ECAI 2012
Luc De Raedt et al. (Eds.)

doi:10.3233/978-1-61499-098-7-456

A Ranking Semantics for First-Order Conditionals

Gabriele Kern-Isberner[1] and **Matthias Thimm**[2]

Abstract. Usually, default rules in the form of conditional statements are built on propositional logic, representing classes of individuals by propositional variables, as in "Birds fly, but penguins don't". Only few approaches have addressed the problem of giving formal semantics to first-order conditionals that allow (non-monotonic) inferences both for classes and for individuals. In this paper, we present a semantics for first-order conditionals that is based on ordinal conditional (or ranking) functions which are well-known in the area of propositional default reasoning and makes use of representative individuals to establish conditional relationships. We generalize the c-representation approach of [8] for inductive reasoning with first-order conditionals, and evaluate our approach via benchmark examples and a catalogue of general properties.

Introduction

Default reasoning aims at implementing uncertain reasoning in a qualitative way. Frameworks and postulates have been set up to elaborate what a "reasonable" default logic should conclude, and different seminal approaches have been put forward, see, e. g., [15, 9]. In many approaches, conditionals are used as representations of rules that hold in general but may have exceptions, see e. g. [2]. A conditional c has the form $(B \,|\, A)$ and represents the (defeasible) rule "If A then (usually, probably) B". Usually, default logics are propositional, only few advances have been made to build on first-order logics, or fragments thereof. In [2], Delgrande presents a first-order conditional logic that allows for the representation of information both on classes and on individuals in the same framework. An example that illustrates nicely the problem under consideration is the following one (taken from [2]).

Example 1. We consider *elephants* x, $E(x)$, and their *keepers* y, $K(y)$, and let $L(x, y)$ denote that x likes y. The following (open) first-order conditionals represent knowledge about the relationships between elephants and keepers in a fictitious zoo.

$$r_1: \quad (L(x,y) \,|\, E(x) \wedge K(y))$$
$$r_2: \quad (\neg L(x, Fred) \,|\, E(x) \wedge K(Fred))$$
$$r_3: \quad (L(Clyde, Fred) \,|\, E(Clyde) \wedge K(Fred))$$

From the point of view of common sense, this knowledge base makes perfect sense: r_1 expresses that usually, elephants like their keeper. However, keeper *Fred* and elephant *Clyde* are exceptional—usually, elephants do not like *Fred*, but *Clyde* likes (even) *Fred*. Maybe *Clyde* is a particularly good-natured elephant, maybe he is as moody as *Fred* and likes only him. So, *Clyde* is definitely exceptional with respect to r_2, but maybe even with respect to r_1.

The elephant-keeper-example makes the ambiguity inherent to such first-order statements obvious, indeed, their formal interpretation is intricate. Conditional r_1 seems to express knowledge on classes (maybe on populations), while r_3 clearly expresses subjective belief: Considering all situations (possible worlds) involving *Clyde* and *Fred* which are imaginable, we expect worlds in which *Clyde* likes *Fred* to be more plausible. So, we might think of applying different techniques to r_1 and r_3, but r_2 obviously mixes the two types of knowledge, how should r_2 be handled? Moreover, r_1 and r_2 are open conditional statements, we might be tempted to use universal quantification here, but then conflicts between the conditionals will arise concerning *Clyde* and *Fred*. In [2], problems of this kind are solved via a modal approach, making use of actual subuniverses of a common universe and by interpreting conditionals in a preferential way.

In the present paper, we propose an approach to reason with first-order conditional knowledge bases that is thoroughly based on ranking functions which are also quite popular to provide semantics to propositional default logics [6]. The knowledge bases may contain classical and conditional formulas, grounded or not, with a coherent interpretation via ranking functions. As a main contribution of this paper, we develop a novel semantics that generalizes the propositional interpretation of conditionals and makes open conditionals take the role of (first-order) default rules expressing uncertain statements about the usual behaviour of individuals in a population. Furthermore, we make use of *representatives* as most convincing instances to establish such open conditional relationships. In order to define an inductive model-based reasoning mechanism, we generalize the concept of so-called *c-representations* using the algebraic formalism of *conditional structures* for the first-order framework [7, 8]. Inference can then be based on considering all c-representations of a knowledge base, or on selecting a particular c-representation as a "best" model of the knowledge base. We identify several key properties of our approach and compare it to related work. We also apply our approach to benchmark examples adapted from propositional default reasoning and show that it behaves well.

The rest of this paper is organized as follows. We continue with introducing the syntax of our first-order conditional logic and propose a novel ranking semantics for first-order conditionals afterwards. Then we generalize the concept of c-representations to our new semantics and illustrate its use on several examples. Afterwards we investigate the properties of the approach and relate it to other works. Finally, we conclude with a brief summary and discussion of future work. Proofs of technical results have been omitted due to space restrictions.

Syntax of first-order conditionals

Let Σ be a first-order signature consisting of a finite set of predicates P_Σ and a finite set of constant symbols U_Σ but without function symbols of arity > 0. An *atom* is a predicate of arity n together with a list of n constants and/or variables. A *literal* is an atom or

[1] Department of Computer Science, TU Dortmund, Germany
[2] Institute for Web Science and Technologies, Universität Koblenz, Germany

a negated atom. Formulas are built on atoms using conjunction ($\wedge$), disjunction ($\vee$), negation ($\neg$), and quantification ($\forall$, $\exists$). We abbreviate conjunctions by juxtaposition and negations usually by overlining, e. g. AB means $A \wedge B$ and $\overline{A}$ means $\neg A$. A *ground* formula contains no variables. In a *closed* formula, all variables (if they occur) are bound by quantifiers, otherwise, the formula is *open*, and the variables that occur outside of the range of quantifiers are called *free*. If a formula A contains free variables we also use the notation $A(\vec{x})$ where $\vec{x} = (x_1, \ldots, x_n)$ contains all free variables in A. If $\vec{c}$ is a vector of the same length as $\vec{x}$ then $A(\vec{c})$ is meant to denote the instantiation of A with $\vec{c}$. A formula $\forall \vec{x} A(\vec{x})$ ($\exists \vec{x} A(\vec{x})$) is *universal* (*existential*) if A involves no further quantification. Let $\mathcal{L}_\Sigma$ be the first-order language that allows no nested quantification, i.e., all quantified formulas are either universal or existential formulas.

$\mathcal{L}_\Sigma$ is extended by a conditional operator "$|$" to a conditional language $(\mathcal{L}_\Sigma \,|\, \mathcal{L}_\Sigma)$ containing first-order conditionals $(B \,|\, A)$ with $A, B \in \mathcal{L}_\Sigma$, and (universally or existentially) quantified conditionals $\forall \vec{x}(B \,|\, A)$, $\exists \vec{x}(B \,|\, A)$[3]. When writing $(B(\vec{x}) \,|\, A(\vec{x}))$, we assume $\vec{x}$ to contain all free variables occurring in either A or B. Conditionals cannot be nested. When the signature is clear from context, we will also omit the subscript Σ. To exemplify the syntax, consider the rules r_1, r_2, r_3 from Example 1 in the introduction with r_1, r_2 being open first-order conditionals.

A *first-order conditional knowledge base* $\mathcal{R}$ is a (finite) set of (conditional) formulas from $\mathcal{L}_\Sigma \cup (\mathcal{L}_\Sigma \,|\, \mathcal{L}_\Sigma)$ with the restriction that no existential (outer) quantification of conditionals may occur. A *first-order knowledge base* $\mathcal{KB} = \langle \mathcal{F}, \mathcal{R} \rangle$ consists of a first-order conditional knowledge base $\mathcal{R}$, together with a set $\mathcal{F}$ of closed formulas from $\mathcal{L}_\Sigma$, called *facts*.

OCF-based Semantics

For interpretation of first-order aspects we use the Herbrand semantics. The *Herbrand base* $\mathcal{H}^\Sigma$ of a first-order signature Σ is the set of all ground atoms of Σ. A *possible world* ω is any subset of $\mathcal{H}^\Sigma$. A possible world can be concisely represented as a *complete conjunction* or *minterm*, i. e. a conjunction of literals where every atom of $\mathcal{H}^\Sigma$ appears either in positive or in negated form. The set of all possible worlds of Σ is denoted by Ω_Σ. Let $\models$ denote the classical satisfaction relation between possible worlds and formulas from $\mathcal{L}_\Sigma$.

For an open conditional $(B(\vec{x}) \,|\, A(\vec{x})) \in (\mathcal{L}_\Sigma \,|\, \mathcal{L}_\Sigma)$ let $\mathcal{H}^{(B(\vec{x}) \,|\, A(\vec{x}))}$ denote the set of all constant vectors $\vec{a}$ used for proper groundings of $(B(\vec{x}) \,|\, A(\vec{x}))$ from the Herbrand universe $\mathcal{H}^\Sigma$, i. e. $\mathcal{H}^{(B(\vec{x}) \,|\, A(\vec{x}))} = U_\Sigma^{|\vec{x}|}$ where $|\vec{x}|$ is the length of $\vec{x}$.

Just as in the propositional case, the set Ω_Σ of possible worlds can be ranked by an *ordinal conditional function* (*OCF*, also called *ranking function*) that assigns degrees of (im)plausibility resp. disbelief to possible worlds and statements [16].

Definition 1. An *ordinal conditional function (OCF)* κ *on* Ω_Σ is a function $\kappa : \Omega_\Sigma \rightarrow \mathbb{N} \cup \{\infty\}$ with $\kappa^{-1}(0) \neq \emptyset$.

We can now make use of the possible world semantics to assign degrees of disbelief also to formulas and (non-quantified) conditionals. In the following, let $A, B \in \mathcal{L}_\Sigma$ denote closed formulas, and let $A(\vec{x}), B(\vec{x}) \in \mathcal{L}_\Sigma$ denote open formulas.

Definition 2. Let κ be an OCF. The κ-ranks of closed formulas are defined (as in the propositional case [16]) via

$$\kappa(A) = \min_{\omega \models A} \kappa(\omega) \quad \text{and} \quad \kappa(B \,|\, A) = \kappa(AB) - \kappa(A).$$

Furthermore, we define the κ-ranks for open formulas and open, non-quantified conditionals by evaluating most plausible instances:

$$\kappa(A(\vec{x})) = \min_{\vec{a} \in \mathcal{H}^{A(\vec{x})}} \kappa(A(\vec{a}))$$
$$\kappa(B(\vec{x}) \,|\, A(\vec{x})) = \min_{\vec{a} \in \mathcal{H}^{(B(\vec{x}) \,|\, A(\vec{x}))}} \kappa(A(\vec{a})B(\vec{a})) - \kappa(A(\vec{a})).$$

The ranks of first-order formulas are coherently based on the usage of OCFs for propositional formulas. Just as in the propositional case, these degrees of beliefs are used to specify when a formula from $(\mathcal{L}_\Sigma \,|\, \mathcal{L}_\Sigma)$ is accepted, i. e. deemed highly plausible, by a ranking function κ (where acceptance is denoted by $\models$). We will first consider the acceptance of closed (conditional) formulas.

Definition 3. Let κ be an OCF. The acceptance relation between κ and formulas from $\mathcal{L}_\Sigma$ and $(\mathcal{L}_\Sigma \,|\, \mathcal{L}_\Sigma)$ is defined as follows:

- *for closed formulas:*
  - $\kappa \models A$ iff for all $\omega \in \Omega$ with $\kappa(\omega) = 0$, it holds that $\omega \models A$.
  - $\kappa \models (B \,|\, A)$ iff $\kappa(AB) < \kappa(A\overline{B})$.
- *for universal/existential conditionals:*
  - $\kappa \models \forall \vec{x}(B(\vec{x}) \,|\, A(\vec{x}))$ iff $\kappa \models (B(\vec{a}) \,|\, A(\vec{a}))$ for all $\vec{a} \in \mathcal{H}^{(B(\vec{x}) \,|\, A(\vec{x}))}$.
  - $\kappa \models \exists \vec{x}(B(\vec{x}) \,|\, A(\vec{x}))$ iff there is $\vec{a} \in \mathcal{H}^{(B(\vec{x}) \,|\, A(\vec{x}))}$ such that $\kappa \models (B(\vec{a}) \,|\, A(\vec{a}))$.

Acceptance of a sentence by a ranking function is the same as in the propositional case for ground sentences, and interprets the classical quantifiers in a straightforward way. Note that no classical relations hold between universal and existential formulas, as acceptance by ranking functions is three-valued.

The treatment of acceptance of open formulas is more intricate, as such formulas will be used to express default statements, like in "A is plausible", or in "usually, if A holds, then B also holds". The basic idea here is that such (conditional) open statements hold if there are individuals that provide most convincing instances of the respective conditional. These so-called *representatives* should, of course, allow for the acceptance of the instantiated conditional (as in Definition 3) while most plausibly verifying the conditional (i. e. satisfying A and B). Moreover, representatives are expected to be least exceptional with respect to falsifying the conditional. The following definition makes use of the κ-ranks of Definition 2 to formalize this precisely.

Definition 4. Let $r = (B(\vec{x}) \,|\, A(\vec{x})) \in (\mathcal{L}_\Sigma \,|\, \mathcal{L}_\Sigma)$ be a non-quantified conditional involving open formulas from $\mathcal{L}_\Sigma$. We say that $\vec{a} \in \mathcal{H}^{(B(\vec{x}) \,|\, A(\vec{x}))}$ is a *weak representative* of r iff it satisfies the following conditions:

$$\kappa(A(\vec{a})B(\vec{a})) = \kappa(A(\vec{x})B(\vec{x})) \tag{1}$$
$$\kappa(A(\vec{a})B(\vec{a})) < \kappa(A(\vec{a})\overline{B}(\vec{a})) \tag{2}$$

The set of *weak representatives of* r is denoted by $WRep(r)$. We say that $\vec{a} \in \mathcal{H}^{(B(\vec{x}) \,|\, A(\vec{x}))}$ is a *(strong) representative* of r iff it is a weak representative of r and

$$\kappa(A(\vec{a})\overline{B}(\vec{a})) = \min_{\vec{b} \in WRep(r)} \kappa(A(\vec{b})\overline{B}(\vec{b})). \tag{3}$$

The set of all *representatives of* r is denoted by $Rep(r)$.

(Weak) Representatives of a conditional are characterized by being most general and least exceptional. This is expressed by (1) that postulates that representatives are most normal with respect to A's being also B's, and also by (3) that demands that representative individuals should be least specific with respect to violating the link between

[3] These quantifications will often be distinguished as *outer* quantifications in the paper.

A and B; otherwise, this violation might be caused by extraordinary attributes. This can be easily exemplified in the popular penguin scenario. Consider a scenario where we have birds, penguins, and super-penguins. Birds usually fly, whereas penguins are expected not to fly while super-penguins are famous for flying. What is a representative (flying) bird here? It is definitely not a penguin since penguins usually do not fly (violating (2)). While we might more strongly believe that super-penguins fly than to care about the non-specified bird next to us (super-penguins are famous!), super-penguins are too specific to serve as good representatives. Representatives should be general, covering as many species of flying birds as possible. But, due to this generality, we would also be more willing to accept an exception here than for more specific subclasses. Superpenguins might be able to fly because they are equipped with motorized wings, and their failure of flying might be caused by a motor problem, an explanation that certainly does not apply to the failure of flying of a normal bird. This motivates (3). Note that (weak) representatives are only *conditional representatives*, i. e., representatives for the respective conditional relationship, as we do not postulate that representatives (certainly or plausibly) satisfy the premise of the conditional. It might well be the case that individuals may serve as representatives for different conditionals. Now we can base our definition of acceptance of open conditionals on the notion of representatives as follows.

Definition 5. Let κ be an OCF and $r = (B(\vec{x}) \,|\, A(\vec{x}))$ an open (non-quantified) conditional. Then $\kappa \models r$ iff $Rep(r) \neq \emptyset$, and one of the two following (exclusive) conditions is satisfied:

(A) it holds that

$$\kappa(A(\vec{x})B(\vec{x})) < \kappa(A(\vec{x})\overline{B}(\vec{x})); \tag{4}$$

(B) $\kappa(A(\vec{x})B(\vec{x})) = \kappa(A(\vec{x})\overline{B}(\vec{x}))$, and for all $\vec{a}_1 \in Rep((B(\vec{x}) \,|\, A(\vec{x})))$ and for all $\vec{a}_2 \in Rep((\overline{B}(\vec{x}) \,|\, A(\vec{x})))$, it holds that

$$\kappa(A(\vec{a}_1)\overline{B}(\vec{a}_1)) < \kappa(A(\vec{a}_2)B(\vec{a}_2)). \tag{5}$$

The acceptance of an open conditional is based on the existence of a suitable $\vec{a}$ satisfying (2), i. e., on the acceptance of the propositional conditional $(B(\vec{a}) \,|\, A(\vec{a}))$ (note that $Rep((B(\vec{x}) \,|\, A(\vec{x}))) \neq \emptyset$ iff $WRep((B(\vec{x}) \,|\, A(\vec{x}))) \neq \emptyset$). However, conditions (1) and (2) alone are too weak to justify the acceptance of $(B(\vec{x}) \,|\, A(\vec{x}))$ since it might well be the case that there are $\vec{a}$ and $\vec{b}$ fulfilling (1) and (2) for $(B(\vec{x}) \,|\, A(\vec{x}))$ and $(\overline{B}(\vec{x}) \,|\, A(\vec{x}))$, respectively. This means that κ might accept both $(B(\vec{x}) \,|\, A(\vec{x}))$ and $(\overline{B}(\vec{x}) \,|\, A(\vec{x}))$, which would be counterintuitive. Hence, we need to make acceptance unambiguous by giving preference to one of the two conditionals. This can be done either by postulating (4) or (5). Condition (4) looks like a natural prerequisite for the acceptance of $(B(\vec{x}) \,|\, A(\vec{x}))$. However, in the birds scenario with penguins and super-penguins, equalities like $\kappa(A(\vec{x})B(\vec{x})) = \kappa(A(\vec{x})\overline{B}(\vec{x}))$ quite naturally arise since penguins are as normal non-flying birds as doves are normal flying birds (see Example 3 below). In this case, (5) again uses the idea of least exceptionality for specifying proper representatives; it makes $(B(\vec{x}) \,|\, A(\vec{x}))$ acceptable, as opposed to $(\overline{B}(\vec{x}) \,|\, A(\vec{x}))$, if the representatives of the first conditional less exceptionally violate the respective conditional than the representatives of the latter conditional. Note that Definition 5 extends the definition of acceptance in the propositional case, i. e., $\kappa \models (B(\vec{a}) \,|\, A(\vec{a}))$ iff $\kappa(A(\vec{a})B(\vec{a})) < \kappa(A(\vec{a})\overline{B}(\vec{a}))$.

Definitions 4 and 5 can be used to define acceptance of open non-conditional formulas $A(\vec{x})$ by considering them as conditionals with tautological antecedents, i.e., as $(A(\vec{x}) \,|\, \top)$. However, it is crucial to remark here that $(A(\vec{x}) \,|\, \top)$ mandatorily demands for a default reading like "being A is plausible", as opposed to "A certainly holds". This distinction is made in our approach by distinguishing between certain knowledge $\mathcal{F}$ (all elements here are closed formulas of $\mathcal{L}_\Sigma$) and default (conditional) beliefs in $\mathcal{R}$ which may involve both closed and open formulas (well-formed according to our syntax definitions). Formally, this is handled by giving different semantics to the two parts of our knowledge bases.

Definition 6. Let $\mathcal{KB} = \langle \mathcal{F}, \mathcal{R} \rangle$ be a first-order knowledge base, and let κ be an OCF.

1. κ *accepts* $\mathcal{R}$, denoted by $\kappa \models \mathcal{R}$, iff $\kappa \models \varphi$ for all $\varphi \in \mathcal{R}$.
2. κ *accepts* $\mathcal{KB}$, denoted by $\kappa \models \mathcal{KB}$, iff $\kappa(\omega) = \infty$ for all $\omega \not\models \mathcal{F}$, and $\kappa \models \mathcal{R}$.

If $\kappa \models \mathcal{KB}$ then we also say that κ is a *model* of $\mathcal{KB}$. If there is no κ with $\kappa \models \mathcal{KB}$ then $\mathcal{KB}$ is *inconsistent*.

In this way, we can accurately distinguish between the statements "A certainly holds for all individuals" ($\forall x A(\vec{x}) \in \mathcal{F}$), "it is plausible that A holds for all individuals" ($\forall x A(\vec{x}) \in \mathcal{R}$, treated as $(\forall x A(\vec{x}) \,|\, \top)$), and "$A$ is plausible" ($A(\vec{x}) \in \mathcal{R}$, treated as $(A(\vec{x}) \,|\, \top)$). In general, having a classical (i. e., unconditional) formula A in $\mathcal{F}$ expresses "A is certain" while A in $\mathcal{R}$ means "A is plausible". Before illustrating the first-order semantics defined above, we first carry over the idea of (propositional) c-representations [8] to the first-order case. This will endow us with the possibility of constructing proper OCF-models of knowledge bases in an easy way.

First-order conditional reasoning

In the propositional case, conditional structures [7, 8] prove to be a powerful means to rule (inductive) conditional reasoning and belief revision. In this section, we will generalize the concept of conditional structures to the first-order case by defining the conditional structure of a possible world with respect to a ground conditional, a first-order conditional, and a set of first-order conditionals. We recall briefly the basic elements of the theory of conditional structures, generalize this theory to the first-order case, and end up with the definition of c-representations for sets of first-order conditionals. The class of all c-representations of a knowledge base provides a semantics that nicely reflects the interactions between conditionals, and a single c-representation may serve as a model for inductive non-monotonic reasoning. Due to restricted space conditions, we will only mention the main features of this generalization and omit all technical details that are analogous to the propositional case.

Let $\mathcal{R} = \{(B_1(\vec{x}_1) \,|\, A_1(\vec{x}_1)), \ldots, (B_n(\vec{x}_n) \,|\, A_n(\vec{x}_n))\}$ be a finite set of first-order conditionals. These conditionals can either involve open or closed formulas; we may omit the (outer) quantification of conditionals, as no existential conditional may occur, and all universal conditionals can be replaced by the set of all instantiations, according to Definition 3. Moreover, all formulas in $\mathcal{R}$ can be assumed to have a conditional form, according to the remarks around Definition 6 at the end of the preceding section. To each conditional $r_i = (B_i(\vec{x}_i) \,|\, A_i(\vec{x}_i))$, we associate a pair of symbols $\mathbf{a}_i^+, \mathbf{a}_i^-$, symbolizing a positive (negative) effect in case of verification (falsification). In order to make these conditional effects computable, we make use of a group structure, introducing the *free abelian group* $\mathcal{F}_\mathcal{R} = \langle \mathbf{a}_1^+, \mathbf{a}_1^-, \ldots, \mathbf{a}_n^+, \mathbf{a}_n^- \rangle$ with generators $\mathbf{a}_1^+, \mathbf{a}_1^-, \ldots, \mathbf{a}_n^+, \mathbf{a}_n^-$, i.e. $\mathcal{F}_\mathcal{R}$ consists of all elements of the form $(\mathbf{a}_1^+)^{r_1}(\mathbf{a}_1^-)^{s_1} \ldots (\mathbf{a}_n^+)^{r_n}(\mathbf{a}_n^-)^{s_n}$ with integers $r_i, s_i \in \mathbb{Z}$ (the ring of integers). Making use of the functions v_i (for *verifying*) and f_i (for *falsifying*) defined by $v_i(\omega) = \#\{\vec{a}_i \in \mathcal{H}^{(B_i(\vec{x}_i) \,|\, A_i(\vec{x}_i))} \mid \omega \models A_i(\vec{a}_i)B_i(\vec{a}_i)\}$

and $f_i(\omega) = \#\{\vec{a}_i \in \mathcal{H}^{(B_i(\vec{x}_i)\,|\,A_i(\vec{x}_i))} \mid \omega \models A_i(\vec{a}_i)\overline{B_i}(\vec{a}_i)\}$, we define the *conditional structure of* ω *with respect to* $\mathcal{R}$ by

$$\sigma_{\mathcal{R}}(\omega) = \prod_{i=1}^{n} (\mathbf{a}_i^+)^{v_i(\omega)} (\mathbf{a}_i^-)^{f_i(\omega)}.$$

Having defined conditional structures of possible worlds with respect to first-order conditionals, we can proceed in just the same way as in the propositional case. Briefly, we make use of homomorphisms to map the conditional structures to ranks in such a way that sets of worlds that show the same (verifying, falsifying, neutral) all-over behaviour with respect to the conditionals, are assigned the same all-over rank. Ranking functions that comply with conditional structures in this way are called *conditionally indifferent with respect to* $\mathcal{R}$. For the technical details on conditional indifference, we refer to [7, 8]. A simple, but important consequence of conditional indifference is that worlds which are equivalent with respect to the conditionals in $\mathcal{R}$ are mapped onto the same values by ranking functions κ which are conditionally indifferent:

Proposition 1. *If the ranking function* κ *is conditionally indifferent with respect to* $\mathcal{R}$*, then* $\sigma_{\mathcal{R}}(\omega_1) = \sigma_{\mathcal{R}}(\omega_2)$ *implies* $\kappa(\omega_1) = \kappa(\omega_2)$ *for all worlds* $\omega_1, \omega_2 \in \Omega$*.*

The result of this theory of conditional structures is a constructive schema for ranking functions that are both models of the first-order knowledge base and conditionally indifferent with respect to it, so-called c-representations, of which we will only use the simple version here focusing on falsification.

Definition 7 (Simple c-representation)**.** A ranking function κ is a *simple c-representation* iff it satisfies $\kappa \models \mathcal{R}$, and is of the form

$$\kappa(\omega) = \kappa_0 + \sum_{1 \leqslant i \leqslant n} f_i(\omega)\kappa_i^-. \qquad (6)$$

κ_0 is a normalizing factor here so that c-representations comply with the demand $\kappa^{-1}(0) \neq \emptyset$. Two features are characteristic for the approach based on conditional structures. First, and most decisively for the first-order case, the same parameter κ_i^- is associated with all instances of an open conditional since κ_i^- is the numerical realization of the algebraic parameter $\mathbf{a}_i^-$ representing the open conditional. Moreover, κ_i^- does not depend on the specific world ω under consideration but, of course, depends on the other conditionals in $\mathcal{R}$. The interactions between the conditionals in $\mathcal{R}$, each being represented by the respective κ_i^-, is handled in a very flexible way by requiring $\kappa \models \mathcal{R}$ which gives rise to a set of inequalities between the κ_i^-. This intensional handling of interactions of conditionals is the second characteristic feature of our approach. It is different from counting violated conditionals, as in [11], and even different from the similar approach in [5]. Here, we apply this idea consequently to the first-order case. Note that open and quantified conditionals are treated differently: While $\forall\vec{x}(B(\vec{x}) \mid A(\vec{x}))$ is established by considering each instantiation $(B(\vec{a}) \mid A(\vec{a}))$ separately, i.e., each instance is represented by a separate $\mathbf{a}_i^-$, for establishing $(B(\vec{x}) \mid A(\vec{x}))$, all instances $(B(\vec{a}) \mid A(\vec{a}))$ are associated with the same $\mathbf{a}_i^-$. Therefore, the numerical parameters κ_i^- are determined in different ways.

Like system Z [6], or ranking functions in general, c-representations offer the possibility to define model-based default reasoning by selecting a specific model of a knowledge base. However, as all c-representations observe the conditional structures of worlds, they combine semi-qualitative information on ranks with structural information. So, the actual numbers that are assigned to the κ_i's in simple c-representations (cf. (6)) just play a minor role, it is more the relationships between these numbers which result from the constraints in Definition 5 that guide crucial inferences. This allows us to study inferences based on *all* c-representations. In the following examples, we will elaborate on these qualitative relationships, considering both inferences for all c-representations and model-based inferences by *minimal simple c-representations*, for which the constants κ_i^- are chosen in a minimal way to guarantee acceptance.

Example 2 (The elephant zoo)**.** First, we will have a closer look at Example 1 mentioned in the introduction. We abbreviate Clyde by c and Fred by f and consider again the set of conditionals $\mathcal{R} = \{r_1, r_2, r_3\}$ given as in Example 1. Moreover, we assume the set of constants U_Σ to consist of two elephants c, e and two keepers f, k, i.e. $U_\Sigma = \{c, e, f, k\}$. The set $\mathcal{F}$ of factual knowledge is given as $\mathcal{F} = \{E(c), E(e), K(f), K(k),\ \forall x \forall y : L(x, y) \Rightarrow E(x) \wedge K(y)\}$. In particular, all possible worlds that violate the conjunction $E(c)E(e)K(f)K(k)$ will receive an infinite rank (see Definition 6) and can be omitted. We apply the semantics defined in Definition 5 to compute simple c-representations κ that satisfy r_1–r_3 in the following way. Basically, we assume κ to be of the form (6) and determine $\kappa_1^-, \kappa_2^-, \kappa_3^-$ (one for each rule) appropriately so as to comply with all constraints in Definition 5. Moreover, we have to choose the normalizing factor κ_0 so that $\min_\omega \kappa(\omega) = 0$. In the following, we will present the resulting inequalities for the κ_i^- to illustrate the sophisticated handling of conditional information on individuals in general and to show in particular that even exceptions to exceptions are dealt with appropriately in our framework. First, all rules need to have representatives. This is particularly simple for r_3, as only (c, f) can be a representative of r_3: $\kappa \models r_3$ iff $\kappa(L(c, f)) < \kappa(\overline{L}(c, f))$. This amounts to postulating $\kappa_2^- < \kappa_1^- + \kappa_3^-$. r_2 can only have two representatives, namely (c, f), or (e, f). However, (c, f) is already a representative of r_3 which would contradict r_2. So, (e, f) can be the only representative. Indeed, regardless of which κ_i^- are chosen, we find that $\kappa(\overline{L}(e, f)) = \min\{\kappa_1^- + \kappa_2^-, 2\kappa_1^- + \kappa_3^-\} = \min_a \kappa(\overline{L}(a, f)$, making (e, f) a proper candidate for a representative due to (1). Next, (2) demands that $\kappa(\overline{L}(e, f)) < \kappa(L(e, f))$ must hold, which yields $\kappa_1^- < \kappa_2^-$. Again, the precise numbers (also for κ_3^-!) are of no concern. Since we have only one weak representative, $Rep(r_2) = WRep(r_2) = \{(e, f)\}$. Due to the inequalities determined so far, we have that $\kappa(\overline{L}(x, f)) = \kappa(L(x, f))$. Hence, according to Definition 5, we have to compare the representatives of r_2 and its negated counterpart, $(L(x, f) \mid E(x) \wedge K(f))$ (note that this is not the same as r_1, as its range is different). We find that $Rep((L(x, f) \mid E(x) \wedge K(f))) = \{(c, f)\}$. Comparing the rankings for these representatives according to (5) yields $2\kappa_2^- < 2\kappa_1^- + \kappa_3^-$.

Finally, for r_1, we find that $\kappa(L(e, k)) = \kappa(L(x, y))$, and in order to make (e, k) a weak representative, the constraint $\kappa_1^- > 0$ is necessary. Due to structural similarities, we also have $WRep(r_1) = \{(e, k), (c, k), (c, f)\}$, but (c, f) cannot be a proper representative of r_1, as this would inhibit (e, f) being a representative for r_2. This demands for $\kappa_3 > \kappa_2$. So, we have $Rep(r_1) = \{(e, k), (c, k)\}$. Again, we find that $\kappa(L(x, y)) = \kappa(\overline{L}(x, y))$, and we have to compare representatives. Since $Rep((\overline{L}(x, y) \mid E(x) \wedge K(y)) = \{(e, f)\}$, (5) can only be satisfied if $2\kappa_1^- < \kappa_2^-$. All these inequalities now specify c-representations which are models of the elephant-keeper-knowledge base. We may choose the κ_i^- in a minimal way (i.e., $\kappa_1^- = 1, \kappa_2^- = 3, \kappa_3^- = 5$) to obtain a concrete ranking function. If our zoo contains more elephants and keepers, similar computations would yield c-representations that treat all elephants resp. keepers not mentioned in the knowlegde base in the same way as e resp. k, thanks to the structural evaluation of conditionals.

Example 3 (Penguins and super-penguins)**.** In this example, we have *penguins (P), birds (B)* and also *(flying) super-penguins (S)* as well as the classes of *winged things (W)* and *flying things (F)*, and our universe consists of the following objects: $t = \textit{Tweety}, p = \textit{Polly}, s = \textit{Supertweety}$, and other things $o_1, \ldots, o_m$, maybe birds or not. The knowledge base $\mathcal{KB}_{tweety} = \langle \mathcal{F}, \mathcal{R} \rangle$ consists of the facts $\mathcal{F} = \{P(t), S(s), B(p), \forall x S(x) \Rightarrow P(x), \forall x P(x) \Rightarrow B(x)\}$, and the conditional knowledge base $\mathcal{R} = \{r_1, r_2, r_3, r_4\}$ containing four open first-order conditionals:

$$r_1 : (F(x) \,|\, B(x)), \quad r_2 : (W(x) \,|\, B(x))$$
$$r_3 : (\overline{F}(x) \,|\, P(x)), \quad r_4 : (F(x) \,|\, S(x)).$$

Let $\kappa_i^-, i = 1, \ldots, 4$ be the penalties for the four conditionals $r_1, \ldots, r_4$, according to Definition 7. Computing the constraints specified for a simple c-representation κ to be a model of $\mathcal{KB}_{tweety}$ (cf. Definition 6) yield $0 < \kappa_1^- < 2\kappa_1^- < \kappa_3^- < \kappa_4^- < \kappa_1^- + \kappa_4^-$ and $2\kappa_3^- < 2\kappa_1^- + \kappa_4^-$. For instance, choosing $\kappa_1^- = 1 = \kappa_2^-, \kappa_3^- = 3, \kappa_4^- = 5$, would give rise to a (minimal) c-representation of $\mathcal{KB}_{tweety}$. All c-representations handle the conditionals accurately, allowing exceptions (penguins) and exceptions to exceptions (super-penguins), taking *Polly*, *Tweety*, *Supertweety* as a representative bird, penguin, and super-penguin, respectively. If any of the other individuals $o_1, \ldots, o_m$ were found to be a bird, penguin, or super-penguin, it would be treated according to its most specific subclass.

Furthermore, the question whether all birds are expected to have wings is of particular interest. Exceptional birds, like penguins and super-penguins, may also be treated as exceptions with respect to other properties of birds (this has become known as the *drowning problem*). In our approach, we find that *for all individuals a*, be it either a bird, a penguin, a super-penguin, or something else, the difference between $\kappa(B(a)W(a))$ and $\kappa(B(a)\overline{W}(a))$ is the same—namely, determined by κ_2^-—, so if a is known to be a bird then it is expected to have wings. Moreover, we might also conclude $(W(x) \,|\, P(x))$ and $(W(x) \,|\, S(x))$ for each c-representation, as *Tweety* and *Supertweety* would serve as proper representatives for these conditionals, respectively.

Formal properties

A (simple) c-representation κ of a first-order knowledge base $\mathcal{KB} = \langle \mathcal{F}, \mathcal{R} \rangle$ is a way to implement a model-based inductive reasoning mechanism, cf. [13]. However, it is clearly not the only way to implement such a reasoning mechanism and one may raise the question whether it is a "good" reasoning mechanism. In the following, we investigate properties for both the general first-order semantics and inference with c-representations. Some of these properties are inspired by similar properties for first-order probabilistic reasoning [17].

The first property links the first-order to the propositional setting.

Proposition 2 (Existence of Representatives)**.** *If $(B(\vec{x}) \,|\, A(\vec{x})) \in \mathcal{KB}$ and $\kappa \models \mathcal{KB}$ then there is $\vec{a}$ with $\kappa \models (B(\vec{a}) \,|\, A(\vec{a}))$.*

The above proposition states that for each conditional in $\mathcal{KB} = \langle \mathcal{F}, \mathcal{R} \rangle$ and model κ of $\mathcal{KB}$ there is at least one instantiation that is satisfied by κ in the propositional sense. The satisfaction of this property may be considered controversial due to the following reasons. Consider the conditional $r = (yellow(x) \,|\, lemon(x))$ which says that lemons are usually yellow [14]. In order for an OCF κ to satisfy r there has to be at least one actual individual a such that $\kappa \models (yellow(a) \,|\, lemon(a))$ holds. But imagine all lemons are infected with some rare disease that turns them blue. Still, in [2] it is argued that the conditional r should be valid, as lemons are usually yellow despite the fact that all currently present lemons are blue. However, allowing this situation to occur makes it difficult for an approach to distinguish validity of the complementary conditionals $(B(\vec{x}) \,|\, A(\vec{x}))$ and $(\overline{B}(\vec{x}) \,|\, A(\vec{x}))$, see also the discussion on this topic for motivating our OCF-based semantics above. If we are in a world where all lemons are blue how can we accept the conditional "lemons are usually yellow" but not "lemons are usually blue", or even "lemons are usually green"? This problem is avoided by our semantics thanks to the *existence of representatives* and to the following property that has already been discussed above.

Proposition 3 (Coherence of Inference)**.** *For every OCF κ, there is no $(B(\vec{x}) \,|\, A(\vec{x}))$ such that $\kappa \models (B(\vec{x}) \,|\, A(\vec{x}))$ and $\kappa \models (\overline{B}(\vec{x}) \,|\, A(\vec{x}))$.*

Note that the approach of [2] satisfies neither *existence of representatives* nor *coherence of inference*. Our semantics is based on a specific relation between antecedent and consequent of a conditional. Such logics were characterized in [12] as being *entailment preserving* which means that implications imply conditional relationships. This holds in our approach if representatives of the antecedent exist.

Proposition 4 (Entailment Preserving)**.** *For every OCF κ, if $\kappa \models \forall \vec{x} : A(\vec{x}) \Rightarrow B(\vec{x})$ and $\kappa(A(\vec{x}))$ is finite then $\kappa \models (B(\vec{x}) \,|\, A(\vec{x}))$.*

While the above properties describe our general first-order semantics we now turn to properties of c-representations. In general, the beliefs one obtains for specific individuals is of special interest, in particular, if those individuals are special in some respect. However, an important demand to be made is that for indistinguishable individuals the same information should be inferred. That is, if for two individuals c_1, c_2 the same information is expressed in $\mathcal{KB}$ then the rank of a formula A should be the same as the rank of the formula A' in which c_1 and c_2 have been swapped. To formalize this intuition let $\phi[c_1 \leftrightarrow c_2]$ denote the same as ϕ (either being a formula, a knowledge base, or a possible world) but every occurrence of c_1 is replaced by c_2 and vice versa.

Proposition 5 (Indifference of Individuals)**.** *Let κ_c be a (simple) c-representation for the knowledge base $\mathcal{KB} = \langle \mathcal{F}, \mathcal{R} \rangle$. If $\mathcal{F} = \mathcal{F}[c_1 \leftrightarrow c_2]$ and $\mathcal{R}$ does not mention c_1 and c_2 then $\kappa_c(A) = \kappa_c(A[c_1 \leftrightarrow c_2])$ for every ground formula A.*

Note that we have to assume that $\mathcal{R}$ does not mention either c_1 or c_2 in order to establish that ω and $\omega[c_1 \leftrightarrow c_2]$ have the same conditional structure. In Example 3, two objects o_i and o_j are indistinguishable with respect to $\mathcal{KB}_{tweety}$, consequently, inferences are the same for o_i and o_j.

Related work

Although there is plenty of work on propositional conditional logics—to name just a few seminal papers, see e. g. [6, 10]— there is only little work on first-order conditional logics.

The core idea of [2] is that conditionals are interpreted not on the whole range of individuals but, in each possible world, on a subset of "actual individuals". Using the operator $\Rightarrow_x$—which only ranges over actual individuals—default statements can be made that do not consider exceptional individuals. For example, the set $\{Bird(x) \Rightarrow_x Fly(x), Bird(Opus), \neg Fly(Opus)\}$ is consistent because Opus can be regarded as exceptional. We already mentioned that the approach of [2] allows both $(B(\vec{x}) \,|\, A(\vec{x}))$ and $(\overline{B}(\vec{x}) \,|\, A(\vec{x}))$ to be derivable at the same time and therefore fails to satisfy *coherence of inference*. The reason for this is that $\{(B(\vec{x}) \mid A(\vec{x})), (\overline{B}(\vec{x}) \mid A(\vec{x}))\}$ is satisfiable by an interpretation that, in each possible world, takes the empty set to be the set of actual individuals. Besides this, in [2] only the semantics for a first-order conditional logic is proposed and no inference other than semantical entailment is investigated. The semantical entailment relation from [2] suffers from the *drowning problem* which has been discussed above. This means that in the example on penguins and super-penguins, $(W(x) \mid P(x))$ would not be derivable. In [2] individuals have to be separated into actual and exceptional individuals and conditionals are interpreted using only the former. However, [2] describes only a general framework for investigating such situations but gives no hints on how to actually determine exceptional individuals. By using representatives we developed a mechanism that addresses this problem and is able to determine which individuals are exceptional with respect to different aspects.

Our approach to first-order conditional reasoning with ranking functions is inspired by [17] which deals with first-order conditional reasoning with probability functions. There, probabilistic conditionals of the form $(B(\vec{x}) \mid A(\vec{x}))[p]$ (*if* $A(\vec{x})$ *then* $B(\vec{x})$ *with probability* p) are used for knowledge representation, and two novel semantics are proposed. Our motivation for the ranking semantics defined in Def. 5 is the same as for the probabilistic case but differs significantly in its implementation due to the characteristics of qualitative reasoning. The semantics of [17] are defined by considering all instantiations of a conditional and do not differentiate between different types of individuals explicitly. Here, we introduced the particularly important role of representatives into the semantics in order to formalize the intuitive meaning of a conditional, namely, that a conditional is accepted if it has most convincing examples.

The work reported here is also related to works on defeasible reasoning in description logics. For example, the paper [1] defines the rational closure [11] for the description logic $\mathcal{ALC}$. The rational closure is a specific non-monotonic reasoning mechanism for propositional logics that has many desirable properties. While already using the very restrictive first-order language $\mathcal{ALC}$, the work [1] makes a lot of other assumptions on the structure of the knowledge bases. For one, they do not allow for *cyclic* dependencies of rules and assume that for every existential formula such as $(\exists x : R(x,a))$ the individual x is explicitly named, i. e., if $(\exists x : R(x,a))$ is in the knowledge base then there has to be an individual b such that $R(b,a)$ is in the knowledge base as well. Our approach is more general than the approach of [1] as we do not impose such restrictions. The greatest difference, however, to works like [1, 4, 3] is that a linear order of the individuals encoding "normality" is assumed as input in order to be able to reason. Our approach does not need such an order as the representativeness of individuals is elaborated solely by the information encoded in the knowledge base.

Summary and Future Work

In this paper we considered conditional reasoning in a first-order context. We made use of ordinal conditional functions to present a novel semantics for conditional knowledge bases that focuses on the role of representatives. In order to find suitable ranking models for such knowledge bases, we generalized the theory of conditional structures to the first-order case and also extended the notion of c-representations accordingly. We illustrated the properties of first-order c-representation by benchmark examples and identified several formal key properties. These investigations show that the structural theory which our approach is based upon allows an accurate handling of conditional knowledge, both for individuals and for classes. In particular, the combination of making reference to representatives and using conditional structures makes it possible that representatives induce a "normal" behaviour for whole classes of individuals. This is due to assigning a structural impact factor to each conditional and using it coherently for all instantiations. So, if the representative needs this factor to be adjusted, suitable inferences for all comparable individuals will result. The theory of conditional structures [7] first designed for propositional conditional logics proves to be particularly useful in this first-order context. As part of our ongoing work, we investigate the properties of the semantics of first-order c-representations in more detail. Moreover, as conditional structures also provide a basis for powerful belief revision operators satisfying the principle of conditional preservation [8], the semantics presented in this paper may also be used to devise revision operators for first-order belief bases.

The approach proposed here has been prototypically implemented within the *Tweety library for artificial intelligence*[4].

Acknowledgement. The research reported here was supported by the DFG (KE 1413/2-1).

REFERENCES

[1] G. Casini and U. Straccia, 'Rational closure for defeasible description logics', in *Proceedings of JELIA'10*, pp. 77–90, (2010).

[2] J. P. Delgrande, 'On First-order Conditional Logics', *Artificial Intelligence*, **105**(1-2), 105–137, (1998).

[3] N. Friedman, J. Y. Halpern, and D. Koller, 'First-order conditional logic for default reasoning revisited', *ACM Transactions on Computational Logic*, **1**(2), 175–207, (2000).

[4] L. Giordano, V. Gliozzi, N. Olivetti, and G. L. Pozzato, 'Preferential vs rational description logics: which one for reasoning about typicality?', in *Proceedings of ECAI'10*, pp. 1069–1070, (2010).

[5] M. Goldszmidt, P. Morris, and J. Pearl, 'A maximum entropy approach to nonmonotonic reasoning', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**(3), 220–232, (1993).

[6] M. Goldszmidt and J. Pearl, 'Qualitative Probabilities for Default Reasoning, Belief Revision, and Causal Modeling', *Artificial Intelligence*, **84**, 57–112, (1996).

[7] G. Kern-Isberner, *Conditionals in Nonmonotonic Reasoning and Belief Revision*, Springer, 2001.

[8] G. Kern-Isberner, 'A Thorough Axiomatization of a Principle of Conditional Preservation in belief revision', *Annals of Mathematics and Artificial Intelligence*, **40**(1–2), 127–164, (2004).

[9] S. Kraus, D. J. Lehmann, and M. Magidor, 'Nonmonotonic Reasoning, Preferential Models and Cumulative Logics', *Artificial Intelligence*, **44**(1–2), 167–207, (1990).

[10] D. Lehmann, 'What Does a Conditional Knowledge Base Entail?', in *Proceedings of KR'89*, Toronto, Canada, (1989).

[11] D. Lehmann, 'Another perspective on default reasoning', *Annals of Mathematics and Artificial Intelligence*, **15**(1), 61–82, (1995).

[12] D. Nute, *Topics in Conditional Logic*, D. Reidel Publishing Company, Dordrecht, Holland, 1980.

[13] J. B. Paris, *The Uncertain Reasoner's Companion – A Mathematical Perspective*, Cambridge University Press, 1994.

[14] H. Putnam, 'Is semantics possible?', in *Mind, Language and Reality: Philosophical Papers*, volume 2, 139–152, Cambridge, (1975).

[15] R. Reiter, 'A Logic for Default Reasoning', *Artificial Intelligence*, **13**, 81–132, (1980).

[16] W. Spohn, 'Ordinal Conditional Functions: A Dynamic Theory of Epistemic States', in *Causation in Decision, Belief Change, and Statistics*, volume 2, 105–134, Kluwer, (1988).

[17] M. Thimm and G. Kern-Isberner, 'On Probabilistic Inference in Relational Conditional Logics', *Logic Journal of the IGPL*, (2012). 10.1093/jigpal/jzs010.

[4] http://sourceforge.net/projects/tweety/