk^*(\mathcal{E}) = \min_{rk} \mathscr{H}^{rk}_{\mathcal{T}}(\text{ch}_\pi(\mathcal{E}))$.

The following lemma proves that rk^* and *rk* agree for formulas of small enough order:

Lemma 5. *rk^* satisfies the following properties, where $n = |\mathcal{T}'|$ is the number of typicality objects in $\mathcal{T}'$:*

1. *for all classical formulas $\alpha \in \mathcal{L}$, $rk^* \Vdash \alpha$ iff $rk \Vdash \alpha$.*
2. *for all defeasible formulas $\alpha \in \mathcal{L}^{\sim}$ of order $\leq n$, $rk^* \Vdash \alpha$ iff $rk \Vdash \alpha$.*

This lets us define approximations to any given ranked interpretation using a finite subset of typicality objects. In particular, if one only cares about satisfaction for formulas of bounded order, then a finite set suffices to model them. Defining the order of a knowledge base to be the maximum order of any formula contained within it, we have the following corollary:

Corollary 2. *Let $\mathcal{K} \subseteq \mathcal{L} \cup \mathcal{L}^{\sim}$ be any knowledge base of order n. Then $\mathcal{K}$ has a ranked model iff it has a ranked model over a set $\mathcal{T}'$ of typicality objects where $|\mathcal{T}'| = n$.*

5 Defeasible Entailment

A central question that we have left unaddressed 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* (Lehmann and Magidor 1992) for DRFOL. It is, by now, well-established that systems for defeasible reasoning are amenable to multiple forms of defeasible entailment, and the work we present in this section should therefore be viewed as the first step in a larger investigation into defeasible entailment.

Rational Closure is a well-known framework for non-monotonic reasoning that can be viewed as one of the core forms of defeasible entailment in KLM-style reasoning. Due to the so-called *drowning effect* (Benferhat et al. 1993), 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 (Lehmann 1995; Casini and Straccia 2013; Casini et al. 2014; Giordano and Gliozzi 2019). It has gained attention in the framework of DLs (Casini and Straccia 2010; Britz et al. 2020; Giordano et al. 2015; Bonatti et al. 2015). An equivalent semantic construction, System Z (Pearl 1990), has been considered for unary first-order logic (Kern-Isberner and Beierle 2015; Beierle et al. 2016, 2017). Several equivalent definitions of Rational Closure can be found in the literature. Here we refer to the one due to Booth and Paris (1998).

Let a knowledge base $\mathcal{K}$ be a set of propositional defeasible implications $\alpha \mathrel{|\!\sim} \beta$ (see Section 3). Booth and Paris provide a construction with the following two immediate consequences:

1. 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 is such that it assigns to every propositional valuation v the *minimal* rank assigned to it by any of the ranked models of $\mathcal{K}$.
2. Propositional Rational Closure can be characterised using $\mathscr{R}^*$. That is, $\alpha \mathrel{|\!\sim} \beta$ is in the (propositional) Rational Closure of $\mathcal{K}$ iff $\mathscr{R}^* \Vdash \alpha \mathrel{|\!\sim} \beta$. The intuition behind the use of the ranked model $\mathscr{R}^*$ for the definition of entailment is that it formalises the *presumption of typicality* (Lehmann 1995): 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.

Based on Corollary 2 and the other results in Section 4.3, we can define an analogous construction for DRFOL:

Definition 12. Let $\mathcal{K} \subseteq \mathcal{L} \cup \mathcal{L}^{\rightsquigarrow}$ be a DRFOL knowledge base of order n, and take $\mathcal{T}' \subset \mathcal{T}$ to be a finite set of typicality objects of cardinality n. Then the *minimal ranked interpretation* of $\mathcal{K}$, which we denote by $rk_{\mathcal{K}} : \mathscr{H}_{\mathcal{T}'} \to \mathbb{N} \cup \{\infty\}$, is defined as follows:

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

Note that we take $\min \emptyset = \infty$ by convention, and that $rk_{\mathcal{K}}$ is a ranked interpretation over $\mathcal{T}'$, hence $rk_{\mathcal{K}} \in \mathcal{R}_{\mathcal{T}'}$. Intuitively, $rk_{\mathcal{K}}$ is what you get if you let every EHI rank as low as possible amongst the models of $\mathcal{K}$. This minimal ranked interpretation can be used to define a defeasible entailment relation for DRFOL:

Definition 13. For any DRFOL knowledge base $\mathcal{K}$ and formula α, we say that α is in the *Rational Closure* of $\mathcal{K}$, denoted $\mathcal{K} \mathrel{|\!\approx_{rc}} \alpha$, iff $rk_{\mathcal{K}} \Vdash \alpha$.

Example 3. Consider the knowledge base $\mathcal{K}$ from Example 2. We add the unary predicate $\mathsf{purple}(\mathsf{x})$ to PRED . 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 $\mathcal{T}' = \{t_1, t_2\}$. Since $\mathcal{K}$ does not contain classical formulas, there are no EHIs of infinite rank. All the EHIs satisfying $\mathcal{K}$ will be assigned rank 0. That is, all the EHIs in which if t_i is an elephant and t_j is a keeper ($i, j \in \{1, 2\}$), t_i likes t_j but, if fred is a keeper, t_i 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. 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 $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{y}) \rightsquigarrow \mathsf{likes}(\mathsf{x}, \mathsf{y})$ ($\mathcal{E}_2$ is not considered in $rk_{\mathcal{K}}$, since it uses the constant t_3).

It then follows that a desirable form of constrained monotonicity, formalised by (RM), holds. Note that all the EHIs at rank 0 in the minimal model $rk_{\mathcal{K}}$ would either satisfy $\mathsf{purple}(t_i)$ ($i \in \{1, 2\}$) or not, since it is irrelevant to the satisfaction of $\mathcal{K}$. The outcome would be that, while of course satisfying the formula $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{fred}) \rightsquigarrow \neg\mathsf{likes}(\mathsf{x}, \mathsf{fred})$, since it is in $\mathcal{K}$, $rk_{\mathcal{K}}$ would not satisfy $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{fred}) \rightsquigarrow \neg\mathsf{purple}(x)$, while it would satisfy $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{purple}(x) \wedge \mathsf{keeper}(\mathsf{fred}) \rightsquigarrow \neg\mathsf{likes}(\mathsf{x}, \mathsf{fred})$.

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

(INCL) $\alpha \in \mathcal{K}$ implies $\mathcal{K} \mathrel{|\!\approx_{rc}} \alpha$

(SMP) $\mathcal{S} = \{\alpha : \mathcal{K} \mathrel{|\!\approx_{rc}} \alpha\}$ is rational

It is straightforward that our definition of $\mathrel{|\!\approx_{rc}}$ carries over to these properties:

Theorem 4. *The entailment relation* $\mathrel{|\!\approx_{rc}}$ *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 4). 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 propositional valuations in a ranked interpretation to be as typical as possible.

(SMP) requires the set of statements corresponding to the Rational Closure of knowledge base $\mathcal{K}$ to be rational (in the sense of Definition 3). By virtue of Theorem 3, this requires defeasible entailment to be characterised by a *single* ranked interpretation. This accounts for the fact that the property is also referred as the Single Model Property.

6 Comparison

Given that KLM-style defeasible reasoning started off as a propositional endeavour, it makes sense to begin this section with a formal comparison to the propositional case. Note firstly that, when restricted to 0-ary predicates, the language of RFOL reduces to a propositional one. In this case the Herbrand universe becomes superfluous, the Herbrand base is the set of 0-ary predicates (propositional atoms), and a Herbrand interpretation is a subset of propositional atoms. Clearly then, Herbrand interpretations reduce to propositional valuations. For DRFOL we work with enriched Herbrand interpretations in which typicality objects are added to the Herbrand universe. But since the Herbrand universe plays no role in the semantics of 0-ary predicates, it is redundant. The ranked interpretations for DRFOL (Definition 5) then reduce to propositional ranked interpretations (described in Section 3, from which it follows that defeasible implication in DRFOL reduces to propositional defeasible implication (represented by the symbol $\mathrel{|\!\sim}$ in Section 3). More specifically, consider a defeasible propositional language generated from a set of atoms, and take this set to be the 0-ary predicates of a DRFOL language. It follows that for every propositional ranked interpretation $\mathscr{R}$ there is a DRFOL ranked interpretation rk such that for all propositional statements α, β constructed from $\neg, \wedge$ and $\vee$, $rk \Vdash \alpha \rightsquigarrow \beta$ iff $\mathscr{R} \Vdash \alpha \mathrel{|\!\sim} \beta$ and $rk \Vdash \alpha$ iff $\mathscr{R} \Vdash \alpha$. Conversely, for every DRFOL ranked interpretation rk, there is a propositional ranked interpretation $\mathscr{R}$ such that for all propositional statements α, β constructed from $\neg, \wedge$ and $\vee$, $rk \Vdash \alpha \rightsquigarrow \beta$ iff $\mathscr{R} \Vdash \alpha \mathrel{|\!\sim} \beta$ and $rk \Vdash \alpha$ iff $\mathscr{R} \Vdash \alpha$.

A similar result holds when DRFOL is restricted to compounds, implications, and defeasible implications that are all ground. Considering RFOL first, observe that, unlike the case discussed above, the Herbrand universe is used to construct the Herbrand base here, and it is therefore used in the definition of Herbrand interpretations. But since we only consider ground statements, each ground atom in a Herbrand interpretation effectively functions like a propositional atom, which again means that Herbrand interpretations reduce to

propositional valuations (for the propositional language with the ground atoms as its set of propositional atoms). Moving on to DRFOL we note that since we are restricted to ground statements, the substitutions referred to in Definition 6 do not play any role, which means that the typicality objects in enriched Herbrand interpretations are redundant. In summary, consider a defeasible propositional language generated from the ground atoms of a language of DRFOL. It follows that for every propositional ranked interpretation $\mathscr{R}$ there is a DRFOL ranked interpretation *rk* such that for all propositional statements α, β constructed from $\neg, \wedge$ and $\vee$, $rk \Vdash \alpha \rightsquigarrow \beta$ iff $\mathscr{R} \Vdash \alpha \mathrel{|\!\sim} \beta$ and $rk \Vdash \alpha$ iff $\mathscr{R} \Vdash \alpha$. And conversely, for every DRFOL ranked interpretation *rk*, there is a propositional ranked interpretation $\mathscr{R}$ such that for all propositional statements α, β constructed from $\neg, \wedge$ and $\vee$, $rk \Vdash \alpha \rightsquigarrow \beta$ iff $\mathscr{R} \Vdash \alpha \mathrel{|\!\sim} \beta$ and $rk \Vdash \alpha$ iff $\mathscr{R} \Vdash \alpha$.

Space considerations prevent us from providing a detailed comparison of DRFOL with $\mathcal{DALC}$, the defeasible version of the DL $\mathcal{ALC}$ (Britz et al. 2020). Suffice it to note that 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 cases above, with one exception: a defeasible implication of the form $C(x) \rightsquigarrow \bot$ has a meaning that is different than $C \mathrel{\tilde{\sqsubseteq}} \bot$, 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 $\neg\alpha \mathrel{|\!\sim} \bot$ ($\mathscr{R} \Vdash \alpha$ iff $\mathscr{R} \Vdash \neg\alpha \mathrel{|\!\sim} \bot$ for all $\mathscr{R}$) while, for $\mathcal{DALC}$, $C \sqsubseteq \bot$ is equivalent to $C \mathrel{\tilde{\sqsubseteq}} \bot$. But in DRFOL, defeasible implications *cannot* inform us about compounds or classical implications. Formally, rational satisfaction sets do *not* necessarily satisfy the following postulate:

$$\text{(SUB)}\ \frac{A(\vec{x}) \rightsquigarrow \bot \in \mathcal{S}}{A(\vec{x}) \rightarrow \bot \in \mathcal{S}}$$

One way this difference manifests itself is in the way our framework handles the finitary Lottery Paradox (Poole 1991). Consider the DRFOL knowledge base $\mathcal{K} = \{\mathsf{penguin(x)} \rightarrow \mathsf{bird(x)}, \mathsf{cuckoo(x)} \rightarrow \mathsf{bird(x)}, \mathsf{bird(x)} \rightarrow \mathsf{cuckoo(x)} \vee \mathsf{penguin(x)}, \mathsf{bird(x)} \rightsquigarrow \mathsf{flies(x)} \wedge \mathsf{nests(x)}, \mathsf{cuckoo(x)} \rightarrow \neg\mathsf{nests(x)}, \mathsf{penguin(x)} \rightarrow \neg\mathsf{flies(x)}\}$. This can also be modelled as a propositional defeasible knowledge base and as a $\mathcal{DALC}$ knowledge base.

In all three cases KLM rationality dictates that being a bird defeasibly implies a contradiction: $\mathsf{bird(x)} \rightsquigarrow \bot$ in the case of DRFOL, $\mathsf{bird} \mathrel{|\!\sim} \bot$ in the propositional defeasible case, and $\mathsf{Bird} \mathrel{\tilde{\sqsubseteq}} \bot$ in the case of $\mathcal{DALC}$. In the defeasible propositional case this means there are no birds ($\mathsf{bird} \mathrel{|\!\sim} \bot$ is equivalent to $\neg\mathsf{bird}$). Similarly for $\mathcal{DALC}$, where $\mathsf{Bird} \mathrel{\tilde{\sqsubseteq}} \bot$ is equivalent to $\mathsf{Bird} \sqsubseteq \bot$. In DRFOL, however, $\mathsf{bird(x)} \rightsquigarrow \bot$ is *not* equivalent to $\mathsf{bird(x)} \rightarrow \bot$. Rather than stating that there are no birds, $\mathsf{bird(x)} \rightsquigarrow \bot$ means that there are no *typical* birds. This leaves open the possibility of there being only atypical birds, something that is not possible in the propositional and DL cases.

Example 4. Let $\text{CONST} = \{\mathsf{tweety}\}$, $\text{VAR} = \{x\}$, $\text{PRED} = \{\mathsf{bird}, \mathsf{penguin}, \mathsf{cuckoo}, \mathsf{flies}, \mathsf{nests}\}$, with $\mathcal{T} = \{\mathsf{t}_1, \ldots\}$ the set of typicality objects. Let *rk* be the ranked interpretation for which $rk(\mathcal{E}) = 0$ and $rk(\mathcal{E}') = \infty$ for all other EHIs, where $\mathcal{E} = \{\mathsf{bird(tweety)}, \mathsf{penguin(tweety)}\}$. It is easily verified that *rk* satisfies all statements in the DRFOL knowledge base $\mathcal{K}$ above, and also satisfies $\mathsf{bird(x)} \rightsquigarrow \bot$, since $rk \not\Vdash \mathsf{bird}(\mathsf{t}_i)$ for all i. But *rk* does not satisfy $\mathsf{bird(x)} \rightarrow \bot$.

We regard this as a significant advantage of DRFOL over previous KLM-style defeasible formalisms.

As a final remark, observe that this distinction is not in conflict with the claim that DRFOL is a proper generalisation of propositional defeasible logic. 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.

7 Related Work

Defeasible reasoning is part of a broader research programme on conditional reasoning (Arlo-Costa 2019), most of which was developed for propositional logic. This paper falls in the class of approaches aimed at moving beyond propositional expressivity. We pointed out the connection with defeasible DLs (Casini and Straccia 2010, 2013; Casini et al. 2015; Giordano et al. 2013, 2015; Bonatti et al. 2015; Bonatti 2019; Pensel and Turhan 2018) in Section 6, but there have also been proposals to extend this approach to first-order logic. Most of these define a preferential order over the elements of the first-order domain (Schlechta 1995; Brafman 1997; Delgrande and Rantsoudis 2020), in line with some of the DL proposals (Giordano et al. 2015; Britz et al. 2020), and present rationality postulates, but they do not provide characterisations in terms of rationality postulates. Others (Delgrande 1998; Kern-Isberner and Thimm 2012) are formally closer to our work in that they use preference orders over interpretations.

Delgrande (1998) proposes a semantics that is closer to the intuitions behind *circumscription* (McCarthy 1980), giving preference to interpretations that minimise the counterexamples to defeasible conditionals. On the other hand, Kern-Isberner and Thimm (2012) propose a technical solution that is 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 conditionals. 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 (Kern-Isberner 2001, 2004). In contrast, our focus in this paper is on getting the theoretical foundations of defeasible reasoning for restricted first-order logics 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 in Section 5, 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 is also necessary. We leave that for future work.

Kern-Isberner and Beierle (2015); Beierle et al. (2016, 2017) use the same semantic approach of Kern-Isberner and Thimm (2012) to develop an extension of Pearl's System Z (1990) 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 in Section 5.

Brafman (1997) suggests that preference orders over the domain should result in forms of reasoning quite different from the use of preference orders over interpretations, comparable to the difference between statistical and subjective readings of probabilities. We leave a proper investigation of the differences between these two different modelling solutions as future work.

As mentioned, the final goal of our investigation is the development of a defeasible extension of Datalog+/-. To the best of out knowledge there is no research on the introduction of defeasible implication in Datalog+/-. Of course, there is a longstanding tradition of non-monotonic extensions of Disjunctive Datalog with an Answer Set semantics (Leone et al. 2006). Although there are some connections between conditional reasoning (of which defeasible reasoning is a special case) and negation-as-failure (Makinson 1994, 2005), these two approaches are different. Answer Set Programming is a popular solution to model the closed-world assumption, while conditional reasoning is focused on reasoning with the potential conflicts resulting from defeasible pieces of information.

8 Conclusion and Future Work

In this paper we have laid the theoretical groundwork for KLM-style defeasible Datalog (DRFOL). Our primary contribution is a set of rationality postulates describing the behaviour of defeasibility in DRFOL, a typicality semantics for interpreting defeasibility in DRFOL, 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.

There are at least three important avenues for future research. The first one relates to a more detailed investigation of defeasible entailment for DRFOL knowledge bases. While Rational Closure for DRFOL is on par with the analogous notions for propositional logic and DLs (restricted to Tboxes), it is not able to fully manage reasoning about individuals. Going back to Example 3, assume that we add a constant bob to CONST. Since we are not informed of anything atypical about bob, we would like to be able to infer the statement $\mathsf{elephant}(\mathsf{bob}) \wedge \mathsf{keeper}(\mathsf{fred}) \sim\!\!\!\rightarrow \neg\mathsf{likes}(\mathsf{bob}, \mathsf{fred})$. But Rational Closure does not sanction this, since the formula $\mathsf{elephant}(\mathsf{x}) \wedge \mathsf{keeper}(\mathsf{fred}) \sim\!\!\!\rightarrow \neg\mathsf{likes}(\mathsf{x}, \mathsf{fred})$ is evaluated only on the typicality constants, and whether bob behaves in a typical way or not is irrelevant w.r.t. the satisfaction of the knowledge base. Consequently, on rank 0 of $rk_{\mathcal{K}}$ there are EHIs in which $\mathsf{elephant}(\mathsf{bob})$ behaves like an atypical elephant. Rational Closure would therefore need to be refined to model the inferences about individuals properly.

Next we discuss a more general point about defeasible entailment. Based on the theoretical basics we have put in place and the preliminary work on Rational Closure for DRFOL, we conjecture that all appropriate forms of DRFOL defeasible entailment will satisfy the (SMP) property, thereby ensuring that all forms of defeasible entailment are rational. This will be similar to the propositional case (Lehmann 1995; Booth and Paris 1998; Giordano et al. 2015), and unlike the case for DLs (Casini and Straccia 2010; Casini et al. 2013).

With a suitable definition (or definitions) of DRFOL defeasible entailment in place, the next step is to investigate 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 (Casini, Straccia, and Meyer 2019; Giordano et al. 2015; Casini, Meyer, and Varzinczak 2019).

Finally, in line with our stated aim in Section 1, the basic theoretical framework presented in this paper places us in a position to see whether the work on DRFOL can be extended to Datalog +/-.

References

Abiteboul, S.; Hull, R.; and Vianu, V. 1995. *Foundations of Databases*. Addison-Wesley.

Arlo-Costa, H. 2019. The Logic of Conditionals. In *The Stanford Encyclopedia of Philosophy*. Summer 2019 edition.

Beierle, C.; Falke, T.; Kutsch, S.; and Kern-Isberner, G. 2016. Minimal Tolerance Pairs for System Z-Like Ranking Functions for First-Order Conditional Knowledge Bases. In *Proc. of FLAIRS 2016*, 626–631. AAAI Press.

Beierle, C.; Falke, T.; Kutsch, S.; and Kern-Isberner, G. 2017. 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.

Benferhat, S.; Cayrol, C.; Dubois, D.; Lang, J.; and Prade, H. 1993. Inconsistency Management and Prioritized Syntax-based Entailment. In *Proc. of IJCAI-93*, 640–645. Morgan Kaufmann Publishers Inc.

Bonatti, P. A. 2019. Rational closure for all description logics. *Artificial Intelligence* 274: 197–223.

Bonatti, P. A.; Faella, M.; Petrova, I. M.; and Sauro, L. 2015. A new semantics for overriding in description logics. *Artificial Intelligence* 222: 1–48.

Booth, R.; and Paris, J. B. 1998. A Note on the Rational Closure of Knowledge Bases with Both Positive and Negative Knowledge. *J. Log. Lang. Inf.* 7(2): 165–190.

Brafman, R. I. 1997. A First-order Conditional Logic with Qualitative Statistical Semantics. *J. Log. Comput.* 7(6): 777–803.

Britz, K.; Casini, G.; Meyer, T.; Moodley, K.; Sattler, U.; and Varzinczak, I. 2020. Principles of KLM-Style Defeasible Description Logics. *ACM T. Comput. Log.* 22(1).

Calì, A.; Gottlob, G.; and Lukasiewicz, T. 2012. A general Datalog-based framework for tractable query answering over ontologies. *Journal of Web Semantics* 14(C): 57–83.

Calì, A.; Gottlob, G.; Lukasiewicz, T.; Marnette, B.; and Pieris, A. 2010. Datalog+/-: A Family of Logical Knowledge Representation and Query Languages for New Applications. In *Proc. of LICS 2010*, 228–242. IEEE.

Casini, G.; Meyer, T.; Moodley, K.; and Nortjé, R. 2014. Relevant Closure: A New Form of Defeasible Reasoning for Description Logics. In *Proc. of JELIA 2014*, volume 8761 of *LNCS*, 92–106. Springer.

Casini, G.; Meyer, T.; Moodley, K.; Sattler, U.; and Varzinczak, I. 2015. Introducing Defeasibility into OWL Ontologies. In *Proc. of ISWC 2015*, volume 9367 of *LNCS*, 409–426. Springer.

Casini, G.; Meyer, T.; Moodley, K.; and Varzinczak, I. 2013. Nonmonotonic reasoning in Description Logics: Rational Closure for the ABox. In *Proceedings of DL-13*, 600–615. CEUR-WS.org.

Casini, G.; Meyer, T.; and Varzinczak, I. 2019. Taking Defeasible Entailment Beyond Rational Closure. In *Proc. of JELIA 2019*, volume 11468 of *LNCS*, 182–197. Springer.

Casini, G.; and Straccia, U. 2010. Rational Closure for Defeasible Description Logics. In *Proc. of JELIA 2010*, volume 6341 of *LNCS*, 77–90. Springer-Verlag.

Casini, G.; and Straccia, U. 2013. Defeasible Inheritance-Based Description Logics. *Journal of Artificial Intelligence Research* 48: 415–473.

Casini, G.; Straccia, U.; and Meyer, T. 2019. A Polynomial Time Subsumption Algorithm for Nominal Safe $\mathcal{ELO}_\bot$ under Rational Closure. *Information Sciences* 501: 588–620.

Delgrande, J. P. 1998. On first-order conditional logics. *Artificial Intelligence* 105(1): 105 – 137.

Delgrande, J. P.; and Rantsoudis, C. 2020. A Preference-Based Approach for Representing Defaults in First-Order Logic. In *Proc. of NMR 2020*, 120–129.

Giordano, L.; and Gliozzi, V. 2019. Strengthening the Rational Closure for Description Logics: An Overview. In *Proc. of CILC 2019*, 68–81. CEUR-WS.org.

Giordano, L.; Gliozzi, V.; Olivetti, N.; and Pozzato, G. L. 2013. A non-monotonic Description Logic for reasoning about typicality. *Artificial Intelligence* 195: 165–202.

Giordano, L.; Gliozzi, V.; Olivetti, N.; and Pozzato, G. L. 2015. Semantic characterization of rational closure: From propositional logic to description logics. *Art. Int.* 226: 1–33.

Kern-Isberner, G. 2001. *Conditionals in Nonmonotonic Reasoning and Belief Revision - Considering Conditionals as Agents*, volume 2087 of *LNCS*. Springer.

Kern-Isberner, G. 2004. A Thorough Axiomatization of a Principle of Conditional Preservation in Belief Revision. *Ann. Math. Artif. Intell.* 40(1-2): 127–164.

Kern-Isberner, G.; and Beierle, C. 2015. A System Z-like Approach for First-Order Default Reasoning. In *Advances in Knowledge Representation, Logic Programming, and Abstract Argumentation*, volume 9060 of *LNCS*, 81–95. Springer.

Kern-Isberner, G.; and Thimm, M. 2012. A Ranking Semantics for First-Order Conditionals. In *Proc. of ECAI 2012*, 456–461. IOS Press.

Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44: 167–207.

Lehmann, D. 1995. Another perspective on default reasoning. *Ann. Math. Artif. Intell.* 15(1): 61–82.

Lehmann, D.; and Magidor, M. 1992. What does a conditional knowledge base entail? *Art. Intell.* 55: 1–60.

Leone, N.; Pfeifer, G.; Faber, W.; Eiter, T.; Gottlob, G.; Perri, S.; and Scarcello, F. 2006. The DLV System for Knowledge Representation and Reasoning. *ACM T. Comput. Log.* 7(3): 499–562.

Makinson, D. 1994. General Patterns in Nonmonotonic Reasoning. In *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. III*, 35–110. Clarendon Press.

Makinson, D. 2005. *Bridges from Classical to Nonmonotonic Logic*. King's College Publications.

McCarthy, J. 1980. Circumscription, a form of nonmonotonic reasoning. *Art. Intell.* 13(1-2): 27–39.

Pearl, J. 1990. System Z: a natural ordering of defaults with tractable applications to nonmonotonic reasoning. In *Proc. of TARK 1990*.

Pensel, M.; and Turhan, A.-Y. 2018. Reasoning in the Defeasible Description Logic $\mathcal{EL}_\bot$ - computing standard inferences under rational and relevant semantics. *Int. J. Approx. Reason.* 103: 28 – 70.

Poole, D. 1991. The Effect of Knowledge on Belief: Conditioning, Specificity and the Lottery Paradox in Default Reasoning. *Art. Intell.* 49(1-3): 281–307.

Schlechta, K. 1995. Defaults as Generalized Quantifiers. *Journal of Logic and Computation* 5(4): 473–494.