1st International Symposium on Imprecise Probabilities and Their Applications, Ghent, Belgium, 29 June - 2 July 1999

A Logic of Extended Probability

Nic Wilson
School of Computing and Mathematical Sciences
Oxford Brookes University
Gipsy Lane, Headington, Oxford, OX3 0BP, U.K.
pnwilson@brookes.ac.uk

Abstract

This paper shows how the logic of gambles corresponding to Peter Walley's system of Imprecise Probability can be extended to allow gambles involving infinitesimal values and infinite values. This logic can then be used for reasoning with infinitesimal probabilities alongside conventional reasoning with linear constraints on probabilities. The proof theory is shown to be sound and complete for finite input sets.

Keywords. Imprecise Probability, Probabilistic Logic, infinitesimals, order of magnitude probabilistic reasoning.

1 Introduction

Very unlikely propositions can sometimes cause problems for theories of probability. The proposition may, for example, relate to an event that has not occurred before in our experience. If we assign a precise probability to such a proposition, then the value can be very arbitrary, and thus can lead to very arbitrary inferences, and arbitrary decisions. Even if a probability interval $[\delta, \epsilon]$ is assigned for such a proposition, related problems arise; the values δ and ϵ will tend to be also very arbitrary; if we're cautious, giving a wide interval, the inferences will often be disappointly weak; if on the other hand, we give a narrow interval, this suffers from similar problems as assigning a precise probability; an interval of intermediate width can suffer from either (or even both) problems. Similarly, utility values of very large magnitudes can be problematic.

An alternative is to consider such values of probability as essentially infinitesimal, and such values of utility as of infinite magnitude. An earlier paper [11] constructed a theory of probability which allowed infinitesimal values of probability as well as the usual values. As shown, in [9, 10, 11], this theory of probability can be used in a simple way in the interpretation and semantics of a number of theories of uncertain reasoning, for example, non-monotonic consequence relations [1, 4]; NCFs [7], possibility functions [3] and counterfactual probabilities [2].

In this paper, this work is developed into a logic that can reason with both conventional and infinitesimal probability; it is a unified framework that can be used for automated deduction with gambles (or linear constraints on probabilities), and also order of magnitude statements about probabilities. It thus unifies a probabilistic logic and a form of possibilistic reasoning.

Section 2 reiterates from [12] the expression of Imprecise Probability [8] as a logic. Section 3 describes the simple theory of infinitesimals given in [11]. Section 4 brings the two together, constructing a logic that can be used for reasoning with linear constraints on probabilities which can take both real and infinitesimal values; the proof of the main result is sketched in section 5. The proof theory is illustrated with an example in section 6. The logic can be also used for making a parametrised family of conventional probabilistic inferences as shown in section 7. Section 8 summarises the main results and discusses some avenues for further exploration.

2 The Logic of Gambles

This section gives the construction of the logic of gambles and basic properties of the logic, taken from [12] (the results are derived from results in [8].

Let Ω be a finite set of possibilities, exactly one of which must be true. A gamble on Ω is a function from Ω to $I\!R$. If you were to accept gamble X and it was ω that turned out to be true then you would gain $X(\omega)$ utiles (so you would lose if $X(\omega) < 0$). An agent's beliefs are elicited by asking them to tell us a set , of gambles they find acceptable, i.e., gambles they would be happy to accept.

For $\lambda \in I\!R$ we write λ for the constant gamble defined by $\lambda(\omega) = \lambda$ for all $\omega \in \Omega$. If $\lambda > 0$ we should certainly consider gamble λ acceptable since, whatever happens, we gain. If $\lambda < 0$ then we should certainly not accept gamble λ since, whatever happens, we lose. Addition and subtraction of gambles is defined pointwise, so that for gambles X and Y, for each $\omega \in \Omega$, $(X - Y)(\omega) = X(\omega) - Y(\omega)$. For $A \subseteq \Omega$ we define gamble A to be the indicator function of A, so that $A(\omega) = 1$ if $\omega \in A$ and $A(\omega) = 0$ otherwise. Any such gamble A would seem acceptable since you couldn't lose. On the other hand, the gamble $A - 1$ (given by $(A-1)(\omega) = 0$ if $\omega \in A$; $(A-1)(\omega) = -1$ otherwise) would only be acceptable to you if you were certain that A were true. If $\Omega = \{\omega_1, \omega_2\}$ and X is the gamble $\{\omega_1\} - 0.5$ (so that $X(\omega_1) = 0.5$ and $X(\omega_2) = -0.5$) and you considered ω_1 more likely than ω_2 then it seems that you should consider the gamble X acceptable.

Gambles can represent a varied set of probability statements. Another way of viewing gambles is as linear constraints on an unknown chance distribution, and indeed any such linear inequality can be represented as a gamble. For example, the constraint $\Pr(A) \geq 0.6$ would be represented by the gamble $A - 0.6$; the constraint $\Pr(A) \leq \Pr(B) + 0.1$ would be represented by the gamble $B + 0.1 - A$; the constraint $\Pr(A|B) \leq 0.9$ (which is taken to satisfied if $\Pr(B) = 0$) by the gamble $0.9B - (A \cap B)$. In general a linear constraint $\lambda_1 \Pr(A_1) + \cdots + \lambda_k \Pr(A_k) \geq \mu$ is represented by the gamble $X = \lambda_1 A_1 + \cdots + \lambda_k A_k - \mu$. The constraint $\lambda_1 \Pr(A_1) + \cdots + \lambda_k \Pr(A_k) \leq \mu$ is equivalent to the constraint $(-\lambda_1)\Pr(A_1) + \cdots + (-\lambda_k)\Pr(A_k) \geq -\mu$ and so is represented by the gamble $-X$. Hence $\lambda_1 \Pr(A_1) + \cdots + \lambda_k \Pr(A_k) = \mu$ is represented by the pair of gambles $\{X, -X\}$. The language does, however, have some limitations: independence relationships or constraints such as $\Pr(A|B) \leq \Pr(A)$ cannot be represented simply in terms of gambles, since they are non-linear.

The Language

Let $\mathcal{L}$ be the set of gambles on Ω. The constant gamble -1 is written as $\perp$.

Proof Theory

We are going to define an inference relation $\vdash$, where for $, \subseteq \mathcal{L}$ and $X \in \mathcal{L}$, $, \vdash X$ is intended to mean 'Given that I'm prepared to accept any gamble in , then I should also be prepared to accept X'. The axioms and inference rules used to define $\vdash$ will be justified by the semantics (with probability distributions as models) given below. However, they can be justified directly [8].

Axiom Schema: X, for any X with $\min_{\omega \in \Omega} X(\omega) \geq 0$.

For any such X we can't lose so it should be acceptable.

Inference Rule (Schema) 1: For any $\lambda \in I\!R$ such that $\lambda \geq 0$ the inference rule *From X deduce λX.*

This relates to the situation where stakes are multiplied by a factor λ.

Inference Rule 2: *From X and Y deduce $X + Y$.*

This relates to the combination of two gambles.

Then, in the usual way, we say that $, \subseteq \mathcal{L}$ proves $X \in \mathcal{L}$, abbreviated to $, \vdash X$, if there is a finite sequence $X_1, \ldots, X_n$ of gambles in $\mathcal{L}$ with $X_n = X$ and each X_i is either an element of , , an axiom or is produced by an inference rule from earlier elements in the sequence.

For $, \subseteq \mathcal{L}$ define TH(,) to be the set $\{X \in \mathcal{L} : , \vdash X\}$. The relation $\vdash$ is reflexive i.e., $\mathrm{TH}(,) \supseteq ,$; it is monotonic i.e, $, \subseteq \Theta \Rightarrow \mathrm{TH}(,) \subseteq \mathrm{TH}(\Theta)$; it is transitive, i.e, $\mathrm{TH}(\mathrm{TH}(,)) \subseteq \mathrm{TH}(,)$ (and therefore $\mathrm{TH}(\mathrm{TH}(,)) = \mathrm{TH}(,)$) and compact i.e, if $, \vdash X$ then there exists finite $,' \subseteq ,$ with $,' \vdash X$.

Semantics

The set of models $\mathcal{M}$ of $\mathcal{L}$ is defined to be the set of probability distributions on Ω. For probability distribution P and gamble X, $\mathrm{P}(X)$ is defined to be the expected value of X, i.e., $\sum_{\omega \in \Omega} \mathrm{P}(\omega) X(\omega)$. We say that probability distribution P satisfies gamble X (written $\mathrm{P} \models X$) iff $\mathrm{P}(X) \geq 0$, that is, if and only if the expected utility is non-negative, i.e., iff we would expect (in the long run) not to lose money from X if P represented objective chances. Note that P satisfies the constraint $\lambda_1 \mathrm{P}(a_1) + \cdots + \lambda_k \mathrm{P}(a_k) \geq \mu$ if and only if $\mathrm{P} \models X$ where $X = \lambda_1 a_1 + \cdots + \lambda_k a_k - \mu$ is the gamble representing the constraint.

We extend $\models$ in the usual way: P satisfies set of gambles , (written $\mathrm{P} \models ,$) if and only if it satisfies every gamble X in , , and the semantic entailment relation $\models$ is defined by $, \models X$ iff for all models P, $[\mathrm{P} \models , \Rightarrow \mathrm{P} \models X]$.

The relation $\models$ is reflexive, transitive, monotonic, but not compact.

The proof theory is sound with respect to the model theory:

Theorem 1 (Soundness) *For gamble X and set of gambles , , $, \vdash X \Rightarrow , \models X$.*

Completeness holds for finite sets of premises:

Theorem 2 (Finite Completeness) *For gamble X and finite set of gambles , , if , $\models X$ then , $\vdash X$.*

We come close to completeness even for infinite sets of premises:

Theorem 3 (Almost-Completeness) *For gamble X and set of gambles , , if , $\models X$ then , $\vdash X + \epsilon$ for all $\epsilon > 0$.*

Corollary (Consistency Completeness) *For , $\subseteq \mathcal{L}$, , $\vdash \perp \iff$, $\models \perp$.*

3 A Simple Theory of Infinitesimals

This section gives the construction of the simple theory of infinitesimals defined in [11]. The extended reals $I\!R^*$ are formed by adding an 'infinitesimal' element ε, to the real numbers. Probability is then defined using the usual axioms.

3.1 The Extended Reals $I\!R^*$

Let the extended reals $I\!R^*$ be $I\!R(\varepsilon)$, the field of rational functions in (dummy variable) ε over the field $I\!R$ (see [5], page 122). Each element of $I\!R^*$ can be written as a pair p/q where p and q are polynomial functions in ε, and p/q represents the same element of $I\!R^*$ as r/s if and only if ps is the same polynomial as qr. $I\!R^*$ clearly contains a copy of $I\!R$: for $x \in I\!R$, the ratio of constant polynomials $x/1$ is in $I\!R^*$, and we'll denote this element of $I\!R^*$ also by x. In particular the element 0 of $I\!R^*$ is the function which has constant value 0.

Every non-zero element r of $I\!R^*$ can be uniquely expressed as $\bar{r}\varepsilon^{\hat{r}}r'$, where $\bar{r} \in I\!R \setminus \{0\}$, $\hat{r}$ is in Z, the set of integers, and $r' \in I\!R^*$ is such that $r'(0) = 1$. Define $\hat{0} = \infty$. The function $r \mapsto \hat{r}$ gives the order of magnitude (in terms of powers of ε) of element r of $I\!R^*$. Also for $r \in I\!R^*$ let $\underline{r}$ be $\lim_{\varepsilon \to 0} r(\varepsilon)$, so that

$$\underline{r} = \begin{cases} 0 & \text{if } \hat{r} > 0; \\ \infty & \text{if } \hat{r} < 0 \text{ and } \bar{r} > 0; \\ -\infty & \text{if } \hat{r} < 0 \text{ and } \bar{r} < 0; \\ \bar{r} & \text{if } \hat{r} = 0. \end{cases}$$

If $r \in I\!R$ then $\underline{r} = r$. For $r \in I\!R^*$, $\underline{r}$ will be referred to as 'the real part of r'. Extended reals r and s will be said to be approximately equal if $\underline{r - s} = 0$.

3.2 The Ordering on $I\!R^*$

If $r = p/q \in I\!R^*$ where p, q are polynomials in ε, then, for $x \in I\!R$, $r(x)$ (the value of r when x is substituted for ε) is a real number, if $q(x) \neq 0$. For $r, s \in I\!R^*$, define relation $>$ by $r > s$ if and only if there exists strictly positive real number y such that $r(x) > s(x)$ for all real x with $0 < x < y$.

ε is intended to be a very small positive number, so we define $r > s$ iff r is bigger than s for small enough ε. Relations $<$, $\geq$ and $\leq$ are defined from relation $>$ in the usual way, e.g., $r \leq s$ if and only if $s > r$ or $s = r$. This ordering makes $I\!R^*$ an ordered field (see [5], page 261).

Though they are defined as functions, elements of $I\!R^*$ should be thought of as numbers; ε is a positive number smaller than any strictly positive real number, ε^2 is an even smaller positive number, $\varepsilon - \varepsilon^2$ is between the two, though much closer to ε, and so on.

The one axiom of the real numbers that $I\!R^*$ lacks is the completeness axiom, that any set bounded above has a least upper bound. However, this is the case for any ordered field extension of $I\!R$ containing an infinitesimal. Consider any ordered field extension R of the reals $I\!R$. Let $\mathcal{E}$ be the set of positive infinitesimals in R, i.e., all $r \in R$ such that $r > 0$ and for all strictly positive reals x, $r < x$. Assume $\mathcal{E} \neq \emptyset$ (i.e., that R contains an infinitesimal). $\mathcal{E}$ is bounded above by any strictly positive real. Suppose it has a least upper bound a. a must be in $\mathcal{E}$ because if for some strictly positive $x \in I\!R$, $x \leq a$ then $x/2 < a$ and $x/2$ is an upper bound for $\mathcal{E}$, which contradicts a being the least upper bound. However $a \in \mathcal{E}$ implies that $2a \in \mathcal{E}$, contradicting a being an upper bound for $\mathcal{E}$.

3.3 Extended Probability

To define extended probability and utility, the usual definitions suffice, except using $I\!R^*$ instead of $I\!R$.

An extended utility function (or extended gamble) on Ω is a function from Ω to $I\!R^*$. An extended probability measure P over Ω is defined to be a function from 2^Ω to $I\!R^*$, satisfying $\mathrm{P}(\Omega) = 1$, and for $A, B \subseteq \Omega$ such that $A \cap B = \emptyset$, $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$.

Extended probability theory has many of the same properties as probability theory, since it shares all the axioms, apart from the completeness axiom of the real numbers.

An interpretation of extended probability is given in [11]. Briefly, a proposition E with very small but unknown probability is assigned the value ε. The statement 'the probability of A is ε' is then interpreted as that A and E are (precisely) equiprobable. The meaning of statements of the form 'The probability of A is r', for various extended reals r, is constructed using a sequence of thought experiments using independent, mutually exclusive, or conditional propositions; for example, 'the probability of A is ε^2' is interpreted as meaning that A is equiprobable with $E \wedge F$ where F is some proposition independent of E and equiprobable with E.

4 A Logic of Extended Gambles

We can extend the language, proof theory and semantics of the logic of gambles to allow infinitesimals and infinities just by replacing $I\!R$ by $I\!R^*$ in the various definitions.

Hence the language $\mathcal{L}^*$ is defined to be the set of extended gambles, i.e., the set of functions from Ω to $I\!R^*$. As before, $\bot$ is an abbreviation for the constant extended gamble -1.

Proof theory

The proof theory is generated by the following axiom schema and inference rules:

Axiom Schema: X, for any X with $\min_{\omega \in \Omega} X(\omega) \geq 0$.

Inference Rule (Schema) 1: For any $\lambda \in I\!R^*$ such that $\lambda \geq 0$ the inference rule *From* X *deduce* λX.

Inference Rule 2: *From* X *and* Y *deduce* $X + Y$.

So, in the usual way, we say that , $\subseteq \mathcal{L}^*$ proves $X \in \mathcal{L}^*$, abbreviated to , $\vdash^* X$, if there is a finite sequence $X_1, \ldots, X_n$ of extended gambles in $\mathcal{L}^*$ with $X_n = X$ and each X_i is either an element of , , an axiom, or is produced by an inference rule from earlier elements in the sequence.

Like $\vdash$, the relation $\vdash^*$ is reflexive, monotonic, transitive and compact.

Semantics

The set of models $\mathcal{M}^*$ of $\mathcal{L}^*$ is defined to be the set of extended probability distributions on Ω, i.e., functions P from Ω to $I\!R^*$ such that for all $\omega \in \Omega$, $\mathrm{P}(\omega) \geq 0$ and $\sum_{\omega \in \Omega} \mathrm{P}(\omega) = 1$. For extended probability distribution P and extended gamble X, $\mathrm{P}(X)$ is defined, as before, to be the expected value of X, i.e., $\sum_{\omega \in \Omega} \mathrm{P}(\omega) X(\omega)$; we say that probability distribution P satisfies extended gamble X (written $\mathrm{P} \models^* X$) iff $\mathrm{P}(X) \geq 0$. The relation $\models^*$ is extended in the usual way: P satisfies set of extended gambles , , written $\mathrm{P} \models^*$, , if and only if it satisfies every extended gamble X in , , and the semantic entailment relation $\models^*$ is defined by , $\models^* X$ iff for all models P, $[\mathrm{P} \models^* , \Rightarrow \mathrm{P} \models^* X]$.

Set of extended gambles , is said to be consistent if it has a model, i.e., if there exists $\mathrm{P} \in \mathcal{M}^*$ such that $\mathrm{P} \models^*$, . It can be seen that , is inconsistent if and only if , $\models^* \bot$.

Like $\models$, the relation $\models^*$ is reflexive, transitive and monotonic, and not compact. A counter-example to compactness is when , $= \{(A - 1 + \varepsilon^n) : n \in I\!N\}$, where A is some non-empty proper subset of Ω. If $\mathrm{P} \models^*$, then for all n, $\mathrm{P}(A - 1 + \varepsilon^n) \geq 0$, so for all n, $\mathrm{P}(A) \geq 1 - \varepsilon^n$. This implies that $\mathrm{P}(A) = 1$, and so we have , $\models^* A - 1$. However, if ,$'$ is any finite subset of , , then it is not the case that ,$' \models^* A - 1$ (contradicting compactness)—since we can choose P such that $\mathrm{P}(A) = 1 - \varepsilon^n$ for some n large enough so that $\mathrm{P} \models^*$,$'$, but P isn't a model of $A - 1$, as $\mathrm{P}(A - 1) = -\varepsilon^n < 0$.

We still have Soundness:

Theorem 4 (Soundness of $\vdash^*$ w.r.t. $\models^*$) *For extended gamble* X *and set of extended gambles* , , , $\vdash^* X \Rightarrow$, $\models^* X$.

The proof is straight-forward: we just check the soundness of the axioms and inference rules; if Z is one of the axiom schema, then we certainly have $\mathrm{P} \models^* Z$ for any model P; if $\lambda \geq 0$ and $\mathrm{P} \models^* X$ and $\mathrm{P} \models^* Y$ then $\mathrm{P} \models^* \lambda X$ and $X + Y$. The result then follows using induction on the length of a proof in the logic.

As before, we cannot hope to have full completeness of $\vdash^*$ with respect to $\models^*$ because $\models^*$ is not compact. However we do have completeness for finite sets , . The following is the main result of the paper:

Theorem 5 (Finite Completeness of $\vdash^*$ w.r.t. $\models^*$) *For extended gamble* X *and finite set of extended gambles* , , *if* , $\models^* X$ *then* , $\vdash^* X$.

The proof is sketched in the next section.

Unfortunately we do not have *Almost-Completeness* (c.f. section 2) as the following example shows.

Counter-Example to Almost-Completeness

Let $I\!R^*_+$ be the set of strictly positive elements of $I\!R^*$, i.e., $\{r \in I\!R^* : r > 0\}$. We say that $\delta \in I\!R^*_+$ is a positive infinitesimal in $I\!R^*$ if $\delta < x$ for all $x \in I\!R$ with $x > 0$. Let $\mathcal{E}$ be the set of positive infinitesimals in $I\!R^*$.

Let A be some non-empty proper subset of Ω. Define the set of extended gambles ,$_1$ to be $\{A - \delta : \delta \in \mathcal{E}\}$ and ,$_2$ to be $\{r - A : r \in I\!R^*_+ - \mathcal{E}\}$.

Now , = ,$_1 \cup$,$_2$ is inconsistent, i.e., it has no model: for suppose $\mathrm{P} \models^*$, , and let $s = \mathrm{P}(A)$; then $\mathrm{P} \models^*$,$_1$ which implies $s \geq \delta$ for all $\delta \in \mathcal{E}$; also $\mathrm{P} \models^*$,$_2$, so $s \leq r$ for all $r \in I\!R^*_+ - \mathcal{E}$; but there is no value $s \in I\!R^*$ satisfying these constraints (the second constraint implies $s \in \mathcal{E}$; but then $s < 2s \in \mathcal{E}$ which contradicts the first constraint).

However, it can be shown that it is not the case that , $\vdash^* \bot$.

This example thus shows that it not the case that

$, \vdash^* \bot \iff , \models^* \bot$, so it is a counter-example to Consistency Completeness.

If Almost-Completeness held, then $, \models^* \bot$ would imply that $, \vdash^* \bot + \frac{1}{2}$; however, $, \vdash^* \bot + \frac{1}{2}$ implies (using inference rule 1) $, \vdash^* \bot$ which, as observed above, is not the case. Therefore Almost Completeness does not hold.

5 Sketch of Proof of Finite Completeness (Theorem 5)

A subset of $\mathcal{L}^*$ is said to be a *convex cone* (over $I\!R^*$) if it is closed under multiplication by non-negative scalars (in $I\!R^*$) and closed under addition, i.e., iff it is closed under the two inference rule schemas of the proof theory. For $, \subseteq \mathcal{L}^*$, define $C(,)$ to be the unique smallest convex cone containing , . Every element of $C(,)$ can then be expressed as a linear combination of the elements in , with non-negative co-efficients. A convex cone Δ is said to be *finite* if $\Delta = C(,)$ for some finite , .

For $T, X \in \mathcal{L}^*$ define TX to be $\sum_{\omega \in \Omega} T(\omega)X(\omega)$. For $, \subseteq \mathcal{L}^*$, define $,^+$ (known as the *dual cone* of ,) to be $\{T \in \mathcal{L}^* : TX \geq 0 \text{ for all } X \in , \}$.

Theorem 6 (Finite $I\!R^*$-convex cones are reflexive) *If Δ is a finite convex cone over $I\!R^*$ then it is reflexive, i.e.,* $(\Delta^+)^+ = \Delta$.

This is the analogue of a fundamental property of convex cones over $I\!R$, which requires a lengthy proof; the proof of this in [6] was used to prove Theorem 6: each step of the proof was examined (where necessary breaking it down into components) to confirm that the step is valid also for $I\!R^*$-convex cones (the key point to check was that the completeness property of $I\!R$ was not used, only the properties of an ordered field, which are shared by $I\!R^*$).

For $, \subseteq \mathcal{L}^*$, define $\mathrm{TH}(,)$ to be the syntactic consequences of , , i.e., $\{X : , \vdash^* X\}$, and define $\mathrm{Th}(,)$ to be the semantic consequences of , , i.e., $\{X : , \models^* X\}$. Theorem 4 (Soundess) is equivalent to the statement that for all , , $\mathrm{TH}(,) \subseteq \mathrm{Th}(,)$. Theorem 5 (Finite Completeness) is equivalent to: for all finite , , $\mathrm{Th}(,) \subseteq \mathrm{TH}(,)$.

Proposition 1

Let $\mathcal{R} = \{R \in \mathcal{L}^* : \text{ for all } \omega \in \Omega, R(\omega) \geq 0\}$. Let , be a subset of $\mathcal{L}^*$. Define $[,]$ to be $,^+ \cap \mathcal{R}$. For each $\omega \in \Omega$ define the indicator gamble I_ω by $I_\omega(\omega') = 0$ if $\omega \neq \omega'$, and $I_\omega(\omega) = 1$. Let $\mathcal{I} = \{I_\omega : \omega \in \Omega\}$.

(i) $\mathrm{Th}(,) = [,]^+$;

(ii) $[,] = [\mathrm{TH}(,)]$;

(iii) $\mathrm{TH}(,) = C(, \cup \mathcal{I})$;

(iv) $[C(, \cup \mathcal{I})] = \big(C(, \cup \mathcal{I})\big)^+$;

(v) $\mathrm{Th}(,) = \big((\mathrm{TH}(,))^+\big)^+$.

Sketch of proof: (i) follows easily using the observation that the models of , are normalised elements of $[,]$. (ii) is a version of Soundness and can be proved in the same way. (iii) is immediate from the definitions. (iv) holds because if $TX \geq 0$ for all $X \in \mathcal{I}$ then $T \in \mathcal{R}$. (v): applying, in turn, equalities (i), (ii), (iii), (iv) and (iii) gives $\mathrm{Th}(,) = [,]^+ = [\mathrm{TH}(,)]^+ = [C(, \cup \mathcal{I})]^+ = \big(\big(C(, \cup \mathcal{I})\big)^+\big)^+ = \big((\mathrm{TH}(,))^+\big)^+$.

Proof of Theorem 5: As , is finite, by Proposition 1(iii), $\mathrm{TH}(,)$ is a finite convex cone, so is reflexive by Theorem 6: $\mathrm{TH}(,) = \big((\mathrm{TH}(,))^+\big)^+$. Hence, by Proposition 1(v), $\mathrm{Th}(,) = \mathrm{TH}(,)$, proving Finite Completeness.

6 Example

The following is an example to illustrate the mechanics of the proof theory.

Suppose we are given the following constraints on extended probabilities:

$\mathrm{P}(A|B) \leq \varepsilon^3$, $\quad \mathrm{P}(B|C) \leq \varepsilon$, $\quad \mathrm{P}(A|A \cup \bar{C}) \geq 0.1$, and the logical piece of information $A \subseteq B \cap C$,

where $\bar{C}$ is the negation of the proposition C, and a constraint is taken to be satisified if the probability is undefined (owing to the denominator being zero).

What upper bounds can we give for $\mathrm{P}(A|C)$?

Firstly we must represent propositions A, B and C as subsets of a set of possibilities Ω. A natural way of doing this that doesn't assume anything more than is given in the problem is to set $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ and $A = \{\omega_1\}$, $B = \{\omega_1, \omega_2, \omega_3\}$, $C = \{\omega_1, \omega_2, \omega_4\}$. (Other representations of Ω, A, B and C are of course possible, but the choice doesn't affect derived probabilities. Although the choice of Ω etc. is not part of the logic, it could be automated.)

The following notation is used for gambles: $(9, 8, 7, 6, 5)$ is taken to be the gamble X with $X(\omega_1) = 9$, $X(\omega_2) = 8$ etc. The above inequalities must be converted into gambles. The constraint $\mathrm{P}(A|B) \leq \varepsilon^3$ is equivalent to $\varepsilon^3 \mathrm{P}(B) - \mathrm{P}(A \cap B) \geq 0$, and hence can be expressed as the gamble $X_1 = \varepsilon^3 B - A \cap B$, which may be written more explicitly as $(\varepsilon^3 - 1, \varepsilon^3, \varepsilon^3, 0, 0)$. Thus P satisfies the first constraint if and only if $\mathrm{P} \models^* X_1$.

Similarly, the second constraint can be represented by the gamble $X_2 = (\varepsilon - 1, \varepsilon - 1, 0, \varepsilon, 0)$, and the third constraint can be represented by $X_3 = (0.9, 0, -0.1, 0, -0.1)$. Let $\Gamma = \{X_1, X_2, X_3\}$.

From X_3 can be derived $Y_1 = (0.9, 0, -0.1, 0, 0)$: this is because $Y_1 - X_3$ is non-negative, so is one of the axioms in the axiom schema, and we can then derive using inference rule 2, $Y_1 = X_3 + (Y_1 - X_3)$. The first inference rule then enables us to derive $10Y_1$ and $\frac{1}{\varepsilon^3}X_1$, so we can derive (using inference rule 2) $Y_2 = 10Y_1 + \frac{1}{\varepsilon^3}X_1$, which equals $(10 - \frac{1}{\varepsilon^3}, 1, 0, 0, 0)$. We can add this to X_2 to derive $Y_3 = (\varepsilon + 9 - \frac{1}{\varepsilon^3}, \varepsilon, 0, \varepsilon, 0)$, and we're essentially there (since we've made $Y_3(\omega_3) = Y_3(\omega_5) = 0$ and $Y_3(\omega_2) = Y_3(\omega_4)$).

For any model P of Γ, by the soundness of the proof theory, $\mathrm{P} \models^* Y_3$, i.e., $(9 - \frac{1}{\varepsilon^3})\mathrm{P}(\omega_1) + \varepsilon(\mathrm{P}(\omega_1) + \mathrm{P}(\omega_2) + \mathrm{P}(\omega_4)) \geq 0$, so $(9 - \frac{1}{\varepsilon^3})\mathrm{P}(A \cap C) + \varepsilon \mathrm{P}(C) \geq 0$, which can be rearranged to give

$$\mathrm{P}(A|C) \leq \frac{\varepsilon^4}{1 - 9\varepsilon^3} \quad \text{(or } \mathrm{P}(C) = 0\text{).}$$

This upper bound is just slightly more than ε^4, and less than e.g., $1.001\varepsilon^4$.

In fact, this is best upper bound we can put on $\mathrm{P}(A|C)$ because it can be shown that there is a model P of Γ with $\mathrm{P}(A|C) = \frac{\varepsilon^4}{1-9\varepsilon^3}$.

7 Substituting Real Values for ε

Since the interpretation of extended probability given in [11] involves considering ε as a small positive unknown real number, one might wonder, if we take an inference $\Gamma \vdash^* X$, whether we can replace all occurrences of ε by a real number x and obtain a sound (conventional) probabilistic inference (as defined in section 2).

Let $X[\varepsilon/x]$ be the gamble obtained when extended gamble X has each occurrence of ε replaced by a positive real number x. Similarly, if Γ is a set of extended gambles, let $\Gamma[\varepsilon/x] = \{Y[\varepsilon/x] : Y \in \Gamma\}$. In fact $X[\varepsilon/x]$ may not be well-defined, since some component may have a zero denominator; for example if X is the constant extended gamble $\varepsilon/(1-\varepsilon)$ then $X[\varepsilon/1]$ is not well-defined. However, it is always the case that for all small enough x, $X[\varepsilon/x]$ is well-defined (because Ω is finite). Hence, if Γ is finite, for all small enough x, $\Gamma[\varepsilon/x]$ is a set of well-defined gambles.

Lemma 1 *Let Γ be a finite set of extended gambles, and suppose that $\Gamma \vdash^* X$ for some extended gamble X. Then there exists strictly positive $y \in \mathbb{R}$ such that for all $0 < x < y$, $\Gamma[\varepsilon/x] \vdash X[\varepsilon/x]$.*

This result also holds for infinite Γ, as long as there exists some strictly positive $y' \in \mathbb{R}$ such that for all $Y \in \Gamma$, $Y[\varepsilon/x]$ is well-defined for all $0 < x < y'$; this condition is required to ensure that for small enough x, $\Gamma[\varepsilon/x]$ is a well-defined set of gambles.

Sketch of proof: By definition of $\vdash^*$ there exists a sequence of extended gambles $X_1, X_2, \ldots, X_n$ with $X_n = X$, where each X_i is either an element of Γ, an axiom, or is produced from earlier elements in the sequence using one of the two inference rules. It can be shown inductively that, if x is small enough, $X_1[\varepsilon/x], X_2[\varepsilon/x], \ldots, X_n[\varepsilon/x]$ is a proof of $X[\varepsilon/x]$ from $\Gamma[\varepsilon/x]$ in the logic of gambles.

In fact, for any particular proof of X from Γ, by considering each step, one can generate (often easily) a range of values of x which lead (via the substitution) to sound conventional probabilistic inferences. The logic of extended gambles can thus be used for making a parametrised family of probabilistic inferences of a more usual kind.

This can be illustrated using the example of section 6, where it can be easily checked that each inference in the proof is valid if ε is substituted for any strictly positive value, and so the whole proof is valid for any such substitution; however, the final rearrangement is only valid if $1 - 9\varepsilon^3 > 0$, i.e., if $\varepsilon < 1/\sqrt[3]{9}$.

Lemma 2 *Let $\Gamma \cup \{X\}$ be a finite set of extended gambles. Suppose there exists strictly positive $y \in \mathbb{R}$ such that for all $0 < x < y$, $\Gamma[\varepsilon/x] \models X[\varepsilon/x]$. Then $\Gamma \models^* X$.*

Note that this result does not hold for infinite Γ. Consider, for example, $\Gamma = \{(1 - n\varepsilon) : n = 1, 2, 3, \ldots\}$, and any X which is not an axiom (e.g., $X = \bot$). For any $x > 0$, $\Gamma[\varepsilon/x]$ has no model, so $\Gamma[\varepsilon/x] \models X[\varepsilon/x]$ (when $X[\varepsilon/x]$ is well-defined). However, Γ is satisfied by *every* model, so it is not the case that $\Gamma \models^* X$.

Sketch of proof of Lemma 2: Suppose $\mathrm{P} \in \mathcal{M}^*$ is such that $\mathrm{P} \models^* \Gamma$. We must show that $\mathrm{P} \models^* X$. It can be shown that for all sufficiently small x, $\mathrm{P}[\varepsilon/x] \models \Gamma[\varepsilon/x]$; so then, by the hypothesis, for sufficiently small x, $\mathrm{P}[\varepsilon/x] \models X[\varepsilon/x]$, i.e., $\sum_{\omega \in \Omega} \mathrm{P}[\varepsilon/x](\omega) X[\varepsilon/x](\omega) \geq 0$. Hence $\left(\sum_{\omega \in \Omega} \mathrm{P}(\omega) X(\omega)\right)[\varepsilon/x] \geq 0$ for all sufficiently small x, which implies, using a basic property of the ordering $\geq$, that $\sum_{\omega \in \Omega} \mathrm{P}(\omega) X(\omega) \geq 0$, i.e., $\mathrm{P} \models^* X$, as required.

Theorem 7 *Let $\Gamma \cup \{X\}$ be a finite set of extended gambles. The following statements are equivalent:*

(i) $\Gamma \vdash^* X$;

(ii) $\Gamma \models^* X$;

(iii) there exists strictly positive $y \in \mathbb{R}$ such that for all $0 < x < y$, $\Gamma[\varepsilon/x] \vdash X[\varepsilon/x]$;

(iv) there exists strictly positive $y \in \mathbb{R}$ *such that for all* $0 < x < y$, , $[\varepsilon/x] \models X[\varepsilon/x]$.

Proof: by Lemma 1, (i) $\Rightarrow$ (iii); by Theorem 1, (iii) $\Rightarrow$ (iv); Lemma 2 gives (iv) $\Rightarrow$ (ii); finally, by Theorem 5, (ii) $\Rightarrow$ (i).

Theorem 7 means that one can give an alternative interpretation of this logic of extended gambles: X follows from finite , if and only if for all sufficiently small x, one can infer (in a more conventional probabilistic way) $X[\varepsilon/x]$ from , $[\varepsilon/x]$.

8 Summary and Discussion

A logic has been constructed which can reason with both conventional and infinitesimal probabilities. Surprisingly, perhaps, for such an expressive logic, it has a very simple proof theory which is complete for finite input sets.

The inferences made by this proof theory are guaranteed to be correct (in the sense of conventional probabilistic reasoning) if we replace the parameter ε by real-valued positive x, so long as x is small enough.

Use of the proof theory, however, may well often not be the most efficient method for deduction; it would be interesting to see to what extent linear programming techniques, for example, could be adapted for this logic.

It is clear that this logic doesn't completely solve the problem of very unlikely propositions; for example, if we want to represent two very small values, we have to precisely state their ratio. A natural extension of this work is therefore to add several incomparable infinitesimals to the reals as briefly discussed in [10]. Also, assuming that a probability of a proposition is arbitrarily small can lead to undesirable inferences: that you should bet against such a proposition at any finite odds. However, propagating an interval $(0, y)$ of valid real substitutions for ε (as discussed in section 7) will partly solve this problem by clarifying how small we're assuming ε to be for any given inference.

A technical issue is whether it is possible to bring proof ($\vdash^*$) and truth ($\models^*$) closer together. However we can't strengthen $\vdash^*$ by adding extra inference rules: a consequence of Finite Completeness is that any sound inference rule is a derived inference rule of $\vdash^*$, so any additional inference rules (or axioms) are either unsound or redundant. An alternative idea is to weaken $\models^*$ by increasing the set of models; a natural way to do so is to increase the range of the probability distributions: i.e., to allow probability values in some larger set $\mathbb{R}^{**}$ (without increasing the language $\mathcal{L}^*$). Possible choices for $\mathbb{R}^{**}$ include rational functions in $\sqrt{\varepsilon}$ or even power series in $\sqrt{\varepsilon}$. Soundness is maintained even though the associated semantic entailment relations are weakened. However, it is not yet clear whether or not Almost-Completeness can be achieved in such a manner.

The close relationship (see e.g., [9]) between order of magnitude probabilities and possibility theory [3] means that this logic can be used as a possibilistic logic. Upper bounds on the possibility of A can be translated to sets of gambles of the form $\{A - N\varepsilon^n : N = 1, 2, \ldots\}$, for a particular value of n; lower bounds get translated to sets of gambles of the form $\{\frac{1}{N}\varepsilon^n - A : N = 1, 2, \ldots\}$. Because these are infinite sets, we lose the completeness of the proof theory. However, a trick can be used to represent each of these infinite sets as a single statement, but this complicates the logic and it is not clear that completeness is restored.

Acknowledgements

Thanks to Peter Walley and the referees for their helpful comments.

References

[1] E. Adams. Probability and the Logic of Conditionals. In *Aspects of Inductive Logic*, J. Hintikka and P. Suppes (eds.). North Holland, Amsterdam, 265–316, 1966.

[2] C. Boutilier. The Probability of Possibility: Adding Uncertainty to Default Rules. *Proceedings of the Ninth Conference of Uncertainty in Artificial Intelligence (UAI93)*, David Heckerman and Abe Mamdani (eds.). Morgan Kaufmann Publishers, San Mateo, California, 461–468, 1993.

[3] D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing and Uncertainty*. Plenum Press, New York, 1988.

[4] S. Kraus, D. Lehmann and M. Magidor. Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence* 44: 167–207, 1990.

[5] S. Maclane and G. Birkhoff. *Algebra* (second edition). Macmillan, New York, 1979.

[6] D. E. Nering. *Linear Algebra and Matrix Theory* (second edition). John Wiley and Sons, 1970.

[7] W. Spohn. A General Non-Probabilistic Theory of Inductive Reasoning. *Uncertainty in Artificial*

Intelligence 4, R. D. Shachter et al. (eds.). Elsevier Science Publishers, 1990.

[8] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[9] N. Wilson. An Order of Magnitude Calculus. *Proceedings of the Eleventh Conference of Uncertainty in Artificial Intelligence (UAI95)*, Montreal, Canada, August 95. Morgan Kaufmann Publishers, San Mateo, California, 1995.

[10] N. Wilson. *Extended Probability*. Unpublished report, available from the author on request, 1995.

[11] N. Wilson. Extended Probability. *Proceedings of the 12th European Conference on Artificial Intelligence (ECAI96)*. John Wiley and Sons, 667-671, 1996.

[12] N. Wilson and S. Moral. A Logical View of Probability. *Proceedings of the 11th European Conference on Artificial Intelligence, (ECAI-94)*, Amsterdam, The Netherlands, August 8–12, 386–390, 1994.