Characterizing and Reasoning about Probabilistic and Non-Probabilistic Expectation

JOSEPH Y. HALPERN
Cornell University
and
RICCARDO PUCELLA
Northeastern University

Expectation is a central notion in probability theory. The notion of expectation also makes sense for other notions of uncertainty. We introduce a propositional logic for reasoning about expectation, where the semantics depends on the underlying representation of uncertainty. We give sound and complete axiomatizations for the logic in the case that the underlying representation is (a) probability, (b) sets of probability measures, (c) belief functions, and (d) possibility measures. We show that this logic is more expressive than the corresponding logic for reasoning about likelihood in the case of sets of probability measures, but equi-expressive in the case of probability, belief, and possibility. Finally, we show that satisfiability for these logics is NP-complete, no harder than satisfiability for propositional logic.

Categories and Subject Descriptors: F.4.1 [**Mathematical Logic and Formal Languages**]: Mathematical Logic—*Modal logic*; I.2.3 [**Artificial Intelligence**]: Deduction and Theorem Proving—*Uncertainty, "fuzzy", and probabilistic reasoning*; I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods—*Modal logic*; G.3 [**Probability and Statistics**]:

General Terms: Theory

Additional Key Words and Phrases: Expectation, probability theory, Dempster-Shafer belief functions, possibility measures

1. INTRODUCTION

One of the most important notions in probability theory is that of *expectation*. The expected value of a random variable is, in a sense, the single number that best describes the random variable. While probability is certainly still the most

A preliminary version of this paper appeared in the *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence* [Halpern and Pucella 2002b]. We thank the UAI and JACM reviewers for their comments. This work was supported in part by NSF under grant IIS-0090145, by ONR under grants N00014-00-1-0341, N00014-01-1-0511, and N00014-02-1-0455, by the DoD Multidisciplinary University Research Initiative (MURI) program administered by the ONR under grants N00014-97-0505 and N00014-01-1-0795. In addition, while on sabbatical in 2001-02, Halpern was supported by a Guggenheim and a Fulbright Fellowship. Sabbatical support from CWI and the Hebrew University of Jerusalem is also gratefully acknowledged. This work was done while Pucella was at Cornell University. Authors' address: J. Y. Halpern, Department of Computer Science, Cornell University, Ithaca, NY 14853, email: `halpern@cs.cornell.edu`, home page: `http://www.cs.cornell.edu/home/halpern`, R. Pucella, College of Computer and Information Science, Northeastern University, Boston, MA 02115, email: `riccardo@ccs.neu.edu`.

dominant approach to representing uncertainty, in the past two decades there has been a great deal of interest in alternative representations of uncertainty, both from a normative and descriptive point of view. This may seem somewhat surprising to those familiar with the many arguments that have been made showing that probability is the only rational approach to representing uncertainty (see, for example, Cox [1946], Ramsey [1931], de Finetti [1931], Savage [1954]). However, all these arguments depend on assumptions, perhaps the most controversial of which is that the uncertainty of an event can be completely characterized by a single number. (See Walley [1991] for a good summary of the arguments for the need to occasionally go beyond probabilistic expectation.) Some alternatives to probability in the literature include sets of probability measure [Huber 1981; Walley 1991], *Dempster-Shafer belief functions* [Shafer 1976] and the closely related *nonadditive measures* [Schmeidler 1989], and possibility measures [Dubois and Prade 1990].

In this paper, we consider the notion of expectation for all these representations of uncertainty. We do not take a stand here on what the "right" way is to represent uncertainty; we simply investigate characterizations of expectation and reasoning about expectation, both for probability and for other representations of uncertainty.

It is well known that a probability measure determines a unique expectation function that is linear (i.e., $E(aX + bY) = aE(X) + bE(Y)$), monotone (i.e., $X \leq Y$ implies $E(X) \leq E(Y)$), and maps constant functions to their value. Conversely, given an expectation function E (that is, a function from random variables to the reals) that is linear, monotone, and maps constant functions to their value, there is a unique probability measure μ such that $E = E_\mu$. That is, there is a 1-1 mapping from probability measures to (probabilistic) expectation functions. One of the goals of this paper is to provide similar characterizations of expectation for other representations of uncertainty.

Some work along these lines has already been done, particulary with regard to sets of probability measures [Huber 1981; Walley 1991; 1981].[1] However, there seems to be surprisingly little work on characterizing expectation in the context of other measures of uncertainty, such as belief functions [Shafer 1976] and possibility measures [Dubois and Prade 1990]. We provide characterizations here.

Having characterized expectation functions, we then turn to the problem of reasoning about them. We define a logic similar in spirit to that introduced by Fagin et al. [1990] (FHM from now on) for reasoning about likelihood expressed as either probability or belief. The same logic is used by Halpern and Pucella [2002a] (HP from now on) for reasoning about upper probabilities. The logic for reasoning about expectation is strictly more expressive than its counterpart for reasoning about likelihood if the underlying semantics is given in terms of sets of probability measures (so that upper probabilities and upper expectations are used, respectively); it turns out to be equi-expressive in the case of probability, belief functions, and possibility measures. This is somewhat surprising, especially in the case of belief functions. In all cases, the fact that expectations are at least as expressive is immediate, since the expectation of φ (viewed as an indicator function, that is, the random variable that is 1 in worlds where φ is true and 0 otherwise)

[1] Walley [1991] actually characterizes lower and upper *previsions*; but these are essentially lower and upper expectations with respect to sets of probability measures.

is equal to its likelihood. However, it is not always obvious how to express the expectation of a random variable in terms of likelihood.

We then provide a sound and complete axiomatization for the logic with respect to each of the interpretations of expectation that we consider, using our characterization of expectation. Finally, we show that, just as in the case of the corresponding logic for reasoning about likelihood, the complexity of the satisfiability problem is NP-complete. This is clear when the underlying semantics is given in terms of probability measures, belief functions, or possibility measures, but it is perhaps surprising that, despite the added expressiveness in the case of sets of probability measures, reasoning in the logic remains NP-complete.

To the best of our knowledge, there is only one previous attempt to express properties of expectation in a logical setting. Wilson and Moral [1994] take as their starting point Walley's notion of lower and upper previsions. They consider when acceptance of one set of gambles implies acceptance of another gamble. This is a notion that is easily expressible in our logic when the original set of gambles is finite, so our logic subsumes theirs in the finite case.

This paper is organized as follows. In the next section, the characterizations of expectation for probability measures and sets of probability measures are reviewed, and the characterizations of expectation for belief functions and possibility measures are provided. In Section 3, we introduce a logic for reasoning about expectation with respect to all these representations of uncertainty. In Section 4, we compare the expressive power of our expectation logic to that of the logic for reasoning about likelihood. In Section 5, we derive sound and complete axiomatizations for the logic in Section 3, with respect to different representations of uncertainty. In Section 6, we prove that the decision problem for the expectation logic is NP-complete for each of the representations of uncertainty we consider. Finally, in Section 7, we discuss an axiomatization of gamble inequalities, which is assumed by the axiomatizations given in Section 5. The proofs of the more technical results are given in the appendix.

2. EXPECTATION FUNCTIONS

Recall that a *random variable* X on a sample space (set of possible worlds) W is a function from W to some range. Let $\mathcal{V}(X)$ denote the image of X, that is, the possible values of X. A *gamble* is a random variable whose range is the reals. In this paper, we focus on the expectation of gambles. Additionally, we restrict to finite sample spaces; most of the issues of interest already arise in the finite sample space setting. (Most of the results in this section extend in a straightforward way to the infinite sample space setting, by adding suitable continuity assumptions on the measures defined. See Halpern [2003] for more detail.) Note that if the sample space is finite, the range $\mathcal{V}(X)$ of a gamble X is finite. This allows us to define expectation using summation rather than integration.

2.1 Expectation for Probability Measures

Given a finite sample space W, and a probability measure μ and gamble X over W, the *expected value of* X (or the *expectation of* X) with respect to μ, denoted

$E_\mu(X)$, is just

$$\sum_{w \in W} \mu(w)X(w). \tag{1}$$

Thus, the expected value of a gamble is essentially the "average" value of the variable.

Actually, (1) makes sense only if every singleton is measurable (i.e., in the domain of μ). If singletons are not necessarily measurable, the standard assumption is that X is *measurable* with respect to the algebra $\mathcal{F}$ on which μ is defined; that is, for each value $x \in \mathcal{V}(X)$, the set of worlds $X = x$ where X takes on value x is measurable.[2] (In general, a function $f : W \to W'$ is measurable with respect to $\mathcal{F}$ if $f^{-1}(w') \in \mathcal{F}$ for all $w' \in W'$.) Then

$$E_\mu(X) = \sum_{x \in \mathcal{V}(X)} x\mu(X = x). \tag{2}$$

Note that this definition makes sense even if W is not finite, so long as $\mathcal{V}(X)$ is finite. It is easy to check that (1) and (2) are equivalent if W is finite and all singletons are measurable.

As is well known, probabilistic expectation functions can be characterized by a small collection of properties. If X and Y are gambles on W and a and b are real numbers, define the gamble $aX + bY$ on W in the obvious way: $(aX + bY)(w) = aX(w) + bY(w)$. Say that $X \leq Y$ if $X(w) \leq Y(w)$ for all $w \in W$. Let $\tilde{c}$ denote the constant function which always returns c; that is, $\tilde{c}(w) = c$. Let μ be a probability measure on W.

PROPOSITION 2.1. *The function E_μ has the following properties for all measurable gambles X and Y.*

(a) E_μ is additive*: $E_\mu(X + Y) = E_\mu(X) + E_\mu(Y)$.*

(b) E_μ is affinely homogeneous*: $E_\mu(aX + \tilde{b}) = aE_\mu(X) + b$ for all $a, b \in \mathbb{R}$.*

(c) E_μ is monotone*: if $X \leq Y$, then $E_\mu(X) \leq E_\mu(Y)$.*

PROOF. See any standard text in probability or discrete mathematics [Billingsley 1995] ☐

The next result shows that the properties in Proposition 2.1 essentially characterize probabilistic expectation functions. It too is well known. We provide the proof here, just to show how the assumptions are used. In the proof (and throughout the paper) we make use of a special type of random variable. Let X_U denote the gamble such that $X_U(w) = 1$ if $w \in U$ and $X_U(w) = 0$ if $w \notin U$. A gamble of the form X_U is traditionally called an *indicator function*.

THEOREM 2.2. *Suppose that E maps gambles measurable with respect to some algebra $\mathcal{F}$ to $\mathbb{R}$ and E is additive, affinely homogeneous, and monotone. Then there is a (necessarily unique) probability measure μ on $\mathcal{F}$ such that $E = E_\mu$.*

[2]Recall that an algebra $\mathcal{F}$ over a sample space W is a set of subsets of W that includes W itself and is closed under complementation and union, so that if $U, V \in \mathcal{F}$, then so is $\overline{U}$ and $U \cup V$.

PROOF. Define $\mu(U) = E(X_U)$. Note that $X_W = \tilde{1}$, so $\mu(W) = 1$, since E is affinely homogeneous. Since $X_\varnothing$ is $\tilde{0}$ and E is affinely homogeneous, it follows that $\mu(\varnothing) = E(X_\varnothing) = 0$. $X_\varnothing \le X_U \le X_W$ for all $U \subseteq W$; since E is monotone, it follows that $0 = E(X_\varnothing) \le E(X_U) = \mu(U) \le E(X_W) = 1$. If U and V are disjoint, then it is easy to see that $X_{U\cup V} = X_U + X_V$. By additivity,

$$\mu(U \cup V) = E(X_{U\cup V}) = E(X_U) + E(X_V) = \mu(U) + \mu(V).$$

Thus, μ is indeed a probability measure.

To see that $E = E_\mu$, note that it is immediate from (2) that $\mu(U) = E_\mu(X_U)$ for $U \in \mathcal{F}$. Thus, E_μ and E agree on all measurable indicator functions. Every measurable gamble X can be written as a linear combination of measurable indicator functions. For each $a \in \mathcal{V}(X)$, let $U_{X,a} = \{w : X(w) = a\}$. Since X is a measurable gamble, $U_{X,a}$ must be in $\mathcal{F}$. Moreover, $X = \sum_{a\in\mathcal{V}(X)} aX_{U_{X,a}}$. By additivity and affine homogeneity, $E_\mu(X) = \sum_{a\in\mathcal{V}(X)} aE(X_{U_{X,a}})$. By Proposition 2.1, $E_\mu(X) = \sum_{a\in\mathcal{V}(X)} aE_\mu(X_{U_{X,a}})$. Since E and E_μ agree on measurable indicator functions, it follows that $E(X) = E_\mu(X)$. Thus, $E = E_\mu$ as desired.

Clearly, if $\mu(U) \ne \mu'(U)$, then $E_\mu(X_U) \ne E_{\mu'}(X_U)$. Thus, μ is the unique probability measure on $\mathcal{F}$ such that $E = E_\mu$. □

2.2 Expectation for Sets of Probability Measures

If $\mathcal{P}$ is a set of probability measures on a space W, define

$$\begin{array}{l}\mathcal{P}_*(U) = \inf\{\mu(U) : \mu \in \mathcal{P}\} \text{ and}\\ \mathcal{P}^*(U) = \sup\{\mu(U) : \mu \in \mathcal{P}\}.\end{array}$$

$\mathcal{P}_*(U)$ is called the *lower probability* of U and $\mathcal{P}^*(U)$ is called the *upper probability* of U. Lower and upper probabilities have been well studied in the literature (see, for example, Borel [1943], Smith [1961]).

There are straightforward analogues of lower and upper probability in the context of expectation. If $\mathcal{P}$ is a set of probability measures such that X is measurable with respect to each probability measure $\mu \in \mathcal{P}$, then define $E_\mathcal{P}(X) = \{E_\mu(X) : \mu \in \mathcal{P}\}$. $E_\mathcal{P}(X)$ is a set of numbers. Define the *lower expectation* and *upper expectation* of X with respect to $\mathcal{P}$, denoted $\underline{E}_\mathcal{P}(X)$ and $\overline{E}_\mathcal{P}(X)$, as the inf and sup of the set $E_\mathcal{P}(X)$, respectively.

The properties of $\underline{E}_\mathcal{P}$ and $\overline{E}_\mathcal{P}$ are not so different from those of probabilistic expectation functions. Note that $E_\mu(X_U) = \mu(U)$. Similarly, it is easy to see that $\mathcal{P}_*(U) = \underline{E}_\mathcal{P}(X_U)$ and $\mathcal{P}^*(U) = \overline{E}_\mathcal{P}(X_U)$. Moreover, we have the following analogue of Propositions 2.1.

PROPOSITION 2.3. *The functions $\underline{E}_\mathcal{P}$ and $\overline{E}_\mathcal{P}$ have the following properties for all gambles X and Y.*

(a) $\underline{E}_\mathcal{P}(X+Y) \ge \underline{E}_\mathcal{P}(X) + \underline{E}_\mathcal{P}(Y)$ (superadditivity);
$\overline{E}_\mathcal{P}(X+Y) \le \overline{E}_\mathcal{P}(X) + \overline{E}_\mathcal{P}(Y)$ (subadditivity).

(b) $\underline{E}_\mathcal{P}$ *and* $\overline{E}_\mathcal{P}$ *are both* positively affinely homogeneous: $\underline{E}_\mathcal{P}(aX+\tilde{b}) = a\underline{E}_\mathcal{P}(X)+b$ *and* $\overline{E}_\mathcal{P}(aX+\tilde{b}) = a\overline{E}_\mathcal{P}(X)+b$ *if* $a, b \in \mathbb{R}$, $a \ge 0$.

(c) $\underline{E}_\mathcal{P}$ *and* $\overline{E}_\mathcal{P}$ *are monotone.*

(d) $\overline{E}_\mathcal{P}(X) = -\underline{E}_\mathcal{P}(-X)$.

PROOF. This result is also well-known; see Walley [1991, Section 2.6.1]. (Walley proves the result for what he calls coherent lower and upper previsions, but then proves [Walley 1991, Section 3.3.4] that these are equivalent to lower and upper expectations, respectively.) □

Superadditivity (resp., subadditivity), positive affine homogeneity, and monotonicity in fact characterize $\underline{E}_{\mathcal{P}}$ (resp., $\overline{E}_{\mathcal{P}}$).

THEOREM 2.4. [Huber 1981] *Suppose that E maps gambles measurable with respect to $\mathcal{F}$ to $\mathbb{R}$ and is superadditive (resp., subadditive), positively affinely homogeneous, and monotone. Then there is a set $\mathcal{P}$ of probability measures on $\mathcal{F}$ such that $E = \underline{E}_{\mathcal{P}}$ (resp., $E = \overline{E}_{\mathcal{P}}$).*[3]

The set $\mathcal{P}$ constructed in Theorem 2.4 is not unique. It is not hard to construct sets $\mathcal{P}$ and $\mathcal{P}'$ such that $\mathcal{P} \neq \mathcal{P}'$ but $\underline{E}_{\mathcal{P}} = \underline{E}_{\mathcal{P}'}$. However, there is a canonical largest set $\mathcal{P}$ such that $E = \underline{E}_{\mathcal{P}}$; $\mathcal{P}$ consists of all probability measures μ such that $E_\mu(X) \geq E(X)$ for all gambles X. This set $\mathcal{P}$ can be shown to be closed and convex. Indeed, it easily follows that Theorem 2.4 actually provides a 1-1 mapping from closed, convex sets of probability measures to lower/upper expectations. Moreover, in a precise sense, this is the best we can do. If $\mathcal{P}$ and $\mathcal{P}'$ have the same convex closure (where the *convex closure* of a set is the smallest closed, convex set containing it), then $\underline{E}_{\mathcal{P}} = \underline{E}_{\mathcal{P}'}$.

As Walley [1991] shows, what he calls *coherent lower/upper previsions* are also lower/upper expectations with respect to some set of probability measures. Thus, lower/upper previsions can be identified with closed, convex sets of probability measures.

2.3 Expectation for Belief Functions

As is well known, a belief function [Shafer 1976] Bel is a function from subsets of a state space W to $[0, 1]$ satisfying the following three properties:

B1. $\mathrm{Bel}(\varnothing) = 0$.

B2. $\mathrm{Bel}(W) = 1$.

B3. For $n = 1, 2, 3, \ldots$,
$$\mathrm{Bel}(\textstyle\bigcup_{i=1}^n U_i) \geq \sum_{i=1}^n \sum_{\{I \subseteq \{1,\ldots,n\} : |I| = i\}} (-1)^{i+1} \mathrm{Bel}(\bigcap_{j \in I} U_j).$$

Given a belief function Bel, there is a corresponding *plausibility function* Plaus, where $\mathrm{Plaus}(U) = 1 - \mathrm{Bel}(\overline{U})$. It follows easily from B3 that $\mathrm{Bel}(U) \leq \mathrm{Plaus}(U)$ for all $U \subseteq W$. $\mathrm{Bel}(U)$ can be thought of as a lower bound of a set of probabilities and $\mathrm{Plaus}(U)$ can be thought of as the corresponding upper bound. This intuition is made precise in the following well-known result.

THEOREM 2.5. [Dempster 1967] *Given a belief function* Bel *defined on W, let $\mathcal{P}_{\mathrm{Bel}} = \{\mu : \mu(U) \geq \mathrm{Bel}(U)$ for all $U \subseteq W\}$. Then* $\mathrm{Bel} = (\mathcal{P}_{\mathrm{Bel}})_*$.[4]

[3] There is an equivalent characterization of $\underline{E}_{\mathcal{P}}$, due to Walley [1991]. He shows that $E = \underline{E}_{\mathcal{P}}$ for some set $\mathcal{P}$ of probability measures iff E is superadditive, $E(cX) = cE(X)$, and $E(X) \geq \inf\{X(w) : w \in W\}$. There is an analogous characterization of $\overline{E}_{\mathcal{P}}$.

[4] Dempster [1967] defines $\mathcal{P}_{\mathrm{Bel}}$ as $\{\mu : \mathrm{Plaus}(U) \geq \mu(U) \geq \mathrm{Bel}(U)$ for all $U \subseteq W\}$, but his definition is easily seen to be equivalent to that given here. For if $\mu(U) \geq \mathrm{Bel}(U)$ for all $U \subseteq W$ then, in particular, $\mu(\overline{U}) \geq Bel(\overline{U})$, so $\mathrm{Plaus}(U) = 1 - \mathrm{Bel}(\overline{U}) \geq 1 - \mu(\overline{U}) = \mu(U)$.

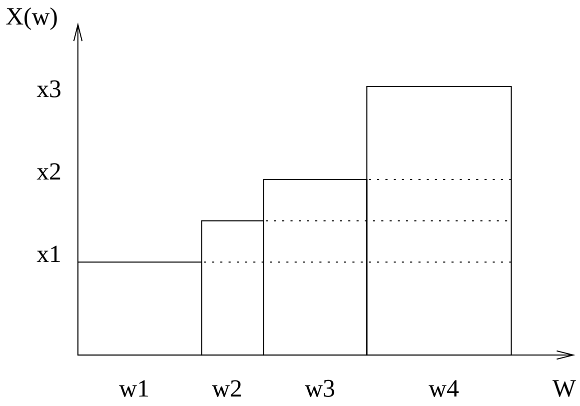

Fig. 1. Two equivalent definitions of probabilistic expectation.

Similarly, we can check that $\mathrm{Plaus}(U) = 1 - \mathrm{Bel}(\overline{U}) = (\mathcal{P}_{\mathrm{Bel}})^*$. There is an obvious way to define a notion of expectation based on belief functions, using the identification of Bel with $(\mathcal{P}_{\mathrm{Bel}})_*$. Given a belief function Bel, define $E_{\mathrm{Bel}} = \underline{E}_{\mathcal{P}_{\mathrm{Bel}}}$. Similarly, for the corresponding plausibility function Plaus, define $E_{\mathrm{Plaus}} = \overline{E}_{\mathcal{P}_{\mathrm{Bel}}}$. (These definitions are in fact used by Dempster [1967]).

This is well defined, but it seems more natural to get a notion of expectation for belief functions that is defined purely in terms of belief functions, without reverting to probability. One way of doing so is due to Choquet [1953].[5]

It takes as its point of departure the following alternate definition of expectation in the case of probability. Suppose that X is a gamble such that $\mathcal{V}(X) = \{x_1, \ldots, x_n\}$, with $x_1 < \ldots < x_n$.

Proposition 2.6. $E_\mu(X) = x_1 + (x_2 - x_1)\mu(X > x_1) + \cdots + (x_n - x_{n-1})\mu(X > x_{n-1})$.

Proof. Figure 1 should help make clear why this result is true, where we assume for simplicity that the worlds $W = \{w_1, w_2, \ldots\}$ are ordered such that if $X(w_i) < X(w_j)$, then $i < j$. If we assume that the probability of a world w_i is the width of the vertical rectangle over w_i, then it should be clear that the total area of the rectangles represent the expectation of X. Notice that the vertical rectangles determine the expectation using the standard definition (2), while the horizontal rectangles determine the expectation using the formula in this proposition.

For a more formal proof, suppose that $\mathcal{V}(X) = \{x_1, \ldots, x_n\}$, where $x_1 < \ldots < x_n$. We proceed by induction on n. If $n = 1$, the result is trivial, since clearly $E_\mu(X) = x_1$. If $n > 1$, then note that $X = X_1 + X_2$, where $X_1(w) = X(w)$ if $X(w) \neq x_n$, and $X_1(w) = x_{n-1}$ if $X(w) = x_n$, and $X_2(w) = 0$ if $X(w) \neq x_n$, and $X(w) = x_n - x_{n-1}$ if $X(w) = x_n$. Clearly $X = X_1 + X_2$, so by additivity, $E(X) = E(X_1) + E(X_2)$. Note that $\mathcal{V}(X_1) = \{x_1, \ldots, x_{n-1}\}$, so by the induction hypothesis, $E_\mu(X_1) = x_1 + (x_2 - x_1)\mu(X > x_1) + \cdots + (x_{n-1} - x_{n-2})\mu(X > x_{n-2})$.

[5]Choquet actually talked about k-monotone capacities, which are essentially functions that satisfy B3 where $n = 1, \ldots, k$. Belief functions are infinitely monotone capacities.

Finally, it is immediate from the definition of X_2 that $E_\mu(X_2) = (x_n - x_{n-1})\mu(X > x_{n-1})$. The result follows immediately. □

Define

$$E'_{\mathrm{Bel}}(X) = x_1 + (x_2 - x_1)\mathrm{Bel}(X > x_1) + \cdots + (x_n - x_{n-1})\mathrm{Bel}(X > x_{n-1}). \quad (3)$$

An analogous definition holds for plausibility:

$$E'_{\mathrm{Plaus}}(X) = x_1 + (x_2 - x_1)\mathrm{Plaus}(X > x_1) + \cdots + (x_n - x_{n-1})\mathrm{Plaus}(X > x_{n-1}). \quad (4)$$

PROPOSITION 2.7. [Schmeidler 1989] *$E_{\mathrm{Bel}} = E'_{\mathrm{Bel}}$ and $E_{\mathrm{Plaus}} = E'_{\mathrm{Plaus}}$.*

Schmeidler [1986; 1989] actually used Choquet's definition to define a notion of expectation for what he called *nonadditive probabilities*, where a nonadditive probability ν maps subsets of a space W to $[0,1]$ such that $\nu(\varnothing) = 0$, $\nu(W) = 1$, and $\nu(U) \leq \nu(V)$ if $U \subseteq V$. He proved an analogue of Proposition 2.7 for arbitrary nonadditive probabilities. Since belief functions and plausibility functions are both nonadditive probabilities in Schmeidler's sense, Proposition 2.7 is actually a special case of Schmeidler's result.

Proposition 2.7 shows that (3) gives a way of defining expectation for belief functions without referring to probability. There is yet a third way of defining expectation for belief functions, which also does not use probability; see the proof of Theorem 6.2 in Appendix A.4.

Since E_{Bel} can be viewed as a special case of the lower expectation $\underline{E}_{\mathcal{P}}$ (taking $\mathcal{P} = \mathcal{P}_{\mathrm{Bel}}$), it is immediate from Proposition 2.3 that E_{Bel} is superadditive, positively affinely homogeneous, and monotone. (Similar remarks hold for E_{Plaus}, except that it is subadditive. For ease of exposition, we focus on E_{Bel} in the remainder of this section, although analogous remarks hold for E_{Plaus}.)

Since it is immediate from the definition that $E_{\mathrm{Bel}}(X_U) = \mathrm{Bel}(U)$, the inclusion-exclusion property B3 of belief functions can be expressed in terms of expectation (just by replacing all instances of $\mathrm{Bel}(V)$ in B3 by $E_{\mathrm{Bel}}(X_V)$). Moreover, it does not follow from the other properties, since it can be shown not to hold for arbitrary lower probabilities. This restatement of the inclusion-exclusion property applies only to the expectation of indicator functions. But, in fact, a more general inclusion-exclusion property holds for E_{Bel}. Given gambles X and Y, define the gambles $X \wedge Y$ and $X \vee Y$ as the minimum and maximum of X and Y, respectively; that is, $(X \wedge Y)(w) = \min(X(w), Y(w))$ and $(X \vee Y)(w) = \max(X(w), Y(w))$. Consider the following inclusion-exclusion property for expectation functions:

$$E(\vee_{i=1}^n X_i) \geq \sum_{i=1}^{n} \sum_{\{I \subseteq \{1,\ldots,n\} : |I| = i\}} (-1)^{i+1} E(\wedge_{j \in I} X_j). \quad (5)$$

Since it is immediate that $X_{U \cup V} = X_U \vee X_V$ and $X_{U \cap V} = X_U \wedge X_V$, (5) generalizes B3. We recover B3 from (5) by taking X_i to be the indicator function X_{U_i}.

There is yet another property satisfied by expectation functions based on belief functions. Two gambles X and Y are said to be *comonotonic* if it is not the case that one increases while the other decreases. Formally, this means that there do not exist worlds w and w' such that $(X(w) - X(w'))(Y(w) - Y(w')) < 0$. The property satisfied by expectation based on belief functions is called *comonotonic additivity*:

$$\text{If } X \text{ and } Y \text{ are comonotonic, then } E(X+Y) = E(X) + E(Y). \qquad (6)$$

The fact that E_{Bel} satisfies this property was essentially recognized by Dellacherie [1970].

PROPOSITION 2.8. *The function E_{Bel} is superadditive, positively affinely homogeneous, monotone, and satisfies (5) and (6).*[6]

PROOF. See Appendix A.1. □

THEOREM 2.9. *Suppose that E is an expectation function that is positively affinely homogeneous, monotone, and satisfies (5) and (6). Then there is a (necessarily unique) belief function* Bel *such that $E = E_{\mathrm{Bel}}$.*

PROOF. See Appendix A.1. □

Schmeidler [1986] proved that expectation functions for nonadditive measures are characterized by positive affine homogeneity, monotonicity, and comonotonicity. Theorem 2.9 shows that the inclusion-exclusion property (5) is what distinguishes expectation for belief functions from expectation for arbitrary nonadditive measures. Note that superadditivity was not assumed in the statement of Theorem 2.9. Indeed, it is a consequence of Theorem 2.9 that superadditivity follows from the other properties. In fact, the full stength of positive affine homogeneity is not needed either in Theorem 2.9. It suffices to assume that $E(\tilde{b}) = b$.

COROLLARY 2.10. *Given a belief function* Bel, *E_{Bel} is the unique expectation function E that is superadditive, positively affinely homogeneous, monotone, and satisfies (5) and (6) such that $E(X_U) = \mathrm{Bel}(U)$ for all $U \subseteq W$.*

PROOF. Proposition 2.8 shows that E_{Bel} has the required properties. If E' is an expectation function that has these properties, by Theorem 2.9, $E' = E_{\mathrm{Bel}'}$ for some belief function Bel'. Since $E'(X_U) = \mathrm{Bel}'(U) = \mathrm{Bel}(U)$ for all $U \subseteq W$, it follows that $\mathrm{Bel} = \mathrm{Bel}'$. □

Corollary 2.10 is somewhat surprising. While it is almost immediate that an additive, affinely homogeneous expectation function (the type that arises from a probability measure) is determined by its behavior on indicator functions, it is not at all obvious that a superadditive, positively affine homogeneous expectation function should be determined by its behavior on indicator functions. In fact, in general it is not; the inclusion-exclusion property is essential. Corollary 2.10 says that Bel and E_{Bel} contain the same information. Thus, so do $(\mathcal{P}_{\mathrm{Bel}})_*$ and $\underline{E}_{\mathcal{P}_{\mathrm{Bel}}}$ (since $\mathrm{Bel} = (\mathcal{P}_{\mathrm{Bel}})_*$ and $E_{\mathrm{Bel}} = \underline{E}_{\mathcal{P}_{\mathrm{Bel}}}$). However, this is not true for arbitrary sets $\mathcal{P}$ of probability measures, as the following example shows.

EXAMPLE 2.11. Let $W = \{1, 2, 3\}$. A probability measure μ on W can be characterized by a triple (a_1, a_2, a_3), where $\mu(i) = a_i$. Let $\mathcal{P}$ consist of the three

[6]In a preliminary version of this paper [Halpern and Pucella 2002b] we used a different characterization of expectation functions, based on belief functions. The current characterization is much simpler.

probability measures $(0, 3/8, 5/8)$, $(5/8, 0, 3/8)$, and $(3/8, 5/8, 0)$. It is almost immediate that $\mathcal{P}_*$ is 0 on singleton subsets of W and $\mathcal{P}_* = 3/8$ for doubleton subsets. Let $\mathcal{P}' = \mathcal{P} \cup \{\mu_4\}$, where $\mu_4 = (5/8, 3/8, 0)$. It is easy to check that $\mathcal{P}'_* = \mathcal{P}_*$. However, $\underline{E}_{\mathcal{P}} \neq \underline{E}_{\mathcal{P}'}$. In particular, let X be the gamble such that $X(1) = 1$, $X(2) = 2$, and $X(3) = 3$. Then $\underline{E}_{\mathcal{P}}(X) = 13/8$ but $\underline{E}_{\mathcal{P}'}(X) = 11/8$. Thus, although $\underline{E}_{\mathcal{P}}$ and $\underline{E}_{\mathcal{P}'}$ agree on indicator functions, they do not agree on all gambles. In light of the discussion above, it should be no surprise that $\mathcal{P}_*$ is not a belief function. ∎

2.4 Expectation for Possibility Measures

A *possibility measure* Poss is a function from subsets of W to $[0, 1]$ such that

Poss1. $\mathrm{Poss}(\varnothing) = 0$.

Poss2. $\mathrm{Poss}(W) = 1$.

Poss3. $\mathrm{Poss}(U \cup V) = \max(\mathrm{Poss}(U), \mathrm{Poss}(V))$ if U and V are disjoint.

It is not hard to show that Poss3 implies that $\mathrm{Poss}(U \cup V) = \max(\mathrm{Poss}(U), \mathrm{Poss}(V))$ even when U and V are not disjoint.

It is well known [Dubois and Prade 1982] that possibility measures are special cases of plausibility functions. Thus, (4) can be used to define a notion of possibilistic expectation; indeed, this has been done in the literature [Dubois and Prade 1987]. It is also straightforward to see from Poss3 that the expectation function E_{Poss} defined from a possibility measure Poss in this way satisfies the following *max property* defined in terms of indicator functions:

$$E_{\mathrm{Poss}}(X_{U \cup V}) = \max(E_{\mathrm{Poss}}(X_U), E_{\mathrm{Poss}}(X_V)). \tag{7}$$

Proposition 2.12. *The function E_{Poss} is positively affinely homogeneous, monotone, and satisfies (6) and (7).*

Proof. See Appendix A.1. □

Theorem 2.13. *Suppose that E is an expectation function that is positively affinely homogeneous, monotone, and satisfies (6) and (7). Then there is a (necessarily unique) possibility measure* Poss *such that $E = E_{\mathrm{Poss}}$.*

Proof. See Appendix A.1. □

Note that, although Poss is a plausibility measure, and thus satisfies the analogue of (5) with $\geq$ replaced by $\leq$, there is no need to state (5) explicitly; it follows from (7). Moreover, just as with expectation for belief functions, it follows from the other properties that E_{Poss} is subadditive. (Since a possibility measure is a plausibility function, not a belief function, the corresponding expectation function is subadditive rather than superadditive.)

3. A LOGIC FOR REASONING ABOUT EXPECTATION

We now consider a logic for reasoning about expectation. To set the stage, we briefly review the FHM logic for reasoning about likelihood.

3.1 Reasoning about Likelihood

The syntax of the FHM logic is straightforward. Fix a set $\Phi_0 = \{p_1, p_2, \ldots\}$ of *primitive propositions*. The choice of primitive propositions is application dependent. The set of primitive propositions could include statements such as "the patient has cancer" and "the patient has fever" if we are reasoning in a medical domain, or statements such as "the price of IBM stock is over \$80" if we are considering a financial domain. The set Φ of *propositional formulas* is the closure of Φ_0 under $\wedge$ and $\neg$. We can define $\vee$ and $\Rightarrow$ in the usual way; we use the operators freely throughout the paper. We assume a special propositional formula *true*, and abbreviate $\neg$*true* as *false*. A *basic likelihood formula* has the form $a_1\ell(\varphi_1) + \cdots + a_k\ell(\varphi_k) \geq b$, where $a_1, \ldots, a_k, b$ are integers and $\varphi_1, \ldots, \varphi_k$ are propositional formulas.[7] The ℓ stands for *likelihood*. Thus, a basic likelihood formula talks about a linear combination of likelihood terms of the form $\ell_i(\varphi)$. A *likelihood formula* is a Boolean combination of basic likelihood formulas. Let $\mathcal{L}^{QU}$ be the language consisting of likelihood formulas. (The QU stands for *quantitative uncertainty*. The name for the logic is taken from Halpern [2003].)

We use standard abbreviations such as $-\ell(\varphi)$ for $(-1)\ell(\varphi)$, and formulas $\ell(\varphi_1) \geq \ell(\varphi_2)$ for $\ell(\varphi_1) - \ell(\varphi_2) \geq 0$. In addition, we write $a_1\ell(\varphi_1) + \ldots + a_k\ell(\varphi_k) \leq b$ for $-a_1\ell(\varphi_1) - \ldots - a_k\ell(\varphi_k) \geq -b$, $a_1\ell(\varphi_1) + \ldots + a_k\ell(\varphi_k) > b$ for $\neg(a_1\ell(\varphi_1) + \ldots + a_k\ell(\varphi_k) \leq b)$, $a_1\ell(\varphi_1) + \ldots + a_k\ell(\varphi_k) < b$ for $-a_1\ell(\varphi_1) - \ldots - a_k\ell(\varphi_k) > -b$, and $a_1\ell(\varphi_1) + \ldots + a_k\ell(\varphi_k) = b$ for $(a_1\ell(\varphi_1) + \ldots + a_k\ell(\varphi_k) \geq b) \wedge (a_1\ell(\varphi_1) + \ldots + a_k\ell(\varphi_k) \leq b)$.

The semantics of $\mathcal{L}^{QU}$ depends on how ℓ is interpreted. In FHM, it is interpreted as a probability measure and as a belief function; in HP, it is interpreted as an upper probability (determined by a set of probability measures). Depending on the interpretation, $\ell(\varphi)$ is the probability of φ (i.e., more precisely, the probability of the set of worlds where φ is true), the belief in φ, etc. For example, in the case of probability, define a *probability structure* to be a tuple $M = (W, \mathcal{F}, \mu, \pi)$, where W is a (possibly infinite) set of worlds, μ is a probability measure whose domain is the algebra $\mathcal{F}$ of subsets of W, and π is an *interpretation*, which associates with each state (or world) in W a truth assignment on the primitive propositions in Φ_0. Thus, $\pi(s)(p) \in \{\mathbf{true}, \mathbf{false}\}$ for $s \in W$ and $p \in \Phi_0$. We require that primitive propositions be measurable, that is, that $\{s \in W : \pi(s)(p) = \mathbf{true}\} \in \mathcal{F}$ for all $p \in \Phi_0$. Extend $\pi(s)$ to a truth assignment on all propositional formulas in the standard way, and associate with each propositional formula the set $[\![\varphi]\!]_M = \{s \in W : \pi(s)(\varphi) = \mathbf{true}\}$. Note that $[\![\varphi]\!]_M$ is measurable for all φ since $\mathcal{F}$ is an algebra. Then

$$M \models a_1\ell(\varphi_1) + \cdots + a_n\ell(\varphi_n) \geq b \text{ iff } a_1\mu([\![\varphi_1]\!]_M) + \cdots + a_n\mu([\![\varphi_n]\!]_M) \geq b.$$

The semantics of Boolean combinations of basic likelihood formulas is given in the

[7] As observed in FHM, we gain no further generality by allowing the coefficients $a_1, \ldots, a_k$ to be rational numbers, since for any formula with rational coefficients, we can easily find an equivalent formula with coefficients that are integers by clearing the dominator. There is no difficulty giving semantics to formulas where the coefficients are arbitrary real numbers (and, indeed, this is what we did in a preliminary version of the paper [Halpern and Pucella 2002b]), but allowing real numbers causes problems in the complexity results. (In the preliminary version, we restricted to integer coefficients for the complexity results.)

obvious way.

We can similarly give semantics to ℓ using lower (or upper) probability. Define a *lower probability structure* to be a tuple $M = (W, \mathcal{F}, \mathcal{P}, \pi)$, W, $\mathcal{F}$ and π are, as before, a (possibly infinite) set of worlds, an algebra of subsets of W, and an interpretation that makes each primitive proposition measurable, and $\mathcal{P}$ is a set of probability measures over $\mathcal{F}$. Likelihood is interpreted as lower probability in lower probability structures:[8]

$$M \models a_1\ell(\varphi_1) + \cdots + a_n\ell(\varphi_n) \geq b \text{ iff } a_1\mathcal{P}_*([\![\varphi_1]\!]_M) + \cdots + a_n\mathcal{P}_*([\![\varphi_n]\!]_M) \geq b.$$

A *belief structure* has the form $M = (W, \mathrm{Bel}, \pi)$, where Bel is a belief function. We can interpret likelihood formulas with respect to belief structures in the obvious way. Similarly, a *possibility structure* has the form $M = (W, \mathrm{Poss}, \pi)$, where Poss is a possibility measure. Again, we interpret likelihood formulas with respect to possibility structures in the obvious way.

Let $\mathcal{M}^{prob}$, $\mathcal{M}^{lp}$, $\mathcal{M}^{bel}$, and $\mathcal{M}^{poss}$ denote the set of all probability structures, lower probability structures, belief structures, and possibility structures, respectively.

3.2 Reasoning about Expectation

Our logic for reasoning about expectation is similar in spirit to $\mathcal{L}^{QU}$. The idea is to interpret a propositional formula φ as the indicator function $X_{[\![\varphi]\!]_M}$, which is 1 in worlds where φ is true, and 0 otherwise. We can then take linear combinations of such gambles. Formally, we again start with a set Φ_0 of primitive propositions. A *(linear) propositional gamble* has the form $b_1\varphi_1 + \cdots + b_n\varphi_n$, where $b_1, \ldots, b_n$ are integers and $\varphi_1, \ldots, \varphi_n$ are propositional formulas that mention only the primitive propositions in Φ_0. We use γ to represent propositional gambles. An *expectation inequality* is a statement of the form $a_1e(\gamma_1) + \cdots + a_ke(\gamma_k) \geq b$, where $a_1 \ldots, a_k$ are integers, $k \geq 1$, and b is an integer. An *expectation formula* is a Boolean combination of expectation inequalities. We use f and g to represent expectation formulas. We define $\leq$, $>$, $<$, and $=$ just as in Section 3.1, Let $\mathcal{L}^E$ be the language consisting of expectation formulas.

Given a model M, we associate with a propositional gamble γ the gamble $\{\!|\gamma|\!\}_M$, where $\{\!|b_1\varphi_1 + \cdots + b_n\varphi_n|\!\}_M = b_1X_{[\![\varphi_1]\!]_M} + \cdots + b_nX_{[\![\varphi_n]\!]_M}$. Of course, the intention is to interpret $e(\gamma)$ in M as the expected value of the gamble $\{\!|\gamma|\!\}_M$, where the notion of "expected value" depends on the underlying semantics. In the case of probability structures, it is probabilistic expectation; in the case of belief structures, it is expected belief; in the case of lower probability structures, it is lower expectation; and so on. For example, if $M \in \mathcal{M}^{prob}$, then

$$M \models a_1e(\gamma_1) + \cdots + a_ke(\gamma_k) \geq b \text{ iff } a_1E_\mu(\{\!|\gamma_1|\!\}_M) + \cdots + a_kE_\mu(\{\!|\gamma_k|\!\}_M) \geq b.$$

Again, Boolean combinations are defined in the obvious way. We leave the obvious semantic definitions in the case of belief structures and lower probability structures to the reader.

[8] In HP, we interpreted likelihood as upper probability. We interpret it here as lower probability to bring out the connections to belief, which is an instance of lower probability. It is easy to translate from upper probabilities to lower probabilities and vice versa, since $\mathcal{P}_*(U) = 1 - \mathcal{P}^*(\overline{U})$.

4. EXPRESSIVE POWER

It is easy to see that $\mathcal{L}^E$ is at least as expressive as $\mathcal{L}^{QU}$. Since the expected value of an indicator function is its likelihood, for all the notions of likelihood we are considering, replacing all occurrences of $\ell(\varphi)$ in a formula in $\mathcal{L}^{QU}$ by $e(\varphi)$ gives an equivalent formula in $\mathcal{L}^E$. Is $\mathcal{L}^E$ strictly more expressive than $\mathcal{L}^{QU}$? That depends on the underlying semantics.

In the case of probability, it is easy to see that it is not. Using additivity and affine homogeneity, it is easy to take an arbitrary formula $f \in \mathcal{L}^E$ and find a formula $f' \in \mathcal{L}^E$ that is equivalent to f (with respect to structures in $\mathcal{M}^{prob}$) such that e is applied only to propositional formulas. Then using the equivalence of $e(\varphi)$ and $\ell(\varphi)$, we can find a formula $f^T \in \mathcal{L}^{QU}$ equivalent to f with respect to structures in $\mathcal{M}^{prob}$. It should be clear that the translation f to f^T causes at most a linear blowup in the size of the formula.

The same is true if we interpret formulas with respect to $\mathcal{M}^{bel}$ and $\mathcal{M}^{poss}$. In both cases, given a formula $f \in \mathcal{L}^E$, we can use (6) to find a formula $f' \in \mathcal{L}^E$ equivalent to f such that e is applied only to propositional formulas (see Lemma A.5 in the appendix). It is then easy to find a formula $f^T \in \mathcal{L}^{QU}$ equivalent to f' with respect to structures in $\mathcal{M}^{bel}$ and $\mathcal{M}^{poss}$. However, now the translation from f to f^T can cause an exponential blowup in the size of the formula; we do not know if there is a shorter translation.

What about lower expectation/probability? In this case, $\mathcal{L}^E$ is strictly more expressive than $\mathcal{L}^{QU}$. It is not hard to construct two structures in $\mathcal{M}^{lp}$ that agree on all formulas in $\mathcal{L}^{QU}$ but disagree on formulas in $\mathcal{L}^E$ such as $e(p+q) > 1/2$. That means that there cannot be a formula in $\mathcal{L}^{QU}$ equivalent to $e(p+q) > 1/2$.

The following theorem summarizes this discussion.

THEOREM 4.1. *$\mathcal{L}^E$ and $\mathcal{L}^{QU}$ are equivalent in expressive power with respect to $\mathcal{M}^{prob}$, $\mathcal{M}^{bel}$, and $\mathcal{M}^{poss}$. $\mathcal{L}^E$ is strictly more expressive than $\mathcal{L}^{QU}$ with respect to $\mathcal{M}^{lp}$.*

PROOF. See Appendix A.2. □

5. AXIOMATIZING EXPECTATION

In FHM, a sound and complete axiomatization is provided for $\mathcal{L}^{QU}$ both with respect to $\mathcal{M}^{prob}$ and $\mathcal{M}^{bel}$; in HP, a sound and complete axiomatization is provided for $\mathcal{L}^{QU}$ with respect to $\mathcal{M}^{lp}$. Here we provide a sound and complete axiomatization for $\mathcal{L}^E$ with respect to these structures, as well as with respect to $\mathcal{M}^{poss}$.

The axiomatization for $\mathcal{L}^{QU}$ given in FHM splits into three parts, dealing respectively with propositional reasoning, reasoning about linear inequalities, and reasoning about likelihood. We follow the same pattern here. The following axioms characterize propositional reasoning:

Taut. All instances of propositional tautologies in the language $\mathcal{L}^E$.

MP. From f and $f \Rightarrow g$ infer g.

Instances of **Taut** include all formulas of the form $f \vee \neg f$, where f is an expectation formula. We could replace **Taut** by a simple collection of axioms that characterize propositional reasoning (see, for example, Mendelson [1964]), but we have chosen to focus on aspects of reasoning about expectations.

The following axiom characterizes reasoning about linear inequalities:

Ineq. All instances in $\mathcal{L}^E$ of valid formulas about linear inequalities.

This axiom is taken from FHM. There, an inequality formula is taken to be a Boolean combination of formulas of the form $a_1x_1 + \cdots + a_nx_n \geq c$, over variables $x_1, \ldots, x_n$. Such a formula is valid if the resulting inequality holds under every possible assignment of real numbers to variables. To get an instance of **Ineq**, we replace each variable x_i that occurs in a valid formula about linear inequalities by a primitive expectation term of the form $e(\gamma_i)$ (naturally each occurrence of the variable x_i must be replaced by the same primitive expectation term $e(\gamma_i)$). As with **Taut**, we can replace **Ineq** by a sound and complete axiomatization for Boolean combinations of linear inequalities. One such axiomatization is given in FHM. It is described in Section 7; the details do not matter for the discussion in this section.

The following axioms characterize probabilistic expectation in terms of the properties described in Proposition 2.1.

E1. $e(\gamma_1 + \gamma_2) = e(\gamma_1) + e(\gamma_2)$,

E2. $e(a\varphi) = ae(\varphi)$ for all $a \in \mathbb{R}$,

E3. $e(\mathit{false}) = 0$,

E4. $e(\mathit{true}) = 1$,

E5. $e(\gamma_1) \leq e(\gamma_2)$ if $\gamma_1 \leq \gamma_2$ is an instance of a valid formula about propositional gamble inequality (see below).

Axiom **E1** is simply additivity of expectations. Axioms **E2**, **E3**, and **E4**, in conjunction with additivity, capture affine homogeneity. Axiom **E5** captures monotonicity. A propositional gamble inequality is a formula of the form $\gamma_1 \leq \gamma_2$, where γ_1 and γ_2 are propositional gambles. Examples of valid propositional gamble inequalities are $p = p \wedge q + p \wedge \neg q$, $\varphi \leq \varphi + \psi$, and $\varphi \leq \varphi \vee \psi$. We define the semantics of gamble inequalities more carefully in Section 7, where we provide a complete axiomatization for them. As in the case of **Ineq**, we can replace **E5** by a sound and complete axiomatization for Boolean combinations of gamble inequalities.[9]

Let $\mathbf{AX}^{prob}$ be the axiomatization $\{\mathbf{Taut}, \mathbf{MP}, \mathbf{Ineq}, \mathbf{E1}, \mathbf{E2}, \mathbf{E3}, \mathbf{E4}, \mathbf{E5}\}$. As usual, given an axiom system **AX**, we say that a formula f is **AX**-*provable* if it can be proved in finitely many steps using the axioms and rules of inferences of **AX**. **AX** is sound with respect to a class $\mathcal{M}$ of structures if every **AX**-provable formula is valid in $\mathcal{M}$. **AX** is *complete* with respect to $\mathcal{M}$ if every formula that is valid in $\mathcal{M}$ is **AX**-provable.

THEOREM 5.1. $\mathbf{AX}^{prob}$ *is a sound and complete axiomatization of* $\mathcal{L}^E$ *with respect to* $\mathcal{M}^{prob}$.

PROOF. See Appendix A.3. □

[9] We could have taken a more complex language that contains both expectation formulas and gamble inequalities. We could then merge the axiomatizations for expectation formulas and gamble inequalities. For simplicity, and to clarify the relationship between reasoning about expectation versus reasoning about likelihood (see Section 4), we consider only the restricted language in this paper.

Despite the fact that we allow structures that have infinitely many worlds, we do not have axioms capturing the continuity properties of expectations that hold for expectations over infinite spaces. Roughly speaking, this is because the logic cannot distinguish between infinite sample spaces and finite ones. We make this intuition precise in Section 6.

The characterizations of Theorems 2.4 and 2.9 suggest the appropriate axioms for reasoning about lower expectations and expected beliefs. The following axioms capture the properties specified in Proposition 2.3:

E6. $e(\gamma_1 + \gamma_2) \geq e(\gamma_1) + e(\gamma_2)$,

E7. $e(a\gamma + b\ \mathit{true}) = ae(\gamma) + b$, for all $a, b \in \mathbb{R}$, $a \geq 0$,

E8. $e(a\gamma + b\ \mathit{false}) = ae(\gamma)$, for all $a, b \in \mathbb{R}$, $a \geq 0$.

Axiom **E6** is superadditivity of the expectation. Axioms **E7** and **E8** capture positive affine homogeneity. Note that because we do not have additivity, we cannot get away with simpler axioms as in the case of probability. Monotonicity is captured, as in the case of probability measures, by axiom **E5**. Let $\mathbf{AX}^{lp}$ be the axiomatization $\{\mathbf{Taut}, \mathbf{MP}, \mathbf{Ineq}, \mathbf{E5}, \mathbf{E6}, \mathbf{E7}, \mathbf{E8}\}$.

THEOREM 5.2. *$\mathbf{AX}^{lp}$ is a sound and complete axiomatization of $\mathcal{L}^E$ with respect to $\mathcal{M}^{lp}$.*

PROOF. See Appendix A.3. □

Although it would seem that Theorem 5.2 should follow easily from Proposition 2.3, this is, unfortunately, not the case. As usual, soundness is straightforward, and to prove completeness, it suffices to show that if a formula f is consistent with $\mathbf{AX}^{lp}$, it is satisfiable in a structure in $\mathcal{M}^{lp}$. Indeed, it suffices to consider formulas f that are conjunctions of expectation inequalities and their negations. However, the usual approach for proving completeness in modal logic, which involves considering maximal consistent sets and canonical structures does not work. The problem is that there are maximal consistent sets of formulas that are not satisfiable. For example, there is a maximal consistent set of formulas that includes $e(\gamma) > 0$ and $e(\gamma) \leq 1/n$ for $n = 1, 2, \ldots$; this is clearly unsatisfiable. A similar problem arises in the completeness proofs for $\mathcal{L}^{QU}$ given in FHM and HP, but the techniques used there do not seem to suffice for dealing with expectations.

Of course, it is the case that any expectation function that satisfies the constraints in the formula f and also every instance of axioms **E6**, **E7**, and **E8** must be a lower expectation, by Theorem 2.4. The problem is that, *a priori*, there are infinitely many relevant instances of the axioms. To get completeness, we must reduce this to a finite number of instances of these axioms. It turns out that this can be done, using techniques from linear programming and Walley's [1991] notion of *natural extension*.

It is also worth noting that, although $\mathcal{L}^E$ is a more expressive language than $\mathcal{L}^{QU}$ in the case of lower probability/expectation, the axiomatization for $\mathcal{L}^E$ in this case is much more elegant than the corresponding axiomatization for $\mathcal{L}^{QU}$ given in HP.

We next consider expectation with respect to belief. As expected, the axioms capturing the interpretation of belief expectation rely on the properties pointed out in Proposition 2.8. Stating these properties in the logic requires a way to express the max and min of two propositional gambles. It turns out that we can view the notation $\gamma_1 \vee \gamma_2$ as an abbreviation for a more complex expression. Given a propositional gamble $\gamma = b_1\varphi_1 + \cdots + b_n\varphi_n$, we construct an equivalent gamble γ' as follows. First define a family ρ_A of propositional formulas indexed by $A \subseteq \{1, \ldots, n\}$ by taking $\rho_A = \bigwedge_{i\in A} \varphi_i \wedge (\bigwedge_{j\notin A} \neg\varphi_j)$. Thus, ρ_A is true exactly if the φ_i's for $i \in A$ are true, and the other φ_j's are false. Note that the formulas ρ_A are mutually exclusive. Define b_A for $A \subseteq \{1, \ldots, n\}$ by taking $b_A = \sum_{i\in A} b_i$. Define $\gamma' = \sum_{A\subseteq\{1,\ldots,n\}} b_A\rho_A$. It is easy to check that the propositional gambles γ and γ' are equal. Given two propositional gambles, say γ_1 and γ_2, we can assume without loss of generality that the involve the same primitive propositions $\varphi_1, \ldots, \varphi_n$. (If not, we can always add "dummy" terms of the form 0ψ.) Form the gambles γ_1' and γ_2' as above. Since all the formulas mentioned in γ_1' and γ_2' are mutually exclusive, it follows that $\max(\gamma_1', \gamma_2') = \sum_{A\subseteq\{1,\ldots,n\}} \max(b_A, b_A')\rho_A$. We take $\gamma_1 \vee \gamma_2$ to be an abbreviation for this gamble. (Note that if γ_1 and γ_2 are propositional formulas, then $\gamma_1 \vee \gamma_2$ really is a gamble equivalent to the propositional formula $\gamma_1 \vee \gamma_2$, given our identification of propositional formulas with indicator functions, so the use of $\vee$ is justified here.) Of course, we can similarly define $\gamma_1 \wedge \gamma_2$, simply by taking min instead of max.

With these definitions, the following axiom accounts for property (5):

E9. $e(\gamma_1 \vee \cdots \vee \gamma_n) \geq \sum_{i=1}^{n} \sum_{\{I\subseteq\{1,\ldots,n\}:|I|=i\}} (-1)^{i+1} e(\bigwedge_{j\in I} \gamma_j)$.

To deal with the comonotonic additivity property (6), it seems that comonotonicity must be expresssed in the logic. It turns out that it suffices to capture only a restricted form of comonotonicity. Note that if $\varphi_1, \ldots, \varphi_m$ are pairwise mutually exclusive, $a_1 \leq \ldots \leq a_m$, $b_1 \leq \ldots \leq b_m$, $\gamma_1 = a_1\varphi_1 + \cdots + a_m\varphi_m$, and $\gamma_2 = b_1\varphi_1 + \cdots + b_m\varphi_m$, then in all structures M, the gambles $\{\!|\gamma_1|\!\}_M$ and $\{\!|\gamma_2|\!\}_M$ are comonotonic. Thus, by (6), it follows that $E_{\mathrm{Bel}}(\{\!|\gamma_1 + \gamma_2|\!\}_M) = E_{\mathrm{Bel}}(\{\!|\gamma_1|\!\}_M) + E_{\mathrm{Bel}}(\{\!|\gamma_2|\!\}_M)$. The proof that E_{Bel} satisfies comonotonic additivity (see the proof of Theorem 2.9 in the appendix) shows that it suffices to restrict to gambles of this form. These observations lead to the following axiom:

E10. $e(\gamma_1+\gamma_2) = e(\gamma_1)+e(\gamma_2)$ if $\gamma_1 = a_1\varphi_1+\cdots+a_m\varphi_m$, $\gamma_2 = b_1\varphi_1+\cdots+b_m\varphi_m$, $a_1 \leq \ldots \leq a_m$, $b_1 \leq \ldots \leq b_m$, and $\neg(\varphi_i \wedge \varphi_j)$ is a propositional tautology for all $i \neq j$.

Let $\mathbf{AX}^{bel}$ be the axiomatization $\{\mathbf{Taut}, \mathbf{MP}, \mathbf{Ineq}, \mathbf{E5}, \mathbf{E7}, \mathbf{E8}, \mathbf{E9}, \mathbf{E10}\}$.

THEOREM 5.3. *$\mathbf{AX}^{bel}$ is a sound and complete axiomatization of $\mathcal{L}^E$ with respect to $\mathcal{M}^{bel}$.*

PROOF. See Appendix A.3. □

Finally, we consider expectation with respect to possibility. The axioms capturing the interpretation of possibilistic expectation E_{Poss} rely on the properties pointed out in Proposition 2.12. The max property (7) is captured by the following axiom:

E11. $(e(\varphi_1) \geq e(\varphi_2)) \Rightarrow (e(\varphi_1 \vee \varphi_2) = e(\varphi_1))$.

E11 essentially says that $e(\varphi_1 \vee \varphi_2) = \max(e(\varphi_1), e(\varphi_2))$.

Let $\mathbf{AX}^{poss}$ be the axiomatization $\{\mathbf{Taut}, \mathbf{MP}, \mathbf{Ineq}, \mathbf{E5}, \mathbf{E7}, \mathbf{E8}, \mathbf{E10}, \mathbf{E11}\}$.

THEOREM 5.4. *$\mathbf{AX}^{poss}$ is a sound and complete axiomatization of $\mathcal{L}^E$ with respect to $\mathcal{M}^{poss}$.*

PROOF. See Appendix A.3. □

6. DECISION PROCEDURES

In FHM, it was shown that the satisfiability problem for $\mathcal{L}^{QU}$ is NP-complete, both with respect to $\mathcal{M}^{prob}$ and $\mathcal{M}^{bel}$; in HP, NP-completeness was also shown with respect to $\mathcal{M}^{lp}$. Here we prove similar results for the language $\mathcal{L}^E$. In the case of $\mathcal{M}^{prob}$, this is not at all surprising, given Theorem 4.1 and the fact that the translation from $\mathcal{L}^E$ to $\mathcal{L}^{QU}$ causes only a linear blowup in the case of $\mathcal{M}^{prob}$. However, we cannot get the result for $\mathcal{M}^{bel}$ or $\mathcal{M}^{poss}$ from Theorem 4.1, since the translation causes an exponential blowup. Of course, in the case of $\mathcal{M}^{lp}$, no translation exists at all. Nevertheless, in all these cases, we can get NP-completeness using techniques very much in the spirit of the linear programming techniques used in FHM.

We define $|f|$ to be the length of f, that is, the number of symbols required to write f, where each coefficient is counted as one symbol. Define $||f||$ to be the length of the longest coefficient appearing in f, when written in binary. The size of a rational number $\frac{a}{b}$, denoted $||\frac{a}{b}||$, where a and b are relatively prime, is defined to be $||a|| + ||b||$.

The next theorems, helpful to establish the complexity of the decision procedures, show that the logic is not expressive enough to distinguish between infinite sample spaces and infinite ones; if a formula is satisfiable at all, it is satisfiable in a (relatively small) finite structure. In other words, we could have restricted to structures with only finitely many worlds without loss of generality. We prove this result for $\mathcal{M}^{lp}$, $\mathcal{M}^{bel}$, and $\mathcal{M}^{poss}$; it follows from Fagin et al. [1990, Theorem 2.4] (which shows that a formula in the language $\mathcal{L}^{QU}$ is satisfiable in $\mathcal{M}^{prob}$ iff it is satisfiable in a finite structure in $\mathcal{M}^{prob}$) and Theorem 4.1 (which shows that every formula in $\mathcal{L}^E$ is equivalent to a formula in $\mathcal{L}^{QU}$ for structures in $\mathcal{M}^{prob}$) that the result holds for $\mathcal{M}^{prob}$ as well.

THEOREM 6.1. *Suppose that $f \in \mathcal{L}^E$ is satisfied in some structure in $\mathcal{M}^{lp}$. Then f is satisfied in a structure $(W, \mathcal{P}, \pi)$ such that $|W| \leq |f|^2$, $|\mathcal{P}| \leq |f|$, $\mu(w)$ is a rational number such that $||\mu(w)||$ is $O(|f|^2||f|| + |f|^2 \log(|f|))$ for every world $w \in W$ and $\mu \in \mathcal{P}$, and $\pi(w)(p) =$ **false** for every world $w \in W$ and every primitive proposition p not appearing in f.*

PROOF. See Appendix A.4. □

To prove a small-model theorem for belief functions (and possibility measures), we need a representation of belief functions that is particularly compact. Given a set W of worlds, a *mass function* m is a function from 2^W to $[0, 1]$ such that $m(\varnothing) = 0$, and $\sum_{U \subseteq W} m(U) = 1$ [Shafer 1976]. As is well known, in finite spaces, there is a one-to-one correspondence between belief functions and mass functions [Shafer 1976]. Given a mass function m, let Bel_m be the set function defined by setting $\mathrm{Bel}_m(U) = \sum_{V \subseteq U} m(V)$. It can be shown that Bel_m is a belief function (in particular, it satisfies B3). Conversely, given a belief function Bel, consider the set function m_{Bel} defined by setting $m_{\mathrm{Bel}}(\varnothing) = 0$, and $m_{\mathrm{Bel}}(U) = \mathrm{Bel}(U) - \sum_{V \subset U} m_{\mathrm{Bel}}(V)$. It can be shown that m_{Bel} is a mass function (in particular, $m_{\mathrm{Bel}}(U) \geq 0$ for all U) and that $\mathrm{Bel}_{m_{\mathrm{Bel}}} = \mathrm{Bel}$. It easily follows that the mappings $\mathrm{Bel} \mapsto m_{\mathrm{Bel}}$ and $m \mapsto \mathrm{Bel}_m$ give a one-to-one correspondence between belief function and mass functions.

Theorem 6.2. *Suppose that $f \in \mathcal{L}^E$ is satisfied in some structure in $\mathcal{M}^{bel}$ (resp., $\mathcal{M}^{poss}$). Then f is satisfied in a structure (W, ν, π) such that $|W| \leq |f|^2$, ν is a belief function (resp., possibility measure) whose corresponding mass function is positive on at most $|f|$ subsets of W and the mass of each of these $|f|$ sets is a rational number of size $O(|f|\,||f|| + |f|\log(|f|))$, and $\pi(w)(p) =$ **false** for every world $w \in W$ and every primitive proposition p not appearing in f.*

Proof. See Appendix A.4. □

Theorems 6.1 and 6.2 yield the following complexity result for the decision problem for $\mathcal{L}^E$.

Theorem 6.3. *The problem of deciding whether a formula in $\mathcal{L}^E$ is satisfiable in $\mathcal{M}^{prob}$ (resp., $\mathcal{M}^{lp}$, $\mathcal{M}^{bel}$, $\mathcal{M}^{poss}$) is NP-complete.*

Proof. See Appendix A.4. □

7. REASONING ABOUT GAMBLE INEQUALITIES

Ineq presumes an oracle for all valid formulas about linear inequalities; as we said earlier, this oracle can be replaced by the complete axiomatization of linear inequalities provided in FHM. Similarly, **E5** assumes an oracle for valid formulas about gamble inequalities. In this section, we provide an axiomatization that can replace that oracle. One of the axioms involves reasoning about linear inequalities over real-valued functions, so we axiomatize this as well.

Let $\mathcal{L}^g$ consist of all Boolean combinations of gamble inequalities $\gamma \geq \tilde{c}$, where γ is a propositional gamble, as defined in Section 3, and c is an integer. As in Section 3.1, we write $\gamma_1 \leq \gamma_2$ as an abbreviation for $\gamma_1 - \gamma_2 \leq \tilde{0}$, which is the form of gamble inequality used in axiom **E5**. We also use the abbreviations we defined in Section 3.1 for $\leq$, $<$, and $=$. However, we now take $\gamma_1 > \gamma_2$ to be an abbreviation for $\gamma_1 \geq \gamma_2 \wedge \neg(\gamma_1 \leq \gamma_2)$. The analogue of the conjunct $\gamma_1 \geq \gamma_2$ is not necessary when reasoning about likelihood or expectation, since likelihood and expectation terms are interpreted as real numbers, and thus are totally ordered. (For real numbers b_1 and b_2, $\neg(b_1 \leq b_2)$ implies $b_1 > b_2$). However, this is not the case for gamble expressions, which are interpreted as real-valued functions.

We can provide a semantics for gamble formulas by considering structures $M = (W, \pi)$ where W is a nonempty set of worlds and π associates with each world in W a truth assignment on the primitive propositions. Let $\mathcal{M}^g$ be the class all such structures. (Clearly, every structure in $\mathcal{M}^{prob}$, $\mathcal{M}^{lp}$, $\mathcal{M}^{bel}$, and $\mathcal{M}^{poss}$ can be interpreted as a structure in $\mathcal{M}^g$ by simply "forgetting" the uncertainty measure over the worlds. Note that W is nonempty for such structures, since $\mu(W) \neq \mu(\varnothing)$ for all the measures of uncertainty considered.) Then

$$M \models \gamma \geq \tilde{c} \text{ iff for all } w \in W,\ \{\!|\gamma|\!\}_M(w) \geq c.$$

Again, Boolean combinations are given semantics in the obvious way.

We can characterize gamble formulas axiomatically as follows. As before, we have **Taut** and **MP** (although now we consider the instances of valid propositional tautologies in the language $\mathcal{L}^g$). Instead of **Ineq**, we have a similar axiom that captures reasoning about linear inequalities interpreted over *real-valued functions* with nonempty domains:

IneqF. All instances in $\mathcal{L}^g$ of valid formulas about linear inequalities over real-valued functions with nonempty domains.

A formula about linear inequalities over real-valued functions is a Boolean combination of formulas of the form $a_1 v_1 + \cdots a_n v_n \geq \tilde{c}$, where $v_1, \ldots, v_n$ are variables ranging over real-valued functions. This formula is valid if, for every domain X and every assignment for real-valued functions with domain X to $v_1, \ldots, v_n$, the resulting inequality holds when $\tilde{c}$ is interpreted as the constant function with domain X that always returns $c \in \mathbb{R}$ and $\geq$ is the pointwise ordering on functions. To get an instance of **IneqF**, we replace each function variable v_i that occurs in a valid formula about linear inequalities by a propositional formula φ_i. (Again, each occurence of the function variable v_i must be replaced by the same propositional formula φ_i.) As with **Ineq**, we can replace **IneqF** by a sound and complete axiomatization for Boolean combinations of linear inequalities over real-valued functions. We provide one such axiomatization at the end of this section.[10] We also consider the following axioms, which capture properties of indicator functions, such as: for any $w \in W$, $X_U(w) + X_V(w) = X_{U \cup V}(w)$ if $U \cap V = \varnothing$, and $X_U(w) \leq X_V(w)$ if $U \subseteq V$.

G1. $a((\varphi \wedge \psi) \vee (\varphi \wedge \neg\psi)) = a(\varphi \wedge \psi) + a(\varphi \wedge \neg\psi)$,

G2. $a(\varphi \wedge \psi) + b(\varphi \wedge \neg\psi) \geq \tilde{0} \Leftrightarrow a(\varphi \wedge \psi) \geq \tilde{0} \wedge b(\varphi \wedge \neg\psi) \geq \tilde{0}$

G3. $\mathit{true} = \tilde{1}$,

G4. $\varphi - \psi = \tilde{0}$ if $\varphi \Leftrightarrow \psi$ is a propositional tautology.

Intuitively, **G1** and **G3** say that propositional formulas act like indicator functions, and **G2** says that the $\geq$ order is the pointwise order on gambles. Axiom **G4** simply says that equivalent formulas yield the same indicator functions. Let $\mathbf{AX}^g$ be the axiomatization $\{\mathbf{Taut}, \mathbf{MP}, \mathbf{IneqF}, \mathbf{G1}, \mathbf{G2}, \mathbf{G3}, \mathbf{G4}\}$.

THEOREM 7.1. *$\mathbf{AX}^g$ is a sound and complete axiomatization of $\mathcal{L}^g$ with respect to $\mathcal{M}^g$.*

PROOF. See Appendix A.5. □

Finally, we consider the complexity of the decision procedure for $\mathcal{L}^g$. As we did in Section 6, let $\mathcal{L}_1^g$ be the restriction of $\mathcal{L}^g$ to inequalities with integer coefficients. Perhaps not surprisingly, the logic for reasoning about gamble inequalities is NP-complete.

THEOREM 7.2. *The problem of deciding whether a formula of $\mathcal{L}_1^g$ is satisfiable in $\mathcal{M}^g$ is NP-complete.*

[10] In a preliminary version of this paper [Halpern and Pucella 2002b], we mistakenly used an axiomatization that relied on **Ineq**. **Ineq** is inappropriate for reasoning about linear inequalities over real-valued functions because, intuitively, **Ineq** captures properties of inequalities where $\geq$ is a linear order. However, the pointwise ordering on functions is only a partial order. Furthermore, we did not have axioms **G2** and **G3** in that version of the paper; they seem to be necessary.

Proof. See Appendix A.5. □

We now provide a sound and complete axiomatization for reasoning about linear inequalities over real-valued functions. As before, start with **Taut** and **MP**, and consider the following axioms:

F1. $v - v \geq \tilde{0}$,

F2. $(a_1v_1 + \cdots + a_kv_k \geq \tilde{c}) \Leftrightarrow (a_1v_1 + \cdots + a_kv_k + 0v_{k+1} \geq \tilde{c})$,

F3. $(a_1v_1 + \cdots + a_kv_k \geq \tilde{c}) \Leftrightarrow (a_{j_1}x_{j_1} + \cdots + a_{j_k}v_{j_k} \geq \tilde{c})$, if $j_1, \ldots, j_k$ is a permutation of $1, \ldots, k$,

F4. $(a_1v_1 + \cdots + a_kv_k \geq \tilde{c}) \wedge (a'_1v_1 + \cdots + a'_kv_k \geq \tilde{c}')$
$\Rightarrow (a_1 + a'_1)v_1 + \cdots + (a_k + a'_k)v_k \geq \widetilde{(c + c')}$,

F5. $(a_1v_1 + \cdots + a_kv_k \geq \tilde{c}) \Leftrightarrow (da_1v_1 + \cdots + da_kv_k \geq \widetilde{dc})$, if $d > 0$,

F6. $(a_1v_1 + \cdots + a_kv_k \geq \tilde{c}) \Rightarrow (a_1v_1 + \cdots + a_kv_k > \tilde{d})$, if $c > d$.

The axiomatization for Ineq for reasoning about linear inequalities over *reals* presented in FHM is identical to this axiomatization, except that it includes one additional axiom: $x \geq c \vee x \leq c$. (Note that now c represents a real number, as opposed to $\tilde{c}$, which represents a constant real-valued function.) The absence of this axiom is not surprising; although $\geq$ is a total order over the reals, it is only a partial order over real-valued functions.

Let $\mathbf{AX}^f$ be the axiomatization $\{\mathbf{Taut}, \mathbf{MP}, \mathbf{F1}, \mathbf{F2}, \mathbf{F3}, \mathbf{F4}, \mathbf{F5}, \mathbf{F6}\}$.

Theorem 7.3. $\mathbf{AX}^f$ *is sound and complete for reasoning about formulas about linear inequalities over real-valued functions with nonempty domain.*

Proof. See Appendix A.5. □

We can now replace axiom **IneqF** in $\mathbf{AX}^g$ by **F1**–**F6** (since $\mathbf{AX}^g$ already has **Taut** and **MP**). Following the discussion above, this means that we replace the v's appearing in **F1**–**F6** by propositional formulas φ's. Note that axiom **F1** becomes redundant in $\mathbf{AX}^g$, since it becomes *false* $\geq$ *false*, which follows from **G2**.

The complexity of reasoning about linear inequalities over real-valued functions (with integer coefficients, following what was said in Section 6) is NP-complete, just like reasoning about linear inequalities over the reals.

Theorem 7.4. *The problem of deciding whether a formula about linear inequalities (with integer coefficients) is satisfiable over real-valued functions with nonempty domain is NP-complete.*

Proof. See Appendix A.5. □

8. CONCLUSION

The notion of expectation is critical in many areas, ranging from physics to decision theory (where much of modern theory of decision making under risk is based on the notion of maximizing expected utility [Savage 1954]), and theoretical computer science (where it arises, for example, when establishing the correctness and efficiency of randomized algorithms [Motwani and Raghavan 1995]). For the most part, expectation has been considered in the context of probability. However, it has been considered for other representations of uncertainty. For example, various notions of decision making are based on taking expectations with respect to other representations of uncertainty [Schmeidler 1986; Walley 1991]. Indeed, it is possible to consider almost all decision rules as instances of a generalized expected utility, for different representations of uncertainty [Chu and Halpern 2003; 2004].

After reviewing characterizations of expectation functions with respect to various notions of uncertainty (and, in a few cases, providing characterizations when none existed), we considered a propositional logic for reasoning about uncertainty. Such a logic should prove useful to reason about decision making with respect to various notions of uncertainty. Our logic is at least as expressive as the corresponding logic for reasoning about uncertainty for all representations we considered, and in some cases it is more expressive. For each representation of uncertainty, we provided a sound and complete axiomatization for the logic, and showed that the satisfiability problem is NP-complete, as it is for reasoning about probabilibity, lower probabilities, and belief functions. Thus, the added expressiveness of being able to reason about expectation of sets of probability measures comes at no computational-complexity cost. Moreover, the axiomatization is, in some cases, much more elegant.

A. PROOFS

A.1 Proofs for Section 2

Proposition 2.6 defines the expectation of X in terms of the values $\{x_1, \ldots, x_n\}$ that make up $\mathcal{V}(X)$. The following lemma, which turns out to be useful in a number of technical results, shows that we can use the same calculation for expectation for any set of values that includes $\mathcal{V}(X)$. The lemma holds for all the representations of uncertainty ν considered in this paper. It uses only the fact that $\nu(W) = 1$ and $\nu(\varnothing) = 0$ for all representations that we are considering.

LEMMA A.1. *Let X be a gamble, let $s_1 \leq \ldots \leq s_m$ be reals such that $\mathcal{V}(X) \subseteq \{s_1, \ldots, s_m\}$, and let ν be either a probability measure, belief function, or possibility measure. Then $E_\nu(X) = s_1 + (s_2 - s_1)\nu(X > s_1) + \cdots + (s_m - s_{m-1})\nu(X > s_{m-1})$.*

PROOF. Clearly, it suffices to show that the result holds for values such that $s_1 < \ldots < s_m$. Indeed, if $s_i = s_{i+1}$ for some i, then clearly the term $(s_{i+1} - s_i)\nu(X > s_i)$ does not contribute to the sum. For a set $S = \{s_1, \ldots, s_m\}$ such that $S \supseteq \mathcal{V}(X)$ and $s_1 < \ldots < s_m$, define

$$E_\nu^S(X) = s_1 + (s_2 - s_1)\nu(X > s_1) + \cdots + (s_m - s_{m-1})\nu(X > s_{m-1}).$$

We proceed by induction on $|S - \mathcal{V}(X)|$. If $|S - \mathcal{V}(X)| = 0$, then $S = \mathcal{V}(X)$. Observe that $E_\nu(X)$ is just $E_\nu^{\mathcal{V}(X)}(X)$, so the base case is immediate. For the inductive step, suppose that $S \supseteq \mathcal{V}(X)$, $|S - \mathcal{V}(X)| = k+1$, and the claim holds for all sets S' such that $S' \supseteq \mathcal{V}(X)$ and $|S' - \mathcal{V}(X)| = k$. Choose $s \in S - \mathcal{V}(X)$ and let $S' = S - \{s\}$. We show that $E_\nu^S(X) = E_\nu^{S'}(X)$. There are three cases.

(1) $s = \min(S)$: Since $s \notin \mathcal{V}(X)$, we must have $(X > s) = W$. Thus,

$$\begin{aligned} & E_\nu^S(X) \\ = \ & s + (s_1 - s)\nu(X > s) + (s_2 - s_1)\nu(X > s_1) + \cdots + (s_n - s_{n-1})\nu(X > s_{n-1}) \\ = \ & s + (s_1 - s) + (s_2 - s_1)\nu(X > s_1) + \cdots + (s_n - s_{n-1})\nu(X > s_{n-1}) \\ = \ & s_1 + (s_2 - s_1)\nu(X > s_1) + \cdots + (s_n - s_{n-1})\nu(X > s_{n-1}) \\ = \ & E_\nu^{S'}(X). \end{aligned}$$

(2) $s = \max(S)$. Let $s_n = \max(S')$. Since $s \notin \mathcal{V}(X)$, $(X > s_n) = \varnothing$. Thus,

$$\begin{aligned} & E_\nu^S(X) \\ = \ & s_1 + (s_2 - s_1)\nu(X > s_1) + \cdots + (s_n - s_{n-1})\nu(X > s_{n-1}) + (s - s_n)\nu(X > s_n) \\ = \ & s_1 + (s_2 - s_1)\nu(X > s_1) + \cdots + (s_n - s_{n-1})\nu(X > s_{n-1}) \\ = \ & E_\nu^{S'}(X). \end{aligned}$$

(3) $S = \{s_1, \ldots, s_k, s, s_{k+1}, \ldots, s_n\}$ for some $s_1, \ldots, s_n$, such that $s_1 < \ldots < s_k < s < s_{k+1} < \ldots < s_n$. Since $s \notin \mathcal{V}(X)$ and $s > s_k$, we have $(X > s) = (X > s_k)$. Thus,

$$\begin{aligned} & (s_{k+1} - s_k)\mu(X > s_k) = (s_{k+1} - s)\mu(X > s_k) + (s - s_k)\mu(X > s_k) \\ & = (s_{k+1} - s)\mu(X > s) + (s - s_k)\mu(X > s_k). \end{aligned}$$

It is immediate that $E_\nu^S(X) = E_\nu^{S'}(X)$.

This completes the proof of the inductive step. $\square$

Lemma A.2. *The functions E_{Bel} and E_{Plaus} satisfy (6).*

Proof. Suppose that X and Y are comonotonic gambles. We claim that there exist pairwise disjoint sets $U_1, \ldots, U_n$ and reals $a_1, \ldots, a_n, b_1, \ldots, b_n$ such that

(a) $X = a_1 X_{U_1} + \cdots + a_n X_{U_n}$,
(b) $Y = b_1 X_{U_1} + \cdots + b_n X_{U_n}$,
(c) $a_1 \leq \ldots \leq a_n$, and
(d) $b_1 \leq \ldots \leq b_n$.

To see this, suppose that $\mathcal{V}(X) = \{a'_1, \ldots, a'_m\}$, where $a'_1 < \ldots < a'_m$, and $\mathcal{V}(Y) = \{b'_1, \ldots, b'_{m'}\}$, where $b'_1 < \ldots < b'_{m'}$. Define $U'_i = X^{-1}(a'_i)$ for $i = 1, \ldots, m$, and $U''_j = Y^{-1}(b'_j)$ for $j = 1, \ldots, m'$. Then clearly (a) $X = a'_1 X_{U'_1} + \cdots + a'_m X_{U'_m}$, (b) $Y = b'_1 X_{U''_1} + \cdots + b'_{m'} X_{U''_{m'}}$, (c) the sets $U'_1, \ldots, U'_m$ are pairwise disjoint, (d) the sets $U''_1, \ldots, U''_{m'}$ are pairwise disjoint, and (e) $W = \cup_{i=1}^m U'_i = \cup_{i=1}^{m'} U''_i$. Let $V_{ij} = U'_i \cap U''_j$. Note that if V_{ij} and $V_{i'j'}$ are both nonempty and $i < i'$, then $j \leq j'$. For suppose not; then $i < i'$ and $j > j'$. Thus, if $w \in V_{ij}$ and $w' \in V_{i'j'}$, then $(X(w) - X(w'))(Y(w) - Y(w')) = (a'_i - a'_{i'})(b'_j - b'_{j'}) < 0$, contradicting the comonotonicity of X and Y. It follows that we can take the claimed sets $U_1, \ldots, U_n$ to be the nonempty sets V_{ij} ordered lexicographically (so that if $U_k = V_{ij}$, $U_{k'} = V_{i'j'}$, and $k < k'$, then either $i < i'$ or $i = i'$ and $j < j'$). Moreover, if $U_k = V_{ij}$, take $a_k = a'_i$ and $b_k = b'_j$. It is clear that these choices have the required properties.

For this choice of $U_1, \ldots, U_n$, it follows that $X + Y = (a_1 + b_1)X_{U_1} + \cdots + (a_n + b_n)X_{U_n}$, with $a_1 + b_1 \leq \ldots \leq a_n + b_n$. By construction, we have $\mathcal{V}(X + Y) =$

$\{a_1 + b_1, \ldots, a_n + b_n\}$. By Lemma A.1, we have the following, where ν is either Bel or Plaus:

$$\begin{array}{ll} & E_\nu(X+Y) \\ = & (a_1+b_1) + (a_2+b_2-a_1-b_1)\nu(X+Y > a_1+b_1) + \cdots + \\ & \quad (a_n+b_n-a_{n-1}-b_{n-1})\nu(X+Y > a_{n-1}+b_{n-1}) \\ = & (a_1+b_1) + (a_2+b_2-a_1-b_1)\nu(U_2 \cup \ldots \cup U_n) + \cdots + \\ & \quad (a_n+b_n-a_{n-1}-b_{n-1})\nu(U_n) \\ = & a_1 + (a_2-a_1)\nu(U_2 \cup \ldots \cup U_n) + \cdots + (a_n-a_{n-1})\nu(U_n) + \\ & \quad b_1 + (b_2-b_1)\nu(U_2 \cup \ldots \cup U_n) + \cdots + (b_n-b_{n-1})\nu(U_n) \\ = & a_1 + (a_2-a_1)\nu(X > a_1) + \cdots + (a_n-a_{n-1})\nu(X > a_{n-1}) + \\ & \quad b_1 + (b_2-b_1)\nu(Y > b_1) + \cdots + (b_n-b_{n-1})\nu(Y > b_{n-1}) \\ = & E_\nu(X) + E_\nu(Y), \end{array}$$

where the last equality follows from Lemma A.1 and the fact that $\mathcal{V}(X) = \{a_1, \ldots, a_n\}$ and $\mathcal{V}(Y) = \{b_1, \ldots, b_n\}$. □

Proposition 2.8. *The function E_{Bel} is superadditive, positively affinely homogeneous, monotone, and satisfies (5) and (6).*

Proof. The fact that E_{Bel} is superadditive, positively affinely homogeneous, and monotone follows immediately from Theorem 2.5 and Proposition 2.3. The fact that E_{Bel} satisfies (5) follows essentially from property B3 of belief functions and Proposition 2.6. First, it is easily checked that for gambles X and Y, $X \vee Y > x = (X > x) \cup (Y > x)$ and $X \wedge Y > x = (X > x) \cap (Y > x)$. Consider the gamble $X_1 \vee \ldots \vee X_n$. Let $S = \mathcal{V}(X_1) \cup \ldots \cup \mathcal{V}(X_n) = \{s_1, \ldots, s_m\}$, with $s_1 < \ldots < s_m$. Let s_0 be an arbitrary real less than every number in S. Clearly, $\mathcal{V}(X_1 \vee \ldots \vee X_n) \subseteq S$. Moreover, for every subset $I \subseteq \{1, \ldots, n\}$, $\mathcal{V}(\wedge_{i \in I} X_i) \subseteq S$. By Proposition 2.6 and Lemma A.1, we have:

$$\begin{array}{ll} & E_{\mathrm{Bel}}(X_1 \vee \ldots \vee X_n) \\ = & s_1 + \sum_{k=1}^m (s_k - s_{k-1})\mathrm{Bel}(X_1 \vee \ldots \vee X_n > s_{k-1}) \\ = & s_1 \mathrm{Bel}(X_1 \vee \ldots \vee X_n > s_0) + \sum_{k=1}^m (s_k - s_{k-1})\mathrm{Bel}(X_1 \vee \ldots \vee X_n > s_{k-1}) \\ = & s_1 \mathrm{Bel}(X_1 > s_0 \cup \ldots \cup X_n > s_0) + \\ & \quad \sum_{k=1}^m (s_k - s_{k-1})\mathrm{Bel}(X_1 > s_{k-1} \cup \ldots \cup X_n > s_{k-1}) \\ \geq & s_1 \sum_{i=1}^n \sum_{I \subseteq \{1,\ldots,n\}, |I|=i} (-1)^{i+1} \mathrm{Bel}(\cap_{j \in I} X_j > s_0) + \\ & \quad \sum_{k=1}^m (s_k - s_{k-1}) \sum_{i=1}^n \sum_{I \subseteq \{1,\ldots,n\}, |I|=i} (-1)^{i+1} \mathrm{Bel}(\cap_{j \in I} X_j > s_{k-1}) \\ = & \sum_{i=1}^n \sum_{I \subseteq \{1,\ldots,n\}, |I|=i} (-1)^{i+1} (s_1 \mathrm{Bel}(\cap_{j \in I} X_j > s_0) + \\ & \quad \sum_{k=1}^m (s_k - s_{k-1})\mathrm{Bel}(\cap_{j \in I} X_j > s_{k-1})) \\ = & \sum_{i=1}^n \sum_{I \subseteq \{1,\ldots,n\}, |I|=i} (-1)^{i+1} (s_1 + \sum_{k=1}^m (s_k - s_{k-1})\mathrm{Bel}(\wedge_{j \in I} X_j > s_{k-1})) \\ = & \sum_{i=1}^n \sum_{I \subseteq \{1,\ldots,n\}, |I|=i} (-1)^{i+1} E_{\mathrm{Bel}}(\wedge_{j \in I} X_j). \end{array}$$

Finally, the fact that E_{Bel} satisfies (6) follows from Lemma A.2. □

The following property turns out to be useful in various proofs involving expectation for belief functions.

Lemma A.3. *If E is an expectation function that satisfies (6) and positive affine homogeneity, then it satisfies the following property on indicator functions:*

$$\begin{array}{l} \text{If } U_1 \supseteq U_2 \supseteq \ldots \supseteq U_n \text{ and } a_1, \ldots, a_n \geq 0, \\ \text{then } E(a_1 X_{U_1} + \cdots + a_n X_{U_n}) = a_1 E(X_{U_1}) + \cdots + a_n E(X_{U_n}). \end{array} \tag{8}$$

Proof. Define $S_k = a_1 X_{U_1} + \cdots + a_k X_{U_k}$, for $k = 1, \ldots, n$. Since $a_i > 0$ for $i = 1, \ldots, n$, it is not hard to show that S_k and $a_{k+1} X_{U_{k+1}}$ are comonotonic for all k. For if $S_k(w) > S_k(w')$, since $U_1 \supseteq \ldots \supseteq U_k$, there must be some $j \leq k$ such that $w \in U_j$ and $w' \notin U_j$. Since $U_k \supseteq U_{k+1}$, it must be the case that $w' \notin U_{k+1}$. Hence $a_{k+1} X_{U_{k+1}}(w') = 0$ and $a_{k+1} X_{U_{k+1}}(w) \geq 0$. Thus, we have comonotonicity. Since E satisfies (6), it follows that

$$\begin{aligned} & E(a_1 X_{U_1} + \cdots + a_n X_{U_n}) \\ = \ & E(S_n) = E(S_{n-1} + a_n X_{U_n}) = E(S_{n-1}) + E(a_n X_{U_n}) = \ldots \\ = \ & E(a_1 X_{U_1}) + \cdots + E(a_n X_{U_n}) \end{aligned}$$

The result is now immediate, since E satisfies positive homogeneity. □

Theorem 2.9. *Suppose that E is an expectation function that is positively affinely homogeneous, monotone, and satisfies (5) and (6). Then there is a (necessarily unique) belief function* Bel *such that $E = E_{\mathrm{Bel}}$.*

Proof. Define $\mathrm{Bel}(U) = E(X_U)$. Just as in the case of probability, it follows from positive affine homogeneity and monotonicity that $\mathrm{Bel}(\varnothing) = 0$, $\mathrm{Bel}(W) = 1$, and $0 \leq \mathrm{Bel}(U) \leq 1$ for all $U \subseteq W$. By (5) (specialized to indicator functions), it follows that Bel satisfies B3. Thus, Bel is a belief function. Now if X is a gamble such that $\mathcal{V}(X) = \{x_1, \ldots, x_n\}$ and $x_1 < x_2 < \ldots < x_n$, then it is easy to check that

$$X = x_1 X_W + (x_2 - x_1) X_{X > x_1} + \cdots + (x_n - x_{n-1}) X_{X > x_{n-1}}.$$

Clearly $W \supseteq (X > x_1) \supseteq \ldots \supseteq (X > x_{n-1})$. Thus, by Lemma A.3,

$$\begin{aligned} E(X) &= x_1 E(X_W) + (x_2 - x_1) E(X_{X > x_1}) + \cdots + (x_n - x_{n-1}) E(X_{X > x_{n-1}}) \\ &= x_1 + (x_2 - x_1) \mathrm{Bel}(X > x_1) + \cdots + (x_n - x_{n-1}) \mathrm{Bel}(X > x_{n-1}) \\ &= E_{\mathrm{Bel}}(X). \end{aligned}$$

For uniqueness, observe that if $E_{\mathrm{Bel}} = E_{\mathrm{Bel}'}$, then $\mathrm{Bel}(U) = E_{\mathrm{Bel}}(X_U) = E_{\mathrm{Bel}'}(X_U) = \mathrm{Bel}'(U)$ for all $U \subseteq W$. □

Proposition 2.12. *The function E_{Poss} is positively affinely homogeneous, monotone, and satisfies (6) and (7).*

Proof. Recall from Section 2.4 that a possibility measure is just a plausibility function. Therefore, by Theorem 2.5 and Proposition 2.7, we have $E_{\mathrm{Poss}} = \overline{E}_{\mathcal{P}_{\mathrm{Bel}}}$, where Bel is the belief function whose corresponding plausibility function is Poss. By Proposition 2.3, we immediately have that E_{Poss} is positively affinely homogeneous, and monotone. By Lemma A.2, we have that E_{Poss} satisfies (6). Showing that E_{Poss} satsifies (7) is straightforward: $E_{\mathrm{Poss}}(X_{\cup_i U_i}) = \mathrm{Poss}(\cup_i U_i) = \max_i \mathrm{Poss}(U_i) = \max_i E_{\mathrm{Poss}}(X_{U_i})$. □

Theorem 2.13. *Suppose that E is an expectation function that is positively affinely homogeneous, monotone, and satisfies (6) and (7). Then there is a (necessarily unique) possibility measure* Poss *such that $E = E_{\mathrm{Poss}}$.*

Proof. Define $\mathrm{Poss}(U) = E(X_U)$. Just as in the case of probability and belief functions, it follows from positive affine homogeneity and monotonicity that $\mathrm{Poss}(\varnothing) = 0$, $\mathrm{Poss}(W) = 1$, and $0 \leq \mathrm{Poss}(U) \leq 1$ for all $U \subseteq W$. By (7), it follows that Poss satisfies Poss3. Thus, Poss is a possibility measure. The proof that $E = E_{\text{Poss}}$ is the same as that of Theorem 2.9, which is not surprising since Poss is a plausibility function. More precisely, let X be a gamble such that $\mathcal{V}(X) = \{x_1, \ldots, x_n\}$ with $x_1 < \ldots < x_n$, and write X as $x_1 X_W + (x_2 - x_1)X_{X > x_1} + \cdots + (x_n - x_{n-1})X_{X > x_{n-1}}$. Since $W \supseteq (X > x_1) \supseteq \ldots \supseteq (X > x_{n-1})$, we can apply Lemma A.3 to get

$$\begin{aligned} E(X) &= x_1 E(X_W) + (x_2 - x_1)E(X_{X>x_1}) + \cdots + (x_n - x_{n-1})E(X_{X>x_{n-1}}) \\ &= x_1 + (x_2 - x_1)\text{Poss}(X > x_1) + \cdots + (x_n - x_{n-1})\text{Poss}(X > x_{n-1}) \\ &= E_{\text{Poss}}(X). \end{aligned}$$

□

A.2 Proofs for Section 4

The following transformations on formulas of $\mathcal{L}^E$ will be used in the proofs of Theorems 4.1, 5.1, 5.3, and 5.4. We prove here that these transformations preserve satisfiability of the formulas, with respect to the appropriate structures.

First, define for a formula f of $\mathcal{L}^E$, the transformation f^{T_1} inductively on the structure of f, taking $(a_1 e(\gamma_1) + \cdot + a_n e(\gamma_n) \geq b)^{T_1}$ to be $t_1 + \cdots + t_n \geq b$, where t_i is obtained from a_i and $\gamma_i = c_{1,i}\varphi_{1,i} + \cdots c_{k,i}\varphi_{k,i}$ as $a_i c_{1,i} e(\varphi_{1,i}) + \cdots + a_i c_{k,i} e(\varphi_{k,i})$. We define $(f_1 \wedge f_2)^{T_1}$ and $(\neg f_1)^{T_1}$ as $f_1^{T_1} \wedge f_2^{T_1}$ and $\neg(f_1^{T_1})$, respectively.

Lemma A.4. *If $M \in \mathcal{M}^{prob}$, then $M \models f$ if and only if $M \models f^{T_1}$.*

Proof. We proceed by induction on the structure of f. For the base case,

$$\begin{array}{ll} & M \models a_1 e(\gamma_1) + \cdots \geq b \\ \text{iff} & a_1 E_\mu(X_{\gamma_1}) + \cdots \geq b \\ \text{iff} & a_1 E_\mu(c_{1,1} X_{[\![\varphi_{1,1}]\!]_M} + \cdots + c_{k,1} X_{[\![\varphi_{k,1}]\!]_M}) + \cdots \geq b \\ \text{iff} & a_1 c_{1,1} E_\mu(X_{[\![\varphi_{1,1}]\!]_M}) + \cdots + a_1 c_{k,1} E_\mu(X_{[\![\varphi_{k,1}]\!]_M}) + \cdots \geq b \\ \text{iff} & M \models a_1 c_{1,1} e(\varphi_{1,1}) + \cdots + a_1 c_{k,1} e(\varphi_{k,1}) + \cdots \geq b. \end{array}$$

The inductive cases are immediate. □

The second transformation is a bit more complicated. As before, we define a transformation taking a formula f of $\mathcal{L}^E$ to a formula f^{T_2} by induction on the structure of f. Intuitively, we translate every expectation term $e(\gamma)$ appearing in f into an expectation term of the form given by Proposition 2.6, namely $d_0 + (d_1 - d_0)e(\psi_1) + \ldots + (d_m - d_{m-1})e(\psi_m)$ for formulas $\psi_1, \ldots, \psi_m$. As before, take $(f_1 \wedge f_2)^{T_2}$ and $(\neg f_1)^{T_2}$ to be $f_1^{T_2} \wedge f_2^{T_2}$ and $\neg(f_1^{T_2})$, respectively. Take $(a_1 e(\gamma_1) + \cdots + a_n e(\gamma_n) \geq b)^{T_2}$ to be $t_1 + \cdots + t_n \geq b$, where t_i is obtained from a_i and $\gamma_i = c_{1,i}\varphi_{1,i} + \cdots + c_{k,i}\varphi_{k,i}$ as follows. Let $p_1, \ldots, p_N$ be the primitive propositions that appear in f. Let an *atom over* $p_1, \ldots, p_N$ be a formula of the form $q_1 \wedge \ldots \wedge q_N$, where q_i is either p_i or $\neg p_i$. There are clearly 2^N atoms over $p_1, \ldots, p_N$. Let $\delta_1, \ldots, \delta_{2^N}$ be an arbitrary ordering of these atoms. It is easy to see that atoms are pairwise disjoint. It is also easy to see that any formula over $p_1, \ldots, p_N$ can be written in a unique way as a disjunction of atoms. Using these atoms, we construct a gamble γ_i' equivalent to γ_i as follows. For $j \in \{1, \ldots, 2^N\}$, let $c_j = \sum_{\{l \,:\, \delta_j \Rightarrow \varphi_{l,i} \text{ is a prop. tautology}\}} c_{l,i}$. It is straightforward to check that $\gamma_i = \sum_{j=1}^{2^n} c_j \delta_j$. Let $d_0, \ldots, d_m$ be the distinct elements of $\{c_j\}_{j \in \{1, \ldots, 2^N\}}$, with $d_0 <$

$\ldots < d_m$. For $j \in \{1,\ldots,m\}$, let $\psi_j = \bigvee_{l\in\{1,\ldots,2^N\},c_l\geq d_j} \delta_l$. Note that ψ_0 is logically equivalent to *true*; moreover, $\psi_{j+1} \Rightarrow \psi_j$ is a propositional tautology for all $j \in \{0,\ldots,m-1\}$. Finally, let $\gamma_i' = d_0+(d_1-d_0)\psi_1+\cdots+(d_m-d_{m-1})\psi_m$. One can check that γ_i and γ_i' are equivalent gambles. Indeed, γ_i' is provably equivalent to $d_0(\sum_{l=1}^{2^N}\delta_l)+(d_1-d_0)(\sum_{c_l\geq d_1}\delta_l)+\ldots+(d_m-d_{m-1})(\sum_{c_l\geq d_m}\delta_l)$. Distributing the coefficients and putting like terms together, we see that the coefficient of δ_l is just $d_0+(d_1-d_0)+\ldots+(d_{r_\ell}-d_{r_\ell-1}) = d_{r_\ell}$, where $r_\ell = \max\{r : c_l \geq d_r\}$. By definition of $d_1,\ldots,d_m$, $d_{r_\ell} = c_l$, and therefore γ_i and γ_i' are equivalent. Note that, by Lemma A.3, if E is an expectation function that satisfies comonotonicity and positive homogeneity, then $E(X_{\gamma_i'}) = d_1 + (d_1-d_0)E(X_{[\![\psi_1]\!]_M})+\cdots+(d_m-d_{m-1})E(X_{[\![\psi_m]\!]_M})$. Let t_i be $a_i d_0 + a_i(d_1-d_0)e(\psi_1)+\cdots+a_i(d_m-d_{m-1})e(\psi_m)$.

Lemma A.5. *If $M \in \mathcal{M}^{prob}\cup\mathcal{M}^{bel}\cup\mathcal{M}^{poss}$, then $M \models f$ if and only if $M \models f^{T_2}$.*

Proof. We proceed by induction on the structure of f, for $M \in \mathcal{M}^{bel}$; the result follows for structures in $\mathcal{M}^{prob}$ and $\mathcal{M}^{poss}$, since these can be viewed as structures in $\mathcal{M}^{bel}$. For the base case, we have

$$\begin{array}{ll}
 & M \models a_1 e(\gamma_1)+\cdots a_n e(\gamma_n) \geq b \\
\text{iff} & a_1 E_{\mathrm{Bel}}(X_{\gamma_1})+\cdots a_n E_{\mathrm{Bel}}(X_{\gamma_n}) \geq b \\
\text{iff} & a_1 E_{\mathrm{Bel}}(X_{\gamma_1'})+\cdots+ a_n E_{\mathrm{Bel}}(X_{\gamma_n'}) \geq b \\
\text{iff} & a_1 E_{\mathrm{Bel}}(d_{0,1}+(d_{1,1}-d_{0,1})X_{[\![\psi_{1,1}]\!]_M}+\cdots+(d_{m_1,1}-d_{m_1-1,1})X_{[\![\psi_{m_1,1}]\!]_M})+ \\
 & \quad \cdots + a_n E_{\mathrm{Bel}}(d_{0,n}+(d_{1,n}-d_{0,n})X_{[\![\psi_{1,n}]\!]_M}+\cdots+ \\
 & \qquad (d_{m_n,n}-d_{m_n-1,n})X_{[\![\psi_{m_n,n}]\!]_M}) \geq b \\
\text{iff} & a_1 d_{0,1} + a_1(d_{1,1}-d_{0,1})E_{\mathrm{Bel}}(X_{[\![\psi_{1,1}]\!]_M})+\cdots+ \\
 & \quad a_1(d_{m_1,1}-d_{m_1-1,1})E_{\mathrm{Bel}}(X_{[\![\psi_{m_1,1}]\!]_M})+\cdots+ \\
 & \qquad a_n d_{0,n} + a_1(d_{1,n}-d_{0,n})E_{\mathrm{Bel}}(X_{[\![\psi_{n,1}]\!]_M})+\cdots+ \\
 & \qquad\quad a_n(d_{m_n,n}-d_{m_n-1,n})E_{\mathrm{Bel}}(X_{[\![\psi_{m_n,n}]\!]_M}) \geq b \\
\text{iff} & M \models a_1 d_{0,1} + a_1(d_{1,1}-d_{0,1})e(\psi_{1,1})+\cdots+a_1(d_{m,1}-d_{m-1,1})e(\varphi_{m,1})+\cdots+ \\
 & \quad a_n d_{0,n} + a_n(d_{1,n}-d_{0,n})e(\psi_{1,n})+\cdots a_n(d_{m,n}-d_{m-1,n})e(\varphi_{m,n}) \geq b.
\end{array}$$

The fact that we can distribute the expectation E_{Bel} of the gamble into a sum of expectation of indicator functions follows from comonotonicity, using Lemma A.3.

The inductive cases are immediate. □

Theorem 4.1. *$\mathcal{L}^E$ and $\mathcal{L}^{QU}$ are equivalent in expressive power with respect to $\mathcal{M}^{prob}$, $\mathcal{M}^{bel}$, and $\mathcal{M}^{poss}$. $\mathcal{L}^E$ is strictly more expressive than $\mathcal{L}^{QU}$ with respect to $\mathcal{M}^{lp}$.*

Proof. The proof follows the lines laid out in Section 4. First, we show that $\mathcal{L}^E$ is at least as expressive as $\mathcal{L}^{QU}$ with respect to $\mathcal{M}^{prob}$, $\mathcal{M}^{bel}$, $\mathcal{M}^{lp}$, and $\mathcal{M}^{poss}$. Formally, we show that every formula f of $\mathcal{L}^{QU}$ is equivalent to the formula $f^T \in \mathcal{L}^E$ that results by replacing $\ell(\varphi)$ by $e(\varphi)$. Define f^T by induction on the structure of f, taking $(a_1\ell(\varphi_1)+\cdots+a_n\ell(\varphi_n) \geq b)^T$ to be $a_1 e(\varphi_1)+\cdots+a_n\ell(\varphi_n) \geq b$, $(f_1 \wedge f_2)^T$ to be $f_1^T \wedge f_2^T$, and $(\neg f_1)^T$ to be $\neg(f^T)$.

It suffices to show that for all M in $\mathcal{M}^{prob}$, $\mathcal{M}^{bel}$, $\mathcal{M}^{lp}$, or $\mathcal{M}^{poss}$, we have $M \models f$ if and only if $M \models f^T$. First, note if ν is a probability measure, belief function, or set of probability measures, then $E_\nu(X_U) = \nu(U)$, where X_U is the indicator function of U. Let M be an arbitrary structure in $\mathcal{M}^{prob}$, $\mathcal{M}^{bel}$, $\mathcal{M}^{lp}$, or $\mathcal{M}^{poss}$, with associated measure of uncertainty μ. Thus, if ν is the measure associated with structure M, we have that $\nu([\![\varphi]\!]_M) = E_\nu(\{\!|X_{[\![\varphi]\!]_M}|\!\}_M)$. (If ν is the set $\mathcal{P}$ or probability measures, we take $E_\nu = \underline{E}_{\mathcal{P}}$.) The result now follows immediately by a straightforward induction on the structure of f.

We next show that $\mathcal{L}^{QU}$ is as expressive as $\mathcal{L}^E$ with respect to $\mathcal{M}^{prob}$, $\mathcal{M}^{bel}$, and $\mathcal{M}^{poss}$. Let $\mathcal{L}^{E'}$ be the sublanguage of $\mathcal{L}^E$ where the expectation operator e is applied only to propositional formulas, as opposed to arbitrary gambles. It is immediate from the construction in the previous paragraph that if $f \in \mathcal{L}^{QU}$, then $f^T \in \mathcal{L}^{E'}$. Moreover, it is easy to see that every formula $f' \in \mathcal{L}^{E'}$ is of the form f^T (where f is the result of replacing all occurrences of $e(\varphi)$ in f' by $\ell(\varphi)$). Since we showed above that f is equivalent to f^T, to prove that $\mathcal{L}^{QU}$ is as expressive to $\mathcal{L}^E$ with respect to $\mathcal{M}^{prob}$, $\mathcal{M}^{bel}$, and $\mathcal{M}^{poss}$, it suffices to show that every formula in $\mathcal{L}^E$ is equivalent to a formula in $\mathcal{L}^{E'}$ in structures in $\mathcal{M}^{prob} \cup \mathcal{M}^{bel} \cup \mathcal{M}^{poss}$. But this is immediate from Lemma A.5, since the formula $f^{T_2} \in \mathcal{L}^{E'}$.

Finally, we show that $\mathcal{L}^E$ is strictly more expressive than $\mathcal{L}^{QU}$ with respect to $\mathcal{M}^{lp}$. To establish this, it is sufficient to exhibit two structures in $\mathcal{M}^{lp}$ that satisfy the same formulas in $\mathcal{L}^{QU}$, but can be distinguished by a formula f in $\mathcal{L}^E$. (This means that f in $\mathcal{L}^E$ cannot be equivalent to any formula in $\mathcal{L}^{QU}$.) We use a variant of Example 2.11. Consider the two $\mathcal{M}^{lp}$ structures $M_1 = (W, \mathcal{P}_1, \pi)$ and $M_2 = (W, \mathcal{P}_2, \pi)$, where $W = \{w_1, w_2, w_3\}$, π is such that p is true at w_1, both p and q are true at w_2, and neither p nor q is true at w_3, and the sets of probability measures $\mathcal{P}_1$ and $\mathcal{P}_2$ are given as follows. We can describe a probability measure on W by a tuple (a_1, a_2, a_3), where a_i gives the probability at w_i. Define $\mathcal{P}_1 = \{(1/3, 2/3, 0), (0, 1/3, 2/3), (2/3, 0, 1/3)\}$, and $\mathcal{P}_2 = \mathcal{P}_1 \cup \{(1/3, 0, /2, 3)\}$. It is easy to check that $(\mathcal{P}_1)_*$ and $(\mathcal{P}_2)_*$ are both 0 on singleton subsets of W and 1/3 on doubleton subsets. Hence, $(\mathcal{P}_1)_* = (\mathcal{P}_2)_*$. Therefore, no formula involving only lower probability can distinguish these two structures, and M_1 and M_2 satisfy the same formulas of $\mathcal{L}^{QU}$. Now, consider the gamble $p + q$. It is easy to see that $\{\!|p+q|\!\}_{M_1} = \{\!|p+q|\!\}_{M_2}$ is the random variable X defined by $X(w_1) = 1$, $X(w_2) = 2$, and $X(w_3) = 0$. It is not hard to check that $\underline{E}_{\mathcal{P}_1}(X) = 2/3$, and $\underline{E}_{\mathcal{P}_2}(X) = 1/3$. Hence, $M_1 \models e(p+q) > 1/2$, and $M_2 \models e(p+q) < 1/2$, so that $M_2 \models \neg e(p+q) > 1/2$. $\square$

A.3 Proofs for Section 5

The proofs of Theorems 5.1–5.4 use an approach similar to the one taken in FHM. We first need some definitions. We say a formula σ is *consistent* with an axiom system **AX** (or simply **AX**-consistent) if $\neg\sigma$ is not provable from **AX**. To show that **AX** is a complete axiomatization with respect to some class of structure $\mathcal{M}$, we must show that every formula that is valid in every structure in $\mathcal{M}$ is provable in **AX**. This is equivalent to showing that every **AX**-consistent formula σ is satisfiable in $\mathcal{M}$.

The proofs of Theorems 5.1–5.4 are all structured as follows. Soundness is straightforward in all cases. Completeness is obtained by showing that a consistent formula f is satisfiable in the appropriate set of structures $\mathcal{M}$. More specifically, assume that formula f is not satisfiable in a structure in $\mathcal{M}$; we must show that f is inconsistent. We first reduce f to a canonical form. Let $g_1 \vee \ldots \vee g_r$ be a disjunctive normal form expression for f (where each g_i is a conjunction of expectation inequalities and their negations). Using propositional reasoning, we can show that f is provably equivalent to this disjunction. Since f is unsatisfiable, each g_i must also be unsatisfiable. Thus, it is sufficient to show that any unsatisfiable conjunction of expectation inequalities and their negations is inconsistent. Let f be such a formula. To show that it is inconsistent, we essentially construct a system of inequalities $\hat{f}$ from f by replacing every term $e(\gamma_i)$ in f by a variable x_i, with the property that f is satisfiable over structures in $\mathcal{M}$ if and only if $\hat{f}$ has a solution over the reals. Since f is unsatisfiable, $\hat{f}$ has no solution, so that $\neg f$ must be an instance of **Ineq**. (That $\hat{f}$ has no solution means that $\hat{f}$ is not satisfiable as a formula about linear inequalities, so that $\neg\hat{f}$ is a valid formula about linear inequalities.) Therefore, $\neg f$ is provable, and f is inconsistent. The details of how to construct $\hat{f}$ differ for each representation of uncertainty.

THEOREM 5.1. *$\mathbf{AX}^{prob}$ is a sound and complete axiomatization of $\mathcal{L}^E$ with respect to $\mathcal{M}^{prob}$.*

PROOF. Soundness is straightforward. For completeness, we proceed as above. Without loss of generality, assume that f is a conjunction of expectation inequalities and their negations. Using axioms **E1** and **E2**, we can convert f into the equivalent formula f^{T_1} (Lemma A.4) where e is applied only to propositional formulas. For every propositional formula φ in f^{T_1}, φ is equivalent to $\vee_{j=1}^{k}\delta_{i_j}$, where $\delta_{i_1},\ldots,\delta_{i_k}$ are the atoms over the propositions in f such that $\delta_{i_j} \Rightarrow \varphi$ for all $1 \leq j \leq k$. Since $\neg(\delta_{i_j} \wedge \delta_{i_l})$ is a propositional tautology for all $j \neq l$, $\varphi = \delta_{i_1} + \cdots + \delta_{i_k}$ is a valid formula about propositional gamble inequalities (see Section 7), and axiom **E5** yields $e(\varphi) = e(\delta_{i_1} + \cdots + \delta_{i_k})$. This means that we can find a formula f' provably equivalent to f^{T_1}, where f' is formed by replacing each term $ae(\varphi)$ of f^{T_1} by the sum $ae(\delta_{i_1}) + \cdots + ae(\delta_{i_k})$, and then collecting like terms. Let f'' be obtained from f' by adding as conjuncts to f' all the expectation inequality formulas $e(\delta_j) \geq 0$, for $1 \leq j \leq 2^N$, $e(\delta_1) + \cdots + e(\delta_{2^N}) \geq 1$, and $-e(\delta_1) - \cdots - e(\delta_{2^N}) \geq -1$ (which together essentially say that the sum of the probabilities of the atoms is 1). It is not hard to see that these formulas are provable, hence f'' is provably equivalent to f', and hence to f. We therefore only have to show that f'' is satisfiable.

The negation of an expectation inequality $a_1e(\gamma_1) + \cdots + a_ne(\gamma_n) \geq b$ can be written $-a_1e(\gamma_1) - \cdots - a_ne(\gamma_n) > -b$. Thus, without loss of generality, we can assume that f'' is the conjunction of the formulas

$$\begin{aligned}
e(\delta_1) + \cdots + e(\delta_{2^N}) &\geq 1 \\
-e(\delta_1) - \cdots - e(\delta_{2^N}) &\geq -1 \\
e(\delta_1) &\geq 0 \\
&\vdots \\
e(\delta_{2^N}) &\geq 0 \\
a_{1,1}e(\delta_1) + \cdots + a_{1,2^N}e(\delta_{2^N}) &\geq b_1 \\
&\vdots \\
a_{r,1}e(\delta_1) + \cdots + a_{r,2^N}e(\delta_{2^N}) &\geq b_r \\
-a'_{1,1}e(\delta_1) - \cdots - a'_{1,2^N}e(\delta_{2^N}) &> b'_1
\end{aligned}$$

$$\vdots$$
$$-a'_{s,1}e(\delta_1) - \cdots - a'_{s,2^N}e(\delta_{2^N}) > b'_s.$$

Consider the following system of inequations, obtained by replacing $e(\delta_i)$ by the variable x_i:

$$\begin{aligned}
x_1 + \cdots + x_{2^N} &\geq 1 \\
-x_1 - \cdots - x_{2^N} &\geq -1 \\
x_1 &\geq 0 \\
&\vdots \\
x_{2^N} &\geq 0 \\
a_{1,1}x_1 + \cdots + a_{1,2^N}x_{2^N} &\geq b_1 \\
&\vdots \\
a_{r,1}x_1 + \cdots + a_{r,2^N}x_{2^N} &\geq b_r \\
-a'_{1,1}x_1 - \cdots - a'_{1,2^N}x_{2^N} &> b'_1 \\
&\vdots \\
-a'_{s,1}x_1 - \cdots - a'_{s,2^N}x_{2^N} &> b'_s.
\end{aligned}$$

Clearly if f'' is satisfiable, say in a structure M with associated probability measure μ, then the system is satisfiable, by taking $x_i = \mu([\![\delta_i]\!]_M)$. Conversely, if the system of inequations has a solution, then f'' is satisfied in a structure M with associated probability measure μ such that $x_i = \mu([\![\delta_i]\!]_M)$.

Returning to the main argument, suppose by way of contradiction, that f is not satisfiable. Then this system of inequations has no solution. It follows that $\neg f''$ is an instance of the axiom **Ineq**. Since f'' is provably equivalent to f, it follows that $\neg f$ is provable, so that f is inconsistent, a contradiction. □

The proofs of Theorems 5.2, 5.3, and 5.4 require some additional terminology. Let f be a formula of $\mathcal{L}^E$, and let $p_1, \ldots, p_N$ be the primitive propositions that appear in f. Observe that there are 2^{2^N} inequivalent propositional formulas over $p_1, \ldots, p_N$. The argument goes as follows. Recall the notion of an atom over $p_1, \ldots, p_N$, introduced before Lemma A.5. It is easy to see that any formula over $p_1, \ldots, p_N$ can be written in a unique way as a disjunction of atoms. There are 2^{2^N} such disjunctions, so the claim follows. Let $\rho_1, \ldots, \rho_{2^{2^N}}$ be some canonical listing of the inequivalent formulas over $p_1, \ldots, p_N$. Without loss of generality, we assume that ρ_1 is equivalent to *true*, and $\rho_{2^{2^N}}$ is equivalent to *false*. Following FHM, we call these formulas ρ_i, $i = 1, \ldots, 2^{2^N}$ *regions*.

THEOREM 5.2. $\mathbf{AX}^{lp}$ *is a sound and complete axiomatization of* $\mathcal{L}^E$ *with respect to* $\mathcal{M}^{lp}$.

PROOF. Soundness is straightforward. For completeness, we proceed as above. Again, without loss of generality, assume that f is a conjunction of expectation inequalities and their negations.

The first step of the proof is to find a formula f' provably equivalent to f, where the gambles in f' are expressed in terms of atoms $\delta_1, \ldots, \delta_{2^N}$. As observed above, every propositional formula φ in f is equivalent to some region ρ_i, $1 \leq i \leq 2^{2^N}$, which is of the form $\vee_{j=1}^{k} \delta_{i_j}$, where $\delta_{i_1}, \ldots, \delta_{i_k}$ are the atoms over the propositions in f such that $\delta_{i_j} \Rightarrow \varphi$ for all $1 \leq j \leq k$. Since $\delta_{i_j} \wedge \delta_{i_l} \Leftrightarrow \mathit{false}$ is a propositional tautology for all $j \neq l$, $\varphi = \delta_{i_1} + \cdots + \delta_{i_k}$ is valid. It easily follows that a propositional gamble $a_1\varphi_1 + \cdots + a_n\varphi_n$ is equal to a propositional gamble of the form $a'_1\delta_1 + \cdots + a'_{2^N}\delta_{2^N}$. By axiom **E5**, $e(a_1\varphi_1 + \cdots + a_n\varphi_n) = e(a'_1\delta_1 + \cdots + a'_{2^N}\delta_{2^N})$ is provable. This means we can find a formula f' provably equivalent to f, where each term $ae(\gamma)$ of f is replaced by a term $ae(a'_1\delta_1 + \cdots + a'_{2^N}\delta_{2^N})$. In fact, the following arguments shows that, without loss of generality, we can take the a'_is to be nonnegative. Given a gamble $a'_1\delta_1 + \ldots + a'_{2^N}\delta_{2^N}$, let $b = \min_{1 \leq i \leq 2^N}(a'_i)$. We can easily show that $(a'_1 + b)\delta_1 + \ldots + (a'_{2^N} + b)\delta_{2^N} - b = a'_1\delta_1 + \ldots + a'_{2^N}\delta_{2^N}$ is a valid gamble equality. Thus, every expectation inequality formula $c_1e(\gamma_1) + \ldots + c_ke(\gamma_k) \geq d$ in f' is provably equivalent to a formula of the form $c_1e(\gamma'_1) + c_1b_1 + \ldots + c_ke(\gamma'_j) + c_kb_k \geq d$, which in turn is provably equivalent to $c_1e(\gamma'_1) + \ldots + c_ke(\gamma'_k) \geq d - c_1b_1 - \ldots - c_kb_k$. Thus, without loss of generality, we can take the coefficients of the gambles in f' to be nonnegative.

The next step of the proof is to derive a finite system of inequations $\hat{f}$ corresponding to f' such that f' is satisfiable if and only $\hat{f}$ has a solution. Suppose that we have found such a formula $\hat{f}$ and that f' is unsatisfiable. Then $\hat{f}$ has no solution. Thus, the formula $\neg f'$ is an instance of **Ineq**. Since f' is provably equivalent to f, $\neg f$ is provable, and thus f is inconsistent, a contradiction. Therefore, f', and hence f, is satisfiable. The remainder of the proof consists of coming up with $\hat{f}$, and showing that it has a solution if and only if f' is satisfiable.

We would like $\hat{f}$ to be a system of inequalities that yields a solution that can be used to construct a lower expectation function that satisfies f'. We can force a solution to $\hat{f}$ to be a lower expectation function by adding to $\hat{f}$ inequalities corresponding to all the instances of axioms **E6**–**E8**. Unfortunately, there are infinitely many such instances. The aim now is therefore to derive a finite set of instances of **E6**–**E8** that is sufficient to constrain a solution to $\hat{f}$ to be a lower expectation function.

To do this, let $\gamma_1, \ldots, \gamma_n$ be the propositional gambles γ such that $e(\gamma)$ appears in f', together with the propositional gambles *false* and *true* (if they do not already appear in f'). For ease of exposition, take $\gamma_{n-1} = \mathit{false}$ and $\gamma_n = \mathit{true}$. As we saw earlier, each γ_i is provably equivalent to a formula of the form $a_{i,1}\delta_1 + \cdots + a_{i,2^N}\delta_{2^N}$, where $\delta_1, \ldots, \delta_{2^N}$ are the atoms over the primitive propositions appearing in f, and $a_{i,j} \geq 0$. We now construct, for every propositional gamble γ_i $(1 \leq i \leq n-1)$, a finite set B_i of vectors. (We do not need a vector for the gamble $\gamma_n = \mathit{true}$.) We give the construction of B_1, and then describe the minor modifications needed to construct B_i for $i = 2, \ldots, n-1$. We then show how to use these sets to identify a finite set of instances of **E6**–**E8**.

Let A_1 be the $(n-1) \times 2^N$ matrix of real numbers whose ith row consists of the coefficients of $\delta_1, \ldots, \delta_{2^N}$ in γ_{i+1}, $i = 1, \ldots, n-1$. Let B'_1 be the set of vectors $\vec{b} = \langle b_2, \ldots, b_n\rangle$ such that $b_2, \ldots, b_{n-1} \geq 0$ (b_n can be negative) and

$$\vec{b}A_1 \leq \langle a_{1,1}, \ldots, a_{1,2^N}\rangle. \tag{9}$$

It is easy to see that B_1' is a closed convex set: it is clearly closed, and if $\vec{b}_1, \vec{b}_2 \in B_1'$, then so is $a\vec{b}_1 + (1-a)\vec{b}_2$, for all $a \in [0,1]$. Thus, for each vector $\vec{y} = \langle y_1, \ldots, y_{n-1} \rangle$ such that $y_1, \ldots, y_{n-2} \geq 0$, there exists a $\vec{b}_y \in B_1'$ such that $\vec{b}_y \cdot \vec{y}$ is maximal (where $\cdot$ is the inner product). It is a consequence of the Krein-Milman Theorem (see Rudin [1991], for instance) that the set of such $\vec{b}_y$ is a finite subset B_1 of B_1'.

We use the set B_1 to derive a finite number of instances of **E6**–**E8** as follows. Note that if $\vec{b} = \langle b_2, \ldots, b_n \rangle$ satisfies (9), then $b_2\gamma_2 + \cdots + b_n\gamma_n \leq \gamma_1$ is a valid formula about gamble inequalities. Therefore, by **E5**, $e(b_2\gamma_2 + \cdots + b_n\gamma_n) \leq e(\gamma_1)$. By applying **E6**, **E7** and **E8**, we can derive $b_2 e(\gamma_2) + \cdots b_{n-1}e(\gamma_{n-1}) + b_n \leq e(\gamma_1)$. Let $F_{\vec{b}}$ be the conjunction of the instances of axioms **E6**, **E7**, and **E8** that are used to perform this derivation. Let f_1 be the conjunction of $F_{\vec{b}}$ for all $\vec{b} \in B_1$.

This process defines a set B_1 of vectors, and from these a formula f_1. Repeat this process for the gambles $\gamma_2, \ldots, \gamma_{n-1}$, by interchanging the role of γ_1 and γ_i, for $2 \leq i \leq n-1$. After doing this for all gambles, we obtain sets $B_1, \ldots, B_{n-1}$ of vectors, and formulas $f_1, \ldots, f_{n-1}$.

Let f'' be the conjunction of $f_1, \ldots, f_{n-1}, f', e(\gamma_{n-1}) \geq 0, -e(\gamma_{n-1}) \geq 0, e(\gamma_n) \geq 1$, and $-e(\gamma_n) \geq -1$ (that is, $e(\mathit{false}) = 0$ and $e(\mathit{true}) = 1$, since $\gamma_{n-1} = \mathit{false}$ and $\gamma_n = \mathit{true}$, by assumption). Let $\hat{f}$ be the conjunction of inequalities obtained by replacing every instance of a term of the form $e(\gamma)$ in f'' by the variable x_γ. Note that, besides $e(\gamma_1), \ldots, e(\gamma_n)$, there are other terms of the form $e(\gamma)$ that arise in $f_1, \ldots, f_{n-1}$; the gambles γ in these terms are linear combinations of $\gamma_1, \ldots, \gamma_n$. We then obtain a system of inequalities over the variables x_γ. We now show that $\hat{f}$ is our required system of inequalities, that is, f' is satisfiable if and only if $\hat{f}$ has a solution.

It is straightforward to show that if f' is satisfiable, then $\hat{f}$ has a solution. If f' is satisfiable, then there exists a lower probability structure $(W, \mathcal{P}, \pi)$ such that $(W, \mathcal{P}, \pi) \models f'$, and hence $(W, \mathcal{P}, \pi) \models f''$ (since the instances of **E6**–**E8** are valid in lower probability structures). Clearly, taking $x_\gamma = \underline{E}_{\mathcal{P}}(\{\!|\gamma|\!\}_M)$ gives a solution to $\hat{f}$.

The interesting direction is showing that if $\hat{f}$ has a solution, then f' is satisfiable in $\mathcal{M}^{lp}$. Suppose that $\hat{f}$ has a solution in which $x_{\gamma_1}, \ldots, x_{\gamma_n}$ taken on values $x_1^*, \ldots, x_n^*$, respectively. (Note that by the construction of $\hat{f}$, we must have $x_{n-1}^* = 0$ and $x_n^* = 1$.) Take $W = \{\delta_1, \ldots, \delta_{2^N}\}$, and let $\pi(\delta_i)(p) = \mathbf{true}$ if and only if $\delta_i \Rightarrow p$ is a propositional tautology. We need to show that there exists a set of probability measures $\mathcal{P}$ such that $(W, \mathcal{P}, \pi) \models f'$. To do this, we start with the solution to $\hat{f}$, which intuitively gives us a lower expectation function defined on a subset of gambles (those that appear in f), and show that it can be extended to a lower expectation function defined on all gambles. This lower expectation function will give us the set $\mathcal{P}$ of probability measures necessary for satisfiability. More precisely, corresponding to the propositional gambles $\gamma_1, \ldots, \gamma_n$ appearing in f (along with *false* and *true*, as described above), let X_i be the gamble corresponding to $\gamma_i = a_{i,1}\delta_1 + \ldots + a_{i,2^N}\delta_{2^N}$ defined on W by $X_i(\delta_j) = a_{i,j}$, for all $1 \leq j \leq 2^N$. Define the function $\underline{P}^*$ on $\{X_1, \ldots, X_n\}$ by $\underline{P}^*(X_i) = x_i^*$. Our goal is to extend $\underline{P}^*$ to a lower expectation function $\underline{E}^*$ defined on all gambles, such that $\underline{E}^*(Y) = \underline{P}^*(Y)$ for $Y \in \{X_1, \ldots, X_n\}$.

The first thing we have to check is that $\underline{P}^*$ is a potential candidate to be a partial lower expectation function. This is made precise by Walley's [1991, p.72] notion of coherence. A function $\underline{P}$ defined on a set $\{X_1,\ldots,X_n\}$ of gambles is *coherent* if and only if for all reals $b_1,\ldots,b_n \geq 0$ and $i^* \in \{1,\ldots,n\}$, we have[11]

$$\sup_{w\in W}\left\{\left[(\sum_{j\neq i} b_j(X_j - \widetilde{\underline{P}(X_j)})) - b_{i^*}(X_{i^*} - \widetilde{\underline{P}(X_{i^*})})\right](w)\right\} \geq 0. \qquad (10)$$

If E is a lower expectation function (that is, if $E = \underline{E}_{\mathcal{P}}$ for some set $\mathcal{P}$ of probability measures) then (10) holds (with $\underline{P}$ replaced by E). To see this, first note that for any gambles $Y, Y_1, \ldots, Y_n$ and nonnegative reals $b, b_1, \ldots, b_n$, by Proposition 2.3, $E(b(Y - \widetilde{E(Y)})) = 0$, and $E(b_1(Y_1 - \widetilde{E(Y_1)}) + \cdots + E(Y_n - \widetilde{E(Y_n)}) \geq 0$. Now suppose that

$$\sup_{w\in W}\left\{\left[(\sum_{j=1}^{n} b_j(Y_j - \widetilde{E(Y_j)})) - bE(Y - \widetilde{E(Y)})\right](w)\right\} = c < 0.$$

Then $\sum_{j=1}^n b_j(Y_j - \widetilde{E(Y_j)}) \leq b(E(Y - \widetilde{E(Y)}) + \tilde{c}$. Thus

$$E(\sum_{j=1}^{n} b_j(Y_j - \widetilde{E(Y_j)})) \leq E(b(Y - \widetilde{E(Y)}) + \tilde{c}) = E(b(Y - \widetilde{E(Y)})) + c = c < 0.$$

But this is a contradiction. It follows that (10) holds for E.

This shows that coherence, that is, (10), is a necessary condition for $\underline{P}$ to be extendible to a lower expectation. Walley's *Natural Extension Theorem*, which we now state, shows that it is sufficient as well.

THEOREM A.6. [Walley 1991, p.123] *Suppose that $\underline{P}$ is a coherent real-valued function defined on a set $\mathcal{G}$ of gambles on W. Define the* natural extension $\underline{E}$ *of $\underline{P}$ as*

$$\underline{E}(Y) = \sup\{b \;:\; Y - \tilde{b} \geq \sum_{j=1}^{m} \lambda_j(Y_j - \widetilde{\underline{P}(Y_j)}), m \geq 0, Y_j \in \mathcal{G}, \lambda_j \geq 0, b \in \mathbb{R}\}.$$

Then $\underline{E}$ is a lower expectation function (that is, $\underline{E} = \underline{E}_{\mathcal{P}}$ for some set $\mathcal{P}$ of probability measures on W) and $\underline{E}$ agrees with $\underline{P}$ on the gambles in $\mathcal{G}$.

It follows from Theorem A.6 that to show that $\underline{P}^*$ can be extended to a lower expectation, it suffices to show that it is coherent. This is done in the next lemma.

LEMMA A.7. *The function $\underline{P}^*$ is coherent on $\{X_1,\ldots,X_n\}$.*

[11] Walley has an apparently stronger requirement, namely, that for all $Y_1,\ldots,Y_m \subseteq \{X_1,\ldots,X_n\}$ and all reals $b_1,\ldots,b_m$, we have $\sup_{w\in W}\left\{[(\sum_{j=2}^m b_j(Y_j - \widetilde{\underline{P}^*(Y_j)})) - b_1(Y_1 - \widetilde{\underline{P}^*(Y_1)})](w)\right\} \geq 0$. However, we can assume that all the Y_j's are distinct without loss of generality (since if there is a repetition of X_i's, we can simply add the coefficients b_i), and if any X_i does not appear among the Y_i's, we can add it, taking $b_i = 0$.

PROOF. Observe that (10) is equivalent to the condition that, for all nonnegative reals $b_1, \ldots, b_{n-1}$, reals b_n, and all i^*, it is *not* the case that

$$\sum_{j \neq i^*} b_j(X_j - \tilde{x}_j^*) < b_{i^*}(X_{i^*} - \tilde{x}_{i^*}^*). \tag{11}$$

Suppose, by way of contradiction, that there exist such nonnegative reals $b_1, \ldots, b_{n-1}$ and real b_n. Note that, without loss of generality, we can assume that $i^* \neq n$ and that $b_{i^*} = 1$. For if $b_{i^*} \neq 0$, then we can divide both all coefficients on both sides by b_{i^*} to get an equivalent instance of (11) where $b_{i^*} = 1$. Moreover, observe that, since $\gamma_{n-1} = \mathit{false}$ and $\gamma_n = \mathit{true}$, we have that $X_{n-1} = \tilde{0}$, $x_{n-1}^* = 0$, $X_n = \tilde{1}$, and $x_n^* = 1$. Thus, if $i^* = n$ or if $b_{i^*} = 0$, then we can take $i^* = n - 1$ and $b_{i^*} = 1$, to get an equivalent instance of (11). Thus, if there exist $b_1, \ldots, b_n$ and i^* such that (11) holds, then we can assume without loss of generality that $i^* \neq n$ and $b_{i^*} = 1$.

For simplicity, in the remainder of the argument, we take $i^* = 1$ (the argument is the same for every choice of $i^* \neq n$) and assume that $b_1 = 1$. Since $X_n = \tilde{1}$ and $x_n^* = 1$, there exist nonnegative $b'_2, \ldots, b'_{n-1}$ and $\epsilon > 0$ such that

$$\sum_{j=2}^{n-1} b'_j(X_j - \tilde{x}_j^*) + \tilde{\epsilon} \leq X_1 - \tilde{x}_1^*.$$

Let $b'_n = \epsilon + x_1^* - \sum_{j=2}^{n-1} b'_j x_j^*$ (b'_n can be negative). Note that $\sum_{j=2}^{n} b'_j x_j^* = x_1^* + \epsilon$, since $x_n^* = 1$. Since $X_n - \tilde{x}_n^* = \tilde{0}$, it follows that

$$\sum_{j=2}^{n} b'_j(X_j - \tilde{x}_j^*) + \tilde{\epsilon} \leq X_1 - \tilde{x}_1^*.$$

Thus, $\sum_{j=2}^{n} b'_j X_j \leq X_1$. By definition of B_1, there exists $\langle b_2^*, \ldots, b_n^* \rangle \in B_1$ such that $\sum_{j=2}^{n} b_j^* X_j \leq X_1$ and $\sum_{j=2}^{n} b_j^* x_j^* \geq \sum_{j=2}^{n} b'_j x_j^* = x_1^* + \epsilon$. In other words, we can assume without loss of generality that $b'_2, \ldots, b'_n$ are in B_1, since otherwise, we can always replace them by $b_2^*, \ldots, b_n^*$. By the above, we have that $E(b'_2 X_2 + \ldots + b'_n X_n) \leq E(X_1)$. By the constraints corresponding to f_1 in $\hat{f}$, we can therefore derive $b'_2 E(X_2) + \cdots + b'_{n-1} E(X_{n-1}) + b'_n \leq E(X_1)$. (The constraints corresponding to f_1 where chosen so that exactly this derivation could be performed.) However, replacing b'_n by $\epsilon + x_1^* - \sum_{j=2}^{n-1} b'_j x_j^*$ and replacing $E(X_i)$ by x_i^*, we get that $\epsilon < 0$, a contradiction. Thus, there cannot exist $b_1, \ldots, b_n$ of the above form. Hence, $\underline{P}^*$ is coherent, as desired. □

Continuing with the proof of Theorem 5.2, by Lemma A.7 and the Natural Extension Theorem, there exists a set of probability measures $\mathcal{P}$ such that $\underline{E}_{\mathcal{P}}$ is defined on all gambles, and $\underline{E}_{\mathcal{P}}$ agrees with $\underline{P}^*$ on $\{X_1, \ldots, X_n\}$. It is routine to check that $(W, \mathcal{P}, \pi) \models f'$, and hence $(W, \mathcal{P}, \pi) \models f$, as required.

THEOREM 5.3. $\mathbf{AX}^{bel}$ *is a sound and complete axiomatization of* $\mathcal{L}^E$ *with respect to* $\mathcal{M}^{bel}$.

PROOF. Soundness is straightforward. For completeness, we proceed as in Theorem 5.1. Assume without loss of generality that f is a conjunction of expectation inequalities and their negations. Using axioms **E7**, **E8**, and **E10** we can convert

f into the equivalent formula $f^{T_2} \in \mathcal{L}^{E'}$ (Lemma A.5) where e is applied only to propositional formulas. Every propositional formula φ in f^{T_2} is provably equivalent to a region ρ_i for some $1 \leq i \leq 2^{2^N}$. Since $\varphi = \rho_i$ is valid, $e(\varphi) = e(\rho_i)$ is provable by **E5**. This means that we can find a formula f' provably equivalent to f^{T_2}, where e is applied only to formulas $\rho_1, \ldots, \rho_{2^{2^N}}$. Let f'' be obtained from f' by adding as conjuncts to f' all the instances of **E9** involving the regions $\rho_1, \ldots, \rho_{2^{2^N}}$, as well as the inequalities $e(\rho_i) \geq 0$ for all $1 \leq i \leq 2^{2^N}$, $e(\rho_1) \geq 1$, $-e(\rho_1) \geq -1$ and $-e(\rho_{2^{2^N}}) \geq 0$. (Recall that by assumption, ρ_1 is *true* and $\rho_{2^{2^N}}$ is *false*.) It is not hard to see that these formulas are provable, hence f'' is provably equivalent to f', and hence to f.

As before, the negation of an expectation inequality $a_1 e(\gamma_1) + \cdots + a_n e(\gamma_n) \geq b$ can be written $-a_1 e(\gamma_1) - \cdots - a_n e(\gamma_n) > -b$. Thus, without loss of generality, we can assume that f'' is the conjunction of formulas of the form $a_1 e(\rho_1) + \cdots + a_{2^{2^N}} e(\rho_{2^{2^N}}) \geq b$ and $-a'_1 e(\rho_1) - \cdots - a'_{2^{2^N}} e(\rho_{2^{2^N}}) > b'$. Let $\hat{f}$ be the system of inequations obtained by replacing the terms $e(\rho_i)$ by the variable x_i, for $1 \leq i \leq 2^{2^N}$. We claim that f'' in satisfiable in $\mathcal{M}^{bel}$ iff the system $\hat{f}$ of equations has a solution. One direction is straightforward. Suppose that f'' is satisfiable. Thus, there exists a belief structure (W, Bel, π) such that $(W, \mathrm{Bel}, \pi) \models f''$. Clearly, taking $x_i = E_{\mathrm{Bel}}([\![\rho_i]\!]_M) = \mathrm{Bel}(\rho_i)$ gives a solution to $\hat{f}$. Conversely, suppose that $\hat{f}$ has a solution, say $x^*_1, \ldots, x^*_{2^{2^N}}$. Let $p_1, \ldots, p_N$ be the primitive propositions appearing in f (and, hence, in f''). Define the belief structure $M = (W, \mathrm{Bel}, \pi)$ as follows:

—$W = \{\delta_1, \ldots, \delta_{2^N}\}$, the set of atoms over $p_1, \ldots, p_N$;

—$\mathrm{Bel}(\{\delta_{i_1}, \ldots, \delta_{i_k}\}) = x^*_i$, where i is the unique index such that $\delta_{i_1} \vee \ldots \vee \delta_{i_k}$ is logically equivalent to the region ρ_i;

—$\pi(\delta_i)(p) = \mathbf{true}$ if and only if $\delta_i \Rightarrow p$ is a propositional tautology.

It is straightforward to show that Bel defined in this way is a belief function. That Bel satisfies B1 and B2 follows from the fact that we must have $x^*_1 = 0$ and $x^*_{2^{2^N}} = 1$. Similarly, that Bel satisfies B3 follows from the observation that any instance of B3 corresponds to an instance of **E9** that was added as an inequality in $\hat{f}$, and thus is satisfied by the solution $x^*_1, \ldots, x^*_{2^{2^N}}$. Finally, that $M \models f''$ holds also follows from the construction of $\hat{f}$. Thus, f'' is satisfiable. This completes the proof of the claim.

Returning to the proof of the theorem, we want to show that if f is consistent, then f is satisfiable. So suppose that f is consistent and, by way of contradicition, that it is not satisfiable. Then f'' is not satisfiable. Thus, by the claim, the system of equations $\hat{f}$ has no solution. Thus, $\neg f''$ is an instance of the axiom **Ineq**. Since f'' is provably equivalent to f, it follows that $\neg f$ is provable, so that f is inconsistent, a contradiction. $\square$

THEOREM 5.4. *$\mathbf{AX}^{poss}$ is a sound and complete axiomatization of $\mathcal{L}^E$ with respect to $\mathcal{M}^{poss}$.*

PROOF. Soundness is straightforward. For completeness, we proceed almost exactly as in Theorem 5.3. As before, we can reduce to showing that a formula f consistent with $\mathbf{AX}^{poss}$ that is a conjunction of formulas of the form

$a_1e(\rho_1)+\cdots+a_{2^{2^N}}e(\rho_{2^{2^N}}) \geq b$ and $-a'_1e(\rho_1)-\cdots-a'_{2^{2^N}}e(\rho_{2^{2^N}}) > b'$ is satisfiable. We add as conjuncts to this formula all the expectation inequalities corresponding to **E11** involving the regions $\rho_1,\ldots,\rho_{2^{2^N}}$, as well as the inequalities $e(\rho_i) \geq 0$ for all $1 \leq i \leq 2^{2^N}$, $e(\rho_1) \geq 1$, $-e(\rho_1) \geq -1$ and $-e(\rho_{2^{2^N}}) \geq 0$. Again, let $\hat{f}$ be the system of inequations formed by replacing $e(\rho_i)$ by x_i. Arguments similar in spirit to those of Theorem 5.3 can be used to show that $\hat{f}$ has a solution iff there is a structure in $\mathcal{M}^{poss}$ satisfying f; we leave details to the reader. The proof is completed using **Ineq**, just as the earlier completeness proofs. □

A.4 Proofs for Section 6

As in FHM and HP, the small model theorems we prove rely on the following lemma, which can be derived from Cramer's rule [Shores 1999] and simple estimates on the size of the determinant (see also Chvátal [1983] for a simpler variant):

Lemma A.8. *If a system of r linear equalities and/or inequalities with integer coefficients, each of length at most l, has a nonnegative solution, then it has a nonnegative solution with at most r positive entries, where the size of each element of the solution is $O(rl + r\log(r))$.*

Before getting to the small model theorems, we first establish a *finite model* theorem, that is, we show that, for all the representations of uncertainty we consider in this paper, if a formula in $\mathcal{L}^E$ is satisfiable, it is in fact satisfiable in a structure with finitely many states. This is a consequence of the completeness proofs in Section 5.

Lemma A.9. *Suppose $f \in \mathcal{L}^E$ is satisfied in some structure in $\mathcal{M}^{prob}$ (resp., $\mathcal{M}^{lp}$, $\mathcal{M}^{bel}$, $\mathcal{M}^{poss}$). Then f is satisfied in a structure in $\mathcal{M}^{prob}$ (resp., $\mathcal{M}^{lp}$, $\mathcal{M}^{bel}$, $\mathcal{M}^{poss}$) with finitely many states.*

Proof. In each of the proofs of Theorems 5.1–5.4, we show that a formula is satisfiable iff a certain system of inequations has a solution. The system of inequations involves only finitely many variables. Our argument showing that if the system of inequations is satisfiable, then there is a structure where the formula is satisfied actually shows that the satisfying structure has only finitely many states, with no more than one state per variable in the system. □

Note for future reference that the satisfying structures in $\mathcal{M}^{prob}$ and $\mathcal{M}^{lp}$ constructed in the proofs of Theorems 5.1 and 5.2 are such that all subsets are measurable.

Theorem 6.1. *Suppose that $f \in \mathcal{L}^E$ is satisfied in some structure in $\mathcal{M}^{lp}$. Then f is satisfied in a structure $(W,\mathcal{P},\pi)$ such that $|W| \leq |f|^2$, $|\mathcal{P}| \leq |f|$, $\mu(w)$ is a rational number such that $||\mu(w)||$ is $O(|f|^2||f|| + |f|^2\log(|f|))$ for every world $w \in W$ and $\mu \in \mathcal{P}$, and $\pi(w)(p) =$ **false** for every world $w \in W$ and every primitive proposition p not appearing in f.*

Proof. The first step in the proof involves showing that if $\mathcal{P}$ is a set of probability measures defined on a finite space W (assuming all subsets are measurable), and if $X_1,\ldots,X_n$ are gambles over W, then we can assume without loss of generality that for each gamble X_i, there is a probability measure $\mu_{X_i} \in \mathcal{P}$ such that $E_{\mu_{X_i}}(X_i) = \underline{E}_{\mathcal{P}}(X_i)$ (rather than $\underline{E}_{\mathcal{P}}(X_i)$ just being the inf of $E_\mu(X_i)$ for $\mu \in \mathcal{P}$). A similar result is proved in HP for upper probabilities. More specifically, in HP, it is shown that, given $\mathcal{P}$, there exists a set $\mathcal{P}' \supseteq \mathcal{P}$ such that, for each $S \subseteq W$, $\underline{E}_{\mathcal{P}}(X_S) = \underline{E}_{\mathcal{P}'}(X_S)$ and, moreover, there exists some $\mu_S \in \mathcal{P}'$ such that $\mu_S(S) = \underline{E}_{\mathcal{P}}(X_S)$. The next result can be viewed as a generalization of this result to arbitrary gambles. It is proved using essentially the same technique as the corresponding result in HP.

Lemma A.10. *Let $\mathcal{P}$ be a set of probability measures defined on a finite set W of worlds, and let $X_1, \ldots, X_n$ be gambles over W. Then there exists a set $\mathcal{P}'$ of probability measures such that, for each gamble X_i, (a) $\underline{E}_{\mathcal{P}}(X_i) = \underline{E}_{\mathcal{P}'}(X_i)$, and (b) there is a probability measure $\mu_{X_i} \in \mathcal{P}'$ such that $E_{\mu_{X_i}}(X_i) = \underline{E}_{\mathcal{P}}(X_i)$.*

Proof. To show that $\mathcal{P}'$ exists, it clearly suffices to show that, for each gamble $X \in \{X_1, \ldots, X_n\}$, there is a probability measure μ_X such that $E_{\mu_X}(Y) \geq \underline{E}_{\mathcal{P}'}(Y)$ for all $Y \in \{X_1, \ldots, X_n\}$ and $E_{\mu_X}(X) = \underline{E}_{\mathcal{P}}(X)$.

Given a gamble $X \in \{X_1, \ldots, X_n\}$, if there exists $\mu \in \mathcal{P}$ such that $E_\mu(X) = \underline{E}_{\mathcal{P}}(X)$, then we are done. If not, there must be a sequence $\mu_1, \mu_2, \ldots$ of measures in $\mathcal{P}$ such that $\lim_{n\to\infty} E_{\mu_n}(X) = \underline{E}_{\mathcal{P}}(X)$. Suppose that $W = \{w_1, \ldots, w_n\}$. By the Bolzano-Weierstrass theorem [Rudin 1976] (which says that every sequence of real numbers has a convergent subsequence), the sequence $\mu_1(w_1)$, $\mu_2(w_1)$, $\mu_3(w_1)$, ... has a convergent subsequence. Suppose, inductively, that we have found a subsequence $\mu_{j,1}, \mu_{j,2}, \ldots$ of $\mu_1, \mu_2, \ldots$ such that $\mu_{j,1}(w), \mu_{j,2}(w), \mu_{j,3}(w), \ldots$ converges for $w \in \{w_1, \ldots, w_j\}$. By applying the Bolzano-Weierstrass Theorem again, there is a subsequence $\mu_{j+1,1}, \mu_{j+2,1}, \ldots$ of $\mu_{j,1}, \mu_{j,2}, \ldots$ such that $\mu_{j+1,1}(w_{j+1})$, $\mu_{j+2,1}(w_{j+1})$, ... converges. It follows that $\mu_{j+1,1}(w), \mu_{j+2,w}), \ldots$ converges for all $w \in \{w_1, \ldots, w_{j+1}\}$. By induction, we can find a subsequence $\mu_{n,1}, \mu_{n,2}, \ldots$ of the original sequence such that $\mu_{n,1}(w), \mu_{n,2}(w), \ldots$ converges for all $w \in W$. Suppose that $\mu_{n,1}(w_i), \mu_{n,2}(w_i), \ldots$ converges to p_i. It is easy to see that $p_1 + \cdots + p_n = 1$, since $\sum_{i=1}^n \mu_{n,j}(w_i) = 1$ for all j. Let μ_X be the probability measure such that $\mu(w_i) = p_i$. Since $\mu_{n,j} \to \mu_X$, it must be the case that $E_{\mu_X}(X) = \underline{E}_{\mathcal{P}}(X)$. Moreover, since $\mu_{n,j} \in \mathcal{P}$ for all j, it must be the case that $E_{\mu_{n,j}}(Y) \geq \underline{E}_{\mathcal{P}}(Y)$ for all j. Thus, $E_{\mu_X}(Y) \geq \underline{E}_{\mathcal{P}}(Y)$, as desired.

Now let $\mathcal{P}' = \mathcal{P} \cup \{\mu_{X_1}, \ldots, \mu_{X_n}\}$. $\mathcal{P}'$ clearly has the desired properties. $\square$

Continuing with the proof of Theorem 6.1, suppose that f is satisfiable in $\mathcal{M}^{lp}$. By Lemma A.9, f is satisfied in a lower probability structure with a finite set W of worlds. Thus, by Lemma A.10, f is satisfied in a structure $M = (W, \mathcal{P}, \pi)$ such that, for all $X \in \{X_1, \ldots, X_n\}$, there exists $\mu_X \in \mathcal{P}$ such that $E_{\mu_X}(X) = \underline{E}_{\mathcal{P}}(X)$.

The rest of the proof also continues in much the same spirit as the proof of the analogous result in HP. A straightforward induction on structures shows that we can find a formula equivalent to f in disjunctive normal form, where each disjunct has length at most $|f|$. (The inductive hypothesis states that we can find both a DNF formula equivalent to f where each disjunct has length at most $|f|$, and a CNF formula equivalent to f where each conjunct has length at most $|f|$.) Since a formula in DNF is satisfiable iff one of its disjuncts is, we can thus assume without loss of generality that f is a conjunction of inequality formulas and negations of inequality formulas. Let $p_1, \ldots, p_N$ be the primitive formulas appearing in f. Let

$\delta_1, \ldots, \delta_{2^N}$ be the atoms over $p_1, \ldots, p_N$. As in the proof of completeness, we derive a system of equalities and inequalities from f, but it is a slightly more complicated system than that used in the completeness proof. Recall that each propositional formula over $p_1, \ldots, p_N$ is a disjunction of atoms. Let $\gamma_1, \ldots, \gamma_k$ be the propositional gambles that appear in g. Notice that $k < |f|$ (since there are some symbols in f, such as the coefficients, that are not in the propositional formulas). The system of equations and inequalities we construct involve variables x_{ij}, where $i = 1, \ldots, k$ and $j = 1, \ldots, 2^N$. Intuitively, x_{ij} represents $\mu_{\gamma_i}([\![\delta_j]\!]_M)$, where $\mu_{\gamma_i} \in \mathcal{P}$ is such that $E_{\mu_{\gamma_i}}(\{\!|\gamma_i|\!\}_M) = \underline{E}_{\mathcal{P}}(\{\!|\gamma_i|\!\}_M)$. Thus, the system includes $k < |f|$ equations of the form

$$x_{i1} + \cdots + x_{i2^N} = 1,$$

for $i = 1, \ldots, k$. Suppose that γ_i is equivalent to $b_{i1}\delta_1 + \cdots + b_{i2^N}\delta_{2^N}$. Then $b_{i1}x_{i1} + \cdots + b_{i2^N}x_{i2^N}$ represents $E_{\mu_{\gamma_i}}(\{\!|\gamma_i|\!\}_M)$. Since $E_{\mu_{\gamma_i}}(\{\!|\gamma_i|\!\}_M) \leq E_{\mu}(\{\!|\gamma_i|\!\}_M)$ for all $\mu \in \mathcal{P}$, the system includes $k2 - k$ inequalities of the form

$$b_{i1}x_{i1} + \cdots + b_{i2^N}x_{i2^N} \leq b_{i1}x_{i'1} + \cdots + b_{i2^N}x_{i'2^N},$$

for each pair i, i' such that $i \neq i'$. For each conjunct in g of the form $a_1 e(\gamma_1) + \cdots + a_n e(\gamma_k) \geq b$, there is a corresponding inequality where, roughly speaking, we replace $e(\gamma_i)$ by $E_{\mu_{\gamma_i}}(\{\!|\gamma_i|\!\}_M)$.[12] Since $E_{\mu_{\gamma_i}}\{\!|\gamma_i|\!\}_M$ corresponds to $b_{i1}x_{i1} + \cdots b_{i2^N}x_{i2^N}$, the appropriate inequality is

$$\sum_{i=1}^{k} a_i(b_{i1}x_{i1} + \cdots + b_{i2^N}x_{i2^N}) \geq b.$$

Negations of such formulas correspond to a negated inequality formula; as before, this is equivalent to a formula of the form

$$-(\sum_{i=1}^{k} a_i(b_{i1}x_{i1} + \cdots + b_{i2^N}x_{i2^N}) > -b.$$

Notice that there are at most $|f|$ inequalities corresponding to the conjuncts of f. Thus, altogether, there are at most $k(k-1) + 2|f| < |f|^2$ equations and inequalities in the system (since $k < |f|$). We know that the system has a nonnegative solution (taking x^i_j to be $\mu_{\gamma_i}([\![\delta_j]\!]_M)$). It follows from Lemma A.8 that the system has a solution $x^* = (x^*_{11}, \ldots, x^*_{12^N}, \ldots, x^*_{k1}, \ldots, x^*_{k2^N})$ with $t \leq |f|^2$ entries positive, and with each entry of size $O(|f|^2||f|| + |f|^2 \log(|f|))$.

We use this solution to construct a small structure satisfying the formula f. Let $I = \{i : x^*_{ij}$ is positive, for some $j\}$; suppose that $I = \{i_1, \ldots, i_{t'}\}$ for some $t' \leq t$. Let $M = (W, \mathcal{P}, \pi)$, where W has t' worlds, say $s_1, \ldots, s_{t'}$. Let $\pi(s_h)$ be the truth assignment corresponding to the formula δ_{i_h}, that is, $\pi(s_h)(p) = \textbf{true}$ if and only if $\delta_{i_h} \Rightarrow p$ is a propositional tautology (and where $\pi(s_h)(p) = \textbf{false}$ if p does not appear in f). Define $\mathcal{P} = \{\mu_j : 1 \leq i \leq k\}$, where $\mu_j(s_h) = x^*_{i_h j}$. It is clear from the construction that $M \models f$. Since $|\mathcal{P}| = k < |f|$, $|W| = t' \leq t \leq |f|^2$ and

[12] For simplicity here, we are implicitly assuming that each of the formulas γ_i appears in each conjunct of f. This is without loss of generality, since if γ_i does not appear, we can put it in, taking $a_i = 0$.

$\mu_j(s_h) = x^*_{i_h j}$, where, by construction, the size of $x^*_{i_h j}$ is $O(|f|^2||f|| + |f|^2 \log(|f|))$, the theorem follows.

THEOREM 6.2. *Suppose that $f \in \mathcal{L}^E$ is satisfied in some structure in $\mathcal{M}^{bel}$ (resp., $\mathcal{M}^{poss}$). Then f is satisfied in a structure (W, ν, π) such that $|W| \leq |f|^2$, ν is a belief function (resp., possibility measure) whose corresponding mass function is positive on at most $|f|$ subsets of W and the mass of each of these $|f|$ is a rational number of size $O(|f|\,||f||+|f|\log(|f|))$, and $\pi(w)(p) =$ **false** for every world $w \in W$ and every primitive proposition p not appearing in f.*

PROOF. First, consider expectation for belief functions. The proof is similar in spirit to the proof of the small model theorem for reasoning about belief functions given in FHM. In FHM, the complexity result used the representation of belief functions in terms of mass functions. We do that here too; to do so, it is helpful to have an alternate characterization of expectation for belief that uses mass functions.

Given a belief function Bel over a finite set W of worlds, let m be the corresponding mass function. For a given gamble X, let $v_U = \min_{w \in U} X(w)$.

LEMMA A.11. $E_{\mathrm{Bel}}(X) = \sum_{U \subseteq W} m(U) v_U$.

PROOF. Suppose that Bel is a belief function on W with corresponding mass function m, and X is a gamble on a finite set $W = \{w_1, \ldots, w_n\}$, where the elements of W are ordered so that $i \leq j$ implies $X(w_i) \leq X(w_j)$. For each $U \subseteq W$, let $i_U = \min\{i : w_i \in U\}$; note that $v_U = X(w_{i_U})$.

$$\underline{E}_{\mathrm{Bel}}(X) = \sum_{U \subseteq W} m(U) v_U.$$

We want to show that $\underline{E}_{\mathrm{Bel}}(X) = E_{\mathrm{Bel}}(X)$.

Recall from Section 2.3 that $E_{\mathrm{Bel}} = \underline{E}_{\mathcal{P}_{\mathrm{Bel}}}$, where $\mathcal{P}_{\mathrm{Bel}}$ is the set $\{\mu : \mu(U) \geq \mathrm{Bel}(U) \text{ for all } U \subseteq W\}$. As a first step, let μ_0 be a probability measure on W such that $\mu_0(X = x) = \sum_{\{U : v_U = x\}} m(U)$. (Note that μ_0 is a probability measure, since

$$\sum_{x \in \mathcal{V}(X)} \mu_0 X = x = \sum_{x \in \mathcal{V}(X)} \sum_{\{U : v_U = x\}} m(U) = \sum_{U \subseteq W} m(U) = 1,$$

since m is a mass function.) Clearly

$$\begin{aligned} E_{\mu_0}(X) &= \sum_{x \in \mathcal{V}(X)} x \mu_0(X = x) = \sum_{x \in \mathcal{V}(X)} x \sum_{\{U : v_U = x\}} m(U) \\ &= \sum_{U \subseteq W} m(U) v_U = \underline{E}_{\mathrm{Bel}}(X). \end{aligned}$$

Moreover, for all $U \subseteq W$, since $\mu_0(U) = \sum_{\{V : w_{i_V} \in U\}} m(V)$, and $w_{i_V} \in U$ for all $V \subseteq U$, it follows that $\mu_0(U) \geq \sum_{\{V : w_{i_V} \in U\}} m(V) = \mathrm{Bel}(U)$. It follows that $\mu_0 \in \mathcal{P}_{\mathrm{Bel}}$. Thus,

$$\underline{E}_{\mathrm{Bel}}(X) = E_{\mu_0}(X) \geq \inf_{\mu \in \mathcal{P}_{\mathrm{Bel}}} E_\mu(X) = E_{\mathcal{P}_{\mathrm{Bel}}}(X) = E_{\mathrm{Bel}}(X).$$

To show that $\underline{E}_{\mathcal{P}_{\mathrm{Bel}}}(X) \geq \underline{E}_{\mathrm{Bel}}(X)$, let $x_1 \leq \cdots \leq x_k$ be the values of X in increasing order, and let $U_j = \cup_{i=1}^{j}(X = x_j) = \{w \in W : X(w) \leq x_j\}$. We claim that $\mu_0(U_j) = \mathrm{Plaus}(U_j)$, $j = 1, \ldots, k$. This is almost immediate from the definition, since

$$\mu_0(U_j) = \sum_{\{U: v_U \le x_j\}} m(U) = \sum_{\{U: U \cap U_j \neq \varnothing\}} m(U) = \mathrm{Plaus}(U_j).$$

It follows that if $\mu \in \mathcal{P}_{\mathrm{Bel}}$, then $\mu_0(U_j) \ge \mu(U_j)$, for $j = 1, \ldots, k$ (since $\mu(U_j) \le \mathrm{Plaus}(U_j)$). For convenience, define $U_0 = \varnothing$. Then we get that

$$\begin{array}{rcl} E_\mu(X) - E_{\mu_0}(X) & = & \sum_{i=i}^k x_i(\mu(X = x_i) - \mu_0(X = x_i)) \\ & = & \sum_{i=1}^k x_i((\mu(U_i) - \mu(U_{i-1})) - (\mu_0(U_i) - \mu_0(U_{i-1}))) \\ & = & x_k(\mu(U_k) - \mu_0(U_k)) + \sum_{i=1}^{k-1}(x_{i+1} - x_i)(\mu_0(U_i) - \mu(U_i)) \\ & \ge & 0. \end{array}$$

The last inequality follows from the fact that $\mu(U_k) = \mu_0(U_k) = \mu_0(W) = 1$, $x_{i+1} > x_i$, and $\mu_0(U_i) \ge \mu(U_i)$ for $i = 1, \ldots, k-1$. Thus, $E_{\mu_0}(X) = \underline{E}_{\mathcal{P}_{\mathrm{Bel}}}(X)$. It follows that $E_{\mathrm{Bel}}(X) = \underline{E}_{\mathcal{P}_{\mathrm{Bel}}}(X)$. $\square$

Continuing with the proof of Theorem 6.2, suppose that $f \in \mathcal{L}^E$ is satisfiable in structure M. As in the proof of Theorem 6.1, we can assume without loss of generality that f is a conjunction of expectation formulas and their negations. Let $\delta_1, \ldots, \delta_{2^N}$ be the atoms over $p_1, \ldots, p_N$. Let $\rho_1, \ldots, \rho_{2^{2^N}}$ be all the inequivalent propositional formulas over $p_1, \ldots, p_N$.

As in the proof of completeness, we derive a system of equalities and inequalities from f. First, note that every propositional formula appearing in the gambles of f is provably equivalent to some region ρ_i. Therefore, we can replace every gamble in f by an equivalent gamble where the propositional formulas are the regions $\rho_1, \ldots, \rho_{2^{2^N}}$. Let f' be resulting formula. Clearly, $M \models f'$. Construct the following system of linear inequalities over the variables $x_1, \ldots, x_{2^{2^N}}$, where, intuitively, the variable x_i stands for the mass corresponding to the set ρ_i. For every gamble γ that appears in f' every ρ_i, we can compute values $v_{\gamma,i} = \min_{w \in [\![\rho_i]\!]_M}(\{\!\!\{\gamma\}\!\!\}_M(w))$. For every gamble γ, replace every term $e(\gamma)$ in f' by $\sum_{i=1}^{2^{2^N}} x_i v_{\gamma,i}$. Consider the system $\hat{f}'$ of at most $|f|$ inequalities (over unknowns $x_1, \ldots, x_{2^{2^N}}$) resulting from this process, after putting together like terms, along with the inequalities $x_1 + \cdots + x_{2^{2^N}} \ge 1$ and $-x_1 - \cdots - x_{2^{2^N}} \ge -1$.

We now show that any nonnegative solution $x_1^*, \ldots, x_{2^{2^N}}^*$ of $\hat{f}'$ can be used to construct a belief function Bel^* such that $(W, \mathrm{Bel}^*, \pi) \models f'$. (Note that we keep the same worlds and interpretation as in M.) Let $x_1^*, \ldots, x_{2^{2^N}}^*$ be a nonnegative solution of $\hat{f}$. Define the mass function $m^*([\![\rho_i]\!]_M) = x_i^*$, and define $m^*(U) = 0$ for all $U \subseteq W$ such that U is not of the form $[\![\rho]\!]_M$ for some region ρ. Because $x_1^* + \ldots + x_{2^{2^N}}^* = 1$ (by choice of $\hat{f}'$), m^* is indeed a mass function. Let Bel^* be the belief function corresponding to m^*, and let $M^* = (W, \mathrm{Bel}^*, \pi)$. Since the interpretation of propositional formulas and gambles depends only on the set of worlds and the interpretation π, for all propositional formulas ρ_i, we have $[\![\rho_i]\!]_M = [\![\rho_i]\!]_{M^*}$, and for all gambles γ appearing in f', we have $\{\!\!\{\gamma\}\!\!\}_M = \{\!\!\{\gamma\}\!\!\}_{M^*}$. We now show that $M^* \models f'$. Let $a_1 e(\gamma_1) + \ldots + a_k e(\gamma_k) \ge b$ be a conjunct in f'. (A similar argument works for the negation of expectation inequalities.) We know from the construction of $\hat{f}'$ and the fact that $x_1^*, \ldots, x_{2^{2^N}}^*$ is a solution of $\hat{f}'$ that

$$a_1 \sum_{i=1}^{2^{2^N}} x_i^* v_{\gamma_1,i} + \ldots + a_k \sum_{i=1}^{2^{2^N}} x_i^* v_{\gamma_k,i} \geq b.$$

By definition, this is just

$$a_1 \sum_{i=1}^{2^{2^N}} m^*([\![\rho_i]\!]_M) \min_{w \in [\![\rho_i]\!]_M} (\{\!|\gamma_1|\!\}_M(w)) + \ldots + \\ a_k \sum_{i=1}^{2^{2^N}} m^*([\![\rho_i]\!]_M) \min_{w \in [\![\rho_i]\!]_M} (\{\!|\gamma_k|\!\}_M(w)) \geq b.$$

Since every subset of W is of the form $[\![\rho_i]\!]_M$ for some i, we have

$$a_1 \sum_{U \subseteq W} m^*(U) \min_{w \in U} (\{\!|\gamma_1|\!\}_M(w)) + \ldots + a_k \sum_{U \subseteq W} m^*(U) \min_{w \in U} (\{\!|\gamma_k|\!\}_M(w)) \geq b.$$

But by Lemma A.11, this is just $a_1 E_{\mathrm{Bel}^*}(\{\!|\gamma_1|\!\}_M) + \ldots + a_k E_{\mathrm{Bel}^*}(\{\!|\gamma_k|\!\}_M) \geq b$, which is equivalent to $a_1 E_{\mathrm{Bel}^*}(\{\!|\gamma_1|\!\}_{M^*}) + \ldots + a_k E_{\mathrm{Bel}^*}(\{\!|\gamma_k|\!\}_{M^*}) \geq b$, and thus $M^* \models a_1 e(\gamma_1) + \ldots + a_k e(\gamma_k) \geq b$. It follows that $M^* \models f'$.

Because f' is satisfiable in M, there is in fact a nonnegative solution to the system $\hat{f}'$, where $x_i = m([\![\rho_i]\!]_M)$ and m is the mass function corresponding to the belief function in M. By Lemma A.8, there is a small nonnegative solution $x_1^*, \ldots, x_{2^{2^N}}^*$, that is, one with at most $|f|$ positive entries and each entry of size $O((|f|)||f|| + (|f|)\log(|f|))$. By the argument above, $x_1^*, \ldots, x_{2^{2^N}}^*$ can be used to construct a model $M^* = (W, \mathrm{Bel}^*, \pi)$ such that $M^* \models f'$. We are not quite done yet; while we have a small mass function, we still have potentially too many worlds. We now show how to cut down the number of worlds in the model to get a small enough structure M'.

Let $\{i_1, \ldots, i_{|f|}\} = \{i : m^*([\![\rho_i]\!]_{M^*}) > 0\}$. Let $\gamma_1, \ldots, \gamma_k$ be the gambles in f'. For $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, |f|\}$, there is some $w_{i,j} \in [\![\rho_{i_j}]\!]_{M^*}$ such that $\{\!|\gamma_i|\!\}_M(w_{i,j}) = \min_{w \in [\![\rho_{i_j}]\!]_M} (\{\!|\gamma_i|\!\}_M(w))$; that is, $w_{i,j}$ is a world where $\{\!|\gamma_i|\!\}_M$ attains its minimum value. Define $M' = (W', \mathrm{Bel}', \pi')$ as follows. Let the set of states W' be $\{w_{i,j} : i \in \{1, \ldots, k\}, j \in \{1, \ldots, |f|\}\}$. Let Bel' be the belief function whose corresponding mass function m' is defined by setting $m'(\{w_{1,j}, \ldots, w_{k,j}\}) = m^*([\![\rho_{i_j}]\!]_{M^*})$ for $j = 1, \ldots, |f|$, and $m'(U) = 0$ if U is not $[\![\rho_{i_j}]\!]_{M^*}$ for some $j \in \{1, \ldots, |f|\}$. The interpretation π' is simply the restriction of π to W', with $\pi(w)(p) = \mathbf{false}$ for primitive propositions p not appearing in f. This model satisfies the size conditions of the theorem.

We now check that f' (and hence f) is satisfiable in M'. It is clearly sufficient to show that for every gamble γ_i in f', we have $E_{\mathrm{Bel}^*}(\{\!|\gamma_i|\!\}_{M^*}) = E_{\mathrm{Bel}'}(\{\!|\gamma_i|\!\}_{M'})$. To do this, we show that

$$\{\!|\gamma_i|\!\}_{M^*}(w_{i,j}) = \min_{1 \leq r \leq k} (\{\!|\gamma_i|\!\}_{M'}(w_{r,j})). \tag{12}$$

By choice of $w_{i,j}$, $\{\!|\gamma_i|\!\}_{M^*}(w_{i,j}) = \min_{w \in [\![\rho_{i_j}]\!]_{M^*}} (\{\!|\gamma_i|\!\}_{M^*}(w))$. Since $\{w_{1,j}, \ldots, w_{k,j}\}$ is a subset of $[\![\rho_{i_j}]\!]_{M^*}$, it follows that $\{\!|\gamma_i|\!\}_{M^*}(w_{i,j}) = \min_{1 \leq r \leq k}(\{\!|\gamma_i|\!\}_{M^*}(w_{r,j}))$.

Next note that for all $w \in \{w_{1,j}, \ldots, w_{k,j}\}$, we have that $\{\!|\gamma_i|\!\}_{M^*}(w) = \{\!|\gamma_i|\!\}_{M'}(w)$. For if $\gamma_i = c_1\rho_{j_1} + \ldots + c_l\rho_{j_l}$, then $\{\!|\gamma_i|\!\}_{M^*} = c_1 X_{[\![\rho_{j_1}]\!]_{M^*}} + \ldots + c_l X_{[\![\rho_{j_l}]\!]_{M^*}}$ and $\{\!|\gamma_i|\!\}_{M'} = c_1 X_{[\![\rho_{j_1}]\!]_{M'}} + \ldots + c_l X_{[\![\rho_{j_l}]\!]_{M'}}$. Since π' is the restriction of π^* to W', it follows that $[\![\rho_{j_i}]\!]_{M^*} = [\![\rho_{j_i}]\!]_{M'} \cap W'$. Since $\{w_{1,j}, \ldots, w_{k,j}\} \subseteq W'$, it follows that $X_{[\![\rho_{j_i}]\!]_{M^*}}(w) = X_{[\![\rho_{j_1}]\!]_{M'}}(w)$ for all $w \in \{w_{1,j}, \ldots, w_{k,j}\}$. Thus, $\{\!|\gamma_i|\!\}_{M^*}(w) = \{\!|\gamma_i|\!\}_{M'}(w)$ for all $w \in \{w_{1,j}, \ldots, w_{k,j}\}$. We can now show that $E_{\mathrm{Bel}^*}(\{\!|\gamma_i|\!\}_{M^*}) = E_{\mathrm{Bel}'}(\{\!|\gamma_i|\!\}_{M'})$:

$$\begin{aligned}
& E_{\mathrm{Bel}^*}(\{\!|\gamma_i|\!\}_{M^*}) \\
= & \textstyle\sum_{i=1}^{2^{2^N}} m^*([\![\rho_i]\!]_{M^*}) \min_{w \in [\![\rho_i]\!]_{M^*}}(\{\!|\gamma_i|\!\}_{M^*}(w)) \\
= & \textstyle\sum_{j=1}^{|f|} m^*([\![\rho_{i_j}]\!]_{M^*}) \min_{w \in [\![\rho_i]\!]_{M^*}}(\{\!|\gamma_i|\!\}_{M^*}(w)) \quad (\text{since } m^*(U) = 0 \text{ if } U \neq [\![\rho_{i_j}]\!]_{M^*}) \\
= & \textstyle\sum_{j=1}^{|f|} m^*([\![\rho_{i_j}]\!]_{M^*}) \{\!|\gamma_i|\!\}_{M^*}(w_{i,j}) \quad (\text{by choice of } w_{i,j}) \\
= & \textstyle\sum_{j=1}^{|f|} m'(\{w_{1,j}, \ldots, w_{k,j}\}) \{\!|\gamma_i|\!\}_{M^*}(w_{i,j}) \\
= & \textstyle\sum_{j=1}^{|f|} m'(\{w_{1,j}, \ldots, w_{k,j}\}) \min_{1 \le r \le k}(\{\!|\gamma_i|\!\}_{M'}(w_{r,j})) \quad (\text{by } (12)) \\
= & \textstyle\sum_{U \subseteq W'} m'(U) \min w \in U(\{\!|\gamma_i|\!\}_{M'}(w)) \quad (\text{adding sets for which } m'(U) = 0) \\
= & E_{\mathrm{Bel}'}(\{\!|\gamma_i|\!\}_{M'}).
\end{aligned}$$

From this result, it is easy to see that $M' \models f'$, and hence $M' \models f$. This establishes the small-model result for $\mathcal{L}^E$ interpreted over belief functions.

Essentially the same argument works in the case of possibility measures. Recall that a possibility measure Poss is just a plausibility function. whose corresponding mass function m is *consonant*; that is, for all sets U, V such that $m(U) > 0$ and $m(V) > 0$, we have either $U \subseteq V$ or $V \subseteq U$ [Dubois and Prade 1982]. In other words, the sets of positive mass $U_1, \ldots, U_k$ can be ordered such that $U_1 \subseteq \ldots \subseteq U_k$. We then proceed much as for belief functions. We construct a system of inequalities over the variables $x_1, \ldots, x_k$, where, Using Lemma A.8 again, because f' is satisfiable in M, there is a small nonnegative solution $x_1^*, \ldots, x_k^*$, with at most $|f|$ positive entries, each of small size. We can then use this solution to build a structure satisfying f where the mass function is consonant and positive on at most $|f|$ sets. This follows directly from the fact that the only sets of positive mass will be among $U_1, \ldots, U_k$, which are already such that $U_1 \subseteq \ldots \subseteq U_k$.) The remainder of the proof goes through as in the belief function case.

THEOREM 6.3. *The problem of deciding whether a formula in $\mathcal{L}^E$ is satisfiable in $\mathcal{M}^{prob}$ (resp., $\mathcal{M}^{lp}$, $\mathcal{M}^{bel}$, $\mathcal{M}^{poss}$) is NP-complete.*

PROOF. The result for $\mathcal{M}^{prob}$ is immediate from the proof that the satisfiability problem for $\mathcal{L}_1^{QU}$ (the restriction of $\mathcal{L}^{QU}$ to rational coefficents) is NP-complete, together with the argument in Theorem 4.1 showing that every formula in $\mathcal{L}^E$ is equivalent to a formula in $\mathcal{L}_1^{QU}$ of the same length (where the formula is essentially given by the translation of Lemma A.4).

Consider $\mathcal{M}^{lp}$. The lower bound follows from the fact that we can reduce propositional satisfiability to the decision problem for $\mathcal{L}^E$; hence the problem is NP-hard (by replacing each proposition p by the formula $e(p) = 1$). The upper bound follows from Theorem 6.1. Given a formula f, first guess a small model $M = (W, \mathcal{P}, \pi)$, of the form guaranteed to exist by Theorem 6.1. (The fact that $\pi(w)(p) =$ **false** for every world $w \in W$ and every primitive proposition p not appearing in f means that we must describe π only for propositions that appear in f.) We can verify that $M \models f$ inductively. For inequality formulas, let $e(\gamma)$ be an arbitrary expectation term in the formula, with γ of the form $b_1\varphi_1 + \cdots + b_n\varphi_n$. For each φ in γ, we compute $[\![\varphi]\!]_M$ by checking the truth assignment of each world in W and seeing whether this truth assignment makes φ true. We then replace each occurrence of $e(\gamma)$ by $\min_{\mu\in\mathcal{P}}\{\sum_{i=1}^n \sum_{w\in[\![\varphi_i]\!]_M} b_i\mu(w)\}$ and verify that the resulting inequality holds. It is easy to see that this verification can be done in time polynomial in $|f|$ and $||f||$. Therefore, the decision problem is in NP, and hence is NP-complete.

Finally, consider $\mathcal{M}^{bel}$ and $\mathcal{M}^{poss}$. As earlier, the lower bound follows from the fact that we can reduce propositional satisfiability to the decision problem for $\mathcal{L}^E$; hence, the problem is NP-hard. Again, the upper bound follows from Theorem 6.2. Given a formula f, guess a small model $M = (W, \nu, \pi)$ of the form guaranteed to exist by Theorem 6.2, along with sets $U_1, \ldots, U_s$ ($s \leq |f|$) such that the mass function corresponding to the belief function (respectively, possibility measure) ν is positive only on $U_1, \ldots, U_s$. We then verify that f is indeed true at some (and hence all) states in the model, just as in the case of $\mathcal{M}^{lp}$. □

A.5 Proofs for Section 7

The following lemmas are useful in the proof of Theorem 7.1.

Lemma A.12. *The formula $\varphi \geq \tilde{0}$ is provable in $\mathbf{AX}^g$.*

Proof. Here is a sketch of the derivation:

(1) $\mathit{true} = \tilde{1}$ (**G3**)
(2) $\varphi \vee \neg\varphi = \mathit{true}$ (**G4**)
(3) $\varphi \vee \neg\varphi = \varphi + \neg\varphi$ (**G1**)
(4) $\varphi + \neg\varphi = \tilde{1}$ (1, 2, 3, **IneqF**, **Taut**, **MP**)
(5) $\varphi + \neg\varphi \geq \tilde{0}$ (4, **IneqF**)
(6) $\varphi \geq \tilde{0}$ (5, **G2**).

□

Lemma A.13. *The formula $(a\varphi + b\neg\varphi \geq \tilde{0}) \Rightarrow (\varphi = \tilde{0})$ is provable in $\mathbf{AX}^g$, for $a < 0$.*

Proof. Here is a sketch of the derivation:

(1) $a\varphi + b\neg\varphi \geq \tilde{0} \Rightarrow a\varphi \geq \tilde{0}$ (**G2**,**Taut**,**MP**)
(2) $a\varphi \geq \tilde{0} \Rightarrow \varphi \leq \tilde{0}$ (IneqF, since $a < 0$)
(3) $\varphi \geq \tilde{0}$ (Lemma A.12)
(4) $a\varphi + b\neg\varphi \geq \tilde{0} \Rightarrow \varphi = \tilde{0}$ (1, 2, 3, **Taut**, **MP**, definition of =).

□

Theorem 7.1. *$\mathbf{AX}^g$ is a sound and complete axiomatization of $\mathcal{L}^g$ with respect to $\mathcal{M}^g$.*

Proof. Soundness is straightforward. For completeness, we show that an unsatisfiable formula f is inconsistent. We first reduce f to a canonical form. Let $g_1 \vee \ldots \vee g_r$ be a disjunctive normal form expression for f (where each g_i is a conjunction of gamble inequalities and their negations). Using propositional reasoning (axioms **Taut** and **MP**), we can show that f is provably equivalent to this disjunction. Since f is unsatisfiable, each g_i must also be unsatisfiable. Thus, it is sufficient to show that any unsatisfiable conjunction of gamble inequalities and their negations is inconsistent.

Let f be such an unsatisfiable conjunction of gamble inequalities and their negations. Let $p_1, \ldots, p_N$ be the primitive propositions appearing in f, and let $\delta_1, \ldots, \delta_{2^N}$ be a canonical listing of the atoms over $p_1, \ldots, p_N$. We first show that any gamble inequality $\gamma \geq \tilde{c}$ is provably equivalent to a gamble inequality $a_1\delta_1 + \cdots + a_{2^N}\delta_{2^N} \geq \tilde{0}$. Consider a term $a\varphi$ appearing in $\gamma \geq \tilde{c}$. Since φ is equivalent to a disjunction $\delta_{i_1} \vee \ldots \vee \delta_{i_k}$, we have that $\varphi = \delta_{i_1} \vee \ldots \vee \delta_{i_k}$, and hence $a\varphi = a(\delta_{i_1} \vee \ldots \vee \delta_{i_k})$, is provable by **G4** and **IneqF**. By repeated applications of **G1**, we have that $a(\delta_{i_1} \vee \ldots \vee \delta_{i_k}) = a\delta_{i_1} + \cdots + a\delta_{i_k}$ is provable. (Note that if φ and φ' are mutually exclusive, then $\varphi \vee \varphi'$ is equivalent to $\varphi \wedge \varphi \vee \varphi' \wedge \neg\varphi$, so by **G1**, $\varphi \vee \varphi' = \varphi + \varphi'$.) By **IneqF**, $a\varphi = a\delta_{i_1} + \cdots + a\delta_{i_k}$ is provable too. Using **IneqF** again, it follows that $a\varphi = a_1\delta_1 + \cdots + a_{2^N}\delta_{2^N}$ is provable, where $a_i = a$ if $i \in \{i_1, \ldots, i_k\}$, and $a_i = 0$ otherwise. Doing this to every term in γ shows that here exist $b_1, \ldots, b_{2^n}$ such that $\gamma = b_1\delta_1 + \cdots + b_{2^N}\delta_{2^N}$ is provable. Hence, using **IneqF**, so is $a_1\delta_1 + \cdots + a_{2^N}\delta_{2^N} \geq \tilde{c}$. By **G3**, $\mathit{true} = \tilde{1}$ is provable, hence so is $c\,\mathit{true} = \tilde{c}$. It then easily follows using **IneqF** that $c\delta_1 + \cdots + c\delta_{2^N} = \tilde{c}$ is provable. Thus, $\gamma \geq \tilde{c}$ is provably equivalent to $(a_1 - c)\delta_1 + \cdots + (a_{2^N} - c)\delta_{2^N} \geq \tilde{0}$, as required.

It immediately follows that f is provably equivalent to a formula f' that is a conjunction of gamble inequalities of the form $a_1\delta_1 + \cdots + a_{2^N}\delta_{2^N} \geq \tilde{0}$ and their negations. Say that f' consists of r gamble inequalities and s negations of gamble inequalities. Consider two arrays of coefficients of f'': $P = (a_{i,j})$, where $a_{i,j}$ is the coefficient of δ_j in $a_{i,1}\delta_1 + \cdots + a_{i,2^N}\delta_{2^N} \geq \tilde{0}$ $(1 \leq i \leq r)$, and $N = (b_{i,j})$, where $b_{i,j}$ is the coefficient of δ_j in $\neg(b_{i,1}\delta_1 + \cdots + b_{i,2^N}\delta_{2^N} \geq \tilde{0})$ $(1 \leq i \leq s)$. Let $I = \{1, \ldots, 2^N\}$. Note that, since f is unsatisfiable, so is f'. If $a_{i,j} < 0$ for some i, j, then, by Lemma A.13, we must have $\delta_j = \tilde{0}$. Let $I'' = \{j : a_{i,j} < 0 \text{ for some } i\}$; let $I' = I - I''$. Let P' and N' be the result of setting all entries in column j of P (resp., N) to 0, for all $j \in I''$. The formula f'' corresponding to the matrices P' and N' is provably equivalent to the original f', by **IneqF**: since $\delta_j = \tilde{0}$ for all $j \in I''$, it must be the case that $a\delta_j = \tilde{0}$. By construction, all entries in P' and N' are nonnegative; moreover, N' is nonempty (since $\neg(-\delta_1 - \cdots - \delta_{2^N} \geq \tilde{0})$ is in f'').

There are now two cases. Taking $N' = (b'_{i,j})$, note that if N' has a row i with all entries nonnegative, then f'' provably implies that $\neg(0\delta_1 + \cdots + 0\delta_{2^N} \geq \tilde{0})$. But using **IneqF**, it is easy to show that $0\delta_1 + \cdots + 0\delta_{2^N} = \tilde{0}$. This shows that f'', and hence f' and f, is inconsistent. On the other hand, if all the rows in N' have a negative entry, then the formula corresponding to P', N' is in fact satisfiable. We can construct a structure M satisfying f' by taking $M = (\{\delta_i : i \in I'\}, \pi)$, where $\pi(\delta_i)(p) = \mathbf{true}$ if $\delta_i \Rightarrow p$ is a propositional tautology. This contradicts the assumption that f' is unsatisfiable. □

The proof of Theorem 7.2 relies on the following small-model result.

LEMMA A.14. *Suppose that $f \in \mathcal{L}_1^g$ is satisfied in some structure in $\mathcal{M}^g$. Then f is satisfied in a structure (W, π) where $|W| \leq |f|$, and $\pi(w)(p) =$ **false** for every world $w \in W$ and every primitive proposition p not appearing in f.*

PROOF. Suppose that $f \in \mathcal{L}_1^g$ is satisfied in a structure $M = (W, \pi)$. conjunction of gamble inequalities and their negations. We want to show that f is in fact satisfied in a small structure. As usual, we can assume without loss of generality that f is a conjunction of gamble inequalities and their negations.

Suppose that there are r gamble inequalities in f and s negations of gamble inequalities. We consider two cases. If $s = 0$, then pick any $w \in W$ and let $M' = (W', \pi')$, with $W' = \{w\}$, and π' is the restriction of π to $\{w\}$ (setting $\pi(w)(p) =$ **false** for every primitive proposition p not appearing in f. It is easy to check that $M' \models f$, since $w \in W$. If $s > 0$, then for every negation of gamble inequality $\neg(\gamma \geq \tilde{c})$ appearing in f, there exists a world $w \in W$ such that $\{\!|\gamma|\!\}_M(w) < \tilde{c}$. Let $w_1, \ldots, w_s$ be such worlds, one corresponding to each of the s negation of gamble inequalities appearing in f. Let $M' = (W', \pi')$, where $W' = \{w_1, \ldots, w_s\}$, and π' is the restriction of π to W' (setting $\pi'(w)(p) =$ **false** for primitive propositions p not appearing in f). Clearly, every gamble inequality in f is satisfied in M', since W' is a subset of W. Moreover, by choice of W', every negation of gamble inequality in f is also satisfied in M' (since our construction guarantees that there is world in M' that is a witness to the falsity of all the negated gamble inequalities in f). Since $|W'| = s < |f|$, the result follows. □

THEOREM 7.2. *The problem of deciding whether a formula of $\mathcal{L}_1^g$ is satisfiable in $\mathcal{M}^g$ is NP-complete.*

PROOF. For the lower bound, observe that $\mathcal{L}_1^g$ includes propositional reasoning. Hence, the decision problem for $\mathcal{L}_1^g$ is at least as hard as propositional reasoning.

For the upper bound, let f be a satisfiable formula of $\mathcal{L}_1^g$. We first guess a small model $M = (W, \pi)$ for the formula, of the form guaranteed to exist by Lemma A.14. (As usual, the fact that $\pi(w)(p) =$ **false** for every world $w \in W$ and every primitive proposition p not appearing in f means that we must describe π only for propositions that appear in f.) We verify that $M \models f$ inductively. For basic formulas of the form $\gamma \geq c$, we must check that for all $w \in W$, $\{\!|\gamma|\!\}_M(w) \geq c$. If γ is of the form $b_1\varphi_1 + \cdots + b_n\varphi_n$, then we can compute $\{\!|\gamma|\!\}_M(w)$ by summing all the b_i such that φ_i is true at w (i.e., $\pi(w)(\varphi_i) =$ **true**). It is easy to see that this verification can all be done in time polynomial in $|f|$ and $||f||$. □

The following notation is useful for the proofs of Theorem 7.3 and Lemma A.18. Given a linear inequality formula f (over real-valued functions) $t \geq \tilde{c}$, we write $\hat{f}$ for the linear inequality formula $t \geq c$ over the reals. We extend this to Boolean combinations of linear inequality formulas in the obvious way.

For the sake of our proof of completeness of $\mathbf{AX}^f$, we need also to show that the following formula is provable:

$$0v_1 + \ldots + 0v_n \geq \tilde{0} \tag{13}$$

This formula can be viewed as saying that the right implication of axiom **I5** holds when $d = 0$.

LEMMA A.15. *The formula (13) is provable from $\mathbf{AX}^f$.*

PROOF. By **F1**, $v_1 \geq v_1$, that is, $v_1 - v_1 \geq \tilde{0}$, is provable. By axiom **F3**, so is $-v_1 + v_1 \geq \tilde{0}$. If we add these latter two inequalities by **F4**, and delete a 0 term by **F2**, we obtain $0v_1 \geq \tilde{0}$. By using **F2** to add 0 terms, it follows that $0v_1 + \ldots + 0v_n \geq \tilde{0}$ is provable, as desired. □

THEOREM 7.3. $\mathbf{AX}^f$ *is sound and complete for reasoning about formulas about linear inequalities over real-valued functions with nonempty domain.*

PROOF. Soundness is straightforward. For completeness, we show that an unsatisfiable formula f is inconsistent. So suppose that f is unsatisfiable. As usual, without loss of generality, we can assume that f is a conjunction of inequalities and their negations, say, with r inequalities and s negations of inequalities. We prove the result by reducing satisfiability of inequalities over real-valued functions to satisfiability of inequalities over real numbers, and then apply techniques from FHM.

There are two cases. First, suppose that $s = 0$, so that there are no negations of inequalities in f. It is easy to see that since f is unsatisfiable over functions, $\hat{f}$ must be unsatisfiable over the reals. For if $\hat{f}$ were satisfiable over the reals with a solution $x_1^*, \ldots, x_k^*$, then f would be satisfied by taking x_i to be the constant function that always returns x_i^*.

Write $\hat{f}$ in matrix form as $A\vec{x} \geq \vec{b}$, where A is the $r \times k$ matrix of coefficients on the left-hand side of the inequalities, $\vec{x}$ is the column vector $(x_1, \ldots, x_k)$, and $\vec{b}$ is the column vector of the right-hand sides of the inequalities. Since $\hat{f}$ is unsatisfiable, $A\vec{x} \geq \vec{b}$ is unsatisfiable. As in FHM, we make use of the following variant of Farkas' lemma [Farkas 1902] (see Schrijver [1986, page 89]) from linear programming.

LEMMA A.16. *If $A\vec{x} \geq \vec{b}$ is unsatisfiable, then there exists a row vector $\vec{\sigma}$ such that*

(1) $\vec{\sigma} \geq \vec{0}$;
(2) $\vec{\sigma}A = \vec{0}$;
(3) $\vec{\sigma} \cdot \vec{b} > 0$.

Intuitively, $\vec{\sigma}$ is a "witness" or "blatant proof" of the fact that $Ax \geq b$ is unsatisfiable. This is because if there were a vector $\vec{x}$ satisfying $A\vec{x} \geq \vec{b}$, then $0 = (\vec{\sigma}A)\vec{x} = \vec{\sigma}(A\vec{x}) \geq \vec{\sigma}\vec{b} > 0$, a contradiction.

We now show that f must be inconsistent. Let $\vec{\sigma} = (\sigma_1, \ldots, \sigma_r)$ be the row vector guaranteed to exist by Lemma A.16. Either by **F5** or by Lemma A.15 (depending on whether $\sigma_j > 0$ or $\sigma_j = 0$), we can multiply both sides of the j^{th} conjunct of f by σ_j (for $1 \leq j \leq r$), and then use **F4** to add the resulting inequality formulas together. The net result (after deleting some 0 terms by **F2**) is the formula $(0v_1 \geq \tilde{c})$, where $c = \vec{\sigma} \cdot \vec{b} > 0$. From this formula, by **F6**, we can conclude $(0v_1 > \tilde{0})$, which by the definition of $>$ implies $\neg(0v_1 \leq \tilde{0})$, which is in turn an abbreviation for $\neg(-0v_1 \geq \tilde{-0})$, that is, $\neg(0v_1 \geq \tilde{0})$. Thus $f \Rightarrow \neg(0v_1 \geq \tilde{0})$ is provable. However, by Lemma A.15, $(0v_1 \geq \tilde{0})$ is also provable. It follows by propositional reasoning that $\neg f$ is provable, that is, f is inconsistent, as required.

Now suppose that $s > 0$. Let f^+ be the conjunction of the inequalities in f, and let $g_1, \ldots, g_s$ be the negations of inequalities in f. That is, $f = f^+ \wedge g_1 \wedge \ldots \wedge g_s$. We first show that, $\hat{f}^+ \wedge \hat{g}_i$ must be unsatisfiable over the reals for some $i \in \{1, \ldots, s\}$.

Assume by way of contradiction that this is not the case, that is, for all $1 \leq i \leq s$, $\hat{f}^+ \wedge \hat{g}_i$ is satisfiable over the reals. Let $x^*_{1,i}, \ldots, x^*_{k,i}$ be the real number solution to the inequalities $\hat{f}^+ \wedge \hat{g}_i$, and let $D = \{d_1, \ldots, d_s\}$. Define the functions $F^*_1, \ldots, F^*_k$ over D by taking $F^*_j(d_i) = x^*_{j,i}$. It is easy to verify that $f = f^+ \wedge g_1 \wedge \ldots \wedge g_s$ is satisfied by those functions. Intuitively, for every element d_i of the domain, $\hat{f}^+$ is satisifed by $F^*_1(d_i), \ldots, F^*_k(d_i)$, so that f^+ is satisfied by $F^*_1, \ldots, F^*_k$; moreover, each g_i is also satisfied, since by the choice of d_i, we have $a_1 F^*_1(d_i) + \ldots + a_k F^*_k(d_i) < c$. Hence, f is satisfiable over real-valued functions, a contradiction. Therefore, $\hat{f}^+ \wedge \hat{g}_{i_0}$ is unsatisfiable over the real numbers for some i_0. As in the case $s = 0$, we use this fact to show that f is inconsistent.

As before, we can write $\hat{f}^+$ in matrix form as $A\vec{x} \geq \vec{b}$. Similarly, the formula $\hat{g}_{i_0} = \neg(a_1 x_1 + \ldots + a_k x_k \geq c)$ can be written in matrix form as $A'\vec{x} > -c$, where A' is the $1 \times s$ matrix $[-a_1, \ldots, -a_k]$, and $\vec{x}$ is the column vector $(x_1, \ldots, x_k)$. Since $\hat{f}^+ \wedge \hat{g}_{i_0}$ is unsatisfiable, the system $A\vec{x} \geq \vec{b}, A'\vec{x} > -c$ must be unsatisfiable. Farkas' lemma does not apply, but a variant of it, called Motzkin's transposition theorem, which is due to Fourier [1826], Kuhn [1956], and Motzkin [1956] (see Schrijver [1986, page 94]), does.

Lemma A.17. *If the system $A\vec{x} \geq \vec{b}, A'\vec{x} > -c$ is unsatisfiable, then there exist a row vector $\vec{\sigma}$ and a real σ' such that*

(1) $\vec{\sigma} \geq \vec{0}$ and $\sigma' \geq \vec{0}$;
(2) $\vec{\sigma}A + \sigma'A' = \vec{0}$;
(3) either
  (a) $\sigma' = 0$ and $\vec{\sigma} \cdot \vec{b} > 0$, or
  (b) $\sigma' > 0$ and $\vec{\sigma} \cdot \vec{b} - \sigma' c \geq 0$

Since $Ax \geq b, A'x > -c$ is unsatisfiable, let $\vec{\sigma} = (\sigma_1, \ldots, \sigma_r)$ and σ' be the row vector and real guaranteed to exist by Lemma A.17.

If case 3(a) of Lemma A.17 applies, then the situation is identical to that of Lemma A.16, and the same argument shows that f is inconsistent. If case (3b) of Lemma A.17 applies, then $\sigma' > 0$. As before, either by axiom **F5** or by Lemma A.15, we can multiply both sides of the j^{th} conjunct in the formula f^+ by σ_j, for $1 \leq j \leq r$. This results in the following system of inequalities:

$$\begin{array}{rcl} \sigma_1 a_{1,1} v_1 + \cdots + \sigma_1 a_{1,k} v_k & \geq & \widetilde{\sigma_1 c_1} \\ & \cdots & \\ \sigma_r a_{r,1} v_1 + \cdots + \sigma_r a_{r,k} v_k & \geq & \widetilde{\sigma_r c_r}. \end{array} \tag{14}$$

Similarly, by **F5**, we can multiply both sides of g_{i_0} by σ' to get $\neg(\sigma' a_1 v_1 + \ldots \sigma' a_k v_k \geq \widetilde{\sigma' c'})$.

Let $a''_1 v_1 + \cdots + a''_k v_k \geq \tilde{d}$ be the result of "adding" all the inequalities in (14). This inequality is provable from f using **F4**. Since $\vec{\sigma}A + \sigma'A' = \vec{0}$, and $A' = [-a_1, \ldots, -a_k]$, we must have that $-\sigma' a_j = -a''_j$, for $j = 1, \ldots, k$. Thus, $g_{i_0} \Rightarrow \neg(a''_1 v_1 + \cdots + a''_k v_k \geq \sigma'\tilde{c}')$ is provable. Since $\sigma b - \sigma' c' \geq 0$, it follows that $d \geq \sigma' c'$. Therefore, by **F6**, $a''_1 v_1 + \cdots a''_k v_k \geq \tilde{d} \Rightarrow a''_1 v_1 + \cdots a''_k v_k \geq \widetilde{\sigma' c'}$ is provable. Taking the contrapositive, we get that $g_{i_0} \Rightarrow \neg(a''_1 v_1 + \cdots + a''_k v_k \geq \tilde{d})$ is provable. But

$f^+ \Rightarrow a''_1 v_1 + \cdots + a''_k v_k \geq \tilde{d}$ is also provable. Since $f \Rightarrow f^+ \wedge g_{i_0}$ is obviously provable, it follows by propositional reasoning that $\neg f$ is provable, that is, f is inconsistent, as desired. □

As with $\mathcal{L}_1^g$, the proof of NP-completeness for the decision problem of formulas about linear inequalities (with integer coefficients) over real-valued functions with nonempty domain relies on the following small-model result. Here, a "small model" for a formula is an assignment of functions with a small domain. This can be done quickly, as we shall see.

LEMMA A.18. *Suppose that f is a satisfiable inequality formula. Then f has a satisfying assignment where there are at most $|f|$ functions in the assignment with a non-zero range, the functions in the assignment have a domain of size at most $|f|$, and every value in the range of the functions is a rational number with size $O(|f|||f|| + |f|\log(|f|))$.*

PROOF. As usual, we can assume without loss of generality that f is a conjunction of inequality formulas and their negations. Suppose that f has a satisfying assignment $v_1^*, \ldots, v_k^*$, over a domain D. Let f^+ be the conjunction of the inequality formulas in r, and let $g_1, \ldots, g_s$ be the negations of inequality formulas in f. Consider two cases, depending on whether $s = 0$. If $s = 0$, pick some element d of D. Clearly, $v_1^*(d), \ldots, v_k^*(d)$ is a solution to $\hat{f}$, and thus $\hat{f}$ is satisfiable over the reals. By Theorem 4.9 in FHM, $\hat{f}$ is satisfiable with a solution $x_1^*, \ldots, x_k^*$ where at most $|f|$ entries are nonzero, and each nonzero value is a rational number of size $O(|f|||f||+|f|\log(|f|))$. From this solution, we can construct a solution $F_1^*, \ldots, F_k^*$ on the domain $D' = \{d\}$ satisfying the conditions of the theorem, by simply taking $F_i^*(d) = x_i^*$.

Now suppose that $s > 0$. Assume f is satisfiable over functions, with a solution $F_1^*, \ldots, F_k^*$ over a domain D. Clearly, for any $d \in D$, $F_1^*(d), \ldots, F_k^*(d)$ is a solution to $\hat{f}^+$. For any $1 \leq i \leq s$, consider $g_i = \neg(a_1 v_1 + \ldots + a_k v_k \geq \tilde{c})$. There must be a $d \in D$ such that $a_1 F_1^*(d) + \ldots + a_k F_k^*(d) < c$. Thus, $F_1^*(d), \ldots, F_k^*(d)$ is a solution to $\hat{g}_i$, and by the above is also a solution to $\hat{f}^+$. Thus, $\hat{f}^+ \wedge \hat{g}$ is satisfiable over the reals. Again by Theorem 4.9 in FHM, we have a solution $x_{1,i}^*, \ldots, x_{k,i}^*$ of the inequalities $\hat{f}^+ \wedge \hat{g}_i$ such that at most $|f|$ entries in the solution are nonzero, and each nonzero value is of size $O(|f|||f|| + |f|\log(|f|))$. Let $D = \{d_1, \ldots, d_s\}$. We construct a solution to $f^+ \wedge g_1 \wedge \cdots \wedge g_s$ from those solutions. Define the functions $F_1^*, \ldots, F_k^*$ over D such that $F_j^*(d_i) = x_{j,i}^*$. it is easy to verify that $f = f^+ \wedge g_1 \wedge \ldots \wedge g_s$ is satisfied by those functions. Intuitively, for every element d_i of the domain, $\hat{f}^+$ is satisfied by $F_1^*(d_i), \ldots, F_k^*(d_i)$; moreover, each g_i (of the form $\neg(a_1 F_1 + \ldots + a_k F_k \geq \tilde{c})$ must be satisfied, since by choice of d_i, we have $a_1 F_1^*(d_i) + \ldots + a_k F_k^*(d_i) < c$. Clearly this assignment $F_1^*, \ldots, F_k^*$ satisfies the statement of the lemma. □

THEOREM 7.4. *The problem of deciding whether a formula about linear inequalities (with integer coefficients) is satisfiable over real-valued functions with nonempty domain is NP-complete.*

PROOF. For the lower bound, note that we can reduce propositional satisfiability to satisfiability in the logic of linear inequalities by simply replacing each primitive proposition p_i in a propositional formula by the inequality $v_i > 0$.

For the upper bound, let f be a satisfiable inequality formula. We guess a small satisfying assignment where the domain D of the functions has size at most $|f|$; such an assignment is guaranteed to exist if f is satisfiable, by Lemma A.18. It is easy to verify that this assignment does indeed satisfy f in time polynomial in $|f|$ and $||f||$. $\square$

REFERENCES

Billingsley, P. 1995. *Probability and Measure.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons.

Borel, E. 1943. *Les Probabilités et la Vie.* Presses Universitaires de France, Paris. English translation *Probabilities and Life,* New York: Dover, 1962.

Choquet, G. 1953. Theory of capacities. *Annales de l'Institut Fourier (Grenoble) 5*, 131–295.

Chu, F. and Halpern, J. Y. 2003. Great expectations. Part I: On the customizability of generalized expected utility. In *Proc. Eighteenth International Joint Conference on Artificial Intelligence (IJCAI '03).* 291–296.

Chu, F. and Halpern, J. Y. 2004. Great expectations. Part II: Generalized expected utility as a universal decision rule. *Artificial Intelligence 159*, 207–229.

Chvátal, V. 1983. *Linear Programming.* W. Freeman and Co., San Francisco.

Cox, R. 1946. Probability, frequency, and reasonable expectation. *American Journal of Physics 14,* 1, 1–13.

de Finetti, B. 1931. Sul significato soggestivo del probabilità. *Fundamenta Mathematica 17*, 298–329.

Dellacherie, C. 1970. Quelques commentaires sure les prolongements de capacités. In *Séminaire Probabilités, Strasbourg.* Lecture Notes in Mathematics, Volume 191. Springer-Verlag, Berlin and New York.

Dempster, A. P. 1967. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics 38*, 325–339.

Dubois, D. and Prade, H. 1982. On several representations of an uncertain body of evidence. In *Fuzzy Information and Decision Processes*, M. M. Gupta and E. Sanchez, Eds. North-Holland, Amsterdam, 167–181.

Dubois, D. and Prade, H. 1987. The mean value of a fuzzy number. *Fuzzy Sets and Systems 24*, 297–300.

Dubois, D. and Prade, H. 1990. An introduction to possibilistic and fuzzy logics. In *Readings in Uncertain Reasoning*, G. Shafer and J. Pearl, Eds. Morgan Kaufmann, San Francisco, 742–761.

Fagin, R., Halpern, J. Y., and Megiddo, N. 1990. A logic for reasoning about probabilities. *Information and Computation 87,* 1/2, 78–128.

Farkas, J. 1902. Theorie der enfachen ungleichungen. *J. Reine und Angewandte Math. 124*, 1–27.

Fourier, J. B. J. 1826. Solution d'une question particulière du calcul des inégalités. *Nouveau Bulletin des Sciences par la Société philomathique de Paris*, 99–100.

Halpern, J. Y. 2003. *Reasoning About Uncertainty.* MIT Press, Cambridge, Mass.

Halpern, J. Y. and Pucella, R. 2002a. A logic for reasoning about upper probabilities. *Journal of A.I. Research 17*, 57–81.

Halpern, J. Y. and Pucella, R. 2002b. Reasoning about expectation. In *Proc. Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI 2002).* 207–215.

Huber, P. J. 1981. *Robust Statistics.* Wiley Interscience.

Kuhn, H. W. 1956. Solvability and consistency for linear equations and inequalities. *American Mathematical Monthly 63*, 217–232.

Mendelson, E. 1964. *Introduction to Mathematical Logic.* Van Nostrand, New York.

Motwani, R. and Raghavan, P. 1995. *Randomized Algorithms.* Cambridge University Press, Cambridge, England.

Motzkin, T. S. 1956. The assignment problem. In *Numerical Analysis*, J. H. Curtiss, Ed. Proc. Symposia in Applied Mathematics, Volume VI. McGraw-Hill, New York, 109–125.

Ramsey, F. P. 1931. Truth and probability. In *The Foundations of Mathematics and Other Logical Essays*, R. B. Braithwaite, Ed. Routledge and Kegan Paul, London, 156–198.

Rudin, W. 1976. *Principles of Mathematical Analysis*, Third ed. McGraw-Hill.

Rudin, W. 1991. *Functional Analysis*, Second ed. McGraw-Hill.

Savage, L. J. 1954. *Foundations of Statistics*. Wiley, New York.

Schmeidler, D. 1986. Integral representation without additivity. *Proc. of the Amer. Math. Soc. 97,* 2, 255–261.

Schmeidler, D. 1989. Subjective probability and expected utility without additivity. *Econometrica 57*, 571–587.

Schrijver, A. 1986. *Theory of Linear and Integer Programming*. Wiley, New York.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, N.J.

Shores, T. 1999. *Applied Linear Algebra and Matrix Analysis*, Second ed. McGraw-Hill.

Smith, C. A. B. 1961. Consistency in statistical inference and decision. *Journal of the Royal Statistical Society, Series B 23*, 1–25.

Walley, P. 1981. Coherent lower (and upper) probabilities. Manuscript, Department of Statistics, University of Warwick.

Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall.

Wilson, N. and Moral, S. 1994. A logical view of probability. In *Proc. 11th European Conference on Artificial Intelligence (ECAI-94)*. 71–95.