# Basing Probabilistic Logic on Gambles

**Peter R. Gillett**
Rutgers Business School
—Newark and New Brunswick
gillett@rbsmail.rutgers.edu

**Richard B. Scherl**
Monmouth University
rscherl@monmouth.edu

**Glenn Shafer**
Rutgers Business School
—Newark and New Brunswick
gshafer@andromeda.rutgers.edu

## Abstract

This article presents a probabilistic logic whose sentences can be interpreted as asserting the acceptability of gambles described in terms of an underlying logic. This probabilistic logic has a concrete syntax and a complete inference procedure, and it handles conditional as well as unconditional probabilities. It synthesizes Nilsson's probabilistic logic and Frisch and Haddawy's anytime inference procedure with Wilson and Moral's logic of gambles.

Two distinct semantics can be used for our probabilistic logic: (1) the measure-theoretic semantics used by the prior logics already mentioned and also by the more expressive logic of Fagin, Halpern, and Meggido and (2) a behavioral semantics. Under the measure-theoretic semantics, sentences of our probabilistic logic are interpreted as assertions about a probability distribution over interpretations of the underlying logic. Under the behavioral semantics, these sentences are interpreted only as asserting the acceptability of gambles, and this suggests different directions for generalization.

**Keywords.** Probabilistic Logic, Anytime Deduction, Gambles, Measure-theoretic, Behavioral Semantics.

## 1 Introduction

This article presents a probabilistic logic whose sentences can be interpreted as asserting the acceptability of gambles. This logic has a concrete syntax and a complete inference procedure, and it handles conditional as well as unconditional probabilities. It synthesizes the probabilistic logic of Nils J. Nilsson [7] and the anytime inference procedure of Alan M. Frisch and Peter Haddawy [3] with the logic of gambles of Nic Wilson and Serafín Moral [13].

Nilsson and Frisch and Haddawy build their probabilistic logics, which we designate by $\mathcal{L}_N$ and $\mathcal{L}_{FH}$, respectively, on top of an underlying logic. According to their semantics, which we call the *measure-theoretic semantics* for probabilistic logic, each sentence says something about the probability of a sentence in the underlying logic. Our probabilistic logic, which we designate by $\mathcal{L}$, also has an underlying logic, and can use measure-theoretic semantics. When it does use this semantics, it is a strict generalization of Nilsson's and Frisch and Haddawy's logics: a sentence in $\mathcal{L}_N$ or $\mathcal{L}_{FH}$ translates into $\mathcal{L}$ with no change in meaning. $\mathcal{L}$ also contains more complex sentences. Instead of merely saying something about the probability of an individual sentence of the underlying logic, a sentence in $\mathcal{L}$ may say something about the expected value of a gamble whose payoff depends on the truth values of several sentences in the underlying logic.

Moreover, whereas Nilsson only discusses how to reason with models, and Frisch and Haddawy do not demonstrate completeness for their set of inference rules, we give a complete set of inference rules for $\mathcal{L}$.

In generalizing from probabilities to expected values, we are following Wilson and Moral, and our demonstration of the completeness of our logic uses the completeness of their logic, which we designate by $\mathcal{L}_{WM}$. We go well beyond their results, however, because we handle conditional as well as unconditional probabilities and we insist on a concrete syntax.

There is a large body of other important related work on probabilistic logic that is not mentioned in this paper ([1] and [4] include reviews of this literature). Fagin, Halpern and Meggido [2], have formulated probabilistic logics that use measure-theoretic semantics, and have complete inference procedures. Fagin, Halpern, and Meggido consider only the case where the underlying logic is propositional logic, but in this case, their probabilistic logics are more expressive than ours. In the case where our probabilistic logic $\mathcal{L}$ uses measure-theoretic semantics and uses propositional logic as its underlying logic, it can be regarded as a relatively small fragment of one of Fagin, Halpern, and Meggido's logics, but it is still of some

interest, because it enables complete inference about relatively elementary probability statements (including those considered by Nilsson and by Frisch and Haddawy) without the greater complexity of Fagin, Halpern, and Meggido's logics. In [5], Halpern and Pucella consider upper probability measures, and in [6] they add reasoning about expectation.

What we find most interesting about $\mathcal{L}$, however, is an alternative semantics that suggests paths for generalization different from the paths followed by other authors. This alternative semantics, which we call *behavioral semantics*, hews more closely to the notion of acceptability and can be formalized using the concept of lower prevision developed by Peter Walley [12]. In measure-theoretic semantics, an interpretation for a probabilistic logic is a probability distribution over interpretations for the underlying logic. A gamble is acceptable if this probability distribution gives the gamble nonnegative expected value. But this implies that if a gamble is not acceptable, then the opposite gamble (the gamble with the signs of all the payoffs reversed) is acceptable. No such implication is inherent in the notion of a gamble being acceptable, and no such implication is built into our probabilistic logic (the syntax considered in this article does not even provide for negation of acceptability statements). Behavioral semantics avoids these implications and therefore can be extended to martingale trees [8] and probability games [11]; see [10].

## 1.1 Comparative Summary

Table 1 summarizes how our logic compares with the others we have been discussing. As the table indicates, our approach is a synthesis, which combines the best features of the other approaches.

|  | $\mathcal{L}_N$ | $\mathcal{L}_{FH}$ | $\mathcal{L}_{WM}$ | $\mathcal{L}$ |
|---|---|---|---|---|
| Complete inference procedure | Yes | No | Yes | Yes |
| Anytime inference procedure | No | Yes | Yes | Yes |
| Elaboration tolerant | No | Yes | No | Yes |
| Concrete syntax | No | Yes | No | Yes |
| Handles conditional probabilities | No | Yes | No | Yes |
| Hyperplane expressive | No | No | Yes | Yes |

Table 1: Comparison of Related Work

When we say that Nilsson's system does not have an anytime inference procedure, we mean that the linear program must be run to completion. In contrast our language, like that of Frisch and Haddaway, is modular, and interim inferences are valid even though they may not have computed the tightest possible bounds.

Another advantage of modularity is elaboration tolerance: additional premises can be introduced and additional questions can be asked without discarding or repeating work already done. Our language, like Frisch and Haddaway's, is elaboration tolerant in this sense. Notice, however, that we have labeled Wilson and Moral's language as elaboration intolerant, even though it seems to have an anytime inference procedure. This is because its inference procedure takes for granted that the sample space has already been set up. When we introduce a new sentence, whether as a premise or a goal, old possibilities may split, according to whether the new sentence is true or false. So no system that takes the sample space for granted is elaboration tolerant.

Our assertion that Wilson and Moral do not have a concrete syntax refers to the fact that they do not specify any particular symbolic representation for their gambles. They specify syntax neither for their probabilistic logic nor for an underlying logic $\mathcal{L}_0$. Nilsson, in contrast, does insist on a concrete syntax for the underlying logic $\mathcal{L}_0$, although he does not specify a syntax for his probabilistic logic.

When we say that Wilson and Moral do not handle conditional probabilities, we mean only that they do not do so explicitly. A bound on a conditional probability can easily be re-expressed as a statement of the type they do handle.

When we say that our logic and that of Wilson and Moral are hyperplane expressive, we are referring to the fact that these logics can express an arbitrary linear constraint on a vector of probabilities $\boldsymbol{P}$, which requires $\boldsymbol{P}$ to lie on one side of a hyperplane. As we have noted, such linear constraints can be much more general than bounds on individual probabilities and conditional probabilities.

As we will see, each of the other three logics can be seen as a simplification of the probabilistic logic $\mathcal{L}$ developed in this article. $\mathcal{L}$ is not, however, the most expressive probabilistic logic possible. It can bound the vector of probabilities $\boldsymbol{P}$ by hyperplanes, and hence it can express the statement that $\boldsymbol{P}$ is in a given simplex, but it cannot express more complicated restrictions on $\boldsymbol{P}$. For example, it cannot express statements about the probability $p$ of a single sentence in the underlying logic such as "$0.3 \leq p \leq 0.5$ or $0.5 \leq p \leq 0.7$." Probabilistic logics that can express such statements include those of Fagin, Halpern, and Megiddo [2], which make use of rich inferential machinery including all instances of propositional tautologies, modus ponens, all instances of valid formulae about linear inequalities, and four axioms for probability. However, their probability axioms are derivable within our logic $\mathcal{L}$.

## 2  Syntax and Inference for $\mathcal{L}$

We designate our probabilistic logic by $\mathcal{L}$. In this section, we describe $\mathcal{L}$'s syntax and inference procedure. This description involves some informal explanation of $\mathcal{L}$'s semantics. In the next two sections, we formalize the semantics in two different ways and demonstrate the soundness and completeness of the inference procedure with respect to both formalizations.

The sentences of $\mathcal{L}$ are the form

$$\langle (\alpha_1, a_1) \ldots (\alpha_n, a_n) \mid \delta \rangle, \tag{1}$$

where $n$ is a nonnegative integer, $\alpha_1, \ldots, \alpha_n$ and $\delta$ are sentences of an underlying logic $\mathcal{L}_0$, and $a_1, \ldots, a_n$ are real numbers. The list $(\alpha_1, a_1) \ldots (\alpha_n, a_n)$ represents a gamble, which pays the sum of those $a_i$ for which the corresponding sentence $\alpha_i$ turns out to be true. Sentence (1) means that this gamble is acceptable to an agent when his knowledge relative to the sentences in $\mathcal{L}_0$ consists of knowledge that $\delta$ is true. In the next two sections, we make this idea into a formal semantics in two different ways. In §3, we formalize it as the condition that the payoff of the gamble has nonnegative expected value conditional on $\delta$; because the expected value has to be computed relative to some probability distribution, this constitutes a use of what we have already called *measure-theoretic semantics* for probabilistic logic. In §4, we formalize it in terms of our *behavioral semantics*, in which probability distributions are replaced by lower previsions.

If the sentence $\delta$ in $\mathcal{L}_0$ is a tautology, then the sentence (1) in $\mathcal{L}$ means that the gamble $(\alpha_1, a_1) \ldots (\alpha_n, a_n)$ is acceptable to the agent *a priori*. This is the special case of unconditional acceptability of gambles, considered by Wilson and Moral. But even here we differ from Wilson and Moral by representing gambles in terms of sentences of an underlying logic rather than merely as functions on a sample space.

### 2.1  The Underlying Logic $\mathcal{L}_0$

We assume that the underlying logic $\mathcal{L}_0$ is two-valued and uses the symbols $\neg$, $\wedge$, $\vee$, $\bot$, and $\top$ in the usual way together with propositional symbols $p$, $p'$, $p''$, etc. We use $\Rightarrow_0$ and $\Leftrightarrow_0$, respectively, for derivability and logical equivalence: $\alpha_1 \Rightarrow_0 \alpha_2$ means that $\alpha_2$ can be derived from $\alpha_1$ (i.e., $\alpha_1 \vdash_{\mathcal{L}_0} \alpha_2$), and $\alpha_1 \Leftrightarrow_0 \alpha_2$ means that either can be derived from the other (i.e., $\alpha_1 \vdash_{\mathcal{L}_0} \alpha_2$ and $\alpha_2 \vdash_{\mathcal{L}_0} \alpha_1$).

We assume that $\mathcal{L}_0$ has a sound and complete inference procedure, so that $\alpha_1 \Leftrightarrow_0 \alpha_2$ holds whenever the two are semantically equivalent. We assume that $\mathcal{L}_0$'s inference procedure is complete only because we need

this assumption in order to show completeness for $\mathcal{L}$. It is not needed in order for $\mathcal{L}$'s inference procedure to be well-defined and sound.

We make no further assumptions about $\mathcal{L}_0$, but reasoning within $\mathcal{L}_0$ is part of reasoning within $\mathcal{L}$ and so details about $\mathcal{L}_0$ are relevant to implementation. Unlike Fagin, Halpern, and Megiddo's logics, however, propositional reasoning takes place only within $\mathcal{L}_0$.

Let us write $wff_0$ for the set consisting of all sentences of $\mathcal{L}_0$. Given a truth assignment $M$ for $\mathcal{L}_0$, let us designate by $\omega_M$ the interpretation it determines—this is a mapping from $wff_0$ to $\{true, false\}$. And let us write $\Omega_0$ for the set consisting of all such interpretations:

$$\Omega_0 := \{\omega_M \mid M \text{ is a truth assignment for } \mathcal{L}_0\}.$$

We call $\Omega_0$ the *sample space for $\mathcal{L}_0$*. This concept should be contrasted with the notion of the sample space for a finite set of sentences in $\mathcal{L}_0$, used in Nilsson's work. Whereas we might explicitly construct the sample space for a few hundred sentences, there is no reasonable sense in which we can explicitly construct $\Omega_0$. If $\mathcal{L}_0$ is undecidable, explicit construction of $\Omega_0$ is not even theoretically possible. But as a theoretical (rather than a computational) object, $\Omega_0$ will be very useful in our mathematical reasoning about $\mathcal{L}$.

We call any real-valued function on $\Omega_0$ a *variable*. Given a variable $X$ on $\Omega_0$ and a subset $A \subseteq \Omega_0$, we define a variable $X^A$ by

$$X^A(\omega) := \begin{cases} X(\omega) & \text{if } \omega \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

We call $X^A$ the *restriction* of $X$ to $A$. We can write $X^A = X \cdot I_A$, where $I_A$ is $A$'s *indicator variable*:

$$I_A(\omega) := \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

This makes it clear that $(X_1 + X_2)^A = X_1^A + X_2^A$, $(rX)^A = rX^A$, and $(X^A)^B = X^{A \cap B}$.

Given $\alpha \in wff_0$, let $[\alpha]$ be the subset of $\Omega_0$ consisting of those truth assignments that assign $\alpha$ the value *true*:

$$[\alpha] := \{\omega \in \Omega_0 \mid \omega(\alpha) = true\}.$$

The set

$$\mathcal{A}_0 := \{[\alpha] \mid \alpha \in wff_0\}$$

is a field of subsets of $\Omega_0$; it is isomorphic to $\mathcal{L}_0$'s Lindenbaum-Tarski field. We call a finitely additive probability measure on the field $\mathcal{A}_0$ a *probability distribution* on $\Omega_0$.

## 2.2 Gambles in $\mathcal{L}_0$

We call an ordered pair $(\alpha, a)$, where $\alpha \in \text{wff}_0$ and $a$ is a real number, a *ticket*. We call a list $(\alpha_1, a_1) \ldots (\alpha_n, a_n)$, where $n$ is a nonnegative integer and the $(\alpha_i, a_i)$ are tickets, a *gamble*. The $\alpha_i$ are the *sentences* of the gamble; the $a_i$ are the *payoffs*. The integer $n$ may be zero; in this case the gamble is an empty list. Notice also that a ticket may occur in a gamble more than once. We write *Gamble* for the set consisting of all gambles. We use meta-variables such as $G, G'$, etc. to designate gambles without specifying the tickets they contain.

Given a gamble $G = (\alpha_1, a_1) \ldots (\alpha_n, a_n)$, we define a variable $X_G$ by

$$X_G(\omega) := \sum\{a_i | 1 \le i \le n \text{ and } \omega(\alpha_i) = true$$

$$= \sum_{i=1}^{n} a_i \cdot I_{[\alpha_i]}(\omega). \tag{4}$$

We call $X_G$ the *variable representation* for the gamble $G$. Many different gambles can be represented by the same variable. We call a variable $X$ *simple* if $X = X_G$ for some gamble $G$. A variable is simple if and only if (1) it takes only a finite number of values, and (2) for each real number $r$, there exists $\alpha \in \text{wff}_0$ such that $\{\omega \mid X(\omega) = r\} = [\alpha]$ (in the language of probability theory, $X$ is measurable with respect to $\mathcal{A}_0$).

Here is our notation for manipulating gambles:

- Given a gamble $G = (\alpha_1, a_1) \ldots (\alpha_n, a_n)$ and a sentence $\beta$ in $\mathcal{L}_0$, we write $G^\beta$ for the result of conjoining each $\alpha_i$ with $\beta$:

$$G^\beta := (\alpha_1 \wedge \beta, a_1) \ldots (\alpha_n \wedge \beta, a_n).$$

- Given a gamble $G = (\alpha_1, a_1) \ldots (\alpha_n, a_n)$ and a real number $r$, we write $rG$ for the result of multiplying each $a_i$ by $r$:

$$rG := (\alpha_1, ra_1) \ldots (\alpha_n, ra_n).$$

- Given gambles $G = (\alpha_1, a_1) \ldots (\alpha_n, a_n)$ and $G' = (\beta_1, b_1) \ldots (\beta_m, b_m)$, we write $GG'$ for the result of concatenating $G$ and $G'$:

$$GG' := (\alpha_1, a_1) \ldots (\alpha_n, a_n)(\beta_1, b_1) \ldots (\beta_m, b_m).$$

These manipulations affect the variable representation in obvious and straightforward ways: $X_{G^\beta} = X_G^{[\beta]}$ (see equation (2)), $X_{rG} = rX_G$, and $X_{GG'} = X_G + X_{G'}$.

We sometimes want to *append* a ticket to a gamble or *remove* an instance of a ticket from a gamble:

- Appending a ticket to a gamble means adding it at the end, without regard to whether it already occurs in the gamble. For example, the result of appending $(\beta, b)$ to $(\alpha, a)(\beta, b)(\gamma, c)$ is $(\alpha, a)(\beta, b)(\gamma, c)(\beta, b)$.

- Removing an instance of $(\beta, b)$ from $(\alpha, a)(\beta, b)(\gamma, c)(\beta, b)$ can result in either $(\alpha, a)(\beta, b)(\gamma, c)$ or $(\alpha, a)(\gamma, c)(\beta, b)$.

Here are ten important ways of changing a gamble. We call them *elementary moves*:

| | |
|---|---|
| 0 | permute the order of the tickets |
| 1 | append $(\alpha, 0)$ |
| 2 | remove an instance of $(\alpha, 0)$ |
| 3 | append $(\perp, a)$ |
| 4 | remove an instance of $(\perp, a)$ |
| 5 | remove an instance of $(\alpha, a)$, and append $(\alpha, a_1)$ and $(\alpha, a_2)$, where $a_1 + a_2 = a$ |
| 6 | remove an instance of $(\alpha, a_1)$ and an instance of $(\alpha, a_2)$, and append $(\alpha, a_1 + a_2)$ |
| 7 | remove an instance of $(\alpha, a)$, and append $(\beta, a)$ and $(\gamma, a)$, where $\beta \wedge \gamma \Leftrightarrow_0 \perp$ and $\beta \vee \gamma \Leftrightarrow_0 \alpha$ |
| 8 | remove an instance of $(\beta, a)$ and an instance of $(\gamma, a)$, and append $(\alpha, a)$, where $\beta \wedge \gamma \Leftrightarrow_0 \perp$ and $\beta \vee \gamma \Leftrightarrow_0 \alpha$ |
| 9 | remove an instance of $(\alpha, a)$ and append $(\beta, a)$, where $\beta \Leftrightarrow_0 \alpha$ |

We say that two gambles $G$ and $G'$ *equivalent* if we can get from one to the other by elementary moves—i.e., if there is a finite sequence of gambles $G_1, \ldots, G_k$ such that $G_1 = G$, $G_k = G'$, and $G_{i+1}$ can be obtained from $G_i$ by an elementary move, for $i = 1, \ldots, k - 1$. This is evidently an equivalence relation. Elementary moves in part are similar to Fagin, Halpern, and Meggido's valid formulae about linear equalities and in part capture intuitions regarding the decomposition of gambles and the nature of probability.

**Proposition 2.1** *Gambles $G$ and $G'$ are equivalent if and only if $X_G = X_{G'}$.*

It is easy to see that $X_G = X_{G'}$ when $G'$ is obtained from $G$ by an elementary move, and this implies that $X_G = X_{G'}$ whenever $G$ and $G'$ are equivalent. So our task is to show that $X_G = X_{G'}$ implies the equivalence of $G$ and $G'$.

To this end, we introduce some additional concepts. We say that two of $\mathcal{L}_0$'s sentences $\alpha$ and $\beta$ are *disjoint* if $\alpha \wedge \beta \Leftrightarrow_0 \perp$. We say that a gamble $(\alpha_1, a_1) \ldots (\alpha_n, a_n)$ is *in standard form* if the following conditions are satisfied:

- The sentences are disjoint: $\alpha_i \wedge \alpha_j \Leftrightarrow_0 \perp$ for $i \ne j$.

- No sentence is absurd: $\alpha_i \not\Leftrightarrow_0 \bot$ for all $i$.

- No payoff is zero: $a_i \neq 0$ for all $i$.

- The payoffs are distinct and in increasing order: $a_1 < \cdots < a_n$.

**Lemma 2.2** *Any gamble is equivalent to a gamble in standard form.*

**Proof:** Consider a gamble $G = (\alpha_1, a_1) \ldots (\alpha_n, a_n)$. By repeated elementary moves of type 7, we can reduce $G$ to an equivalent gamble $G_1$ in which every ticket's sentence has the form

$$\beta_1 \wedge \cdots \wedge \beta_n, \tag{5}$$

where for each $i$, either $\beta_i = \alpha_i$ or else $\beta_i = \neg\alpha_i$. (We ignore the placement of parentheses in the expression (5), but we assume that these parentheses are placed in some canonical way, using if necessary elementary moves of type 9.) Any two sentences in $G_1$ are either disjoint or equivalent. Using elementary moves of type 6 and 9, we can consolidate the tickets with equivalent sentences, reducing $G_1$ to an equivalent gamble $G_2$ whose sentences are disjoint. Using repeated elementary moves of type 8, we can reduce $G_2$ to an equivalent gamble $G_3$ whose sentences are still disjoint and whose payoffs are all distinct. Elementary moves of types 0, 2, and 4 will then reduce $G_3$ to an equivalent gamble in standard form. ∎

**Lemma 2.3** *If $G$ and $G'$ are in standard form, and $X_G = X_{G'}$, then $G$ and $G'$ are equivalent.*

**Proof:** As $X_G = X_{G'}$, the two gambles must have the same list of payoffs $a_1, \ldots, a_n$. Since $G$ and $G'$ are in standard form, $G = (\alpha_1, a_1) \ldots (\alpha_n, a_n)$ and $G' = (\beta_1, a_1) \ldots (\beta_n, a_n)$, where $[\alpha_i] = [\beta_i]$ for $i = 1, \ldots, n$. So $\alpha_i \Leftrightarrow_0 \beta_i$, and $G$ can be transformed into $G'$ by elementary moves of type 9. ∎

We complete the proof of Proposition 2.1 by considering two gambles $G_1$ and $G_2$ such that $X_{G_1} = X_{G_2}$ and showing that they are equivalent. By Lemma 2.2, we have gambles $G'_1$ and $G'_2$ that are in standard form and are equivalent to $G_1$ and $G_2$ respectively. The equivalence of $G'_i$ and $G_i$ implies that $X_{G'_i} = X_{G_i}$, and hence that $X_{G'_1} = X_{G'_2}$. By Lemma 2.3, $G'_1$ and $G'_2$ are equivalent. Hence $G_1$ and $G_2$ are equivalent.

### 2.3 The Syntax of $\mathcal{L}$

As we have already explained, a sentence of $\mathcal{L}$ is any expression of the form (1), where $n$ is a nonnegative integer, $\delta$ and $\alpha_1, \ldots, \alpha_n$ are sentences of $\mathcal{L}_0$, and $a_1, \ldots, a_n$ are real numbers. Notice that $n$ is allowed to be zero, so that $\langle \mid \delta \rangle$ is a sentence of $\mathcal{L}$.

We write *wff* for the set consisting of all sentences of $\mathcal{L}$. We speak of the *gamble* and the *condition* of an element of *wff*; $(\alpha_1, a_1) \ldots (\alpha_n, a_n)$ is $\langle (\alpha_1, a_1) \ldots (\alpha_n, a_n) \mid \delta \rangle$'s gamble, and $\delta$ is its condition. We use meta-variables such as $S, S'$, etc. to designate elements of *wff* without specifying their gambles or conditions.

Notice that we do not form negations of the sentences in $\mathcal{L}$: when $S$ is a sentence of $\mathcal{L}$, $\neg S$ is *not* a sentence in $\mathcal{L}$. Nor do we form conjunctions or disjunctions.

### 2.4 Inference in $\mathcal{L}$

We now define an inference relation $\vdash$ for $\mathcal{L}$.

We adopt one axiom schema and five inference rules. The axiom schema applies to any $\alpha, \delta \in \textit{wff}_0$ while the inference rules apply to any $G, G' \in$ *Gamble*, and any $\delta, \epsilon \in \textit{wff}_0$.

***Acceptability***

RATIONALITY $\vdash \langle (\alpha, 1) \mid \delta \rangle$.

SUBSTITUTION $\langle G \mid \delta \rangle \vdash \langle G' \mid \delta \rangle$ if $G$ and $G'$ are equivalent.

COMBINATION $\{ \langle G \mid \delta \rangle, \langle G' \mid \delta \rangle \} \vdash \langle GG' \mid \delta \rangle$.

SCALING $\langle G \mid \delta \rangle \vdash \langle rG \mid \delta \rangle$ if $r \geq 0$.

***Conditioning***

CONTINGENCY $\langle G \mid \delta \rangle \vdash \langle G^\delta \mid \epsilon \rangle$ if $\delta \Rightarrow_0 \epsilon$.

UPDATING $\langle G^\delta \mid \epsilon \rangle \vdash \langle G \mid \delta \rangle$ if $\delta \Rightarrow_0 \epsilon$.

The axiom and first three rules capture our notion of the acceptability of gambles, consistent with Wilson and Moral's approach or Walley's sense of desirability as applied to our formal notion of gambles. The conditioning rules, on the other hand, capture a particular interpretation of conditional probabilities (the "called-off bet" interpretation implicit in Frisch and Haddawy's system). Under this interpretation, conditional probabilities can be defined in terms of unconditional probabilities, and so our logic could be formulated more simply without using conditioning sentences; our intention in future work, however, is to study variants of our logic $\mathcal{L}$ in which the acceptability axiom and rules are retained in conjunction with weaker conditioning rules, and conditioning sentences remain fundamental.

Inference proceeds in the usual way. One starts with a set of premises and enlarges it in steps, including at each either an axiom or a sentence whose inference is authorized from sentences already in the set by one of

the inference rules. If $\Gamma \subseteq$ *wff*, $S \in$ *wff*, and we can infer $S$ from $\Gamma$ in this way, then we write $\Gamma \vdash S$.

Implementing the inference rules involves, of course, using the inference procedure of the underlying logic $\mathcal{L}_0$. In order to use Contingency, for example, we must show that $\epsilon$ can be inferred from $\delta$. Inference in $\mathcal{L}_0$ enters even into the use of Substitution, since we need to demonstrate equivalence or implication in $\mathcal{L}_0$ in order to prove the equivalence of two gambles.

The following proposition lists some elementary consequences of our inference rules.

**Proposition 2.4** *1.* $\langle G \mid \delta \rangle \vdash \langle G^\delta \mid \delta \rangle$ *and* $\langle G^\delta \mid \delta \rangle \vdash \langle G \mid \delta \rangle$.

*2.* $\langle G \mid \delta \rangle \vdash \langle G^\delta \mid \top \rangle$ *and* $\langle G^\delta \mid \top \rangle \vdash \langle G \mid \delta \rangle$.

*3. If* $\delta \Leftrightarrow_0 \epsilon$, *then* $\langle G \mid \delta \rangle \vdash \langle G \mid \epsilon \rangle$.

**Proof:** We obtain the two inferences in Statement 1 by setting $\epsilon$ equal to $\delta$ in Contingency and Updating, respectively. We similarly obtain Statement 2 by setting $\epsilon$ equal to $\top$. To derive Statement 3, we start with $\langle G \mid \delta \rangle$, use Contingency to get $\langle G^\delta \mid \epsilon \rangle$, use the equivalence of the gambles $G^\delta$ and $G^\epsilon$, together with Substitution, to get $\langle G^\epsilon \mid \epsilon \rangle$, and then use Updating to get $\langle G \mid \epsilon \rangle$. ∎

It is noteworthy that the logic $\mathcal{L}$ includes an absurdity—a sentence from which any other sentence can be inferred. This is the sentence $\langle (\top, -1) \mid \top \rangle$. This sentence says that our agent is willing to give away \$1 *a priori*, and the inference rules allow us to infer from this that he will be willing to give away any amount of money under any other state of knowledge $\delta$. If $\Gamma \vdash \langle (\top, -1) \mid \top \rangle$, then we say that $\Gamma$ is *incoherent*. More generally, if $\Gamma \vdash \langle (\delta, -1) \mid \delta \rangle$, then we say that $\Gamma$ is *incoherent* in $\delta$.

## 3 Measure-Theoretic Semantics for $\mathcal{L}$

As we have already explained informally, we can adapt the measure-theoretic semantics for our language $\mathcal{L}$ by using the notion of conditional expected value:

- An *interpretation* of $\mathcal{L}$ is a probability distribution $P$ on the sample space $\Omega_0$.

- An interpretation $P$ *satisfies* a sentence $\langle G \mid \delta \rangle$ if $P$'s expected value for the variable $X_G$ (see equation (4)), conditional on $\delta$, is nonnegative.

This definition of satisfaction is only informal. Our formal definition will resolve the indeterminacy of conditional expected value when the condition $\delta$ has probability zero in a way consistent with Frisch and

Haddawy: the sentence is satisfied by the interpretation in this case.

In this section, we study the entailment relation for $\mathcal{L}$ based on this measure-theoretic semantics and then show that the inference procedure we described in the preceding section is sound and complete with respect to this relation. We explain our alternative semantics, behavioral semantics in the next section.

### 3.1 Entailment Under Measure-Theoretic Semantics

Formally, we say that $P$ *satisfies* $\langle (\alpha_1, a_1) \ldots (\alpha_n, a_n) \mid \delta \rangle$ if

$$\sum_{i=1}^{n} a_i \cdot P([\alpha_i] \cap [\delta]) \geq 0. \qquad (6)$$

This inequality is equivalent to

$$\sum_{i=1}^{n} a_i \cdot \frac{P([\alpha_i] \cap [\delta])}{P([\delta])} \geq 0, \qquad (7)$$

provided that we agree to the convention that the ratio $P([\alpha_i] \cap [\delta])/P([\delta])$ is equal to zero (and hence inequality (7) is satisfied) whenever the denominator, $P([\delta])$, is equal to zero. The left-hand side of inequality (7) is the conditional expected value of the variable corresponding to the gamble $(\alpha_1, a_1) \ldots (\alpha_n, a_n)$, conditional on the event $[\delta]$. This justifies the informal definition we offered a moment ago: $P$ satisfies $\langle G \mid \delta \rangle$ if the expected value of $X_G$, conditional on $\delta$, is nonnegative. This treatment of conditional probability is mandated by the called-off bet interpretation of conditional probability discussed earlier.

If we write $G$ for the gamble $(\alpha_1, a_1) \ldots (\alpha_n, a_n)$, then we can rewrite the inequality (6) as a condition on the variable representation for $G$:

$$E_P X_G^{[\delta]} \geq 0, \qquad (8)$$

where $E_P$ represents the expected value operator for $P$.

We write $\models^m$ for the measure-theoretic entailment relation for $\mathcal{L}$: $\Gamma \models^m S$ if and only if $P$ satisfies $S$ whenever $P$ satisfies $S'$ for all $S' \in \Gamma$. As usual, we abbreviate $\Gamma \models^m S$ to $\models^m S$ when $\Gamma$ is empty; this means that every interpretation $P$ satisfies $S$.

### 3.2 Soundness Under Measure-Theoretic Semantics

Now we verify that $\mathcal{L}$'s inference procedure is sound with respect to measure-theoretic semantics: if $\Gamma \vdash G$, then $\Gamma \models^m G$. It suffices to show that the axioms and the inference rules are sound.

RATIONALITY $\vdash \langle(\alpha,1) \mid \delta\rangle$.

Every interpretation $P$ satisfies $\langle(\alpha,1) \mid \delta\rangle$, because the inequality (6) reduces to $P([\alpha]\cap[\delta]) \geq 0$ for this sentence, and a probability is always nonnegative.

SUBSTITUTION $\langle G \mid \delta\rangle \vdash \langle G' \mid \delta\rangle$ if $G$ and $G'$ are equivalent.

Soundness follows from the fact that the inequality (6) can be put in the form (8) and the fact that equivalent gambles have the same variable representation (Proposition 2.1).

COMBINATION $\{\langle G \mid \delta\rangle, \langle G' \mid \delta\rangle\} \vdash \langle GG' \mid \delta\rangle$.

Soundness follows from the inequality (8) and the relation $X_{GG'}^{[\delta]} = X_G^{[\delta]} + X_{G'}^{[\delta]}$.

SCALING $\langle G \mid \delta\rangle \vdash \langle rG \mid \delta\rangle$ if $r \geq 0$.

Soundness follows from the inequality (8) and the relation $X_{rG}^{[\delta]} = rX_G^{[\delta]}$.

CONTINGENCY $\langle G \mid \delta\rangle \vdash \langle G^\delta \mid \epsilon\rangle$ if $\delta \Rightarrow_0 \epsilon$.

Here we use the inequality (8) and the calculation $X_{G^\delta}^{[\epsilon]} = (X_G^{[\epsilon]})^{[\delta]} = X_G^{[\epsilon]\cap[\delta]}$; when $\delta \Leftrightarrow_0 \epsilon$, $[\epsilon] \cap [\delta] = [\delta]$.

UPDATING $\langle G^\delta \mid \epsilon\rangle \vdash \langle G \mid \delta\rangle$ if $\delta \Rightarrow_0 \epsilon$.

Soundness follows by the same argument as for CONTINGENCY.

### 3.3 Completeness Under Measure-Theoretic Semantics

Now we verify that $\mathcal{L}$'s inference procedure is complete under measure-theoretic semantics: if $\Gamma$ is finite and $\Gamma \models^m S$, then $\Gamma \vdash S$.

**Lemma 3.1** If $X_G^{[\delta]} = X_{G'}^{[\delta']}$, then $\langle G \mid \delta\rangle \vdash \langle G' \mid \delta'\rangle$.

**Proof:** We have $\langle G \mid \delta\rangle \vdash \langle G^\delta \mid \delta \vee \delta'\rangle$ by CONTINGENCY. Our hypothesis $X_G^{[\delta]} = X_{G'}^{[\delta']}$ implies $X_{G^\delta} = X_{G'^{\delta'}}$. So by Proposition 2.1, $G^\delta$ and $G'^{\delta'}$ are equivalent, and therefore $\langle G^\delta \mid \delta \vee \delta'\rangle \vdash \langle G'^{\delta'} \mid \delta \vee \delta'\rangle$ by SUBSTITUTION. Finally, $\langle G'^{\delta'} \mid \delta \vee \delta'\rangle \vdash \langle G' \mid \delta'\rangle$ by UPDATING. ∎

**Lemma 3.2** If $X$ is simple and $\min X \geq 0$, then there exist $G \in$ *Gamble* and $\delta \in wff_0$ such that $X_G^{[\delta]} = X$ and $\vdash \langle G \mid \delta\rangle$.

**Proof:** Suppose $a_1, \ldots, a_n$ are the values taken by $X$, and for each $i$ choose $\alpha_i \in wff_0$ such that $[\alpha_i] = \{\omega \mid X(\omega) = a_i\}$. Then $X = X_G$, where $G := (\alpha_1, a_1)\ldots(\alpha_n, a_n)$. So $X = X_G^{[\top]}$. Because the $a_i$ are nonnegative, $\vdash \langle(\alpha_i, a_i) \mid \top\rangle$ for all $i$ by RATIONALITY and SCALING, and so $\vdash \langle G \mid \top\rangle$ by COMBINATION. ∎

**Lemma 3.3** If $r > 0$ and $X_{G'}^{[\delta']} = rX_G^{[\delta]}$, then $\langle G \mid \delta\rangle \vdash \langle G' \mid \delta'\rangle$.

**Proof:** We have $\langle G \mid \delta\rangle \vdash \langle rG \mid \delta\rangle$ by SCALING. And because our hypothesis can be written in the form $X_{G'}^{[\delta']} = X_{rG}^{[\delta]}$, we have $\langle rG \mid \delta\rangle \vdash \langle G' \mid \delta'\rangle$ by Lemma 3.1. ∎

**Lemma 3.4** If $X_G^{[\delta]} = X_{G_1}^{[\delta_1]} + X_{G_2}^{[\delta_2]}$, then $\{\langle G_1 \mid \delta_1\rangle, \langle G_2 \mid \delta_2\rangle\} \vdash \langle G \mid \delta\rangle$.

**Proof:** Because $X_{G_i}^{[\delta_i]} = X_{G_i\delta_i} = X_{G_i\delta_i}^{[\delta_1 \vee \delta_2]}(i = 1, 2)$, we can infer $\langle G_1^{\delta_1} \mid \delta_1 \vee \delta_2\rangle$ and $\langle G_2^{\delta_2} \mid \delta_1 \vee \delta_2\rangle$ by Lemma 3.1. We can then infer $\langle G_1^{\delta_1} G_2^{\delta_2} \mid \delta_1 \vee \delta_2\rangle$ by COMBINATION. Because $X_{G_1\delta_1 G_1\delta_1}^{[\delta_1 \vee \delta_2]} = X_{G_1\delta_1}^{[\delta_1 \vee \delta_2]} + X_{G_2\delta_2}^{[\delta_1 \vee \delta_2]} = X_{G_1}^{[\delta_1]} + X_{G_2}^{[\delta_2]} = X_G^{[\delta]}$, we can then infer $\langle G \mid \delta\rangle$ by Lemma 3.1. ∎

To complete our demonstration of completeness, we use the completeness of Wilson and Moral's logic of gambles $\mathcal{L}_{WM}$. Suppose $\Gamma \models^m S$, where

$$\Gamma = \{\langle G_1 \mid \delta_1\rangle, \ldots, \langle G_k \mid \delta_k\rangle\} \quad \text{and} \quad S = \langle G \mid \delta\rangle.$$

By virtue of the representation (8), this means that an analogous entailment holds in $\mathcal{L}_{WM}$:

$$\{X_{G_1}^{[\delta_1]}, \ldots, X_{G_k}^{[\delta_k]}\} \models_{WM} X_G^{[\delta]}.$$

So, by the completeness of $\mathcal{L}_{WM}$,

$$\{X_{G_1}^{[\delta_1]}, \ldots, X_{G_k}^{[\delta_k]}\} \vdash_{WM} X_G^{[\delta]}.$$

This means that we can get from the variables $\{X_{G_1}^{[\delta_1]}, \ldots, X_{G_k}^{[\delta_k]}\}$ to the variable $X_G^{[\delta]}$ in a finite number of steps using Wilson and Moral's axiom and two inference rules [13]. In other words, $X_G^{[\delta]}$ is a positive linear combination (a linear combination with positive coefficients) of $X_{G_1}^{[\delta_1]}, \ldots, X_{G_k}^{[\delta_k]}$ together with a finite number of nonnegative simple variables. But by Lemma 3.2, for any nonnegative simple variable $X$, there is a sentence representing $X$ that can be inferred in $\mathcal{L}$. And by Lemmas 3.3 and 3.4, we can infer from a finite number of sentences with given representing variables a sentence whose representing variable is any positive linear combination of these variables. This gets us to a sentence $\langle G' \mid \delta'\rangle$ such that $X_{G'}^{[\delta']} = X_G^{[\delta]}$. Finally, by Lemma 3.1, we can then infer $\langle G \mid \delta\rangle$.

## 4 Behavioral Semantics for $\mathcal{L}$

In our behavioral semantics for $\mathcal{L}$, which we study in this section, we replace the concept of a probability distribution with a concept of lower prevision adapted from Walley [12].

The context for this section is the syntax and inference procedure for $\mathcal{L}$ that we developed in §2. In particular, an underlying logic $\mathcal{L}_0$ is in place, and the sample space $\Omega_0$ is defined. We leave aside, however, the semantics developed in §3 to study a different semantics for the same syntax and inference procedure.

## 4.1 Lower Previsions

Suppose $\Omega$ is a nonempty set and $\mathcal{A}$ is a field of subsets of $\Omega$. Write $\mathcal{X}$ for the linear space consisting of all real-valued functions on $\Omega$ that are measurable with respect to $\mathcal{A}$ and take only finitely many values. We call a real-valued function $\underline{E}$ on $\mathcal{X}$ a *lower prevision* for $\Omega$ if it satisfies these three conditions:

1. $\underline{P}(X) \geq \inf\{X(\omega) \mid \omega \in \Omega\}$ for all $X \in \mathcal{X}$.

2. $\underline{P}(X_1 + X_2) \geq \underline{P}(X_1) + \underline{P}(X_2)$ for all $X_1, X_2 \in \mathcal{X}$.

3. $\underline{P}(rX) = r\underline{P}(X)$ for all $r \geq 0$ and $X \in \mathcal{X}$.

We call a lower prevision satisfying condition 2 with equality for all $X_1, X_2 \in \mathcal{X}$ a *linear prevision*.

The following proposition gives some insight into the concepts of lower prevision and linear prevision by relating them to more familiar concepts.

**Proposition 4.1** *1. A real-valued function $\underline{P}$ on $\mathcal{X}$ is a linear prevision if and only if there is a finitely additive probability measure $P$ on $(\Omega, \mathcal{A})$ such that*

$$\underline{P}(X) = E_P X \text{ for all } X \in \mathcal{X}.$$

*(Here $E_P$ is the expected value operator for $P$.)*

*2. A real-valued function $\underline{P}$ on $\mathcal{X}$ is a lower prevision if and only if it is the lower envelope of the expected value operators for a set of finitely additive probability measures—i.e., if and only if there is a set $\Lambda$ of finitely additive probability measures on $(\Omega, \mathcal{A})$ such that*

$$\underline{P}(X) = \inf_{P \in \Lambda} E_P X \text{ for all } X \in \mathcal{X}.$$

Statement 1 is proven in §3.2 of Walley [12] and Statement 2 in §3.3. Notice, however, that our terminology is not quite the same as Walley's. He calls any real-valued function on any set of real-valued functions on $\Omega$ a lower prevision, and he relates the three conditions that we have used as the definition of lower prevision to a concept that he calls coherence. For an explanation of our disagreement with Walley regarding coherence, see [10].

Our motivation for considering lower previsions does not derive from their relation to probability measures. On the contrary, we regard lower previsions as more fundamental than probability measures, because they emerge more directly from the idea of the acceptability of gambles. For the moment, let us follow Wilson and Moral by thinking of the elements of $\mathcal{X}$ as gambles, and let us write $\mathcal{C}$ for the subset of $\mathcal{X}$ consisting of the gambles we consider acceptable. What should $\mathcal{C}$ look like? According to the intuition that underlies both our inference procedure for $\mathcal{L}$ and our definition of lower prevision, $\mathcal{C}$ should satisfy three conditions (see also Walley [12]):

1. If $X_1, \ldots, X_n \in \mathcal{C}$, and $r_1, \ldots, r_n$ are positive real numbers, then $r_1 X_1 + \cdots + r_n X_n \in \mathcal{C}$.

2. If $\inf_{\omega \in \Omega} X(\omega) \geq 0$, then $X \in \mathcal{C}$.

3. $-1 \notin \mathcal{C}$.

But it is easy to verify that if $\mathcal{C}$ verifies these conditions, and we set

$$\underline{P}(X) = \sup\{a \mid X - a \in \mathcal{C}\} \qquad (9)$$

for all $X \in \mathcal{X}$, then $\underline{P}$ qualifies as a lower prevision, and

$$\underline{P}(X) \geq 0 \text{ if and only if } X \in \mathcal{C}.$$

(Accepting $X - a$ is the same as paying $a$ for $X$, and so (9) is the most one will pay for $X$. Thus $\underline{P}(X) \geq 0$ if and only if one will pay 0 for $X$.)

Conversely, starting with a lower prevision $\underline{P}$, we can set $\mathcal{C} := \{X \mid \underline{P} \geq 0\}$, and $\mathcal{C}$ will satisfy our three conditions. So we could define our behavioral semantics directly in terms of $\mathcal{C}$ rather than in terms of $\underline{P}$. This would make behavioral semantics much more transparent. Using lower previsions has the advantage, however, of making the similarities with measure-theoretic semantics transparent.

## 4.2 Entailment Under Behavioral Semantics

Here are the definitions of interpretation and satisfaction in our behavioral semantics for $\mathcal{L}$:

- An *interpretation* is a lower prevision on $\Omega_0$.

- An interpretation $\underline{P}$ *satisfies* a sentence $\langle G \mid \delta \rangle$ if

$$\underline{P}(X_G^{[\delta]}) \geq 0. \qquad (10)$$

This is quite parallel to measure-theoretic semantics, where an interpretation, a probability distribution $P$ on $\Omega_0$, satisfies $\langle G \mid \delta \rangle$ if $E_P(X_G^{[\delta]}) \geq 0$. Because

the expected value operators for probability distributions are a special kind of lower prevision (namely, linear previsions), we may say that behavioral semantics generalizes measure-theoretic semantics.

We write $\models^b$ for the entailment relation for logic $\mathcal{L}$ determined by this new definition of satisfaction.

### 4.3 Soundness Under Behavioral Semantics

The demonstration that $\mathcal{L}$'s axioms and inference rules are sound with respect to behavioral semantics proceeds just like the demonstration with respect to measure-theoretic semantics (§3.2).

### 4.4 Completeness Under Behavioral Semantics

The completeness of $\mathcal{L}$ under behavioral semantics follows easily from its completeness under measure-theoretic semantics.

Suppose, indeed, that $\Gamma \models^b S$, where

$$\Gamma = \{\langle G_1 \mid \delta_1 \rangle, \ldots, \langle G_k \mid \delta_k \rangle\} \quad \text{and} \quad S = \langle G \mid \delta \rangle.$$

This means that if $\underline{P}$ is a lower prevision and $\underline{P}(X_{G_i}^{[\delta_i]}) \geq 0$ for $i = 1, \ldots, k$, then $\underline{P}(X_G^{[\delta]}) \geq 0$. Because a linear prevision is a lower prevision, this means in particular that if $E_P X_{G_i}^{[\delta_i]} \geq 0$ for $i = 1, \ldots, k$, then $E_P X_G^{[\delta]} \geq 0$. In other words, $\Gamma \models^m S$, where $\models^m$ is the entailment relation under measure-theoretic semantics. So we obtain $\Gamma \vdash S$ from completeness under measure-theoretic semantics.

## 5 Translating $\mathcal{L}_N$ and $\mathcal{L}_{FH}$ into $\mathcal{L}$

Nilsson, in $\mathcal{L}_N$, began with sentences of the form $\mathbb{P}(\alpha) = a$ and inferred sentences of the more general forms $a \leq \mathbb{P}(\alpha)$ and $\mathbb{P}(\alpha) \leq b$. In order to translate these sentences into $\mathcal{L}$, we can be guided by a fact about the common semantics: the probability of an event $A$ under an interpretation $P$ is the same as the expected value under $P$ of the indicator variable $I_A$. This produces the translations shown in Table 2.

| Sentence in $\mathcal{L}_N$ | Equivalent condition on $P$ | Sentence in $\mathcal{L}$ |
|---|---|---|
| $a \leq \mathbb{P}(\alpha)$ | $E_P(I_{[\alpha]} - a) \geq 0$ | $\langle (\alpha, 1)\, (\top, -a) \mid \top \rangle$ |
| $\mathbb{P}(\alpha) \leq b$ | $E_P(b - I_{[\alpha]}) \geq 0$ | $\langle (\top, b)\, (\alpha, -1) \mid \top \rangle$ |

Table 2: Translating from Nilsson's logic $\mathcal{L}_N$ to $\mathcal{L}$.

In Frisch and Haddawy's logic $\mathcal{L}_{FH}$, a sentence simultaneously expresses a lower and an upper bound on a conditional probability:

$$\mathbb{P}(\alpha \mid \delta) \in [a, b]. \tag{11}$$

How should we translate this sentence into $\mathcal{L}$? The most natural approach might be to add the condition $\delta$ to the translations of $a \leq \mathbb{P}(\alpha)$ and $\mathbb{P}(\alpha) \leq b$ in Table 2. This produces two sentences:

$$\langle (\alpha, 1)\, (\top, -a) \mid \delta \rangle \text{ and } \langle (\top, b)\, (\alpha, -1) \mid \delta \rangle. \tag{12}$$

Another approach is to translate the sentence (11) directly into a condition on an interpretation $P$:

$$a \leq \frac{P([\alpha] \cap [\delta])}{P([\delta])} \leq b. \tag{13}$$

Under the convention that the ratio is zero when its denominator is zero, condition (13) is equivalent to the two conditions

$$P([\alpha] \cap [\delta]) - a \cdot P([\delta]) \geq 0 \text{ and } b \cdot P([\delta]) - P([\alpha] \cap [\delta]) \geq 0,$$

and these two conditions are naturally expressed in $\mathcal{L}$ by the two sentences

$$\langle (\alpha \wedge \delta, 1)\, (\delta, -a) \mid \top \rangle \text{ and } \langle (\delta, b)\, (\alpha \wedge \delta, -1) \mid \top \rangle \tag{14}$$

Both (12) and (14) are correct; the sentences $\langle (\alpha, 1)\, (\top, -a) \mid \delta \rangle$ and $\langle (\alpha \wedge \delta, 1)\, (\delta, -a) \mid \top \rangle$ are equivalent to each other, and the sentences $\langle (\top, b)\, (\alpha, -1) \mid \delta \rangle$ and $\langle (\delta, b)\, (\alpha \wedge \delta, -1) \mid \top \rangle$ are equivalent to each other.

Because the sentences (11) in $\mathcal{L}_{FH}$ have the same interpretation as the sentences (14) in $\mathcal{L}$ (both mean that the condition (13) holds under the convention that the ratio is zero when the denominator is zero), this translation is in fact a translation of $\mathcal{L}_{FH}$ into our probabilistic logic. So we may say that we have extended Frisch and Haddawy's logic and added a sound and complete inference procedure. Together with the soundness of their inference rules, this implies that their inference rules are consequences of ours.

There is one complication, deriving from the fact one of Frisch and Haddawy's is represented by two of our sentences. Because of their representation, Frisch and Haddawy introduce the following inference rule:

$$\frac{\mathbb{P}(\alpha \mid \delta) \in [x, y] \quad \mathbb{P}(\alpha \mid \delta) \in [u, v]}{\mathbb{P}(\alpha \mid \delta) \in [max(x, u), min(y, v)]}$$

This rule can be seen as an embodiment of the any-time character of their logic $\mathcal{L}_{FH}$; if we apply this rule whenever it can be applied, we always know the tightest bounds on the probability of $\alpha$ given $\delta$ that are justified by our computation so far. Because we express the lower and upper bounds separately, we have no need for such an inference rule, but in an implementation we could, of course, track the largest $a$ for which $\langle (\alpha \wedge \delta, 1)\, (\delta, -a) \mid \top \rangle$ is in our database and the smallest $b$ for which $\langle (\delta, b)\, (\alpha \wedge \delta, -1) \mid \top \rangle$ is in it.

# 6  Summary and Conclusions

Why stop with a language that is only hyperplane expressive? Why not further expand the language so that it can say anything one pleases about the probabilities of sentences in the underlying logic? There are two obvious ways to answer this question:

- If we are really only interested in the probabilities of individual sentences, as Nilsson and Frisch and Haddawy appeared to be, then there is no reason to generalize further.

- We might feel that we want more than bounds on individual probabilities, but that hyperplane expressiveness is enough. Most practical work with probabilities is directed towards decision-making, and we might argue that decisions depend only on the acceptability of gambles.

We might also challenge the ontological role of probabilities. Is a probability something with a reality of its own, or is it only a way of expressing our attitudes? If it is only a way of expressing our attitudes, and if the attitudes in question come down to the willingness to accept gambles, then we step outside what is meaningful when we go beyond hyperplane expressiveness. This is the view taken by Walley.

Walley's view throws into question, of course, the presumption that a probabilistic logic should use probability measures as interpretations. If the reality to which we are referring has only acceptable gambles, then perhaps these, not probabilities, should be our interpretations. Perhaps also some of our inference rules should be reconsidered. In particular, the conditioning rule UPDATING can be called into question, for it does not have a clear direct justification in terms of the acceptability of gambles.

Our work in this area is motivated by our interest in moving beyond standard probability measures as a semantics for probabilistic logic. In addition to relaxing inference rules such as UPDATING to investigate alternative formulations for conditional probability, we are also interested in shifting away from the static framework of a sample space. This would move probabilistic logic in the direction of temporal and causal logic, and would make contact with our earlier work on basing logic on the concept of an event space [9]. See [10].

## Acknowledgements

## References

[1] Vijay Chandru and John N. Hooker. *Optimization Methods for Logical Inference*. John Wiley & Sons, Inc., New York, 1999.

[2] Ronald Fagin, Joseph Y. Halpern, and Nimrod Megiddo. A logic for reasoning about probabilities. *Information and Control*, 87:78–128, 1990.

[3] Alan Frisch and Peter Haddawy. Anytime deduction for probabilistic logic. *Artificial Intelligence*, 69:93–122, 1994.

[4] Joseph Y. Halpern. *Reasoning about Uncertainty*. The MIT Press, Cambridge, Massachusetts, 2003.

[5] Joseph Y. Halpern and Riccardo Pucella. A logic for reasoning about upper probabilities. In *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence*, pages 203–210, 2001.

[6] Joseph Y. Halpern and Riccardo Pucella. Reasoning about expectation. In *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*, pages 207–215, 2002.

[7] Nils J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28:71–87, 1986.

[8] Glenn Shafer. *The Art of Causal Conjecture*. The MIT Press, Cambridge, Massachusetts, 1996.

[9] Glenn Shafer, Peter R. Gillett, and Richard Scherl. The logic of events. *Annals of Mathematics and Artificial Intelligence*, 28:315–389, 2000.

[10] Glenn Shafer, Peter R. Gillett, and Richard B. Scherl. A new understanding of subjective probability and its generalization to lower and upper prevision. *International Journal of Approximate Reasoning*, 33:1–49, 2003.

[11] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's only a Game!* John Wiley & Sons,Inc., New York, 2001.

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

[13] Nic Wilson and Serafín Moral. A logical view of probability. In *Proceedings of the 11th European Conference on Artificial Intelligence*. John Wiley & Sons, Ltd, 1994.