ELSEVIER Annals of Pure and Applied Logic 92 (1998) 217–234

ANNALS OF PURE AND APPLIED LOGIC

In conjunction with qualitative probability

Tim Fernando *

IMS, University of Stuttgart, D-70174 Stuttgart, Germany

Received December 1996

Communicated by Ph.G. Kolaitis

Abstract

Numerical probabilities (associated with propositions) are eliminated in favor of qualitative notions, with an eye to isolating what it is about probabilities that is essential to judgements of acceptability. A basic choice point is whether the conjunction of two propositions, each (separately) acceptable, must be deemed acceptable. Concepts of acceptability closed under conjunction are analyzed within Keisler's weak logic for generalized quantifiers – or more specifically, filter quantifiers. In a different direction, the notion of a filter is generalized so as to allow sets with probability non-infinitesimally below 1 to be acceptable. © 1998 Elsevier Science B.V. All rights reserved.

AMS Classification: 03B48; 03C80

Keywords: Qualitative probability; Probabilistic logic; Generalized quantifiers; Non-monotonic reasoning

1. Introduction: Weighing the evidence

Let $\mathscr{L}$ be a set of "formulas" $(\varphi, \psi, \ldots)$, and V be a set of "possibilities" $(a, b, \ldots)$, connected to $\mathscr{L}$ by a relation $\models$, capturing the intuition that for all $a \in V$ and $\varphi \in \mathscr{L}$,

$$a \models \varphi \quad \text{iff} \quad a \text{ "supports" } \varphi.$$

We can, for concreteness, take the case of predicate logic, although we can also proceed more abstractly, assuming only some relation $\models \subseteq V \times \mathscr{L}$. In any case, the point is to pick out a set $\mathscr{A} \subseteq \mathscr{L}$ of "acceptable" formulas on the basis of $\models$, so that for every $\varphi \in \mathscr{L}$,

φ is *acceptable* iff the "evidence" $\{a \in V\colon a \models \varphi\}$ supporting φ "has enough weight"

* E-mail: fernando@ims.uni-stuttgart.de.

0168-0072/98/$19.00
PII S0168-0072(98)00007-4

or, in other words,

$$\varphi \in \mathscr{A} \quad \text{iff} \quad \{a \in V \colon a \models \varphi\} \in \mathscr{H} \tag{1}$$

for some family $\mathscr{H} \subseteq \mathsf{Pow}(V)$ of "heavy" subsets of V. Clearly, the safest choice is to equate $\mathscr{A}$ with the $\models$-validities (i.e., the formulas φ such that for every $a \in V$, $a \models \varphi$), which is to say, to set $\mathscr{H} = \{V\}$. But suppose we took a *chance* on some other choice of $\mathscr{A}$ (and $\mathscr{H}$) that tolerates failures or exceptions. What are the rules for doing so *rationally*?

"Logic" suggests the closure condition

$$(\mathrm{Up}) \quad \frac{\varphi \in \mathscr{A} \qquad \varphi \vdash \psi}{\psi \in \mathscr{A}} \qquad \frac{A \in \mathscr{H} \qquad A \subseteq B \ (\subseteq V)}{B \in \mathscr{H}}$$

where $\varphi \vdash \psi$ means that for every a, if $a \models \varphi$ then $a \models \psi$. (Up) alone introduces no element of risk into $\mathscr{A}$ (or $\mathscr{H}$) insofar as (Up) does not yield, from the assumption that $V \in \mathscr{H}$, any more elements of $\mathscr{H}$. On the other hand, (Up) does suggest how to proceed – namely, by weakening $\vdash$ to some binary relation $\mid\!\sim$ on $\mathscr{L}$ or (turning to $\mathscr{H}$) $\subseteq$ to some binary relation $\preccurlyeq$ on $\mathsf{Pow}(V)$

$$(\mathrm{Up})_{\mid\sim} \quad \frac{\varphi \in \mathscr{A} \qquad \varphi \mid\!\sim \psi}{\psi \in \mathscr{A}} \qquad (\mathrm{Up})_{\preccurlyeq} \quad \frac{A \in \mathscr{H} \qquad A \preccurlyeq B}{B \in \mathscr{H}}.$$

It is natural to expect that the intuitions that come to play in developing the rule $(\mathrm{Up})_{\mid\sim}$ are syntactic (or proof-theoretic), whereas those for $(\mathrm{Up})_{\preccurlyeq}$ are semantic. An example where this distinction matters concerns the operation $\wedge$ of conjunction on $\mathscr{L}$, for which it is understood that for every $a \in V$, $a \models \varphi \wedge \psi$ iff $a \models \varphi$ and $a \models \psi$. It is largely on this simple example that the present paper turns.

1.1. Content of paper

There is a certain plausibility to asserting

(And) if φ and ψ are both acceptable, then $\varphi \wedge \psi$ is acceptable

if only because

(i) it takes a bit of sophistication to even sense the difference between "φ and ψ" and "$\varphi \wedge \psi$";

and

(ii) after such sophistication is acquired, we learn that the difference does not (in a sense) matter, if acceptability is construed as validity (semantic or syntactic).

(And) becomes problematic, however, as soon as we accept some exceptions. This is brought out most clearly perhaps by H. Kyburg's "lottery paradox": the proposition that one in say, a million tickets in a lottery will win is acceptable, as are each of a million propositions asserting that a particular ticket will not win. The "paradox" vanishes after a moment's thought on the underlying semantics: were we to agree that

$$\mathscr{A} = \{\varphi \in \mathscr{L} \colon \text{there is at most one } a \in V \text{ for which } a \not\models \varphi\},$$

we may run up against a pair φ and ψ of formulas in $\mathscr{A}$ such that the one counterexample to φ is different from the one counterexample to ψ, whence $\varphi \wedge \psi \notin \mathscr{A}$.

But as long as we concentrate on the syntactic side $\mathscr{A}$, choosing the semantics $\mathscr{H}$ to support the manipulations on $\mathscr{A}$ that we have decided to legitimize, there is hope for (And). In this regard, it is worthwhile noting what Pearl [17] calls a "long-standing tension between the logical and probabilistic approaches to dealing with such exceptions": whereas the former is "prescriptive" (insofar as logic is simply a record of "conversational conventions"), probabilities are "descriptive" (be they measures of objective frequencies or subjective beliefs). Now, the question is how could probabilities *describe* the *logical* rule (And)? A natural way to proceed is to accept precisely the formulas with probability greater than some fixed threshold $\alpha \in [0, 1]$ (say, 0.999), given a probability function *pr* from $\mathscr{L}$ to the unit interval $[0, 1]$

$$\mathscr{A}_{pr,\alpha} = \{\varphi \in \mathscr{L}\colon \ pr(\varphi) > \alpha\}.$$

The hitch is that exceptions add up: from $pr(\varphi) = 1 - \delta$ and $pr(\psi) = 1 - \varepsilon$, one cannot, in general, do better than predict $pr(\varphi \wedge \psi) \geqslant 1 - (\delta + \varepsilon)$. This suggests, assuming we stick with pobability measures (rather than some alternative where, for example, conjunction is interpreted by the greatest lower bound operation in some lattice), that we

(I) replace the condition that $pr(\varphi) > \alpha$ by the requirement that $1 - pr(\varphi)$ be *infinitesimal*, where infinitesimals are assumed to be closed under addition: if δ and ε are infinitesimals, then so is $\delta + \varepsilon$.

The notion of an infinitesimal here is exactly that introduced by A. Robinson in his non-standard reconstruction of calculus. Reversing chronological order, we could, as an alternative to (I),

(II) explain away infinitesimals by $\varepsilon\delta$-type limits (à la Bolzano–Weierstrass).

Pearl [18] describes the work of Adams and Spohn in much this way, though without the emphasis on (And). I have decided here to focus on (And) because, together with (Up), it supports a very direct and general analysis of approaches (I) and (II), using no more structure than that implicated by line (1) above. This is made precise by Theorems 1 and 2 in Sections 2.1 and 2.2, respectively, where, in particular, no appeal is made to numbers, be they in the standard unit interval $[0, 1]$, or some non-standard copy thereof. This is not to say that the analysis given is incompatible with a numerical approach; only that it allows us to avoid all kinds of arithmetical complications – not to mention the somewhat embarrassing question: what probability function?

Without specifying a particular probability function, we return, in Section 3, to probabilities, beefing up the rule (Up) to a rule $(\mathrm{Up})_{\preccurlyeq}$, where $\preccurlyeq$ means "at least as probable as". In this section, we refrain from making any commitment to the soundness or, for that matter, unsoundness of (And). $(\mathrm{Up})_{\preccurlyeq}$ applies not only to models of (And), but also to $\mathscr{A}_{pr,\alpha} = \{\varphi \in \mathscr{L} : pr(\varphi) > \alpha\}$, where *pr* is a probability function, and α is some number in $[0, 1]$. (It is easy enough to introduce a rule restricting α say, to be above $\frac{1}{2}$). As weak as the rule $(\mathrm{Up})_{\leqslant}$ might be, it nevertheless constitutes a step to understanding what is involved *qualitatively* in accepting formulas with probabilities that fall non-infinitesimally short of 1.

1.2. Related work

There is an evidently widespread belief that it is a non-trivial (if not hopeless) enterprise to reason *qualitatively* about formulas with probabilities greater than some fixed threshold non-infinitesimally short of 1. Pearl [16] puts the matter as follows

> Probabilities that are infinitesimally close to 0 and 1 are very rare in the real world. Most default rules used in ordinary discourse maintain a certain percentage of exceptions, simply because the number of objects in every meaningful class is finite. Thus, a natural question to ask is, why study the properties of a logic that applies only to extreme probabilities? Why not develop a logic that characterizes moderately high probabilities, say probabilities higher than 0.5 or 0.9 – or more ambitiously, higher than α, where α is a parameter chosen to fit the domains of the predicates involved?
> The answer is that any such alternative logic would be extremely complicated and probably would need to invoke many axioms of arithmetic. [pp. 493, 494]

A similar view can be found in Halpern and Rabin [6], where "a logic to reason about likelihood" is developed relative to a *non*-probabilistic semantics. The probabilistic approach to likelihood presented in Section 3 (below) is not only qualitative but quite possibly simpler than might have been feared. It proceeds along lines similar to Scott [19] and Segerberg [20], as discussed further below.

The literature on non-monotonic formalisms that incorporate (And) as a basic or derived rule is vast and, for the uninitiated, downright bewildering. (See, for instance, Pearl [17] and the references cited therein,[1] plus Kyburg [13] for earlier work.) What the present paper offers is a logical approach that departs minimally (if at all) from standard practice in classical mathematical logic. The one possible point of departure is the appeal to the *weak logic of generalized quantifiers* in Keisler [9] (which arguably belongs to the mainstream of logic) for formalizing line (1) above,[2] and even then, the introduction of generalized quantifiers can be eliminated according to Theorems 1 and 2 below, resulting in ordinary predicate logic. Some specialists in the logic of generalized quantifiers seem to consider Theorem 1 part of the subject's folklore. It is implicit in van Lambalgen [22], and appears in a disguised form as Theorem 5 of Alechina and van Lambalgen [1], the inessential notational differences being due to that work's somewhat novel syntax (involving modality) and semantics (motivated by proof theory).

A crucial syntactic point that ought to be stressed is the expulsion of numbers from formulas below – in contrast, that is, to the quantitative approaches in Keisler [10] and

---

[1] A well-known example is Kraus et al. [12], in which the rule $(\mathrm{Up})_{\mid\sim}$ described above leads dangerously to monotonicity (p. 180). Instead, a weak system of cumulative reasoning is developed there, from which a $\mid\sim$-form of (And) but not $(\mathrm{Up})_{\mid\sim}$ can be derived

[2] A fine point about (1) and weak logic is that the extension of $\mathscr{A}$ to a binary predicate $\mid\sim$ on $\mathscr{L}$ can be treated by passing from unary to binary generalized quantifiers, at the cost only of notational clutter; see Section 2.3.

Halpern [7], where numbers appear explicitly in formulas. The idea behind minimizing (explicit) reference to probabilities is to isolate what it is about probabilities that is essential to judgements of acceptability; but by opening the door to alternative non-probabilistic interpretations of formulas, the challenge then becomes showing that only the probabilistic semantics need matter. (More concretely, the problem in establishing completeness is how to define a probability measure from syntactic entities that do not mention numbers.)

2. Reasoning according to preference: filters

A straightforward formalization of line (1) is provided by the *weak logic* $L(Q)$ for generalized quantifiers of Keisler [9] where

(i) L is a first-order language and Q is a generalized quantifier symbol, inducing formulas $Qx\varphi$ (in addition to the usual closure rules on first-order L-formulas)

and

(ii) an $L(Q)$-model is a pair (M,q) consisting of a first-order L-model M and a family $q \subseteq \mathsf{Pow}(|M|)$ of subsets of the universe $|M|$ of M, so that (relativizing (1) to (M,q))

$$(M,q) \models Qx\varphi[f] \quad \text{iff} \quad \{a \in |M| \colon (M,q) \models \varphi[f^x_a]\} \in q$$

for every function f mapping variables to objects in $|M|$ (and where f^x_a is the function that maps x to a, but is otherwise identical to f).

That is to say, (1) is analyzed by building an L-model M around the set V of possibilities so that the acceptability of a formula φ can be evaluated by exposing (as it were) the "hidden variable" x, the instantiations of which are measured relative to q $(=\mathscr{H}_{M,q})$

$$\varphi \in \mathscr{A}_{M,q} \quad \text{iff} \quad (M,q) \models Qx\varphi\,.$$

$L(Q)$ offers not only the expressive power of predicate logic (as well as the possibility of nested judgments of acceptability through iterations of Q), but also a natural model theory, relative to which a complete proof system can be obtained from a simple extension of one for first-order logic by axiom schemes for α-equivalence and extensionality (Keisler [9]).

2.1. Filter quantifiers

Given a set V, a *filter on* V is a non-empty family $\mathscr{H}$ of subsets of V satisfying (Up) and closed under intersections – viz., $V \in \mathscr{H}$; whenever $A \in \mathscr{H}$ and $A \subseteq B \subseteq V$, $B \in \mathscr{H}$; and for all $A \in \mathscr{H}$ and $B \in \mathscr{H}$, $A \cap B \in \mathscr{H}$ (thereby supporting (And)). These properties translate in $L(Q)$ to the (filter) schemes

(Q1) $Qx\ x = x$,

(Q2) $\forall x(\varphi \supset \psi) \supset (Qx\varphi \supset Qx\psi)$,

(Q3) $Qx\varphi \wedge Qx\psi \supset Qx(\varphi \wedge \psi)$.

Let us write $F[Q]$ for the $L(Q)$-theory induced by (Q1), (Q2) and (Q3), where φ and ψ are $L(Q)$-formulas with the same set of free variables. Note that $F[Q]$ holds for $Q = \forall$, or restricted universal quantification. The converse is not quite true: consider the standard model M of arithmetic, and let q be the family

$$\{A \subseteq \{0,1,\ldots\} : \{0,1,\ldots\} - A \text{ is finite}\}$$

of *co-finite* sets of natural numbers; then (M,q) is a model of $F[Q]$, even though q is a *non-principal* filter. (A filter q on V is *principal* if $\bigcap q \in q$, in which case $\bigcap q$ is said to *generate* q.) Nevertheless, the converse can be approximated through elementary extensions and some slick bookkeeping due to van Lambalgen [22]. Fix a relation symbol R not in L that accepts any finite positive number of arguments, and call an $L(Q)$-model (M,q) *relatively principal* if under some expansion of (M,q) to R,

$$Qx\varphi \equiv \forall x(R(x,\overline{y}) \supset \varphi) \tag{2}$$

holds for every $L(Q)$-formula φ with free variables $x, \overline{y}$ (where $\overline{y}$ is say, ordered according to some fixed well-ordering of variables). Line (2) says that Q can be taken to be universal quantification $\forall$ restricted to some set R of "generic" (or "normal") elements (modulo $\overline{y}$). These generic elements are "transcendental" in $\overline{y}$: assuming $\forall y Qx(x \neq y)$, they cannot be named by L-terms with free variables drawn from $\overline{y}$. (Hence, the necessity of adding $\overline{y}$ to R.)

Theorem 1.[3] *Every $L(Q)$-model of $F[Q]$ can be elementarily extended to a relatively principal $L(Q)$-model.*

Proof. Fix an $L(Q)$-model (M,q) of $F[Q]$, and a finite set Φ_0 of instances of (2). By the compactness theorem of weak logic (Keisler [9]), it suffices to show how to expand (M,q) to a model of Φ_0. The idea is to interpret R as the set

$$\{(F_\psi(\overline{a}), \overline{a}) \colon \psi \text{ is an } L(Q)\text{-formula, } \overline{a} \in dom(F_\psi)\},$$

for certain partial functions F_ψ (to be defined presently) from the set of finite sequences of (the universe) $|M|$ (of M) to $|M|$. Let n be the number of free variables in $Qx\psi$. The domain of F_ψ consists exactly of the n-tuples $\overline{a} \in |M|^n$ such that

$$(*) \quad (M,q) \not\models Qx\psi[\overline{a}].$$

Given such a sequence $\overline{a}$, let φ be the conjunction

$$\bigwedge\{\varphi_0 : Qx\varphi_0 \text{ has the same set of free variables as } Qx\psi,\\ (M,q) \models Qx\varphi_0[\overline{a}] \quad \text{and} \quad `Qx\varphi_0 \equiv \forall x(R(x,\overline{y}) \supset \varphi)' \in \Phi_0\}$$

[3] See the note in Section 1.2 for bibliographic information.

(appealing here to the finiteness of Φ_0). It is understood that an empty conjunction is some tautology. Using (Q3) in case the set is non-empty, it follows that

$$(M,q) \models Qx\varphi[\bar{a}].$$

Hence, by assumption (*) and $Qx\varphi \wedge \neg Qx\psi \supset \exists x(\varphi \wedge \neg\psi)$, to which (Q2) can, as pointed out to me by N. Alechina, be rewritten,

$$(M,q) \models \exists x(\varphi \wedge \neg\psi)[\bar{a}].$$

It remains to choose some such witness for the value of $F_\psi(\bar{a})$. □

Remark. The interpretation of R suggested in the proof of Theorem 1 can be described roughly as follows: for every $L(Q)$-formula ψ such that $(M,q) \not\models Qx\psi$, throw in a witness to the set

$$\{\neg\psi\} \cup \{\varphi \in L(Q) : (M,q) \models Qx\varphi\}$$

of formulas. What is "rough" about this description is that (a) we should be more careful to specify what variables and constants to allow in the formulas, and (b) the required witnesses may not exist in M. The first point is a simple matter of book-keeping, while the second can be handled by appealing to the existence of elementary extensions that are *ω-saturated* – suggesting a restatement of Theorem 1 as

Proposition 1′. *Every ω-saturated $L(Q)$-model of $F[Q]$ is relatively principal.*

In fact, Proposition 1′ can be sharpened to so-called *recursively saturated* models (Barwise [3]), as the sets of formulas that must be realized can be given effectively as follows:

$$\{(\neg Qx\psi) \supset \neg\psi\} \cup \{(Qx\varphi) \supset \varphi\colon\ \varphi \in L_n(Q)\}$$

for every $\psi \in L_n(Q)$, where L_n is an expansion of the language L to n fresh constants (abusing notation in identifying $L_n(Q)$ with the set of $L_n(Q)$-formulas with one free variable x). Note the similarity of $(\neg Qx\psi) \supset \neg\psi$ to Henkin expansions that witness existential statements. The next section adapts notions from a celebrated method, the *force* of which is to *omit* rather than to *realize types.*

2.2. Graded normality

The proper extensions mentioned in Theorem 1 can be avoided, provided (2) is weakened to allow for varying grades (rather than an absolute, either/or, notion) of genericity: add an argument place to R, for a relation symbol $\sqsubset$ to be used with the intuition that

$$u \sqsubset v \quad \text{iff} \quad u \text{ is "more generic" than } v.$$

More precisely, let L be a countable first-order language, $\sqsubset$ be a fresh binary relation symbol (not in L), and let $y \sqsubseteq z$ abbreviate $(y \sqsubset z) \vee (y = z)$. Given an L-model M, let L_M be L together with a fresh constant symbol for every object in M. To simplify the notation, we will identify an object a in M with its constant symbol in L_M, and write M for the L_M-model obtained by expanding the L-model M to L_M, with every a in M interpreted as a.

Theorem 2. *For every countable (or finite) $L(Q)$-model (M,q) of $F[Q]$, there is a transitive binary relation $\prec \subseteq |M| \times |M|$ such that for every $L_M(Q)$-formula φ with exactly one free variable x,*

$$Qx\varphi \equiv (\forall z)(\exists y \sqsubseteq z)(\forall x \sqsubset y)\varphi \tag{3}$$

holds in $(M,q,\prec)$.

Remarks. 1. Line (3) describes an $\varepsilon\delta$-limit/asymptotic/cofinal-type quantification, with (2) falling out as the special case given by

$$u \sqsubset v \;\equiv\; R(u)$$

(i.e., the second argument v in $\sqsubset$ is vacuous).

2. The converse of Theorem 2 (the soundness of $F[Q]$ under (3)) is trivial: the schemes (Q1) and (Q2) follow from (3) alone, while (Q3) is a consequence of the transitivity of $\prec$ (as well as (3)). (I have not investigated what generality (3) buys beyond that of (2), in the absence of the assumption that $\sqsubset$ is transitive.)

3. The restriction to one free variable is inessential, and is made only to simplify notation, allowing us to suppress the subscripts $\bar{y}$ on $\sqsubset$ in (3).

4. An interpretation of $\sqsubset$ validating (3) in Theorem 2 is more complicated to describe than an interpretation of R supporting (2) in Theorem 1. In this connection, it is interesting to note the sentiment

> any fool can realize a type, but it takes a model-theorist to omit one

expressed in Sacks [18]. The twist in Theorem 2 is that the "ontological promiscuity" in saturation arguments is avoided by a purely combinatorial argument (without resorting to any of the model-theorist's tools, such as completeness).

Proof of Theorem 2. Fix an $L(Q)$-model (M,q) of $F[Q]$, and partition the set Φ of $L_M(Q)$-formulas with exactly one free variable x as follows:

$$\Phi_+ = \{\varphi \in \Phi \colon (M,q) \models Qx\varphi\},$$
$$\Phi_- = \{\psi \in \Phi \colon (M,q) \not\models Qx\psi\}.$$

If for every $\varphi \in \Phi_+$, $(M,q) \models \forall x\varphi$, then we can set $\prec$ to $\{(a,a) : a \in |M|\}$ and we are done. Otherwise, choose a $\varphi_0 \in \Phi_+$ and $a_0 \in |M|$ such that

$$(M,q) \not\models \varphi_0\,[a_0],$$

and let $\{\varphi_0, \varphi_1, \varphi_2, \varphi_3, \ldots\}$ be an enumeration of Φ. We will define $\prec$ by finite approximations $\prec^i$ (for $i \geqslant 0$), with

$$\prec \; = \bigcup_{i \geqslant 0} \prec^i .$$

The plan is to construct for each $i \geqslant 0$, a finite transitive relation $\prec^i$ such that the following four conditions hold, with the understanding that

$$\Phi^i_+ = \Phi_+ \cap \{\varphi_j : j \leqslant i\}, \qquad \Phi^i_- = \Phi_- \cap \{\varphi_j : j \leqslant i\},$$

$$E^i = \{a \in dom(\prec^i) : (\forall b \prec^i a)\; a \prec^i b\} \quad (= \text{the set of } \prec^i\text{-minimal points}).$$

(C1) $\prec^i \subseteq \prec^{i+1}$.
(C2) $dom(\prec^{i+1} - \prec^i) \subseteq E^{i+1}$.
(C3) For every $a \in E^i$, $(M, q) \models \bigwedge \Phi^i_+[x/a]$.
(C4) There is a "witness" map $w_i : \Phi^i_- \to E^i$ such that for every $\psi \in \Phi^i_-$,

$$(M, q) \models \neg\psi \; [x/w_i(\psi)],$$

and for every $j > i$, and every $a \in |M|$,

$$a \preccurlyeq^j w_i(\psi) \text{ implies } w_j(\psi) \prec^j a.$$

What makes (C1)–(C4) interesting is

Lemma A. *If $\prec$ is the union $\bigcup_i \prec^i$ of finite transitive relations $\prec^i$ satisfying* (C1)–(C4), *then the required equivalence* (3) *holds.*

Proof. Suppose (first) that $(M, q) \models Qx\varphi$. Given $a \in |M|$, either there is some $a' \in |M|$ such that $a' \prec a$, or not. If not, then

$$(M, q, \prec) \models (\exists y \sqsubseteq a)(\forall x \sqsubset y)\varphi$$

holds vacuously (by the definition of $\sqsubseteq$ as the disjunction of $\sqsubset$ with equality). Otherwise, choose an i such that $\varphi \in \Phi^i_+$ and an $a' \preccurlyeq^i a$ which, by the transitivity and finiteness of $\prec^i$, belongs to E^i. Then by (C2) and (C3), $(M, q) \models (\forall x \sqsubset a')\varphi$.

Next, assume $(M, q) \not\models Qx\varphi$. Choose an i such that $\varphi \in \Phi^i_-$, and conclude from (C1) and (C4) that

$$(M, q, \prec) \models (\forall y \sqsubseteq z)(\exists x \sqsubset y)\neg\varphi[z/w_i(\varphi)]. \qquad \square$$

To push the construction of $\prec$ through, the following will be useful.

Lemma B. *For all $\varphi \in \Phi_+$, $\psi \in \Phi_-$, and every finite subset A_0 of*

$$\{a \in |M| \; : \; |M| - \{a\} \in q\},$$

$$\{a \in |M| \; : \; (M, q) \models (\varphi \wedge \neg\psi)[x/a]\} \not\subseteq A_0.$$

Proof. Repeated applications of (Q3) give

$$(M,q) \models Qx\left(\varphi \wedge \bigwedge_{a\in A_0} x\neq a\right),$$

and therefore, if $\psi \in \Phi_-$ then (Q2) implies

$$(M,q) \not\models \forall x\left(\left(\varphi \wedge \bigwedge_{a\in A_0} x\neq a\right)\supset\psi\right),$$

i.e.,

$$(M,q) \not\models \forall x\left((\varphi\wedge\neg\psi)\supset \bigvee_{a\in A_0} x=a\right),$$

as required. □

Looking more closely at $\{a\in|M| : |M|-\{a\}\in q\}$, observe that (Q2) implies

$$\{a\in|M| : |M|-\{a\}\notin q\} = \bigcap q$$

and

Lemma C. *For every* $a\in\bigcap q$ *and every* $\varphi\in\Phi_+$, $(M,q)\models\varphi[x/a]$.

Proof. If $(M,q)\not\models\varphi[x/a]$, then

$$\{b\in|M| : (M,q)\models\varphi[x/b]\} \subseteq |M|-\{a\},$$

so that $\varphi\in\Phi_+$ and (Q2) imply $|M|-\{a\}\in q$ (i.e., $a\notin\bigcap q$). □

Next, define

$$I^i = \{a\in|M|-E^i : (\exists b\in|M|)b\prec^i a\} \quad (= \text{the image of } \prec^i \text{ minus } E^i)$$

and add to the list (C1)–(C4)

(C5) For every $a\in I^i$, there is a $\varphi\in\Phi^i_+$ such that $(M,q)\not\models\varphi[a]$.
(C6) The witness map w_i mentioned in (C4) is surjective (onto E^i) and has the additional property that for all $\psi,\psi'\in\Phi^i_-$ such that $\psi\neq\psi'$, if $w_i(\psi)=w_i(\psi')$, then $w_i(\psi)\in\bigcap q$.

Let us turn finally to the definition of $\prec^i$ (and w_i). The initial stage $i=0$ is trivial: since $\varphi_0\in\Phi_+$, we can set $\prec^0=\emptyset$ (whence $E^0=\emptyset=w_0$). Now, consider stage $i+1$.

Case 1: $\varphi_{i+1}\in\Phi_+$. Let $N=\{a\in E^i : (M,q)\not\models\varphi_{i+1}[x/a]\}$. By Lemma C, for every $a\in N$, $a\notin\bigcap q$. Hence by (C6), each $a\in N$ has a unique $\psi_a\in\Phi^i_-$ such that $w_i(\psi_a)=a$. Next, apply (Q3) and Lemma B to define a function $new: N\to|M|$ such that for every $a\in N$,

$$(M,q)\models\left((\neg\psi_a)\wedge\bigwedge\Phi^{i+1}_+\right)[x/new(a)]$$

(whence $new(a) \notin I^i$ by (C5)) and

$$new(a) \in (E^i - N) \cup \{new(a') : a' \in N - \{a\}\} \quad \text{implies} \quad new(a) \in \bigcap q.$$

Then set $\prec^{i+1}$ to be the transitive closure of

$$\prec^i \cup \{(new(a), a) : a \in N\} \cup \{(new(a), new(a)) : a \in N\}$$

and define $w_{i+1} : \Phi_-^{i+1} \to E^{i+1}$ by

$$w_{i+1}(\psi) = \begin{cases} new(w_i(\psi)) & \text{if } w_i(\psi) \in N, \\ w_i(\psi) & \text{otherwise} \end{cases}$$

for every $\psi \in \Phi_-^{i+1}$ $(= \Phi_-^i$ by the case assumption).

Case 2: $\varphi_{i+1} \in \Phi_-$. Writing ψ for φ_{i+1}, choose, by (Q3) and Lemma B, an a such that

$$(M,q) \models \left((\neg\psi) \wedge \bigwedge \Phi_+^i\right)[x/a],$$

$$a \in \{w_i(\psi) : \psi \in \Phi_-^i\} \quad \text{implies} \quad a \in \bigcap q.$$

Then set

$$\begin{aligned} \prec^{i+1} &= \prec^i \cup \{(a, a_0), (a, a)\}, \\ w_{i+1} &= w_i \cup \{(\psi, a)\} \end{aligned}$$

(recalling that a_0 was the element chosen at the outset satisfying $\neg\varphi_0$, and that $a \notin I^i$, by (C5)).

These two cases together yield the following picture of $\prec^i$, for $i > 0$. At the top is a_0 (chosen to satisfy $\neg\varphi_0$ where $\varphi_0 \in \Phi_+$). Coming out of a_0 are branches, each with exactly one tip (i.e., an element of E^i). Each tip satisfies all of Φ_+^i. Moreover, each $\psi \in \Phi_-^i$ is satisfied at some tip (given by w_i). For different ψ and ψ' in Φ_-^i, either the corresponding branches meet only at a_0 or else the corresponding branches are the same, and will not grow further (in $\prec^j$, for $j > i$) because the tip of the branch satisfies all of Φ_+ (Lemma C). A branch witnessing ψ in Φ_-^i will grow further in $\prec^j$, where $j > i$, only in case the tip of that branch violates some $\varphi \in \Phi_+^j$ (which is not in Φ_+^i). With this picture, verifying conditions (C1) through (C6) becomes routine. □

2.3. *The binary case (from $\mathscr{A}$ to $\mid\!\sim$)*

Theorems 1 and 2 generalize to binary quantifiers (with only notational complications) as follows. To step from a unary quantifier up to a binary quantifier, an L-model M is paired with a binary relation $q \subseteq \mathsf{Pow}(|M|) \times \mathsf{Pow}(|M|)$ such that

$$(M,q) \models Qx(\varphi, \psi)[f] \quad \text{iff} \quad \{a \in |M| : (M,q) \models \varphi[f_a^x]\} \; q \; \{a \in |M| : (M,q) \models \psi[f_a^x]\}.$$

The weak completeness and compactness theorems of Keisler [9] lift immediately to this setting (as worked out, for instance, in Westerståhl [23]). The filter schemes (Q1)–(Q3) turn into

$$Qx(\varphi,\varphi),$$
$$\forall x(\psi \supset \psi') \supset (Qx(\varphi,\psi) \supset Qx(\varphi,\psi')),$$
$$Qx(\varphi,\psi) \wedge Qx(\varphi,\psi') \supset Qx(\varphi,\psi \wedge \psi'),$$

respectively, with line (2) becoming

$$Qx(\varphi,\psi) \equiv \forall x((\varphi \wedge R_{\varphi,x}(x,\overline{y})) \supset \psi), \tag{4}$$

where $\overline{y}$ lists the free variables $\neq x$ in φ as well as ψ. Line (4) supports a reading of the formula $Qx(\varphi,\psi)$ as "for all relevant φ-x's, ψ." Theorem 1 can be lifted to these binary forms, under the additional condition that R is "extensionalized" so that

$$\forall x\overline{y}(\varphi \equiv \varphi') \supset \forall x\overline{y}(R_{\varphi,x}(x,\overline{y}) \equiv R_{\varphi',x}(x,\overline{y}))$$

for all $L(Q)$-formulas φ and φ' with the same free variables $x, \overline{y}$. Similar remarks apply to Theorem 2. (I.e., the relation symbol $\sqsubset$ must also be relativized to the antecedent φ and the quantified variable x, although its extension can be arranged to depend on φ only up to $\equiv$.)

3. Between preferences and probabilities: quasi-filters

Having upgraded a "normality" predicate R into a "preference" relation $\sqsubset$, let us proceed further, into probability measures, concentrating on $\mathscr{H}$ (rather than on $\mathscr{A}$). Fix a non-empty set V, and a field $\mathscr{F}$ of subsets of V (containing V and closed under $\cup$ and $\bar{\cdot}$). Recall that a *(finitely additive) probability function on* $\mathscr{F}$ is a function $\mu : \mathscr{F} \to [0,1]$ such that $\mu(V) = 1$, and for all $A, B \in \mathscr{F}$,

$$\mu(A \cup B) = \mu(A - B) + \mu(B - A) + \mu(A \cap B).$$

3.1. Weakening filters non-conservatively

A family $\mathscr{H} \subseteq \mathscr{F}$ of sets in $\mathscr{F}$ is *sizable* if there is a probability function μ on $\mathscr{F}$ such that for every $A \in \mathscr{F}$,

$$A \in \mathscr{H} \quad \text{iff} \quad \mu(A) > \mu(\overline{A}) \quad (\text{i.e., } \mu(A) > \tfrac{1}{2}).$$

Given $\alpha \in [0,1]$, call a family $\mathscr{H} \subseteq \mathscr{F}$ *α-sizable* if there is a probability function μ on $\mathscr{F}$ such that for every $A \in \mathscr{F}$,

$$A \in \mathscr{H} \quad \text{iff} \quad \mu(A) > \alpha.$$

Let us record some properties of α-sizable families differentiating them from filters.

Proposition 3. (i) *Sizable families verify*

$$\text{(Half)} \quad \frac{\overline{A} \notin \mathscr{H} \quad A \cup \overline{B} \notin \mathscr{H}}{A \cup B \in \mathscr{H}}.$$

(ii) *For* $n \geqslant 1$, *let*

$$(\mathrm{Cov}_n) \quad \bigvee_{\delta:\{1,2,\ldots,n\} \to \{+,-\}} \left(\left(\bigcup_{i=1}^{n} A_i^{\delta(i)} \right) \in \mathscr{H} \right),$$

where $A_i^+ = A_i$ *and* $A_i^- = \overline{A}_i$. *Then* (Cov_n) *is valid for* α*-sizable families iff* $\alpha < 1 - 2^{-n}$.

Proof. Part (i) is trivial: $\mu(\overline{A}) \leqslant \frac{1}{2}$ and $\mu(A \cup \overline{B}) \leqslant \frac{1}{2}$ imply $\mu(A \cup B) = 1$, since

$$\begin{aligned} \mu(A \cup B) &= \mu(A - B) + \mu(B) \\ &= \mu(\overline{B}) - \mu(\overline{B} - A) + \mu(B) \\ &= 1 - \mu(\overline{B} - A), \end{aligned}$$

where $\mu(\overline{B} - A) = 0$ (because $\mu(\overline{A}) \leqslant \frac{1}{2}$ and $\mu(A \cup \overline{B}) \leqslant \frac{1}{2}$).

Part (ii) can be proved by the following clever argument I owe to N. Alechina (considerably simplifying my original proof). The set V can be partitioned into the family

$$\{A_1^{\delta(1)} \cap \cdots \cap A_n^{\delta(n)} | \delta : \{1,2,\ldots,n\} \to \{+,-\}\}$$

of 2^n disjoint pieces. Hence, for any probability function μ, there must be at least one function δ such that $\mu(A_1^{\delta(1)} \cap \cdots \cap A_n^{\delta(n)} \leqslant 2^{-n}$ – i.e., $\mu(A_1^{-\delta(1)} \cup \cdots \cup A_n^{-\delta(n)}) \geqslant 1 - 2^{-n}$ where $-+ = -$ and $-- = +$). $\square$

3.2. A generalization of filters

To avoid (Half) and (Cov_n), we quantify away α as follows. Call $\mathscr{H}$ *additive* if it is α-sizable for some α. That is, $\mathscr{H}$ is additive iff for some probability function $\mu : \mathscr{F} \to [0,1]$, and $\alpha \in [0,1]$, $\mathscr{H} = \{A \in \mathscr{F} : \mu(A) > \alpha\}$. Rewriting $(\mathrm{Up})_{\preccurlyeq}$ (from Section 1), with $\mu(A) \leqslant \mu(B)$ in place of $A \preccurlyeq B$, we get

$$(\mathrm{Up})_\mu \quad \frac{A \in \mathscr{H} \qquad \mu(A) \leqslant \mu(B)}{B \in \mathscr{H}}.$$

Note, however, that unless we know what μ is, we cannot assert $(\mathrm{Up})_\mu$. Could it be then that the best we could do to characterize additive families is to assert (Up)?

The following counterexample shows that there is more structure in additive families to account for.

(†) Let $V \supseteq \{1,2,3,4\}$ and $\mathscr{H} = \{A \subseteq V : \{1,3\} \subseteq A$ or $\{2,4\} \subseteq A\}$. Although $\mathscr{H}$ verifies (Up), it is not additive: were it induced by μ, then as $\{1,3\} \in \mathscr{H}$ and $\{1,2\} \notin \mathscr{H}$, $\mu(\{3\}) > \mu(\{2\})$; but as $\{2,4\} \in \mathscr{H}$ and $\{3,4\} \notin \mathscr{H}$, $\mu(\{2\}) > \mu(\{3\})$.

To strengthen (Up), some notation is useful saying (roughly) that a sequence $A_1, A_2, \ldots, A_n$ of sets in $\mathscr{F}$ is heavier than a sequence $B_1, B_2, \ldots, B_n$ (of equal length). With that in mind, let us write

$$\sum_{i=1}^{n} A_i \geqslant \sum_{i=1}^{n} B_i$$

to mean that for every $a \in V$,

$$\sum_{i=1}^{n} \chi^{A_i}(a) \geqslant \sum_{i=1}^{n} \chi^{B_i}(a),$$

where χ^A is the characteristic function of A (mapping elements of A to 1, and elements of $V - A$ to 0). Now, consider the condition

(σ) for all sequences $A_1, \ldots, A_n$ and $B_1, \ldots, B_n \in \mathscr{F}$,

$$\sum_{i=1}^{n} A_i \geqslant \sum_{i=1}^{n} B_i \quad \text{implies} \quad (\exists i \in \{1, 2, \ldots, n\})\ A_i \in \mathscr{H} \text{ or } B_i \notin \mathscr{H}.$$

(Up) is just the case $n = 1$ (as $A_1 \geqslant B_1$ just means $A_1 \supseteq B_1$). Example (†) above violates

$$\frac{A \cup C \notin \mathscr{H} \qquad B \cup (C - A) \in \mathscr{H} \qquad A \cup (D - B) \in \mathscr{H}}{B \cup D \in \mathscr{H}}$$

which follows from the case $n = 2$ of (σ).

Theorem 4. *Assume $\mathscr{F}$ is finite. Then a family $\mathscr{H} \subseteq \mathscr{F}$ is additive iff it validates (σ).*

Proof. Let $<_{\mathscr{H}}$ be the inequality $(\mathscr{F} - \mathscr{H}) \times \mathscr{H}$. $\mathscr{H}$ is additive iff for all $a \in V$, there are real numbers $x_a \geqslant 0$ such that whenever $A <_{\mathscr{H}} B$,

$$\sum_{a \in A} x_a < \sum_{b \in B} x_b.$$

Without loss of generality, let $\mathscr{F} = \mathsf{Pow}(V)$ and $V = \{1, 2, \ldots, \hat{n}\}$. Fix an enumeration $(\hat{A}_1, \hat{B}_1,), \ldots, (\hat{A}_{\hat{m}}, \hat{B}_{\hat{m}})$ of $(\mathsf{Pow}(V) - \mathscr{H}) \times \mathscr{H}$, and for every $i \in \{1, \ldots, \hat{m}\}$ and $j \in \{1, \ldots, \hat{n}\}$, define

$$\alpha_{ij} = \begin{cases} 1 & \text{if } j \in \hat{B}_i - \hat{A}_i, \\ -1 & \text{if } j \in \hat{A}_i - \hat{B}_i, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathscr{H}$ is additive iff

(P) there exist real numbers $x_1, \ldots, x_{\hat{n}} \geqslant 0$ such that

$$\sum_{j=1}^{\hat{n}} \alpha_{ij} x_j > 0 \quad \text{for every } i \in \{1, \ldots, \hat{m}\},$$

Let $\mathbf{A}$ be the $\hat{m}\times\hat{n}$ matrix (α_{ij}), and appeal to the following well-known fact from linear algebra (e.g., Strang [21, p. 333]):

$$(*)\quad \mathbf{Ax}\geqslant\mathbf{b}\ \text{ has a nonnegative solution }\ \mathbf{x}\ \text{ iff }\ \mathbf{yA}\geqslant\mathbf{0},\ \ \mathbf{yb}<\mathbf{0}\ \text{ has } no \text{ solution }\ \mathbf{y}\leqslant\mathbf{0}.$$

But $\mathbf{Ax} > \mathbf{0}$ has a solution iff $\mathbf{Ax}\geqslant\mathbf{1}$ has a solution (since there are only finitely many variables $x_1,\dots,x_{\hat{n}}$ constituting $\mathbf{x}$). So, setting $\mathbf{b}$ to $\mathbf{1}$ and multiplying the right hand side of $(*)$ by -1, we can associate with (P) the "dual" (i.e., its negation)

(D) there exist real numbers $y_1,\dots,y_{\hat{m}}\geqslant 0$, not all 0, such that

$$\sum_{i=1}^{\hat{m}} y_i\alpha_{ij}\leqslant 0\quad \text{for every } j\in\{1,\dots,\hat{n}\}.$$

The reals in (P) and (D) can be assumed to be rationals, and after clearing the denominators to make the y_i's in (D) positive integers, the negation of (D) can be expressed as

(σ') for all sequences $A_1,\dots,A_n$ and $B_1,\dots,B_n\in\mathscr{F}$,

$$(\forall i\in\{1,2,\dots,n\})\ A_i<_{\mathscr{H}}B_i\quad\text{implies}\quad\text{not}\ \sum_{i=1}^{n}A_i\geqslant\sum_{i=1}^{n}B_i$$

(with y_i the number of copies of the pair $\hat{A}_i,\hat{B}_i$ in the sequence $A_1,B_1,\dots,A_n,B_n$). Now, (σ') is just the contrapositive of (σ). □

Remarks. 1. Theorem 4 is reminiscent of a theorem due to Kraft et al. [11], characterizing binary relations on finite Boolean algebras that can be read "at least as probable as". That characterization is established again in Scott [19], employing methods similar to that used in the proof of Theorem 4 above, although I do not see how to reduce Theorem 4 to the Kraft et al. result, as the notion of an additive family quantifies away even more information than the comparative relation "at least as probable as". (Of course, it is easy enough to express the unary predicate "probable" as "at least as probable as A", for some fixed A; but our goal has been to isolate the bare minimum.) The "primal/dual" argument above appears to apply more generally than "the general method" (Theorem 1.1) of Scott [19].

2. Clearly, (σ) is satisfied by every filter $\mathscr{H}$: taking the contrapositive (σ') of (σ), suppose that for every $i\in\{1,2,\dots,n\}$, $A_i\notin\mathscr{H}$ and $B_i\in\mathscr{H}$; then $\bigcap_i B_i\in\mathscr{H}$, and for every $k\in\{1,2,\dots,n\}$, $\bigcap_i B_i\not\subseteq A_k$, whence $\sum_i A_i\not\geqslant\sum_i B_i$.

3. To force $\alpha\geqslant\frac{1}{2}$, it suffices to add a rule of non-empty intersection (considerably weaker than (And))

$$\text{(NEI)}\quad \frac{A\in\mathscr{H}\qquad B\in\mathscr{H}}{A\cap B\neq\emptyset}.$$

For $\alpha = m/n$, with $m < n$, add the rule that for any sequence $A_1, \ldots, A_n$ that partitions V, there is a subset I of $\{1, \ldots, n\}$ with cardinality m such that

$$\bigcup_{i \in I} A_i \in \mathscr{H}.$$

In yet another direction, the rule $(\mathrm{Cov})_n$, from Proposition 3(ii), adds the restriction that $\alpha < 1 - 2^{-n}$.

4. To see that each case $n-1$ of (σ) is insufficient, let $V = \{1, 2, \ldots, n^2\}$, and, recalling the definition of $<_{\mathscr{H}}$ in the proof above of Theorem 4, consider

$$\begin{aligned}
\{1,2,3,\ldots,n\} &<_{\mathscr{H}} \{1, n+1, 2n+1, \ldots, (n-1)n+1\},\\
\{n+1, n+2, n+3, \ldots, n+n\} &<_{\mathscr{H}} \{2, n+2, 2n+2, \ldots, (n-1)n+2\},\\
\{2n+1, 2n+2, 2n+3, \ldots, 2n+n\} &<_{\mathscr{H}} \{3, n+3, 2n+3, \ldots, (n-1)n+3\},\\
&\vdots\\
\{(n-1)n+1, (n-1)n+2, \ldots, n^2\} &<_{\mathscr{H}} \{n, n+n, 2n+n, \ldots, n^2\}.
\end{aligned}$$

(Note that if the left-hand sides are packaged as an $n \times n$ matrix $\mathbf{A}$, then the right-hand sides form the transpose $\mathbf{A}^{\mathrm{T}}$ of $\mathbf{A}$.)

4. Conclusion: directions from and applications to natural language

Although the approaches pursued in Sections 2 and 3 are compatible (insofar as (Up) and (And) imply (σ)), the thrust of Section 2 – viz., extending (And) – runs opposite that of Section 3, the canonical models of which, $\mathscr{A}_{pr,\alpha}$, fail to verify (And). Were we to adopt (And), then Section 3 would become uninteresting; were we to reject it, then Section 2 would become pointless. My reason for keeping the two sections in one paper is that I see no compelling argument that can (once and for all) settle the matter for or against (And). Intuitions about acceptability simply differ – intuitions expressed in languages, begging to be investigated. Among these are natural languages such as English, the semantics of which might be explored in the hope, for example, of shedding light on the acceptability of (And). The skeptical reader is bound to protest at such naiveté: natural languages are not formal; surely!

There is no denying that first-order logic is formal in a way that English is not. Even so, there is a growing body of work inspired by Montague's [15] slogan "English as a formal language", rephrased in more recent years as "Natural language as a programming language" (motivated by investigations into so-called "dynamic semantics", going back to, among other papers, Kamp [8]). More concretely, *generic* sentences, such as "Birds fly", provide a natural testing ground for questions about acceptability. And, in a slightly different direction, the very idea of formalizing, if not mechanizing, natural language leads to the problem of accepting exceptions. Briefly put, there is, I think, room for both Sections 2 and 3 in exploring the semantics of natural language.

4.1. Defeasibility in natural language interpretation

The transformation of "English as a formal language" into "natural language as a programming language" can be depicted by the equation

Natural language interpretation
= logic (e.g. model-theoretic semantics in Montague grammar)
+ control (e.g. choice of formula/disambiguation),

the point being that some notion of "control" is required to formalize a natural language utterance before it can be subjected to a logical analysis. Moreover, that logical analysis may, in turn, feed back into how to formalize the natural language utterance (and so on). This suggests a broader construal of "logic" (as applied to natural language) than that traditionally associated with model theory. For instance, the resulting logic should, many argue, allow for defeasibility, a particular preferential approach to which is pursued in Asher [2] and Lascarides and Asher [14], that verifies (And). Integrating this with statistical/probabilistic information is an obvious task crying out for attention. Stepping back a bit, basic logical issues raised by the problem of control are discussed in [4], where the challenge to compositionality posed by the context-dependence of disambiguation is taken up, and a logical approach based on generalized quantifiers is outlined.

4.2. A case study: Back to many (naturally)

Having derived much inspiration from Keisler [9], let us close by noting that a linguistically motivated analysis of *many* not only calls the filter properties (sound for *uncountably many*) into question, but also introduces (at least) two other dimensions to the classical picture in Keisler [9] (or, for that matter, in Westerståhl [23]): vagueness and intensionality. Passing from cardinalities to (vague subjective) probabilities over "possible worlds", Fernando and Kamp [5] propose to define there to be *many* φ-x's iff the number of φ-x's is more than what might be expected for every natural number $n \geqslant 1$,

$$(\exists_n x)(\varphi) \quad \supset \quad ((\mathsf{Many}\text{-}x)(\varphi) \equiv \mathsf{Prob}((\exists_{<n} x)(\varphi)))$$

where Prob is to be interpreted as an additive family $\mathscr{H}$. The preceding formulas are to be understood in an intensionalized form of weak logic within which the condition (σ) for additive families can be encoded in the manner of Segerberg [20]). (An extended account is in preparation with H. Kamp.)

Acknowledgements

I am very grateful to N. Alechina, H. Kamp, M. van Lambalgen and D. Westerståhl for helpful discussions, and to the two anonymous referees of this paper for a number of useful suggestions.

References

[1] N. Alechina, M. van Lambalgen, Generalized quantification as substructural logic, J. Symbolic Logic 61 (3) (1996).
[2] N. Asher, Reference to Abstract Objects in Discourse, Kluwer Academic Publishers, Dordrecht, 1993.
[3] J. Barwise, Admissible Sets and Structures, Springer, Berlin, 1975.
[4] T. Fernando, Ambiguity under changing contexts, Linguistics Philosophy 20(6) 1997.
[5] T. Fernando, H. Kamp, Expecting many, in: Proc. Semantics and Linguistic Theory VI, Cornell University, 1996.
[6] J.Y. Halpern, M.O. Rabin, A logic to reason about likelihood, Artif. Intel. 32 (1987).
[7] J.Y. Halpern, An analysis of first-order logics of probability, Artif. Intel. 46 (1990).
[8] H. Kamp, A theory of truth and semantic representation, in: J. Groenendijk, T. Janssen, M. Stokhof, (Eds.), Formal Methods in the Study of Language. Mathematical Centre Tracts 135, Amsterdam, 1981.
[9] H.J. Keisler, Logic with the quantifier "there exist uncountably many", Ann. Math. Logic 1 (1970).
[10] H.J. Keisler, Probability quantifiers, in: J. Barwise, S. Feferman (Eds.), Model-Theoretic Logics, Perspectives in Mathematical Logic, Springer, Berlin, 1985.
[11] C.H. Kraft, J.W. Pratt, A. Seidenberg, Intuitive probability on finite sets, Ann. Math. Statist. 30 (1959).
[12] S. Kraus, D. Lehmann, M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artif. Intell. 44 (1990).
[13] H.E. Kyburg, Jr., Conjunctivitis, in: M. Swain (Ed.), Induction, Acceptance, and Rational Belief, Reidel, Dordrecht, 1970.
[14] A. Lascarides, N. Asher, Temporal interpretation, discourse relations and commonsense entailment, Linguistics Philosophy 16(5) (1993).
[15] R. Montague, English as a formal language, in: R. Thomason (Ed.), Formal Philosophy, Yale University Press, New Haven, 1974.
[16] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Mateo, CA, 1988.
[17] J. Pearl, From Adams' conditionals to default expressions, causal conditionals and counterfactuals, in: E. Eells, B. Skyrms (Eds.), Probability and Conditionals, Cambridge Studies in Probability, Induction and Decision Theory, Cambridge University Press, Cambridge, 1994.
[18] G.E. Sacks, Saturated Model Theory, Benjamin, Reading, MA, 1972.
[19] D. Scott, Measurement structures and linear inequalities, J. Math. Psychol. 1 (1964) 233–247.
[20] K. Segerberg, Qualitative probability in a modal setting, in: J.E. Fenstad (Ed.), Proc. 2nd Scandinavian Logic Symp., North-Holland, Amsterdam, 1971, pp. 341–352.
[21] G. Strang, Linear Algebra and its Applications, Academic Press, Orlando, 1980.
[22] M. van Lambalgen, Natural deduction for generalized quantifiers, in: J. van der Does, J. van Eijck (Eds.), Generalized Quantifier Theory and Applications, Dutch Network for Language, Logic and Information, Amsterdam, 1991; CSLI Lecture Notes, to appear.
[23] D. Westerståhl, Quantifiers in formal and natural languages, in: D. Gabbay, F. Guenthner (Eds.), Handbook of Philosophical Logic, vol. IV, Reidel, Dordrecht, 1989.