A Preference-Based Approach for Representing Defaults in First-Order Logic

James Delgrande, **Christos Rantsoudis**
Simon Fraser University, Canada
first_last@sfu.ca

Abstract

A major area of knowledge representation concerns representing and reasoning with defaults such as "birds fly". In this paper we introduce a new, preference-based approach for representing defaults in first-order logic (FOL). Our central intuition is that an individual (or tuple of individuals) is not simply normal or not, but rather is normal with respect to a particular predicate. Thus an individual that satisfies $Bird$ may be normal with respect to Fly but not $BuildNest$. Semantically we associate a total preorder over n-tuples with each n-ary relation in the domain. Syntactically, a predicate-forming construct is introduced into FOL that lets us assert properties of minimal elements in an ordering that satisfy a given condition. Default inference is obtained by (informally) asserting that a tuple in an ordering is ranked as "low" as consistently possible. The approach has appealing properties: specificity of defaults is obtained; irrelevant properties are handled appropriately; and one can reason about defaults within FOL. We also suggest that the approach is more expressive than extant approaches and present some preliminary ideas for its use in Description Logics.

1 Introduction

One of the major challenges of Artificial Intelligence (AI) has been in representing and reasoning with *defaults* such as "birds fly". Since the early days of AI, researchers in the field have recognized the importance of intelligent systems being able to draw *default assertions*, where one would conclude by default that a bird flies, while allowing for exceptional conditions and non-flying birds. Of the early approaches to nonmonotonic reasoning, *default logic* (Reiter 1980) and *autoepistemic logic* (Moore 1985) were based on the notion of a fixed-point construction in order to expand the set of obtained consequences, while *circumscription* (McCarthy 1980) was based on the idea of minimizing the extension of a predicate. In these approaches, desirable properties (such as *specificity*) had to be hand-coded in a theory (Reiter and Criscuolo 1981; McCarthy 1986). About a decade later, approaches based on *conditional logics* (Delgrande 1987; Lamarre 1991; Boutilier 1994; Fariñas del Cerro, Herzig, and Lang 1994) and nonmonotonic consequence relations (Kraus, Lehmann, and Magidor 1990; Lehmann and Magidor 1992) represented defaults as objects (binary modal operators in conditional logics) in a theory. In such approaches, the semantics was based on an ordering over possible worlds. While properties such as specificity followed directly from the semantics, other properties, such as handling irrelevant properties, were not obtained. Arguably, at present there is no generally-accepted approach that adequately handles inference of default properties, reasoning in the presence of irrelevant information, and reasoning about default properties of an individual known to be exceptional with respect to another default property.

In this paper, we present a new account of defaults. Consider the default assertions "birds fly" and "birds build nests". The usual interpretation is that a normal bird flies and it builds nests. Our interpretation is that, with regards to flying, a normal bird flies, and with regards to nest building, a normal bird builds nests. That is, normality is given with respect to some property. Consequently, "birds fly" would be interpreted as saying that, with respect to the property of flight, an individual that is a bird in fact flies. Similarly, a penguin, as concerns flight, does not fly.

Semantically, for each n-ary relation in the domain, we associate a total preorder over n-tuples of individuals, where the preorder gives the relative normality of a tuple with respect to that relation. Syntactically, we introduce a "predicate-forming construct" into the language of FOL that lets us identify those individuals that satisfy a certain condition (like $Bird$) and that are minimal in a given ordering (like that corresponding to Fly); one can then state assertions regarding such (minimal-in-the-ordering) individuals, for example that they indeed satisfy Fly. Notably, an individual abnormal in one respect (like flight) may be normal in another respect (like nest building). These orderings allow us to naturally specify a wide class of default assertions, including on predicates of arity >1. Default inference, in which an individual is concluded to have a given property "by default", is specified via a preference ordering over models. Then inferences that follow by default are just those that obtain in the minimal models. In the approach we avoid a modal semantics on the one hand and fixed-point constructions on the other. We also show how a "predicate-forming construct" can be translated into a standard first-order theory and argue that the approach presents various advantages: it satisfies a set of broadly-desirable properties; it is perspicuous, and presents a more nuanced and expressive account of defaults than previous approaches; and it is couched within classical FOL.

In the next section we informally describe our framework. In Section 3 we present the syntax and semantics of our logic. After presenting some examples in Section 4, we look at various properties of our logic as well as provide a characterization result in Section 5. We briefly present our treatment of nonmonotonic inferences in Section 6. In Section 7 we compare our approach with related work as well as discuss about future directions. Section 8 concludes.

2 The Approach: Intuitions

A common means for specifying the meaning of a default is via a *preference order* over models or possible worlds in which worlds or models are ordered with respect to their *normality*. Then, something holds normally (typically, defeasibly, etc.) just when it holds in the most preferred models or possible worlds. For example, in a conditional logic, "birds fly" can be represented propositionally as $Bird \Rightarrow Fly$. This assertion is true just when, in the minimal $Bird$-worlds, Fly is also true. In circumscription "birds fly" can be represented as $\forall x.Bird(x) \wedge \neg Ab_f(x) \rightarrow Fly(x)$, so a bird that is not abnormal with respect to flight flies. Then models are ordered based on the extensions of the Ab predicates (with smaller extensions preferred), and a bird a flies by default just if $Fly(a)$ is satisfied in the minimal models that satisfy $Bird(a)$.

Our approach belongs to the preference-based paradigm, but with significant differences from earlier work. Our preferences are expressed *within* FOL models, and not between models (or possible worlds, as in a modal framework). Preferences are given by a total preorder over n-tuples of individuals for each n-ary relation in the domain; these orderings give the relative normality of a tuple with respect to the underlying relation. Defaults are then expressed by making assertions concerning sets of minimal (tuples of) individuals in an ordering.

Consider again the assertion that birds normally fly. We interpret this as, for a bird that is normal with respect to the unary relation fly,[1] that bird flies. In a model, the relative normality of individuals with respect to flight is given by a total preorder associated with the relation fly. Then we can say that "birds fly" is true in a model just when, in the order associated with fly, the minimal $bird$ individuals satisfy fly. Similarly, "penguins do not fly" is true in a model just when, in the order associated with fly, the minimal $penguin$ individuals do not satisfy fly. The ranking of an individual with respect to one relation (like fly) is not related to the ranking associated with another relation (like $build\_nest$).

These considerations extend to relations of arity > 1. Consider "elephants normally like their keepers". Semantically we would express this by having, in the total preorder associated with the relation $likes$, that the most normal *pairs* of individuals (d_1, d_2), in which d_1 is an elephant and d_2 is a keeper, satisfy $likes$. Analogously we could go on and express that "elephants normally do not like (keeper) Fred".

[1]We use the notation that a lower case string like fly is used for a relation in a model whereas upper case, like Fly, is used for a predicate symbol in the language (in this case denoting fly).

Syntactically, we introduce a new construct into the language of FOL that, for an ordering associated with a relation, enables us to specify minimal domain elements in the ordering that satisfy a given condition. This construct has two parts, a predicate P and a formula ϕ; it is written $\{P(\vec{y}), \phi(\vec{y})\}$. The construct stands for a (new) predicate that denotes a domain relation which holds for just those (tuples of) individuals that satisfy ϕ and that are minimal in the ordering corresponding to P (i.e., the ordering associated with the denotation of P).

Given this construct, one can make assertions regarding individuals that satisfy this relation. For example, to express "birds normally fly" we use:

$$\forall x \left(\{Fly(y), Bird(y)\}(x) \rightarrow Fly(x)\right) \tag{1}$$

whereas for "penguins normally do not fly" we use:

$$\forall x \left(\{Fly(y), Penguin(y)\}(x) \rightarrow \neg Fly(x)\right) \tag{2}$$

So those individuals that satisfy $bird$ and that are minimal in the ordering associated with fly also satisfy fly whereas the minimal elements in the fly ordering that satisfy $penguin$ do not satisfy fly. That is, "birds fly" and "penguins do not fly" both concern the property of flight and so are respect to the same (fly) ordering.[2]

The fact that we deal with orderings over individuals means that our approach is irreducibly first-order. This is in contrast to most work in default logic, in which default theories are very often expressed in propositional terms, and where a rule with variables is treated as the set of corresponding grounded instances. It is also in contrast to work in conditional logics and nonmonotonic inference relations, which are nearly always expressed in propositional terms. As we suggest later, for many domains, it may well be that a first-order framework is essential for an adequate expression of defaults assertions.

For our earlier example "elephants (E) normally like (L) their keepers (K)", we have the following:

$$\forall x_1, x_2\big(\{L(y_1, y_2), E(y_1) \wedge K(y_2)\}(x_1, x_2) \rightarrow L(x_1, x_2)\big) \tag{3}$$

whereas for "elephants normally do not like (keeper) Fred":

$$\forall x_1, x_2\big(\{L(y_1, y_2), E(y_1) \wedge K(y_2) \wedge y_2 = Fred\}(x_1, x_2) \rightarrow \neg L(x_1, x_2)\big) \tag{4}$$

In addition, we suggest that our approach leads to a reconsideration of how some defaults are best expressed. Consider the assertion "adults are normally employed at a company". In a conditional approach, one might express this as:

$$Adult(x) \Rightarrow_x \exists y\big(EmployedAt(x, y) \wedge Company(y)\big)$$

where, without worrying about details, $\Rightarrow_x$ is a variable-binding connective (Delgrande 1998). But the interpretation

[2]One way of viewing this is that a relation such as fly gives a *partition* of a domain, into those elements that belong to the relation and those that do not. We also note that the interpretation of "birds fly" as a default conditional (like $Bird \Rightarrow Fly$) is somewhat superficial. A more nuanced approach would assert that for birds, *flight* is a default means of locomotion, perhaps along with others such as *bipedal walking*. We return to this point later.

that "the most normal adults are employed at a company" is unsuitable, since an abnormal adult here would be abnormal with respect to other normality assertions regarding adults. As well, it does not seem to make much sense to say that normality now refers to the full consequent. Instead, it seems that the best way of interpreting this assertion is that we have a normality ordering associated with $employed\_at$, giving the relative normality of pairs of domain elements with respect to this relation. Then for the most normal pairs (d_1, d_2) where d_1 is an adult, there is some pair (d_1, d_3) among them for which d_1 is employed at d_3 and d_3 is a company. Consequently, this suggests that simple conditionals, at least in a first-order framework, may not be adequate to represent general default information.

The preceding sketches our intuitions regarding how we intend to *represent* and *interpret* default information. With regards to *inferring* default information in a knowledge base (KB), we define preferences between models in a similar manner to those of other *preferential logics* (McCarthy 1980; Shoham 1987; Kraus, Lehmann, and Magidor 1990). Again, what is new in our approach is that we have multiple orderings inside our models and so we can define more nuanced preferences between models. As we will see, although we only briefly treat nonmonotonic inferring of assertions, our ordering between the models will result in desirable properties with respect to defeasibility. Specifically, we satisfy the following principles:

1. *Specificity*: Properties are ascribed on the basis of most specific applicable information. Hence a penguin will not fly by default whereas a bird will.
2. *Inheritance*: Individuals will inherit all typical properties of the classes to which they belong, except for those we know are exceptional. Hence, by default, a penguin may be concluded to not fly, but will be concluded to have feathers, etc.
3. *Irrelevance*: Default inference is not affected by irrelevant information. Hence, by default, a yellow bird will be concluded to fly.

As we have noted, there is no generally-accepted approach that fully captures these properties. Default logic, autoepistemic logic and circumscription do not satisfy specificity, while the rational closure mechanism of the KLM framework does not satisfy inheritance.

3 Language and Semantics

As discussed, a first-order setting is required for our investigation. Thus, the language we employ is based on standard FOL enhanced with the aforementioned minimality operators. We start with some formal preliminaries, including the syntax of our new logic, and finish the section by presenting the semantics.

3.1 Formal Preliminaries

We assume that the reader has some familiarity with standard FOL (Enderton 1972; Mendelson 2015). Let $\mathcal{L}$ be a first-order language containing a set of *predicate symbols* $\mathcal{P} = \{P, Q, \dots\}$, a set of *constant symbols* $\mathcal{C} = \{c_i \mid i \in \mathbb{N}\}$ and a set of *variables* $\mathcal{V} = \{x, y, z \dots\}$.[3] Predicate symbols and variables may be subscripted, as may other entities in the language. The constants and variables make up the set of *terms*, which are denoted by t_i, $i \in \mathbb{N}$. A tuple of variables $x_1, \dots, x_n$ is denoted by $\vec{x}$, and similarly for terms. For any formula ϕ, the expression $\phi(\vec{x})$ indicates that the free variables of ϕ are among those in $\vec{x}$. Our language $\mathcal{L}_N$ is given in the following definition, with $\mathcal{L}$ given in Items 1-3.

Definition 1. *The well-formed formulas (wffs) of $\mathcal{L}_N$ are defined inductively as follows:*

1. *If P is an n-ary predicate symbol and $t_1, \dots, t_n$ are terms then $P(t_1, \dots, t_n)$ is a wff.*
2. *If t_1 and t_2 are terms then $t_1 = t_2$ is a wff.*
3. *If ϕ and ψ are wffs and x is a variable then $(\neg\phi)$, $(\phi \rightarrow \psi)$ and $(\forall x\ \phi)$ are wffs.*
4. *If P is an n-ary predicate symbol, $\vec{y}$ is a tuple of n variables, $\phi(\vec{y})$ is a wff and $\vec{t}$ is a tuple of n terms then $\{P(\vec{y}), \phi(\vec{y})\}(\vec{t})$ is a wff.*

Parentheses may be omitted if no confusion results. The connectives $\wedge$, $\vee$, $\equiv$ and $\exists$ are introduced in the usual way. For a wff of the form $\{P(\vec{y}), \phi(\vec{y})\}(\vec{t})$, the part $\{P(\vec{y}), \phi(\vec{y})\}$ can be thought of as a self-contained *predicate-forming construct (pfc)*. The first part, $P(\vec{y})$, specifies that the ordering is with respect to predicate P; it also provides names for the n variables of P, in $\vec{y}$. The second part $\phi(\vec{y})$ will in general be true of some substitutions for $\vec{y}$ and false for others. The denotation of $\{P(\vec{y}), \phi(\vec{y})\}$ is just those n-tuples of domain elements that satisfy ϕ and are minimal in the ordering corresponding to P. So $\{P(\vec{y}), \phi(\vec{y})\}$ behaves just like any predicate symbol. Thus $\{Fly(y), Bird(y)\}(x)$ can be thought of as analogous to an atomic formula, which in a model will be true of some individuals (viz. those that belong to $bird$ and that are minimal in the fly ordering) and false of others. Similarly, $\{Fly(y), Bird(y)\}(Tweety)$ will assert that $Tweety$ is a minimal $bird$ element in the fly ordering.

In the wff $\{P(\vec{y}), \phi(\vec{y})\}(\vec{t})$, there is a one-to-one correspondence between the terms in $\vec{t}$ and the variables $\vec{y}$ inside the pfc; but otherwise they are unrelated. Hence for the expression $\{Fly(x), Bird(x)\}(x)$ the occurrences of variable x within $\{\dots\}$ are distinct from the third occurrence. For $\{P(\vec{y}), \phi(\vec{y})\}$, the variables in $\vec{y}$ are local to $\{\dots\}$, and can be thought of as effectively bound within the expression. In the following, we use the term *predicate expression* to refer to both predicate symbols and pfcs.

We next remind the reader of some terminology regarding orderings. A *total preorder* on a set is a transitive and connected relation on the elements of the set. A *well-founded* order is one that has no infinitely-descending chains of elements. Formally, for a set S and a total preorder $\preceq$ on S, $\preceq$ is well-founded iff:

$$(\forall T \subseteq S)\big(T \neq \emptyset \rightarrow (\exists x \in T)(\forall y \in T)\, x \preceq y\big)$$

[3]For simplicity, except for constants, we exclude function symbols. Note that this does not affect expressiveness, since any n-ary function can be encoded by a $(n+1)$-place predicate.

We will work only with well-founded total preorders. Given well-foundedness, for a total preorder $\preceq$ we can define the minimal S-elements of $\preceq$, as follows:

$$\min(\preceq, S) = \{x \in S \mid \forall y \in S : x \preceq y\}$$

3.2 Semantics

We next present the formal semantics, which will interpret the terms and formulas in $\mathcal{L}_N$ with respect to a model.

Definition 2. *A model is a triple $\mathcal{M} = \langle \mathcal{D}, \mathcal{I}, \mathcal{O} \rangle$ where $\mathcal{D} \neq \emptyset$ is the domain, $\mathcal{I}$ is the interpretation function, and $\mathcal{O}$ is a set containing, for each n-ary relation r in $\mathcal{D}$, a well-founded total preorder $\preceq_r$ on $\mathcal{D}^n$. Specifically:*

1. *$\mathcal{I}$ interprets the predicate and constant symbols into $\mathcal{D}$ as follows:*
   - *$P^{\mathcal{I}} \subseteq \mathcal{D}^n$, for each n-ary predicate symbol $P \in \mathcal{P}$*
   - *$c^{\mathcal{I}} \in \mathcal{D}$, for each constant symbol $c \in \mathcal{C}$*
2. *$\mathcal{O} = \{ \preceq_r \subseteq \mathcal{D}^n \times \mathcal{D}^n \mid r \subseteq \mathcal{D}^n$ and $\preceq_r$ is a well-founded total preorder on $\mathcal{D}^n \}$*

A *variable map* $v : \mathcal{V} \mapsto \mathcal{D}$ assigns each variable $x \in \mathcal{V}$ an element of the domain $v(x) \in \mathcal{D}$.

Definition 3. *Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{I}, \mathcal{O} \rangle$ be a model and v a variable map. The denotation of a term t, written as $t^{\mathcal{I},v}$, is defined as follows:*

1. *$t^{\mathcal{I},v} = t^{\mathcal{I}}$, if t is a constant*
2. *$t^{\mathcal{I},v} = v(t)$, if t is a variable*

The satisfaction relation $\models$ is defined below. We first give some preliminary terminology and notation. Assume we have a model $\mathcal{M}$, a variable map v and a wff ϕ. When $\mathcal{M}$ satisfies ϕ under v we write $\mathcal{M}, v \models \phi$. When $\mathcal{M}$ satisfies ϕ under v where the free variable x of ϕ is assigned to d we write $\mathcal{M}, v \models \phi(x/d)$. For $\vec{x}$ a tuple of variables $x_1, \ldots, x_n$ and $\vec{d}$ a tuple of domain elements $d_1, \ldots, d_n$, we denote by $\vec{x}/\vec{d}$ the one-to-one assignment $x_1/d_1, \ldots, x_n/d_n$. Similarly for $\vec{x}/\vec{y}$ and $\vec{x}/\vec{t}$. Last, given a tuple of n variables $\vec{y}$ and a formula ϕ with free variables among $\vec{y}$, the values of $\vec{y}$ for which ϕ can be satisfied are given by the set:

$$\phi(\vec{y})^{\mathcal{M},v} = \{\vec{d} \in \mathcal{D}^n \mid \mathcal{M}, v \models \phi(\vec{y}/\vec{d})\} \tag{5}$$

We can now define the denotation of a pfc $\{P(\vec{y}), \phi(\vec{y})\}$, written $\{P(\vec{y}), \phi(\vec{y})\}^{\mathcal{M},v}$, as the set of domain tuples that:

1. belong to the denotation of ϕ, as given in Equation 5 and
2. are the minimal such tuples in the ordering associated with $P^{\mathcal{I}}$, viz. $\preceq_{P^{\mathcal{I}}}$.

Definition 4. *Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{I}, \mathcal{O} \rangle$ be a model and v a variable map. The denotation of $\{P(\vec{y}), \phi(\vec{y})\}$ is defined as the set:*

$$\{P(\vec{y}), \phi(\vec{y})\}^{\mathcal{M},v} = \min(\preceq_{P^{\mathcal{I}}}, \phi(\vec{y})^{\mathcal{M},v})$$

Finally, for each predicate symbol $P \in \mathcal{P}$ we define $P^{\mathcal{M},v} = P^{\mathcal{I}}$. The satisfaction relation is given as follows (recall that a predicate expression is either a predicate symbol or a pfc).

Definition 5. *Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{I}, \mathcal{O} \rangle$ be a model, v a variable map and P a predicate expression.*

1. *$\mathcal{M}, v \models \mathsf{P}(t_1, \ldots, t_n)$ iff $(t_1^{\mathcal{I},v}, \ldots, t_n^{\mathcal{I},v}) \in \mathsf{P}^{\mathcal{M},v}$*
2. *$\mathcal{M}, v \models t_1 = t_2$ iff $t_1^{\mathcal{I},v} = t_2^{\mathcal{I},v}$*
3. *$\mathcal{M}, v \models \neg\phi$ iff $\mathcal{M}, v \not\models \phi$*
4. *$\mathcal{M}, v \models \phi \rightarrow \psi$ iff $\mathcal{M}, v \not\models \phi$ or $\mathcal{M}, v \models \psi$*
5. *$\mathcal{M}, v \models \forall x\phi$ iff $\mathcal{M}, v \models \phi(x/d)$ for all $d \in \mathcal{D}$*

As usual, if $\mathcal{M}, v \models \phi$ for all $\mathcal{M}$ and v, then ϕ is *valid* in $\mathcal{L}_N$. If ϕ is a sentence (i.e., without free variables) then $\mathcal{M}, v \models \phi$ iff $\mathcal{M}, v' \models \phi$ for all variable maps v, v'; thus we just write $\mathcal{M} \models \phi$. If Φ is a set of sentences then $\mathcal{M} \models \Phi$ iff $\mathcal{M} \models \phi$ for all $\phi \in \Phi$, and we say that $\mathcal{M}$ *is a model of* Φ. Finally, we write $\Phi \models \phi$ when all models of Φ are models of ϕ, and we say that Φ *logically entails* ϕ.

4 Examples

We have already seen some wffs in Equations 1–4 of Section 2. We now give some more examples that illustrate the range and application of our approach. As we have described, the first part of a pfc denotes an ordering associated with a given predicate. The second part is a formula that specifies minimal (tuples of) individuals in the ordering. The order of the variables in the two parts is important, as the next two equations illustrate:

$$\forall x_1, x_2 \left(\{L(y_1, y_2), P(y_1, y_2)\}(x_1, x_2) \rightarrow L(x_1, x_2)\right) \tag{6}$$

$$\forall x_1, x_2 \left(\{L(y_1, y_2), P(y_2, y_1)\}(x_1, x_2) \rightarrow L(x_1, x_2)\right) \tag{7}$$

When L abbreviates $Likes$ and P abbreviates $ParentOf$, it is easy to see that 6 states that "parents normally like their children" while 7 states that "children normally like their parents". Recall also that the tuples of domain elements belonging to the denotation of a pfc $\{P(\vec{y}), \phi(\vec{y})\}$ do not necessarily have to satisfy the predicate P. See, for instance, Equations 2 and 4.

On another note, predicates in ϕ may have a higher arity than P in a pfc $\{P(\vec{y}), \phi(\vec{y})\}$. For instance, consider the statement "people that trust (T) themselves are normally daring (D)". We would express that using the following wff:

$$\forall x \left(\{D(y), T(y, y)\}(x) \rightarrow D(x)\right)$$

Furthermore, we can express statements about specific individuals by directly replacing variables with constants. For example, for constant $John$, we can express that "John's pets are normally happy (H)" by:

$$\forall x \left(\{H(y), HasPet(John, y)\}(x) \rightarrow H(x)\right)$$

The reading of our new wffs can sometimes be a bit cumbersome. Consider the earlier example that "adults (A) are normally employed (Em) at a company (C)". According to the discussion in Section 2, this statement can be expressed by the following wff:

$$\forall x_1, x_2 \big(\{Em(y_1, y_2), A(y_1)\}(x_1, x_2) \rightarrow \exists x_3$$
$$\big(\{Em(y_1, y_2), A(y_1)\}(x_1, x_3) \wedge Em(x_1, x_3) \wedge C(x_3)\big)\big)$$

This contains two instances of the pfc $\{Em(y_1, y_2), A(y_1)\}$. As a possible solution, if we introduce the predicate NAE (for *Normal Adults* wrt the Em ordering) we can rewrite the previous into the more compact and perspicuous formula:

$$\forall x, y \big(NAE(x, y) \rightarrow \exists z \big(NAE(x, z) \wedge Em(x, z) \wedge C(z)\big)\big)$$

This method of abbreviating pfcs via smaller "predicate names" could be used at the outset in order to make KBs more readable.

A key point is that our approach is highly versatile, and can express nuances that (arguably) other approaches cannot. Consider for example the ambiguous statement "undergraduate students attend undergraduate courses".[4] Let UGS stand for "undergrad student" and UGC for "undergrad course". Among other possibilities, we have the following interpretations:

1. "Normally, the things UGSs attend are UGCs"

That is, for the most normal pairs (d_1, d_2) according to the $attend$ relation, such that d_1 is an UGS that attends d_2, d_2 is an UGC. In $\mathcal{L}_N$:

$$\forall x_1, x_2\big(\{\mathit{Attend}(y_1, y_2), UGS(y_1) \wedge \mathit{Attend}(y_1, y_2)\}(x_1, x_2) \rightarrow UGC(x_2)\big)$$

2. "Normal UGSs attend only UGCs"

That is, for the most normal pairs (d_1, d_2) according to the $attend$ relation, such that d_1 is an UGS, everything d_1 attends is an UGC. In $\mathcal{L}_N$:

$$\forall x_1, x_2\big(\{\mathit{Attend}(y_1, y_2), UGS(y_1)\}(x_1, x_2) \rightarrow \forall x_3\big(\mathit{Attend}(x_1, x_3) \rightarrow UGC(x_3)\big)\big)$$

3. "Normal UGSs attend some UGC"

That is, for the most normal pairs (d_1, d_2) according to the $attend$ relation, such that d_1 is an UGS, there exists an UGC that d_1 attends. In $\mathcal{L}_N$:

$$\forall x_1, x_2\big(\{\mathit{Attend}(y_1, y_2), UGS(y_1)\}(x_1, x_2) \rightarrow \exists x_3\big(\mathit{Attend}(x_1, x_3) \wedge UGC(x_3)\big)\big)$$

4. "Normally UGSs attend UGCs"

Analogous to Equation 3, we have:

$$\forall x_1, x_2\big(\{\mathit{Attend}(y_1, y_2), UGS(y_1) \wedge UGC(y_2)\}(x_1, x_2) \rightarrow \mathit{Attend}(x_1, x_2)\big)$$

These examples illustrate the wealth of expressiveness in our logic and present a contrast to the more limited expressiveness of current approaches in the literature.

5 Characterization and Properties

In this section we provide a characterization of our new logic through a translation into standard FOL. As well, we present some notable properties and briefly compare our approach to other well-known systems from the literature.

First, we show how to *encode* our approach in standard FOL, via the introduction of a new set of predicate symbols representing the preference orderings. Then, we *express* the pfcs inside the language using these new predicate symbols. This translation then serves as a syntactic counterpart to the more semantic approach of Section 3; and our equivalence result (Theorem 1) provides a counterpart to a standard soundness and completeness result.[5] More precisely:

1. For each n-ary predicate symbol P we introduce a new predicate symbol $P^{\preceq}$ of arity $2n$. We use these new predicate symbols to express the preference orderings instead of embedding them directly into the models. That is, each $P^{\preceq}$ will be used in the place of $\preceq_{P^{\mathcal{I}}}$.
2. After having interpreted the new predicate symbols $P^{\preceq}$ in the aforementioned way, we translate each wff $\{P(\vec{y}), \phi(\vec{y})\}(\vec{t})$ to the first-order formula:

$$\phi(\vec{y}/\vec{t}) \wedge \forall \vec{z}\left(\phi(\vec{y}/\vec{z}) \rightarrow P^{\preceq}(\vec{t}, \vec{z})\right)$$

So the variables from $\vec{y}$ that appear free in ϕ are assigned to the respective terms and variables from $\vec{t}$ and $\vec{z}$, with the latter being employed in order to ensure the minimality of the former through the new predicate symbols $P^{\preceq}$. The following list shows some of the examples of Sections 2 and 4 expressed in FOL using this translation:

1. *Birds (B) normally fly (F)*

$$\forall x\big(B(x) \wedge \forall z\big(B(z) \rightarrow F^{\preceq}(x, z)\big) \rightarrow F(x)\big)$$

2. *Penguins (P) normally do not fly*

$$\forall x\big(P(x) \wedge \forall z\big(P(z) \rightarrow F^{\preceq}(x, z)\big) \rightarrow \neg F(x)\big)$$

3. *Elephants normally like their keepers*

$$\forall x_1, x_2\big(E(x_1) \wedge K(x_2) \wedge \forall z_1, z_2\big(E(z_1) \wedge K(z_2) \rightarrow L^{\preceq}(x_1, x_2, z_1, z_2)\big) \rightarrow L(x_1, x_2)\big)$$

4. *Children normally like their parents*

$$\forall x_1, x_2\big(P(x_2, x_1) \wedge \forall z_1, z_2\big(P(z_2, z_1) \rightarrow L^{\preceq}(x_1, x_2, z_1, z_2)\big) \rightarrow L(x_1, x_2)\big)$$

5. *People that trust themselves are normally daring*

$$\forall x\big(T(x, x) \wedge \forall z\big(T(z, z) \rightarrow D^{\preceq}(x, z)\big) \rightarrow D(x)\big)$$

6. *John's pets are normally happy*

$$\forall x\big(HasPet(John, x) \wedge \forall z\big(HasPet(John, z) \rightarrow H^{\preceq}(x, z)\big) \rightarrow H(x)\big)$$

As we see, everything presented so far can be expressed using standard FOL without the need to enhance the models with preference orderings or the syntax with pfcs. Instead, a new set of predicate symbols together with a translation of pfcs suffice.

Next, we present the formal translation as well as a characterization theorem.

[4] This example is a type of assertion that might occur unconditionally in a description logic TBox. The fact that there are (at least) four corresponding nonmonotonic (normality) assertions indicates that a fully general approach to defeasibility in description logics may require substantial expressive power. See Section 7 for a further discussion.

[5] Alternatively, we could have provided an axiomatisation of our new construct and directly proven a soundness and completeness result. This is done in (Brafman 1997), where a conditional logic is developed based on orderings over individuals, but for each $n \in \mathbb{N}$ there is a single ordering on n-tuples; see Section 7. We feel that the given translation is at least as informative as an axiomatisation, while being more straightforward to obtain.

5.1 Translation into FOL

We first extend $\mathcal{P}$ with a new set of predicate symbols $\mathcal{P}^{\preceq} = \{P^{\preceq} \mid P \in \mathcal{P}\}$. Let $\mathcal{P}^+ = \mathcal{P} \cup \mathcal{P}^{\preceq}$ and let $\mathcal{L}^+$ be the extension of $\mathcal{L}$ with $\mathcal{P}^+$.

Definition 6. *Given the syntax of $\mathcal{L}_N$, the translation $\tau : \mathcal{L}_N \to \mathcal{L}^+$ is defined as follows:*

1. $\big(P(t_1, \ldots, t_n)\big)^\tau = P(t_1, \ldots, t_n)$
2. $(t_1 = t_2)^\tau = (t_1 = t_2)$
3. $(\neg\phi)^\tau = \neg\phi^\tau$
4. $(\phi \to \psi)^\tau = \phi^\tau \to \psi^\tau$
5. $(\forall x\phi)^\tau = \forall x\phi^\tau$
6. $\big(\{P(\vec{y}), \phi(\vec{y})\}(\vec{t})\big)^\tau = \big(\phi(\vec{y}/\vec{t})\big)^\tau \wedge \forall\vec{z}\Big(\big(\phi(\vec{y}/\vec{z})\big)^\tau \to P^{\preceq}(\vec{t}, \vec{z})\Big)$

As for the semantics, the language $\mathcal{L}^+$ is interpreted over the usual models of FOL. Regarding the relation between the models of $\mathcal{L}_N$ and the models of $\mathcal{L}^+$, we can define a similar translation τ from the former into the latter.

Definition 7. *For a given model $\mathcal{M} = \langle\mathcal{D}, \mathcal{I}, \mathcal{O}\rangle$ of $\mathcal{L}_N$, the model $\mathcal{M}^\tau = \langle\mathcal{D}^\tau, \mathcal{I}^\tau\rangle$ of $\mathcal{L}^+$ is defined as follows:*

1. $\mathcal{D}^\tau = \mathcal{D}$
2. *for every constant symbol $c \in \mathcal{C}$: $c^{\mathcal{I}^\tau} = c^{\mathcal{I}} \in \mathcal{D}$*
3. *for every n-ary predicate symbol $P \in \mathcal{P}$:*
   - $P^{\mathcal{I}^\tau} = P^{\mathcal{I}} \subseteq \mathcal{D}^n$
   - $(P^{\preceq})^{\mathcal{I}^\tau} = \{(\vec{d}, \vec{e}) \in \mathcal{D}^n \times \mathcal{D}^n \mid \vec{d} \preceq_{P^{\mathcal{I}}} \vec{e}\}$

It is easy to see that $\mathcal{M}^\tau$ interprets all new predicate symbols $P^{\preceq}$ as (well-founded) total preorders, in the following sense.

Proposition 1. *Let $\mathcal{M}^\tau$ be a model of $\mathcal{L}^+$ according to Definition 7 and $P^{\preceq} \in \mathcal{P}^+$. The following hold:*

1. $\mathcal{M}^\tau \models \forall\vec{x}\, P^{\preceq}(\vec{x}, \vec{x})$
2. $\mathcal{M}^\tau \models \forall\vec{x}, \vec{y}, \vec{z}\,\big(P^{\preceq}(\vec{x}, \vec{y}) \wedge P^{\preceq}(\vec{y}, \vec{z}) \to P^{\preceq}(\vec{x}, \vec{z})\big)$
3. $\mathcal{M}^\tau \models \forall\vec{x}, \vec{y}\,\big(P^{\preceq}(\vec{x}, \vec{y}) \vee P^{\preceq}(\vec{y}, \vec{x})\big)$

Given this translation τ on formulas and models, we obtain the following characterization of $\mathcal{L}_N$ through $\mathcal{L}^+$.

Theorem 1. *Let ϕ and $\mathcal{M}$ be a wff and a model of $\mathcal{L}_N$, respectively, and let v be a variable map. Then:*

$$\mathcal{M}, v \models \phi \text{ iff } \mathcal{M}^\tau, v \models \phi^\tau$$

Proof. The proof follows by induction on the construction of ϕ. We only present the step for pfcs:

Consider $\phi = \{P(\vec{y}), \psi(\vec{y})\}(\vec{t})$ and that the Induction Hypothesis (IH) holds for ψ. We have that $\mathcal{M}, v \models \phi$ iff:

$$\vec{t}^{\,\mathcal{I},v} \in \min(\preceq_{P^{\mathcal{I}}}, \psi(\vec{y})^{\mathcal{M},v})$$

Let us also assume that $\vec{y}$ is a tuple of n variables. By definition then:

1. $\vec{t}^{\,\mathcal{I},v} \in \{\vec{d} \in \mathcal{D}^n \mid \mathcal{M}, v \models \psi(\vec{y}/\vec{d})\}$
2. $\forall\vec{e} \in \mathcal{D}^n$ if $\mathcal{M}, v \models \psi(\vec{y}/\vec{e})$ then $\vec{t}^{\,\mathcal{I},v} \preceq_{P^{\mathcal{I}}} \vec{e}$

By (IH) and the definition of $\mathcal{M}^\tau$ then we also have:

3. $\vec{t}^{\,\mathcal{I}^\tau,v} \in \{\vec{d} \in (\mathcal{D}^\tau)^n \mid \mathcal{M}^\tau, v \models \big(\psi(\vec{y}/\vec{d})\big)^\tau\}$
4. $\forall\vec{e} \in (\mathcal{D}^\tau)^n$ if $\mathcal{M}^\tau, v \models \big(\psi(\vec{y}/\vec{e})\big)^\tau$ then $(\vec{t}^{\,\mathcal{I}^\tau,v}, \vec{e}) \in (P^{\preceq})^{\mathcal{I}^\tau}$

From 3. it immediately follows that:

5. $\mathcal{M}^\tau, v \models \big(\psi(\vec{y}/\vec{t})\big)^\tau$

Next, let $\vec{z}$ be a random tuple of variables and let:

6. $\mathcal{M}^\tau, v \models \big(\psi(\vec{y}/\vec{z})\big)^\tau$

Let us also assume that $v(\vec{z}) = \vec{e} \in (\mathcal{D}^\tau)^n$. It immediately follows:

7. $\mathcal{M}^\tau, v \models \big(\psi(\vec{y}/\vec{e})\big)^\tau$

From 4., 7. and the fact that $\vec{e} = v(\vec{z})$ then $(\vec{t}^{\,\mathcal{I}^\tau,v}, v(\vec{z})) \in (P^{\preceq})^{\mathcal{I}^\tau}$ which is equivalent to:

8. $\mathcal{M}^\tau, v \models P^{\preceq}(\vec{t}, \vec{z})$

From 6., 8. and the fact that $\vec{z}$ was a random tuple, we have that:

9. $\mathcal{M}^\tau, v \models \forall\vec{z}\Big(\big(\psi(\vec{y}/\vec{z})\big)^\tau \to P^{\preceq}(\vec{t}, \vec{z})\Big)$

Finally, 5. and 9. give:

$$\mathcal{M}^\tau, v \models \big(\psi(\vec{y}/\vec{t})\big)^\tau \wedge \forall\vec{z}\Big(\big(\psi(\vec{y}/\vec{z})\big)^\tau \to P^{\preceq}(\vec{t}, \vec{z})\Big)$$

By definition of τ then we get $\mathcal{M}^\tau, v \models \big(\{P(\vec{y}), \psi(\vec{y})\}(\vec{t})\big)^\tau$, i.e., $\mathcal{M}^\tau, v \models \phi^\tau$. The reverse procedure gives the other direction as well: if $\mathcal{M}^\tau, v \models \phi^\tau$ we end up with 3. and 4. which, by (IH) and the definition of $\mathcal{M}^\tau$, are equivalent to $\mathcal{M}, v \models \phi$. □

Through this characterization result we can move from $\mathcal{L}_N$ into $\mathcal{L}^+$ and use the known machinery of standard FOL when evaluating formulas in $\mathcal{L}_N$.

5.2 Properties

We now examine some properties of our logic, starting with the fact that we can reason about defaults directly within our framework. A representative example of this property is showcased in the next proposition.

Proposition 2. *Let $\Phi = \{\phi_1, \phi_2, \phi_3\}$ be a KB where:*

1. $\phi_1 = \forall x\,\big(P(x) \to B(x)\big)$
   "All penguins are birds"
2. $\phi_2 = \forall x\,\big(\{F(y), B(y)\}(x) \to F(x)\big)$
   "Birds normally fly"
3. $\phi_3 = \forall x\,\big(\{F(y), P(y)\}(x) \to \neg F(x)\big)$
   "Penguins normally do not fly"

Furthermore, consider the following sentence:

4. $\psi = \forall x\,\big(P(x) \to \neg\{F(y), B(y)\}(x)\big)$
   "Penguins are not normal birds with respect to flying"

Then $\Phi \models \psi$ is derivable in $\mathcal{L}_N$.

This is quite an important characteristic of $\mathcal{L}_N$ since reasoning about defaults within the logic is not possible with many other approaches, e.g. default logic or circumscription.

Next, we move on to compare the logic $\mathcal{L}_N$ to the well-known KLM systems (Kraus, Lehmann, and Magidor 1990; Lehmann and Magidor 1992). We start by noting that, like most of the approaches that employ preference orderings (either between worlds or between elements of a domain), KLM rely on a single ordering. This is in contrast to our multiple orderings and the fact that we can use multiple pfcs, each one associated with a different ordering, inside the same expression. Consider, e.g., the following instance of the KLM postulate of Right Weakening: from $\models Fly \to Mobile$ and $Bird \mid\!\sim Fly$ infer $Bird \mid\!\sim Mobile$. The way we would express this instance of RW in $\mathcal{L}_N$ would be the following:

$$\forall x\left(F(x) \to M(x)\right) \wedge \forall x\left(\{F(y), B(y)\}(x) \to F(x)\right) \to \forall x\left(\{M(y), B(y)\}(x) \to M(x)\right)$$

where F abbreviates Fly, M abbreviates $Mobile$ and B abbreviates $Bird$. This formula is not valid in our logic since the two pfcs refer to two different orderings (corresponding to Fly and $Mobile$). The same holds for any other KLM postulate apart from Reflexivity. This is because we have not imposed any relationship between the different orderings or attempted to combine them in any way. We could impose, e.g., the following condition between two orderings:

whenever $\forall\vec{x}\,(P(\vec{x}) \to Q(\vec{x}))$ *we also have that* $\preceq_{P^{\mathcal{I}}} \subseteq \preceq_{Q^{\mathcal{I}}}$

which would make the previous formula valid in our logic. One could propose such restrictions in our models (and more specifically in the set $\mathcal{O}$) but this is not our intention here. However, if we introduce a new predicate symbol G that corresponds to a *global* ordering we get the following.

Proposition 3. *The KLM postulates articulated using only the (global) ordering associated with predicate G are valid in $\mathcal{L}_N$.*

It immediately follows that our approach is at least as expressive as the KLM systems.

Corollary 1. *Any proof wrt the KLM systems can be transformed into a proof in $\mathcal{L}_N$.*

We end this section by presenting some properties of pfcs in the next proposition, with the names suggesting similar properties that have appeared in the literature.

Proposition 4. *The following formulas are valid in $\mathcal{L}_N$:*

1. *REF:* $\forall\vec{x}\left(\{P(\vec{y}), \phi(\vec{y})\}(\vec{x}) \to \phi(\vec{y}/\vec{x})\right)$
2. *RCE:* $\forall\vec{x}\,(\phi(\vec{x}) \to \psi(\vec{x})) \to \forall\vec{x}\left(\{P(\vec{y}), \phi(\vec{y})\}(\vec{x}) \to \psi(\vec{y}/\vec{x})\right)$
3. *LLE/RCEC:* $\forall\vec{x}\,(\phi(\vec{x}) \equiv \psi(\vec{x})) \to \forall\vec{x}\left(\{P(\vec{y}), \phi(\vec{y})\}(\vec{x}) \equiv \{P(\vec{y}), \psi(\vec{y})\}(\vec{x})\right)$
4. *AND:* $\forall\vec{x}\left(\{P(\vec{y}), \phi(\vec{y})\}(\vec{x}) \wedge \{P(\vec{y}), \psi(\vec{y})\}(\vec{x}) \to \{P(\vec{y}), (\phi \wedge \psi)(\vec{y})\}(\vec{x})\right)$
5. *OR:* $\forall\vec{x}\left(\{P(\vec{y}), (\phi \vee \psi)(\vec{y})\}(\vec{x}) \to \{P(\vec{y}), \phi(\vec{y})\}(\vec{x}) \vee \{P(\vec{y}), \psi(\vec{y})\}(\vec{x})\right)$

6 Default Inference in $\mathcal{L}_N$

To this point, in presenting $\mathcal{L}_N$, we have dealt with a monotonic formalism. We now examine nonmonotonic reasoning in $\mathcal{L}_N$ and explore how default inferences can be obtained. Our investigations are still preliminary and are on the semantic level, i.e., we work with models. Nevertheless, a syntactic approach is also in the works and employs an extension of the Closed World Assumption for nonmonotonic reasoning. The goal then will be to provide a correspondence between the two approaches (syntactic and semantic). We present the latter here which, similar to (McCarthy 1980; Shoham 1987; Kraus, Lehmann, and Magidor 1990), employs *preferences* between the models of $\mathcal{L}_N$.

We start by restricting our models a bit further so that they only contain orderings without infinitely-ascending chains of elements, i.e., our orderings are "upwards" well-founded as well. We then proceed with the following definitions.

Definition 8. *Let $\mathcal{M}$ be a model of $\mathcal{L}_N$ and $\preceq_r \in \mathcal{O}$. The set $\min_k(\preceq_r)$ is defined inductively as follows:*

1. $\min_1(\preceq_r) = \{\vec{d} \in \mathcal{D}^n \mid \forall\vec{e} \in \mathcal{D}^n : \vec{d} \preceq_r \vec{e}\}$
2. $\min_{k+1}(\preceq_r) = \{\vec{d} \in \mathcal{D}^n \mid \forall\vec{e} \in \mathcal{D}^n \setminus \bigcup_{n=1}^{k} \min_n(\preceq_r) : \vec{d} \preceq_r \vec{e}\}$

Intuitively, the set $\min_k(\preceq_r)$ denotes the k-th least set of $\preceq_r$-equivalent elements in the ordering $\preceq_r$. Using these sets, we can define a preference between orderings on the same relation r as follows.

Definition 9. *Let $\mathcal{M} = \langle\mathcal{D}, \mathcal{I}, \mathcal{O}\rangle$ and $\mathcal{M}' = \langle\mathcal{D}, \mathcal{I}', \mathcal{O}'\rangle$ be two models of $\mathcal{L}_N$ with $\preceq_r \in \mathcal{O}$ and $\preceq'_r \in \mathcal{O}'$. We say that $\preceq_r$ is lexicographically preferred to $\preceq'_r$ iff $\exists n \in \mathbb{N}$ such that:*

1. $\min_k(\preceq'_r) = \min_k(\preceq_r)\ \ \forall\, k \in \{1, \ldots, n-1\}$
2. $\min_n(\preceq'_r) \subset \min_n(\preceq_r)$

Given the lexicographic preference between two orderings, we can now generalize our definition to a preference between models.

Definition 10. *Let $\mathcal{M} = \langle\mathcal{D}, \mathcal{I}, \mathcal{O}\rangle$ and $\mathcal{M}' = \langle\mathcal{D}, \mathcal{I}', \mathcal{O}'\rangle$ be two models of $\mathcal{L}_N$. We say that $\mathcal{M}$ is preferred to $\mathcal{M}'$, viz. $\mathcal{M} < \mathcal{M}'$, iff for every $P \in \mathcal{P}$ we have that $\preceq_{P^{\mathcal{I}}}$ is lexicographically preferred to $\preceq'_{P^{\mathcal{I}'}}$.*

Next, we define the *minimal* models of a KB, which will be our main tool for drawing default inferences.

Definition 11. *Let Φ and $\mathcal{M}$ be a KB and a model of $\mathcal{L}_N$, respectively. $\mathcal{M}$ is a minimal model of Φ iff $\mathcal{M}$ is a model of Φ and there is no model $\mathcal{M}'$ of Φ such that $\mathcal{M}' < \mathcal{M}$.*

Using the above, the next definition shows how to obtain default inferences in $\mathcal{L}_N$.

Definition 12. *Let Φ and ϕ be a KB and a sentence of $\mathcal{L}_N$, respectively. We say that Φ entails ϕ by default iff $\mathcal{M} \models \phi$ for all minimal models $\mathcal{M}$ of Φ.*

Before moving to a final example we note that, since the orderings $\preceq_r$ are well-founded in both directions, the following proposition holds.

Proposition 5. *Let $\mathcal{M}$ be a model of $\mathcal{L}_N$ and $\preceq_r \in \mathcal{O}$. Then:*

1. *either* $\forall k \in \mathbb{N}\ \min_k(\preceq_r) \neq \emptyset$
2. *or* $\exists n \in \mathbb{N}$:
   - $\forall k \in \{1, \ldots, n\}\ \min_k(\preceq_r) \neq \emptyset$
   - $\forall k > n\ \min_k(\preceq_r) = \emptyset$

This means that the sets $\min_k(\preceq_r)$, the preference between two orderings, and the preference between two models are all well-defined. Furthermore, a KB has a model iff it has a minimal one and, similar to monotonic inferences, inconsistent KBs entail all sentences by default. We conclude this section with a showcase of how *specificity*, *inheritance* and *irrelevance*, the three principles that we highlighted in Section 2, are handled in $\mathcal{L}_N$.

Corollary 2. *Let $\Phi = \{\phi_i \mid 1 \leq i \leq 6\}$ be a KB where:*

1. $\phi_1 = B(Tweety) \wedge Y(Tweety)$
   "Tweety is a yellow (Y) bird"
2. $\phi_2 = P(Opus)$
   "Opus is a penguin"
3. $\phi_3 = \forall x\big(P(x) \rightarrow B(x)\big)$
   "All penguins are birds"
4. $\phi_4 = \forall x\big(\{F(y), B(y)\}(x) \rightarrow F(x)\big)$
   "Birds normally fly"
5. $\phi_5 = \forall x\big(\{W(y), B(y)\}(x) \rightarrow W(x)\big)$
   "Birds normally have wings (W)"
6. $\phi_6 = \forall x\big(\{F(y), P(y)\}(x) \rightarrow \neg F(x)\big)$
   "Penguins normally do not fly"

Then Φ entails the following sentences by default:

1. $\psi_1 = \neg F(Opus)$
   "Opus does not fly"
2. $\psi_2 = W(Opus)$
   "Opus has wings"
3. $\psi_3 = F(Tweety) \wedge W(Tweety)$
   "Tweety flies and has wings"

So Opus, being both a bird and a penguin, is concluded to not fly by ψ_1 (specificity) but to have wings by ψ_2 (inheritance) since it is an exceptional bird wrt flying but inherits any other typical property of birds. Then ψ_3 (irrelevance) shows that Tweety, being a yellow bird, is still concluded to fly and have wings since being yellow is irrelevant wrt those two properties.

7 Related and Future Work

7.1 Related Work

Our view that normality is relative to a property such as fly was anticipated by work in circumscription, in particular in its use of Ab predicates (McCarthy 1986). (Otherwise the approaches have little in common.)

Conditional approaches to assertions of normality are generally propositional; first-order approaches include (Delgrande 1998; Kern-Isberner and Thimm 2012). A predecessor to our work, in a full first-order setting, is Brafman's (1997) approach to conditional statements. There, conditional statements of the form "*if ϕ then normally ψ*" are written as "$\phi \rightarrow_{\vec{x}} \psi$" with the intuition being that the minimal tuples of the domain that make ϕ true also make ψ true. There are two main differences between the "$\phi \rightarrow_{\vec{x}} \psi$" notation and our corresponding "$\forall \vec{x}\ \big(\{P(\vec{y}), \phi(\vec{y})\}(\vec{x}) \rightarrow \psi\big)$" notation that make the latter more expressive.

The first difference comes from the fact that we employ multiple orderings, which gives a more nuanced approach. In (Brafman 1997) it is not possible to have an individual that is normal in some respect (say, nest building) while abnormal in another (like flying).

Secondly, our approach allows more expressive formulas, as we have seen in the sequence of examples in Section 4. As well, consider the "adults are normally employed at a company" example, which Brafman would write as:

$$Adult(x) \rightarrow_x \exists y\ \big(EmployedAt(x, y) \wedge Company(y)\big)$$

Such a conditional does not seem to capture accurately the meaning of the original expression, as argued in Section 2.

However, we are able to capture Brafman's approach in ours, provided the formula ϕ of "$\phi \rightarrow_{\vec{x}} \psi$" has free variables only among $\vec{x}$ and there are no iterated occurrences of "$\rightarrow_{\vec{x}}$". More precisely, we can consider the class of models in our approach in which there is a single ordering for each arity n, say $\preceq_{\mathcal{U}_n}$. Then, we can use the formula:

$$\forall \vec{y}\ \big(\{\mathcal{U}_n(\vec{x}), \phi(\vec{x})\}(\vec{y}) \rightarrow \psi(\vec{x}/\vec{y})\big)$$

to represent Brafman's assertion $\phi(\vec{x}) \rightarrow_{\vec{x}} \psi$. Furthermore, combining this translation with the method described in Section 5.1 implies that the approach of (Brafman 1997) can also be expressed in standard FOL.

More recently there has been work in Description Logic (Baader et al. 2007) that deals with the representation and reasoning of defeasible assertions. The literature on so-called *defeasible DLs* is large and most of the established approaches to nonmonotonic reasoning (like default logic, circumscription or the rational closure) have been adapted for the DL setting; see for instance (Baader and Hollunder 1992; Bonatti, Lutz, and Wolter 2009; Giordano et al. 2015). Nevertheless, there have recently been interesting new proposals that relate to our work here.

First, driven by the need to overcome problems like the inheritance of properties in the presence of exceptions, multiple orderings have also been considered in (Gliozzi 2016; Giordano and Gliozzi 2019) to account for different rankings between individuals, each corresponding to a particular aspect (like Fly or $BuildNest$). However, although multiple orderings are considered in the semantics, only one "typicality" (in practice *minimality*) operator is employed in the syntax and there is no corresponding syntactic construct like our pfcs. Furthermore, their multiple orderings are employed only among individuals (and not tuples) and the use of typicality operators is limited, being only allowed on the left side of a subsumption axiom. This results in an interesting, but less expressive, representation of defaults, as opposed to that developed here.

Similar to the previous approach, but closer to ours, is the work in (Gil 2014), where the author takes into account multiple typicality operators. This work however suffers from similar limitations regarding the scope of the orderings and the limited employment of these typicality operators. A further limitation is the lack of any association between its operators/orderings and any relations or aspects.

A third line of work, originating from an approach in (Britz and Varzinczak 2016) to define orderings not only among individuals but also among tuples, culminated in interesting recent developments regarding defeasible reasoning in DLs (Varzinczak 2018; Britz and Varzinczak 2019). An important characteristic of this work is that the orders on individuals are derived from the ones specified by the roles, i.e., they do not correspond to any concepts like in the previously mentioned (and our) work. This results in (contextual) defeasible subsumption needing specific role names to be subscripted in order to specify the origin of the order that will be employed, something that we do within the language by means of the pfcs.

7.2 Future Work

One goal in our work is to extend the approach to a DL setting. In the following we present some preliminary ideas behind such an extension. Consider again the assertion "birds fly" which in a (non-defeasible) DL language is expressed by the concept inclusion $Bird \sqsubseteq Fly$, whereas in a defeasible DL it could be expressed as $\mathsf{T}(Bird) \sqsubseteq Fly$ or $Bird \precsim Fly$ among others. We propose to express the same concept inclusion, perhaps through some extended syntax, in such a way that its structure will invoke the use of the pfcs from our setting. In this specific assertion, e.g., the "new" DL expression would be semantically equivalent to the $\mathcal{L}_N$-formula:

$$\forall x\,\big(\{Fly(y), Bird(y)\}(x) \to Fly(x)\big) \tag{8}$$

That is, we are interested in the minimal elements that satisfy the left side of the inclusion, similar to some of the aforementioned DL approaches, while also specifying the preference ordering we want to employ. This means that the domain of any given interpretation would once again be enhanced with preference orderings and the new syntax would somehow indicate the preference ordering that is used inside a (default) concept inclusion. In other words, whereas the inclusion $Bird \sqsubseteq Fly$ would be interpreted as $Bird^{\mathcal{I}} \subseteq Fly^{\mathcal{I}}$, its default version would translate into and follow the same semantics of Equation 8, being instead interpreted (roughly) as $\min(\preceq_{Fly^{\mathcal{I}}}, Bird^{\mathcal{I}}) \subseteq Fly^{\mathcal{I}}$.

As for more complex and ambiguous statements, consider the "undergraduate students attend undergraduate courses" example that we saw at the end of Section 4. No approach in the literature can adequately handle the various interpretations we gave in Section 4, especially in a DL setting. Our goal then for future work is to try and express these interpretations in DL terms in a way that would semantically correspond to the intended formulas of $\mathcal{L}_N$. Whereas the question of how to syntactically express such assertions is certainly non-trivial, we believe that the current framework could provide a basis for (more elaborately) dealing with defeasibility in DLs. The biggest advantage perhaps will be that the properties we presented, both the KLM postulates as well as defeasible principles like specificity, inheritance and irrelevance, will continue to hold in any DL language (consider, e.g., Corollary 2 adapted for such a DL). The combination of employing multiple orderings in the domain of any interpretation together with using $\mathcal{L}_N$ and its pfcs to interpret the new default concept inclusions seems to overcome the difficulties of the established approaches as well as allow a more "informed" representation of defaults in any DL language.

Apart from DLs, we plan to expand on this work in the future in a number of directions. First, a natural extension would be to allow quantifying into a pfc. This would allow an assertion such as "each elephant normally likes its keeper", which is somewhat different from our previous example. Moreover, by allowing quantifying into a pfc, we would be able to encode *nested* default assertions, such as "profs that (normally) give good lectures are (normally) liked by their students". Second, we plan to allow for complex expressions in a pfc, and so allow a predicate expression in place of P in $\{P(\vec{y}), \phi(\vec{y})\}$. Last, as we already mentioned in Section 6, a thorough treatment and examination of nonmonotonic reasoning in $\mathcal{L}_N$ is also in the works.

8 Conclusion

We have presented a new and well-behaved approach to representing default assertions through an expressive language and novel formalism. This approach takes the position that normality is not an absolute characteristic of an individual, but instead is relative to a property (or, in general, relation). This is achieved via an extension to the language of FOL, along with an enhancement to models in FOL; a subsequent result however shows that the approach may be embedded in standard FOL. The approach allows for a substantially more expressive language for representing default information than previous approaches. Moreover, we show that the approach possesses quite natural and desirable features and satisfies the standard KLM properties. With a variety of future directions and promising possible applications, like the one we briefly discussed for the DL setting, we believe the current framework presents an interesting new approach to representing and reasoning about defaults as well as obtaining "well-behaved" nonmonotonic reasoning in general.

Acknowledgements

We thank the reviewers for their helpful comments. Financial support was gratefully received from the Natural Sciences and Engineering Research Council of Canada.

References

Baader, F., and Hollunder, B. 1992. Embedding defaults into terminological knowledge representation formalisms. In Nebel, B.; Rich, C.; and Swartout, W., eds., *Proceedings of the Third International Conference on the Principles of Knowledge Representation and Reasoning*, 306–317.

Baader, F.; Calvanese, D.; McGuinness, D.; Nardi, D.; and Patel-Schneider, P., eds. 2007. *The Description Logic Handbook*. Cambridge: Cambridge University Press, second edition.

Bonatti, P. A.; Lutz, C.; and Wolter, F. 2009. The complexity of circumscription in description logic. *Journal of Artificial Intelligence Research* 35:717–773.

Boutilier, C. 1994. Conditional logics of normality: A modal approach. *Artificial Intelligence* 68(1):87–154.

Brafman, R. I. 1997. A first-order conditional logic with qualitative statistical semantics. *Journal of Logic and Computation* 7(6):777–803.

Britz, K., and Varzinczak, I. J. 2016. Introducing role defeasibility in description logics. In *Logics in Artificial Intelligence - 15th European Conference, JELIA*, 174–189.

Britz, K., and Varzinczak, I. 2019. Contextual rational closure for defeasible $\mathcal{ALC}$. *Annals of Mathematics and Artificial Intelligence* 87(1-2):83–108.

Delgrande, J. 1987. A first-order conditional logic for prototypical properties. *Artificial Intelligence* 33(1):105–130.

Delgrande, J. 1998. On first-order conditional logics. *Artificial Intelligence* 105(1-2):105–137.

Enderton, H. 1972. *A Mathematical Introduction to Logic*. Academic Press.

Fariñas del Cerro, L.; Herzig, A.; and Lang, J. 1994. From ordering-based nonmonotonic reasoning to conditional logics. *Artificial Intelligence* 66(2):375–393.

Gil, O. F. 2014. On the non-monotonic description logic alc+t_{min}. *CoRR* abs/1404.6566.

Giordano, L., and Gliozzi, V. 2019. Reasoning about exceptions in ontologies: An approximation of the multipreference semantics. In Kern-Isberner, G., and Ognjanovic, Z., eds., *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU*, volume 11726, 212–225. Springer.

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

Gliozzi, V. 2016. Reasoning about multiple aspects in rational closure for DLs. In Adorni, G.; Cagnoni, S.; Gori, M.; and Maratea, M., eds., *AI*IA 2016: XVth International Conference of the Italian Association for Artificial Intelligence*, volume 10037 of *Lecture Notes in Computer Science*, 392–405. Genova, Italy: Springer.

Kern-Isberner, G., and Thimm, M. 2012. A ranking semantics for first-order conditionals. In *Proceedings of the European Conference on Artificial Intelligence*, 456–461. IOS Press.

Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44(1-2):167–207.

Lamarre, P. 1991. S4 as the conditional logic of nonmonotonicity. In *Proceedings of the Second International Conference on the Principles of Knowledge Representation and Reasoning*, 357–367.

Lehmann, D., and Magidor, M. 1992. What does a conditional knowledge base entail? *Artificial Intelligence* 55(1):1–60.

McCarthy, J. 1980. Circumscription – a form of nonmonotonic reasoning. *Artificial Intelligence* 13:27–39.

McCarthy, J. 1986. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence* 28:89–116.

Mendelson, E. 2015. *Introduction to Mathematical Logic*. CRC Press, 6th edition.

Moore, R. 1985. Semantical considerations on nonmonotonic logic. *Artificial Intelligence* 25:75–94.

Reiter, R., and Criscuolo, G. 1981. On interacting defaults. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 270–276.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence* 13(1-2):81–132.

Shoham, Y. 1987. A semantical approach to nonmonotonic logics (extended abstract). In *Symposium on Logic in Computer Science*, 275–279.

Varzinczak, I. 2018. A note on a description logic of concept and role typicality for defeasible reasoning over ontologies. *Logica Universalis* 12(3-4):297–325.