10

In conjunction with qualitative probability

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 judgments 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.

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

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

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

$$\varphi \text{ is } \textit{acceptable} \quad \text{iff} \quad \begin{array}{l} \text{the "evidence" } \{a \in V : a \models \varphi\} \\ \text{supporting } \varphi \text{ "has enough weight"} \end{array}$$

or, in other words,

(1) $$\varphi \in \mathcal{A} \quad \text{iff} \quad \{a \in V : a \models \varphi\} \in \mathcal{H}$$

for some family $\mathcal{H} \subseteq \mathsf{Pow}(V)$ of "*h*eavy" subsets of V. Clearly, the safest choice is to equate $\mathcal{A}$ 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 $\mathcal{H} = \{V\}$. But suppose we took a *chance* on some other choice of $\mathcal{A}$ (and $\mathcal{H}$) that tolerates failures or exceptions. What are the rules for doing so *rationally*?

"Logic" suggests the closure condition

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

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 $\mathcal{A}$ (or $\mathcal{H}$) insofar as (Up) does not yield, from the assumption that $V \in \mathcal{H}$, any more elements of $\mathcal{H}$. On the other hand, (Up) does suggest how to proceed — namely, by weakening $\vdash$ to some binary relation $\mid\!\sim$ on $\mathcal{L}$ or (turning to $\mathcal{H}$) $\subseteq$ to some binary relation $\preceq$ on $\mathsf{Pow}(V)$

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

Although the correspondence between $\mathcal{A}$ and $\mathcal{H}$ described by (1) extends to pairs $\mathcal{A}$ and $\mathcal{H}$ closed respectively under $(\mathrm{Up})_{\mid\sim}$ and $(\mathrm{Up})_{\preceq}$, for some relations $\mid\!\sim$ and $\preceq$, the intuitions that most naturally come to play with the rule $(\mathrm{Up})_{\mid\sim}$ differ from those behind $(\mathrm{Up})_{\preceq}$ — the former being syntactic (or proof-theoretic), the latter, semantic. An example where this distinction matters concerns the operation $\wedge$ of conjunction on $\mathcal{L}$, 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 milion 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

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

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

But as long as we concentrate on the syntactic side $\mathcal{A}$, choosing the semantics $\mathcal{H}$ to support the manipulations on $\mathcal{A}$ that we have decided to legitimize, there is hope for (And). In this regard, it is worthwhile noting what Pearl 1994 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[1]).

[1]In the case of objective probabilities, it is useful to equate $\mathcal{L}$ with the set of *terms* of a formal language (ideally: typed), whereas in the case of subjective probabilities, it is

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, .999), given a probability function pr from $\mathcal{L}$ to the unit interval $[0, 1]$

$$\mathcal{A}_{pr,\alpha} \;=\; \{\varphi \in \mathcal{L} \;:\; pr(\varphi) > \alpha\}\ .$$

The hitch is that exceptions add up: from $pr(\varphi) = 1 - \delta$ and $pr(\psi) = 1 - \epsilon$, one cannot, in general, do better than predict $pr(\varphi \wedge \psi) \geq 1 - (\delta + \epsilon)$. This suggests 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 $\delta + \epsilon$.

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 $\epsilon\delta$-type limits (à la Bolzano-Weierstrass).

Pearl 1994 describes the work of Adams and Spohn in much this way, though with*out* 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 §§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 §3, to probabilities, beefing up the rule (Up) to a rule (Up)$_{\preceq}$, where $\preceq$ means "at least as probable." In this section, we refrain from making any commitment to the soundness or, for that matter, unsoundness of (And). (Up)$_{\preceq}$ applies not only to models of (And), but also to $\mathcal{A}_{pr,\alpha} = \{\varphi \in \mathcal{L} : pr(\varphi) > \alpha\}$, 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 (Up)$_{\preceq}$ 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 1988 puts the matter as follows

convenient to take $\mathcal{L}$ to be the language's set of formulas (V the language's collection of models, and $\models$ its satisfaction relation).

> 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 meaningul 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. [pages 493, 494]

A similar view can be found in Halpern and Rabin 1987, where "a logic to reason about likelihood" is developed relative to a *non*-probabilistic semantics. The probabilistic approach to likelihood presented in §3 (below) is not only qualitative but quite possibly simpler than might have been feared. It proceeds along lines similar to Scott 1964 and Segerberg 1971, 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 1994 and the references cited therein, plus Kyburg 1970 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 1970 (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 1991, and appears in a disguised form as Theorem 5 of Alechina and van Lambalgen 1996, 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 1985 and Halpern 1990, 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 judgments of acceptability; but by opening the door to alternative non-probabilistic interpretations of formulas, the challenge then becomes showing that only the

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

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 1970 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))

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

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).

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

$$\begin{array}{ll} (\mathrm{Q1}) & Qx\ x = x \\ (\mathrm{Q2}) & \forall x(\varphi \supset \psi)\ \supset\ (Qx\varphi \supset Qx\psi) \\ (\mathrm{Q3}) & Qx\varphi\ \wedge\ Qx\psi\ \supset\ Qx(\varphi \wedge \psi)\ . \end{array}$$

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

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

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 the following bit of bookkeeping due to van Lambalgen 1991. 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,

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

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). (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 1970), it suffices to show how to expand (M, q) to a model of Φ_0. The idea is to interpret R as the set

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

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

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

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

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

(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

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

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

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

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

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 bookkepping, 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.*

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

In fact, Proposition $1'$ can be sharpened to so-called *recursively saturated* models (Barwise 1975), as the sets of formulas that must be realized can be given effectively as follows

$$\{(\neg Qx\psi) \supset \neg\psi\} \quad \cup \quad \{(Qx\varphi) \supset \varphi \; : \varphi \in L_n(Q)\}$$

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,*

(3) $$Qx\varphi \quad \equiv \quad (\forall z)(\exists y \sqsubseteq z)(\forall x \sqsubset y)\ \varphi$$

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

Remarks.

1. (3) describes an $\epsilon\delta$-limit/asymptotic/cofinal-type quantification, with (2) falling out as the special case given by
$$u \sqsubset v \quad \equiv \quad R(u)$$
(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 $\overline{y}$ on $\sqsubset$ in (3).
4. It is considerably more complicated to describe an interpretation of $\sqsubset$ validating (3) in Theorem 2 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 1972. The twist in Theorem 2 is that the "ontological promiscuity" in saturation arguments is avoided by a purely combinatorial argument (with*out* 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

$$\begin{aligned} \Phi_+ &= \{\varphi \in \Phi \ : \ (M,q) \models Qx\varphi\} \\ \Phi_- &= \{\psi \in \Phi \ : \ (M,q) \not\models Qx\psi\} \ . \end{aligned}$$

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

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

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

$$\prec \quad = \quad \bigcup_{i\geq 0} \prec^i \quad .$$

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

$$\begin{aligned} \Phi^i_+ &= \Phi_+ \cap \{\varphi_j : j \leq i\} \\ \Phi^i_- &= \Phi_- \cap \{\varphi_j : j \leq i\} \\ E^i &= \{a \in dom(\prec^i) : (\forall b \prec^i a)\ a \prec^i b\} \qquad (= \text{ the set of } \prec^i\text{-minimal points}). \end{aligned}$$

(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 $\psi \in \Phi^i_-$,

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

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

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

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), (C2), (C3) and (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

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

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 appeal to (C1), (C2) and (C3) for an $a' \in E^i$ such that $a' \prec^i a$.

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

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

□ (Lemma A)

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

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

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

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

i.e.,

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

as required. □ (Lemma B)

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., $a \notin \bigcap q$). □ (Lemma C)

Next, let I^i be the image of $\prec^i$ minus E^i

$$I^i = \{a \in |M| - E^i : (\exists b \in |M|)\ b \prec^i a\}$$

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 (*on*to 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 $a \in N$,

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

(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

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

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

$$(M,q) \quad \models \quad ((\neg\psi) \wedge \bigwedge \Phi^i_+)[x/a]$$

and

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

Then set

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

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

Verifying conditions (C1) through (C6) is routine. $\square$

2.3 The binary case (from $\mathcal{A}$ 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

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

The weak completeness and compactness theorems of Keisler 1970 lift immediately to this setting (as worked out, for instance, in Westerståhl 1989). The filter schemes (Q1), (Q2) and (Q3) turn into

$$\begin{aligned}
&Qx(\varphi, \varphi) \\
&\forall x(\psi \supset \psi') \ \supset \ (Qx(\varphi, \psi) \supset Qx(\varphi, \psi')) \\
&Qx(\varphi, \psi) \wedge Qx(\varphi, \psi') \ \supset \ Qx(\varphi, \psi \wedge \psi')
\end{aligned}$$

respectively, with line (2) becoming

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

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 $\mathcal{H}$ (rather than on $\mathcal{A}$). Fix a non-empty set V, and a field $\mathcal{F}$ of subsets of V (containing V and closed under $\cup$ and $\bar{\cdot}$). Recall that a *(finitely additive) probability function on* $\mathcal{F}$ is a function $\mu : \mathcal{F} \rightarrow [0, 1]$ such that $\mu(V) = 1$, and for all $A, B \in \mathcal{F}$,

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

3.1 Weakening filters non-conservatively. A family $\mathcal{H} \subseteq \mathcal{F}$ of sets in $\mathcal{F}$ is *sizable* if there is a probability function μ on $\mathcal{F}$ such that for every $A \in \mathcal{F}$,

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

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

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

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

Proposition 3.

(i) *Sizable families verify*

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

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

$$(\mathrm{Cov}_n) \qquad \bigvee_{\delta:\{1,2,\ldots,n\}\to\{+,-\}} \Big((\bigcup_{i=1}^{n} A_i^{\delta(i)}) \in \mathcal{H} \Big)$$

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: $\mu(\overline{A}) \leq \frac{1}{2}$ and $\mu(A \cup \overline{B}) \leq \frac{1}{2}$ 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 $\mu(\overline{A}) \leq \frac{1}{2}$ and $\mu(A \cup \overline{B}) \leq \frac{1}{2}$).

Part (ii) requires a little more work, the point being to prove that for every $n \geq 1$, and sequence $A_1, \ldots, A_n \in \mathcal{F}$,

$$(5) \qquad \max_{\delta:\{1,2,\ldots,n\}\to\{+,-\}} \mu(\bigcup_{i=1}^{n} A_i^{\delta(i)}) \;\geq\; 1 - 2^{-n}$$

where equality can be arranged for suitable $A_1, \ldots, A_n$. To establish the equality in (5), we induct on n, with the induction hypothesis strengthened as follows (taking μ to be the uniform distribution over the power set of $V = \{1, 2, \ldots, 2^n\}$)

IH_n: there is a sequence $A_1, \ldots, A_n$ of subsets of $\{1, 2, \ldots, 2^n\}$ and an enumeration $\delta_1, \delta_2, \ldots, \delta_{2^n}$ of functions from $\{1, 2, \ldots, n\}$ to $\{+, -\}$ such that for each $j \in \{1, 2, \ldots, 2^n\}$,

$$\{1, 2, \ldots, 2^n\} - \bigcup_{i=1}^{n} A_i^{\delta_j(i)} \;=\; \{j\}\ .$$

The base case $n = 1$ is immediate. Next, assuming IH_n, define, for each $i \in \{1, 2, \ldots, n\}$,

$$B_i \;=\; A_i \;\cup\; \{2^n + j \;:\; j \in A_i\}$$

and $B_{n+1} = \{1, 2, \ldots, 2^n\}$. Then for each $j \in \{1, 2, \ldots, 2^n\}$,

$$\{1, 2, \ldots, 2^{n+1}\} - (\overline{B}_{n+1} \cup \bigcup_{i=1}^{n} B_i^{\delta_j(i)}) \;=\; \{j\}$$

$$\{1, 2, \ldots, 2^{n+1}\} - (B_{n+1} \cup \bigcup_{i=1}^{n} B_i^{\delta_j(i)}) \;=\; \{2^n + j\}$$

from which IH_{n+1} easily follows, and hence, the equality in (5). As for the inequality in (5), a function $\delta : \{1, 2, \ldots, n\} \to \{+, -\}$ verifying it (given any fixed sequence $A_1, \ldots, A_n$) can be constructed by choosing $\delta(1)$ so that

$\mu(A_1^{\delta(1)}) \geq \frac{1}{2}$, and then $\delta(2)$ so that $\mu(A_2^{\delta(2)} - A_1^{\delta(1)}) \geq \frac{3}{4} - \mu(A_1^{\delta(1)})$, and so forth, the general pattern being that for every $k \in \{1, 2, \ldots, n\}$, $\delta(k)$ satisfies

$$\mu(A_k^{\delta(k)} - \bigcup_{i=1}^{k-1} A_i^{\delta(i)}) \ \geq \ 1 - 2^{-k} - \mu(\bigcup_{i=1}^{k-1} A_i^{\delta(i)}) \ .$$

□

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

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

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 counter-example shows that there is more structure in additive families to account for.

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

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

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

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

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

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 $B_1, \ldots, B_n \in \mathcal{F}$,

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

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

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

which follows from the case $n = 2$.

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

Proof. $\mathcal{H}$ is additive iff the inequality $<_{\mathcal{H}}$ given by $(\mathcal{F} - \mathcal{H}) \times \mathcal{H}$ can be translated into a satisfiable system of linear inequalities as follows. Without loss of generality, let $\mathcal{F} = \mathsf{Pow}(V)$ 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 $(\mathsf{Pow}(V) - \mathcal{H}) \times \mathcal{H}$, and observe that $\mathcal{H}$ is additive iff

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

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

where for every $i \in \{1, \ldots, \hat{m}\}$ and $j \in \{1, \ldots, \hat{n}\}$,

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

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 1980, p. 333)

> Either $\mathbf{Ax} \geq \mathbf{b}$ has a nonnegative solution $\mathbf{x}$, or else there is a $\mathbf{y}$ such that $\mathbf{yA} \geq \mathbf{0}, \mathbf{yb} < \mathbf{0}$ and $\mathbf{y} \leq \mathbf{0}$.

Passing from $\geq$ to $>$, (P) has the "dual"

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

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

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, y_i can be taken to be the number of copies of the pair $\hat{A}_i, \hat{B}_i$ in

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

$$(\forall i \in \{1, 2, \ldots, n\}) \ A_i <_{\mathcal{H}} B_i \quad \text{implies}$$
$$\text{not } \sum_{i=1}^{n} A_i \geq \sum_{i=1}^{n} B_i \ .$$

Now, (σ') is just the contrapositive of (σ). □

Remarks.

1. Theorem 4 is reminiscent of a theorem due to Kraft, Pratt and Seidenberg 1959, characterizing binary relations on finite Boolean algebras that can be read "at least as probable as." That characterization is established again in Scott 1964, 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, Pratt, Seidenberg/Scott result, as the notion of an additive family quantifies away even more information than the comparative relation "at least as probable as." Also, the "primal/dual" argument above applies more generally than "the general method" (Theorem 1.1) of Scott 1964.
2. Clearly, (σ) is satisfied by every filter $\mathcal{H}$: taking the contrapositve (σ') of (σ), suppose that for every $i \in \{1, 2 \ldots, n\}$, $A_i \notin \mathcal{H}$ and $B_i \in \mathcal{H}$; then $\bigcap_i B_i \in \mathcal{H}$, and for every $k \in \{1, 2, \ldots, n\}$, $\bigcap_i B_i \not\subseteq A_k$, whence $\sum_i A_i \not\geq \sum_i B_i$.
3. To force $\alpha \geq \frac{1}{2}$, it suffices to add a rule of non-empty intersection (considerably weaker than (And))

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

For $\alpha = \frac{m}{n}$, 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

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

In yet another direction, the rule $(\mathrm{Cov})_n$, from Proposition 3(ii), adds the restriction that $\alpha < 2^{-n}$.
4. To see that each case $n - 1$ of (σ) is insufficient, let $V = \{1, 2, \ldots, n^2\}$, and, recalling the definition of $<_{\mathcal{H}}$ in the proof above of Theorem 4, consider

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

(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 $\mathbf{A}^T$ of $\mathbf{A}$.)

4 Conclusion: directions from and applications to natural language. Although the approaches pursued in §§2 and 3 are compatible (insofar as (Up) and (And) imply (σ)), the thrust of §2 — viz., extending (And) — runs opposite that of §3, the canonical models of which, $\mathcal{A}_{pr,\alpha}$, fail to verify (And). Were we to adopt (And), then §3 would become uninteresting; were we to reject it, then §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 R. Montague's slogan "English as a formal language" (1970/4), 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 1981). 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 §§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

$$\begin{aligned} \textit{Natural language interpretation} \;&= \\ \textit{logic}\ [\text{e.g. model-theoretic}& \\ \text{semantics in Montague grammar}]& \\ +\ \ \textit{control}\ [\text{e.g. choice of}& \\ \text{formula/disambiguation}]& \quad , \end{aligned}$$

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, verifying (And), is pursued in Asher 1993 and Lascarides and Asher 1993. 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 Fernando 1996, 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 1970, 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 1970 (or, for that matter, in Westerståhl 1989): vagueness and intensionality. Passing from cardinalities to (vague subjective) probabilities over "possible worlds", Fernando and Kamp 1996 propose to define there to be *many* φ-x's iff the number of φ-x's is more than what might be expected:

$$(\mathsf{Many}\text{-}x)(\varphi) \quad \text{iff} \quad \bigvee_{n\geq 1} \Big((\exists_{\geq n} x)(\varphi) \ \wedge \ \mathsf{Prob}((\exists_{<n} x)(\varphi)) \Big) \tag{6}$$

where Prob is to be interpreted as an additive family $\mathcal{H}$. A formal study of (6) within an intensionalized form of weak logic is currently being written up (with H. Kamp), drawing in no small measure on Theorem 4 above (which can be encoded within modal logic, in the manner of Segerberg 1971).

Acknowledgments. I am very grateful to N. Alechina, H. Kamp and M. van Lambalgen for helpful discussions.

References

N. Alechina and M. van Lambalgen. Generalized quantification as substructural logic. *Journal of Symbolic Logic*, 61(3), 1996.

Nicholas Asher. *Reference to Abstract Objects in Discourse.* Kluwer Academic Publishers, Dordrecht, 1993.

Jon Barwise. *Admissible Sets and Structures.* Springer-Verlag, Berlin, 1975.

T. Fernando and H. Kamp. Expecting many. In *Proc. Semantics and Linguistic Theory VI.* Cornell University, 1996.

Tim Fernando. Ambiguity under changing contexts. Manu-script, 1996. Preliminary version presented in *Mathematics of Language 4* (Philadelphia, October 1995), and submitted for publication.

J.Y. Halpern and M.O. Rabin. A logic to reason about likelihood. *Artificial Intelligence*, 32, 1987.

Joseph Y. Halpern. An analysis of first-order logics of probability. *Artificial Intelligence*, 46, 1990.

Hans Kamp. A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language.* Mathematical Centre Tracts 135, Amsterdam, 1981.

H. Jerome Keisler. Logic with the quantifier "there exist uncountably many". *Annals of Mathematical Logic*, 1, 1970.

H. Jerome Keisler. Probability quantifiers. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, Perspectives in Mathematical Logic. Springer-Verlag, 1985.

C.H. Kraft, J.W. Pratt, and A. Seidenberg. Intuitive probability on finite sets. *Ann. Math. Statist.*, 30, 1959.

Henry E. Kyburg, Jr. Conjunctivitis. In M. Swain, editor, *Induction, Acceptance, and Rational Belief.* D. Reidel, Dordrecht, 1970.

Michiel van Lambalgen. Natural deduction for generalized quantifiers. In J. van der Does and J. van Eijck, editors, *Generalized Quantifier Theory and Applications.* Dutch Network for Language, Logic and Information (Amsterdam), 1991. To appear in CSLI Lecture Notes.

A. Lascarides and N. Asher. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, 16(5), 1993.

Richard Montague. English as a formal language. In R. Thomason, editor, *Formal philosophy.* Yale University Press, New Haven, 1974.

Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, San Mateo, California, 1988.

Judea Pearl. From Adams' conditionals to default expressions, causal conditionals and counterfactuals. In E. Eells and B. Skyrms, editors, *Probability and Conditionals*, Cambridge Studies in Probability, Induction and Decision Theory. Cambridge University Press, 1994.

Gerald E. Sacks. *Saturated Model Theory.* W.A. Benjamin, Reading, Mass., 1972.

Dana Scott. Measurement structures and linear inequalities. *Journal of Mathematical Psychology*, 1:233–247, 1964.

Krister Segerberg. Qualitative probability in a modal setting. In J.E. Fenstad, editor, *Second Scandinavian Logic Symposium*, pages 341–352. North-Holland, 1971.

Gilbert Strang. *Linear Algebra and Its Applications.* Academic Press, Orlando, 1980.

Dag Westerståhl. Quantifiers in formal and natural languages. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume IV. D. Reidel, Dordrecht 1989.