# A First-order Conditional Logic with Qualitative Statistical Semantics

RONEN I. BRAFMAN, *Computer Science Department, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada.*
*E-mail: brafman@cs.ubc.ca*

## Abstract

We define a first-order conditional logic in which conditionals, such as $\alpha \rightarrow \beta$, are interpreted as saying that normal/common/typical objects which satisfy $\alpha$ satisfy $\beta$ as well. This qualitative 'statistical' interpretation is achieved by imposing additional structure on the domain of a *single* first-order model in the form of an ordering over domain elements and tuples. $\alpha \rightarrow \beta$ then holds if all objects with property $\alpha$ whose ranking is minimal satisfy $\beta$ as well. These minimally ranked objects represent the typical or common objects having the property $\alpha$. This semantics differs from that of the more common subjective interpretation of conditionals, in which conditionals are interpreted over *sets* of standard first-order structures. Our semantics provides a more natural way of modelling qualitative statistical statements, such as 'typical birds fly', or 'normal birds fly'. We provide a sound and complete axiomatization of this logic, and we show that it can be given probabilistic semantics.

*Keywords*: Conditional logic, first-order logic, qualitative statistical reasoning, nonmonotonic logic.

## 1 Introduction

Conditional logics have been the focus of much AI research in recent years because of their important connection to default reasoning and belief revision. Most work on conditional logic has concentrated on the propositional case, for example [18, 6, 15, 16]. First-order conditional logics have received less attention, and there does not seem to be agreement on their formulation. This is understandable given that first-order logic is considerably more expressive than propositional logic. This additional expressivity requires one to deal with issues that do not arise in the propositional case, such as the interaction between quantifiers and the conditional operator.

The language of propositional conditional logics contains, in addition to the standard Boolean operators, a binary conditional operator $\rightarrow$. These logics are interpreted over a structure that consists of an ordered set $S$ of standard propositional models. The formula $\alpha \rightarrow \beta$ is satisfied by such a structure if among standard propositional models in $S$ that satisfy $\alpha$, all those that are minimal with respect to (abbreviated w.r.t.) the ordering satisfy $\beta$ as well. While this is just a rough description, for instance, the ordering may be relative to each world or partial, it captures the essence of the semantics. This semantics justifies an intuitive reading of $\alpha \rightarrow \beta$ as 'in the most normal $\alpha$ worlds, $\beta$ holds'.

A natural generalization of this idea to first-order conditional logic is to use a similar interpretation over an ordered set of standard first-order models. There are various issues that this interpretation raises. For instance, should the domain of all these models be the same? Or, should constants be interpreted as rigid designators? Regardless of these choices, the essence of the interpretation is the same, and as before, an informal reading of the conditional $\alpha \rightarrow \beta$ would be 'in the most normal $\alpha$ worlds, $\beta$ holds'. The main difference is that now, $\alpha$ and $\beta$

 

are part of a much richer language.

However, the structure of first-order models provides an alternative interpretation for conditionals. Roughly, we can think of a first-order model as a collection of propositional structures, each describing the properties of a single object within the domain (or more generally, the properties of some particular tuple of objects). Thus, another possible interpretation of first-order conditional logics is w.r.t. a *single* first-order structure that is supplemented with an ordering over *domain elements*. Now, $\alpha \rightarrow \beta$ is interpreted as saying that 'all objects in the structure's domain that are normal w.r.t. the property $\alpha$, satisfy $\beta$ as well'.

There are different manners in which one can enhance a first-order structure. In particular, our choice of an ordering relation over the structure's domain is not the only one. For example, Schlechta's work on defaults as generalized quantifiers [20], developed independently,[1] adds weak filters to a first-order structure to obtain an interpretation of defaults (or conditionals) that is similar in flavour to ours. We discuss the view of first-order conditionals as generalized quantifiers in Section 4 and Schlechta's work in Section 5.

The two approaches to first-order conditional logics outlined above closely resemble two interpretations of probabilistic statements: the frequentist interpretation and the subjectivist interpretation. While these interpretations have a long history, two recent works on probabilistic logic brought them to the attention of the formal reasoning community in AI. Bacchus [4] and Halpern [14] examine two types of statements about probabilities: statements about subjective beliefs, such as 'the probability that Tweety flies is 0.9', and statistical statements, such as '90% of birds fly'. They suggest that the first type of statement is naturally modeled by a probability measure over a possible-worlds structure, while the second statement can be interpreted by imposing a probability assignment over a single domain. The choice between these interpretations is akin to the choice between the two possible interpretations of first-order conditionals. Indeed, conditional sentences can be viewed as qualitative counterparts of probabilistic sentences. This is a consequence of the probabilistic semantics of conditionals provided by Adams [1] and by Goldszmidt and Pearl [13]. According to this semantics, and depending on which interpretation of probabilities one chooses, $\alpha \rightarrow \beta$ can be roughly understood as a statement of subjective belief, i.e. 'in those $\alpha$ worlds I consider most likely, $\beta$ is the case', or as qualitative statistical statements, i.e. 'in the actual world, most objects with property $\alpha$ have property $\beta$ as well'.

In this paper, we formalize this second, statistical interpretation of first-order conditional logics. We interpret conditionals over a standard first-order structure to which we add an ordering over domain elements. More precisely, for every natural number $n$, there is a total pre-order over the set of $n$-tuples of domain elements. A conditional of the form $\varphi \rightarrow_{(\bar{x})} \psi$ is interpreted as saying that all the minimal tuples $\bar{c}$ of length $|\bar{x}|$ that satisfy $\varphi$ when $\bar{c}$ is substituted for $\bar{x}$, must also satisfy $\psi$ under this substitution. A natural extension of this semantics imposes restrictions on the possible pre-orders and on the relationship between pre-orders of tuples of different lengths. We explore three variants of this semantics and show that the most powerful of them can be given a probabilistic semantics akin to Goldszmidt and Pearl's $\epsilon$-semantics [13].

Our work is closely related to Bacchus's work on representing statistical information in first-order logic [4]. Our logic can be thought of as a qualitative counterpart to Bacchus's quantitative approach. Indeed, the form of quantification we use resembles Bacchus' notation, and some of our axioms can be viewed as qualitative counterparts of similar axioms that appear in his work. However, an important difference between the two logics is their expressive

[1] The fundamental ideas of our work first appeared in [7].

power. Bacchus's logic has numbers as objects in the language, and he allows quantification over them, addition, multiplication, and other operations. Our counterpart of probabilities, ranks, are not objects in the language, and we do not manipulate them within the logic.

We believe that making the distinction between the subjectivist and statistical (frequentist) interpretations is as important in qualitative reasoning as it is important in quantitative reasoning. We regularly reason with qualitative statistical information and it is important to distinguish it from subjective beliefs. Qualitative statistical information summarizes our experience in the world. For instance, rather then memorize various instances of flying birds and non-flying birds, we may store our knowledge in the form of a qualitative statistical statement—'common/typical birds fly'. Although this statement may give rise to subjective beliefs about whether or not a particular bird flies, it is not about our subjective beliefs. It is a qualitative assertion about the ratio of birds and flying birds or about the properties of prototypical birds, and it should be interpreted accordingly.

Once the distinction between subjective and statistical conditionals is understood, it can help us approach problems such as the lottery paradox. It is perfectly consistent to state that typical/common tickets will not win the lottery, yet some ticket will win. This has been called a paradox because in some subjective conditional logics, e.g. Delgrande's [9], the following theory is inconsistent:

$$\forall x(true \rightarrow \neg Winner(x)) \ , \ true \rightarrow \exists x Winner(x).$$

It is easy to see why this is the case. Interpreted subjectively, the first sentence says that in our most normal worlds, all tickets will not win, while the second sentence claims that in our most normal worlds a winner exists. This, of course, is a contradiction. Moreover, the first statement does not seem to capture our information about lotteries. The statistical statement that most tickets will not win the lottery is a more plausible one. More importantly, we see that the two interpretations are fundamentally different.

Another example of a theory that is inconsistent under a subjective interpretation is the following:

$$\forall x, y(Pet(x, y) \rightarrow (Dog(y) \vee Cat(y))) \ , \ \forall y(Pet(John, y) \rightarrow (Snake(y))).$$

The subjective interpretation of this theory is that in our most normal worlds, everybody's pets are either dogs or cats, while John's pets are snakes. But clearly, John is someone, so that, based on the first sentence, his pets should be dogs or cats. A statistical interpretation makes more sense here. Under such an interpretation, we would understand these formulas as saying that most/typical people's pets are dogs or cats, but John's typical pets are snakes. As we shall see, both statements are consistent under the qualitative statistical semantics we present.

In Section 2 we provide a formal account of the semantics of first-order statistical conditional logic followed, in Section 3, by a description of a sound and complete axiomatization of this logic. In Section 4, we look at different syntactic and semantic variants of this logic. In particular, we show that our conditional logic can be given probabilistic semantics. We also show that the conditional operator can be replaced by a class of quantifiers or alternatively, by a set of new predicates. This latter formulation leads to a semi-standard first-order language with a much simpler axiom system. We end with a discussion of the logic and its relation with other work. The paper contains two appendices: Appendix A contains the proofs of the theorems presented in the paper, and Appendix B contains a list of symbols. Throughout this paper, we assume familiarity with the basic syntax and semantics of first-order logic.

## 2 Language and semantics

We now proceed with a description of the language of statistical conditional logic, a language which extends the language of first-order logic (FOL). This will be followed by a definition of the language's formal semantics.

We assume the existence of an underlying first-order language $\mathcal{L}$, containing (possibly infinite) sets of predicate symbols $P$, constant symbols $C$, function symbols $F$, and variables $V$. Our language, $\mathcal{L}_C$ consists of a set of well formed formulas (wff) defined below using $P, C, F$ and $V$.

DEFINITION 2.1
The set of *well formed formulas* of $\mathcal{L}_C$ is defined inductively as follows:

- Atomic formulas of $\mathcal{L}_C$ are wffs (an atomic formula is a predicate symbol with an appropriate number of terms).
- If $\varphi$ and $\psi$ are wffs, then so are $\neg\varphi, \varphi \Rightarrow \psi$.
- If $\varphi$ is a wff and $x$ is a variable, then $\forall x \varphi$ is a wff.
- If $\varphi$ and $\psi$ are wffs then $\varphi \rightarrow_{(\bar{x})} \psi$ is a wff.

We shall use the symbols $\vee, \wedge, \Leftrightarrow$ and $\exists$ freely, with the understanding that they are defined by $\neg, \Rightarrow$ and $\forall$. We shall use the shorthand $\bar{x}$ for a sequence of variables $x_1, \ldots, x_n$ for some fixed $n$, when $n$ is clear from the context. Notice that we use $\Rightarrow$ to denote material implication and a subscripted $\rightarrow$ to denote conditionals.

Intuitively, the conditional $\varphi \rightarrow_{(\bar{x})} \psi$ can be read as saying that normal, or typical tuples that can be substituted for $\bar{x}$ and have the property $\varphi$, have the property $\psi$ as well. Hence, in this formula, the variables $\bar{x}$ are implicitly quantified by the conditional operator.

Let us consider a number of examples of wffs and their intuitive interpretation. $instructor(x, y)$ holds when $x$ is $y$'s instructor, $like(x, y)$ holds when $x$ likes $y$, and $clear(x)$ holds when $x$'s lessons are clear.

1. $instructor(x, y) \rightarrow_{(x,y)} like(x, y)$. Here, $x$ and $y$ are implicitly quantified by the conditional operator. This sentence can be read as 'normal instructor/student pairs are such that the instructor likes her student'.
2. $\forall x(instructor(x, y) \rightarrow_{(y)} like(x, y))$. Here $x$ is universally quantified, while $y$ is implicitly quantified by the conditional operator. This sentence can be read as 'all instructors like their normal (or typical) students'.
3. $\forall y(instructor(x, y) \rightarrow_{(x)} like(x, y))$ can be read as 'all students are liked by their typical instructors'.
4. $(instructor(x, y) \rightarrow_{(y)} like(x, y)) \rightarrow_{(x)} clear(x)$ can be read as 'typically, instructors who are liked by their typical students are clear'.

Next, we extend the standard definition of free-variables to $\mathcal{L}_C$.

DEFINITION 2.2
We define the variables of a term as follows:

- If $t = v \in V$ then $var(t) = v$.
- If $t = f(t')$ then $var(t) = var(t')$.
- If $t = c \in C$ then $var(t) = \emptyset$.

The free variables of a wff $\psi$ (fv($\psi$)) are defined inductively:

- If $P(t_1, \ldots, t_n)$ is an atomic formula, then $\text{fv}(P(t_1, \ldots, t_n)) = \cup_{i=1}^{n} var(t_i)$.
- If $\varphi$ and $\psi$ are wffs, then fv($\neg\varphi$)=fv($\varphi$) and fv($\varphi \Rightarrow \psi$)=fv($\varphi$)$\cup$ fv($\psi$).
- If $\varphi$ is a wff and $x$ is a variable, then $\text{fv}(\forall x\varphi) = \text{fv}(\varphi) \setminus \{x\}$.
- If $\varphi$ and $\psi$ are wffs and $\bar{x}$ is a set of variables, then $\text{fv}(\varphi \rightarrow_{(\bar{x})} \psi) = [\text{fv}(\varphi)\cup\text{fv}(\psi)] \setminus \{\bar{x}\}$.

In order to define our models, we need the following:

DEFINITION 2.3
$R$ is a *ranking function* on $D$ if $R : D \mapsto \Omega$, where $\Omega$ is a totally ordered set.

We interpret the language $\mathcal{L}_C$ over a class of structures consisting of first-order models whose domain elements are ranked.

DEFINITION 2.4
A ranked first-order structure is a pair $\mathcal{M}=\langle \mathbf{M}, \mathbf{R}\rangle$

- $\mathbf{M}$ is a standard first order structure.
- $\mathbf{R} = \{R_n | n \in N\}$, where for each $n$, $R_n$ is a ranking function on $|\mathbf{M}|^n$.[2]

We shall use the shorthand notation $(a,b) =_{\mathbf{R}} (c,d)$ for $R_2((a,b)) = R_2((c,d))$ (and similarly for other relations). We shall also talk about minimal elements in $|\mathbf{M}|^n$, with the understanding that minimality is w.r.t. the ranking $R_n$.

DEFINITION 2.5
Let $s : Vars \mapsto ||M||$ be an assignment function. We define the notion of *satisfiability* of a wff $\alpha$ under $s$ in a ranked first-order structure $\mathcal{M}=\langle \mathbf{M}, \mathbf{R}\rangle$, written $\mathcal{M} \models \alpha[s]$, as follows:

- If $\alpha$ is atomic then $\mathcal{M} \models \alpha[s]$ if $\mathbf{M} \models \alpha[s]$.
- If $\alpha = \neg\beta$ then $\mathcal{M} \models \alpha[s]$ if $\mathcal{M} \not\models \beta[s]$.
- If $\alpha = \beta \Rightarrow \gamma$ then $\mathcal{M} \models \alpha[s]$ if $\mathcal{M} \not\models \beta[s]$ or $\mathcal{M} \models \gamma[s]$.
- If $\alpha = \forall x\beta$ then $\mathcal{M} \models \alpha$ if $\mathcal{M} \models \alpha[s][^x_a]$ for all $a \in |\mathbf{M}|$.
  ($[s][^x_d]$ is the same as $s$, except that it assigns $d$ to $x$.)
- If $\alpha = (\beta \rightarrow_{(\bar{x})} \gamma)$ then $\mathcal{M} \models \alpha[s]$ if for each $\bar{d} \in min_{\bar{x}}(\beta)$ we have that $\mathcal{M} \models \gamma[s][^{\bar{x}}_{\bar{d}}]$, where $min_{\bar{x}}\beta$ is the set of minimal elements of the set $\{\bar{d} \in |\mathbf{M}|^{|\bar{x}|} : \mathcal{M} \models \beta[s][^{\bar{x}}_{\bar{d}}]\}$.

That is, let $B$ be the set of all tuples $\bar{c}$ such that $\beta(\bar{x})$ is satisfied when we substitute $\bar{c}$ for $\bar{x}$. $\beta \rightarrow_{(\bar{x})} \gamma$ holds if all those tuples that are minimal members of $B$ (w.r.t. $R_{|\bar{x}|}$) are such that $\gamma$ is also satisfied when they are substituted for $\bar{x}$. For example, $bird(x) \rightarrow_{(x)} fly(x)$ would be satisfied if all those domain objects in the extension of $bird$ (often denoted by $bird^{\mathbf{M}}$) that are minimal w.r.t. $R_1$, are also in $fly^{\mathbf{M}}$.

The set $min_{\bar{x}}\beta$ can be empty even when $\exists\bar{x}\beta(\bar{x})$ holds (when the set $B$ defined above contains an infinite descending set of tuples.) However, we would like to exclude this case. Following [15], we consider only *smooth* rankings.

DEFINITION 2.6
$\mathbf{R}$ is **smooth** (w.r.t. $\mathcal{M}$) if for all $\beta$ and $\bar{x}$, $\{\bar{d} \in |\mathbf{M}|^{|\bar{x}|} : \mathcal{M} \models \beta[^{\bar{x}}_{\bar{d}}]\}$ is empty, has a minimal element, or equals $|\mathbf{M}|^{|\bar{x}|}$. We define $\mathbf{NS}$ to be the class of smooth ranked structures.

[2] Recall that $|\mathbf{M}|$ is the domain of $\mathbf{M}$.

We follow the convention that conditionals of the form $false \rightarrow_{(\bar{x})} \beta$ are always satisfied. Also, notice that a conditional such as $bird(Tweety) \rightarrow_{(x)} fly(Tweety)$ holds iff the corresponding material implication, $bird(Tweety) \Rightarrow fly(Tweety)$, holds. This is because the satisfiability of $bird(Tweety)$ and $fly(Tweety)$ does not depend on $s$. Hence, $bird(Tweety)$ is satisfied under one assignment function iff it is satisfied under all assignment functions.

So far, we have not imposed any additional requirements on the functions $R_n$. However, there are two properties which we consider quite natural for some interpretations of the ranking $\mathbf{R}$. The first requirement, *permutation* says that the ranking is indifferent to the order of the elements in the tuple. In essence, this implies that the ranking is defined over bags (or multi-sets) of domain elements, rather then over tuples. The second requirement, *concatenation*, says that preferences are closed under concatenation. That is, if we prefer $\bar{c}$ over $\bar{d}$ and $\bar{c}'$ over $\bar{d}'$, then we prefer $\bar{c} \circ \bar{c}'$ over $\bar{d} \circ \bar{d}'$. The definition has two parts, one covering the case of strict preference and another covering the case of non-strict preference.

*Permutation* and *concatenation* assert certain independence properties of the ranking $\mathbf{R}$. The former asserts that the ranking of a multi-set of objects is independent of their ordering, while the latter relates the ranking of a tuple to the ranking of its components. We believe that together, these properties imply that the ranking of a multi-set is a function of the rank of its components (see Conjecture 4.5).

DEFINITION 2.7
A ranked structure $\mathcal{M}=\langle \mathbf{M}, \mathbf{R}\rangle$ satisfies *permutation* if for every $n \in N$ and for every permutation $\pi$ over $\{1, \ldots, n\}$, we have that $R_n(d_1, \ldots, d_n) = R_n(d_{\pi(1)}, \ldots, d_{\pi(n)})$. $\mathcal{M}$ satisfies *concatenation* if $R_n(\bar{c}) \geq R_n(\bar{d})$ and $R_n(\bar{c}') \geq R_n(\bar{d}')$, then $R_n(\bar{c} \circ \bar{c}') \geq R_n(\bar{d} \circ \bar{d}')$, and if in addition $R_n(\bar{c}) > R_n(\bar{d})$ then $R_n(\bar{c} \circ \bar{c}') > R_n(\bar{d} \circ \bar{d}')$.

EXAMPLE 2.8
Consider a language with three unary predicates: $L, M, P$. Individuals with the property $L$ (respectively, $M$,$P$) are good in logic (resp. mathematics, physics). We shall loosely refer to persons with the above properties as logicians, mathematicians, and physicists. We state that $L(x) \rightarrow_{(x)} M(x)$ and $L(x) \rightarrow_{(x)} \neg P(x)$ (i.e. logicians are normally good in math and normally not good in physics). Consider the following model $\mathcal{M}$: there are four individuals, Alice, Bob, Craig, and Diane. They are ranked as follows: Alice is the most typical, Bob is second, Craig is third, and Diane is fourth. The ordering over pairs, triplets, etc., is arbitrary. $L^{\mathcal{M}} = \{\text{Bob, Craig}\}$, $M^{\mathcal{M}} = \{\text{Bob}\}$, $P^{\mathcal{M}} = \{\text{Craig,Diane}\}$. We see that $\mathcal{M} \models L(x) \rightarrow_{(x)} M(x)$ because the only minimal object with property $L$, Bob, has the property $M$. $\mathcal{M} \models L(x) \rightarrow_{(x)} \neg P(x)$ because all minimal objects with property $L$ satisfy $\neg P$. However, notice that there is a logician, Craig, that is good in physics. Hence, $\mathcal{M} \not\models \forall x(L(x) \Rightarrow \neg P(x))$. Also, notice that in $\mathcal{M}$, normally, a person is not a logician, a mathematician, or a physicist because Alice, the most typical person in the domain, is none of the above.

Now suppose we have an additional binary predicate $W(x, y)$, with the intended meaning that $x$ is wealthier than $y$. Suppose we state that $(L(x) \wedge P(y)) \rightarrow_{(x,y)} W(y, x)$. That is, typical logician/physicist pairs are such that the physicist is wealthier than the logician. We can constrain the ranking over pairs in $M$ as follows in order to satisfy this assertion: (1) (Alice,Bob), (Alice,Craig), and (Alice, Diane), (2) (Bob,Craig) and (Craig,Diane), (3) (Bob,Diane), and (4) all other ordered pairs. If $W^{\mathcal{M}} = \{$(Alice,Diane), (Diane,Craig), (Craig,Bob), (Alice,Craig), (Alice,Bob), (Diane,Bob)$\}$, then $\mathcal{M} \models (L(x) \wedge P(y)) \rightarrow_{(x,y)} W(y, x)$. This follows from the fact that the most typical logician/physicist pairs, namely (Bob,Craig) and (Craig,Diane), are within the extension of $W$.

The ranking of pairs above does not satisfy the permutation property, e.g. (Alice,Bob) and

(Bob,Alice) do not have the same ranking. Similarly, this ranking does not satisfy the concatenation property. For instance, (Bob,Diane) is less typical than (Bob,Craig), despite the fact that Craig is more normal than Diane. A ranking (of bags) that satisfies both properties is the following: (1) {Alice,Alice} (2) {Alice,Bob} (3) {Alice,Craig}, {Bob,Bob} (4) {Alice,Diane},{Bob,Craig} (5) {Bob,Diane},{Craig,Craig} (6) {Craig,Diane}.

## 3 Axiomatization

This section describes a sound and complete set of axioms for the class **NS**. We use the following shorthand notation:

$$\alpha \leq_{\bar{x}} \beta \stackrel{\text{def}}{=} \neg((\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\alpha)$$

$$\alpha <_{\bar{x}} \beta \stackrel{\text{def}}{=} (\alpha \vee \beta \rightarrow_{(\bar{x})} \neg\beta).$$

Notice that $\alpha \leq_{\bar{x}} \beta$ is read as 'some normal tuples w.r.t. the property $\alpha \vee \beta$ satisfy $\alpha$', and $\alpha <_{\bar{x}} \beta$ is read as 'all normal tuples w.r.t. the property $\alpha \vee \beta$ satisfy $\neg\beta$'. As we shall see, the former implies that those tuples that are normal for $\beta$ are no more normal than those tuples normal for $\alpha$, while the latter implies that those tuples normal for $\alpha$ are strictly more normal than those tuples normal for $\beta$. For this reason, we use the $\leq$ and $<$ symbols.

The axiom schema for our logic are displayed in Figure 1.

$$\frac{\alpha\ ,\ \alpha \Rightarrow \beta}{\beta} \qquad \textbf{Modus Ponens}$$

is the only inference rule in our logic.

Here are intuitive interpretations of these axiom schema.

**Reflexivity**—the normal $\alpha$ objects satisfy $\alpha$. Hence, normal birds are a subset of the set of birds.

**Left Equivalence**—normal objects for logically equivalent formulas have identical properties. Hence, objects can now be thought of as normal w.r.t. a set, or the property whose extension is that set.

**Right Weakening & And**—properties common to all normal objects for some formula are closed under logical implication. Hence, if small birds normally have small beaks, then **RW** implies that they normally have beaks. If they normally fly as well, then **And** implies that they normally have small beaks and fly.

**Cautious Monotony**—if all normal objects for $\alpha$ satisfy $\beta$ then the normal objects for $\alpha$ are also normal for $\alpha \wedge \beta$. For example, if Tweety is a normal bird and normal birds fly, then Tweety is a normal flying bird.

**Or**—the normal objects for $\alpha \vee \beta$ are the union of normal objects for $\alpha$ and $\beta$.

**Rational Monotony**—if normal objects of $\alpha$ are $\gamma$, but some normal objects for $\alpha \wedge \beta$ are not $\gamma$, then normal objects for $\alpha$ are not $\beta$. For example, if normal birds fly, but it is not the case that normal birds weighing over 50 kg fly, then normal birds weigh less than 50 kg.

**Weakening**—if all objects have the property $\beta$ then in particular, normal $\alpha$ objects satisfy $\beta$.

**Instantiation**—the set of normal objects for a property is nonempty unless there are no objects with this property.

**Renaming**—invariance under renaming of the bound variables.

**Interchange**—if we have a pair of normal and universal quantifiers quantifying over disjoint variables, then their order can be exchanged. For example, 'for any colour, normal children

(3.1) All instances of FOL tautologies

(3.2) $\alpha \rightarrow_{(\bar{x})} \alpha$ **(Reflexivity)**

(3.3) $\forall \bar{y}(\alpha(\bar{y}) \Leftrightarrow \beta(\bar{y})) \Rightarrow ((\alpha \rightarrow_{(\bar{x})} \gamma) \Leftrightarrow (\beta \rightarrow_{(\bar{x})} \gamma))$ **(Left Equivalence)**

(3.4) $\forall \bar{y}(\alpha(\bar{y}) \Rightarrow \beta(\bar{y})) \Rightarrow (\gamma \rightarrow_{(\bar{x})} \alpha) \Rightarrow (\gamma \rightarrow_{(\bar{x})} \beta)$ **(Right Weakening)**

(3.5) $(\alpha \rightarrow_{(\bar{x})} \beta) \wedge (\alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow (\alpha \wedge \beta \rightarrow_{(\bar{x})} \gamma)$ **(Cautious Monotony)**

(3.6) $(\alpha \rightarrow_{(\bar{x})} \beta) \wedge (\alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow (\alpha \rightarrow_{(\bar{x})} (\beta \wedge \gamma))$ **(And)**

(3.7) $(\alpha \rightarrow_{(\bar{x})} \gamma) \wedge (\beta \rightarrow_{(\bar{x})} \gamma) \Rightarrow ((\alpha \vee \beta) \rightarrow_{(\bar{x})} \gamma)$ **(Or)**

(3.8) $(\alpha \rightarrow_{(\bar{x})} \gamma) \wedge \neg(\alpha \wedge \beta \rightarrow_{(\bar{x})} \gamma) \Rightarrow \alpha \rightarrow_{(\bar{x})} \neg\beta$ **(Rational Monotony)**

(3.9) $\forall \bar{x}\beta \Rightarrow \alpha \rightarrow_{(\bar{x})} \beta$ **(Universal Weakening)**

(3.10) $(\alpha \rightarrow_{(\bar{x})} \beta) \Rightarrow (\exists \bar{x}\alpha \Rightarrow \exists \bar{x}(\alpha \wedge \beta))$ **(Instantiation)**

(3.11) $(\alpha \rightarrow_{(\bar{x})} \beta) \Rightarrow (\alpha \rightarrow_{(\bar{y})} \beta)[^{\bar{x}}_{\bar{y}}]$ **(Renaming)**
where $\bar{y}$ does not occur in $\alpha$ and $\beta$

(3.12) $\forall \bar{y}(\alpha \rightarrow_{(\bar{x})} \beta) \Leftrightarrow (\alpha \rightarrow_{(\bar{x})} \forall \bar{y}\beta)$ **(Universal Interchange)**
whenever $\bar{y} \cap (\mathrm{fv}(\alpha) \cup \bar{x}) = \emptyset$

(3.13) $(\alpha \rightarrow_{(\bar{x},\bar{y})} \beta) \Leftrightarrow (\alpha \rightarrow_{(\bar{y},\bar{x})} \beta)$ **(Permutation)**

(3.14) $(\alpha \leq_{\bar{x}} \beta) \wedge (\alpha' \leq_{\bar{y}} \beta') \Rightarrow ((\alpha \wedge \alpha') \leq_{\bar{x},\bar{y}} (\beta \wedge \beta'))$ **(Weak-Concat)**
where $\bar{y} \cap \bar{x} = \emptyset, \bar{x} \cap \mathrm{fv}(\alpha' \vee \beta') = \emptyset, \bar{y} \cap \mathrm{fv}(\alpha \vee \beta) = \emptyset$.

(3.15) $(\alpha <_{\bar{x}} \beta) \wedge (\alpha' \leq_{\bar{y}} \beta') \Rightarrow ((\alpha \wedge \alpha') <_{\bar{x},\bar{y}} (\beta \wedge \beta'))$ **(Strong-Concat)**
where $\bar{y} \cap \bar{x} = \emptyset, \bar{x} \cap \mathrm{fv}(\alpha' \vee \beta') = \emptyset, \bar{y} \cap \mathrm{fv}(\alpha \vee \beta) = \emptyset$.

(3.16) $(\alpha \wedge \alpha' \rightarrow_{(\bar{x},\bar{y})} \gamma) \Rightarrow (\alpha \rightarrow_{(\bar{x})} (\alpha' \rightarrow_{(\bar{y})} \gamma))$ **(Distribution)**
where $\bar{y} \cap \bar{x} = \emptyset\ \bar{x} \not\in \mathrm{fv}(\alpha'), \bar{y} \not\in \mathrm{fv}(\alpha)$

(3.17) $\alpha \rightarrow_{(\bar{x},\bar{y})} ((\alpha \rightarrow_{(\bar{x})} \beta) \Rightarrow \beta)$ **(Projection)**
where $\bar{y} \cap \bar{x} = \emptyset$

FIG. 1. The axiom schema

can name that colour' if and only if 'normal children can name all colours'. This shows that there are certain problems with the interpretation of the conditional $\alpha \rightarrow_{(\bar{x})} \beta$ as 'most $\alpha$s are $\beta$'. Indeed, it does not seem to be necessarily the case that if for any colour, most children can name that colour, then most children can name all colours. As we show later on, the 'most' interpretation can be justified to a certain extent via a probabilistic interpretation of conditionals. However, this latter notion of 'most' may be better represented by the expression 'virtually all'.

The role of Axiom (3.12) in the proof of the completeness theorem is worth mentioning. This axiom guarantees that the models we construct are smooth. Suppose that there is no minimal rank for tuples satisfying $\alpha$ and that the domain is infinite. We can have a model that satisfies $\forall y(\alpha \rightarrow_{(x)} \beta)$, where for each domain element $d$ there is some rank $r_d$ such that (1) there are elements $e$ whose rank is lower than $r_d$ such that $\alpha[^x_e]$ is satisfied, and (2) $(\alpha \Rightarrow \beta)[^{x,y}_{c,d}]$ is satisfied iff $c$ belongs to a rank lower than $r_d$. If we construct the model so that $\{r_d : d \in \|M\|\}$ forms an infinite descending chain, the wff $(\alpha \rightarrow_{(x)} \forall y\beta)$ will not be satisfied.

**Permutation**—we can permute the variables bound by the conditional operator.

**Weak-Concat**—if tuples normal for $\alpha$ (resp. $\alpha'$) are as normal as tuples normal for $\beta$ (resp. $\beta'$) then tuples normal for $\alpha \wedge \alpha'$ are as normal as tuples normal for $\beta \wedge \beta'$. Thus, if Alice is a normal female and Bill is a normal male then (Alice,Bill) is a normal female–male pair.

**Strong-Concat**—similar to (3.14), only with strict inequality.

**Distribution**—we can minimize components separately. Suppose that members of all normal female/male pairs look differently. Hence, it is true that given a normal female, all normal males look different from her.

**Projection**—provides another form of minimization by component. Suppose that $(c, d)$ is a normal pair w.r.t. the property $\alpha$. If all objects that together with $d$ are normal w.r.t. $\alpha$ satisfy $\beta$, then $c$ should satisfy $\beta$ as well. That is, fixing $d$ as one component of the pair, $c$ should be normal for $\alpha$ together with $d$ if $(c, d)$ is normal for $\alpha$.

We can now state the following results:

THEOREM 3.1
Axioms (3.1)–(3.12) are sound w.r.t. the class **NS**.

THEOREM 3.2
Axiom (3.13) is sound w.r.t. the class of **NS** structures satisfying *permutation* and axioms (3.14)–(3.17) are sound w.r.t. the class of **NS** structures satisfying *concatenation*.

THEOREM 3.3
For a countable language $\mathcal{L}_C$, axioms (3.1)–(3.12) are complete with respect to the class **NS**.

The proof of the completeness theorem is rather long and appears in the Appendix together with the other proofs. Our construction employs Henkin style witnesses (see [10]) which are repeatedly added to the language in order to guarantee the existence of the non-normal objects stipulated by the theory.

THEOREM 3.4
For a countable language $\mathcal{L}_C$

1. Axioms (3.1)–(3.13) are complete w.r.t. the class of **NS** structures satisfying *permutation*.
2. Axioms (3.1)–(3.17) are complete w.r.t. the class of **NS** structures satisfying *permutation* and *concatenation*.

## 4 Alternative formulations

We present a number of syntactic and semantic variants of our logic. First, we show that the logic can be given probabilistic semantics, and then we show how different syntactic constructs can be used to replace the conditional operator.

### *4.1 Probabilistic semantics*

The existence of probabilistic semantics for propositional conditional logics has been pointed out by a number of authors, for example Adams [1], Goldszmidt and Pearl [13], and Lehmann and Magidor [16]. We claim that conditional first-order theories can be given probabilistic semantics similar to that of $\epsilon$-semantics [1, 12]. The intuition behind this semantics is as follows: suppose there is a probability distribution $Pr$ defined over the domain of a first-order model **M**, and let us abuse notation and write $Pr(\varphi(x))$ for $Pr(\{d : \mathbf{M} \models \varphi(x)[^x_d]\})$. We would like to say that a conditional $\varphi \rightarrow_{(x)} \psi$ is satisfied by this model when $Pr(\psi(x)|\varphi(x))$ is almost 1. That is, with high probability, any element that has the property $\varphi$, has the property $\psi$. In order to formalize this intuition we have to say what we mean by 'almost 1' and how we would treat tuples of domain elements, as opposed to single domain elements. We resolve the first question by using the ideas of Goldszmidt *et al.* [12] in their formulation of $\epsilon$-semantics. Instead of looking at a single probability assignment, we look at a sequence of probability assignments and replace the requirement $Pr(\psi(x)|\varphi(x))$ is 'almost' 1 with $Pr(\psi(x)|\varphi(x)) = 1$

in the limit. (An alternative approach is to use non-standard probabilistic measures, where 'almost' 1 means 'infinitesimally close to 1', see [16].) We call this a *parameterized probabilistic model* (PPF). The second problem can be solved by defining separate probability distributions for each tuple size, much like our use of different rankings for different tuple sizes. A nicer solution would be to use the probability distribution over the domain to induce a probability distribution over tuples, i.e. $Pr((d_1, \ldots, d_k)) = Pr(d_1) \cdots Pr(d_k)$. In that case, we say that the PPF is *strong*. Unfortunately, our language is not strong enough to force such an interpretation. However, we can obtain this condition by resorting to some form of $\omega$-consistency. We conjecture that this property of $\omega$-consistency follows from regular consistency when the theory is finite.

What follows is a formalization of the discussion above.

DEFINITION 4.1
A parameterized probabilistic first-order model (PPF) is a pair $\mathcal{M}_P = \langle \mathbf{M}, P \rangle$, where

- $\mathbf{M}$ is a standard first-order structure.
- $P = \{Pr^k : k \in N\}$, where $Pr^k = \{Pr^k_n : n \in N\}$ is a sequence of probability measures on $|\mathbf{M}|^k$.

Satisfiability of a formula $\varphi$ in a PPF $\mathcal{M}_P$ is defined in the standard fashion, except for the following case:

- $\mathcal{M}_P \models \beta \rightarrow_{(\bar{x})} \gamma$ if $\lim_{n\to\infty} Pr^{|\bar{x}|}_n(\{\bar{c} : \mathcal{M}_P \models \gamma[^{\bar{x}}_{\bar{c}}]\} \mid \{\bar{c} : \mathcal{M}_P \models \beta[^{\bar{x}}_{\bar{c}}]\}) = 1$. (We abbreviate this as $\lim_{n\to\infty} Pr(\gamma|\beta) = 1$.)

That is, the conditional $\beta \rightarrow_{(\bar{x})} \gamma$ is satisfied whenever the probability of the set of substitutions for $\bar{x}$ under which $\beta \wedge \gamma$ is satisfied, given the set of substitutions under which $\beta$ is satisfied, approaches 1.

We shall also need the following two properties:

DEFINITION 4.2
We say that a PPF $\mathcal{M}_P$ is *smooth* w.r.t. $\mathcal{L}_C$ if for every (possibly infinite) set B of wffs and every wff $\alpha$: if $\lim_{n\to\infty} Pr_n(\beta|\alpha) = 0$ for every $\beta \in B$ then $\lim_{n\to\infty} Pr_n(\bigvee_{\beta\in B} \beta|\alpha) = 0$.

We say that a PPF $\mathcal{M}_P$ is pointed if for all formulas $\alpha, \beta \in \mathcal{L}_C$ $\lim_{n\to\infty} Pr_n(\beta|\alpha)$ exists.

We chose the term *smooth* because of its similarity to the smoothness requirement in the context of our standard model: both conditions make Axiom (3.12) valid. The requirement that the PPF be pointed is needed to ensure that Axiom (3.8) (Rational Monotony) will hold.

THEOREM 4.3
For a countable language **L**, axioms (3.1)–(3.12) are sound and complete w.r.t. the class of smooth, pointed PPF models.

DEFINITION 4.4
A PPF is **strong** if for all $k, n \in N$ and for all $d_1, \ldots, d_k \in ||M||$, we have that $Pr^k_n((d_1, \ldots, d_k)) = Pr^1_n(d_1) \cdots Pr^1_n(d_k)$.

We conjecture the following:

CONJECTURE 4.5
Given a countable language $L$, for finite theories $\Gamma$, $\Gamma$ has a model in **NS** satisfying *permutation* and *concatenation* iff $\Gamma$ has a strong PPF model.

While it is easy to define a ranked structure based on a PPF model, we have not been able to prove the other direction.

We remark that Lehmann and Magidor[16] give a slightly different semantics to propositional conditional logics using non-standard probability measures. There is a strong correspondence between non-standard reals and infinite converging sequences of reals. In fact, one way of defining non-standard reals is by means of equivalent classes of converging sequences of standard real numbers. Thus, we believe that the two methods are equivalent.

## 4.2 Syntactic variants

The syntax of our system of statistical conditionals can be changed to highlight different aspects of these conditionals. While we chose to use a formal language similar to that used in conditional logics, we see two other possibilities.

As we mentioned earlier, the conditional $\varphi \rightarrow_{(\bar{x})} \psi$ tells us something about the properties of a subset of the objects that can be assigned to $\bar{x}$ such that $\varphi$ is satisfied. Thus, it can be viewed as a quantifier that is slightly weaker than the universal quantifier, but much stronger than the existential quantifier. Indeed, [7] presents the logic of this paper in this form. Rather than use conditionals, it uses *normality* quantifiers of the form $\mathcal{N}_{\bar{x}}^{\varphi}$, where $\mathcal{N}_{\bar{x}}^{\varphi}$ quantifies over all the most normal $\bar{x}$ satisfying $\varphi$. Thus, instead of $\varphi \rightarrow_{(\bar{x})} \psi$, one writes $\mathcal{N}_{\bar{x}}^{\varphi}\psi$. In fact, the universal and existential quantifiers can be shown to be special cases of normal quantifiers (and hence of conditionals). Instead of $\forall x \neg\alpha$ we can write $\mathcal{N}_{\bar{x}}^{\alpha} false$ (or $\alpha \rightarrow_{(x)} false$).[3]

While the use of normality quantifiers presents only a slight syntactical variant to our conditional language, one can take a completely different approach that has a much simpler axiomatization and gives a system that is much closer to standard first-order logic. This approach is based on adding *normality* predicates, which are counterparts of McCarthy's abnormality predicates [19]. In this approach we add a normality predicate for each wff $\varphi$ and each subset of variables free in $\varphi$. In principle this requires at least a countable number of additional predicates in the language. However, any finite theory will refer only to a finite subset of these predicates.

Formally, if $\mathcal{L}$ is a first-order language, then $\mathcal{L}_N$ is the minimal extension of $\mathcal{L}$ such that if $\varphi$ is a wff with free variables $x_1, \ldots, x_n$ and $t_1, \ldots, t_k$ are terms then $Normal_{x_{i_1},\ldots,x_{i_k}}^{\varphi(x_1,\ldots,x_n)}(t_1, \ldots, t_k)$ is a wff, where $x_{i_1}, \ldots, x_{i_k} \in \{x_1, \ldots, x_n\}$.[4]

The intuitive reading of this formula is best illustrated by an example. To say that Alice and Bob are a normal mother and son pair, we write: $Normal_{x,y}^{Mother(x,y)}(Alice, Bob)$. The idea is that the subscript $x, y$ (whose order is important) maps the terms to the appropriate free variables in the wff $Mother(x, y)$. Equivalently, we could write $Normal_{y,x}^{Mother(x,y)}(Bob, Alice)$. These formulas are equivalent because in both *Bob* gets mapped to $y$ and *Alice* gets mapped to $x$. We can represent conditional statements, such as $\varphi \rightarrow_{(x)} \psi$ by writing $\forall x(Normal_y^{\varphi(y)}(x) \Rightarrow \psi(x))$.

Formally, we can revise our definition of satisfiability within a ranked first-order structure by replacing the clause on conditionals by:

- $\mathcal{M} \models Normal_{x_{i_1},\ldots,x_{i_k}}^{\varphi(x_1,\ldots,x_n)}(t_1, \ldots, t_k)[s]$ if $\bar{s}'(x_1), \ldots, \bar{s}'(x_n)$ is in $min_{x_1,\ldots,x_n}\varphi(x_1, \ldots, x_n)$;

[3] This is related to Lewis'[18] notions of inner modality, corresponding to the normality quantifier.

[4] This formulation of the normality quantifier was suggested to me by Daniel Lehmann.

where $s'$ is $s[^{x_{i_1},\ldots,x_{i_k}}_{t_1,\ldots t_k}]$ and $\bar{s}'$ is its extension to the set of all terms.

Using this notation, instead of the first 12 axiom schemas of NPC, we require only the following seven-axiom schema:

(i) All instances of FOL tautologies

(ii) $\exists\bar{y}\alpha(\bar{y}) \Leftrightarrow \exists\bar{y}Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y})$

(iii) $\forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \Rightarrow \alpha(\bar{y}))$

(iv) $(\alpha \equiv \beta) \Rightarrow \forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \equiv Normal_{\bar{x}}^{\beta(\bar{x})}(\bar{y}))$

(v) $\forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})\vee\beta(\bar{x})}(\bar{y}) \Rightarrow (Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \vee Normal_{\bar{x}}^{\beta(\bar{x})}(\bar{y})))$

(vi) $\exists\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \wedge \beta(\bar{y})) \Rightarrow \forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})\wedge\beta(\bar{x})}(y) \Rightarrow Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}))$

(vii) $Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \Rightarrow Normal_{\bar{z}}^{\alpha(\bar{z})}(\bar{y})$

It is possible to weaken the system by replacing (vi) with

(vi′) $\forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \Rightarrow \beta(\bar{y})) \Rightarrow \forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})\wedge\beta(\bar{x})}(\bar{y}) \Rightarrow Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}))$.

It may seem that this logic is more expressive than our conditional logic. We have shown that beyond expressing conditionals it allows us to say of a particular object that it is normal. However, we can say the same thing in our original language if it contains equality. To say that Tweety is a normal bird we simply write: $\neg(bird(x) \rightarrow_{(x)} \neg(x = Tweety))$. That is, it is not the case that all normal birds are not Tweety. Hence, for a first-order language that contains equality, we can show the following:

THEOREM 4.6
The axioms (i)–(vii) are sound and complete w.r.t. the class of NS structures.

## 5 Discussion

We presented a first-order conditional logic in which conditionals are interpreted as qualitative statistical statements. Our main contributions are in pointing out this alternative interpretation for conditionals, one that underlies much common-sense knowledge; in providing a formal language for making such statements, including a sound and complete axiomatization; and in pointing out its probabilistic semantics.

Our particular approach is motivated by the fact that people seem to possess much information about typical, normal, or expected properties of objects; for example, birds typically fly, the bus usually arrives on time, or physicists are usually good in math. In addition, people seem to possess representations of prototypical objects. (See [21] for an interesting discussion of prototypes and empirical studies of their use.) Much of this information is statistical in nature: most birds we encounter fly, most times we wait for the bus it is on schedule, and most physicists we know are good in math. However, most of this information is qualitative, rather than quantitative. Our logic allows us to reason with such statements and our axiomatization and semantics is intended to capture and model them.

Tversky [22] suggests a model of how prototypes are used. In Tversky's model the degree of 'typicality' of an individual w.r.t. a class is determined by its distance from the prototype of this class in the space of properties, where different properties can have a different influence on this metric. While we do not claim that our model wholly captures the common-sense use of prototypes—current theories of prototypes and their use are more involved than Tversky's model—an intuitive interpretation of our semantics is based on these ideas; namely, that objects in our model are ranked based on their similarity to the class prototype. Objects which possess all the typical properties of class $\alpha$ are minimally ranked w.r.t. that class and they may differ from each other on various properties to which the prototype is indifferent. The prototype itself is only implicitly specified as the set of properties common to all minimally ranked objects of the class. Objects from class $\alpha$ that are not minimally ranked have some atypical property, and the degree to which some property $\beta$ is atypical w.r.t. class $\alpha$ is manifested by the distance between the ranking of the most normal $\alpha$ objects versus the ranking of the most normal $\alpha \wedge \beta$ objects.

Still, there are problems with logics of prototypes that our proposal suffers from. For example, if $\beta_i$ for $i = 1, \ldots, 10$ is a typical property of class $\alpha$, it does not follow that $\beta_1 \wedge \cdots \wedge \beta_{10}$ is a typical property of class $\alpha$, although this conclusion is sanctioned by all systems which have the **And** property. It remains to be seen whether a more satisfactory solution that does not suffer from these problems exists. One possibility is to allow a richer structure in which there are different rankings with respect to different properties.[5] In that case, the **And** property would not necessarily hold. However, it is not clear what relationship should hold between the different rankings, and this approach deserves further study.

As we have seen, the conditional operator can be viewed as a quantifier. The idea of non-standard quantification is not a new one. In [5] Barwise and Cooper define and treat the subject of generalized quantifiers. Their treatment is very broad and general, having a linguistic aim. The idea also appears in the field of model theoretic logic, but on a different level and with different aims (for example, quantifiers that quantify over enumerable sets).

A closer approach is presented in Altham's work on formalizing the concept 'many' and 'nearly all' [2]. Altham adds a new quantifier **M**. In his approach $(\mathbf{M}x)Fx$ (i.e. many $x$ are $F$) is satisfied, if at least $n$ distinct individuals in the domain are $F$. Nearly all $x$ are $F$ (written $(\mathbf{N}x)Fx$ ), if fewer than $n$ members of the domain are not $F$.

The choice of $n$ that corresponds with the notion of *many* is different in different contexts, e.g. compare the idea of *many* Christians to that of *many* Quakers. To overcome this Altham introduces an infinite set of quantifiers $\{\mathbf{M}^k \ k \in \mathrm{N}\}$, where $(\mathbf{M}^k x)Fx$ if at least $k$ individuals are $F$. With every predicate he associates an index that determines how many is *many* (i.e. which $k$ should be used to say *many* $x$s are $y$s). However, one cannot infer the index of a wff from the indices of its subformulas, but only obtain bounds on its size (e.g. if the index of $\alpha$ is $k$ and of $\beta$ is $j$ then the index of $\alpha \vee \beta$ is bounded below by the maximum of $j$ and $k$).

Altham's approach is more quantitative than ours, actually counting the number of objects with a given property. Our conditional operator is more flexible in that it leaves this number unspecified and allows for interpretations that are different than those suggested by the notion of *many*.

Perhaps the work most closely related to ours is Schlechta's work on defaults as generalized quantifiers [20]. Schlechta interprets defaults of the form 'birds normally fly' as saying that 'most birds fly' or that 'the elements of a large or important subset of the set of birds fly'. Hence, as in our logic, defaults (or conditionals) are interpreted w.r.t. a single first-order model. Schlechta formalizes the notion of 'important subset' using a weak notion of a *filter*. A filter w.r.t. a set $A$ is a set of subsets of $A$ that contains $A$, that is closed under the superset relation and under intersections. Schlechta replaces this last property with the requirement of non-empty intersections. Hence, Schlechta's semantics consists of a first-order structure sup-

[5] This was suggested to me by David Poole and an anonymous referee.

plemented by weak filters that tell us for each set what its important subsets are. Schlechta provides a sound and complete axiomatization for his logic.

Schlechta's work and our work attempt to formalize similar ideas, but they differ on a number of issues. Schlechta's language contains unary predicates only. Hence, it provides only a limited formalization of first-order conditionals. Consequently, it is less expressive than our logic and does not address issues such as the interaction between normal tuples of different sizes. Schlechta's axiom system is weak and, consequently, more flexible. In fact, restricted to unary predicates, our semantics specializes Schlechta's semantics, since rankings implement a (real) filter which, in turn, is a weak filter. Hence, we have formalized a more specific approach to conditionals. Indeed, while our conditional operator can be given probabilistic semantics as stating statements that are almost certainly true, one possible interpretation of Schlechta's defaults is as stating facts that are true with probability greater than 0.5. Our more specific semantics has the benefit of a simpler representation: whereas one must specify a filter for every subset of the domain in order to handle conditionals in Schlechta's logic, a single ranking on the domain suffices in our case. Finally, it should be mentioned that Schlechta's work is concerned with issues which we have not considered, such as specificity of defaults.

Delgrande [9] was the first to offer a first-order version of conditional logic. His language is restricted to flat formulas, i.e. formulas in which conditionals are not nested. Delgrande interprets his language over a set of possible worlds, all of which are models of first-order logic with a common domain and valuation function (a function that assigns an individual to each term). A conditional $\varphi(t) \rightarrow \psi(t)$ is satisfied if in the least exceptional worlds in which $\varphi(t)$ holds, $\psi(t)$ holds as well. This type of interpretation is characteristic of the subjective approach. The model tells us what worlds the agent perceives as more reasonable, and conditionals are interpreted as talking about the properties of the most likely worlds or scenarios.

Another approach to first-order conditionals is offered by Lehmann and Magidor [17]. Their approach is somewhat different in motivation from the other approaches, and cannot be classified as distinctly subjective or statistical, but rather as a combination of both. Lehmann and Magidor are interested in the study of nonmonotonic consequence relations. A consequence relation is a function which, given a theory, returns the theory containing the consequence of the original theory. In [17] Lehmann and Magidor look at consequence relations defined over first-order theories. Their approach can be embedded in a conditional logic with the restriction that conditionals are flat, and that conditionals do not appear under the scope of a quantifier. Thus, $\forall x(bird(x) \rightarrow fly(x))$ is not a wff, while $(\forall x bird(x)) \rightarrow (\forall x fly(x))$ is a wff. This approach extends their earlier work on propositional consequence relations [15, 16]. In particular they are interested in studying when quantifiers can be introduced or eliminated. For instance, one rule they suggest is:

$$\frac{\alpha \rightarrow \beta}{\exists x\alpha \rightarrow \exists x\beta} \qquad (\exists - \mathbf{Intr})$$

The semantics Lehmann and Magidor provide for their theories takes the form of a partial order over pairs of the form $\langle M, f\rangle$, where $M$ is a standard first-order model, and $f$ is a function assigning a domain element to each variable. The conditional $bird(x) \rightarrow fly(x)$ is satisfied if in those minimal $\langle M, f\rangle$ pairs in which the object assigned to $x$ by $f$ has the property *bird*, that object also has the property *fly*. Thus, the model can be viewed as saying what models are most normal, but also what assignments to objects are most normal. While the ordering over models fits the subjective approach, the fact that there is an ordering over the assignment functions $f$ is more in line with the statistical approach: if one fixes the model $M$, the ordering over assignment functions can be understood as an ordering over infinite tuples of objects which is closer to the statistical approach. Motivated by the work of Lehmann and Magidor, [7] uses such orders over assignment functions to define the semantics of statistical conditionals.

Compared with these approaches, our approach provides a purely statistical interpretation of conditional first-order logic, employing a general language that does not restrict us to flat formulas, nor does it restrict the use of quantifiers.

In [11] Friedman *et al.* discuss a number of problems that arise in the logic of Delgrande and in other formulations of first-order conditional logic. We mentioned these problems in the introduction, and here we demonstrate our ability to deal with them. The first problem arises when we substitute terms for universally quantified variable: in Delgrande's logic the following is inconsistent:

$$\forall x, y(Pet(x, y) \rightarrow (Dog(y) \vee Cat(y)) \;,\; \forall y(Pet(John, y) \rightarrow (Snake(y)).$$

This inconsistency stems from our ability to substitute $John$ for the variable $x$ in the first sentence, obtaining a sentence that says that John's typical pets are dogs or cats. This contradicts the second statement which says that John's pets are typically snakes (assuming, of course, that snakes are not dogs or cats). In our language, in order to express similar information, we would write $Pet(x, y) \rightarrow_{(x,y)} (Dog(y) \vee Cat(y))$. That is, normally, a person's pet is either a dog or a cat. Because this form of quantification is weaker than universal quantification, substitutions for $x$ or $y$ are not allowed.

An instance of a similar problem is the lottery paradox. We know that most tickets do not win the lottery, but there is a winning ticket. The theory:

$$\forall x(true \rightarrow \neg Winner(x)) \;,\; true \rightarrow \exists x Winner(x)$$

is inconsistent in Delgrande's logic and Friedman *et al.* overcome this problem using the properties of plausibility spaces. However, if we use statistical conditionals to represent similar information we obtain the following theory:

$$true \rightarrow_{(x)} \neg Winner(x) \;,\; \exists x Winner(x).$$

It is easy to construct an NS model of this theory.

Indeed, in many cases, the information we have about the world is best interpreted as rough, qualitative statistical statements and should be modelled accordingly. If we are careful to make the distinction between subjective and statistical information, we can prevent some of the above pitfalls. However, it is important to note that the lottery paradox still possess a challenge to the subjective approach: it is not unreasonable to argue that the following (subjective) statement is consistent: 'For every given ticket, I believe that *that* ticket will not win the lottery. Yet, I believe that some ticket will win the lottery'. It seems that these statements, again, correspond to the following theory:

$$\forall x(true \rightarrow \neg Winner(x)) \;,\; true \rightarrow \exists x Winner(x)$$

with conditionals interpreted subjectively. Only now, we have made this challenge more crisp because it is not conflated with statements such as 'most tickets will not win the lottery'.

Our logic does suffer from a number of weaknesses. On the technical side, it would be nice to do away with the smoothness requirement, especially in the context of the probabilistic semantics. This may not be too difficult, since this restriction has been overcome in the propositional case [6].[6] On the more pragmatic side, the fact that we cannot substitute terms for

[6] The situation in the first-order case is slightly more complicated because of the existence of quantifiers.

variables bound by the conditional operator also means that we cannot conclude $fly(Tweety)$ from $bird(x) \rightarrow_{(x)} fly(x)$. That is, although we know that most/normal birds fly, we have no mechanism to deduce that Tweety is one of these normal birds. This is not surprising, since we cannot conclude that Tweety flies given a statistical statement such as '99% of all birds fly'. In order to handle this we would need additional nonmonotonic machinery. For instance, one would like to deduce that an object satisfying a formula is normal for that formula unless the opposite is known, i.e. Tweety is a normal bird unless we can prove otherwise. Such deductions are not possible within our logic, which is a monotonic formalism.

The problem we just alluded to is a fundamental problem at the heart of statistical reasoning: Given statistical information, what is it that we should believe about particular individuals? That is, if we know that 99% of birds fly, what should we believe about Tweety. The problem we are facing now is simply the qualitative counterpart of this old problem. Recently, Bacchus *et al.* [3] have proposed a solution to this problem in the quantitative case, i.e. they suggest a method for moving from statistics to beliefs. At this stage it is not clear whether their ideas can be used to solve the qualitative version of this problem, and this remain an important challenge. This paper did not attempt to resolve this question. However, it supplies a necessary first step to the resolution of this problem by making explicit the distinction between statistical and subjective conditionals and by supplying a logic for reasoning about statistical statements. We believe that a natural next step is to provide a semantically appealing logic that combines both statistical and subjective conditionals, using which the relationship between both types of information can be studied.

## Acknowledgements

This work is based on [7], conducted under the guidance of Menachem Magidor, whose technical expertise, time, and support were essential for its success. Numerous important comments and suggestions regarding that work were provided by Daniel Lehmann. Craig Boutilier, Joe Halpern, David Poole, and the two anonymous referees provided useful comments on the current version of this work, which owes much of its existence to the encouragement of Nir Friedman, who also supplied many important comments during numerous discussions.

## References

[1] E. Adams. *The Logic of Conditionals*. D. Reidel, Dordrecht, 1975.

[2] J. E. J. Altham. *The Logic of Plurality*. Methuen, London, 1971.

[3] F. Bacchus, A. J. Grove, J. Y. Halpern and D. Koller. Statistical foundations for default reasoning. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pp. 563–569. Morgan Kaufmann, San Mateo, 1995.

[4] F. Bacchus. *Representing and Reasoning With Probabilistic Knowledge*. The MIT Press, Cambridge, MA, 1990.

[5] J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, **4**, 159–219, 1981.

[6] C. Boutilier. Conditional logics of normality: a modal approach. *Artificial Intelligence*, **68**, 87–154, 1994.

[7] R. I. Brafman. *A Logic of Normality: Predicate Calculus Incorporating Assertions*. Master's thesis, Hebrew University of Jerusalem, 1991.

[8] C. C. Chang and H. J. Keisler. *Model Theory*. North Holland, Amsterdam, 1990.

[9] J. P. Delgrande. A first-order logic for prototypical properties. *Artificial Intelligence*, **33**, 105–130, 1987.

[10] H. B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, New York, 1972.

[11] N. Friedman, J. Y. Halpern and D. Koller. Conditional first-order logic revisited. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*. AAAI Press, Menlo Park, 1996.

[12] M. Goldszmidt, P. Morris and J. Pearl. A maximum entropy approach to nonmonotonic reasoning. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 15, 220–232, 1993.

[13] M. Goldszmidt and J. Pearl. Rank-based systems: A simple approach to belief revision, belief update and reasoning about evidence and actions. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Confernece (KR'92)*, pp. 661–672. Morgan Kaufmann, San Mateo, 1992.

[14] J. Halpern. An analysis of first-order logics of probability. *Artificial Intelligence*, 46, 311–350, 1990.

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

[16] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55, 1–60, 1992.

[17] D. Lehmann and M. Magidor. Preferential logics: the predicate calculus case. In *Proceedings of the Third Conference on Theoretical Aspects of Reasoning About Knowledge*, San Francisco, CA, March 1990. pp. 57–72. Morgan Kaufmann, Los Altos, CA, 1990.

[18] D. K. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, MA, 1973.

[19] J. McCarthy. Circumscription, a form of non monotonic reasoning. *Artificial Intelligence*, 13, 27–39, 1980.

[20] K. Schlechta. Defaults as generalized quantifiers. *Journal of Logic and Computation*, 5, 473–494, 1995.

[21] E. E. Smith. Concepts and induction. In *Foundations of Cognitive Science*, M. I. Posner, ed. pp. 501–526. The MIT Press, Cambridge, MA, 1989.

[22] A. Tversky. Features of similarity. *Psychological Review*, 84, 327–352, 1977.

# A Proofs

- The proofs of some of the lemmas and derived rules appearing in this section have been lifted to the first-order case from similar proofs appearing in Lehmann and Magidor's [16]. To acknowledge this, we have marked each such claim with (LM$\star$).
- We observe that the proof of the deduction theorem trivially generalizes to our logic, and hence we will (implicitly) allow ourselves to use it in our proofs.
- We shall use $\Longrightarrow$ to denote meta-level implication. As before, object-level implication is denoted by $\Rightarrow$.

## *Some additional derived rules*

(A.1) $\neg(\alpha \rightarrow_{(\bar{x})} \beta) \Rightarrow \exists\bar{x}(\alpha \wedge \neg\beta)$

PROOF. Assume $\neg\exists\bar{x}(\alpha \wedge \neg\beta)$. This is equivalent to $\forall\bar{x}(\neg\alpha \vee \beta)$. Using (3.9) we have $\neg\alpha \rightarrow_{(\bar{x})} (\neg\alpha \vee \beta)$. By **Reflexivity** $\alpha \rightarrow_{(\bar{x})} \alpha$. Using **(And)**, we have $\alpha \rightarrow_{(\bar{x})} ((\neg\alpha \vee \beta) \wedge \alpha)$. Using **(R.W.)**, we now have $\alpha \rightarrow_{(\bar{x})} \beta$. ∎

(A.2) $\neg(\alpha \rightarrow_{(\bar{x})} \beta) \Rightarrow \exists\bar{x}[(\alpha \wedge \neg\beta) \wedge ((\alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow (\alpha \Rightarrow \gamma))]$

PROOF. We will show that the negation of the consequence, $\exists\bar{x}[(\alpha \wedge \neg\beta) \wedge ((\alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow (\alpha \Rightarrow \gamma))]$ implies that $\alpha \rightarrow_{(\bar{x})} \beta$. The negation of the consequence is logically equivalent to $\forall\bar{x}[\neg\alpha \vee \beta \vee ((\alpha \rightarrow_{(\bar{x})} \gamma) \wedge \alpha \wedge \neg\gamma)]$. By following the steps taken in the proof of A.1, we get $\alpha \rightarrow_{(\bar{x})} [\beta \vee ((\alpha \rightarrow_{(\bar{x})} \gamma) \wedge \alpha \wedge \neg\gamma)]$. Reasoning by cases, suppose that $\alpha \rightarrow_{(\bar{x})} \gamma$, using **And** (conjoining $\gamma$ to the consequence of the conditional) and a few PC deductions we get $\alpha \rightarrow_{(\bar{x})} \beta$. Alternatively, if $\neg(\alpha \rightarrow_{(\bar{x})} \gamma)$, then using (3.9) we have $\alpha \rightarrow_{(\bar{x})} \neg(\alpha \rightarrow_{(\bar{x})} \gamma)$. As in the previous case, using **(And)** and a number of PC deductions we can deduce $\alpha \rightarrow_{(\bar{x})} \beta$. ∎

(A.3) $\neg(\alpha \rightarrow_{(\bar{x})} \beta) \wedge (\alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow \exists\bar{x}(\alpha \wedge \gamma \wedge \neg\beta)$

PROOF. By A.1 it is enough to show that
$(\neg(\alpha \rightarrow_{(\bar{x})} \beta) \wedge \alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow \neg(\alpha \rightarrow_{(\bar{x})} (\neg\gamma \vee \beta))$, but this is equivalent to
$(\alpha \rightarrow_{(\bar{x})} (\neg\gamma \vee \beta) \wedge \alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow (\alpha \rightarrow_{(\bar{x})} \beta)$ which we get using **And** and **R.W.**. ∎

(A.4) $\exists\bar{x}\alpha(\bar{x}) \Rightarrow \alpha \rightarrow_{(\bar{x})} \forall\bar{y}(\gamma(\bar{x},\bar{y}) \vee \neg(\alpha \rightarrow_{(\bar{x})} \gamma(\bar{x},\bar{y})))$ $(\bar{y} \cap (\bar{x} \cup \mathrm{fv}(\alpha)) = \emptyset)$.

PROOF. First, notice that the following schema is derivable:

$$\alpha \rightarrow_{(\bar{x})} (\gamma \vee \neg(\alpha \rightarrow_{(\bar{x})} \gamma)).$$

First, $(\alpha \rightarrow_{(\bar{x})} \gamma) \vee \neg(\alpha \rightarrow_{(\bar{x})} \gamma)$ is an instance of an FOL tautology. Next, notice that $\psi \rightarrow_{(\bar{x})} \varphi$ is derivable from $\varphi$ using (3.9) (because from $\varphi$ we get $\forall\bar{x}\varphi$ using FOL). Hence, we have that $(\alpha \rightarrow_{(\bar{x})} \gamma) \vee (\alpha \rightarrow_{(\bar{x})} \neg(\alpha \rightarrow_{(\bar{x})} \gamma))$. Using Right Weakening on both disjuncts, we obtain $(\alpha \rightarrow_{(\bar{x})} (\gamma \vee \neg(\alpha \rightarrow_{(\bar{x})} \gamma))$. Assuming $(\bar{y} \cap \bar{x} = \emptyset)$ and using FOL, we obtain $\forall\bar{y}(\alpha \rightarrow_{(\bar{x})} (\gamma \vee \neg(\alpha \rightarrow_{(\bar{x})} \gamma)))$. Assuming in addition $\bar{y} \cap \mathrm{fv}(\alpha) = \emptyset$, we obtain $\alpha \rightarrow_{(\bar{x})} \forall y(\gamma \vee \neg(\alpha \rightarrow_{(\bar{x})} \gamma))$, using **Interchange**. The result then immediately follows. ∎

(A.5) $(\alpha \wedge \beta) \rightarrow_{(\bar{x})} \gamma \Rightarrow \alpha \rightarrow_{(\bar{x})} (\beta \Rightarrow \gamma)$ $(LM\star)$

PROOF. Using **R.W.** we derive $(\alpha \wedge \beta) \rightarrow_{(\bar{x})} \gamma \Rightarrow (\alpha \wedge \beta) \rightarrow_{(\bar{x})} (\beta \Rightarrow \gamma)$.
Using **Reflexivity** and **R.W.**, we obtain $(\alpha \wedge \neg\beta) \rightarrow_{(\bar{x})} (\beta \Rightarrow \gamma)$.
Hence, $(\alpha \wedge \beta) \rightarrow_{(\bar{x})} \gamma \vdash (\alpha \wedge \beta) \rightarrow_{(\bar{x})} (\beta \Rightarrow \gamma) \wedge (\alpha \wedge \neg\beta) \rightarrow_{(\bar{x})} (\beta \Rightarrow \gamma)$.
Using **Or** we now have $(\alpha \wedge \beta) \rightarrow_{(\bar{x})} \gamma \vdash (\alpha \wedge (\beta \vee \neg\beta)) \rightarrow_{(\bar{x})} (\beta \Rightarrow \gamma)$.
We conclude by **L.E.** ∎

(A.6) $((\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\beta) \Rightarrow (\alpha \rightarrow_{(\bar{x})} \neg\beta)$ $(LM\star)$

PROOF. $(\alpha \vee \beta) \rightarrow_{(\bar{x})} (\alpha \vee \beta)$ implies that $(\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\beta \vdash (\alpha \vee \beta) \rightarrow_{(\bar{x})} (\alpha \wedge \neg\beta)$ (by **And**).
Thus $(\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\beta \vdash (\alpha \vee \beta) \rightarrow_{(\bar{x})} \alpha$ (by **R.W.**). Using **C.M.** one gets that $((\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\beta) \Rightarrow (((\alpha \vee \beta) \wedge \alpha) \rightarrow_{(\bar{x})} \neg\beta)$. Using (**L.E.**) the consequence in this implication implies $\alpha \rightarrow_{(\bar{x})} \neg\beta$. ∎

(A.7) $(\alpha \rightarrow_{(\bar{x})} \mathbf{false}) \Rightarrow ((\alpha \wedge \beta) \rightarrow_{(\bar{x})} \mathbf{false})$ $(LM\star)$

PROOF. We have that $\alpha \rightarrow_{(\bar{x})} \mathbf{false} \Leftrightarrow \alpha \rightarrow_{(\bar{x})} \beta \wedge \neg\beta$ and using **R.W.** we obtain $\alpha \rightarrow_{(\bar{x})} \mathbf{false} \vdash \alpha \rightarrow_{(\bar{x})} \beta$.
Using **C.M.** we now have $\alpha \rightarrow_{(\bar{x})} \mathbf{false} \wedge \alpha \rightarrow_{(\bar{x})} \beta \vdash (\alpha \wedge \beta) \rightarrow_{(\bar{x})} \mathbf{false}$. ∎

(A.8) $(\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} (\neg\alpha \wedge \neg\beta) \vdash (\beta \vee \gamma) \rightarrow_{(\bar{x})} \neg\beta$ $(LM\star)$

PROOF. By **Reflexivity** we have $\vdash (\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} (\alpha \vee \beta \vee \gamma)$ and by **R.W.**
$(\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} (\neg\alpha \wedge \neg\beta) \vdash (\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} \neg\alpha$.
Applying **And** and then **R.W.** we have $(\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} (\neg\alpha \wedge \neg\beta) \vdash (\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} (\beta \vee \gamma)$.
Again, by **Reflexivity** and **R.W.**, we have $(\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} (\neg\alpha \wedge \neg\beta) \vdash (\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} \neg\beta$.
Using **C.M.** one obtains:
$(\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} (\neg\alpha \wedge \neg\beta) \vdash ((\alpha \vee \beta \vee \gamma) \wedge (\beta \vee \gamma)) \rightarrow_{(\bar{x})} \neg\beta$ and **L.E.** gives us the desired result. ∎

(A.9) $(\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\alpha \vdash (\alpha \vee \beta \vee \gamma) \rightarrow_{(\bar{x})} \neg\alpha$ $(LM\star)$

PROOF. From PC, we have $(\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\alpha \vdash (\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\alpha$.
By **Reflexivity** we have $\vdash (\gamma \wedge \neg\alpha) \rightarrow_{(\bar{x})} (\gamma \wedge \neg\alpha)$
and using **R.W.** we have $\vdash (\gamma \wedge \neg\alpha) \rightarrow_{(\bar{x})} \neg\alpha$.
Using **Or** we have $(\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\alpha \vdash ((\alpha \vee \beta) \vee (\gamma \wedge \neg\alpha)) \rightarrow_{(\bar{x})} \neg\alpha)$. Since $(\alpha \vee \beta) \vee (\gamma \wedge \neg\alpha)$ is logically equivalent to $(\alpha \vee \beta \vee \gamma)$, by applying **L.E.** we have the desired result. ∎

In what follows the notation $\hat{\beta}_{\bar{x}}$ (or $\hat{\beta}$ when $\bar{x}$ is clear from the context) will be used to denote the set $\{\bar{c} \in |\mathbf{M}|^{|\bar{x}|} : \mathcal{M} \models \beta[^{\bar{x}}_{\bar{c}}]\}$. We will use the letters $f, g$ to refer to tuples of domain elements (i.e. members of $|\mathbf{M}|^n$) and the following notation: $\mathcal{M} \models_f \alpha \overset{\mathrm{def}}{=} \mathcal{M} \models \alpha[^{\bar{x}}_{f}]$ (where $|f| = |\bar{x}|$).[7]

THEOREM 3.1
Axioms (3.1)–(3.12) are sound w.r.t. the class **NS**.

PROOF. • (3.1) is inherited from standard **PC** due to the identity of the definition of satisfaction with respect to wffs of the form $\neg\alpha, \alpha \wedge \beta$ and $\forall x\alpha$. The same applies to Modus Ponens.

[7] Recall that $|f|$ stands for the cardinality of $|f|$. Unfortunately, when $\mathbf{M}$ is a first-order structure, $|\mathbf{M}|$ stands for the domain of $\mathbf{M}$. However, this is the only exception.

- (3.2)–(3.8) are all easily proven as in the propositional case (e.g. see [15, 16]). For instance, consider (3.5): suppose that $\mathcal{M} \models (\alpha \rightarrow_{(\bar{x})} \beta) \wedge (\alpha \rightarrow_{(\bar{x})} \gamma)$. Thus, all minimally ranked tuples satisfying $\alpha$ satisfy $\beta$. Clearly, all minimally ranked tuples satisfying $\alpha \wedge \beta$ satisfy $\beta$. We conclude that the minimally ranked tuples of $\hat{\alpha}$ and of $\widehat{\alpha \wedge \beta}$ are the same. Consequently, since $\mathcal{M} \models \alpha \rightarrow_{(\bar{x})} \gamma$, they must satisfy $\gamma$.
- (3.9) is also trivial.
- (3.10) Immediate from definition.
- (3.11) Immediate.
- (3.12) Assume $\mathcal{M} \models \forall \bar{x}(\alpha \rightarrow_{(\bar{y})} \beta)$ and that $\{\bar{x} \cap \bar{y} = \emptyset$ and $\mathrm{fv}(\alpha) \cap \bar{x} = \emptyset\}$.
  Then, for each $\bar{s} \in \|M\|$, $\mathcal{M} \models \alpha \rightarrow_{(\bar{y})} \beta[^{\bar{x}}_{\bar{s}}]$.
  This implies that for each $\bar{s}$ in $\mathcal{M}$ it is the case that for every $f \in |M|^{|\bar{y}|}$ minimal in $\hat{\alpha}$, $\mathcal{M} \models_f \beta[^{\bar{x}}_{\bar{s}}]$.
  Therefore $\mathcal{M} \models \alpha \rightarrow_{(\bar{y})} \forall \bar{x}\beta$.
  The other direction is as easy. ∎

THEOREM 3.2
Axioms (3.13) is sound w.r.t. the class of NS structures satisfying *permutation*, and axioms (3.14)–(3.17) are sound w.r.t. the class of NS structures satisfying *concatenation*.

PROOF. First, we show that axiom (3.13) is sound w.r.t. the class of NS structures satisfying *permutation*.

Assume that *permutation* holds and that $\neg(\alpha \rightarrow_{(\bar{y},\bar{x})} \beta)$ is satisfied. Thus, there is some minimal $\bar{c} \circ \bar{d}$ that satisfies $\alpha[^{\bar{y},\bar{x}}_{\bar{c},\bar{d}}]$ and does not satisfy $\beta[^{\bar{y},\bar{x}}_{\bar{c},\bar{d}}]$. This implies that $\alpha \wedge \neg\beta[^{\bar{x},\bar{y}}_{\bar{d},\bar{c}}]$ is satisfied. Moreover, $\bar{d} \circ \bar{c}$ must be a minimal tuple to satisfy $\alpha[^{\bar{x},\bar{y}}_{\bar{d},\bar{c}}]$, otherwise, we use the *permutation* property to show that $\bar{c} \circ \bar{d}$ is not a minimal element satisfying $\alpha[^{\bar{y},\bar{x}}_{\bar{c},\bar{d}}]$. Therefore, $\neg(\alpha \rightarrow_{(\bar{x},\bar{y})} \beta)$ holds as well.

Next, we show that axioms (3.14)–(3.17) are sound w.r.t. the class of NS structures satisfying *concatenation*.

(3.14) Assume that *concatenation* holds in the model. Notice that $\alpha \leq_{\bar{x}} \beta$ is satisfied in $\mathcal{M}$ iff the minimal rank of tuples in $\hat{\alpha}$ is no higher than that of tuples in $\hat{\beta}$. Thus, the sentence $(\alpha \leq_{\bar{x}} \beta) \wedge (\alpha' \leq_{\bar{y}} \beta')$ is satisfied only if the minimal tuples in $\hat{\alpha}$ are no higher than the minimal tuples in $\hat{\beta}$ and the minimal tuples in $\hat{\alpha}'$ are no higher than the minimal tuples in $\hat{\beta}'$.

Next, notice that if $\bar{c} \circ \bar{c}'$ is minimal in $(\widehat{\alpha \wedge \alpha'})$ and the conditions on the variables in Axiom (3.14) are satisfied, then $\bar{c}$ is minimal in $\alpha$ and $\bar{c}'$ is minimal in $\alpha'$: clearly, $\bar{c}$ satisfies $\alpha$ and $\bar{c}'$ satisfies $\alpha'$. If some $\bar{d}$ satisfies $\alpha$ and is ranked lower than $\bar{c}$, then using *concatenation*, we will have a tuple satisfying $\alpha \wedge \alpha'$ whose rank is lower.

It is now easy to obtain the soundness of Axiom (3.14): let $\bar{c}$ be minimal in $\hat{\alpha}$ and $\bar{c}'$ minimal in $\hat{\alpha}'$. We know that $\bar{c} \circ \bar{c}'$ must be minimal in $\widehat{\alpha \wedge \alpha'}$. (Otherwise, another tuple that is ranked lower than $\bar{c} \circ \bar{c}'$ is minimal in this set, and as we have seen this implies that either $\bar{c}$ is not minimal for $\hat{\alpha}$, or $\bar{c}'$ is not minimal for $\hat{\alpha}'$.) If $\bar{d}$ is minimal in $\hat{\beta}$ and $\bar{d}'$ is minimal in $\hat{\beta}'$, then we know that $\bar{d} \circ \bar{d}'$ is minimal for $\widehat{\beta \wedge \beta'}$, and that $\bar{c}$ is ranked no higher than $\bar{d}$ and that $\bar{c}'$ is ranked no higher than $\bar{d}'$. Consequently, *concatenation* implies that $\bar{c} \circ \bar{c}'$ is ranked no higher than $\bar{d} \circ \bar{d}'$, which is what we had to prove.

(3.15) Similar to (3.14).

(3.16) Suppose that $\mathcal{M} \models \alpha \wedge \alpha' \rightarrow_{(\bar{x},\bar{y})} \gamma$, we must show that $\mathcal{M} \models \alpha \rightarrow_{(\bar{x})} (\alpha' \rightarrow_{(\bar{y})} \beta)$.

Choose some $\bar{d}$ normal for $\alpha$ and let $\bar{d}'$ be any tuple normal for $\alpha'[^{\bar{x}}_{\bar{d}}]$ and assume that $\alpha \wedge \alpha' \rightarrow_{(\bar{x},\bar{y})} \gamma$ We must show that $\bar{d} \circ \bar{d}'$ is in $\hat{\gamma}$. We notice that because $\bar{x} \cup fv(\alpha') = \emptyset$ $\bar{d}'$ is in fact normal for $\alpha'$.

Given our assumptions, to show that $\bar{d} \circ \bar{d}' \in \hat{\gamma}$, it is sufficient to prove that $\bar{d} \circ \bar{d}'$ is minimal in $(\widehat{\alpha \wedge \alpha'})$. Let $\bar{c} \circ \bar{c}'$ be minimal in $(\widehat{\alpha \wedge \alpha'})$. As we have argued above, given the restrictions on the free variables of $\alpha$ and $\alpha'$ and on $\bar{x}$ and $\bar{y}$, this implies that $\bar{c}$ is minimal in $\hat{\alpha}$ and that $\bar{c}'$ is minimal in $\hat{\alpha}'$. However, $\bar{d}$ cannot be ranked higher than $\bar{c}$, and similarly, $\bar{d}'$ cannot be ranked higher than $\bar{c}'$. Hence, by *concatenation*, $\bar{d} \circ \bar{d}'$ cannot be ranked higher than $\bar{c} \circ \bar{c}'$, and must therefore be minimal in $(\widehat{\alpha \wedge \alpha'})$.

(3.17) First, suppose that $\bar{c} \circ \bar{c}'$ is minimal$_{\bar{x},\bar{y}}$ in $\hat{\alpha}$. From *concatenation*, we get that $\bar{c}$ must be minimal$_{\bar{x}}$ in $\widehat{\alpha[^{\bar{y}}_{\bar{c}'}]}$. Thus, if $(\alpha \rightarrow_{(\bar{x})} \beta)[^{\bar{x},\bar{y}}_{\bar{c},\bar{c}'}]$ then all minimal$_{\bar{x}}$ tuples in $\widehat{\alpha[^{\bar{y}}_{\bar{c}'}]}$ satisfy $\beta[^{\bar{y}}_{\bar{c}'}]$. In particular, we have that $\bar{c}$ in $\widehat{\beta[^{\bar{y}}_{\bar{c}'}]}$. ∎

THEOREM 3.3
For a countable language $\mathcal{L}_C$, axioms (3.1)–(3.12) are complete with respect to the class NS.

Before we start the proof, we mention that we will need a syntactic notion of substitution. We will use the same notation as in the case of the (semantic) assignment function, but which one is intended would be clear from the context. The definition will be identical to the one used in first-order logic, where basically, $\alpha[^{x}_{c}]$ says that we must substitute the constant $c \in \mathcal{L}_C$ for all those instances of $x$ in the wff $\alpha$, where $x$ appears free in $\alpha$. So $Man(x)[{x \atop c}] = Man(c)$, but $Man(x) \rightarrow_{(\bar{x})} Tall(x)[{x \atop c}]$ is $Man(x) \rightarrow_{(\bar{x})} Tall(x)$. Thus, syntactic substitution defines a mapping from $\mathcal{L}_C$ to $\mathcal{L}_C$.

PROOF. Let $\mathcal{L}_C$ be a countable language and $\hat{\mathbf{\Gamma}}$ a consistent set of closed wffs in $\mathcal{L}_C$. We shall build a model $\mathcal{M}$ s.t. (such that) $\mathcal{M} \models \hat{\mathbf{\Gamma}}$. We shall construct the model in $\omega$ stages. Each stage will deal with one negated conditional $\neg(\alpha \rightarrow_{(\bar{x})} \beta)$.

At each stage we will define five sets $\mathbf{L}_n, \mathbf{\Gamma}_n, \mathbf{F}_n, \mathbf{E}_n$ and $\mathbf{E}^n$. We initialize $\mathbf{L}_0, \mathbf{\Gamma}_0, \mathbf{F}_0$ and $\mathbf{E}_0$ as follows:

- $\mathbf{L}_0 = \mathcal{L}_C$.
- $\mathbf{\Gamma}_0 \supseteq \hat{\mathbf{\Gamma}}$, is a maximal consistent set of closed wffs. (We remark that consistency is defined as usual and if $\mathbf{\Gamma}$ is maximal consistent, then for all $\alpha \in \mathbf{L}$ we have $\alpha \in \mathbf{\Gamma}$ or $\neg\alpha \in \mathbf{\Gamma}$.)
- $\mathbf{F}_0 = \emptyset$.
- $\mathbf{E}_0$ is an enumeration of pairs of wffs in $\mathbf{L}_0$ defined as $\{\langle\alpha,\beta\rangle \mid \neg(\alpha \rightarrow_{(\bar{x})} \beta) \in \mathbf{\Gamma}_0\}$.
- $\mathbf{E}^0 = \mathbf{E}_0$.

DEFINITION A.1
A **type** $\Gamma(x_1, \ldots, x_k)$ in the variables $x_1, \ldots, x_k$ is a maximal consistent set of formulas of $\mathbf{L}$ in these variables. [8]

Assume that we have defined $\mathbf{L}_n, \mathbf{\Gamma}_n, \mathbf{F}_n, \mathbf{E}_n$ and $\mathbf{E}^n$. Let $\langle\alpha,\beta\rangle$ be the $(n+1)$th pair in the enumeration $\mathbf{E}^n$. Define $\mathrm{NC}(\alpha)_n^k = \{\gamma : (\alpha \rightarrow_{(\bar{x})} \gamma) \in \mathbf{\Gamma}_n \text{ and } |\bar{x}| = k\}$ and let $\tau$ be a type in the variables $\bar{x}$ with respect to $\mathbf{\Gamma}_n$ *in* $\mathbf{L}_n$ (i.e. containing $\mathbf{\Gamma}_n$), such that $\mathrm{NC}(\alpha)_n^k \cup \{\neg\beta\} \subset \tau$.

LEMMA A.2
Such a type $\tau$ exists

PROOF. We must show that $\mathbf{\Gamma}_n \cup \{\neg\beta\} \cup \mathrm{NC}(\alpha)_n^k$ is consistent. We shall prove that
$\mathbf{\Gamma}_n \vdash \exists\bar{x}(\neg\beta \wedge \gamma_1 \wedge \ldots \wedge \gamma_m)$ for any finite subset $\{\gamma_1, \ldots, \gamma_m\} \subset \mathrm{NC}(\alpha)_n^k$.
We know that $\alpha \rightarrow_{(\bar{x})} \gamma_1, \ldots, \alpha \rightarrow_{(\bar{x})} \gamma_m \in \mathbf{\Gamma}_n$. Therefore by using **And** we obtain $\alpha \rightarrow_{(\bar{x})} (\gamma_1 \wedge \ldots \wedge \gamma_n) \in \mathbf{\Gamma}_n$ (by maximal consistency). $\langle\alpha,\beta\rangle \in \mathbf{E}_j$ for some $j \leq n \implies \neg(\alpha \rightarrow_{(\bar{x})} \beta) \in \mathbf{\Gamma}_n$ (by definition of $\mathbf{E}_j$).
By A.2 and $\mathbf{\Gamma}_n$'s maximal consistency we have $\exists\bar{x}(\neg\beta \wedge \gamma_1 \wedge \ldots \wedge \gamma_m) \in \mathbf{\Gamma}_n$. ∎

We can now define the $(n+1)$th stage in our construction.

- $\mathbf{L}_{n+1} = \mathbf{L}_n \cup c_0^{n+1}, c_1^{n+1}, \ldots, c_{k-1}^{n+1}$ where $c_m^{n+1} \notin \mathbf{L}_n$ for all $m < k$.
- $\mathbf{\Gamma}_{n+1} = \mathbf{\Gamma}_n \cup \tau\left[{x_0, x_1, \ldots, x_{k-1} \atop c_0^{n+1}, \ldots c_{k-1}^{n+1}, \ldots, c_{k-1}^{n+1}}\right]$ where we have assumed $x_0, \ldots, x_{k-1}$ are the variables occurring in $\tau$.
- $\mathbf{F}_{n+1} = \mathbf{F}_n \cup f_{n+1}^{\alpha}$ where $f_{n+1}^{\alpha} = (c_0^{n+1}, \ldots, c_{k-1}^{n+1})$.
- $\mathbf{E}_{n+1}$ is an arbitrary enumeration of pairs $\langle\alpha',\beta'\rangle$ s.t. $(\alpha' \wedge \beta') \in \mathbf{L}_{n+1} \setminus \mathbf{L}_n$ and $\neg(\alpha' \rightarrow_{(\bar{x})} \beta') \in \mathbf{\Gamma}_{n+1}$.

- $\mathbf{E}^{n+1}$ is the enumeration obtained by using Cantor's diagonalization method to order the elements of $\mathbf{E}_0, \ldots, \mathbf{E}_{n+1}$.

Define $\mathbf{L} = \bigcup \mathbf{L}_n$, $\mathbf{\Gamma} = \bigcup \mathbf{\Gamma}_n$, $\mathbf{F} = \bigcup \mathbf{F}_n$. It is clear that $\mathbf{\Gamma}$ is maximal consistent in $\mathbf{L}$, because each of the $\mathbf{\Gamma}_n$ is maximal consistent in $\mathbf{L}_n$ for each $n \in \mathbf{N}$ and the union itself is consistent (since $\mathbf{\Gamma}_n \subset \mathbf{\Gamma}_{n+1}$).

LEMMA A.3

1. For each $\mathbf{\Gamma}$-consistent wff $\alpha$ we have a tuple $f^\alpha \in \mathbf{F}$ such that $|f^\alpha| = |\bar{x}|$ and for any $\gamma$ s.t. $\alpha \rightarrow_{(\bar{x})} \gamma \in \mathbf{\Gamma}$, we have that $\gamma[{\bar{x} \atop f^\alpha}] \in \mathbf{\Gamma}$.
   We shall call $f^\alpha$ **normal**$_{\bar{x}}$ for $\alpha$.
2. For each $\alpha$ and $\beta$ s.t. $\neg(\alpha \rightarrow_{(\bar{x})} \beta) \in \mathbf{\Gamma}$, we have some $f$ that is normal$_{\bar{x}}$ for $\alpha$ such that $\neg\beta[{\bar{x} \atop f}] \in \mathbf{\Gamma}$.

PROOF. If $\alpha$ is $\mathbf{\Gamma}$-consistent then $\neg(\alpha \rightarrow_{(\bar{x})} \neg\alpha) \in \mathbf{\Gamma}$, otherwise by **Reflexivity** and **And** one has $\alpha \rightarrow_{(\bar{x})}$ **false** $\in \mathbf{\Gamma} \implies \forall\bar{x}\neg\alpha \in \mathbf{\Gamma}$, contradicting $\alpha$'s $\mathbf{\Gamma}$-consistency.
Therefore, it is enough to prove the second part of the lemma.
There exist some $n$ s.t. $\alpha, \beta \in \mathbf{L}_n \setminus \mathbf{L}_{n-1}$ and due to our construction process, this means $\neg(\alpha \rightarrow_{(\bar{x})} \beta) \in \mathbf{\Gamma}_n \setminus \mathbf{\Gamma}_{n-1}$. The fact that $\neg(\alpha \rightarrow_{(\bar{x})} \beta) \in \mathbf{\Gamma}_n$ implies that $\langle\alpha,\beta\rangle \in \mathbf{E}_n$. Therefore, there exists some $m > n$ s.t. we have chosen to deal with the pair $\langle\alpha,\beta\rangle$ in the $m$th stage.

---

[8] This definition is from [8].

LEMMA A.4
$f_m^\alpha$ is normal for $\alpha$ and $\neg\beta[^{x}_{f_m^\alpha}] \in \mathbf{\Gamma}$.

PROOF. By the construction process we know that $\neg\beta[^{x}_{f_m^\alpha}] \in \mathbf{\Gamma}_m \subset \mathbf{\Gamma}$ , and that for each $\gamma$ s.t. $\alpha \rightarrow_{(x)} \gamma \in \mathbf{\Gamma}_{m-1}$, we have $\gamma[^{x}_{f_m^\alpha}] \in \mathbf{\Gamma}_m \subset \mathbf{\Gamma}$. Assume $\delta$ is such that $\alpha \rightarrow_{(x)} \delta \in \mathbf{\Gamma} \setminus \mathbf{\Gamma}_m$. Therefore there exists some $k \geq m$ s.t. $\alpha \rightarrow_{(x)} \delta \in \mathbf{\Gamma}_k \setminus \mathbf{\Gamma}_{k-1}$. This means that $\delta = \delta'(\bar{x}, \bar{c})$ where $\bar{c} \in \mathbf{L}_k \setminus \mathbf{L}_m$. But we know by virtue of $\alpha[^{x}_{f_m^\alpha}] \in \mathbf{\Gamma}_m$ that $\exists\bar{x}\alpha \in \mathbf{\Gamma}_n$ (because $\mathbf{\Gamma}_n$ is maximal consistent and $\exists\bar{x}\alpha \in \mathbf{L}_n$). By A.3 we know that $(\alpha \rightarrow_{(x)} \forall\bar{y}(\delta'(\bar{x},\bar{y}) \vee \neg(\alpha \rightarrow_{(x)} \delta'(\bar{x},\bar{y})))) \in \mathbf{\Gamma}_n$ too.

Therefore, $\forall\bar{y}(\delta'(\bar{x},\bar{y}) \vee \neg(\alpha \rightarrow_{(x)} \delta'(\bar{x},\bar{y})))[^{x}_{f_m^\alpha}] \in \mathbf{\Gamma}_m$. By substitution (of FOL) and maximal consistency, we know $\delta'(\bar{x},\bar{c}) \vee \neg(\alpha \rightarrow_{(x)} \delta'(\bar{x},\bar{c}))[^{x}_{f_m^\alpha}]$ is in $\mathbf{\Gamma}_m$, and by maximal consistency and the fact that $\neg(\alpha \rightarrow_{(x)} \delta'(\bar{x},\bar{c})) \notin \mathbf{\Gamma}_k$ we know that $\delta'(\bar{x},\bar{c})[^{x}_{f_m^\alpha}] \in \mathbf{\Gamma}_k \subset \mathbf{\Gamma}$. ∎

This proves the whole lemma. ∎

We can now define our model $\mathcal{M} \stackrel{\text{def}}{=} \langle \mathbf{M}, \mathbf{R} \rangle$.

- $\mathbf{M}$ is the standard FOL structure defined by the standard part of $\mathbf{\Gamma}$ (as in [10]).
- $\mathbf{R} = \{R_n | n \in N\}$ where $R_k$ is a ranking function over $||M||^k$ that we will now construct.

*Fix some integer $n$. We will now construct the ranking $R_n$. The construction is identical for every $n$.*

DEFINITION A.5
Let $\mathbf{S}$ be the set of all $\mathbf{\Gamma}$-consistent wffs (both open and closed), fix some ordered set of $n$ (different) variables $\bar{x}$ and let $\mathcal{R} \in \mathbf{S} \times \mathbf{S}$ where $\alpha\mathcal{R}\beta$ iff $\neg(\alpha \vee \beta \rightarrow_{(x)} \neg\alpha) \in \mathbf{\Gamma}$.

LEMMA A.6
$\mathcal{R}$ is transitive. $(LM\star)$

PROOF. We have to show that $\neg((\alpha \vee \beta) \rightarrow_{(x)} \neg\alpha) \wedge \neg((\beta \vee \gamma) \rightarrow_{(x)} \neg\beta) \vdash \neg((\alpha \vee \gamma) \rightarrow_{(x)} \neg\alpha)$. This is equivalent to $((\alpha \vee \gamma) \rightarrow_{(x)} \neg\alpha) \wedge \neg((\beta \vee \gamma) \rightarrow_{(x)} \neg\beta) \vdash ((\alpha \vee \beta) \rightarrow_{(x)} \neg\alpha)$. By A.8 we have $(\alpha \vee \gamma \rightarrow_{(x)} \neg\alpha) \vdash (\alpha \vee \beta \vee \gamma \rightarrow_{(x)} \neg\alpha)$. By contraposition on A.7 we have $\neg((\beta \vee \gamma) \rightarrow_{(x)} \neg\beta) \vdash \neg((\alpha \vee \beta \vee \gamma) \rightarrow_{(x)} \neg(\alpha \vee \beta))$. By (**R.M.**) we obtain $((\alpha \vee \gamma) \rightarrow_{(x)} \neg\alpha) \wedge \neg((\beta \vee \gamma) \rightarrow_{(x)} \neg\beta) \vdash (((\alpha \vee \beta) \wedge (\alpha \vee \beta \vee \gamma)) \rightarrow_{(x)} \neg\alpha)$, which of course implies the desired result. ∎

LEMMA A.7
For each $\alpha, \beta \in \mathbf{S}$, either $\alpha\mathcal{R}\beta$ or $\beta\mathcal{R}\alpha$ . $(LM\star)$

PROOF. Suppose $\alpha\not\mathcal{R}\beta$ and $\beta\not\mathcal{R}\alpha$ . Then we have $((\alpha \vee \beta) \rightarrow_{(x)} \neg\alpha) \wedge ((\alpha \vee \beta) \rightarrow_{(x)} \neg\beta) \in \mathbf{\Gamma}$. By **And** and **Reflexivity** we have $((\alpha \vee \beta) \rightarrow_{(x)} \neg\alpha) \wedge \neg\beta \wedge (\alpha \vee \beta) \in \mathbf{\Gamma}$. Therefore, $((\alpha \vee \beta) \rightarrow_{(x)} \textbf{false}) \in \mathbf{\Gamma}$. Using A.6 this implies that $(((\alpha \vee \beta) \wedge \beta) \rightarrow_{(x)} \textbf{false}) \in \mathbf{\Gamma}$, and using **L.E.** one has $\beta \rightarrow_{(x)} \textbf{false} \in \mathbf{\Gamma}$. Using **Instantiation** we have $\exists\bar{x}\beta \Rightarrow \textbf{false} \in \mathbf{\Gamma}$. Consequently, $\forall\bar{x}\neg\beta \in \mathbf{\Gamma}$ implying $\beta \notin \mathbf{S}$. ∎

LEMMA A.8
If $\alpha\mathcal{R}\beta$ then for any $f$ normal$_x$ for $\alpha$ s.t. $\beta[^{x}_{f}] \in \mathbf{\Gamma}$, $f$ is normal$_x$ for $\beta$ . $(LM\star)$

PROOF. Suppose $\alpha\mathcal{R}\beta$ , $f$ is normal$_x$ for $\alpha$, and $\beta[^{x}_{f}] \in \mathbf{\Gamma}$. Let $\beta \rightarrow_{(x)} \gamma \in \mathbf{\Gamma}$. We must show that $\gamma[^{x}_{f}] \in \mathbf{\Gamma}$. It is sufficient to prove that $(\alpha \rightarrow_{(x)} (\beta \Rightarrow \gamma)) \in \mathbf{\Gamma}$, because $f$ is normal$_x$ for $\alpha$ . But by **L.E.** we have $\beta \rightarrow_{(x)} \gamma \vdash (\alpha \vee \beta) \wedge \beta \rightarrow_{(x)} \gamma$. By A.4 we have $(\alpha \vee \beta) \rightarrow_{(x)} (\beta \Rightarrow \gamma) \in \mathbf{\Gamma}$. $\alpha\mathcal{R}\beta$ implies that $\neg(\alpha \vee \beta \rightarrow_{(x)} \neg\alpha) \in \mathbf{\Gamma}$ , and **R.M.** implies that $(\alpha \vee \beta) \wedge \alpha \rightarrow_{(x)} (\beta \rightarrow \gamma)$ , giving us the desired result by use of **L.E.**. ∎

DEFINITION A.9
$\alpha \sim \beta$ for $\alpha, \beta \in \mathbf{S}$ iff $\alpha\mathcal{R}\beta$ and $\beta\mathcal{R}\alpha$ .

We can conclude from the previous lemmas that $\sim$ is an equivalence relation. We denote the equivalence class of $\alpha$ by $\bar{\alpha}$, we will denote the set of equivalence classes by $\mathbf{E}$. The relationship $\mathcal{R}$ on formulas induces an ordering over equivalence classes in $\mathbf{E}$ defined below.

DEFINITION A.10

- $\bar{\alpha} \stackrel{\text{def}}{=} \{\gamma \in \mathbf{S} : \alpha \sim \gamma\}$,

- $\mathbf{E} \stackrel{\text{def}}{=} \{\bar{\alpha} : \alpha \in \mathbf{S}\}$,
- $\bar{\alpha} \leq \bar{\beta}$ iff $\alpha \mathcal{R} \beta$,
- $\bar{\alpha} < \bar{\beta}$ iff $\bar{\alpha} \leq \bar{\beta}$ and $\alpha \not\sim \beta$.

The previous lemmas imply that $<$ is a strict total order on $\mathbf{E}$.

LEMMA A.11
If $\alpha, \beta \in \mathbf{S}$ and $\bar{\beta} < \bar{\alpha}$ then $\beta \rightarrow_{(\bar{x})} \neg\alpha$. $(LM\star)$

PROOF. $\bar{\beta} < \bar{\alpha}$ implies $\alpha \not\mathcal{R} \beta$ which implies $(\alpha \vee \beta \rightarrow_{(\bar{x})} \neg\alpha)$. By A.2 we have that $\beta \rightarrow_{(\bar{x})} \neg\alpha$. ∎

LEMMA A.12
Let $\alpha, \beta \in \mathbf{S}$. If there exists some $f \in \mathbf{F}$ such that $f$ is normal$_{\bar{x}}$ for $\alpha$ and $\beta[^{\bar{x}}_{f}] \in \Gamma$, then $\bar{\beta} \leq \bar{\alpha}$. $(LM\star)$

PROOF. If so, then $\neg(\alpha \rightarrow_{(\bar{x})} \neg\beta)$. By Lemma A.11 $\alpha \not< \beta$. ∎

We can now define $\mathbf{R}$ (actually $R_n$). Let $|f^{\alpha}| = |g^{\beta}| = |f| = |g| = n$.

- If $f^{\alpha}, g^{\beta} \in \mathbf{F}$ then $f^{\alpha} <_{\mathbf{R}} g^{\beta}$ iff $\bar{\alpha} < \bar{\beta}$.
- If $f \in \mathbf{F}$, $g \in \bar{\mathbf{F}} \setminus \mathbf{F}$ then $f <_{\mathbf{R}} g$.
- If both $f, g \in \bar{\mathbf{F}} \setminus \mathbf{F}$ then $f =_{\mathbf{R}} g$.

$\mathbf{R}$ is well defined because Lemma A.12 implies that if $f^{\alpha} \equiv g^{\beta}$ (i.e. $f$ was constructed as normal$_{\bar{x}}$ for $\alpha$ and $g$ was constructed as normal$_{\bar{x}}$ for $\beta$ and they are identical) then $\bar{\alpha} = \bar{\beta}$. Also, notice that Axiom (3.11) implies that the choice of the variables $\bar{x}$ in the definition of normal$_{\bar{x}}$ does not matter, and that different rankings obtained using different sets of variables will be identical.

DEFINITION A.13
We say that $f$ satisfies$_{\bar{x}}$ $\alpha$, written $f \models_{\bar{x}} \alpha$ iff $\alpha[^{\bar{x}}_{f}] \in \Gamma$. We say that $f$ is minimal$_{\bar{x}}$ for $\alpha$ iff $f \models_{\bar{x}} \alpha$ and $g \models_{\bar{x}} \alpha$ implies that $f \leq_{\mathbf{R}} g$. When $\bar{x}$ is clear from the context we will simply say that $f$ satisfied $\alpha$ or that $f$ is minimal for $\alpha$ and write $f \models \alpha$.

LEMMA A.14
If $f$ is minimal for $\alpha$, then f∈F. $(LM\star)$

PROOF. Suppose $f$ is minimal for $\alpha$, therefore $f \models \alpha$ which implies $\alpha \in \mathbf{S}$. This means that there is a normal$_{\bar{x}}$ function $g$ for $\alpha$ in $\mathbf{F}$ (by construction of $\mathbf{F}$). Therefore, $f \leq_{\mathbf{R}} g$ and hence, $f \in \mathbf{F}$. ∎

We conclude that any minimal function is normal for some wff.

LEMMA A.15
Let $f \in \mathbf{F}$ be normal for $\alpha$. Then $f$ is minimal in $\hat{\beta}$ iff $f \models \beta$ and $\bar{\beta} = \bar{\alpha}$. $(LM\star)$

PROOF. First, assume $f$ is minimal in $\hat{\beta}$. Because $f$ is normal for $\alpha$ and $f \models \beta$ we have (by Lemma A.12) that $\bar{\beta} \leq \bar{\alpha}$. Since $\beta \in \mathbf{S}$, we have some $g \in \mathbf{F}$ (by construction of $\mathbf{F}$), s.t. $g$ is normal for $\beta$. By minimality of $f$ in $\hat{\beta}$, we have $f \leq_{\mathbf{R}} g$. This, by definition of $\mathbf{R}$ means $\bar{\beta} \not< \bar{\alpha}$ implies $\bar{\beta} = \bar{\alpha}$.

For the other direction, assume $\bar{\beta} = \bar{\alpha}$ and $f \models \beta$. Assume $g \in \mathbf{F}$ is normal for $\gamma$ and $g < f$. Therefore, $\bar{\gamma} < \bar{\alpha}$ which implies $\bar{\gamma} < \bar{\beta}$. By Lemma A.11, $\gamma \rightarrow_{(\bar{x})} \neg\beta$ and thus, $g \notin \hat{\beta}$. ∎

LEMMA A.16
Let $\alpha \in \mathbf{S}$. $\hat{\alpha}$ has a minimal element. $(LM\star)$

PROOF. Let $g \in \mathbf{F}$ be any tuple normal for $\alpha$. Since $g \models \alpha$ and $\bar{\alpha} = \bar{\alpha}$, we have by Lemma A.15 that $g$ is minimal in $\hat{\alpha}$. However, $\bar{\alpha} \leq \bar{\beta}$, hence $g < f$. ∎

LEMMA A.17
If $g \in \mathbf{F}$ is normal for $\alpha$ and minimal in $\hat{\beta}$, then $g$ is normal for $\beta$. $(LM\star)$

PROOF. $g$ is minimal in $\hat{\beta}$. Therefore by Lemma A.15, $\bar{\beta} = \bar{\alpha}$ and hence, $\alpha \mathcal{R} \beta$. We also have that $g$ is normal for $\alpha$ and $g \models \beta$. Therefore (by Lemma A.8), $g$ is normal for $\beta$. ∎

Note that $g \in \mathbf{F} \implies \exists \alpha \in \mathbf{S}$ s.t. $g$ is normal for $\alpha$. Because Lemma A.15 tells us that if $g$ is minimal for $\alpha$ it is in $F$, we can conclude that if $g$ is minimal for $\alpha$ then $g$ is normal for $\alpha$.

**Proof of completeness theorem:** We now show that $\mathcal{M} \models \mathbf{\Gamma}$. Since our original set of wffs $\hat{\mathbf{\Gamma}} \subset \mathbf{\Gamma}$, this is enough. By induction on the structure of $\alpha \in \mathbf{\Gamma}$ we show that $\mathcal{M} \models \alpha \Leftrightarrow \alpha \in \mathbf{\Gamma}$. We have implicitly assumed that $\{=\} \not\in \mathbf{L}$. The generalization to languages containing equality proceeds as in the Henkin's proof for standard first-order logic. (Details of that proof can be found in most texts on mathematical logic. Enderton's [10] is a good choice.) There are a number of cases we must consider, but because we basically augmented Henkin's proof to deal with the conditional operator, the standard steps are conducted as in Henkin's proof. Hence we only mention the additional steps that are required here.

- $\alpha$ is an atomic formula, a negation ($\alpha = \neg\beta$) or an implication ($\alpha = \beta \Rightarrow \gamma$). We proceed as in Henkin's completeness proof for FOL (see [10] (pp. 132–133)).
- $\alpha = \forall x \beta$. To follow Henkin's proof we need to extend the substitution lemma to our language, and to use the following lemma:

  LEMMA A.18

  For each wff $\varphi(x) \in \mathbf{L}$ we have some $c \in \mathbf{L}$ s.t. $\neg\forall x\varphi \rightarrow \neg\varphi[^x_c] \in \mathbf{\Gamma}$.

  PROOF. Assume $\neg\forall x\varphi \in \mathbf{\Gamma}$. $\implies \exists x \neg\varphi \in \mathbf{\Gamma}$. $\implies \neg\varphi \in \mathbf{S}$ (the set of $\mathbf{\Gamma}$-consistent wffs). Hence, $\neg(\neg\varphi \rightarrow_{(x)} \varphi) \in \mathbf{\Gamma}$. Otherwise, using $\exists x \neg\varphi \in \mathbf{\Gamma}$ and (3.10) we would conclude *false*. Therefore, at some stage in the construction we must have added a constant $\bar{c}$ such that $\neg\varphi[^x_{\bar{c}}] \in \mathbf{\Gamma}$. ∎

  The generalization of the substitution lemma is straightforward since the conditional operator behaves much like the universal quantifier.
- $\alpha = (\beta \rightarrow_{(x)} \gamma)$. First, assume that $\beta \rightarrow_{(x)} \gamma \in \mathbf{\Gamma}$. if $\beta \in \mathbf{S}$ we know by Lemma A.2 that $\exists f \in \mathbf{F}$ normal for $\beta$. By Lemma A.17 we know that for any $g \in \mathbf{F}$ minimal in $\hat{\beta}$, it is the case that $g$ is normal for $\beta$. Hence, for all $g \in \hat{\beta}$ that are minimal in $\hat{\beta}$, we have that $\gamma[^x_g] \in \mathbf{\Gamma}$. By the inductive hypothesis $\mathcal{M} \models \gamma[^x_g][\hat{s}]$. We conclude that $\mathcal{M} \models \beta \rightarrow_{(x)} \gamma$. If $\beta \not\in \mathbf{S}$ then $\mathcal{M} \models \beta \rightarrow_{(x)} \gamma$ trivially.

  Assume now that $\beta \rightarrow_{(x)} \gamma \not\in \mathbf{\Gamma}$. By maximal consistency $\neg(\beta \rightarrow_{(x)} \gamma) \in \mathbf{\Gamma}$. This implies (by A.1) that $\beta \in \mathbf{S}$. Lemma A.3 shows that there exists $f \in \mathbf{F}$ normal for $\beta$, s.t. $\neg\gamma[^x_f] \in \mathbf{\Gamma}$. Lemma A.15 implies that $f$ is minimal in $\hat{\beta}$. This implies (by using the inductive hypothesis to show that $\mathcal{M} \models \neg\gamma[^x_f][\hat{s}]$) that $\mathcal{M} \not\models \beta \rightarrow_{(x)} \gamma[\hat{s}]$. $\implies \mathcal{M} \models \neg(\beta \rightarrow_{(x)} \gamma)[\hat{s}]$.

Finally, we note that in order to imitate Henkin's proof we need to establish the semantic and syntactic equivalence of alphabetic variants. The syntactic equivalence is proven using Axiom (3.11) (**Renaming**) for the case of normally quantified wffs. The semantical equivalence follows from soundness. ∎

THEOREM 3.4

For a countable language $\mathcal{L}_C$

1. Axioms (3.1)–(3.13) are complete w.r.t. the class of NS structures satisfying *permutation*.
2. Axioms (3.1)–(3.17) are complete w.r.t. the class of NS structures satisfying *permutation* and *concatenation*.

PROOF. 1. Suppose that $\bar{c} \circ \bar{d}$ is normal$_{\bar{x},\bar{y}}$ for $\alpha(\bar{x}, \bar{y})$. First, we show that $\bar{d} \circ \bar{c}$ is normal$_{\bar{x},\bar{y}}$ for $\alpha(\bar{y}, \bar{x})$. We know that $\alpha(\bar{y}, \bar{x})[^{\bar{x},\bar{y}}_{\bar{d},\bar{c}}] \in \Gamma$. If $\alpha(\bar{y}, \bar{x}) \rightarrow_{(\bar{x},\bar{y})} \beta(\bar{y}, \bar{x})$ then by (3.13) and (3.11) we get $\alpha(\bar{x}, \bar{y}) \rightarrow_{(\bar{x},\bar{y})} \beta(\bar{x}, \bar{y})$. Because $\bar{c} \circ \bar{d}$ is normal$_{\bar{x},\bar{y}}$ for $\alpha(\bar{x}, \bar{y})$ we have $\beta(\bar{x}, \bar{y})[^{\bar{x},\bar{y}}_{\bar{c},\bar{d}}] \in \Gamma$. This, in turn, implies that $\beta(\bar{y}, \bar{x})[^{\bar{y},\bar{x}}_{\bar{d},\bar{c}}] \in \Gamma$.

Next, we show that $\alpha(\bar{y}, \bar{x}) \sim \alpha(\bar{x}, \bar{y})$. If $(\alpha(\bar{x}, \bar{y}) \vee \alpha(\bar{y}, \bar{x})) \rightarrow_{(\bar{x},\bar{y})} \neg\alpha(\bar{x}, \bar{y})$, then using (3.11) and (3.13) we obtain $(\alpha(\bar{x}, \bar{y}) \vee \alpha(\bar{y}, \bar{x})) \rightarrow_{(\bar{x},\bar{y})} \neg\alpha(\bar{y}, \bar{x})$. Which using **And** results in a contradiction to **Reflexivity**.

Since $\bar{c} \circ \bar{d}$ is normal$_{\bar{x},\bar{y}}$ for $\alpha(\bar{x}, \bar{y})$, and $\bar{d} \circ \bar{c}$ is normal$_{\bar{x},\bar{y}}$ for $\alpha(\bar{y}, \bar{x})$, their rank must be identical.

If $\bar{c} \circ \bar{d}$ is not normal for any formula, then it has some rank which is higher than any other rank, as does $\bar{d} \circ \bar{c}$.

2. First, we modify the construction of the ranking described earlier in the proof of Theorem 3.3. In that proof, we said that if $f, g \not\in F$ then $f =_{\mathbf{R}} g$. We now redefine $\mathbf{R}$ (actually $R_n$) as follows: let $|f| = |g| = |f^{\alpha}| = |g^{\beta}| = n$.

- If $f^{\alpha}, g^{\beta} \in \mathbf{F}$ then $f^{\alpha} <_{\mathbf{R}} g^{\beta}$ iff $\hat{\alpha} < \hat{\beta}$.
- If $f \in \mathbf{F}$, $g \in \hat{\mathbf{F}} \setminus \mathbf{F}$ then $f <_{\mathbf{R}} g$.
- If both $f, g \in \hat{\mathbf{F}} \setminus \mathbf{F}$ then let $k$ be the number of domain elements comprising $f$ that are normal for some formula and let $m$ be the number of domain elements comprising $g$ that are normal for some formula:
  - if $k > n$ then $f <_{\mathbf{R}} g$;
  - if $k = n$ then $f =_{\mathbf{R}} g$;
  - if $k < n$ then $f >_{\mathbf{R}} g$.

The relevant lemmas remain unchanged. We now show that the model obtained satisfies *concatenation*. (Proof of *permutation* remains unchanged.)

First, we prove the following claim:

CLAIM A.19
If $\bar{c} \circ \bar{d}$ is normal for some wff $\alpha$ then there is some wff for which $\bar{c}$ is normal.

PROOF. We claim that $\bar{c}$ is normal for $\alpha[^{\bar{y}}_{\bar{d}}]$. Suppose that $\alpha \rightarrow_{(\bar{x})} \gamma[^{\bar{y}}_{\bar{d}}] \in \mathbf{\Gamma}$, we have to show that $\gamma[^{\bar{x},\bar{y}}_{\bar{c},\bar{d}}] \in \mathbf{\Gamma}$. From Axiom (3.17) we have that $\alpha \rightarrow_{(\bar{x},\bar{y})} ((\alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow \gamma)) \in \mathbf{\Gamma}$. We know that $\bar{c} \circ \bar{d}$ is normal for $\alpha$ and therefore, $(\alpha \rightarrow_{(\bar{x})} \gamma) \Rightarrow \gamma[^{\bar{x},\bar{y}}_{\bar{c},\bar{d}}] \in \mathbf{\Gamma}$. Thus, if $\alpha \rightarrow_{(\bar{x})} \gamma[^{\bar{x},\bar{y}}_{\bar{c},\bar{d}}] \in \mathbf{\Gamma}$, we are done. But this is equivalent to $\alpha \rightarrow_{(\bar{x})} \gamma[^{\bar{y}}_{\bar{d}}]$, since $\bar{x}$ is already bound in $\alpha \rightarrow_{(\bar{x})} \gamma$ by the conditional. (See our discussion of syntactical substitution before the proof of Theorem 3.3.) ∎

Notice that the converse also holds.

CLAIM A.20
If $\bar{c} \in \mathbf{F}$ $\bar{d} \in \mathbf{F}$ then $\bar{c} \circ \bar{d} \in \mathbf{F}$.

PROOF. Let $\bar{c}$ be normal for $\alpha$ and $\bar{d}$ normal for $\beta$, we claim that $\bar{c} \circ \bar{d}$ is normal for $\alpha \wedge \beta$. (We must take care to make fv($\alpha$) and fv($\beta$) disjoint.) This follows from Axiom (3.16). Suppose that $\alpha \wedge \beta \rightarrow_{(\bar{x},\bar{y})} \gamma \in \mathbf{\Gamma}$. Hence, we have that $\alpha \rightarrow_{(\bar{x})} (\beta \rightarrow_{(\bar{y})} \gamma) \in \mathbf{\Gamma}$. Because $\bar{c}$ is normal for $\alpha$, we have that $\beta \rightarrow_{(\bar{y})} \gamma[^{\bar{x}}_{\bar{c}}] \in \mathbf{\Gamma}$. Because $\bar{d}$ is normal for $\beta$ (and free-var($\beta$)$\cap$ fv($\alpha$)$= \emptyset$) we have that $\gamma[^{\bar{x},\bar{y}}_{\bar{c},\bar{d}}] \in \mathbf{\Gamma}$. ∎

Suppose that $\bar{c} \leq_{\mathbf{R}} \bar{d}$ and $\bar{c}' \leq_{\mathbf{R}} \bar{d}'$, we must show that $(\bar{c} \circ \bar{c}') \leq_{\mathbf{R}} (\bar{d} \circ \bar{d}')$ and that if $\bar{c} <_{\mathbf{R}} \bar{d}$ then $(\bar{c} \circ \bar{c}') <_{\mathbf{R}} (\bar{d} \circ \bar{d}')$.

First, suppose that either $\bar{d}$ or $\bar{d}'$ are not normal for any formula. Hence, by Claim A.1, neither is $\bar{d} \circ \bar{d}'$ normal for any formula. If both $\bar{c}$ and $\bar{c}'$ are normal for some formula, then by Claim A.2, we are done. Otherwise, if the ranks of $\bar{c}$ and $\bar{c}'$ and those of $\bar{d}$ and $\bar{d}'$ are the same, we immediately get that the ranks of $(\bar{c} \circ \bar{c}')$ and $(\bar{d} \circ \bar{d}')$ are identical, because the number of domain elements in each that are normal for some formula must be the same. If this is not the case, w.l.o g , $\bar{c} <_{\mathbf{R}} \bar{d}$. This implies that the number of normal domain elements in $\bar{c}$ is greater than the corresponding number for $\bar{d}$. Since $\bar{c}' \leq_{\mathbf{R}} \bar{d}'$, the number of normal domain elements in $\bar{c}'$ is greater or equal to the number of normal domain elements in $\bar{d}'$. Consequently, this number in $\bar{c} \circ \bar{c}'$ is greater than in $\bar{d} \circ \bar{d}'$.

Next, assume that both $\bar{d}$ and $\bar{d}'$ are normal for some wff. We know that if $f \leq_{\mathbf{R}} g$ and $g$ is normal for some wff $\alpha$ then $f \in \mathbf{F}$. This follows from the definition of $\leq_{\mathbf{R}}$. Hence, by definition of $\mathbf{F}$, there is some wff for which $f$ is normal. Therefore, we know that all of $\bar{c}, \bar{c}', \bar{d}, \bar{d}'$ are normal for some wff. We continue the proof under this assumption.

Let $\alpha$ be such that $\bar{c}$ is normal$_{\bar{x}}$ for $\alpha$ , and let $\bar{c}'$ be normal$_{\bar{y}}$ for $\alpha'$. Let $\beta$ be such that $\bar{d}$ is normal$_{\bar{x}}$ for $\beta$ , and let $\bar{d}'$ be normal$_{\bar{y}}$ for $\beta$. We claim that $\bar{c} \circ \bar{c}'$ is normal$_{\bar{x},\bar{y}}$ for $\alpha \wedge \alpha'$ and that $\bar{d} \circ \bar{d}'$ is normal$_{\bar{x},\bar{y}}$ for $\beta \wedge \beta'$ . This easily follows from (3.16), as we have shown in Claim A.2 above.

We must now show that $(\alpha \wedge \alpha')\mathcal{R}(\beta \wedge \beta')$. This easily follows by using (3.14).

The stronger relation follows from Axiom 3.15. ∎

THEOREM 4.3
For a countable language $\mathbf{L}$, axioms (3.1)–(3.12) are sound and complete w.r.t. the class of smooth, pointed PPF models.

PROOF. Soundness is easily verified. For Axioms (3.2)–(3.8) the proof proceeds as in the corresponding proof for $\epsilon$-semantics. The pointedness property is needed for Axiom (3.8) which contains a negated conditional. Axioms (3.9)–(3.11) are immediate. The only slightly problematic case is Axiom (3.12) for which the smoothness condition is needed. First, suppose that $\forall \bar{y}(\alpha \rightarrow_{(\bar{x})} \beta)$. To show that $\alpha \rightarrow_{(\bar{x})} \forall \bar{y}\beta$ we need to show that $Pr(\exists \bar{y} \neg \beta | \alpha) = 0$ in the limit. However, from $\forall \bar{y}(\alpha \rightarrow_{(\bar{x})} \beta)$ it follows that for every possible assignment $\bar{c}$ to $\bar{y}$, $Pr(\neg \beta[^{\bar{y}}_{\bar{c}}] | \alpha) = 0$ in the limit. Since the existential can be viewed as an infinite disjunction over these possible assignments, the result follows by smoothness. The other direction is immediate and does not require smoothness.

Recall that we assume no relation between the probability distributions defined on different tuple sizes of domain elements (i.e. the probability over singletons and pairs is unrelated). Therefore, whatever construction we use to construct the probability distribution on one tuple size can be used for any other tuple size. Therefore, we shall concentrate on the construction of a probability distribution over (single) domain elements.

Let $\Gamma$ be some consistent theory. From Theorem 3.3 we know that this theory has a model $\mathcal{M}$ containing a countable number of objects (this follows from the construction of the model). First, we enumerate the elements of this model's domain. This will also induce an enumeration of the domain elements within each rank. We will now construct a sequence of probability distributions, $Pr_1, Pr_2, \ldots$, such that for each $m \in N$, $Pr_m$ will assign positive probability only to a subset of the domain containing the elements of the $m$ most normal ranks. These distributions will always observe the constraint that the probability of each domain element conditional on the set of elements in its rank is equal to $1/2^k$ if it is the $k$th element of its rank and there is an infinite number of elements in this rank. If there is a finite number of elements in a rank, all of them will be equi-probable.

It remains to define the probability of each rank in each of $Pr_1, Pr_2, \ldots$. In $Pr_n$ we set the probability of the most likely rank to be $\beta_n$. We set the probability of all other ranks among the $n$ most normal ranks to be $(1/n)$ of the probability of the previous rank. $\beta_n$ is chosen so that the probabilities of these ranks assigned positive probability sum up to 1. All other ranks are assigned probability 0. From now on we ignore $\beta_n$ since it is irrelevant to the conditional probabilities, in which we are interested.

Notice that this construction has the following property. Let $c$ be some domain object and let $A$ be the set of all domain objects ranked higher than $c$. Then $\lim_{n\to\infty} Pr(A|A \cup \{c\}) = 0$. To see this, suppose that $Pr_n(c) = 1/2^k\beta$, where $\beta$ is the weight of $c$'s rank and $k$ is a constant. But, $Pr_n(A) \leq \beta \cdot \sum_{i=1}^{n} \frac{1}{n^i}$.

We must now show that this PPF models $\Gamma$. First, suppose that $\varphi \rightarrow_{(x)} \psi \in \Gamma$. This means that in $\mathcal{M}$ all minimal elements satisfying $\varphi$ satisfy $\psi$ as well. Suppose that the conditional probability of the minimal elements satisfying $\varphi$ w.r.t. their entire rank is $r$. Notice that this value is fixed and positive for all $Pr_i$, $1 \leq i < \infty$. Since any element satisfying $\varphi \wedge \neg\psi$ must be in a higher rank, we have that $Pr_n(\neg\psi|\varphi) \leq 2r/n$. Thus $\lim_{n\to\infty} Pr(\neg\psi|\varphi) = 0$.

Suppose now that $\varphi \rightarrow_{(x)} \psi \notin \Gamma$. This means that in the minimal rank of elements satisfying $\varphi$ there is one satisfying $\psi$. The conditional probability of this element w.r.t. its rank is some fixed $r > 0$. We have that

$$Pr_n(\neg\psi|\varphi) = Pr_n(\neg\psi|min_x\varphi) \cdot Pr_n(min_x\varphi|\varphi) \geq r \cdot \frac{r}{r + \sum_{i=1}^{\infty} \frac{1}{n^i}}$$

(where $min_x\varphi$ are those minimal elements satisfying $\varphi$). Thus, $\lim_{n\to\infty} Pr(\psi|\varphi) = r > 0$.

Finally, we have to show that our models are smooth and pointed. To see that the model is smooth, suppose that for some wff $\alpha$ $\lim_{n\to\infty} Pr_n(\beta|\alpha) = 0$ for every $\beta$ in some set of wffs $B$. By our construction, for this to be the case, the minimal element with property $\beta \wedge \alpha$ for any $\beta \in B$ is ranked higher than the minimal elements with property $\alpha$. Hence, that is true for the set of all elements with some property in $\{\beta \wedge \alpha | \beta \in B\}$. By our construction, the probability of this set of elements among the set of elements with property $\alpha$ approaches 0 in the limit.

In order to see that the model is pointed, we must show that if $\alpha, \beta \in \mathcal{L}_C$ then $\lim_{n\to\infty} Pr_n(\beta|\alpha)$ exists. Let $A$ be the set of elements with property $\alpha$ and $B$ be the set of elements with property $\alpha \wedge \beta$, and let $mA, mB$ be the respective set of minimal elements in $A$ and $B$. From our construction it is clear that $\lim_{n\to\infty} Pr_n(B|A) = \lim_{n\to\infty} Pr_n(mB|mA)$. It is easy to see that our construction guarantees that this limit exists. ∎

THEOREM 4.6
The axioms (i)–(vii) are sound and complete w.r.t. the class of **NS** structures.

PROOF. This proof relies on the expressive equivalence of $\mathcal{L}_C$ in $\mathcal{L}_N$ discussed in Section 4.2 under the assumption that both languages contain equality. For the completeness theorem this requires extending our completeness theorem to such languages. Once we add the appropriate first-order axioms for equality, our Henkin-style proof proceeds much in the same manner. As in the standard case, what is obtained is a model in which equality really only stands for an equivalence relation. But by moving into a model in which the objects are these equivalence classes instead of the original objects (the quotient structure), we obtain the desired representation. See [10] for additional details.

Completeness–We have seen that we can encode the conditionals of $\mathcal{L}_C$ in $\mathcal{L}_N$. That is, we can express in $\mathcal{L}_N$ statements that are satisfied in a smooth ranked structure iff the conditional statement is satisfied. Hence, if we show that we can derive from these seven axioms sentences that are equivalent to Axioms (3.1)–(3.12), we can conclude that any theory that is consistent w.r.t. (i)–(vii) will also be consistent w.r.t. axioms (3.1)–(3.12) in its $\mathcal{L}_c$ form. By our completeness theorem, this implies that it has a model. Since the definition of satisfiability coincides in both models (that is the conditions under which a $\mathcal{L}_C$ sentence is satisfied are identical to the conditions under which the equivalent $\mathcal{L}_N$ sentence is satisfied), we know that this is a model of our original theory.

(3.1) Identical to $(i)$.

(3.2) Translated, we have

$$\forall\bar{x}(Normal_{\bar{y}}^{\varphi(\bar{y})}(\bar{x}) \Rightarrow \varphi(\bar{x}))$$

which is exactly $(iii)$.

(3.3) Translated, we have

$$\forall\bar{y}(\alpha(\bar{y}) \Leftrightarrow \beta(\bar{y})) \Rightarrow (\forall\bar{x}(Normal_{\bar{y}}^{\alpha(\bar{y})}(\bar{x}) \Rightarrow \gamma(\bar{x})) \Leftrightarrow \forall\bar{x}(Normal_{\bar{y}}^{\beta(\bar{y})}(\bar{x}) \Rightarrow \gamma(\bar{x}))).$$

Hence, it is sufficient to show that:

$$\forall\bar{y}(\alpha(\bar{y}) \Leftrightarrow \beta(\bar{y})) \Rightarrow (\forall\bar{x}(Normal_{\bar{y}}^{\alpha(\bar{y})}(\bar{x}) \Leftrightarrow Normal_{\bar{y}}^{\beta(\bar{y})}(\bar{x}))).$$

But this is precisely $(iv)$.

(3.4) Translated, we have

$$\forall\bar{y}(\alpha(\bar{y}) \Rightarrow \beta(\bar{y})) \Rightarrow ((\forall\bar{x}Normal_{\bar{x}}^{\gamma}(\bar{x}) \Rightarrow \alpha(\bar{x})) \Rightarrow (\forall\bar{x}Normal_{\bar{x}}^{\gamma}(\bar{x}) \Rightarrow \beta(\bar{x}))).$$

This is valid in FOL. (Recall that $Normal_{\bar{x}}^{\gamma}$ is treated as a predicate name.)

(3.5) Using FOL, we get that $\exists\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}))$ implies that $(vi) \Rightarrow (vi')$. Next, we observe that $\exists\bar{y}\alpha(\bar{y})$ implies that $(ii) \wedge (vi) \Rightarrow (vi')$. Since it is easy to show that $\neg\exists\bar{y}\alpha$ implies the same, we conclude that $(ii) \wedge (vi) \Rightarrow (vi')$ is provable.

Given the above, we show the relation between $(vi')$ and cautious monotony (3.5). Assume that $\forall y(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \Rightarrow \beta(\bar{y}))$. Hence

$$\forall y(Normal_{\bar{x}}^{\alpha\wedge\beta(\bar{x})}(\bar{y}) \Rightarrow Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y})).$$

Thus, given that $\forall y(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \Rightarrow \gamma(\bar{y}))$, we obtain $\forall y(Normal_{\bar{x}}^{\alpha\wedge\beta(\bar{x})}(\bar{y}) \Rightarrow \gamma(\bar{y}))$. This is equivalent to the translation of (3.5).

(3.6) Immediate.

(3.7) This easily follows using $(v)$.

(3.8) Translated, we have:

$$[\forall y(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \Rightarrow \gamma(\bar{y})) \quad \wedge\neg\forall y(Normal_{\bar{x}}^{\alpha\wedge\beta(\bar{x})}(\bar{y}) \Rightarrow \gamma(\bar{y}))] \Rightarrow \forall y(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \Rightarrow \neg\beta(\bar{y})).$$

First, notice that the second conjunct in the antecedent of the implication implies that $\exists\bar{y}Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y})$ and $\exists\bar{y}Normal_{\bar{x}}^{\alpha(\bar{x})\wedge\beta(\bar{x})\wedge\neg\gamma(\bar{x})}(\bar{y})$. This follows using $(ii)$ and $(i)$. Next, notice that using $(vi)$, this same conjunct implies that $\forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})\wedge\beta(\bar{x})\wedge\neg\gamma(\bar{x})}(y) \Rightarrow Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}))$. Reasoning by cases, it is sufficient to show that if the antecedent is false we get a contradiction. Hence, suppose that $\exists\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}) \wedge \beta(\bar{y}))$. Using $(vi)$ this implies that $\forall\bar{y}(Normal_{\bar{x}}^{\alpha(\bar{x})\wedge\beta(\bar{x})}(y) \Rightarrow Normal_{\bar{x}}^{\alpha(\bar{x})}(\bar{y}))$. Together with the first conjunct of the antecedent this is inconsistent (since we can show that there exists an indiyidual under which both $\gamma$ and $\neg\gamma$ hold).

(3.9) Is a tautology of FOL.

(3.10) Follows from $(ii)$.

(3.11) Follows from $(vii)$.

(3.12) Follows from the fact that $\forall x, y(\alpha(x) \Rightarrow \beta(x,y)) \Rightarrow \forall x(\alpha(x) \Rightarrow \forall y\beta(x,y))$.

Soundness—It is easy to see that all axioms schema are valid in all NS structures. ∎

# B Notation

| | |
|---|---|
| $\circ$ | concatenation operator |
| $\mathcal{L}_C$ | the language |
| fv($\alpha$) | free variables of $\alpha$ |
| $\mathcal{M}$ | model |
| $\mathbf{M}$ | first-order structure |
| $\lvert\mathcal{M}\rvert, \lvert\mathbf{M}\rvert$ | the domain |
| $\lvert f\rvert, \lvert\bar{x}\rvert$ | the cardinality of tuple $f, \bar{x}$ |
| $\mathbf{R}$ | set of ranking functions |
| $\mathcal{M}_P$ | parameterized probabilistic first-order model |
| $\alpha \leq_{\bar{x}} \beta$ | $\neg((\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\alpha)$ |
| $\alpha <_{\bar{x}} \beta$ | $(\alpha \vee \beta) \rightarrow_{(\bar{x})} \neg\beta$ |
| $\widehat{\beta_{\bar{x}}}$ | $\{\bar{c} \in \lvert\mathbf{M}\rvert^{\lvert\bar{x}\rvert} : \mathcal{M} \models \beta[^{\bar{x}}_{\bar{c}}]\}$ |
| $\mathcal{M} \models_f \alpha$ | $\mathcal{M} \models \alpha[^{\bar{x}}_f]$ |
| $NC(\alpha)^k_n$ | $\{\gamma : (\alpha \rightarrow_{(\bar{x})} \gamma) \in \mathbf{\Gamma}_n$ and $\lvert\bar{x}\rvert = k\}$ |
| $f$ normal$_{\bar{x}}$ for $\alpha$ | $\gamma[^{\bar{x}}_f] \in \mathbf{\Gamma}$ whenever $\alpha \rightarrow_{(\bar{x})} \gamma \in \mathbf{\Gamma}$ |
| $\mathbf{S}$ | set of $\mathbf{\Gamma}$-consistent wffs |
| $\alpha\mathcal{R}\beta$ | $\neg(\alpha \vee \beta \rightarrow_{(\bar{x})} \neg\alpha) \in \Gamma$ |
| $\alpha \sim \beta$ | $\alpha\mathcal{R}\beta$ and $\beta\mathcal{R}\alpha$ |
| $\bar{\alpha}$ | $\{\gamma \in \mathbf{S} : \alpha \sim \gamma\}$ |
| $\mathbf{E}$ | $\{\bar{\alpha} : \alpha \in \mathbf{S}\}$ |
| $\bar{\alpha} \leq \bar{\beta}$ | $\alpha\mathcal{R}\beta$ |
| $\bar{\alpha} < \bar{\beta}$ | $\bar{\alpha} \leq \bar{\beta}$ and $\alpha \not\sim \beta$ |
| $f \models_{\bar{x}} \alpha$ | $\alpha[^{\bar{x}}_f] \in \mathbf{\Gamma}$ |

Received 9 July 1996