CDD: 160

ON A LOGIC FOR 'ALMOST ALL' AND 'GENERIC' REASONING[1]

PAULO A. S. VELOSO[2]

Progr. Eng. Sistemas e Computação, COPPE
Instituto de Matemática, UFRJ
RIO DE JANEIRO, RJ
BRASIL

veloso@cos.ufrj.br

Abstract: *Some arguments use 'generic', or 'typical', objects. An explanation for (some aspects of) this idea in terms of 'almost all' is suggested. The intuition of 'almost all' as 'but for a few exceptions' is rendered precise by means of ultrafilters. A logical system, with generalized quantifiers for 'almost all', is proposed as a basis for generic reasoning. This logic is monotonic, has a simple sound and complete deductive calculus, and is a conservative extension of classical first-order logic, with which it shares several properties. For generic reasoning, generic individuals are introduced and internalized as generic constants, thereby producing conservative extensions where one can reason about generic objects as intended. A many-sorted version of this logic is introduced to handle distinct notions of 'large' subsets. Other possible applications for this logic are indicated.*

---

[1]This paper is the revised text of an invited talk presented at IMLLAI'98 (International Meeting on Language, Logic and Artificial Intelligence), held in July 1998 in Fortaleza, Ceará, Brazil. The author wishes to thank the organizers for their kind invitation and support. Helpful discussions with Walter A. Carnielli and Antonio M. Sette are also gratefully acknowledged.

[2] Research partly sponsored by CNPq (Brazil): grant 301163/91-0.

Key-words: *Ultrafilter logic; 'almost all'; semantic; axiomatization; soundness; completeness; 'typical' object; generic constant; relative 'most', sorted ultrafilter logic; sorted 'almost all' and 'generic' reasoning.*

1 INTRODUCTION

The work examines a logical system with a generalized quantifier, whose intended interpretation is 'almost all'. The primary motivation is a qualitative approach to notions such as 'all but a few' and 'very few'. The intuition of 'almost all' as 'but for a few exceptions' is rendered precise by means of ultrafilters.

This ultrafilter logic is monotonic, has a simple sound and complete deductive system, and is a conservative extension of classical first-order logic, with which it shares several properties. For generic reasoning, generic individuals are introduced and internalized as generic constants, thereby producing conservative extensions where one can reason about generic objects as intended. A sorted version of this logic is introduced to handle distinct notions of 'large' subsets. Other possible applications for this logic are indicated.

The familiar "Tweety example" (Reiter (1980)) may be used to convey the main ideas underlying our approach. Consider the assertions

(1) "Birds 'generally' fly";

(2) "Tweety is a 'typical' bird".

We wish to express such assertions and reason about them in a formal manner. For the moment, let us concentrate on the first assertion.

One usually understands "Birds 'generally' fly" as "All birds, but for 'a few' exceptions, fly". The paraphrase of the former by "Almost all birds fly" and of the latter by "Very few birds do not fly" suggests explicating 'almost all' in terms of 'very few' or having a 'negligible' set of exceptions.

Considering a unary predicate F on a universe B of birds, we can express "All birds fly" [by $\forall$ x F(x)] or "Some birds fly" [by $\exists\nabla$ x F(x)] with the apparatus of classical first-order logic. To *express* such 'most' assertions formally, we introduce the new operator ∇, to express "Almost all birds fly" by ∇ x F(x).

To give a precise *meaning* to 'almost all' assertions such as ∇x F(x), we extend the usual notions by providing a family $\mathcal{U}$ of 'large' sets and stipulate that ∇x F(x) means that the set $\{b \in B : F(b)\}$ is in the family $\mathcal{U}$ as a rigorous counterpart for "the set of flying birds is large".

To *reason* about such 'most' assertions in a formal manner, we will set up a deductive system, extending the classical first-order predicate calculus.

As for the second assertion, we shall later suggest how to express being 'typical'. The idea is using 'almost all' to explain 'typical' and 'generic'. The desideratum is then being able to conclude the following assertion

(3) "Tweety does fly".

Ultrafilter logic is related to default logic (Reiter (1980)) and its variants (Antoniou (1997); Besnard (1989); Brewka (1991); Lukaziewicz (1990); Marek and Truszczynski (1993)) (as well as to belief revision (Gärdenfors (1988); Makinson and Gärdenfors (1991))). Indeed, they do have a large intersection in terms of applications, as indicated by benchmark examples, which was one of its motivations (Carnielli and Sette (1994); Schlechta (1995)). But, they are quite different logical systems, both technically and in terms of intended interpretation (Carnielli and Veloso (1997)).

Concerning technical aspects, ultrafilter logic is, as we shall see, monotonic and a conservative extension of classical first-order logic, in sharp contrast with the nonmonotonic nature of the approaches *via* defaults.

As for intended interpretation, one can perhaps phrase the difference in terms of positive and negative views (Sette, Carnielli and Veloso (1999)). Our approach favors a positive view, in the sense that we wish to express assertions such as (1) and (2). The default approach takes a negative view in the sense of interpreting such assertions as 'in the absence of information to the contrary'.

As a logic with generalized quantifiers, ultrafilter logic is connected to such extensions of first-order logic (Barwise and Feferman (1985); Keisler (1970)). It is also related to the tradition of analysis and formalization of language (Frege (1879); Tarski (1936); Church (1956)).

In the sequel, we first indicate some plausible motivation for the idea of giving precise counterparts for 'almost all' and 'almost none' in terms of (very) large and (very) small sets, respectively. This will provide the basis for interpreting our generalized quantifier. We then set up, both semantically and axiomatically, a logical system based on this idea. This system, shown to be sound and complete, proves to be a conservative extension of classical first-order logic. We then introduce the ideas of 'typical' and 'generic', examine some of its properties, and show how one can reason correctly about them within our formalism. We also indicate the desirability of having several notions of 'large' and how this can be simply formulated in a many-sorted version of our ultrafilter logic. We finally comment on some perspectives and directions for further work, specifically some interesting connections with fuzzy logic and inductive and empirical reasoning, which suggest the possibility of other applications for our ultrafilter logic.

2 ON 'ALMOST ALL' AND 'VERY FEW'

We will now motivate and outline our approach to making precise the notion of 'almost all'. Towards this goal, we will introduce some precise counterparts for 'almost all' and 'almost none' in terms of (very) large and (very) small sets, respectively, indicating their plausibility.

Returning to our initial example, recall that we propose to explicate "Birds 'generally' fly" as "The set $\underline{F} = \{b \in B : F(b)\}$ of flying birds is (very) large". For this purpose, we need a precise account of the notion of '(very) large'.

But, the complement $\underline{F}^c = \{b \in B : \neg F(b)\}$ is the set of non-flying birds, which may be regarded as the set of exceptions. This idea suggests the basic intuition: $\underline{F} \subseteq B$ is '(very) large' exactly when the set $\underline{F}^c$ of exceptions is '(very) small'. We thus require a rigorous notion of '(very) small' or 'tiny'.

We view a 'tiny' set X as one that is 'negligible' with respect to the universe, in the sense of being practically empty: $X \approx \varnothing$. We wish a precise and qualitative account of this notion.

Some – reasonable – criteria for subsets of a universe S to be considered tiny, or negligible with respect to S, are as follows:

($\varnothing$) the empty set $\varnothing$ is tiny $\{\varnothing \approx \varnothing\}$;

($\subseteq$) X is tiny, if $X \subseteq Y$ and Y is tiny $\{X \approx \varnothing$ if $X \subseteq Y$ and $Y \approx \varnothing\}$;

($\cup$) $X \cup Y$ is tiny, if both X and Y are tiny $\{X \cup Y \approx \varnothing$ if $X \approx \varnothing \approx Y\}$.

A family $\mathcal{I}$ of subsets of S satisfying these three properties forms what is called an *ideal* over S (Halmos (1972)). Some examples are the family $\wp_{\omega}(S)$ of the finite subsets of a universe S and families of unlikely subsets as those with measure (or probability) 0.

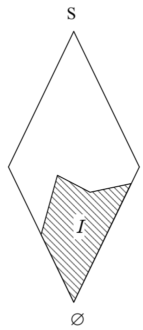

Figure 2.1. Ideal of '(very) small' subsets of universe S

These criteria suggest that a family of tiny subsets of a universe forms an ideal over the universe. Now, we do not wish to add features that may unduly restrict applicability. So, in the absence of other reasonable criteria on (very) small subsets and for the purpose of generality, we also adopt the converse view: any ideal is such a family, providing a notion of tiny subsets of a universe. As a net result, we propose taking the mathematical concept of ideal over a set as precise counterpart for the intuitive, and somewhat vague, idea of "(very) small subsets of a universe".

The dual idea is that of a '(very) large', or simply 'huge', set as one almost as large as the universe.

Complementation gives reasonable criteria for subsets of a universe S to be considered (very) large or huge.

(S) the universe S is huge $\{S \approx S\}$;

($\supseteq$) X is huge, if $X \supseteq Y$ and Y is huge $\{X \approx S$ if $X \supseteq Y$ and $Y \supseteq S\}$;

($\cap$) $X \cap Y$ is huge, if both X and Y are huge $\{X \cap Y \approx S$ if $X \approx S \approx Y\}$.

An intuitive way of understanding these criteria is by viewing a huge subset as one that has all the 'typical' elements of the universe [3].

As a family $\mathcal{F}$ of subsets of S with these three properties is a *filter* over S (Halmos (1972)), these criteria suggest that a notion of (very) large subsets of a universe is provided by a filter over the universe.

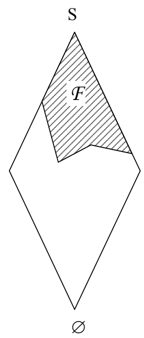

Figure 2.2: Filter of '(very) large' subsets of universe S

To summarize, over a given universe S, a notion of (very) large subsets can be provided by means of a filter $\mathcal{F}$ giving the huge subsets (with X≈S) or by an ideal $\mathcal{T}$ giving the tiny subsets (with Y ≈ ∅). In such case, we can use either the set $\underline{F} = \{b \in B : F(b)\}$ of flying birds or the set $\underline{F}^c = \{b \in B: \neg F(b)\}$ of non-flying birds, to interpret "Birds 'generally' fly" as

3 This intuitive interpretation was suggested by Hans Kamp during the discussion following the presentation at IMLLAI'98, Fortaleza, July 1998.

($\mathcal{F}$) The set of flying birds is (very) large {$\underline{F}\approx S$, i. e. $\underline{F}\in\mathcal{F}$},

($\mathcal{I}$) The set of non-flying birds is (very) small {$\underline{F}^c\approx\varnothing$, i. e. $\underline{F}^c\in\mathcal{I}$}.

Finally, we wish to have a decisive criterion for 'large' subsets. For this purpose, we will also require:

(^{c}) X is 'large' or X^c is 'large' {$X\approx S$ or $X^c\approx S$}.

The net result is that a notion of 'large' subsets of a universe S is provided by a family $\mathcal{U}$ of subsets of S with the following properties:

(S) the universe S is large {$S\in\mathcal{U}$};

($\supseteq$) X is large, if $X\supseteq Y$ and Y is large {$X\in\mathcal{U}$ if $X\supseteq Y$ and $Y\in\mathcal{U}$};

($\cap$) $X\cap Y$ is large, if both X and Y are large {$X\cap Y\in\mathcal{U}$ if X, $Y\in\mathcal{U}$};

(^{c}) X^c is large or X^c is large' {$X^c\in\mathcal{U}$ if $X\notin\mathcal{U}$}.

In other words, a family of large subsets of a universe S is a *proper ultrafilter* over S, which we shall generally call simply an *ultrafilter*.

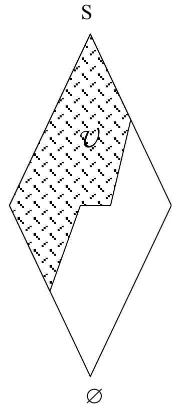

Figure 2.3: Ultrafilter of 'large' subsets of universe S

In general, there are many distinct (proper) ultrafilters over a given universe, each one providing a distinct notion of large.

We recall that an *ultrafilter* $\mathcal{U}$ is a filter that is maximal with respect to inclusion ($\mathcal{U}=\mathcal{F}$ for any filter $\mathcal{U}\subseteq\mathcal{F}$) (Bell and Slomson (1971); Chang and Keisler (1973)). A *proper* ultrafilter is one with $\mathcal{U}\neq\wp(S)$. Also, A *principal* filter is one that $\mathcal{F}$ is generated by some element $X\in\wp(S)$ ($\mathcal{F}=\{Y\subseteq S : X\subseteq Y\}$). So, a non-principal ultrafilter has no singleton.

A collection $\mathcal{F}\subseteq\wp(S)$ is said to have *fip* (short for *finite intersection property*) when $X_1\cap\ldots\cap X_n\neq\varnothing$ for any finite family $\{X_1,\ldots,X_n\}\subseteq\mathcal{F}$. A collection $\mathcal{F}\subseteq\wp(S)$ has fip iff it can be extended to some (proper) ultrafilter over S.

As examples of (non-principal) ultrafilters we mention the so-called *Fréchet ultrafilters*, which have no finite subset of a given (infinite) set.

Ultrafilters are also connected to a notion of measure in that what is outside an ultrafilter has 'measure zero'. A basic intuition underlying our approach is what lies outside an ultrafilter is small.

Returning to the example of flying birds, we understand "Birds 'generally' fly" as the set $\underline{F} = \{b \in B : F(b)\}$ of flying birds is 'large', which we formulate as $\underline{F} \in \mathcal{U}$. Thus, the set $\underline{F}^c = \{b \in B : \neg F(b)\}$ of non-flying birds is 'small' ($\underline{F}^c \notin \mathcal{U}$), but not necessarily empty. So, exceptions (non-flying birds) may very well exist, but the appearance of such exceptions does not force belief revision. The fact that $X \in \mathcal{U}$ implies $X^c \notin \mathcal{U}$ but not $X^c = \varnothing$ is the main reason why nonmonotonicity is not forced in this positive approach.

One should bear in mind that we are proposing properties of large sets that describe the notion of "family of large sets" in an axiomatic manner, rather than defining a specific family of 'large' sets. The situation is much as in, say, Algebra, when one axiomatizes groups or vector spaces.

A more pertinent analogy is, perhaps, with probability. Probability theory is more concerned with obtaining some probabilities from others, by means of properties, than with assigning probabilities to particular events.

The role of our proposed logic is more clearly understood in such a setting. One expresses available knowledge by some axioms, giving in particular some connections among large sets. From these axioms one concludes other properties. In other words, we deal not with a particular notion of large sets, but rather with a class of such notions as embodied in the axioms expressing the pertinent knowledge.

3 A LOGIC FOR ALMOST ALL

Our logic for most adds to classical first-order logic a generalized quantifier ∇, with intended interpretation 'almost all' and whose behavior will be seen to be intermediate between $\forall$ and $\exists$. We now examine this logic: syntax, semantics and axiomatics.

We consider a fixed signature (logical type) ρ with a repertoire of symbols for predicates, functions and constants. We also consider a denumerably infinite set V of new symbols for variables. We let L(ρ) be the usual first-order language (with equality ≡) of signature ρ, closed under the propositional connectives, as well as under the quantifiers ∀ and ∃.

3.1 Syntax of ∇

We use $L^{\nabla}(\rho)$ for the extension of the usual first-order language L(ρ) obtained by adding the new operator ∇.

The formulas of $L^{\nabla}(\rho)$ are built by the usual formation rules and the following new variable-binding *formation rule*

(∇) for each variable v∈V, if φ is a formula in $L^{\nabla}(\rho)$ then so is ∇ v φ.

Other usual syntactic notions, such as sentence, (free) substitution, etc., can be appropriately adapted.

We shall also employ the following notations, for a formula φ in L^{∇} (ρ):

- occ(φ) (fr(φ)) for the set of variables with (free) occurrences in φ,
- φ (v/t) for the result of substituting term t for all the free occurrences of variable v in φ.

When convenient and safe we may resort to a more informal notation for substitution: for a formula of the form φ(v) we write simply φ(t) for φ(v/t).

As an example illustrating the expressive power of such languages, consider a binary predicate L (with L(x, y) standing for x loves y).

We can express some assertions by means of purely first-order sentences, e. g. "Everybody loves somebody" by $\forall x \exists y L(x, y)$.

Some assertions expressed by means of the quantifier ∇ are as follows.

– "Almost everybody loves somebody" by $\nabla x \exists y L(x, y)$.
– "Somebody loves almost everybody" by $\exists x \nabla y L(x, y)$.
– "Everybody loves almost everybody" by $\forall x \nabla y L(x, y)$.
– "Almost everybody loves everybody" by $\nabla x \forall y L(x, y)$.
– "People generally love each other" in the sense of "Almost everybody loves almost everybody" by $\nabla x \nabla y L(x, y)$.

3.2 Semantics of ∇

The semantic interpretation for our logic of 'almost all' is provided by extending the usual first-order definition of satisfaction to the new quantifier ∇. For this purpose, we resort to ultrafilter structures: expansions of first-order structures by ultrafilters.

An *ultrafilter structure* $\mathcal{A}^{\mathcal{U}} = (\mathcal{A}, \mathcal{U})$ for signature ρ consists of a first-order structure $\mathcal{A}$ for signature ρ together with an ultrafilter $\mathcal{U}$ over the universe A of $\mathcal{A}$.

We extend the usual definition of *satisfaction* of a formula φ in a structure under an assignment $s: V \to A$ to variables ($\mathcal{A}^{\mathcal{U}} \models \varphi\ [s]$) as follows

($\models \nabla$) for a most formula $\nabla v\ \varphi$, we define $\mathcal{A}^{\mathcal{U}} \models \nabla v\ \varphi\ [s]$ iff the set $\{a \in A : \mathcal{A}^{\mathcal{U}} \models \varphi\ [s(v{:=}a)]\}$ belongs to the ultrafilter $\mathcal{U}$;

where, as usual, $s(v{:=}a)$ is the assignment agreeing with s on every variable but v, and $s(v{:=}a)(v)=a$.

Other familiar semantic notions, such as reduct, model ($\mathcal{A}^{\mathcal{U}} \vDash \Gamma$), *ultrafilter consequence* ($\Gamma \vDash^{\nabla} \tau$ iff $\mathcal{A}^{\mathcal{U}} \vDash \Gamma$ whenever $\mathcal{A}^{\mathcal{U}} \vDash \Gamma$), validity, etc. are as usual.

As usual, satisfaction of a formula depends only on the meanings of its symbols. In particular, for a formula φ of $\mathsf{L}^{\nabla}(\rho)$ without ∇ (so in $\mathsf{L}(\rho)$), $\mathcal{A}^{\mathcal{U}} \vDash \varphi\ [s]$ iff $\mathcal{A} \vDash \varphi\ [s]$. Also, satisfaction of a formula hinges only on the values assigned to its free variables. So, for a formula φ with no free occurrence of variables other than $v_1,\ldots,v_m$, we can employ the usual notation $\mathcal{A}^{\mathcal{U}} \vDash \varphi\ [a_1,\ldots,a_m]$, for $<a_1,\ldots,a_m>\in A^m$. Such a formula *defines* an m-ary relation:

$$\mathcal{A}^{\mathcal{U}}[\varphi] := \{<a_1,\ldots,a_m>\in A^m : \mathcal{A}^{\mathcal{U}} \vDash \varphi\ [a_1,\ldots,a_m]\}.$$

A convenient manner of presenting some ideas related to satisfaction is by means of extensions of formulas: projections of the defined relations. The *extension* of a formula φ of $\mathsf{L}^{\nabla}(\rho)$ in structure $\mathcal{A}^{\mathcal{U}} = (\mathcal{A}, \mathcal{U})$ for signature ρ under an assignment s:V→A with respect to a variable v ∈V is the set:

$$\mathcal{A}^{\mathcal{U}}[\varphi\,|\,v](s) := \{a\in A : \mathcal{A}^{\mathcal{U}} \vDash \varphi\ [s(v:=a)]\}.$$

These concepts for a formula φ with free variables x and y under an assignment s:V→A with s(x)=a are illustrated in the figure 3.1.

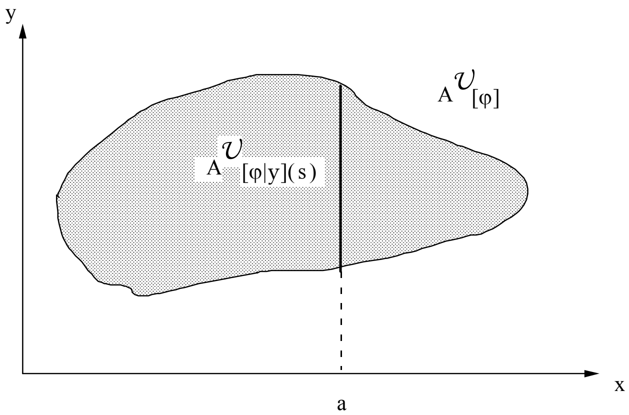

Figure 3.1: Formula φ (x, y): relation defined and extension

Since our definition of satisfaction extends the classical one, we have, e. g.

$$\mathcal{A}^{\mathcal{U}} \models \exists\, v\, \varphi\, [s] \text{ iff, for some } a \in A,\ \mathcal{A}^{\mathcal{U}} \models \varphi\, [s(v{:=}a)];$$

which can be rephrased in terms of extensions as

$$\mathcal{A}^{\mathcal{U}} \models \exists v \varphi\, [s] \text{ iff } \mathcal{A}^{\mathcal{U}}[\varphi \,|\, v](s) \neq \emptyset.$$

By the same token, the satisfaction condition for a most formula $\nabla v\ \varphi$ can be conveniently presented as

$$\mathcal{A}^{\mathcal{U}} \models \nabla v \varphi\, [s] \text{ iff } \mathcal{A}^{\mathcal{U}}[\varphi \,|\, v](s) \in \mathcal{U}.$$

To illustrate, consider the previous signature λ consisting of the binary predicate L. An example of ultrafilter structure for signature λ is the expansion $◐^{\mathcal{U}} = (◐, \mathcal{U})$ of the first-order structure $◐ = <\mathbf{N}, \leq>$ by an ultrafilter $\mathcal{U}$ over the naturals. As formula φ take

L(x, y), which defines the relation ≤. We then have the following extensions

$$\mathbf{◐}^{\mathcal{U}}[\varphi\,|\,x](s) = \{a \in \mathbf{N} : a \leq s(y)\}$$

and

$$\mathbf{◐}^{\mathcal{U}}[\varphi\,|\,y](s) = \{b \in \mathbf{N} : s(x) \leq b\}.$$

In particular, for an assignment s:V→**N** with $s(x)=0=s(y)$, we have

$$\mathbf{◐}^{\mathcal{U}}[\varphi\,|\,x](s) = \{0\} \text{ and } \mathbf{◐}^{\mathcal{U}}[\varphi\,|\,y](s) = \mathbf{N};$$

thus, if $\{0\} \notin \mathcal{U}$, since $\mathbf{N} \in \mathcal{U}$, we have

$$\mathbf{◐}^{\mathcal{U}} \not\models \nabla x\varphi\ [s] \text{ and } \mathbf{◐}^{\mathcal{U}} \models \nabla y\varphi\ [s].$$

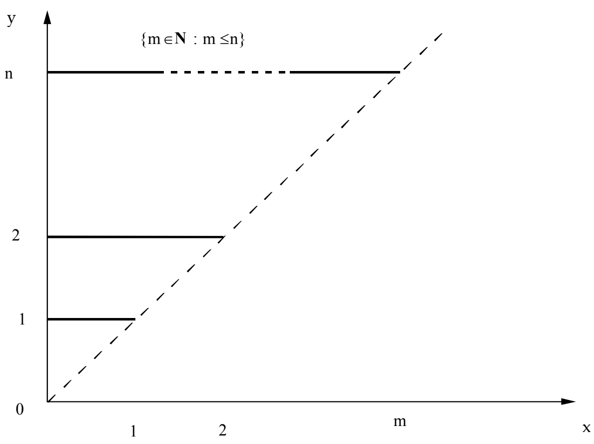

Figure 3.2: Extensions $\mathbf{◐}^{\mathcal{U}}[L(x, y)\,|\,x](s) = \{m \in \mathbf{N} : m \leq s(y)\}$

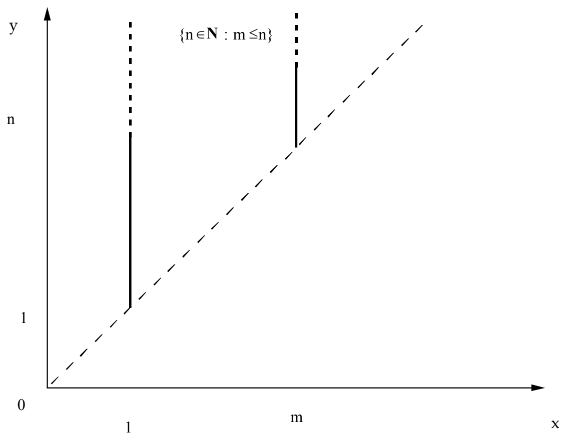

Figure 3.3: Extensions $◐^{\mathcal{U}}[L(x, y) | y](s) = \{n \in \mathbf{N} : s(x) \leq n\}$

In fact, if $\mathcal{U}$ is a Fréchet ultrafilter, having no finite subset, for any assignment $s: V \to \mathbf{N}$, we can see that (see figures 3.2 and 3.3)

$$◐^{\mathcal{U}}[\varphi | x](s) \notin \mathcal{U} \text{ (as } \{a \in \mathbf{N} : a \leq s(y)\} \text{ is finite),}$$

$$◐^{\mathcal{U}}[\varphi | y](s) \in \mathcal{U} \text{ (as } \{b \in \mathbf{N} : s(x) \leq b\} \text{ is cofinite);}$$

and thus, we also have

$$\{b \in \mathbf{N} : ◐^{\mathcal{U}} \vDash \nabla x\, \varphi\, [s(y{:=}b)]\} = \varnothing$$

and

$$\{a \in \mathbf{N} : ◐^{\mathcal{U}} \vDash \nabla y\, \varphi\, [s(x{:=}a)]\} = \mathbf{N},$$

which shows that

$$\mathbf{\circ}^{\mathcal{U}} \not\models \nabla y \nabla x \varphi\ [s] \text{ and } \mathbf{\circ}^{\mathcal{U}} \models \nabla x\, \nabla y \varphi\ [s].$$

Indeed, this is in agreement with the intuitive reading of these sentences:

– $\nabla y \nabla x\, L(x,y)$ asserts "for most $n \in \mathbf{N}$, the interval $\{m \in \mathbf{N} : m \leq n\}$ is large", which is false (since the interval $\{m \in \mathbf{N} : m \leq n\}$ is large for no $n \in \mathbf{N}$);

– $\nabla x \nabla y\, L(x,y)$ asserts "for most $m \in \mathbf{N}$, the interval $\{n \in \mathbf{N} : m \leq n\}$ is large", which is true (since the interval $\{n \in \mathbf{N} : m \leq n\}$ is large for every $n \in \mathbf{N}$).

Thus, the atomic $L(x,y)$ is an example of a formula φ such that

$$\not\models^{\nabla} \nabla y \nabla x \varphi \rightarrow \nabla x\, \nabla\, y\, \varphi.$$

This behavior of the new quantifier ∇ contrasts sharply with the commutativities of the classical quantifiers $\forall$ and $\exists$, since

$$\models^{\nabla} \forall y \forall x \varphi \rightarrow \forall x \forall y\, \varphi \text{ and } \models^{\nabla} \exists y \exists x \varphi \rightarrow \exists x \exists y \varphi.$$

More positive examples of the behavior of new the quantifier ∇ are the following (interderivable) transfers over the classical quantifiers $\forall$ and $\exists$:

$$\models^{\nabla} \exists y\, \nabla x\, \varphi \rightarrow \nabla x\, \exists y\, \varphi \text{ and } \models^{\nabla} \nabla x \forall y\, \varphi \rightarrow \forall\, y\, \nabla x\, \varphi.$$

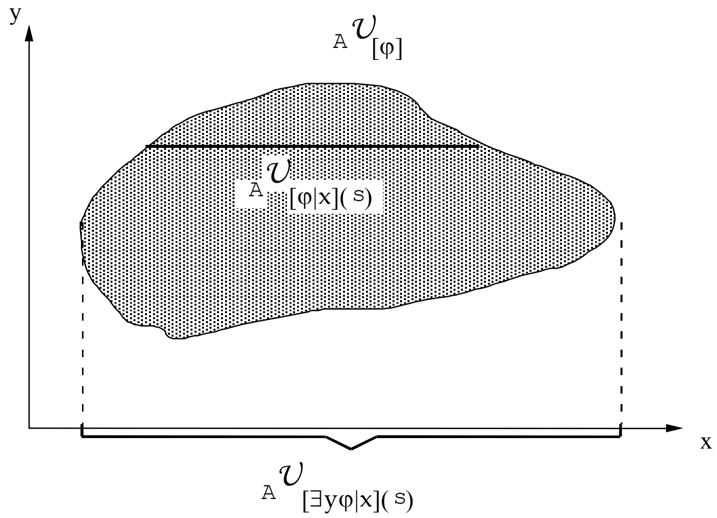

Figure 3.4: Illustration of transfer of ∃ over ∇

As an illustration of the former, take formula φ as L(x, y) and consider an ultrafilter structure $\mathcal{A}^{\mathcal{U}} = (\mathcal{A}, \mathcal{U})$ for signature λ, where $\mathcal{A}$ = <A, R> with $R \subseteq A^2$.

Each assignment s :V→A with s(y)=b gives a subset of A:

$$R_b := \mathcal{A}^{\mathcal{U}}[L(x, y) \mid x](s) = \{a \in A : \langle a,b \rangle \in R\}.$$

The transfer of ∃ over ∇ expresses the following property of the family $\{R_b \subseteq A : b \in A\}$ of subsets of A:

if $R_b \in \mathcal{U}$, for some b∈A, then so is the union $\cup_{b \in A} R_b$ in $\mathcal{U}$.

3.3 Axiomatics of ∇

We will now set up a deductive system for our logic by adding schemata coding properties of ultrafilters to a calculus for classical first-order logic.

We will construct a deductive system for our logic as follows.

– Consider a sound and complete deductive calculus for classical first-order logic, with Modus Ponens as the sole inference rule (and a set $A(\rho)$ of axiom schemata, as in, e. g. (Enderton (1972))).

– Extend it by adding the set $A^{\nabla}(\rho)$ consisting of all the generalizations of a set $B^{\nabla}(\rho)$ of schemata.

Thus, for a set Σ of sentences and a formula φ in $L^{\nabla}(\rho)$, we will have

$$\Sigma \vdash^{\nabla} \varphi \text{ iff } \Sigma \cup A^{\nabla}(\rho) \vdash \varphi.$$

As a clear consequence, we will have monotonicity of derivability $\vdash^{\nabla}$:

$$\text{If } \Gamma \vdash^{\nabla} \varphi \text{ then } \Gamma \cup \Delta \vdash^{\nabla}.$$

We will take as set of schemata the union $B^{\nabla}(\rho) := (\nabla\exists) \cup (\nabla\neg) \cup (\nabla\wedge) \cup (\nabla\alpha)$ of four sets of schemata. We now indicate these four kinds of schemata.

We first take some formulas corresponding to properties of ultrafilters. Consider the following sets of formulas of $L^{\nabla}(\rho)$

$$(\nabla\exists) := \{\nabla v\, \varphi \to \exists v \varphi : \varphi \in L^{\nabla}(\rho)\};$$
$$(\nabla\neg) := \{\neg\nabla v\, \varphi \to \nabla v \neg \varphi : \varphi \in L^{\nabla}(\rho)\};$$
$$(\nabla\wedge) := \{(\nabla v \psi \wedge \nabla v\, \theta) \to \nabla v\, (\psi \wedge \theta) : \psi, \theta \in L^{\nabla}(\rho)\}.$$

Then, the following formulas are seen to be provable

$(\forall\nabla)$ $\forall v\varphi\rightarrow\nabla v\,\varphi$ {from $(\nabla\neg)\cup(\nabla\exists)$};
$(\rightarrow\nabla)$ $\forall v\,(\psi\rightarrow\theta)\rightarrow(\forall v\psi\rightarrow\nabla v\,\theta)$ {from $(\nabla\neg)\cup(\nabla\wedge)\cup(\nabla\exists)$};
$(\neg\nabla)$ $\nabla v\neg\varphi\leftrightarrow\neg\nabla v\varphi$ {from $(\nabla\neg)\cup(\nabla\wedge)\cup(\nabla\exists)$};
$(\wedge\nabla)$ $\nabla v\,(\psi\wedge\theta)\leftrightarrow(\nabla v\,\psi\wedge\nabla v\,\theta)$ {from $(\nabla\wedge)\cup(\forall\nabla)$};
$(\leftrightarrow\nabla)$ $\forall v\,(\psi\leftrightarrow\theta)\rightarrow(\nabla v\psi\leftrightarrow\nabla v\,\theta)$ {from $(\rightarrow\nabla)$}.

As a result, we have substitutivity of equivalents:

$$\text{if } \Sigma\vdash^{\nabla}\psi\leftrightarrow\theta \text{ then } \Sigma\vdash^{\nabla}\nabla\, v\,\psi\leftrightarrow\nabla v\theta.$$

We thus see that, within equivalence, the new quantifier ∇ provably

- commutes with negation $\neg$,

and

- distributes over the binary propositional connectives $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$.

The preceding three schemata code some properties of ultra-filters. For instance, in terms of extensions, schema $(\nabla\wedge)$ expresses

$$\mathcal{A}^{\mathcal{U}}[\psi\wedge\theta\,|\,v](s)\in\mathcal{U}, \text{ whenever } \mathcal{A}^{\mathcal{U}}[\psi\,|\,v](s)\in\mathcal{U}$$

and

$$\mathcal{A}^{\mathcal{F}}[\theta\,|\,v](s)\in\mathcal{U};$$

which (as $\mathcal{A}^{\mathcal{U}}[\psi\wedge\theta\,|\,v](s)=\mathcal{A}^{\mathcal{U}}[\psi\,|\,v](s)\cap\mathcal{A}^{\mathcal{F}}[\theta\,|\,v](s)$) corresponds to the intersection of v-extensions in $\mathcal{U}$ is still in $\mathcal{U}$.

But, it does not cover the intersections of v-extensions with u-extensions.

So, we still need some axioms related to extensionality (Veloso (1999)). Consider also the following set of formulas of $\mathsf{L}^{\nabla}(\nabla)$

$$(\nabla\alpha) := \{\nabla v\varphi \rightarrow \nabla u\, \varphi\, (v/u) : \varphi \in \mathsf{L}^{\nabla}(\rho), u \notin occ(\varphi)\}.$$

We now have our set $B^{\nabla}(\rho) := (\nabla\Xi) \cup (\nabla\neg) \cup (\nabla\wedge) \cup (\nabla\alpha)$ of schemata, and take all the generalizations of the formulas in $B^{\nabla}(\rho)$ to form the set $A^{\nabla}(\rho)$ of *ultrafilter axiom schemata*.

Among the provable formulas we have

$(\nabla^{\alpha}\wedge)(\nabla x\psi \wedge \nabla y\, \theta) \rightarrow \nabla z\, [\psi\, (x/z) \wedge \theta\, (y/z)]$, for $z \notin occ(\Psi \wedge \theta)$
{from $(\nabla\alpha) \cup (\nabla\wedge)$};
$(\alpha\nabla)\nabla v\, \varphi \leftrightarrow \nabla u\, \varphi\, (v/u)$, for $u \notin occ(\varphi)$ {from $(\nabla\alpha) \cup (\forall\nabla)$}.

We thus obtain prenex forms, i.e. every formula φ of $\mathsf{L}^{\nabla}(\rho)$ is provably equivalent to one consisting of a prefix of quantifiers ($\forall, \exists$ and ∇) followed by a quantifier-free matrix μ: $\vdash^{\nabla} \varphi \leftrightarrow Q_1 v_1 \ldots Q_k v_k$.

Other usual deductive notions, such as (maximal) consistent sets, conservative extension, witnesses, etc. can be appropriately adapted.

4 ULTRAFILTER LOGIC

We shall now establish some properties of ultrafilter logic, namely soundness and completeness of the deductive system with respect to ultrafilter structures.

4.1 Soundness

We first examine the soundness of our deductive system with respect to ultrafilter structure. As usual, soundness is easily established.

Indeed, the axioms in $A^{\nabla}(\rho)$ code properties of ultrafilters, so they hold in all ultrafilter structures.

Clearly, the axioms in $(\nabla\exists)\cup(\nabla\wedge)\cup(\nabla\neg)$ code properties of ultrafilters, so they hold in every ultrafilter structure. As for $(\nabla\alpha)$, if variable u does not occur in φ, we have equal extensions $\mathcal{A}^{\mathcal{U}}[\varphi|v](s) = \mathcal{A}^{\mathcal{U}}[\varphi(v/u)|u](s)$.

We thus have soundness of our deductive system with respect to ultrafilter consequence, since Modus Ponens preserves satisfaction.

4.2 Deductive Properties

For the proof of completeness, we will need some properties of our deductive system, which can be established as in classical first-order logic.

• Consider a set $\Gamma\cup\{\sigma\}$ of sentences and a formula φ in $L^{\nabla}(\rho)$.

(1) (Deduction Theorem) If $\Gamma\cup\{\sigma\} \vdash^{\nabla}$ then $\Gamma\vdash^{\nabla} \sigma\to\varphi$.

(2) (Witness) If $fr(\exists v\varphi)=\varnothing$ for constant c not occurring in $\Gamma\cup\{\varphi\}$: $\Gamma\cup\{\exists v\ \varphi\to\varphi(v/c)\}$ is a conservative extension of Γ, i. e. $\Gamma\vdash^{\nabla}\tau$ iff $\Gamma\cup\{\exists v\ \varphi\to\varphi(v/c)\}\ \Gamma\vdash^{\nabla}\tau$.

(3) Set Γ is consistent iff every finite subset of Γ is consistent.

(4) Set $\Gamma\cup\{\sigma\}$ is consistent iff $\Gamma\nvdash^{\nabla}\neg\sigma$.

(5) For a maximal consistent Γ: $\Gamma\ \Gamma\nvdash^{\nabla}\sigma$ iff $\Gamma\vdash^{\nabla}\to\sigma$.

We thus have the extension of consistent sets to maximal consistent sets with witnesses.

• Consider a consistent set Γ of sentences in $\mathsf{L}^{\nabla}(\rho)$.

(1) (Henkin) There exists a consistent extension Δ of Γ by (at most $|\mathsf{L}^{\nabla}(\rho)|$) new constants, where every existential sentence has a witness: if $fr(\exists v\ \varphi)=\varnothing$, then $\Delta \vdash^{\nabla} \exists v\ \varphi \rightarrow \varphi(v/c)$, for some constant c.

(2) (Lindenbaum) There exists a maximal consistent extension Σ of Δ (over the same language).

4.3 Completeness

We now examine the completeness of our deductive system with respect to ultrafilter structure. Completeness is usually harder, but we can adapt Henkin's well-known proof for classical first-order logic (Henkin (1949)), by providing an adequate ultrafilter by means of witnesses.

We proceed to outline how this can be done.

Given a consistent set Γ in $\mathsf{L}^{\nabla}(\rho)$, extend it to a maximal consistent set Σ in $\mathsf{L}^{\nabla}(\rho \cup C)$, with witnesses in set C set of new constants for the existential sentences of $\mathsf{L}^{\nabla}(\rho \cup C)$. Considering the set T of variable-free terms of $\mathsf{L}(\rho \cup C)$, form the canonical structure $\mathcal{U}$, for signature $\rho \cup C$ as usual. It has universe $H := T/\approx$ where $t' \approx t''$ iff $\Sigma \vdash^{\nabla} t' \equiv t''$.

Henkin's inductive proof establishes for a sentence τ of language $\mathsf{L}(\rho \cup C)$

$$\mathcal{U} \vDash \tau \text{ iff } \Sigma \vdash \tau.$$

In our case, we need an extra inductive step to deal with the new quantifier ∇. This can be handled as follows.

Use the provable most sentences to form the family of *large* subsets of H

$$\Sigma^{\nabla} := \{\Sigma\,(v\,|\,\varphi)\subseteq H : \Sigma\vdash^{\nabla} \nabla\, v\;\; \varphi,\ fr(\nabla v\varphi)=\varnothing\};$$

where $\Sigma(v\,|\,\varphi)$ is the set *represented* within Σ by formula φ of $\mathsf{L}^{\nabla}(\rho\cup C)$ with respect to a variable $v\in V$, in the sense

$$\Sigma\,(v\,|\,\varphi) := \{t/\!\approx\,\in H : \Sigma\vdash^{\nabla} \varphi\,(v/t)\}.$$

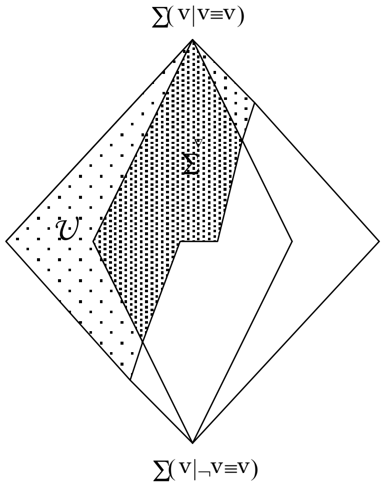

Figure 4.1: Ultrafilter extension $\mathcal{U}\supseteq\Sigma^{\nabla}$

In view of our axioms, the family $\Sigma^{\nabla}\subseteq\wp(H)$ is closed under intersection (by $(\nabla\alpha)$ and $(\nabla\wedge)$) and $\varnothing\notin\Sigma^{\nabla}$ (by $(\nabla\exists)$). Thus, Σ^{∇} has fip and it can be extended to some (proper) ultrafilter $\mathcal{U}\subseteq\wp(H)$. We use this ultrafilter to expand the canonical structure $\boldsymbol{\mathcal{U}}$ to an ultrafilter structure $\boldsymbol{\mathcal{U}}^{\mathcal{U}} := (\boldsymbol{\mathcal{U}}, \mathcal{U})$.

We can now show, by induction, that for a sentence τ of $\mathsf{L}^{\nabla}(\rho\cup C)$

$$\boldsymbol{\mathcal{U}}^{\mathcal{U}} \vDash \tau \text{ iff } \Sigma\vdash^{\nabla}\tau.$$

The inductive steps for the propositional connectives as well as for the classical quantifiers $\forall$ and $\exists$ are as in Henkin's proof.

Now, the inductive step for the new quantifier ∇, namely

$$\mathcal{U}^{\mathcal{U}} \vDash \nabla \text{ v } \varphi \text{ iff } \Sigma \vdash^{\nabla} \nabla \text{v } \varphi,$$

for a *sentence* ∇v φ, follows from the crucial property (due to ($\nabla\neg$))

$$\Sigma\,(\text{v}\,|\,\varphi) \in \Sigma^{\nabla} \text{ iff } \Sigma(\text{v}\,|\,\varphi) \in \mathcal{U}.$$

We thus have a Löwenheim-Skolem Theorem for our logical system.

Löwenheim-Skolem Theorem for ultrafilter logic

Each consistent set Γ of sentences of $\mathsf{L}^{\nabla}(\rho)$ has an ultrafilter model $\mathcal{M}^{\mathcal{U}}$ with cardinality at most that of its language, i.e. $|\text{M}| \leq |\,\mathsf{L}^{\nabla}(\rho)|$.

Hence, we have the desired result for ultrafilter logic.

Completeness of $\vdash^{\nabla}$ *with respect to* $\vDash^{\nabla}$

The deductive system $\vdash^{\nabla}$ is complete with respect to ultrafilter consequence: $\Gamma \vdash^{\nabla} \tau$ iff $\Gamma \vDash^{\nabla} \tau$.

4.4 Metamathematical Properties

We now examine some metamathematical properties of ultrafilter logic.

We have a sound and complete deductive system for ultrafilter logic:

- ultrafilter derivability is sound and complete with respect to ultrafilter consequence: $\Gamma \vdash^{\nabla} \tau$ iff $\Gamma \vDash^{\nabla} \tau$.

As usual, such a result transfers the finitary character of derivability $\vdash^{\nabla}$ to the compactness of the semantical consequence $\vDash^{\nabla}$.

Thus, our logic is a proper extension of classical first-order logic with compactness and Löwenheim-Skolem properties. The apparent conflict with Lindström's theorems (Lindström (1966)) (see, e. g. (Flum (1985))) is explained because we are using a non-standard notion of model (due to the ultrafilters in the models). This feature may confer to ultrafilter logic some independent model-theoretic interest.

Since a first-order structure can be expanded by some ultrafilter over its universe, we have the pleasing fact that our ultrafilter logic is a conservative extension of classical first-order logic.

Conservativeness of ultrafilter logic over classical logic

Ultrafilter logic is a conservative extension of classical first-order logic, i. e., for each set $\Gamma \cup \{\tau\}$ of sentences of $L(\rho)$: $\Gamma \vdash \tau$ iff $\Gamma \vdash^{\nabla} \tau$.

Also, as any nonempty set can be extended to some ultrafilter, the 'almost all' consequences of a pure first-order theory are the universal ones.

Universality of 'almost all' consequences of first-order theory

Given a set Γ of sentences of $L(\rho)$, for every formula φ of $L(\rho)$: $\Gamma \vdash^{\nabla} \nabla v\ \varphi$ iff $\Gamma \vdash^{\nabla} \forall v\ \varphi$.

This result corroborates the feeling that 'almost all' requires positive information, otherwise it reduces to 'all'. This observation will become clearer in the context of generic reasoning, to which we now turn.

5 TYPICAL AND GENERIC

We now wish to argue that our logic also supports a form of typical and generic reasoning, as in the familiar Tweety example, mentioned in the introduction. Namely, from the facts

(1) "Birds 'generally' fly"; and

(2) "Tweety is a 'typical' bird";

we wish to be able to conclude

(3) "Tweety does fly".

5.1 Basic ideas

Our approach to typical and generic reasoning involves two steps:

– formulating 'generally' as ∇; and

– regarding 'typical' and 'generic' as versions of 'prototypical'.

The former – formulating "birds 'generally' fly" as ∇ x F(x) – looks quite natural, in view of our interpretation of ∇ as 'holding almost universally'.

The latter – regarding 'typical' and 'generic' as versions of 'prototypical' – may require some explanation. How should one picture a 'prototypical' bird: with or without wings, with or without beak?

We propose to interpret a 'typical' bird as "a bird that exhibits the properties that *almost all* birds posses", and 'generic' bird as "a bird that has exactly the properties possessed by *almost all* birds".

Notice that "the properties that *all* birds exhibit" would be too strong.

It remains to give a rigorous formulation for these ideas of 'prototypical' objects in terms of "the properties that almost all objects possess". We proceed to explain how this can be done. Our approach can be understood as a symbolic form of typical and generic reasoning, in that the quantifier ∇ can be used to capture precisely the meaning of 'typical' and 'generic'.

5.2 Typical and generic objects

We first examine typical and generic elements in an ultrafilter structure.

Consider an ultrafilter structure $\mathcal{A}^{\mathcal{U}} = (\mathcal{A}, \mathcal{U})$ for a given signature ρ.

First, a formula φ of $L^{\nabla}(\rho)$ with single free variable v defines a property: the set $\mathcal{A}^{\mathcal{U}}[\varphi]$ of elements that satisfy it. Thus, we can express

– an element $a \in A$ has the property (defined by) φ by $a \in \mathcal{A}^{\mathcal{U}}[\varphi]$, i. e. $\mathcal{A}^{\mathcal{U}} \vDash \varphi\,[a]$;

– almost all elements of A have the property (defined by) φ by

$$\mathcal{A}^{\mathcal{U}} \vDash \nabla v\ \varphi, \text{ i. e. } \mathcal{A}^{\mathcal{U}}[\varphi] \in \mathcal{U}.$$

So, given a most sentence $\nabla v\ \varphi$ of $L^{\nabla}(\rho)$, we will call element $a \in A$:

• *typical for sentence* $\nabla v\ \varphi$ iff

$$\mathcal{A}^{\mathcal{U}} \vDash \nabla v\ \varphi \text{ whenever } \mathcal{A}^{\mathcal{U}} \vDash \varphi\,[a];$$

• *generic for sentence* $\nabla v\, \varphi$ iff

$$\mathcal{A}^{\mathcal{U}} \vDash \nabla v\, \varphi \text{ iff } \mathcal{A}^{\mathcal{U}} \vDash \varphi\,[a].$$

We can regard a typical element $a \in A$ as providing a local test for the most sentence $\nabla v\, \varphi$, which is decisive if a is generic.

Also, by the equivalence between $\nabla v \neg \varphi$ and $\neg\nabla v\ \varphi$, we see that

a is generic for $\nabla v\ \varphi$ iff a is typical for $\nabla v\ \varphi$ and $\nabla v \neg \varphi$.

Now, consider a set Φ of formulas of $\mathsf{L}^{\nabla}(\rho)$. We shall call an element $a \in A$ *typical* (or *generic*) *for set* Φ of formulas iff a is typical (or generic) for every most sentence $\nabla v\ \varphi$ in Φ. In particular, by a *typical* (or *generic*) *element* we will mean an element $a \in A$ that is typical (or generic) for every most sentence of $\mathsf{L}^{\nabla}(\rho)$.

Generic elements are indiscernible among themselves, in that they cannot be separated by formulas. Given generic elements g' and g'', for evey formula φ, with single free variable $v \in V$, of $\mathsf{L}^{\nabla}(\rho)$:

$$\mathcal{A}^{\mathcal{U}} \models \varphi\ [g'] \text{ iff } \mathcal{A}^{\mathcal{U}} \models \varphi\ [g''].$$

Generic elements are somewhat reminiscent of Hilbert's ideal elements, or even of Platonic forms. So, it is not surprising that some ultrafilter structures fail to have generic elements.

For instance, in the naturals with zero and successor and a non-principal ultrafilter, containing the cofinite subsets, a typical element, if any, must be non-standard.

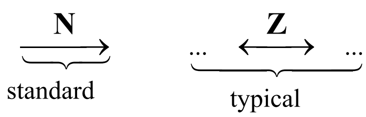

Figure 5.1: Typical elements and non-standard naturals

Indeed, consider the expansion $\mathfrak{S}^{\mathcal{U}} = (\mathfrak{S}, \mathcal{U})$ of the first-order structure $\mathfrak{S} = <\mathbf{N}, S, 0>$ by a Fréchet ultrafilter $\mathcal{U}$ over the naturals.

As formula φ take v ≡ SS0, which defines the singleton {2}. Much as before

$$\mathfrak{S}^{\mathcal{U}} \vDash \neg\, \varphi\, [a] \text{ iff } a \neq 2;$$

and

$$\mathfrak{S}^{\mathcal{U}} [\neg\, \varphi] \in \mathcal{U} \text{ (as } \mathbf{N} - \{2\} \text{ is cofinite).}$$

We thus have

$$\mathfrak{S}^{\mathcal{U}} \vDash \nabla v \neg\, \varphi \text{ but } \mathfrak{S}^{\mathcal{U}} \nvDash \neg\, \varphi\, [2],$$

which shows that the natural 2 is not typical for formula ¬v ≡ SS0.

Similarly, since each natural n∈**N** is defined by an equation v ≡ t, for some ground term t, we can see that no standard natural n ∈ **N** can be typical.

Along similar lines, we notice that, in the case of a definable nonempty set of generic elements the new operator ∇ reduces to a relativized classical quantification.

Interestingly enough, such generic elements have much better behavior in theories, where they can be regarded as generic witnesses.

First, in view of fip, finite sets of sentences have generic elements.

Generic element for finite sets of sentences

A <u>finite</u> set Σ of (most) sentences of $L^{\nabla}(\rho)$ has a generic element in an ultrafilter structure $\mathcal{A}^{\mathcal{U}} = (\mathcal{A}, \mathcal{U})$.

For each sentence ∇v φ in Σ, we have either $\mathcal{A}^{\mathcal{U}}[\varphi] \in \mathcal{U}$ (if $\mathcal{A}^{\mathcal{U}} \vDash \nabla v \varphi$) or else $\mathcal{A}^{\mathcal{U}}[\neg\, \varphi] \in \mathcal{U}$ (if $\mathcal{A}^{\mathcal{U}} \nvDash \nabla v \varphi$). By fip, we have some $a \in A$ such that $a \in \mathcal{A}^{\mathcal{U}}[\varphi]$, when $\mathcal{A}^{\mathcal{U}} \vDash \nabla v\varphi$, and $a \in \mathcal{A}^{\mathcal{U}}[\neg\varphi]$, when $\mathcal{A}^{\mathcal{U}} \nvDash \nabla v\, \varphi$. So, $\mathcal{A}^{\mathcal{U}} \vDash \nabla v\, \varphi$ iff $\mathcal{A}^{\mathcal{U}} \vDash \varphi[\mathbf{a}]$.

5.3 Reasoning with generic constants

We now internalize the previous ideas in extensions by new constants.

We wish to add a new constant c behaving as a crucial witness for most, i. e. constant c has property φ iff φ holds almost universally.

Given a signature ρ and a new constant c *not* in ρ, consider the expansion $\rho[c] := \rho \cup \{c\}$ of type ρ obtained by adding the new constant c.

Given a sentence $\nabla v\, \varphi$ of $L^{\nabla}(\rho)$, we construct the *genericity axiom* $\varpi[\nabla v\, \varphi/c]$ *of* c *for* sentence $\nabla v\varphi$ as the following sentence (of $L^{\nabla}(\rho\, [c])$)

$$\nabla v\varphi \leftrightarrow \varphi\, (v/c).$$

Now, for a set Σ of sentences of $L\nabla(\rho)$, by the *genericity axiom schema on* c *for* a set Σ of sentences, we mean the following set of sentences of $L^{\nabla}(\rho\, [c])$

$$\varpi[\Sigma/c] := \{\varpi\, [\nabla v\, \varphi/c] : \nabla v\, \varphi \in \Sigma\}.$$

In particular, when Σ is the set of all the most sentences of $L^{\nabla}(\rho)$, we use $\varpi[c]$ for the *genericity axiom schema on* c.

These conditions extend conservatively theories in $L^{\nabla}(\rho)$ to $L^{\nabla}(\rho\, [c])$.

Conservative addition of new generic constant

Consider a set Γ of sentences of $L^{\nabla}(\rho)$. For each set Σ of (most) sentences of $L^{\nabla}(\rho)$, $\Gamma[\Sigma/c] := \Gamma \cup \varpi[\Sigma/c]$ is a conservative extension of Γ, such that, for every most sentence $\nabla v\, \varphi \in \Sigma$: $\Gamma \vdash^{\nabla} \nabla v\, \varphi$ iff $\Gamma[\Sigma/c] \vdash^{\nabla} \varphi\, (v/c)$.

This result establishes the correctness of reasoning with a generic constant. In particular, the extension $\Gamma[c]$ of Γ by generic constant c is conservative and $\Gamma[c] \vdash^{\nabla} \varphi(v/c)$ iff $\Gamma \vdash^{\nabla} \nabla v\ \varphi$ for every sentence $\nabla v\ \varphi$ of $\mathsf{L}^{\nabla}(\rho)$.

An example, similar to flying birds and Tweety, may illustrate these ideas.

From "Most swans are white", one concludes "A generic swan is white".

Signature γ has unary predicate W and theory Γ, over $\mathsf{L}^{\nabla}(\gamma)$, has $\nabla v\ W(v)$ {for "Almost all swans are white"}.

Considering a new constant s (for generic swan), generic extension $\Gamma[s]$ has (in $\mathsf{L}^{\nabla}(\gamma\ [s])$ the genericity axiom $\varpi[\nabla v\ W(v)/s]$

$\nabla v\ W(v) \leftrightarrow W(s)$ {"Most all swans are white iff the generic swan is white"}.

Hence, in $\Gamma[s]$, we have the following sentence of $\mathsf{L}^{\nabla}(\gamma\ [s])$

$$W(s)\ \{\text{i. e. "A generic swan is white"}\}.$$

Notice that, if b is a non-white swan, in $\Gamma[s] \cup \{\neg W(b)\}$, one has both $\neg W(b)$ and $W(s)$, so $\neg b \equiv s$ (this non-white swan b is not a generic swan). So, by conservativeness, $\exists x\ W(x) \wedge \exists y \neg W(y)$ and $\exists x\ \exists y \neg x \equiv y$ are consequences of $\Gamma \cup \{\exists y\ \neg W(y)\}$ ("Most swans are white, but there is a non-white swan").

This example illustrates the monotonic nature of our logic: we do not have to retract conclusions in view of new facts. Given that "Most swans are white", we conclude that "a generic swan is white", a conclusion which we may hold even if further evidence indicates some non-white swans.

A variation of this example may serve to illustrate why we do not have multiple extensions. A theory asserting both "generally birds fly", as $\nabla x\, F(x)$, and "generally birds do not fly", as $\nabla v\, \neg F(x)$, would be inconsistent.

The next example illustrates using several (relativized) generic constants.

Assuming that "Generally humans like dogs", one would conclude "A generic human likes a generic dog".

Consider a signature η having a binary predicate L (with L(x,y) standing for x likes y), as well as unary predicates H and D (standing, respectively, for 'is human' and 'is a dog').

Assume that theory Λ has as axiom the following sentence of $\mathsf{L}^{\nabla}(\eta)$

$$\nabla x\, \nabla y\, [H(x) \wedge D(y) \rightarrow L(x,y)] \ \{\text{generally humans like dogs}\}.$$

Considering new constant h (for generic human), the extension $\Lambda[h]$ has among its the genericity axioms the following sentence of $\mathsf{L}^{\nabla}(\eta\, [h])$:

$$\nabla x\ \nabla y\, [H(x) \wedge D(y) \rightarrow L(x,y)] \leftrightarrow \nabla\, y\, [H(h) \wedge D(y) \rightarrow L(h,y)].$$

Thus, as a consequence of $\Lambda[h]$, we have

$$\nabla y\, [H(h) \wedge D(y) \rightarrow L(h,y)] \ \{\text{a generic human likes most dogs}\}.$$

With another new constant d (for generic dog), we form extension $\Lambda[h][d]$ over $\eta[h][d] := \eta[h] \cup \{d\}$, which has as genericity axiom for sentence $\nabla y\, [H(h) \wedge D(y) \rightarrow L(h,y)]$ of $\mathsf{L}^{\nabla}(\eta\, [h]$ the following sentence of $\mathsf{L}^{\nabla}(\eta\, [h][d])$:

$$\nabla y\, [H(h) \wedge D(y) \rightarrow L(h,y)] \leftrightarrow [H(h) \wedge D(d) \rightarrow L(h,d)].$$

Hence, among the consequences of $\Lambda[h][d]$, we have

$H(h) \wedge D(d) \rightarrow L(h, d)$ {a generic human likes a generic dog}.

Notice that $\Lambda[h][d]$ does not commit us to the existence of (generic) humans or dogs; all that it asserts is that "generic humans, if any, like generic dogs, if any". Assuming the existence of generic humans and dogs, then $\Lambda[h][d] \cup \{H(h), D(d)\} \vdash^{\nabla} L(h,d)$.

The kind of reasoning in this example involves several generic constants, which can be introduced by iterating our construction.

Given a signature ρ and a (denumerable) list $\underline{c}$ of new constants not in ρ, we form the iterated expansions $\rho_{n+1} := \rho_n[c_n]$ of $\rho_0 := \rho$ by new constants $c_0, \ldots, c_n$, and set $\rho[\underline{c}] := \cup_{n \geq 0} \rho_n$. Much as before, given sets Φ_n of sentences of $L^{\nabla}(\rho_n)$, we have a set $\varpi[\Phi_n/c_n]$ of formulas in $L^{\nabla}(\rho_{n+1})$ expressing the genericity condition on c_n for Φ_n.

Also, each set Γ of sentences in $L^{\nabla}(\rho)$ has as conservative extensions:

- $\Gamma \cup \varpi[\Phi_0 | v_0/c_0] \cup \ldots \cup \varpi [\Phi_n/c_n]$ in $L^{\nabla}(\rho_{n+1})$, for each $m \geq 0$;

- the union $\Phi \cup \cup_{n \geq 0} \varpi [\Phi_n/c_n]$ in $L^{\nabla}(\rho [\underline{c}])$.

6. RELATIVE MOST

We shall now examine the idea of having a notion of most relative to a universe: how it arises and is formulated, as well as some related issues.

We will first indicate how the proper expression of 'relative most' assertions brings about the idea of a notion of large with respect to each universe, leading to its natural formulation in a sorted version of ultrafilter logic. Then, the need for establishing some connections while blocking others leads to comparing such relative concepts.

Finally, these ideas are incorporated into a sorted framework for 'almost all' and 'generic' reasoning.

6.1 The need for 'relative most'

Our generalized quantifier ∇ may exhibit somewhat unexpected behavior in some cases. We shall now examine these undesirable side-effects and propose a way to overcome this difficulty.

The generalized quantifier ∇ is meant to capture the idea of holding almost universally, i. e. for 'almost all' objects of the universe. Sometimes we wish to express the idea of holding almost universally over a given subset of the universe, i. e. for 'almost all' objects of a given sub-universe.

We now examine the expression of such 'relative most' assertions.

On a universe B of birds, we express "Generally birds have wings" by $\nabla v\ W(v)$ {in the sense "Almost all birds have wings"}.

How are we to express 'relative most' assertions, such as "Winged birds generally fly", "Eagles generally fly", or "Penguins generally have beaks"?

An apparently natural suggestion, by analogy, is as follows.

One can express the relativization of a universal assertion such as

$\forall\ v\ W(v)$ {for "All birds have wings"},

to the sub-universe of eagles by $\forall v\ [E(v) \rightarrow W(v)]$ {for "All eagles have wings"}.

By analogy, one would expect to relativize a 'most' assertion such as $\forall v\ K(v)$ {for "Almost all birds have beaks"}, to the sub-universe of penguins by means of $\nabla v\ [P(v) \rightarrow K(v)]$ {for "Almost all birds have beaks"}.

Is this a reasonable expression or are we misreading the last formula?

For a 'most' formula the form $\nabla v\ [M(v) \rightarrow N(v)]$ the reading "generally M's are N's" is perhaps appropriate. But one must bear in mind that what this does assert is "for most *birds* b, if M(b) then N(b)".

For, given the meaning of the conditional, a formula

$$\nabla v\ [M(v) \rightarrow N(v)]$$

– has as refutatory evidence the set $\underline{M} \cap \underline{N}^c$ of exceptions;

and

– means that the set $\underline{M}^c \cup \underline{N}$ is a large set of *birds*.

This indicates that we may have been misled by the analogy.

Indeed, formula $\nabla v\ W(v)$ means $\underline{W} \approx \underline{B}$ and asserts "Almost all birds have wings", which can be read as "Generally birds have wings".

On the other hand, formula $\nabla v\ [P(v) \rightarrow F(v)]$ means $\underline{P}^c \cup \underline{F} \approx \underline{B}$ and asserts "Almost all birds are non-penguins or fly", which does not seem to convey the idea of "Penguins generally fly".

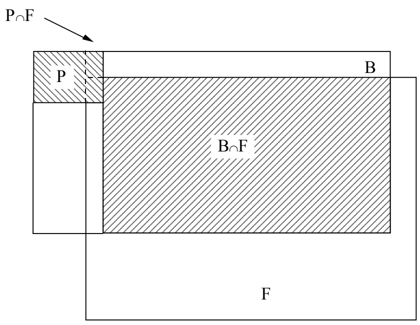

Figure 6.1: Flying birds and penguins

Another reason for seeing that $\nabla v\,[M(v) \rightarrow N(v)]$ is not a reasonable way of expressing the 'relative most' assertion "Almost all M's are N's" stems from some undesirable connections, as we shall illustrate later in 6.3.

This question appears to be connected to the so-called 'Confirmation Paradox' in Philosophy of Science (Hempel (1965)). Each flying eagle is considered as evidence in favor of "Eagles fly", whereas a non-flying non-eagle is not felt so, even though "Eagles are fliers" and "Non-fliers are non-eagles" are logically equivalent.

The classification of evidences for "Eagles fly" as confirmatory or refutatory reads as the truth table of the conditional. This analogy appears to incriminate material implication. It also suggests a possible way out: by means of a new (non truth-functional) connective for 'generally' (from which one would expect to define a generalized quantifier 'most').

We prefer to follow an alternative route, still within the ultrafilter approach, namely distinct notions of 'large' subsets.

The basic idea for expressing 'relative most' assertions is that each given universe has its own 'relative notion of large' subsets.

We now begin to take a closer look at the proposal of employing distinct notions of large subsets. We shall use variations of the preceding examples to introduce this idea and examine some of its features.

Let us first examine the question of adequate expression.

With 'relative notion of large', we can express "Generally M's are N's"

– more precisely as "Almost all M's are N's",
– so as to mean that the set $\{a \in \underline{M} : N(a)\}$ is 'almost as large as' $\underline{M}$: $\underline{M} \cap \underline{N} \approx \underline{M}$.

For instance, we express "Winged birds generally fly"

– as “Almost all winged birds fly”, meaning that
– the set $\underline{W}\cap\underline{F}$ of flying winged birds is a large subset of the universe $\underline{W}$ of winged birds ($\underline{W}\cap\underline{F}\approx\underline{W}$).

In this manner, we can also distinguish, say, “Eagles generally fly” from “Penguins generally fly” as follows.

– “Almost all eagles fly” means that the flying eagles form a large set of *eagles* ($\underline{E}\cap\underline{F}\approx\underline{E}$), whereas
– “Almost all penguins fly” means that the flying penguins form a large set of *penguins* ($\underline{P}\cap\underline{F}\approx\underline{P}$).

The idea of each universe having its own relative notion of large subsets may be formulated by giving an ultrafilter $\mathcal{U}_S$ over each given universe S.

In this manner, the two preceding ‘relative most’ assertions now become:
– $\underline{E}\cap\underline{F}\in\mathcal{U}_{\underline{E}}$ for $\underline{E}\cap\underline{F}\approx\underline{E}$ {“Almost all eagles fly”};
– $\underline{P}\cap\underline{F}\in\mathcal{U}_{\underline{P}}$ for $\underline{P}\cap\underline{F}\approx\underline{P}$ {“Almost all penguins fly”}.

A sorted approach seems to be adequate for the idea that each given universe has its own relative notion of large subsets.

6.2 Sorted ultrafilter logic

A many-sorted approach seems to provide an appropriate framework for formulating the idea of distinct notions of large relative to the universes, by assigning ultrafilters corresponding to these notions of large.

We shall now examine sorted versions of ultrafilter logic. The basic idea is that the previous (unsorted) concepts now become relativized to sorts.

We consider many-sorted signatures, where the extra-logical symbols, as well as variables, come classified according to sorts (Birkhoff and Lipson (1970); Enderton (1972)). Quantifiers are relativized to sorts, as expressed in the formation rules:

- for each variable v over sort s, if φ is a formula in $L^{\nabla}(\rho)$, then so are $(\forall\ v : s)\ \varphi$, $(\exists v : s)\ \varphi$ and $(\nabla v : s)\ \varphi$.

An *ultrafilter structure* $\mathcal{A}^{\mathcal{U}}$ for S-sorted signature ρ is an expansion an S-sorted first-order structure $\mathcal{A}$ for signature ρ, obtained by assigning for each sort s of signature ρ an ultrafilter $\mathcal{U}[s]$ over the universe $\mathcal{A}[s]$ of sort s (giving the large subsets of $\mathcal{A}[s]$).

The extension of *satisfaction* becomes relativized to sorts accordingly:

$(\vDash\nabla)_s$ for a most formula $(\nabla v : s)\ \varphi$, we define $\mathcal{A}^{\mathcal{U}} \vDash (\nabla v{:}s)\ \varphi\ [s]$ iff the set $\{a \in A[s] : \mathcal{A}^{\mathcal{U}} \vDash \varphi[s(v{:=}a)]\}$ is in the ultrafilter $\mathcal{U}[s]$.

The *ultrafilter axiom schemata* in the union $B^{\nabla}(\rho)$ become sorted as well:

$(\nabla\exists)_s : (\nabla v : s)\ \varphi \to (\exists v : s)\ \varphi$;
$(\nabla\neg)_s : \neg\ (\nabla v : s)\varphi \to (\nabla v : s)\ \neg\ \varphi$;
$(\nabla\wedge)_s : ((\nabla\ v : s)\ \psi \wedge (\nabla v : s)\ \theta\ ) \to (\nabla v : s)\ (\psi\wedge\theta)$;
$(\nabla\alpha)_s : (\nabla\ v : s)\ \varphi \to (\nabla u : s)\ \varphi\ (v/u)$, for each variable u : s not occurring in φ.

As an illustration, we examine a possible formulation for the assertion “People generally oppose defeaters of each team supported”.

Consider a signature δ with two sorts $\underline{s}$ (for spectators) and $\underline{t}$ (for teams) as well as binary predicates S, D and O (with S(x,y), D(z,y) and O(x,z) standing, respectively, for x supports y, z defeated y and x opposes z).

Theory Δ has the sorted axiom the following sentence of $\mathsf{L}^{\nabla}(\delta)$:

$$(\nabla x : \underline{s})\ (\forall y : \underline{t})\ (\nabla z : \underline{t})\ [S(x,y) \wedge D(z,y) \rightarrow O(x,z)].$$

By adding a constant b of sort $\underline{t}$ (for the Brazilian team) one forms the expanded signature $\delta' := \delta \cup \{b\}$, where we can express “Most people support the Brazilian team or a team that defeated the Brazilian team” as

$$(\nabla x : \underline{s})\ \ [S(x,b) \vee (\exists y : \underline{t})\ (D(y,b) \wedge S(x,y))].$$

$$\begin{array}{ccccccc} & & & & b & & \\ & & & & \downarrow & & \\ \underline{p} & - & S & - & \underline{t} & = & D \end{array}$$

Figure 6.2: Signature for persons and teams

Much as in classical first-order logic, the sorted and unsorted versions are very similar. So, soundness and completeness carry over to the sorted version, by relativizing to sorts the previous arguments. For completeness, the witnesses introduced for the existential quantifiers inherit the corresponding sort and we now have sorted families of large subsets

$_s\Sigma^\nabla := \{\Sigma\ (v : s\,|\,\varphi) \subseteq H : \Sigma \vDash^\nabla (\nabla v : s)\ \varphi, fr((\nabla v :s)\ \varphi)=\varnothing\}$.

Similarly, the ideas of typical and generic elements carry over to this case.

The next example illustrates the usage of several sorted generic constants.

Assuming that "People generally oppose defeaters of each team supported", one would conclude "A generic supporter of the Brazilian team opposes a generic team that defeated the Brazilian team".

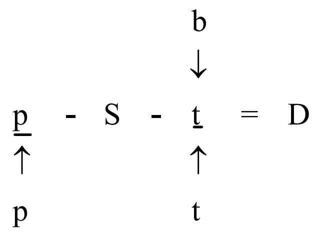

Figure 6.3: Expanded signature with generic persons and teams

Consider theory Δ and expanded signature δ' of the preceding example.

With a new generic constant p of sort $\underline{p}$, form extension $\Delta[p : \underline{p}]$ with genericity axioms such as

$$(\nabla x : \underline{s})\ (\forall y : \underline{t})\ (\nabla z : \underline{t})\ [S(x,y) \wedge D(z,y) \to O(x,z)] \leftrightarrow$$
$$(\forall y : \underline{t})\ (\nabla z : \underline{t})[S(p,y) \wedge D(z, y) \to O(p, z)].$$

So, as a consequence of $\Delta[p : \underline{p}]$, we have the formula of $L^\nabla(\delta'[p : \underline{p}])$:

$$(\nabla z : \underline{t})\ [S(p, b) \wedge D(z, b) \to O(p, z)]$$

for "generic supporters of Brazil oppose most teams that defeated Brazil".

With a further new constant t of sort $\underline{t}$, we form the extension $\Delta[p:\underline{p}][t:\underline{t}]$ over $\delta'[p:\underline{p}][t:\underline{t}] := \delta'[p:\underline{p}] \cup \{t:\underline{t}\}$, which has as genericity axiom for sentence $(\nabla z:\underline{t})\ [S(p, b) \wedge D(z, b) \rightarrow O(p, z)]$ of $\mathsf{L}^{\nabla}(\delta'[p:\underline{p}])$ the sentence of $\mathsf{L}^{\nabla}(\delta'[p:\underline{p}][t:\underline{t}])$:

$$(\nabla z:\underline{t})\ [S(p, b) \wedge D(z, b) \rightarrow O(p, z)] \leftrightarrow [S(p, b) \wedge D(t, b) \rightarrow O(p, t)].$$

Then, as consequence of $\Delta[p:\underline{p}][t:\underline{t}]$, we have the sentence of L^{∇} $(\delta'[p:\underline{p}][t:\underline{t}])$:

$$S(p, b) \wedge D(t, b) \rightarrow O(p, t)$$

{"generic supporters of Brazil oppose generic teams that defeated Brazil"}.

6.3 Comparing 'relative notions of large'

We shall now examine how the need for establishing some connections while blocking others leads to comparing relative notions of large sets.

As mentioned in 6.1, another reason for seeing that relativization does provide an adequate way of expressing 'relative most' assertion comes from the transitive behavior of our generalized quantifier ∇.

The next example illustrates some aspects of this transitive behavior of ∇.

Consider the following expressions for facts on birds:

(K) $\nabla v\ [K(v) \rightarrow F(v)]$ {for "Most birds with beaks fly"};
(W) $\forall v\ [W(v) \rightarrow K(v)]$ {"All winged birds have beaks"},
(P) $\forall v\ [P(v) \rightarrow K(v)]$ {"All penguins have beaks"}.

(A) From the two facts (K) and (W), we conclude ∇v [W(v) $\rightarrow$ F(v)].
(B) From the two facts (K) and (P), we conclude ∇v [P(v) $\rightarrow$ F(v)].

Notice that the two arguments (in A and B) share the same logical form

from $\forall$v [φ(v) $\rightarrow$ W(v)] and ∇ v [W(v) $\rightarrow$ F(v)] infer ∇v [φ(v) $\rightarrow$ F(v)],

being correct, because

$$\{\forall v\ [\varphi(v)\rightarrow\psi(v)],\nabla v\ [\psi(v)\rightarrow\theta(v)] \vdash^{\nabla}\nabla v\ [\varphi(v)\rightarrow\theta(v)].$$

Also, the three given facts, namely "Most birds with beaks fly", "All winged birds have beaks" and "All penguins have beaks", appear quite reasonable. On the other hand, the concluded assertion

– in A, read as "Most winged birds fly", looks acceptable, whereas

– in B, read as "Most penguins fly", is *not* expected.

One can consistently hold "Most birds with beaks fly", "All penguins have beaks" and "Most penguins do not fly" (or even "No penguin flies"). Apparently, one feels that the set of penguins, being a small set of birds (with beaks), does constitute a sizable set of exceptions to the belief that most birds (with beaks) fly (as indicated in figure 6.1).

Also, notice that the assumption $\neg\nabla$vP(v) (i. e. "Very few birds are penguins") would yield ∇v[P(v) $\rightarrow$ E(v)], which, if read as "Most penguins are eagles", would sound somewhat puzzling.

It is not difficult to see that much of this confusion actually comes from faulty expressions and misreading.

Now, let us consider the form with relative notion of large: "generally M's are N's" expressed as "Almost all M's are N's" and meaning the set $\{a \in M : N(a)\}$ is a large subset of $\underline{M}$ ($\underline{M}\cap\underline{N}\approx\underline{M}$).

We can see that the derivation of the undesired conclusion is now blocked.

Indeed, consider the following assertions:

(P') "All penguins have beaks": $\underline{P} \subseteq \underline{K}$.

(K') "Most birds with beaks fly": $\underline{K}\cap\underline{F} \subseteq \underline{K}$ 'almost as large as' $\underline{K}$ ($\underline{K}\cap\underline{F}\approx\underline{K}$).

(A') "Almost all penguins fly": $\underline{P}\cap\underline{F} \subseteq \underline{P}$ is 'almost as large as' $\underline{P}$ ($\underline{P}\cap\underline{F}\approx\underline{P}$).

In the presence of the first assertion,

– neither (K') entails (A') (since we may even have $\underline{P}\cap\underline{F}=\varnothing$, as we expect),
– nor does (A') entail (K') (since $\underline{P}\subset\underline{K}$ may very well be a small set of *birds with beaks*, as we believe); if the notions of large subsets are not connected.

This example illustrates the idea that we may have independent notions of large subsets. If the set of penguins is not a large set of *birds with beaks* ($\underline{P} \subseteq \underline{K}$ not 'almost as large as' $\underline{K}$), then a set $X\subseteq\underline{P}$ may be a large set of *penguins* without being a large set of *birds with beaks*. It is this independence that blocks the undesired conclusion.

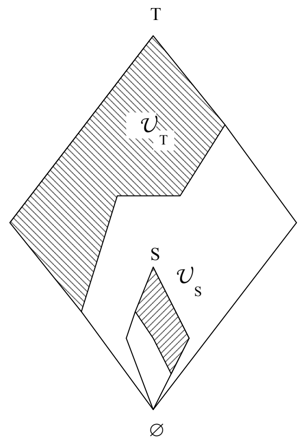

Figure 6.4: Ultrafilters over set T and 'small' subset $S \subseteq T$ ($S \notin \mathcal{U}_T$)

The independent notions of large subsets in terms of ultrafilters is explained as follows. For $S \subseteq T$, given ultrafilters $\mathcal{U}_S$, over S, and $\mathcal{U}_T$, over T,

if $S \notin \mathcal{U}_T$ then, for every $X \subseteq S$: $X \notin \mathcal{U}_T$ (even for those X in $\mathcal{U}_S$).

As an example, take T as the naturals and S as the even naturals. If the set of evens is not considered a large set of naturals ($S \notin \mathcal{U}_T$) then, no set consisting only of evens (including the large sets of evens) is a large set of naturals.

We have just seen how the derivation of the undesired conclusion "Most penguins fly" is blocked by distinct notions of large subsets. By the same token, the derivation of the conclusion "Most winged birds fly" is also blocked. Yet, this conclusion appears to be expected as reasonable.

Consider the following assertions:

(W') "All winged birds have beaks": $\underline{W} \subset \underline{K}$.

(K') "Most birds with beaks fly": $\underline{K}\cap\underline{F}\subset\underline{K}$ 'almost as large as' $\underline{K}$ ($\underline{K}\cap\underline{F}\approx\underline{K}$).

(B') "Most winged birds fly": $\underline{W}\cap\underline{F}\subset\underline{W}$ is 'almost as large as' $\underline{W}$ ($\underline{W}\cap\underline{F}\approx\underline{W}$).

Just as before, the first assertion (W') establishes no connection between (K') and (B'), if the notions of large subsets are not related.

We now have distinct notions large subsets. Is a large set of *winged birds* also a large set of *birds with beaks*? Is a set of winged birds that happens to be a large set of *birds with beaks* a large set of *winged birds* as well?

Yet, here we seem to feel that there may be a kind of coherence between these distinct notions small/large subsets enabling their transferal. Consider a universe T and sub-universe S⊆T with relative notions of small/large subsets. In case S⊆T happens to be 'almost as large as' T, it appears intuitively plausible that the small subsets of S are the subsets of T that are small subsets of T.

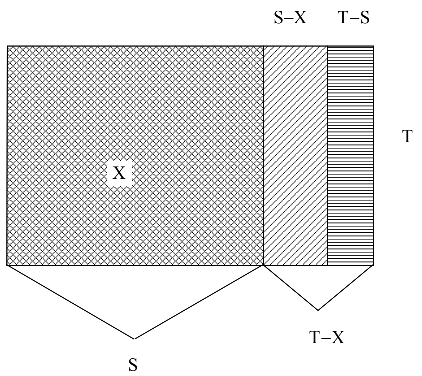

Figure 6.5: Coherent transfer in terms of exceptions

This *coherent transfer* for a subset S⊆T is as follows. If S is a large subset of T (S≈T), then for any set X⊆S: X is a large subset of S (X≈S) iff X is a large subset of T (X≈T).

The application of this coherent transfer to the preceding example yields the desired conclusion. Since $\underline{W} \subseteq \underline{K}$ with $\underline{W} \approx \underline{K}$, we can transfer $\underline{F} \cap \underline{K} \approx \underline{K}$ to $\underline{F} \cap \underline{K} \approx \underline{W}$, whence, as $\underline{F} \cap \underline{K} \subset \underline{F} \cap \underline{W}$, we can conclude $\underline{F} \cap \underline{W} \approx \underline{W}$, as desired.

This coherent transfer for a subset S⊆T is expressed in terms of ultrafilters as follows. For ultrafilters $\mathcal{U}_T$, over T, and $\mathcal{U}_S$, over subset S⊆T,

if $S \in \mathcal{U}_T$ (S≈T), then for every set X⊆S: $X \in \mathcal{U}_S$ (X≈S) iff $X \in \mathcal{U}_T$ (X≈T).

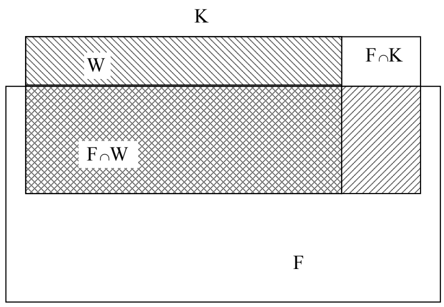

Figure 6.6: Flying birds with beaks and wings

Explanations for these comparisons between relative notions of large subsets in terms of ultrafilters, which justify them, can be provided by means of the concept of relativized family. The *relativization* of family $\mathcal{F} \subseteq \mathcal{P}$ (T) of subsets of T to subset S⊆T is the family $\mathcal{F} \cap \mathcal{P}(S) = \{X \subseteq S : X \in \mathcal{F}\}$ of subsets of S that are in family $\mathcal{F}$.

Now, coherent transfer for a subset S⊆T rests on the following observation about any given ultrafilter $\mathcal{U}_T$ over set T
if S∈$\mathcal{U}_T$ then the relativized family $\mathcal{U}_T \cap \mathcal{P}$(S) is an ultrafilter over S⊆T.

As an example, take T as the naturals and S as the even naturals. If the set of evens is considered a large set of naturals (S∈$\mathcal{U}_T$) then, the large set of naturals consisting only of evens form an ultrafilter over the set of even naturals.

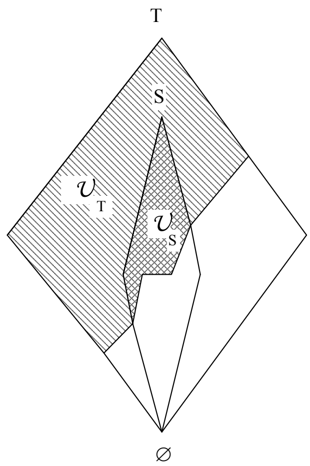

Figure 6.7: Ultrafilter over T and relativized family to large subset S⊆T

We can see that the converse also holds, i. e.

if the relativized family $\mathcal{U}_T \cap \mathcal{P}$(S) is an ultrafilter over S⊆T then S∈$\mathcal{U}_T$;

explaining the need for independent notions of large subsets when S∉$\mathcal{U}_T$.

Summing up the preceding discussions, consider an ultrafilter $\mathcal{U}_T$ over set T and a subset S⊆T.

– In case $S \notin \mathcal{U}_T$, we have $\mathcal{U}_T \cap \mathcal{P}(S)=\varnothing$, and we must have independent notions of large subsets: ultrafilters $\mathcal{U}_S$ and $\mathcal{U}_T$ are unrelated.
– In case $S \in \mathcal{U}_T$, the relativized family $\mathcal{U}_T \cap \mathcal{P}(S)$ is an ultrafilter over S, and we may decide to take $\mathcal{U}_S$ as $\{X \subseteq S : X \in \mathcal{U}_T\}$, thereby enforcing coherent transfer, or prefer to provide an ultrafilter U_S over S.

Notice that such questions and decisions fall outside the realm of our ultrafilter logic. Whether or not most birds with beaks have wings, most winged birds have beaks and most birds (with beaks) are penguins are ornithological, rather than logical, matters.

In the sequel we shall examine these ideas in the context of our sorted ultrafilter logic.

6.4 Sorted universes and 'large' subsets

We shall now consider comparison of universes with distinct notions of large subsets in a sorted framework. We shall examine how to formulate some ideas related to sub-universes as well as coherent transfer in a many-sorted approach. As before, we shall use variations of the preceding examples to introduce the ideas and some of their features.

We shall examine two kinds of comparison between universes:
– when a universe happens to be a subset of another one;
– how the notions of large subsets relative to each universe are related.

Let us first consider the former, the need for which appears as a side-effect of the sorted approach. We shall examine how to formulate the ideas of sub-universes and their intersections in a many-sorted approach.

In our sorted framework, sorts are unrelated: we have equality only over a sort, rather than between distinct sorts. Nevertheless, we can express some relationships among sorts by means of appropriate injections. The idea is that an injection from s to t establishes a bijection from its domain s onto its image i(s), the latter being a real subset of t.

Thus, we can express that a sort is a subsort of sort t as follows (Meré and Veloso (1992), (1995)): we use a unary function i from s to t together with an axiom asserting its injectivity. Thus, we express that sort s as a subsort of sort t ($s \subseteq t$) by a function $i : s \to t$ and the following injectivity axiom

$(i : \subseteq)(\forall x,x' : s)\ [i(x) \equiv i(x') \to x \equiv x']$ {injective $i : s \to t$ (for "all s's are t's)}.

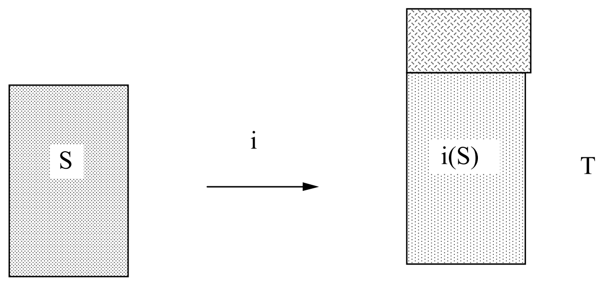

Figure 6.8: Subsort and injective function

As an illustration of subsorts, we examine a possible formulation for the assertions "Eagles and penguins are birds", "Eagles generally fly", and "Penguins generally do not fly".

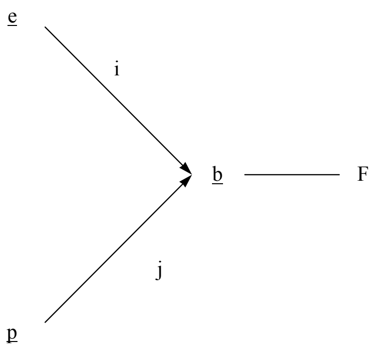

Figure 6.9: Signature for subsorts eagles and penguins of birds

Consider a signature with three sorts $\underline{b}$ (for birds), $\underline{e}$ (for eagles) and $\underline{p}$ (for penguins), as well as a unary predicate F over sort $\underline{b}$ (for flying birds).

We form a theory with sorted axioms as follows.

We express subsort information by injective functions $i : \underline{e} \rightarrow \underline{b}$ and $j : \underline{p} \rightarrow \underline{b}$

$(i : \subseteq)$ $(\forall x, x' : \underline{e})\ [i(x) \equiv i(x') \rightarrow x \equiv x']$ {injective $i : \underline{q} \rightarrow \underline{b}$ (all eagles are birds)},

$(j : \subseteq)$ $(\forall y, y' : \underline{p})\ [j(y) \equiv j(y') \rightarrow y \equiv y']$ {injective $j : \underline{p} \rightarrow \underline{b}$ (all penguins are birds)}.

We now express most information relative to sorts as follows:

$(\underline{e}\nabla F)$ $(\nabla x : \underline{e})\ F(i(x))$ {almost all eagles fly};

$(\underline{p}\nabla F)$ $(\nabla y : \underline{p}) \neg F(j(y))$ {almost all penguins do not fly}.

The extension by new generic constants $e : \underline{e}$ and $p : \underline{p}$ has genericity axioms

$$(\nabla x : \underline{e})\ F(i(x)) \leftrightarrow F(i(e)) \text{ and } (\nabla y : \underline{p})\ \neg F(j(y)) \leftrightarrow \neg F(j(p)).$$

Thus, the consequences of this generic extension include

(e) $F(i(e))$ {generic eagles, if any, do fly};
(p) $\neg F(j(p))$ {generic penguins, if any, do not fly}.

We also have transitivity of subsorts. As "Eagles are winged birds" and "Winged birds are birds" are expressed by injections, their composite gives an injection from the sort of eagles to that of birds, expressing "Eagles are birds".

$$\begin{array}{ccccc} \underline{e} & \xrightarrow{i} & \underline{w} & \xrightarrow{j} & \underline{b} \\ \| & & & & \| \\ \underline{e} & & \xrightarrow{j \circ i} & & \underline{b} \end{array}$$

Figure 6.10: Transitivity of subsorts by composition

We can also use the device of expressing subsort information by injections for introducing intersection of subsorts of a common sort. The idea is based on the so-called pullback representation for the intersection subsets of a common universe (MacLane (1971); Goldblatt (1979)).

The next example illustrates this construction of intersection of subsorts of a common sort.

Given subsets Q (of Quakers) and R (of Republicans) of the set H (of humans), we wish to form the intersection $Q \cap R$ (for the Republican Quakers), assumed nonempty. We may proceed as follows.

Consider a signature ν with three sorts $\underline{h}$ (for humans), $\underline{q}$ (for Quakers) and $\underline{r}$ (for Republicans) as well as functions $i:\underline{q}\rightarrow\underline{h}$ and $j:\underline{r}\rightarrow\underline{h}$.

We express subsort information by injectivity axioms for i and j

$(i : \subseteq)$ $(\forall y,y' : \underline{q})$ $[i(y)\equiv i(y')\rightarrow y\equiv y']$ {all Quakers are humans},
$(j : \subseteq)$ $(\forall z,z' : \underline{r})$ $[j(z) \equiv j(z')\rightarrow z\equiv z']$ {all Republicans are humans)}.

Sorts $\underline{q}$ (of Quakers) and $\underline{r}$ (of Republicans) are not directly connected, but we can resort to their copies within sort $\underline{h}$ (of humans).

Indeed, within sort $\underline{h}$ of humans, we have:

– the image $i(\underline{q})$ as a copy of the sort $\underline{q}$ of Quakers, and
– the image $j(\underline{r})$ as a copy of the sort $\underline{r}$ of Republicans.

So, their intersection $i(\underline{q})\cap j(\underline{r})$, consisting of the humans that are both Quakers and Republicans, can be defined by the formula $\mu(v)$ as follows:

μ (v): $(\exists\ y : \underline{q})$ $(\exists\ z : \underline{r})$ $[i(y)\equiv v \wedge v\equiv j(z)]$ {humans that come from $\underline{q}$ and from $\underline{r}$}.

Now, we wish to introduce a sort $\underline{n}$ (for the intersection $Q\cap R$), i. e. a new sort $\underline{n}$ such that

– sort $\underline{n}$ is a common subsort of sorts $\underline{q}$ and $\underline{r}$;
– sort $\underline{n}$ behaves as $Q\cap R$ i. e. as $i(\underline{q})\cap j(\underline{r})$.

Clearly, the assumption "Some Quaker (human) is a Republican (human)":

(∅) $(\exists y : \underline{q})$ $(\exists z : \underline{r})$ $i(y) \equiv j(z)$ {some Quaker and Republican coincide in $\underline{h}$},

is equivalent to having a nonempty intersection $(i(\underline{q}) \cap j(\underline{r}) \neq \emptyset)$ $(\exists v : \underline{h}) \mu(v)$ {some human is both Quaker and Republican}.

Indeed, the intersection $Q \cap R$ is in one-to-one correspondence with the following set of pairs (the pullback of functions $i{:}\underline{q} \rightarrow \underline{h}$ and $j{:}\underline{r} \rightarrow \underline{h}$)

$N := \{ \langle q,r \rangle \in \underline{q} \times \underline{r} : i(q) = j(r) \}$ {Quakers and Republicans coinciding as humans};

and we shall use this set $N \subseteq \underline{q} \times \underline{r}$ to represent the intersection $Q \cap R$.

Now, the assignment $\langle q,r \rangle \mapsto q$ defines an injective function k from $N \subseteq \underline{q} \times \underline{r}$. Similarly, we have an injection $l{:}N \rightarrow \underline{r}$ given by $\langle q,r \rangle \mapsto r$. Also, for every $n = \langle q,r \rangle \in N$, we have $i(k(n)) = i(q) = j(r) = j(l(n))$.

We can then introduce a new sort $\underline{n}$ (for N representing $Q \cap R$), via (⊆) injections $k{:}\underline{n} \rightarrow \underline{q}$ and $l{:}\underline{n} \rightarrow \underline{r}$ {$\underline{n}$ is a common subsort of $\underline{q}$ and $\underline{r}$}.

We know that we have the commutativity $i \equiv k = j \equiv l$: (=) $(\forall x : \underline{n})$ $i(k(x)) \equiv j(l(x))$ {same image as humans}.

By (=), this common image is included in the intersection $i(\underline{q}) \cap j(\underline{r})$:

$(\forall x : \underline{n})$ $\mu(i(k(x)))$ {every human coming from $\underline{n}$ is Quaker and Republican}.

Now, to make sort $\underline{n}$ a copy of $N \subseteq \underline{q} \times \underline{r}$ we require cover by joint subjectivity

(c) $(\forall y : \underline{q})\ (\forall z : \underline{r})\ [i(y) \equiv j(z) \rightarrow (\exists x : \underline{n})(y \equiv k(x) \wedge z \equiv l(x))]$ {any $n \in N$ comes from a $\underline{n}$}

But, given ($\subseteq$), (=) and (c) are jointly equivalent to the intersection axiom

($\cap$) $(\forall y : \underline{q})\ (\forall z : \underline{r})\ [i(y) \equiv j(z) \leftrightarrow (\exists x : \underline{n})(y \equiv k(x) \wedge z \equiv l(x))]$ {$\underline{n}$ in bijection with N};

so we use ($\subseteq$) and ($\cap$) to characterize sort $\underline{n}$ as the intersection of $\underline{q}$ and $\underline{r}$.

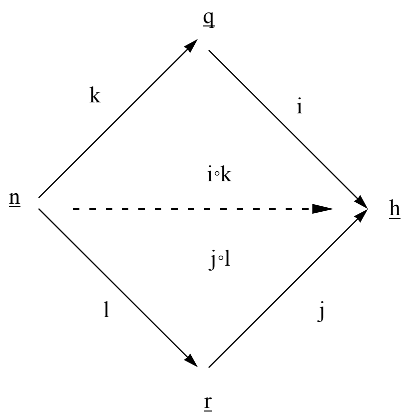

Figure 6.11: Intersection of subsorts of a common sort

Thus, we can circumvent some side-effects of the sorted approach.

Let us now turn to the other comparison, namely how the notions of large subsets relative to each universe are related. Recall that we have introduced distinct notions of large subsets to capture the idea of a set 'almost as large as' its universe.

We can also use the idea of expressing subsort information by an injective function to formulate a sorted version of coherent transfer as follows.

An injection i from s to t, establishes a bijection between each subset $X \subseteq s$ and its image $i(X) \subseteq t$, and between the families $\wp(s)$ and $\{i(X) \subseteq t : X \subseteq s\}$. This suggests regarding as plausible the equivalence between the assertions

– X is a large subset of s ($X \approx s$)

and

– i(X) is a large subset of t ($i(X) \approx t$),

for a subset $X \subseteq s$. In particular, it indicates a sorted version of subsort s is 'almost as large as' t ($s \approx t$), namely $i(s) \approx t$, i. e. $i(s) \in \mathcal{U}_t$.

This leads to the following version of sorted coherent transfer

if $i(s) \in \mathcal{U}_t$ ($i(s) \approx t$), then for any subset $X \subseteq s$: $X \in \mathcal{U}_s$ ($X \approx s$) iff $i(X) \in \mathcal{U}_t$ ($i(X) \approx t$).

Another formulation is suggested by the observation that if $i(s) \approx t$ (image i(s) 'almost covers' t), then the non-image $t - i(s)$ is a small subset of t, and so a subset $Y \subseteq t$ is 'about as large as' its pre-image $i^{-1}(Y) \subseteq s$: $Y \approx i^{-1}(Y)$.

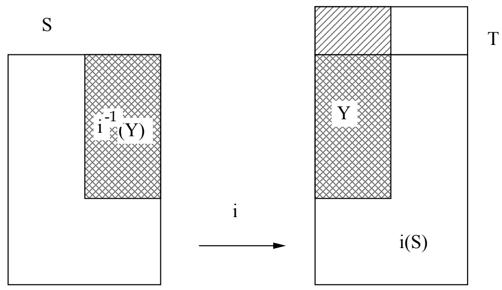

Figure 6.12: Subsort and coherent transfer by injection

So, we formulate *coherent transfer under injection* i : s→t as follows

if $i(s)\in\mathcal{U}_t$, then for any subset $Y\subseteq t$: $i^{-1}(Y)\in\mathcal{U}_s$ ($i^{-1}(Y)\approx s$) iff $Y\in\mathcal{U}_t$ ($Y\approx t$).

As before, explanation and justification can be provided by adapting the concept of relativized family. The *pre-image* of family $\mathcal{F}\subseteq\wp(t)$ of subsets of t under function: s→t is the family $i^{-1}(\mathcal{F}) := \{i^{-1}(Y)\subseteq s : Y\in\mathcal{F}\}$ of subsets of s.

Now, coherent transfer under injection i : s→t reflects the following observation about any given ultrafilter $\mathcal{U}_t$ over sort t

if $i(s)\in\mathcal{U}_t$ then the pre-image family $i^{-1}(\mathcal{U}_t)$ is an ultrafilter over sort s.

We now wish to express this coherent transfer under injection within our sorted ultrafilter logic. For this purpose, we may proceed as follows.

Consider an injection i:s→t (expressing that s is a subsort of t).

First, we see that $i(s) \in \mathcal{U}_t$ $(i(s) \approx t)$ asserts that image i(s) *almost covers* sort t, which we can express by the following formula of $L^{\nabla}(\rho)$

$(\nabla : i)$ $(\nabla\, y : t)(\exists\, x : s)\; y \equiv i(x)$ {Almost every object of t comes from s via i}.

Similarly, we can express that the pre-image of those objects of t with property φ(y) *almost covers* sort s by replacing variable y : t by a new variable x over sort s not occurring in φ(y) as follows

$(\nabla\, x : s)\; \varphi(i(x))$ [where variable x : s does not occur in φ(y)].

Then, we can express *most transfer of formula* $(\nabla\, y : t)\varphi$ of $L^{\nabla}(\rho)$, asserting "Almost every object of t has property φ(y) iff almost every object of s has property φ(i(x))", by the formula of $L^{\nabla}(\rho)$

$(y \mid \varphi : i)$ $(\nabla\, y : t)\; \varphi \leftrightarrow (\nabla\, x : s)\; \varphi(y/(i(x))$ [where $x \notin occ(\varphi)$].

Now, we formulate sorted coherent transfer for formula $(\nabla\, y : t)$ φ of $L^{\nabla}(\rho)$ by forming the *coherence axiom* $(i : y \mid \varphi)$ *of formula* φ as $(\nabla : i) \rightarrow (y \mid \varphi : i)$, i.e.

$(i : y \mid \varphi)(\nabla\, y : t)(\exists\, x : s)\; y \equiv i(x) \rightarrow [(\nabla\, y : t)\; \varphi \leftrightarrow (\nabla\, x : s)\varphi\; (y/(i(x)))$

[where $\notin occ(\varphi)$].

Thus, we formulate sorted coherent transfer for a set Φ of most formulas of $L^{\nabla}(\rho)$ by means of the following *coherence axiom schema for* Φ

$$(i : \Phi) := \{(\nabla : i) \rightarrow (y \mid \varphi : i) : (\nabla\ y : t)\ \varphi \in \Phi\}.$$

In particular, when Φ is the set of all the most formulas of $L^{\nabla}(\rho)$, we have the *coherence axiom schema for injection* I

$$(i : \rho) := \{(i : y \mid \varphi) : (\nabla\ y : t)\ \varphi \in L^{\nabla}(\rho)\}.$$

For instance, in our example of flying eagles and birds in 6.4, we have the following coherence axiom $(i : z \mid F(z))$ of formula $(\nabla\ z : \underline{b})\ F(z)$

$$(\nabla\ z : \underline{b})(\exists\ x : \underline{e})\ y \equiv i(x) \rightarrow [(\nabla\ z : \underline{b})\ F(z) \leftrightarrow (\nabla\ x : \underline{e})\ F(i(x))],$$

expressing the assertion "If almost all birds are eagles then most birds fly iff most eagles fly".

As mentioned, we can see that this formulation is compatible with the idea of independent notions of large subsets in case subsort s of t is not 'almost as large as' t. For, if we do not know the antecedent $(\nabla : i)\ \{i(s) \approx t\}$, the usage of the coherence axioms $(i : \varphi)$ becomes blocked.

6.5 Sorted framework for 'almost all' and 'generic'

We now examine many-sorted ultrafilter logic with coherence transfers as a sorted framework for 'almost all' and 'generic' reasoning.

First, let us examine sorted formulations for variants of our preceding examples on flying and non-flying birds seen in 6.1 and 6.3.

Consider a signature β with three sorts $\underline{p}$ (for penguins), $\underline{w}$ (for winged birds) and $\underline{k}$ (for birds with beaks) as well as a unary predicate F over sort $\underline{k}$ (for flying birds with beaks).

We express subsort information by injective functions $i : \underline{w} \to \underline{k}$ and $j : \underline{p} \to \underline{k}$

$$(i : \subseteq)\ (\forall x, x' : \underline{w})\ [i(x) \equiv i(x') \to x \equiv x']$$ {all winged birds have beaks}.

$$(j : \subseteq)\ (\forall y, y' : \underline{e})\ [j(y \equiv j(y') \to y \equiv y']$$ {all penguins are birds with beaks};

as well as the coherence axiom schemata for injections $i : \underline{w} \to \underline{k}$ and $j : \underline{p} \to \underline{k}$

$$(i : \beta) := \{(\nabla : i) \to (z \mid \psi : i) : x \notin occ(\psi), (\nabla z : \underline{k})\psi \in L^{\nabla}(\beta)\},$$

$$(j : \beta) := \{(\nabla : j) \to (z \mid \theta : j) : x \notin occ(\theta), (\nabla z : \underline{k})\ \theta \in L^{\nabla}(\beta)\}.$$

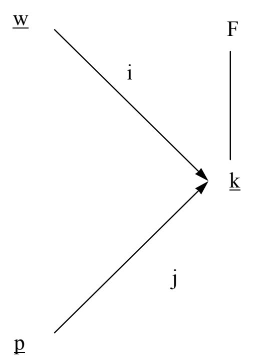

Figure 6.13: Signature for penguins, winged birds and birds with beaks

We first illustrate how desired conclusions can be derived when we know "Most birds with beaks have wings".

We express most information relative to sorts $\underline{w}$ and $\underline{k}$ as almost coverage

$(\nabla : i)$ $(\nabla\, z : \underline{k})(\exists\, x : \underline{w})\; z \equiv i(x)$ {almost all birds with beaks have wings}.

Then, from the instance (with $(\nabla\, z : \underline{k})\, F(z)$ for $(\nabla\, z : \underline{k})\, \psi$)

$$(\nabla\, z : \underline{k})(\exists\, x : \underline{w})\; z \equiv i(x) \rightarrow [(\nabla\, z : \underline{k})\, F(z) \leftrightarrow (\nabla\, x : \underline{w})\, F(i(x))]$$

of schema $(i : \beta)$, we can conclude the equivalence between

$(\underline{k}\nabla F)$ $(\nabla\, z : \underline{k})\, F(z)$ {almost all birds with beaks fly},

and

$(\underline{w}\nabla F)$ $(\nabla\, x : \underline{w})\, F(i(x))$ {almost all winged birds fly}.

Thus, in the presence of $(\nabla : i)$: "Most birds with beaks have wings", from $(\underline{k}\nabla F)$ "Most birds with beaks fly" we can conclude $(\underline{w}\nabla F)$ "Most winged birds fly".

We now illustrate how unexpected conclusions can be blocked when we do not know "Most birds with beaks are penguins".

Assume the absence of information on almost coverage of sort $\underline{k}$ by $\underline{p}$

$(\nabla : j)$ $(\nabla\, z : \underline{k})(\exists\, y : \underline{w})\; z \equiv j(y)$ {almost all birds with beaks are penguins}.

Then, the usage of instances of the axiom schema $(j : \beta)$ is blocked.

In fact, assume we also know the following most information about sort $\underline{p}$:

$(\underline{p}\ \nabla\neg F)\ (\nabla\ y : \underline{p})\ -F(j(y))$ {almost all penguins do not fly}.

Then, the instance of the axiom schema (j:β) with F(z) for θ, i.e.

$$(\nabla\ z : \underline{k})(\exists\ y : \underline{w})\ z\equiv j(y) \rightarrow [(\nabla\ z : \underline{k})\ F(z) \leftrightarrow (\nabla\ y : \underline{p})\ F(j(y))].$$

yields the conclusion

$(\underline{k}\nabla\neg\underline{p})(\nabla z{:}\underline{k})\neg(\exists y{:}\underline{p})\ z\equiv j(y)$ {almost all birds with beaks are not penguins}.

Thus, the fact $(\underline{k}\nabla F)$: "Most birds with beaks fly" does not force the conclusion $(\underline{p}\nabla F)$: "Most penguins fly" in the absence of (j : β): "Most birds with beaks are penguins", forcing instead $(\underline{k}\nabla\neg\ \underline{p})$: "Most birds with beaks are not penguins" in the presence of $(\underline{p}\ \nabla\neg\ F)$: "Most penguins do not fly".

This example illustrates how the coherence axiom schema for an injection provides uniform control based on the relative sizes of the sorts. In addition, finer control can be achieved by selecting a particular set of formulas to be coherently transferred.

In general, the framework is as follows. We consider a sorted theory consisting of the following sets of axioms:

– a set Σ of axioms expressing (basically syntactical) subsort information,
– a set Δ of axioms expressing coherent transfers between some subsorts,
– a set Γ of axioms expressing the remaining available knowledge.

Since set $\Gamma\cup\Sigma\cup\Delta$ specifies a many-sorted ultrafilter theory, we have

$$\Gamma\cup\Sigma\cup\Delta\vDash^{\nabla}\tau \text{ iff } \Gamma\cup\Sigma\cup\Delta\vdash^{\nabla}\tau.$$

We now consider some sorted variations on the so-called "Nixon example".

The situation is as follows. Assuming that "Quakers generally are pacifist", "Republicans generally are not pacifist" and "Nixon is a Republican Quaker", what can one conclude Nixon's attitude regarding pacifism?

Even assuming that Nixon is a typical/generic Republican Quaker, one appears to be left with the question "Does Nixon behave generally as a Republican or as a Quaker"?

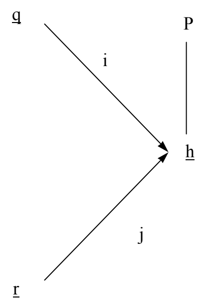

Figure 6.14: Basic signature for "Nixon example"

Consider the expansion $\nu' := \nu \cup\{P : \underline{h}\}$ of signature ν in 6.4 by a unary predicate P over sort $\underline{h}$ (for pacifist human).

We express most information on pacifism for sorts $\underline{q}$ and $\underline{r}$ as follows

$(\underline{q}\nabla P)$ $(\nabla\, y : \underline{q})$ $P(i(y))$ {almost all Quakers are pacifist},

$(\underline{r}\nabla\neg P)$ $(\nabla z : \underline{r})$ $\neg P(j(z))$ {almost all Republicans are not pacifist}.

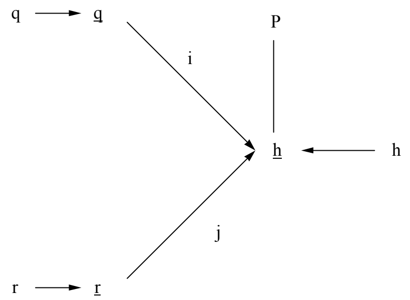

Figure 6.15: Signature for "Nixon example" with (generic) constants

Clearly, by adding new generic constants $q : \underline{q}$ and $r : \underline{r}$, we will have

– P(i(q)) {any generic Quaker is pacifist},
– P(j(r)) {any generic Republican is not pacifist}.

We may also add the coherence axiom schemata for injections i and j

$(i : \nu') := \{(\nabla : i) \to (v \mid \psi : i) : y \notin occ(\psi), (\nabla\, v : \underline{k})\ \psi \in L^{\nabla}(\nu')\}$,

$(j : \nu') := \{(\nabla : j) \to (v \mid \theta : j) : z \notin occ(\theta), (\nabla\, v : \underline{k})\ \theta \in L^{\nabla}(\nu')\}$.

Consider adding a new constant $h : \underline{h}$ for some (maybe generic) human.

Even if we know that such human is a Quaker, i. e. $(\exists y : \underline{q})\ i(y)\equiv h$, we cannot conclude anything about h's pacifist attitude: neither $P(h)$ nor $\neg P(h)$.

Now, assume that such human represents most Quakers, in the strong sense $(\nabla\ y : \underline{q})\ i(y)\equiv h$. We will then have $P(h)$.

But, this assumption is indeed very strong. For, if $h : \underline{h}$ is generic, it yields $(\underline{h}\nabla P)\ (\nabla\ v : \underline{h})\ P(v)$ {almost all humans are pacifist}.

Also, it is equivalent to almost coverage of humans by Quakers $(\nabla : i)\ (\nabla\ v : \underline{h})(\exists\ y : \underline{q})\ v\equiv i(y)$ {almost all humans are Quakers},
(in view of transfer of $\exists$ over ∇ and coherence axiom schema $(i : \nu')$).

On the other hand, if we had assumed other relative sizes of universes

$$(\nabla : j)\ (\nabla\ v : \underline{h})(\exists\ z : \underline{r})\ v\equiv j(z)\ \text{\{almost all humans are Republicans\}},$$

we would conclude

$$(\underline{h}\nabla\neg P)\ (\nabla\ v : \underline{q})\neg\ P(v)\ \text{\{almost all humans are not pacifist\}}.$$

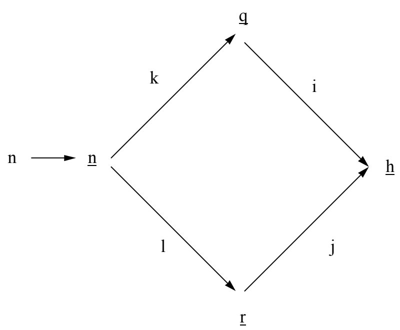

Figure 6.16: "Nixon example" with generic Republican Quaker

Now, let us examine the impact of the, apparently strong, assumption "Nixon is a generic Republican Quaker".

We now know that there is a Republican Quaker. So, we may expand the above signature ν' to ν* by introducing a new sort n (for Republican Quakers) with injections k:n→q and l:n → r, as in 6.4, as well as a new generic constant n : n.

We may also add the coherence axiom schemata for injections k and l

$$(k : \nu^*) := \{(\nabla : k) \rightarrow (y \mid \psi : k) : x \notin occ(\psi), (\nabla\, y : \underline{q})\, \psi) \in L^{\nabla}(\nu^*)\},$$

$$(l : \nu^*) := \{(\nabla : l) \rightarrow (z \mid \theta : l) : x \notin occ(\theta), (\nabla\, z : \underline{r})\, \theta) \in L^{\nabla}(\nu^*)\}.$$

In the absence of information about relative sizes, we cannot conclude anything about the pacifist attitude of such a Republican Quaker.

Assume that we know "Most Quakers are Republican Quakers" (Q∩R≈Q), which expresses the almost coverage of q by n, i. e.

(∇ : k) $(\nabla\, y : \underline{q})(\exists\, x : \underline{n})\, y \equiv k(x)$ {almost all Quakers are Republican Quakers}.

Then, the coherence axiom (∇ : k) → (y | P(i(y)) : k) for injection k:n → q yields (n∇P) $(\nabla\, x : \underline{n})\, P(i(k(x)))$ {almost all Republican Quakers are pacifist}.

Hence, as n is a generic constant for sort n, we can conclude

– P(i(k(n))) {any generic Republican Quaker is pacifist}.

Now, if we knew instead "Most Republicans are Republican Quakers" (Q∩R≈R), i. e. "Almost all Republicans are Republican Quakers", as

(∇ : l) (∇ y : r)(∃ x : n) y≡l(x) {almost all Quakers are Republican Quakers}.

(r∇ q) (∇ z : r)(∃ y : q) i(y)≡j(z) {almost coverage of r by n};

we would have

(n∇¬P) (∇ x : n) ¬P(j(l(x))) {almost all Republican Quakers are not pacifist},

whence, by genericity

– ¬P(j(l(x))) {any generic Republican Quaker is not pacifist}.
Finally, if we assume both pieces of information

– "Most Quakers are Republican Quakers" (Q∩R≈Q);
– "Most Republicans are Republican Quakers" (Q∩R≈R), and
we will have (since, as expected, Q≈Q∩R≈R) conflicting conclusions

– "Almost all Republican Quakers are pacifist" (n∇P), and
– "Almost all Republican Quakers are not pacifist" (n∇¬P);

from which, the commutativity i(k(n))≡j(l(n)) will yield the contradiction

"A generic Republican Quaker is pacifist and not pacifist".

Notice that the conclusions agree with what is intuitively expected.

– Suspended judgement, in the absence of further information.
– Republican Quakers generally behave as Quakers, if one knows $Q \cap R \approx Q$.
– Republican Quakers generally behave as Republicans, if one has $Q \cap R \approx R$.
– Contradictory behavior as Quakers and Republicans, in case both $Q \cap R \approx Q$ and $Q \cap R \approx R$.

This example illustrates once more the importance of knowledge concerning relative sizes in establishing deductive connections.

It may also serve to clarify another point about relative most information.

Consider the assertion "Most Quakers are Republicans".

Does it assert that there are many Republicans among Quakers?

Our more precise reading of it is "Almost all Quakers are Republicans", which implies "Almost all Quakers are Republican Quakers" ($Q \cap R \approx Q$).

An explanation comes from an analysis of the role of the copies of subsorts in expressing relative most assertions.

Recall that, within sort $\underline{h}$ of humans, we have:

– the image $i(\underline{q})$ as a copy of the sort $\underline{q}$ of Quakers, and
– the image $j(\underline{r})$ as a copy of the sort $\underline{r}$ of Republicans.

(In particular, the intersection sort n̲ of Republican Quakers has three copies, within sorts: q̲ of Quakers, r̲ of Republicans and h̲ of humans.)

Now, consider the assertion "Almost all Quakers are Republicans".

We will argue that it can, and should, be understood as "Almost all Quakers are humans that are Republican (humans)".

The former is supposed to mean that the set of Republicans form a large set of Quakers. Since sorts q̲ (of Quakers) and r̲ (of Republicans) are not directly connected, we resort to their copies within sort h̲ (of humans).

This suggests understanding "The set of Republicans form a large set of Quakers" as "The set of Quakers that, as humans, are Republican humans form a large set of Quakers", which is the meaning of "Almost all Quakers are humans that are Republican (humans)".

We now examine this situation in our sorted framework.
Recall that we have the intersection axiom

($\cap$) $(\forall y : \underline{q})\ (\forall z : \underline{r})\ [i(y) \equiv j(z) \leftrightarrow (\exists x : \underline{n})(y \equiv k(x) \wedge z \equiv l(x))]$ {n̲ behaves as $Q \cap R$}.

This intersection axiom asserts the equivalence between

($\underline{q}\nabla\ \underline{r}$) $(\nabla\ y : \underline{q})(\exists\ z : \underline{r})\ i(y) \equiv j(z)$ {almost all Quakers are Republican humans},
(∇ : k) $(\nabla\ y : \underline{q})(\exists\ x : \underline{n})\ y \equiv k(x)$ {almost all Quakers are Republican Quakers}.

The behavior of sort n̲ as the intersection $Q \cap R$ yields the equivalence of

– almost all Quakers are Republican humans (q∇ r), and
– the almost coverage of q by n (∇ : k).

Thus, we can conclude "Almost all Quakers are Republican Quakers" (Q∩R≈Q) from the paraphrase "Almost all Quakers are humans that are Republican humans" of "Almost all Quakers are Republicans".

Similarly, "Almost all Republicans are Quakers", paraphrased as "Almost all Republicans are humans that are Quakers" yields the almost coverage "Almost all Republicans are Republican Quakers" (Q∩R≈R).

7 SOME PROSPECTS

We finally comment on some perspectives and directions for further work, specifically some interesting connections with fuzzy logic and inductive and empirical reasoning, which suggest the possibility of other applications for our ultrafilter logic.

The basic idea is exploiting the expressive power of ultrafilter logic and perhaps extend it by some extra generalized quantifiers such as ∇.

A first possible application is to the realm of imprecise reasoning, in the spirit of fuzzy logic (Turner (1984)). Some common ground is indicated by the basic intuitions of 'large', 'almost all', 'very few' and 'about as large as'.

For instance, a fuzzy concept, such as 'very tall' might be explicated as: a 'very tall' person is a person that is taller than 'almost' everybody (else).

Thus, we may consider extracting (fuzzy) concepts from a binary predicate L, by definitions such as:

– property H(y) as ∇xL(x,y) {for "y is very high (or tall)"};

– property S(y) as $\neg$H(y) {for "x is very short (or low)"};
– relation D(x,y)as∇z[L(x,z)$\wedge$L(z,y)]{for "y is much higher than x"};
– relation F(x,y) as D(x,y)$\vee$D(y,x) {for "x is far from y" (in height};
– relation E(x,y) as $\neg$F(x,y) {for "x is about as high as y"}.

As an illustration, consider an expansion $\mathcal{M}^{\mathcal{U}} = (\mathcal{M}, \mathcal{U})$ of a non-standard first-order structure $\mathcal{M}$ = <**N**+**Z**,< for the naturals by a Fréchet ultrafilter.

– the standard naturals n$\in$**N** are very short: $\mathcal{M}^{\mathcal{U}} \vDash \neg\nabla$xL(x,y) **[**n**]**;
– the non-standard numbers z$\in$**Z** are very high: $\mathcal{M}^{\mathcal{U}} \vDash \nabla$xL(x,y)**[**z**]**;
– a non-standard number z$\in$**Z** is much higher than a standard natural n$\in$**N**: $\mathcal{M}^{\mathcal{U}} \vDash \nabla$v[L(x,v)$\wedge$L(v,y)] **[**n,z**]**;
– numbers in the same copy are about as high, i.e. for a,b$\in$**N** or a,b$\in$**Z**: $\mathcal{M}^{\mathcal{U}} \vDash \neg\nabla$v[L(x,v)$\wedge$L(z,v)] **[**a,b**]**.

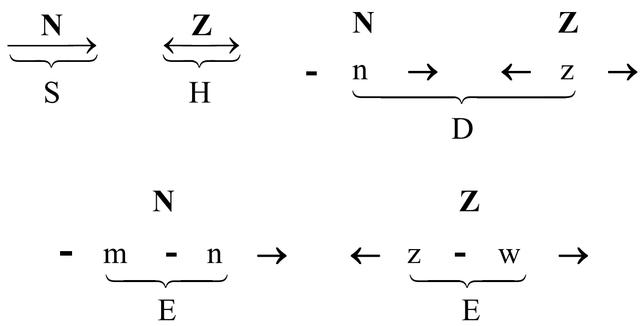

Figure 7.1: Standard and non-standard numbers

Such approach may provide alternative qualitative foundations for (versions of) fuzzy logic.

Another possible application could be to the area of inductive reasoning, as in empirical experiments and tests. This arises from the observation that, whereas laws of pure mathematics may be of the form "All M's are N's", one can argue that laws of natural sciences are really assertions of the – more cautious – form "Most M's are N's", or at least should be regarded in this manner. Here, the expressive power of ∇ may be helpful.

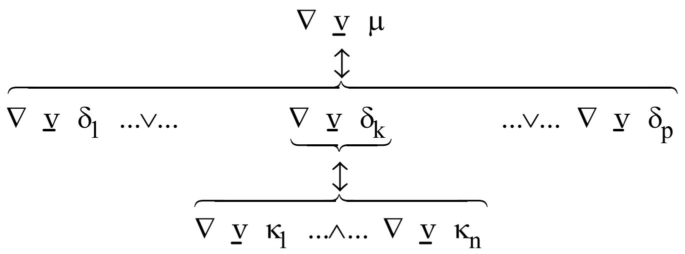

Figure 7.2: Conditions for establishing empirical law $\nabla\, \underline{v}\, \mu$

For instance, consider such an empirical law λ of the form $\nabla\underline{v}\, \mu$, with $\nabla\underline{v}$ being a short hand for the most prefix $\nabla v_1 \ldots \nabla v_m$. Let us assume that μ is quantifier-free – a reasonable assumption in the context of experiments. Then, μ can be put into disjunctive normal form. Thus μ is equivalent to a disjunction $\delta_1 \vee \ldots \vee \delta_p$, so $\nabla\underline{v}\mu$ will be established iff some disjunct $\nabla\underline{v}\, \delta_k$ can be established. Such a disjunct δ_k is in turn equivalent to a conjunction $\kappa \wedge_1 \wedge \ldots \wedge \kappa_n$. Thus, establishing λ is reduced to establishing $\nabla\underline{v}\, \kappa_1, \ldots, \nabla\underline{v}\, \kappa_n$, which are independent tasks, each one involving a literal (an atomic formula, perhaps negated).

Now, consider an inductive jump: having established $\varphi(\underline{v})$ for a (small) set of objects ($\forall\underline{v}\, [\varepsilon(\underline{v}) \rightarrow \varphi\, (\underline{v})]$), one wishes $\nabla\underline{v}\, \varphi(\underline{v})$. The latter would follow from $\nabla \underline{v}\ \varepsilon(\underline{v})$, but this involves large experimental evidence. The case of program testing may be illustrative: one tests the behavior of a program for a (small) set of data and then argues that the program will exhibit this behavior in general. Here, the rationale is that the set of test data is 'representative' in that it covers the possible execution paths.

This idea suggests the following strategy: "experimental evidence that is 'almost' dense for a similarity enables the inductive jump".

Establishing

(τ): $\forall \underline{u} \nabla \underline{v}\ [S(\underline{u},\underline{v}) \rightarrow (\varphi(\underline{u}) \rightarrow \varphi(\underline{v}))]$ {similarity 'almost' transfers property},

(δ): $\nabla \underline{v} \exists \underline{u}\ [\varepsilon(\underline{u}) \wedge S(\underline{u},\underline{v})]$ {experiments 'almost' dense for similarity};

is sufficient for concluding $\nabla \underline{v}\ \varphi(\underline{v})$ from $\forall \underline{v}\ [\varepsilon(\underline{v}) \rightarrow \varphi(\underline{v})]$.

For small experimental evidence ($\neg \nabla \underline{v}\ \varepsilon(\underline{v})$), it is wise to take:

– $\nabla \underline{u}\ \nabla \underline{v}\ [\varepsilon(\underline{u}) \vee S(\underline{u},\underline{v}) \vee \varepsilon(\underline{v})]$ {economic representative set}.

The applications outlined above suggest two other interesting avenues.

The first avenue concerns the weakening of some mathematical concepts.

The idea of 'almost' dense is close to that of 'almost' coverage, which has been found useful in expressing connections between sorts. Along similar lines, concepts such as 'almost' equal or 'almost' disjoint might be useful.

The basic idea is weakening some universal quantifiers to ∇. For instance, one might consider the concept of 'almost equivalence', obtained by weakening symmetry and transitivity and replacing reflexivity by $\nabla y\ \exists x\ R(x,y)$. Analogously, an 'almost partition' would amount to a set of blocks 'almost' covering the universe where intersecting blocks would have 'almost' the same elements. Similar weakening of some mathematical concepts might be of interest. (Note that we are not proposing a program; one can expect that only some such weakenings – typically with qualitative flavor – will be of interest.)

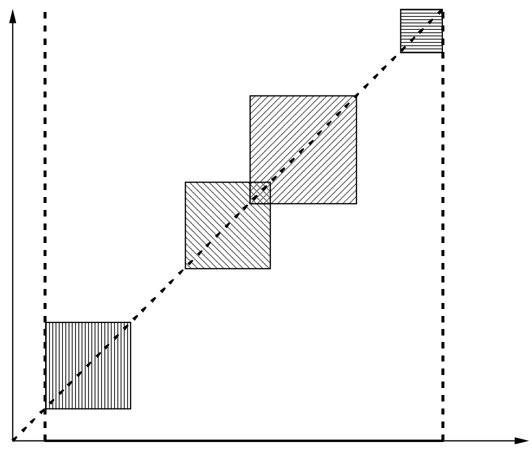

Figure 7.3: 'Almost equivalence' relation

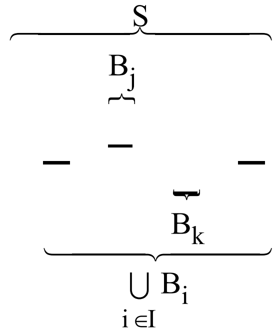

Figure 7.4: 'Almost partition'

The other avenue comes from the idea that we generally wish a fairly small set of experiments. This suggests considering an extension of our logic to deal with finiteness. The idea is considering another quantifier $\circ$, the intended interpretation of $\circ v\, \theta(v)$ being "$\theta(v)$ holds only for fairly few elements". The semantic interpretation involves expanding an ultrafilter structure $\mathcal{A}^{\mathcal{U}} = (\mathcal{A}, \mathcal{U})$ by a family N of fairly small subsets of the universe A. Taking this family N as the ideal $\wp_{\omega}(A)$ of the finite subsets of A might be of interest in case A is infinite. In general, we may wish to have less stringent conditions on such a family $\mathcal{N}$ of fairly small subsets.

8 CONCLUSIONS

We have examined a logical system with generalized quantifiers over ultrafilters, which serves as a rigorous basis for qualitative reasoning with notions such as 'all but a few' and 'very few', as well as 'typical' and 'generic'. This monotonic logical system is a conservative extension of classical first-order logic, with which it shares several properties.

For expressing 'most' assertions a generalized quantifier ∇ has been introduced with the intended interpretation "holding 'almost' universally" and the intuition of 'almost all' as 'but for a few exceptions' has been rendered precise by means of ultrafilters.

This approach leads to an extension of classical first-order logic by the addition of a generalized quantifier for 'almost universally'. Its semantics is given by adding an ultrafilter to a first-order structure. By formulating the characteristic properties of ultrafilters, we obtain an axiomatics, which is shown to be sound and complete with respect to this semantics. As a result, this ultrafilter logic turns out to be a conservative extension of classical first-order logic.

For generic reasoning, we have explicated the ideas of 'typical' and 'generic' individuals in terms of "possessing the properties that almost all individuals have". The addition of generic constants produces a conservative extension, where one can correctly reason about generic objects as intended.

The expression of almost all by means of a generalized quantifier over an ultrafilter captures the idea of holding 'almost universally' in a 'a given universe. Unfortunately, simple relativization fails to express adequately 'relative most' assertions. To circumvent this problem, relative notions of large have been introduced by means of an ultrafilter over each given universe, leading naturally to a many-sorted version of ultrafilter logic. To control deductive connections, comparisons among sub-universes with relative notions of large, given by corresponding ultrafilters, have been introduced. Sorted ultrafilter theories with axioms expressing such connections have been shown to handle correctly relative notions of large subsets and inheritances, thereby providing a sorted framework for 'almost all' and 'generic' reasoning.

We have also commented on some perspectives for further work. Some interesting connections with fuzzy logic and inductive and empirical reasoning suggest the possibility of other applications for our logic (Grácio (1999); Veloso and Carnielli (2001)). Ultrafilter logic extends conservatively classical first-order logic, but still sharing several properties, such as compactness and Löwenheim-Skolem. It appears to merit further investigation (Rentería, Haeusler and Veloso (2002)).

Resumo: Certos argumentos empregam objetos 'genéricos' ou 'típicos'. Sugere-se uma explicação para (alguns aspectos desta) idéia em termos de 'quase todos'. A intuição de 'quase todos' como 'todos exceto por algumas exceções' é tornada precisa através de ultrafiltros. Propõe-se um sistema lógico, com quantificadores generalizados para 'quase todos', como uma base para raciocínio genérico. Esta lógica é monotônica, tem um cálculo dedutivo simples, que é correto e completo, e é uma extensão conservativa da lógica clássica de primeira ordem, com a qual compartilha várias propriedades. Para raciocínio genérico, introduz-se a idéia de indivíduos genéricos, que são internalizados como constantes genéricas, originando extensões conservativas onde se pode raciocinar sobre objetos genéricos conforme almejado. Considera-se também uma versão poli-sortida dessa lógica a fim de tratar distintas noções de subconjuntos 'grandes'. Além disso, indicam-se outras possíveis aplicações para tal lógica.

Palavras chave: *Lógica de ultrafiltros, 'quase todos', semântica, axiomatização, corretude, completude, objetos 'típicos', constantes genéricas, noção relativa de 'quase todos', lógica de ultrafiltros sortida, raciocínio com 'quase todos' e 'genérico' sortidos.*

REFERENCES

ANTONIOU, G. (1997). *Nonmonotonic Reasoning* (Cambridge, Mass., MIT Press).

BARWISE, J. (ed.). (1977). *Handbook of Mathematical Logic* (Amsterdam, North-Holland).

BARWISE, J. & FEFERMAN, S. (eds.) (1985). *Model-Theoretic Logics* (Berlin, Springer-Verlag).

BELL, J. L. & SLOMSON, A B. (1971). *Models and Ultraproducts: an introduction* (Amsterdam, North-Holland).

BESNARD, P. (1989). *An Introduction to Default Logic* (Berlin, Springer-Verlag).

BESNARD, P., BREWKA, G., FROIDEVAUX, C., GRÉGOIRE, E. & SIEGEL, P. (1991). "Nonmonotonicity", *Journal Applied Non-Classical Logics* **1**(2), pp. 267-310.

BREWKA, G. (1991). *Nonmonotonic Reasoning: logical foundations of commonsense.* (Cambridge, Cambridge University Press).

BIRKHOFF, G. & LIPSON, J. D. (1970). "Heterogeneous algebras", *J. Combin. Th.*, ser. A **8**, pp. 115-133.

CARNIELLI, W. A. & SETTE, A. M. (1994). *Default operators*, (Abstr.) Workshop on Logic, Language, Information and Computation, Recife, Brazil.

CARNIELLI, W. A. & VELOSO, P. A. S. (1997). "Ultrafilter logic and generic reasoning". In Gottlob, G., Leitsch, A. and Mundici, D. (eds.) *Computational Logic and Proof Theory* (5th Kurt Gödel Colloquium: KGC'97) {LNCS 1289} (Berlin, Springer-Verlag), pp. 34-53.

CHANG, C. C. & KEISLER, H. J. (1973). *Model Theory* (Amsterdam, North-Holland).

CHURCH, A. (1956). *Introduction to Mathematical Logic,* vol. I. (Princeton, NJ, Princeton University Press).

ENDERTON, H. B. (1972). *A Mathematical Introduction to Logic* (New York, Academic Press).

FLUM, J. (1985). "Characterizing logics". In: Barwise and Feferman (eds.) (1985), p. 77-120.

FREGE, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reines Denken.* Louis Nebert, Halle {English translation in van Heijenoort, J. (ed.) *From Frege to Gödel* (Cambridge, Mass., Harvard University Press),}.

GÄRDENFORS, P. (1988). *Knowledge in Flux: modelling the dynamics of epistemic states* (Cambridge, Mass., MIT Press).

GOLDBLATT, R. (1979). *Topoi: the Categorial Analysis of Logic* (Amsterdam, North-Holland).

GRÁCIO, M.C.G. (1999). *Lógicas Moduladas e Raciocínio sob Incerteza.*Ph. D. Dissertation, UNICAMP, Campinas.

HALMOS, P. R. (1972). *Lectures on Boolean Algebras* (London, van Nostrand Reinhold).

HEMPEL, C. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York, Free Press).

HENKIN, L. (1949). "The completeness of the first-order functional calculus", *Journal of Symbolic Logic* **14**, pp. 159-166.

KEISLER, H. J. (1970). "Logic with the quantifier 'there exist uncountably many'", *Annals of Math. Logic* **1**, pp. 1-93.

LINDSTRÖM, P. (1966). "On extensions of elementary logic", *Theoria* **35**, pp. 1-11.

LUKAZIEWICZ, W. (1990). *Non-monotonic Reasoning: formalization of commonsense reasoning* (Chichester, Ellis Horwood).

MACLANE, S. (1971). *Categories for the Working Mathematician* (Berlin, Springer-Verlag).

MAKINSON, D. and GÄRDENFORS, P. (1991). "Relations between the logic of theory change and non-monotonic logic". In: Fuhrmann, A. and Morreau, M. (eds.) *The Logic of Theory Change* {LNAI} (Berlin, Springer-Verlag).

MAREK, V.W. and TRUSZCZYNSKI, M. (1993). *Nonmonotonic Logic: context-dependent reasoning* (Berlin, Springer-Verlag).

MERÉ, M.C. & VELOSO, P.A.S. (1992). "On extensions by sorts" (Rio de Janeiro, PUC-Rio, Dept. Informática, Res. Rept. MCC 38/92).

———. (1995). "Definition-like extensions by sorts", *Bull. IGPL* **3**(4), pp. 579-595.

REITER, R. (1980). "A logic for default reasoning", *J. Artificial Intelligence* **13**(1), pp. 81-132.

RENTERÍA, C. J., HAEUSLER, E. H. & VELOSO, P. A. S. (2002). "NUL: dedução natural para lógica de ultrafiltros" (Rio de Janeiro, PUC-Rio, Dept. Informática, Res. Rept. MCC 16/02).

SCHLECHTA, K. (1995). "Defaults as generalized quantifiers", *Journal Logic and Computation* **5**(4), pp. 473-494.

SETTE, A. M., CARNIELLI, W. A. & VELOSO, P. A. S. (1999). "An alternative view of default reasoning and its logic". In: Haeusler, E.H. and Pereira, L.C. (eds.) *Pratica: Proofs, Types and Categories* (Rio de Janeiro, PUC-Rio), pp. 127-158.

TARSKI, A. (1930). "Fundamentale Begriffe der Methodologie der deduktiven Wissenschaften". *Monatshefte für Mathematik und*

Physik **37**, pp. 361-404 {English translation in Tarski (1956), pp. 60-109}.

———. (1936). "Der Wahrheitsbegiff in den formalisierten Sprachen". *Studia für Mathematik und Philosophia* **1**, pp. 261-405 {English translation in Tarski (1956), pp. 152-278}.

———. (1956). *Logic, Semantics and Metamathematics: papers from 1923 to 1938*. Woodger, J. H. (transl.) (Oxford, Clarendon).

TURNER, W. (1984). *Logics for Artificial Intelligence* (Chichester Ellis Horwood).

VELOSO, P. A. S. (1999). "On ultrafilter logic and a missing axiom", *Bull. Sect. Logic*, 28(1), pp. 17-26.

VELOSO, P. A. S & CARNIELLI, W. A (1997). "An ultrafilter logic for generic reasoning and some applications". (Rio de Janeiro, COPPE-UFRJ, Progr. Eng, Sist. Comput., Res. Rept. ES-437/97).

———. (2001). "Logics for qualitative reasoning". *CLE e-Prints* **1** {http://www.cle.unicamp.br/e-prints}.