Logics with Lower and Upper Probability Operators

Nenad Savić[a], Dragan Doder[b], Zoran Ognjanović[c]

[a]*Institut of Computer Science, University of Bern*
[b]*Computer Science and Communications, University of Luxembourg*
[c]*Mathematical Institute, Serbian Academy of Sciences and Arts*

Abstract

We present a first-order and a propositional logic with unary operators that speak about upper and lower probabilities. We describe the corresponding class of models, and we discuss decidability issues for the propositional logic. We provide infinitary axiomatizations for both logics and we prove that the axiomatizations are sound and strongly complete. For some restrictions of the logics we provide finitary axiomatic systems.[1]

Keywords: Probabilistic Logic, Upper and Lower Probabilities, Axiomatization, Completeness theorem

1. Introduction

During the last few decades, uncertain reasoning has emerged as one of the main fields in computer science and artificial intelligence. Many different tools are developed for representing and reasoning with uncertain knowledge. One particular line of research concerns the formalization in terms of probabilistic logic. After Nilsson (1986) gave a procedure for probabilistic entailment that, given probabilities of premises, calculates bounds on the probabilities of the derived sentences, researchers from the field started investigations about formal systems for probabilistic reasoning (Fagin et al. (1990); Fagin and Halpern (1994); Fattorosi-Barnaba and Amati (1995); Frisch and

Email addresses: `savic@inf.unibe.ch` (Nenad Savić), `dragan.doder@uni.lu` (Dragan Doder), `zorano@mi.sanu.ac.rs` (Zoran Ognjanović)

[1]This paper is revised and extended version of the conference paper by Savic et al. (2015) presented at Ninth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA 2015).

Haddawy (1994); Halpern (1990); Heifetz and Mongin (2001); Meier (2012); Ognjanovic and Raskovic (1999, 2000)).

However, in many applications, sharp numerical probabilities appear too simple for modeling uncertainty. This calls for developing different imprecise probability models (de Cooman and Hermans (2008); Dubois and Prade (1988); Levi (1980); Miranda (2008); Shafer et al. (1976); Walley (1991, 2000); Zadeh (1978)). In order to model some situations of interest, some approaches use sets of probability measures instead of one fixed measure, and the uncertainty is represented by two boundaries – lower and upper probabilities (Huber (1981); Kyburg (1961)). Consider the following example, essentially taken from Halpern and Pucella (2002).

Example 1. *Suppose that a bag contains 10 marbles and we know that 4 of them are red, and the remaining 6 are either black or green, but we do not know the exact proportion (for example, it is possible that there are no green marbles at all). The goal is to model a situation where the person picks a marble from the bag at random. The cases when person picks up a red marble (red event), when person picks up a black marble (black event) and when person picks up a green marble (green event) will be denoted by R, B and G, respectively. Clearly, the probability of the red event is* 0.4*, but we cannot assign strict probability to black or green event. Therefore, we use the set of probability measures $P = \{\mu_\alpha \mid \alpha \in [0, 0.6]\}$, where μ_α assigns* 0.4 *probability to red event, α to black event, and $0.6 - \alpha$ to green event. We assign two functions to arbitrary set of probability measures P, first one is $P^\star(X) = sup\{\mu(X) \mid \mu \in P\}$ and the second one is $P_\star(X) = inf\{\mu(X) \mid \mu \in P\}$ which will be used to define a range of probabilities, i.e. they will be an upper and a lower probability, respectively.*

One of the main problems in probabilistic logics with non-restricted real-valued semantics is that those formalisms are rich enough to express the type of a proper infinitesimal $\{0 < x < \frac{1}{n} \mid n = 1, 2, 3, \ldots\}$, so the logics are not compact (see Example 10). As an unpleasant logical consequence, for any finitary axiomatic system, there are consistent sets of formulas which are unsatisfiable (van der Hoek (1997)), i.e., the axiomatization is not strongly complete. Halpern and Pucella (2002) provided a finitary axiomatization for propositional reasoning about linear combinations of upper probabilities, and they proved weak completeness (every consistent formula is satisfiable) for the logic. Their formulas are Boolean combinations of the expressions of the form $r_1\ell(\alpha_1) + \ldots + r_n\ell(\alpha_n) \geq r_{n+1}$, where ℓ is the upper probability operator and r_i are real numbers[2], for $i \in \{1, 2, \ldots, n+1\}$.

In this paper, we propose sound and strongly complete (every consistent set of formulas is satisfiable) propositional logic for reasoning about lower and upper probabilities (LUPP), and its first-order extension ($LUPFO$[3]). Our syntax is simpler than the one by Halpern and Pucella (2002), since we don't have the arithmetical operations built into syntax. We extend propositional calculus (in the case of LUPP) and first-order languages (in the case of LUPFO) with modal-like unary operators of the form $U_{\geq s}$ and $L_{\geq s}$, where s ranges over the unit interval of rational numbers. The intended meanings of $U_{\geq s}\alpha$ and $L_{\geq s}\alpha$ are "the upper/lower probability of α is at least s". If $Green$ is a propositional letter, then by the $LUPP$ formula $U_{=0.6}Green$ is a $LUPP$ we can represent the fact that upper probability that a green marble is picked is 0.6 (Example 1). Note that we can also represent the strict probability of choosing a red marble (Red) in LUPP. The formula which assigns the probability 0.4 to that event is $L_{\geq 0.4}Red \wedge U_{\leq 0.4}Red$. In the first-order case, we consider formulas like $L_{\geq 0.5}(\forall x)R(x)$. In natural language, this sentence can represent the statement "the lower probability that it will rain in all the regions of the considered country is at least a half." (Here the lower probability can arise from considering different weather reports, where each report assigns fixed chance of rain to every region and also a fixed chance for raining in all regions.) Now we introduce another example that can be modeled in $LUPFO$.

Example 2. *Authorities of a certain country are worried that Zika virus may be carried into the country across the borders, and have engaged a number of health experts to estimate the probability of at least one infected person entering the country. If the highest estimated probability is over a certain threshold tr, the authorities will institute a restricted border crossing regime. That constraint can be represented in our logic by the formula*

$$U_{\geqslant tr}(\exists)\text{Zika}(x).$$

[2]Halpern and Pucella (2002) define the rich language with formulas with all the reals as coefficients. But, in order to obtain decidability, they have to restrict their language and allow only integer coefficients, i.e. $r_i \in \mathbb{Z}$.

[3]LUP stands for "lower and upper probabilities". The suffixes P and FO indicate that the logic is propositional or first-order, respectively.

The corresponding semantics of our logics consist of special types of Kripke models (possible worlds). In the propositional case, each possible world contains an evaluation of propositional letters, while in the first-order case it contains a first-order structure of a chosen language. In addition, each model is equipped with a set of probability measures defined over the worlds. In order to obtain strong completeness, we use infinitary inference rules. Thus our languages are countable and formulas are finite, while only proofs are allowed to be infinite. We also propose the restricted logics $LUPP^{FR(n)}$ and $LUPFO^{FR(n)}$[4] (for each n in $\mathbb{N} \setminus \{0\}$). For those logics, we achieve compactness using only a finite set of probability values, which is still enough for many practical applications. We propose finitary axiomatization for $LUPP^{FR(n)}$ and $LUPFO^{FR(n)}$.

From the technical point of view, we have modified some of our earlier developed completion methods presented in Ikodinovic et al. (2014); Ilic-Stepic et al. (2014); Ognjanovic and Raskovic (1999); Raskovic et al. (2008). Providing a compete axiom system for the logic is the key issue in formalization of reasoning about upper and lower probabilities. In real-world situations, we usually don't have the complete specifications of systems, but we can obtain probability constrains from different sources. In that way we derive upper and lower probabilities, and the complete axiom system provides tools to deduce formal properties of the considered system.

The contents of this paper are as follows. In Section 2 we recall the notions of lower and upper probability, as well as the representation theorem we use in our axiomatizations. In Section 3 we present the syntax and semantics of the logic $LUPFO$, as well as the axiomatization and we prove some auxiliary propositions. We prove the soundness and completeness of the axiomatization in Section 4. In Section 5 we introduce the logic $LUPP$ and discuss its decidability. We comment the axiomatization of the logic $LUPP$ in Section 6, and the soundness and completeness theorem for that logic in Section 7. In Section 8 we present the finitary logics $LUPP^{FR(n)}$ and $LUPFO^{FR(n)}$, where the probabilities are restricted to a finite set. Section 9 is dedicated to related work and we conclude in Section 10.

[4] FR stands for "finite range"

2. Preliminaries

In this section, we first introduce some basic notions of probability theory that we will use in the paper.

Let $W \neq \emptyset$ and let H be an algebra of subsets of W, i.e., a set of subsets of W such that:

- $W \in H$,
- if $A, B \in H$, then $W \setminus A \in H$ and $A \cup B \in H$.

A function $\mu : H \longrightarrow [0, 1]$ is a finitely additive probability measure, if the following conditions hold:

- $\mu(W) = 1$,
- $\mu(A \cup B) = \mu(A) + \mu(B)$, whenever $A \cap B = \emptyset$.

For a set P of probability measures defined on H, the lower probability measure $P_\star$ and the upper probability measure $P^\star$ are defined by

- $P_\star(X) = \inf\{\mu(X) \mid \mu \in P\}$,
- $P^\star(X) = \sup\{\mu(X) \mid \mu \in P\}$

for every $X \in H$. In the proof of soundness and completeness, we will use the following basic properties of $P_\star$ and $P^\star$:

- $P_\star(X) \leq P^\star(X)$,
- $P_\star(X) = 1 - P^\star(W \setminus X)$,
- $P^\star(X \cup Y) \leq P^\star(X) + P^\star(Y)$, whenever $X \cap Y = \emptyset$.

Now we introduce our running example.

Example 3. *A large data storage manufacturer is in the process of designing a new storage unit. The unit is composed of a large number of hard-disk drives, which the manufacturer can purchase from a number of brands and models.*

For a given drive brand and model, reliability depends on a particular production series. For each model there are statistical data from multiple previous series on the probability of occurrence of a "head crash" (physical contact between the drive head and the magnetic disk) in a given time unit. This event leads to the failure of the affected drive.

Since the storage unit itself is produced in series, each unit may contain a different production series of the chosen drive brand and model. The manufacturer wants to estimate the probability of storage unit failure in the worst case. The failure of one drive has a known probability of causing the failure of the storage unit (a data loss event).

Let the choice (drive brand and model) A have most series with the survival probability of at least 99% (i.e. failure probability 1%), but a few series with probability of 90% (i.e. failure probability 10%). For the choice B, the probability is 94% to 96% for all series. Then the upper probability of drive failure for the choice A is

$$\begin{aligned} P^\star_A(\text{Drive failure}) &= \sup\{\mu((\text{Drive failure})) \mid \mu \text{ ranges over all series of A}\} \\ &= 10\%. \end{aligned}$$

Similarly, the upper probability of drive failure for the choice B is

$$P^\star_B(\text{Drive failure}) = 6\%.$$

If the manufacturer wishes to act cautiously, i.e. to maximize the survival probability in the worst case, they will make the choice B, due to smaller upper probability of drive failure.

In order to axiomatize upper and lower probabilities, we need to completely characterize $P_\star$ and $P^\star$ with a finite number of simple properties that can be represented by some formulas of the logic we will introduce. Many complete characterizations are proposed in the literature, the earliest appears to be by Lorentz (1952). We will use the characterization by Anger and Lembcke (1985) (also used by Halpern and Pucella (Halpern and Pucella, 2002, Theorem 2.3)). We start with the definition of (n, k)-cover.

Definition 1 ((n, k)-cover). *A set A is said to be covered n times by a multiset $\{\{A_1, \ldots, A_m\}\}$ of sets if every element of A appears in at least n sets from $A_1, \ldots, A_m$, i.e., for all $x \in A$, there exist distinct $i_1, \ldots, i_n$ in $\{1, \ldots, m\}$ such that for all $j \leq n$, $x \in A_{i_j}$. An (n, k)-cover of (A, W) is a multiset $\{\{A_1, \ldots, A_m\}\}$ that covers W k times and covers A $n + k$ times.*

We illustrate this definition with the following example:

Example 4. *Let A be a two element set, i.e., $A = \{a, a'\}$. By the previous definition, A is covered 2 times by the (multi)set $\{\{A_1, A_2, A_3\}\}$ if at least two out of these three sets, A_1, A_2, A_3, contain an element a, and also at least two out of these three sets, A_1, A_2, A_3, contain an element a' as well. So, for example, if*

- $A_1 = \{a, b, c\}$,
- $A_2 = \{a, a', d\}$,
- $A_3 = \{a'\}$,

then, the (multi)set $\{\{A_1, A_2, A_3\}\}$ covers the set A two times since $a \in A_1 \cap A_2$ and $a' \in A_2 \cap A_3$.
Let $W = \{a, a', b, c, d\}$. Then the multiset $\{\{A_1, A_2, A_3\}\}$ is the $(1, 1)$-cover of (A, W), since it covers A 2 times, and covers W once. Indeed, every element of W is in at least one of the sets A_1, A_2, A_3, because

$$a \in A_1 \cap A_2$$

$$a' \in A_2 \cap A_3$$

$$b \in A_1$$

$$c \in A_1$$

$$d \in A_2.$$

If, for example

- $A_1 = \{a, b, c\}$,
- $A_2 = \{a, b, c\}$,
- $A_3 = \{a, a'\}$,

then, the multiset $\{\{A_1, A_2, A_3\}\}$ is not the $(1, 1)$-cover of (A, W), because it does not cover the set A two times, since $a' \in A_3$ and $a' \notin A_1 \cup A_2$ (note that also $\{\{A_1, A_2, A_3\}\}$ does not cover W once).

Theorem 1 (Anger and Lembcke (1985)). *Let W be a set, H an algebra of subsets of W, and f a function $f : H \longrightarrow [0,1]$. There exists a set P of probability measures such that $f = P^\star$ iff f satisfies the following three properties:*

(1) $f(\emptyset) = 0$,

(2) $f(W) = 1$,

(3) for all natural numbers m, n, k and all subsets $A_1, \dots, A_m$ in H, if $\{\{A_1, \dots, A_m\}\}$ is an (n,k)-cover of (A, W), then $k + nf(A) \leq \sum_{i=1}^{m} f(A_i)$.

Those three properties are crucial for obtaining the complete axiomatization. The third property is completely described by the axioms (7) and (8) in Section 3.3 (see the explanation after the axiomatization). The first two properties do not have corresponding axioms, but those two simple properties can be derived in the system (Lemma 1 (a) and (b)).

3. The Logic LUPFO

In this section we discuss the syntax and semantics of the logic $LUPFO$. First-order logic is widely recognized as being a fundamental building block in knowledge representation. As usual, if we are interested in a subjective approach to probabilistic first-order reasoning, then we can take the set of possible worlds to be first-order structures.

Let $S = \mathbb{Q} \cap [0,1]$ and $Var = \{x, y, z, \dots\}$ be a denumerable set of variables. Language for the logic $LUPFO$ consists of:

- the elements of the set Var,
- classical propositional connectives $\neg$ and $\wedge$[5],
- universal quantifier $\forall$,
- for every integer $k \geq 0$, denumerably many function symbols $F_0^k, F_1^k, \dots$ of arity k,
- for every integer $k \geq 0$, denumerably many relation symbols $P_0^k, P_1^k, \dots$ of arity k,

[5]Other connectives, $\vee$, $\rightarrow$, $\leftrightarrow$, are defined in the standard way

- the list of upper probability operators $U_{\geq s}$, for every $s \in S$,
- the list of lower probability operators $L_{\geq s}$, for every $s \in S$,
- comma, parentheses.

The function symbols of arity 0 are called constant symbols.

3.1. Formulas

The set of terms is inductively defined as follows:

a) any variable is a term,

b) if $t_1, \ldots t_n$ are terms and F a function symbol, then $F(t_1, \ldots, t_n)$ is a term as well.

Only expressions which can be obtained by finitely many applications of rules *a*) and *b*) are terms.
Classical first-order formulas are defined as usual. We will denote them by $\alpha, \beta, \ldots$, and the set of all classical first-order formulas will be denoted by For_{FO}. Existential quantifier, $\exists x$, is defined as $(\exists x)\alpha := \neg(\forall x)\neg\alpha$. The variable in a formula is free iff it is not quantified.

Example 5. *Continuing Example 3, we introduce the first-order language whose variables represent possible choices of drive brand and model, and the unary relation symbols* Head_Crash, Drive_Death *and* Data_Loss *which represent the head crash, drive failure and storage unit failure (respectively) of a choice (within a given time unit). Then, the fact that the head crash of any drive, certainly leads to the failure of the affected drive, can be represented by the formula*

$$(\forall x)(\text{Head\_Crash}(x) \rightarrow \text{Drive\_Death}(x)).$$

Now we introduce the basic formulas that speak about lower and upper probabilities.

Definition 2. ***(Basic lower and upper probabilistic formulas)*** *If $\alpha \in For_{FO}$ and $s \in S$, then a basic first-order lower probability formula is any formula of the form $L_{\geq s}\alpha$, and a basic first-order upper probability formula is any formula of the form $U_{\geq s}\alpha$.*

Example 6. *Following Example 3, let c be the constant symbol that represent the choice C. Then the basic upper probability formula*

$$U_{\geq 0.05}\text{Head\_Crash}(c)$$

says that the upper probability of head crash for the choice C is at least $0,05$ *(i.e., there is a series of the corresponding model with the probability of head crash at least* $0,05$*).*

Now suppose that $\text{Last\_Longer}(x, y)$ *is a relation that tells that the expected life time of choice* x *is greater than expected life time of choice* y*. Then the formula*

$$U_{\geq 0.5}(\exists x)(\text{Last\_Longer}(x, c))$$

states that the upper probability that there is a choice which lasts longer than C *is at least a half.*

Definition 3 (first-order lower and upper probabilistic formulas). *The set of all first-order lower and upper probabilistic formulas, denoted by* For_{FOP}*, is the smallest set containing all basic first-order lower and upper probability formulas which is closed under Boolean connectives. We will denote formulas from* For_{FOP} *by* $\phi, \psi, \ldots$*, possibly indexed.*

We use the following abbreviations to introduce other types of inequalities: $U_{<s}\alpha$ is $\neg U_{\geq s}\alpha$, $L_{<s}\alpha$ is $\neg L_{\geq s}\alpha$, $U_{\leq s}\alpha$ is $L_{\geq 1-s}\neg\alpha$, $L_{\leq s}\alpha$ is $U_{\geq 1-s}\neg\alpha$, $U_{=s}\alpha$ is $U_{\leq s}\alpha \wedge U_{\geq s}\alpha$, $L_{=s}\alpha$ is $L_{\leq s}\alpha \wedge L_{\geq s}\alpha$, $U_{>s}\alpha$ is $\neg U_{\leq s}\alpha$, $L_{>s}\alpha$ is $\neg L_{\leq s}\alpha$. We also denote both $\alpha \wedge \neg\alpha$ and $\phi \wedge \neg\phi$ by $\perp$ (and similarly for $\top$).

Example 7. *Continuing Example 3, the formula*

$$U_{\geq 0.05}\text{Drive\_Death}(x)) \rightarrow L_{\leq 0.99}\neg\text{Data\_Loss}(x)).$$

says that if upper probability of drive failure of the choice x within given time unit is at least 0.05*, then the lower probability that its data loss will not occur within the time unit is at most* 0.99*.*

Definition 4 (*LUPFO* formulas). *The set of all LUPFO formulas is the set*

$$For = For_{FO} \cup For_{FOP}.$$

LUPFO formulas will be denoted by $\rho, \sigma, \ldots$*, possibly with subscripts.*

Note that formulas are defined in the same style as in the works of Cintula and Noguera (2014); Hájek et al. (1995); Ognjanovic and Raskovic (2000), i.e. neither mixing of pure propositional formulas and lower and upper probabilistic formulas, nor iteration of lower and upper probability operators are allowed.

In the logic $LUPFO$, a sentence is either a classical first-order sentence (a formula with no free variables), or a lower and upper probabilistic formula in which all the operators are applied to classical sentences.

Example 8.

$$(\exists x)P_0^2(x, F_0^0) \wedge P_1^1(F_1^1(x)) \qquad \textit{and} \qquad L_{\geq \frac{1}{2}}(\forall x)P_0^3(F_0^1(x), x, F_1^0),$$

are first-order lower and upper probabilistic formulas, but

$$\alpha \vee U_{<\frac{1}{3}}(\exists x)P_0^1(x) \qquad \textit{and} \qquad L_{=\frac{1}{4}}U_{<1}\beta,$$

are not first-order lower and upper probabilistic formulas.

3.2. Semantics

Definition 5. *A $LUPFO$-model is a tuple $M = \langle W, D, I, \upsilon, H, P\rangle$, where:*

- *W is a non-empty set of objects, i.e., set of (possible) worlds,*
- *D associates with each world w in W a non-empty set $D(w)$ as a domain,*
- *I associates an interpretation of function and relation symbols with every $w \in W$: $I(w)(F_j^n) : D(w)^n \to D(w)$ and $I(w)(R_i^m) \subseteq D(w)^m$,*
- *υ associates a valuation of variables $\upsilon(w)$ with each world $w \in W$, i.e. $\upsilon(w) : Var \to D(w)$,*
- *H is an algebra of subsets of W,*
- *P is a set of finitely additive measures on H.*

It is not hard to see that every world w with its associated domain, variable valuation and interpretation is one classical first-order model. For a given $\alpha \in For_{FO}$, we will use the notation $[\alpha] = \{w \in W \mid$

$\langle D(w), I(w), \upsilon(w)\rangle \models_{FO} \alpha\}$[6]. The class of all $LUPFO$-models M with the property that for every $\alpha \in For_{FO}$, $[\alpha] \in H$, we will denote by $LUPFO_{Meas}$. Elements from $LUPFO_{Meas}$ will be called measurable $LUPFO$-models.

Definition 6 (Satisfiability relation). *The satisfiability relation,* $\models \subseteq LUPFO_{Meas} \times For$, *is defined in the way such that for* $M = \langle W, D, I, \upsilon, H, P\rangle$ *we have:*

- *if* $\alpha \in For_{FO}$, $M \models \alpha$ *iff* $\langle D(w), I(w), \upsilon(w)\rangle \models_{FO} \alpha$, *for every* $w \in W$,
- *if* $\alpha \in For_{FO}$, $M \models U_{\geq s}\alpha$ *iff* $P^{\star}([\alpha]) \geq s$,
- *if* $\alpha \in For_{FO}$, $M \models L_{\geq s}\alpha$ *iff* $P_{\star}([\alpha]) \geq s$,
- $M \models \neg\phi$ *iff it is not the case that* $M \models \phi$,
- $M \models \phi \wedge \psi$ *iff* $M \models \phi$ *and* $M \models \psi$,

where $\phi, \psi \in For_{FO}$.

Example 9. *Suppose that the first-order language contains only one relation symbol, the unary symbol* Head_Crash*. Let* $M = \langle W, D, I, \upsilon, H, P\rangle$ *be a model such that:*

- $W = \{w_1, w_2, w_3\}$.
- $D(w_1) = D(w_2) = D(w_3) = D$, *where* D *is a set of choices which contains the choices* A *and* B.
- I *associates the interpretation* HC *to the relation symbol* Head_Crash *in the following way:* $\mathrm{HC}(w_1) = \{A\}$, $\mathrm{HC}(w_2) = \{B\}$ *and* $\mathrm{HC}(w_3) = \emptyset$.
- υ *is any mapping which associates a valuation of variables with each world* $w \in W$, *such that* $\upsilon(w_1)(x) = B$, $\upsilon(w_2)(x) = A$ *and* $\upsilon(w_3)(x) = B$.
- H *is the algebra of all subsets of* W,
- $P = \{\mu_1, \mu_2\}$, *where* μ_1 *and* μ_2 *are the unique probability measures defined by:*

[6] $\models_{FO}$ stands for the first-order satisfiability relation

– $\mu_1(w_1) = 0.5$, $\mu_1(w_2) = 0.3$, $\mu_1(w_3) = 0.2$,

– $\mu_2(w_1) = 0.2$, $\mu_2(w_2) = 0.4$, $\mu_2(w_3) = 0.4$.

Let us consider satisfiability of the formulas

$\phi = U_{>0}\text{Head\_Crash}(x)$,

$\psi_1 = L_{\geq 0.7}(\exists x)\text{Head\_Crash}(x)$ *and*

$\psi_2 = U_{\geq 0.7}(\exists x)\text{Head\_Crash}(x)$

in the model M*. Note that* $\langle D(w_1), I(w_1), \upsilon(w_1)\rangle \not\models_{FO} \text{Head\_Crash}(x)$*, since* $\upsilon(w_1)(x) = B$ *and* $B \notin \text{HC}(w_1)$*. Similarly,* $\langle D(w_2), I(w_2), \upsilon(w_2)\rangle \not\models_{FO} \text{Head\_Crash}(x)$ *and* $\langle D(w_3), I(w_3), \upsilon(w_3)\rangle \not\models_{FO} \text{Head\_Crash}(x)$*. Consequently,* $[\text{Head\_Crash}(x)] = \emptyset$*, so* $\mu_1([\text{Head\_Crash}(x)]) = \mu_2([\text{Head\_Crash}(x)]) = 0$*. Then* $P^\star([\text{Head\_Crash}(x)]) = \sup\{\mu_1([\text{Head\_Crash}(x)]), \mu_2([\text{Head\_Crash}(x)])\} = 0$*, so we conclude that*

$$M \not\models U_{>0}\text{Head\_Crash}(x).$$

Note that $\langle D(w_1), I(w_1), \upsilon(w_1)\rangle \models_{FO} (\exists x)\text{Head\_Crash}(x)$ *since* $\text{HC}(w_1) = \{A\}$*,* $\langle D(w_2), I(w_2), \upsilon(w_2)\rangle \models_{FO} (\exists x)\text{Head\_Crash}(x)$ *since* $\text{HC}(w_2) = \{B\}$ *and* $\langle D(w_3), I(w_3), \upsilon(w_3)\rangle \not\models_{FO} (\exists x)\text{Head\_Crash}(x)$ *since* $\text{HC}(w_3) = \emptyset$*. Consequently,* $[(\exists x)\text{Head\_Crash}(x)] = \{w_1, w_2\}$*, so* $\mu_1([(\exists x)\text{Head\_Crash}(x)]) = 0.5 + 0.3 = 0.8$ *and* $\mu_2([(\exists x)\text{Head\_Crash}(x)]) = 0.2 + 0.4 = 0.6$*. Then* $P_\star([(\exists x)\text{Head\_Crash}(x)]) = \inf\{0.8, 0.6\} = 0.6$ and $P^\star([(\exists x)\text{Head\_Crash}(x)]) = 0.8$*. Finally, we conclude*

$$M \not\models L_{\geq 0.7}(\exists x)\text{Head\_Crash}(x)$$

and

$$M \models U_{\geq 0.7}(\exists x)\text{Head\_Crash}(x).$$

Definition 7 (Satisfiability of a formula). *A formula* $\rho \in For$ *is satisfiable if there is a measurable LUPFO-model* M *such that* $M \models \rho$*;* ρ *is valid if for every measurable LUPFO-model* M*,* $M \models \rho$*. A set of formulas* T *is satisfiable if there is a measurable LUPFO-model* M *such that* $M \models \rho$ *for every* $\rho \in T$*.*

Example 10. *Consider the set*

$$T = \{\neg U_{=0}\alpha\} \cup \{U_{<\frac{1}{n}}\alpha \mid n \quad is \quad a \quad positive \quad integer\}.$$

Every finite subset of T is satisfiable, but the set T itself is not. Therefore, the compactness theorem, which states that "if every finite subset of T is satisfiable, then T is satisfiable", does not hold for $LUPFO$.

3.3. The Axiomatization Ax_{LUPFO}

Now we are ready to introduce the axiomatic system for the logic $LUPFO$, which we denote by Ax_{LUPFO}. In order to axiomatize upper and lower probabilities, we need to completely characterize $P_\star$ and $P^\star$ with a finite number of properties. Many complete characterizations are proposed in the literature, the earliest appears to be by Lorentz (1952). We will use the characterization by Anger and Lembcke (1985) (also used in (Halpern and Pucella, 2002, Theorem 2.3)).

Axiom schemes

(1) all axioms of the classical propositional logic, separately for formulas from For_{FO} and for formulas from For_{FOP},

(2) $(\forall x)(\alpha \to \beta) \to (\alpha \to (\forall x)\beta)$, where the variable x does not occur free in α and $\alpha, \beta \in For_{FO}$,

(3) $(\forall x)\alpha(x) \to \alpha(t)$, where $\alpha(t)$ is obtained by substitution of all free occurrences of x in the first-order formula $\alpha(x)$ by the term t which is free for x in $\alpha(x)$,

(4) $U_{\leq 1}\alpha \wedge L_{\leq 1}\alpha$

(5) $U_{\leq r}\alpha \to U_{<s}\alpha$, $s > r$

(6) $U_{<s}\alpha \to U_{\leq s}\alpha$

(7) $(U_{\leq r_1}\alpha_1 \wedge \ldots \wedge U_{\leq r_m}\alpha_m) \to U_{\leq r}\alpha$, if $\alpha \to \bigvee_{J\subseteq\{1,\ldots,m\},|J|=k+n} \bigwedge_{j\in J} \alpha_j$ and $\bigvee_{J\subseteq\{1,\ldots,m\},|J|=k} \bigwedge_{j\in J} \alpha_j$ are propositional tautologies, where $r = \frac{\sum_{i=1}^{m} r_i - k}{n}$, $n \neq 0$

(8) $\neg(U_{\leq r_1}\alpha_1 \wedge \ldots \wedge U_{\leq r_m}\alpha_m)$, if $\bigvee_{J\subseteq\{1,\ldots,m\},|J|=k} \bigwedge_{j\in J} \alpha_j$ is a propositional tautology and $\sum_{i=1}^{m} r_i < k$

(9) $L_{=1}(\alpha \to \beta) \to (U_{\geq s}\alpha \to U_{\geq s}\beta)$

Inference Rules

Let $\alpha \in For_{FO}$ and $\rho, \sigma \in For$.

(1) From ρ and $\rho \to \sigma$ infer σ

(2) From α infer $(\forall x)\alpha$

(3) From α infer $L_{\geq 1}\alpha$

(4) From the set of premises

$$\{\phi \to U_{\geq s-\frac{1}{k}}\alpha \mid k \in \mathbb{N}, k \geq \frac{1}{s}\}$$

infer $\phi \to U_{\geq s}\alpha$

(5) From the set of premises

$$\{\phi \to L_{\geq s-\frac{1}{k}}\alpha \mid k \in \mathbb{N}, k \geq \frac{1}{s}\}$$

infer $\phi \to L_{\geq s}\alpha$.

By the axioms $(1)-(3)$, we have that classical first-order logic is sublogic of the $LUPFO$.
Axiom 4 states that the upper bound for upper and lower probabilities is 1. Axioms 5 and 6 state properties of the order of reals; without them even the obvious statements like $U_{\leq\frac{1}{2}}\alpha \to U_{\leq\frac{3}{4}}\alpha$ would not be formally derivable. They are equivalent to Lemma 1 (c) and (d), and they are explicitly used in the proof of Strong Completeness Theorem 5. Axioms 7 and 8 are the logical analogue of the third condition from Theorem 1. Indeed, we can formally write that $\{\{A_1, \ldots, A_m\}\}$ covers a set A n times as

$$A \subseteq \bigcup_{J\subseteq\{1,\ldots,m\},|J|=n} \bigcap_{j\in J} A_j.$$

Therefore, the condition that the formula $\alpha \to \bigvee_{J\subseteq\{1,\ldots,m\},|J|=k+n} \bigwedge_{j\in J} \alpha_j$ is a tautology gives us that $[\alpha]$ is covered $n+k$ times by a multiset $\{\{[\alpha_1], \ldots, [\alpha_m]\}\}$, while the condition that $\bigvee_{J\subseteq\{1,\dots,m\},|J|=k}\bigwedge_{j\in J}\alpha_j$ is a propositional tautology ensures that $W=[\top]$ is covered k times by a multiset $\{\{[\alpha_1],\dots,[\alpha_m]\}\}$. Axiom 9 is crucial for proving that equivalent formulas have equal lower and upper probabilities (see Lemma 1(e)).

Rule 1 is modus ponens, Rule 3 is the lower probability necessitation. Both Rule 4 and Rule 5 are infinitary rules of inference and Rule 4 intuitively says that if upper probability is arbitrary close to s then it is at least s, while Rule 5 intuitively says that if lower probability is arbitrary close to s then it is at least s. They have the role to ensure that some infinitary sets are inconsistent. That is the reason that those two rules are infinite - all formulas from an infinite set should be taken as premises in the corresponding application of a rule. In particular, they are used in the proof of Lemma 2 (b) and (c). This lemma is crucial for proving Lindenbaum Lemma (Theorem 4.2). Finally, the rules 4 and 5 are given in the implicative form to allow a straightforward proof of Deduction theorem.

Definition 8 (Inference relation).

- *$T\vdash\rho$ (ρ is derivable from T) if there is an at most denumerable sequence of formulas $\rho_1,\rho_2,\dots,\rho$, such that every ρ_i is an axiom or a formula from the set T, or it is derived from the preceding formulas by an inference rule;*
- *$\vdash\rho$ (ρ is a theorem) iff $\emptyset\vdash\rho$;*
- *T is consistent if there is at least a formula $\alpha\in For_{FO}$ and a formula $\phi\in For_{FOP}$ that are not deducible from T, otherwise T is inconsistent;*
- *T is maximally consistent set if it is consistent and:*
  (1) for every $\alpha\in For_{FO}$, if $T\vdash\alpha$, then $\alpha\in T$ and $L_{\geq 1}\alpha\in T$
  (2) for every $\phi\in For_{FOP}$, either $\phi\in T$ or $\neg\phi\in T$.
- *T is deductively closed if for every $\rho\in For$, if $T\vdash\rho$, then $\rho\in T$.*

It is easy to check that T is inconsistent iff $T\vdash\bot$. Note that it is not required that for every $\alpha\in For_{FO}$, either α or $\neg\alpha$ belongs to a maximal consistent set (as it is done for formulas from For_{FOP}). Otherwise, it can be proved that in our canonical model, by Rule 3, for each α we would have $L_{\geq 1}\alpha$ or $L_{\geq 1}\neg\alpha$ so the lower (and the upper) probability operator would not make sense.

Example 11. *Now we will use our running Example 3 to illustrate the inference in the introduced axiom system. Suppose that the failure of one drive has a known probability of causing the failure of the storage unit (a data loss event), given by the function f: for the probability p of the drive failure, the probability of the failure of the storage unit is $f(p)$. Suppose also that the statistics showed that for the choice C there are series with the probability of head crash, within a given time unit, at least* 0.05. *Recall that head crash always leads to disk failure. It is easy to conclude that the upper probability of the data loss for the choice C is at least $f(0.05)$. Now we show that this conclusion can be formally derived in the logic $LUPFO$.*

Let T be a set of formulas which contains all the information above. Then, in particular, it contains the following formulas:

$$(\forall x)(\text{Head\_Crash}(x) \to \text{Drive\_Death}(x)) \tag{1}$$

("Head crash certainly leads to the disk failure"),

$$U_{\geq 0.05}\text{Head\_Crash}(c) \tag{2}$$

("Upper probability of head crash for the choice C (whose corresponding constant symbol is c) is at least 0.05*"), and the set of formulas*

$$\{U_{\geq s}\text{Drive\_Death}(c) \to U_{\geq f(s)}\text{Data\_Loss}(c) \mid s \in \mathbb{Q} \cap [0,1]\}. \tag{3}$$

From the formula (1) and Axiom 3 we have

$$T \vdash \text{Head\_Crash}(c) \to \text{Drive\_Death}(c). \tag{4}$$

If we apply Rule 3 to (4), we obtain

$$T \vdash L_{=1}(\text{Head\_Crash}(c) \to \text{Drive\_Death}(c)). \tag{5}$$

From Axiom 9 we obtain

$$\vdash L_{=1}(\text{Head\_Crash}(c) \to \text{Drive\_Death}(c)) \to \tag{6}$$

$$(U_{\geq 0.05}\text{Head\_Crash}(c) \to U_{\geq 0.05}\text{Drive\_Death}(c)). \tag{7}$$

From (5) and (6), using Modus Ponens, we derive

$$T \vdash U_{\geq 0.05}\text{Head\_Crash}(c) \to U_{\geq 0.05}\text{Drive\_Death}(c). \tag{8}$$

Now, by applying Modus Ponens to (2) and (8) we conclude

$$T \vdash U_{\geq 0.05}\text{Drive\_Death}(c). \tag{9}$$

Finally, from (3) and (9) we obtain

$$T \vdash U_{\geq f(0.05)}\text{Data\_Loss}(c). \tag{10}$$

Theorem 2 (Deduction theorem). *Let T be a set of sentences. Then $T \cup \{\rho\} \vdash \sigma$ iff $T \vdash \rho \to \sigma$, where either both $\rho, \sigma \in For_{FO}$, or $\rho, \sigma \in For_{FOP}$.*

Proof. The only interesting case is when $\rho,\sigma \in For_{FOP}$.
($\Leftarrow$) Direct consequence of Rule 1.
($\Rightarrow$) Suppose that $T \cup \{\rho\} \vdash \sigma$. We will use the induction on the length of the inference.

The cases when either $\vdash \sigma$ or $\rho = \sigma$ or σ is obtained by application of Modus Ponens are the same as in the classical propositional case. Thus, let us consider the case where $\sigma = (\forall x)\sigma_1$ is obtained by an application of Rule 2 from $T, \rho \vdash \sigma_1$. In that case we have:

$T \vdash \rho \to \sigma_1$,by the induction hypothesis

$T \vdash (\forall x)(\rho \to \sigma_1)$, by Rule 2

$T \vdash (\forall x)(\rho \to \sigma_1) \to (\rho \to (\forall x)\sigma_1)$, Axiom 2

$T \vdash \rho \to \sigma$, by Rule 1

Now, let $\sigma = L_{\geq 1}\alpha$ be obtained from $T \cup \{\rho\}$ by an application of Rule 3. In that case:

- $T, \rho \vdash \alpha$
- $T, \rho \vdash L_{\geq 1}\alpha$ by Rule 3

However, since $\alpha \in For_{FO}$ and $\rho \in For_{FOP}$, ρ cannot affect the proof of α from $T \cup \{\rho\}$, and we have:

(1) $T \vdash \alpha$

(2) $T \vdash L_{\geq 1}\alpha$ by Rule 3

(3) $T \vdash L_{\geq 1}\alpha \to (\rho \to L_{\geq 1}\alpha)$

(4) $T \vdash \rho \to L_{\geq 1}\alpha$ by Rule 1.

Next, let us consider the case where $\sigma = \sigma_1 \to U_{\geq s}\alpha$ is obtained from $T \cup \{\rho\}$ by an application of Rule 4. Then:

(1) $T, \rho \vdash \sigma_1 \to U_{\geq s-\frac{1}{k}}\alpha$, for all $k \geq \frac{1}{s}$

(2) $T \vdash \rho \to (\sigma_1 \to U_{\geq s-\frac{1}{k}}\alpha)$, by the induction hypothesis

(3) $T \vdash (\rho \wedge \sigma_1) \to U_{\geq s-\frac{1}{k}}\alpha$

(4) $T \vdash (\rho \wedge \sigma_1) \to U_{\geq s}\alpha$, by Rule 4

(5) $T \vdash \rho \to \sigma$.

If the formula is obtained by an application of Rule 5, the proof is similar. □

We will not always explicitly emphasize moments in proofs where we use Deduction theorem.

Proposition 1. *$\vdash U_{\leq r}\alpha \to L_{\leq r}\alpha$.*

Proof. We consider two cases.

(1) $r \neq 1$. From Axiom (8) we obtain that $\neg(U_{\leq r}\alpha \wedge U_{\leq s}\neg\alpha)$, whenever $r + s < 1$. Therefore $U_{\leq r}\alpha \to U_{>s}\neg\alpha$, and because that holds for every $s < 1 - r$, by inference rule (4) we have $U_{\leq r}\alpha \to U_{\geq 1-r}\neg\alpha$, i.e. $U_{\leq r}\alpha \to L_{\leq r}\alpha$.

(2) $r = 1$. Direct consequence of Axiom (4). □

Consequently, we obtain that

$$\vdash L_{\geq r}\alpha \to U_{\geq r}\alpha$$

for each $r \in S$.

Lemma 1.

(a) $\vdash U_{=1}\top$

(b) $\vdash U_{=0}\bot$

(c) $\vdash U_{\geq s}\alpha \to U_{>r}\alpha$, $s > r$

(d) $\vdash U_{>s}\alpha \to U_{\geq s}\alpha$

(e) *If* $T \vdash \alpha \leftrightarrow \beta$ *then* $T \vdash U_{\geq s}\alpha \leftrightarrow U_{\geq s}\beta$

Proof.
The proofs of (a) and (b) are straightforward, (c) and (d) are obtained from Axioms (5) and (6) and contraposition, and (e) is direct consequence of Rule (3) and Axiom (9). □

4. Soundness and Completeness

4.1. Soundness

Theorem 3 (Soundness). *The axiomatic system Ax_{LUPFO} is sound with respect to the class of measurable $LUPFO$-models.*

Proof. Our goal is to show that every instance of an axiom schemata holds in every model and that the inference rules preserve the validity. The soundness theorem for the first-order logic implies that every instance of an axioms (1) − (3) hold in every model, and that inference rule (2) preserve validity.
Now, for example, let us consider Axiom (7). Suppose that

$$\alpha \to \bigvee_{J\subseteq\{1,\ldots,m\},|J|=k+n} \bigwedge_{j\in J} \alpha_j$$

and

$$\bigvee_{J\subseteq\{1,\ldots,m\},|J|=k} \bigwedge_{j\in J} \alpha_j$$

are propositional tautologies, and suppose that

$$(U_{\leq r_1}\alpha_1 \wedge \ldots \wedge U_{\leq r_m}\alpha_m)$$

holds in a model $M = \langle W, D, I, \upsilon, H, P\rangle$. We already explained that this means that a multiset $\{\{[\alpha_1], \ldots, [\alpha_m]\}\}$ is an (n, k)-cover of $([\alpha], [\top])$. Also, the inequalities $P^\star([\alpha_1]) \leq r_1, \ldots, P^\star([\alpha_m]) \leq r_m$ hold, by assumption. Since $P^\star$ is an upper probability measure, by Theorem 1, we know that

$$k + nP^\star([\alpha]) \leq \sum_{i=1}^{m} P^\star([\alpha_i]),$$

so we obtain that

$$P^{\star}([\alpha]) \leq \frac{\sum_{i=1}^{m} r_i - k}{n}, n \neq 0$$

therefore

$$P^{\star}([\alpha]) \leq r, \quad r = \frac{\sum_{i=1}^{m} r_i - k}{n},$$

i.e. $M \models U_{\leq r}\alpha$ as well.

For Axiom (8), suppose that

$$\bigvee_{J \subseteq \{1,\ldots,m\}, |J|=k} \bigwedge_{j \in J} \alpha_j$$

is a propositional tautology, i.e., a multiset $\{\{[\alpha_1], \ldots, [\alpha_m]\}\}$ covers W k-times, and that $\sum_{i=1}^{m} r_i < k$. The goal is to show that

$$\neg(U_{\leq r_1}\alpha_1 \wedge \ldots \wedge U_{\leq r_m}\alpha_m)$$

holds in a model M. Suppose that it is not the case, i.e.,

$$M \models U_{\leq r_1}\alpha_1 \wedge \ldots \wedge U_{\leq r_m}\alpha_m.$$

Then, we have that $P^{\star}([\alpha_1]) \leq r_1, \ldots, P^{\star}([\alpha_m]) \leq r_m$, hence

$$\sum_{i=1}^{m} P^{\star}([\alpha_i]) \leq \sum_{i=1}^{m} r_i < k.$$

By Theorem 1, we have that

$$\sum_{i=1}^{m} P^{\star}([\alpha_i]) \geq k.$$

Contradiction, so $M \models \neg(U_{\leq r_1}\alpha_1 \wedge \ldots \wedge U_{\leq r_m}\alpha_m)$.

Consider now the Axiom (9). If $M \models L_{=1}(\alpha \to \beta)$, we have that

$$P_{\star}([\alpha \to \beta]) = 1,$$

so

$$P^{\star}([\alpha \wedge \neg\beta]) = 1 - P_{\star}([\alpha \to \beta]) = 0.$$

Therefore

$$P^{\star}([\alpha]) = P^{\star}([\alpha \wedge \beta] \cup [\alpha \wedge \neg\beta]) \leq P^{\star}([\alpha \wedge \beta]) + P^{\star}([\alpha \wedge \neg\beta]) \leq P^{\star}([\beta]).$$

Hence, if $P^{\star}([\alpha]) \geq s$, then $P^{\star}([\beta]) \geq s$, so $M \models U_{\geq s}\alpha \to U_{\geq s}\beta$. The other axioms can be proved to be valid in a similar way.

Rule (1) is validity-preserving for the same reason as in classical logic. Rule (2) is similar to the classical case. Rule (3): if α holds in $M = \langle W, D, I, \upsilon, H, P\rangle$, then $[\alpha] = W$, and therefore $\mu([\alpha]) = 1$ for every $\mu \in P$. Then $P_{\star}([\alpha]) = 1$, so $M \models L_{\geq 1}\alpha$. Rule (4): Suppose that $M \models \phi \to U_{\geq s-\frac{1}{k}}\alpha$ whenever $k \geq \frac{1}{s}$. If $M \not\models \phi$, then obviously $M \models \phi \to U_{\geq s}\alpha$. Otherwise $M \models U_{\geq s-\frac{1}{k}}\alpha$ for every $k \geq \frac{1}{s}$, so $M \models U_{\geq s}\alpha$ because of the properties of the set of reals. Rule (5) is validity-preserving for the same reason as Rule (4). □

4.2. Completeness

In order to prove the completeness theorem we start with some auxiliary statements. After that, we show how to extend a consistent set of formulas T to a maximal consistent set of formulas $T^{\star}$. Finally, we construct the canonical model using the set $T^{\star}$ such that $M_{T^{\star}} \models \rho$ iff $\rho \in T^{\star}$.

Lemma 2. *Let $T \subseteq For$ be a consistent set.*

(1) For any formula $\phi \in For_{FOP}$, either $T \cup \{\phi\}$ is consistent or $T \cup \{\neg\phi\}$ is consistent.

(2) If $\neg(\phi \to U_{\geq s}\alpha) \in T$, then there is some $n > \frac{1}{s}$ such that $T \cup \{\phi \to \neg U_{\geq s-\frac{1}{n}}\alpha\}$ is consistent.

(3) If $\neg(\phi \to L_{\geq s}\alpha) \in T$, then there is some $n > \frac{1}{s}$ such that $T \cup \{\phi \to \neg L_{\geq s-\frac{1}{n}}\alpha\}$ is consistent.

Proof.

(1) If $T \cup \{\phi\} \vdash \bot$, and $T \cup \{\neg\phi\} \vdash \bot$, then by Deduction theorem we have $T \vdash \neg\phi$ and $T \vdash \phi$. Contradiction.

(2) Suppose that for all $n > \frac{1}{s}$:

$$T, \phi \to \neg U_{\geq s-\frac{1}{n}}\alpha \vdash \bot.$$

Therefore, by Deduction theorem and propositional reasoning, we have

$$T \vdash \phi \to U_{\geq s-\frac{1}{n}}\alpha,$$

and by application of Rule 4 we obtain $T \vdash \phi \to U_{\geq s}\alpha$. Contradiction with the fact that $\neg(\phi \to U_{\geq s}\alpha) \in T$.

(3) can be proved in a similar way. □

Now we prove the Lindenbaum lemma for the logic $LUPFO$.

Theorem 4. *Every consistent set of formulas can be extended to a maximal consistent set.*

Proof. Consider a consistent set T. By $Cn_{FO}(T)$ we will denote the set of all first-order formulas that are consequences of T. Let $\phi_0, \phi_1, \ldots$ be an enumeration of all formulas from For_{FOP}. We define a sequence of sets T_i, $i = 0, 1, 2, \ldots$ as follows:

(1) $T_0 = T \cup Cn_{FO}(T) \cup \{L_{\geq 1}\alpha \mid \alpha \in Cn_{FO}(T)\}$

(2) for every $i \geq 0$,

  (a) if $T_i \cup \{\phi_i\}$ is consistent, then $T_{i+1} = T_i \cup \{\phi_i\}$, otherwise

  (b) if ϕ_i is of the form $\psi \to U_{\geq s}\beta$, then $T_{i+1} = T_i \cup \{\neg\phi_i, \psi \to \neg U_{\geq s-\frac{1}{n}}\beta\}$, for some positive integer n, so that T_{i+1} is consistent, otherwise

  (c) if ϕ_i is of the form $\psi \to L_{\geq s}\beta$, then $T_{i+1} = T_i \cup \{\neg\phi_i, \psi \to \neg L_{\geq s-\frac{1}{n}}\beta\}$, for some positive integer n, so that T_{i+1} is consistent, otherwise

  (d) $T_{i+1} = T_i \cup \{\neg\phi_i\}$.

(3) $T^\star = \bigcup_{i=0}^{\infty} T_i$.

The set T_0 is obviously consistent because it contains consequences of an consistent set. Note that existence of the natural numbers (n) from the steps $2(b)$ and $2(c)$ of the construction is provided by Lemma 2, and each T_i is consistent.

It still remains to show that $T^\star$ is maximal consistent set. The steps (1) and (2) of the above construction ensure that $T^\star$ is maximal.

$T^\star$ obviously doesn't contain all formulas. If $\alpha \in For_{FO}$, by the construction of T_0, α and $\neg\alpha$ can not be both in T_0. For a formula $\phi \in For_{FOP}$, the set $T^\star$ does not contain both $\phi = \phi_i$ and $\neg\phi = \phi_j$, because the set $T_{\max\{i,j\}+1}$ is consistent.

Let us prove that $T^\star$ is deductively closed. If a formula $\alpha \in For_{FO}$ and $T \vdash \alpha$, then by the construction of T_0, $\alpha \in T^\star$ and $L_{\geq 1}\alpha \in T^\star$. Let

$\phi \in For_{FOP}$. It can be easily proved (induction on the length of the inference) that if $T^\star \vdash \phi$, then $\phi \in T^\star$. Note the fact that, if $\phi = \phi_j$ and $T_i \vdash \phi$ it has to be $\phi \in T^\star$ because $T_{\max\{i,j\}+1}$ is consistent.
Suppose that the sequence $\phi_1, \phi_2, \ldots, \phi$ is the proof of ϕ from $T^\star$. If the mentioned sequence is finite, there must be some set T_i such that $T_i \vdash \phi$, and $\phi \in T^\star$. Therefore, suppose that the sequence is countably infinite. We can show that, for every i, if ϕ_i is obtained by an application of an arbitrary inference rule, and all the premises belong to $T^\star$, then, also $\phi_i \in T^\star$. If the inference rule is finitary one, then there must be a set T_j which contains all the premises and $T_j \vdash \phi_i$. So, we conclude that $\phi_i \in T^\star$.
Now, consider the infinitary Rule 4. Let $\phi_i = \psi \to U_{\geq s}\alpha$ be obtained from the set of premises $\{\phi_i^k = \psi \to U_{\geq s_k}\alpha \mid s_k = s - \frac{1}{k}, k > \frac{1}{s}, k \in \mathbb{N}\}$. By the induction hypothesis, we have that $\phi_i^k \in T^\star$, for every k. If $\phi_i \notin T^\star$, by step (2)(b) of the construction, there are some l and j so that

$$\neg(\psi \to U_{\geq s}\alpha),\ \psi \to \neg U_{\geq s-\frac{1}{l}}\alpha \in T_j.$$

Thus, we have that for some $j' \geq j$:
- $\psi \wedge \neg U_{\geq s}\alpha \in T_{j'}$,
- $\psi \in T_{j'}$,
- $\neg U_{\geq s-\frac{1}{l}}\alpha$, $U_{\geq s-\frac{1}{l}}\alpha \in T_{j'}$.

Contradiction with the consistency of a set $T_{j'}$.
If we consider the infinitary Rule 5, the proof is similar.

Thus, $T^\star$ is deductively closed set which does not contain all formulas, so it is consistent. □

Now we use $T^\star$ to define a measurable structure $M_{T^\star}$. The main idea is that maximality of $T^\star$ ensures that for every formula α and every rational number from the unit interval s, the set $T^\star$ contains one of the following formulas: $U_{>}\alpha$, $U_{<}\alpha$ or $U_{=}\alpha$. Then we can use the density of rational numbers in the set of reals to define $P^\star$. Then the axioms 7 and 8 and Lemma 1 (a) and (b) will guarantee that all the three conditions from Theorem 1 are fulfilled, which provide us with a set of probability measures for the structure $M_{T^\star}$.

Let D be a countably infinite set. We associate each object $w_{(\bar{v},\bar{I})}$ with a pair $(\bar{v}, \bar{I})$, where $\bar{v}$ is a valuation on D and $\bar{I}$ is an interpretation of relation and function symbols.

Definition 9. *If $T^\star$ is the maximally consistent set of formulas, then a canonical model $M_{T^\star} = \langle W, D, I, v, H, P\rangle$ is defined as follows:*

- $W = \{w_{(\bar{v},\bar{I})} \mid w_{(\bar{v},\bar{I})} \models Cn_{FO}(T)\}$ *contains all first-order interpretations that satisfy the set* $Cn_{FO}(T)$,
- $I(w_{(\bar{v},\bar{I})}) = \bar{I}$ *and* $v(w_{(\bar{v},\bar{I})}) = \bar{v}$, *for* $w_{(\bar{v},\bar{I})} \in W$,
- $H = \{[\alpha] \mid \alpha \in For_{FO}\}$, *where* $[\alpha] = \{w_{(\bar{v},\bar{I})} \in W \mid w_{(\bar{v},\bar{I})} \models \alpha\}$,
- P *is any set of probability measures such that* $P^\star([\alpha]) = \sup\{s \mid U_{\geq s}\alpha \in T^\star\}$.

Note that as a consequence of the fact that those two boundaries don't capture all the information contained in the set of probabilities, $M_{T^\star}$ is not unique for a given $T^\star$., i.e., there might be more different sets of probability measures whose upper and lower probability are $P^\star$ and $P_\star$.

Lemma 3. *Let* $T^\star$ *be a maximal consistent set of formulas. Then,* $M_{T^\star} \in LUPFO_{Meas}$.

Proof. First, we prove that $M_{T^\star}$ is well defined.
The proof that H is an algebra is straightforward.
$P^\star([\alpha]) := \sup\{s \mid U_{\geq s}\alpha \in T^\star\}$ is well defined because $[\alpha] = [\beta]$ implies $\sup\{s \mid U_{\geq s}\alpha \in T^\star\} = \sup\{s \mid U_{\geq s}\beta \in T^\star\}$, by Lemma 1(e). Let $f([\alpha]) = \sup\{s \mid U_{\geq s}\alpha \in T^\star\}$. We want to prove that f is an upper probability measure for some set of probability measures P, i.e., there exists P such that $f = P^\star$. It is sufficient to prove the three conditions from Theorem 1. Using Lemma 1 (a), (b), the conditions $f(\emptyset) = 0$ and $f(W) = 1$ become trivial to prove. The only thing left to prove is that if $\{\{[\alpha_1], \ldots, [\alpha_m]\}\}$ is (n, k)-cover of $([\alpha], W)$, then $k + nf([\alpha]) \leq \sum_{i=1}^{m} f([\alpha_i])$.

Let $f([\alpha_i]) = a_i$, i.e. $\sup\{r \mid U_{\geq r}\alpha_i \in T^\star\} = a_i$, $i = 1, \ldots, m$. For arbitrary $\varepsilon > 0$ there exists rational numbers $q_i \in [a_i, a_i + \varepsilon]$ such that $U_{\leq q_i}\alpha_i \in T^\star$ (otherwise $U_{>q_i}\alpha_i \in T^\star$ which is contradiction with the fact that a_i is supremum). Hence, we have

$$T^\star \vdash U_{\leq q_1}\alpha_1 \wedge \ldots \wedge U_{\leq q_m}\alpha_m,$$

and by Axiom 7, we have

$$T^\star \vdash U_{\leq q}\alpha, \quad q = \frac{\sum_{i=1}^{m} q_i - k}{n}, n \neq 0,$$

i.e.,

$$\sup\{r \mid U_{\geq r}\alpha \in T^\star\} \leq q$$

or

$$f([\alpha]) \leq q.$$

Therefore, we have

$$f([\alpha]) \leq \frac{\sum_{i=1}^{m} q_i - k}{n} \leq \frac{\sum_{i=1}^{m} a_i + m\varepsilon - k}{n}, \quad q_i \leq a_i + \varepsilon,$$

and because this holds for every $\varepsilon > 0$ we obtain $k + nf([\alpha]) \leq \sum_{i=1}^{m} f([\alpha_i])$.

If $n = 0$, we need to show that $k \leq \sum_{i=1}^{m} f([\alpha_i])$. Reasoning as above, we have that $T^\star \vdash U_{\leq q_1}\alpha_1 \wedge \ldots \wedge U_{\leq q_m}\alpha_m$, for some $q_i \in [a_i, a_i + \varepsilon]$, and because of Axiom (8), and $\bigvee_{J\subseteq\{1,\ldots,m\},|J|=k} \bigwedge_{j\in J} \alpha_j$ are propositional tautologies, we have that $\sum_{i=1}^{m} q_i \geq k$. Since that holds for every $\varepsilon > 0$, we obtain $\sum_{i=1}^{m} a_i \geq k$.

Directly from the construction of $M_{T^\star}$ we have that $M_{T^\star} \in LUPFO_{Meas}$. $\square$

Now we are ready to prove the main result of this paper.

Theorem 5 (Strong completeness). *A set of formulas T is consistent iff it is satisfiable.*

Proof. Direction from right to left follows from the Soundness Theorem. For the proof of the other direction we extend a consistent set to a maximal consistent set using Theorem 4 and then we construct $LUPFO_{Meas}$-model $M_{T^\star}$ as in Definition 9 (Lemma 3 guarantee that it is a measurable $LUPFO$-model), and show that for every $\rho \in For$, $M_{T^\star} \models \rho$ iff $\rho \in T^\star$. We use the induction on the complexity of the formula.

- $\rho = \alpha \in For_{FO}$. If $\alpha \in Cn_{FO}(T)$, then by definition of $M_{T^\star}$ we have $M_{T^\star} \models \alpha$. Conversely, if $M_{T^\star} \models \alpha$, by the completeness of classical first-order logic we have that $\alpha \in Cn_{FO}(T)$.

- Consider the case when $\rho = U_{\geq s}\alpha$. If $U_{\geq s}\alpha \in T^\star$, then

$$\sup\{r \mid U_{\geq r}\alpha \in T^\star\} = P^\star([\alpha]) \geq s,$$

and so $M_{T^\star} \models U_{\geq s}\alpha$. Now, suppose that $M_{T^\star} \models U_{\geq s}\alpha$, i.e.

$$\sup\{r \mid U_{\geq r}\alpha \in T^\star\} \geq s.$$

a) If $P^\star([\alpha]) > s$, then by the properties of supremum and monotonicity of $P^\star$ (Lemma 1 (c), (d)), we have $U_{\geq s}\alpha \in T^\star$.

b) If $P^\star([\alpha]) = s$, then, as a direct consequence of inference Rule 4, we have that $U_{\geq s}\alpha \in T^\star$.

- Next, let $\rho = L_{\geq s}\alpha$, i.e. $\rho = U_{\leq 1-s}\neg\alpha$. First, suppose that $U_{\leq 1-s}\neg\alpha \in T^\star$. We want to show that

$$\sup\{r \mid U_{\geq r}\neg\alpha \in T^\star\} \leq 1 - s,$$

so suppose that

$$\sup\{r \mid U_{\geq r}\neg\alpha \in T^\star\} > 1 - s.$$

Then, there exist a rational number $q \in (1-s, 1-s+\epsilon]$, for some $\epsilon > 0$, such that

$$U_{\geq q}\neg\alpha \in T^\star.$$

Hence, $U_{>1-s}\neg\alpha \in T^\star$ which leads us to contradiction. So,

$$\sup\{r \mid U_{\geq r}\neg\alpha \in T^\star\} \leq 1 - s,$$

i.e.

$$P^\star([\neg\alpha]) \leq 1 - s$$

and thus we obtain $M_{T^\star} \models L_{\geq s}\alpha$. Now, for the other direction, suppose that $M_{T^\star} \models U_{\leq 1-s}\neg\alpha$, i.e. $\sup\{r \mid U_{\geq r}\neg\alpha \in T^\star\} \leq 1-s$. Consider the following two cases:

(1) $\sup\{r \mid U_{\geq r}\neg\alpha \in T^\star\} < 1 - s$. Then, if $U_{>1-s}\neg\alpha \in T^\star$, then also $U_{\geq 1-s}\neg\alpha \in T^\star$, so $\sup\{r \mid U_{\geq r}\neg\alpha \in T^\star\} \geq 1 - s$. Contradiction.

(2) $\sup\{r \mid U_{\geq r}\neg\alpha \in T^\star\} = 1 - s$. We want to show that then

$$\inf\{r \mid U_{\leq r}\neg\alpha \in T^\star\} = 1 - s$$

must hold as well. First, suppose that

$$\inf\{r \mid U_{\leq r}\neg\alpha \in T^\star\} < 1 - s.$$

Hence, there exist a rational number $q_1 \in [1-s-\epsilon, 1-s)$ such that

$$U_{\leq q_1}\neg\alpha \in T^\star,$$

and so $U_{<1-s}\neg\alpha \in T^\star$, contradiction with the fact that $U_{\geq 1-s}\neg\alpha \in T^\star$ (direct consequence of inference rule (4)). Now, suppose that

$$\inf\{r \mid U_{\leq r}\neg\alpha \in T^\star\} > 1 - s,$$

i.e.

$$\inf\{r \mid U_{\leq r}\neg\alpha \in T^\star\} = 1 - s + \varepsilon.$$

Take an arbitrary rational number $q_2 \in (1-s, 1-s+\varepsilon)$ and then both $U_{\leq q_2}\neg\alpha \in T^\star$ and $U_{\geq q_2}\neg\alpha \in T^\star$ leads us to contradiction (because of the properties of infimum and supremum), which is impossible. Therefore,

$$\inf\{r \mid U_{\leq r}\neg\alpha \in T^\star\} = 1 - s,$$

or equvivalently

$$\inf\{r \mid L_{\geq 1-r}\alpha \in T^\star\} = 1 - s$$

and then, by the inference Rule 5, we obtain that $L_{\geq s}\alpha \in T^\star$.

- Now, let $\rho = \neg\psi \in For_{FOP}$. Then $M_{T^\star} \models \neg\psi$ iff it is not the case that $M_{T^\star} \models \psi$ iff $\psi \notin T^\star$ iff $\neg\psi \in T^\star$.

- Finally, let $\rho = \phi \wedge \psi \in For_{FOP}$. Then, $M_{T^\star} \models \phi \wedge \psi$ iff $M_{T^\star} \models \phi$ and $M_{T^\star} \models \psi$ iff $\phi, \psi \in T^\star$ iff $\phi \wedge \psi \in T^\star$. □

5. The logic *LUPP*

In this section we will describe the syntax and semantics of the logic $LUPP$, and we discuss the decidability problem of satisfiability of $LUPP$-formulas.

5.1. Syntax

Let S be the set of rational numbers from $[0, 1]$ and let $\mathcal{L} = \{p, q, r, \ldots\}$ be a countable set of propositional letters. The language of logic $LUPP$ consists of the elements of set $\mathcal{L}$, classical propositional connectives $\neg$ and $\wedge$ and the lists of upper probability operators $U_{\geq s}$ and $L_{\geq s}$, for every $s \in S$. The set of all classical propositional formulas over $\mathcal{L}$ is defined as usual, and we will denote it by For_C[7]. We will denote the propositional formulas by α, β and γ.

[7] C stands for classical

Definition 10. ***(Lower and upper probabilistic formulas)*** *If* $\alpha \in For_C$ *and* $s \in S$*, then a basic lower probability formula is any formula of the form* $L_{\geq s}\alpha$*, and a basic upper probability formula is any formula of the form* $U_{\geq s}\alpha$*. The set of all lower and upper probabilistic formulas, denoted by* For_P*, is the smallest set containing all basic lower and upper probability formulas which is closed under Boolean connectives.*

We denote the lower and upper probabilistic formulas by ϕ and ψ, possibly indexed. Let

$$For_{LUPP} = For_C \cup For_P.$$

The formulas from the set For_{LUPP} will be denoted by ρ and σ, possibly with subscripts.

We use the same abbreviations for $L_{<s}\alpha$, $L_{\leq s}\alpha$, $L_{=s}\alpha$, $L_{>s}\alpha$, $U_{<s}\alpha$, $U_{\leq s}\alpha$, $U_{=s}\alpha$ and $U_{>s}\alpha$ as for the formulas from For_{LUPFO}.

Example 12. *Continuing Example 1, it is clear that upper and lower probability, for the case that picked marble is green or black, are equal to* 0.6*. If there are no green marbles at all, then we obtain that lower probability for the case that picked marble is not green equals to* 1*. We can express that by the following formula of our language:*

$$U_{=0.6}(G \cup B) \wedge L_{=0.6}(G \cup B) \rightarrow L_{=1}\neg G.$$

Another example of a lower and upper probabilistic formula is

$$U_{<\frac{1}{3}}\alpha \rightarrow L_{\geq\frac{1}{2}}(\alpha \wedge \beta),$$

where α*,* $\beta \in For_C$*.*

Next we state two formulas that are not well defined lower and upper probabilistic formulas of the logic $LUPP$:

$$\alpha \wedge U_{=1}\beta, \qquad U_{\geq s}U_{\geq r}\alpha.$$

The first formula is not well defined since it is a Boolean combination of a pure propositional formula and an upper probabilistic formula, while the second formula is not well defined lower and upper probabilistic formula because it contains nested operators.

5.2. Semantics

The semantics for $LUPP$ is based on the possible-world approach.

Definition 11 ($LUPP$-structure). *An $LUPP$-structure is a tuple $\langle W, H, P, \upsilon \rangle$, where:*

- *W is a nonempty set of worlds.*
- *H is an algebra of subsets of W. The elements of H are called measurable worlds.*
- *P is a set of finitely additive probability measures defined on H*
- *$\upsilon : W \times \mathcal{L} \longrightarrow \{true, false\}$ provides for each world $w \in W$ a two-valued evaluation of the primitive propositions, which is extended to classical propositional formulas as usual.*

For given $\alpha \in For_C$ and $LUPP$-structure M, let $[\alpha]_M = \{w \in W \mid \upsilon(w)(\alpha) = true\}$. We will not write the subscript M when it's clear from context.

Definition 12 (Measurable structure). *The structure $M = \langle W, H, P, \upsilon \rangle$ is measurable if $[\alpha]_M \in H$ for every $\alpha \in For_C$. The class of a measurable structures of the logic $LUPP$ will be denoted by $LUPP_{Meas}$.*

Next we define when a formula holds in a measurable structure. Since our classical propositional formulas represent strict knowledge, we require them to hold in every world of the structure.

Definition 13 (Satisfiability relation). *The satisfiability relation $\models \subseteq LUPP_{Meas} \times For_{LUPP}$ is defined in the way such that for $M = \langle W, H, P, \upsilon \rangle$ we have:*

- *$M \models \alpha$ iff $\upsilon(w)(\alpha) = true$, for all $w \in W$,*
- *$M \models U_{\geq s}\alpha$ iff $P^{\star}([\alpha]) \geq s$,*
- *$M \models L_{\geq s}\alpha$ iff $P_{\star}([\alpha]) \geq s$,*
- *$M \models \neg\phi$ iff it is not the case that $M \models \phi$,*

- *$M \models \phi \wedge \psi$ iff $M \models \phi$ and $M \models \psi$,*

where $\alpha \in For_C$ and $\phi, \psi \in For_P$.

Now we introduce some basic notions related to the satisfiability relation.

Definition 14 (Satisfiability of a formula). *A formula $\rho \in For_{LUPP}$ is satisfiable if there is an $LUPP_{Meas}$-model M such that $M \models \rho$; ρ is valid if for every $LUPP_{Meas}$-model M, $M \models \rho$. A set of formulas T is satisfiable if there is an $LUPP_{Meas}$-model M such that $M \models \rho$ for every $\rho \in T$.*

Note that we can use an argument similar as the one we used in Example 10 to show that the logic $LUPP$ is not compact.

5.3. Decidability

We have already mentioned that Halpern and Pucella (2002) have shown decidability for the formulas which are Boolean combinations of the expressions of the form

$$r_1\ell(\alpha_1) + \ldots + r_n\ell(\alpha_n) \geq r_{n+1},$$

where ℓ is the upper probability operator and r_i are integers, for $i \in \{1, 2, \ldots, n+1\}$. First of all, note that using only integers as coefficients has the same expressive power as using all rational numbers. For example, if we want to express $\frac{3}{7}\ell(\alpha) \geq 1$ by using only integers, we can reformulate the formula as $3\ell(\alpha) \geq 7$, etc. Also note that our formula $U_{\geq s}\alpha$ is satisfiable iff the formula $\ell(\alpha) \geq s$ is satisfiable in the logic of Halpern and Pucella (2002). Similarly, $L_{\geq s}\alpha$ is satisfiable iff the formula $-\ell(\neg\alpha) \geq -(1-s)$ is satisfiable. Then decidability of our logic is a consequence of decidability of the logic provided by Halpern and Pucella (2002). Moreover, since the problem of deciding whether a formula of their language is satisfiable is NP-complete (Halpern and Pucella, 2002, Theorem 5.2), and also for the classical formulas problem is NP-complete, we have an upper bound of the decidability problem for $LUPP$. The lower bound follows from the well known fact that the complexity of the decision problem for classical propositional logic is NP-complete. Thus, the satisfiability problem for $LUPP$-formulas is NP-complete as well.

6. The axiomatization $\boldsymbol{Ax_{LUPP}}$

The axiomatic system for the logic $LUPP$ is very similar to the axiomatic system for the logic $LUPFO$. We will denote the system by Ax_{LUPP}.

Axiom schemes consists of the Axioms (4) – (9) from the logic $LUPFO$ and the following axiom:

(a) all instances of the classical propositional tautologies,

while Inference Rules consists of all the Inference Rules from the logic $LUPFO$ except Rule (2).

It is easy to see that classical propositional logic is sublogic of the $LUPP$. Inference relation, consistent and maximally consistent set are defined in the same way as for the logic $LUPFO$. Also, it is clear that the Deduction theorem holds for the logic $LUPP$ as well.

7. Soundness and Completeness

7.1. Soundness

Theorem 6 (Soundness). *The axiomatic system Ax_{LUPP} is sound with respect to the class of $LUPP_{Meas}$-models.*

Proof. The proof is analogous to the proof for the logic $LUPFO$. □

7.2. Completeness

In order to prove the strong completeness theorem for the logic $LUPP$, we first need to adapt Lemma 2 Theorem 4, since we now deal with a propositional language. Their proofs are the same as the proofs when we considered the logic $LUPFO$.
Using those results, we can define the canonical model.

Definition 15. *If $T^\star$ is the maximally consistent set of formulas, then a tuple $M_{T^\star} = \langle W, H, P, v\rangle$ is defined as:*

- *$W = \{w \mid w \models Cn_C(T)\}$ contains all classical propositional interpretations that satisfy the set $Cn_C(T)$,*
- *$H = \{[\alpha] \mid \alpha \in For_C\}$, where $[\alpha] = \{w \in W \mid w \models \alpha\}$,*

- *P is any set of probability measures such that $P^\star([\alpha]) = \sup\{s \mid U_{\geq s}\alpha \in T^\star\}$,*
- *for every world w and every propositional letter p, $\upsilon(w,p) = true$ iff $w \models p$.*

We can prove that $M_{T^\star}$ is well defined and that it belongs to $LUPP_{Meas}$ analogously as it is done in Lemma 3.

Theorem 7 (Strong completeness). *A set of formulas T is consistent iff it is $LUPP_{Meas}-satisfiable$.*

There is no difference between the proof of this theorem and the proof for the strong completeness theorem for a logic $LUPFO$ except that we consider now classical propositional formulas instead of classical first-order formulas.

8. The logics $LUPP^{FR(n)}$ and $LUPFO^{FR(n)}$

In this section we introduce the two families of finitary logics, $LUPP^{FR(n)}$ and $LUPFO^{FR(n)}$. These logics have the same language as the logics $LUPP$ and $LUPFO$ respectively, only their semantics are changed. Let us discuss first the logic $LUPP^{FR(n)}$ which is similar to $LUPP$. The main difference is that the finitely additive measures map H to $N = \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$, for a fixed positive integer n. Therefore, we obtain countably many different logics, one for each n. Considering the semantics, a model is the tuple $\langle W, H, P, \upsilon\rangle$ defined as above but the set P consists of finitely additive measures with restricted range N, i.e., for each $\mu \in P$, $\mu : H \longrightarrow N$. Hence, for every $X \in H$, $P^\star(X)$ also belongs to N, because N is finite and therefore $\sup\{\mu(x) \mid \mu \in P\} = \max\{\mu(x) \mid \mu \in P\}$. The class of all measurable $LUPP^{FR(n)}$ models is defined analoguosly as for the logic $LUPP$ and will be denoted by $LUPP^{FR(n)}_{Meas}$.

We want to show that there are finitary axiomatizations of these logics and to prove that they are sound and complete with respect to the considered classes of models.

For $s \in [0,1)$, let $s^+ = \min\{r \in N \mid s < r\}$, and if $s \in (0,1]$, let $s^- = \max\{r \in N \mid s > r\}$.
The axiomatization of the logic $LUPP^{FR(n)}$ includes all the axioms from Section 3, plus one more axiom:

(8) $U_{>s}\alpha \to U_{\geq s^+}\alpha$. The inference rules of the axiomatization are rules (1) and (2) from Section 3. Consequently, our axiomatization is finite, and the proofs are finite sequences of formulas.

Lemma 4. *(a)* $\vdash U_{>s}\alpha \leftrightarrow U_{\geq s^+}\alpha$,

(b) $\vdash U_{<s}\alpha \leftrightarrow U_{\leq s^-}\alpha$,

(c) $\vdash \bigvee_{s\in N} U_{=s}\alpha$,

(d) $\vdash \underline{\bigvee}_{s\in N} U_{=s}\alpha$.[8]

Proof. Proofs for (a) and (b) are trivial (direct consequences of Axiom 8 including contrapositive).
(c) Clearly $\vdash (U_{\geq 1}\alpha \vee \neg U_{\geq 1}\alpha) \wedge \neg U_{>1}\alpha$. Therefore

$$\vdash (U_{\geq 1}\alpha \wedge \neg U_{>1}\alpha) \vee (\neg U_{\geq 1}\alpha \wedge \neg U_{>1}\alpha).$$

Since $U_{\geq 1}\alpha \wedge \neg U_{>1}\alpha = U_{=1}\alpha$ and $\vdash U_{<1}\alpha \to U_{\leq 1}\alpha$ we have $\vdash U_{=1}\alpha \vee U_{<1}\alpha$. Furthermore, using a tautology $A \leftrightarrow ((B \vee \neg B) \wedge A)$, we obtain

$$\vdash U_{<1}\alpha \leftrightarrow ((U_{\geq 1^-}\alpha \vee \neg U_{\geq 1^-}\alpha) \wedge U_{<1}\alpha).$$

From that, using the classical propositional tautology $(A \vee B) \wedge C \leftrightarrow (A \wedge C) \vee (B \wedge C)$ and the following equivalence $U_{<1}\alpha \leftrightarrow \neg U_{>1^-}\alpha$, we obtain that

$$\vdash U_{<1}\alpha \leftrightarrow ((U_{\geq 1^-}\alpha \wedge \neg U_{>1^-}\alpha) \vee (U_{<1^-}\alpha \wedge U_{<1}\alpha)).$$

Since $\vdash U_{=1}\alpha \vee U_{<1}\alpha$, using the equivalence above, and the facts that $U_{=1^-}\alpha \leftrightarrow (U_{\geq 1^-}\alpha \wedge \neg U_{>1^-}\alpha)$ and $U_{<1^-}\alpha \leftrightarrow (U_{<1^-}\alpha \wedge U_{<1}\alpha)$ we obtain that

$$\vdash U_{=1}\alpha \vee U_{=1^-}\alpha \vee U_{<1^-}\alpha.$$

Finally, we have that $\vdash (\bigvee_{s\in N} U_{=s}\alpha) \vee U_{<0}\alpha$, so $\vdash (\bigvee_{s\in N} U_{=s}\alpha)$.
(d) $U_{=r}\alpha = U_{\geq r}\alpha \wedge \neg U_{>r}\alpha$, so $\vdash U_{=r}\alpha \to \neg U_{=s}\alpha$, for every $s > r$. Similarly, we can prove that $\vdash U_{=r}\alpha \to \neg U_{=s}\alpha$, for every $s < r$. So, using result proved in (c), as a consequence we obtain $\vdash \underline{\bigvee}_{s\in N} U_{=s}\alpha$. □

The proof of the strong completeness theorem is similar to one presented in Section 4. We will only explain the idea of the proof without going into the details. First, we can prove the soundness theorem:

[8] $\underline{\bigvee}$ stands for the exclusive disjunction, i.e., for the logic operation which is true if and only if exactly one value is true

Theorem 8 (Soundness). *The axiomatic system* $Ax_{LUPP^{FR(n)}}$ *is sound with respect to the class of* $LUPP^{FR(n)}_{Meas}$*-models.*

Next, we prove the deduction theorem:

Theorem 9 (Deduction theorem). *Let* T *be a set of formulas. Then* $T \cup \{\rho\} \vdash \sigma$ *iff* $T \vdash \rho \to \sigma$.

The proof is straightforward. After that, we prove:

Theorem 10. *Every consistent set of formulas can be extended to a maximal consistent set.*

The proof of Theorem 10 is similar to the proof of the Theorem 4. The only difference is that we skip the steps where we use infinitary inference rules, i.e. steps 2(b) and 2(c). One more fact needs some explanation. In the proof of the strong completeness theorem for the logic $LUPFO$ we use that if $\sup\{r \mid U_{\geq r}\alpha \in T^\star\} = s$, and $s \in S$, then $U_{\geq s}\alpha \in T^\star$. Now, we have that s must belong to the set N. Indeed, if $s \notin N$ then there is some $r < s$ such that $r^+ = s^+$ (i.e. $r > s^-$, density of rational numbers), so, $U_{\geq r^+}\alpha \in T^\star$, i.e. $U_{\geq s^+}\alpha \in T^\star$, therefore $T^\star \vdash U_{\geq s^+}\alpha$, but $s < s^+$. Contradiction. Furthermore, $U_{\geq s}\alpha \in T^\star$ because of Lemma 4(d). The rest of the proof of the strong completeness is identical as in Section 7. Thus, the following statement holds for the logic $LUPP$.

Theorem 11 (Strong completeness). *A set of formulas* T *is consistent iff it is* $LUPP^{FR(n)}_{Meas} - satisfiable$.

The logic $LUPFO^{FR(n)}$ can be obtained from $LUPFO$ in the same way as the logic $LUPP^{FR(n)}$ is obtained from $LUPP$. Because of the obvious similarities in the construction and the proofs, we omit details.
Also, note that, because of the simplified semantics, the logics $LUPP^{FR(n)}$ and $LUPFO^{FR(n)}$ are compact.

9. Related work

Arguably, the most important advancement in probability logic, after the work of Leibnitz and Boole, was made by Keisler (1977, 1985), who gave model-theoretic approach to probability theory. Keisler introduced several probability quantifiers appropriate for statistical reasoning, as for example

$Px > r$. The considered models are first-order structures with probabilities on their domains. The formula $(Px > r)\phi(x)$ means that the set $\{a \mid \phi(a)\}$ has probability greater than r. A recursive axiomatization for that kind of logics was given by Hoover (1978). He used admissible and countable fragments of infinitary predicate logic. Later, Bacchus (1990) considered a similar logic, with the constraint that the range of probability functions is required to be the unit interval of a totally ordered field, instead of a particular field, (e.g., $[0, 1]$, or $[0, 1]_Q$). That modification allowed him to obtain a finitary, strongly complete axiomatization, using the standard Henkin style procedure. As it is pointed out by Ognjanovic et al. (2016), there are sentences on the real-valued probability functions that are not provable in that system.

Since the middle of 1980s, researchers attempt to combine probability-based and logic-based approaches to knowledge representation, developing the logical frameworks for modelling uncertainty in which probabilities express degrees of belief. In the first of those papers which resulted from the work on developing an expert system in medicine, Nilsson (1986) presented a logic with probabilistic operators as a well-founded framework for uncertain reasoning. He gave a procedure for probabilistic entailment that calculates bounds on the probabilities of a conclusion, given probabilities of premises. The first probability logic developed along the lines of the Nilsson's research is due to Fagin et al. (1990). They have developed a propositional logic for reasoning about simple and conditional probabilities. Reasoning about probabilities is formally expressed by linear weight formulas, which are Boolean combinations of the basic formulas. An example of the basic linear weight formula is $\frac{1}{2}w(\alpha) + \frac{2}{5}w(\alpha \wedge \beta) \geqslant \frac{5}{6}$, where α and β are formulas in some underlying classical propositional logic and $w(\alpha)$ reads "the probability of α". They provided a finitary axiomatization for the logic, and they proved weak completeness. The standard semantics for this kind of probability logics (Fagin et al. (1990); Fagin and Halpern (1994); Halpern and Pucella (2006)) is a variant of Kripke models, where accessibility relation between the worlds is replaced by a finitely additive probability measure, and it is also used in the fuzzy approach to probabilistic reasoning (Hájek et al. (1995); Flaminio and Godo (2007)).

This Kripke-style semantics is also proposed for probability logics with unary operators by Ognjanovic and Raskovic (1999, 2000). They developed a new technique for proving strong completeness for non-compact probability logics which combines Henkin-style procedures for classical and modal logics and which works with infinitary proofs. That technique allowed them to prove strong completeness for real-valued probabilities. For a detailed overview of the approach we refer the reader to Ognjanovic et al. (2016). In this work, we followed the approach and we modified the technique they developed.

Jaeger (2005) developed a first-order logic for inductive probabilistic reasoning $\mathcal{L}_{ip}$. He obtained weakly complete axiomatization, but he paid the price by allowing probability distributions in the semantic models that use non-standard probability values. Halpern (1990) considered two first-order probability logics. In the first logic probabilities are defined on the domain, while in the second logic probabilities are defined on possible worlds, but with the restriction that the measures $\mu(w)$ in all worlds of a model are equal. Thus, formulas expressing probabilities either hold in every world from a model or they are not satisfiable in that model. Halpern provided axiomatizations for both logics, but completeness can be proved only if the domains are bounded in size by some finite n. Bacchus (1990) argued that it is difficult to justify that assumption even for artificial intelligence applications. In his opinion, while domain may be finite, it is questionable that there is a fixed upper bound on its size. Bacchus also pointed out that there are many domains, interesting in AI applications, that are not finite. Later, Ognjanovic and Raskovic (2000) introduced a strongly complete axiom system for the first-order probability logic with probabilities defined on possible worlds with unbounded domains.

Poole (2003) and Grove et al. (1992) also worked on probabilistic first-order probabilistic logics, but we point out that axiomatization issues are not consiedered in those papers. Grove et al. (1992) considered how to compute the asymptotic conditional probabilities for the first-order formulas, i.e. for two first-order formulas φ and ψ, they considered the number of structures with a finite domain that satisfy ψ, and computed the fraction of them in which φ is true. Poole (2003) worked on first-order probabilistic inference, investigating first-order belief networks and providing an algorithm for reasoning about multiple individuals, where some facts about some of them are known, but the rest is treated as a group.

In all the logics reviewed above, each semantical model is equipped with a unique probability measure (defined on the domain of a first-order structure for the logics with the statistical approach, or on sets of worlds for the subjective probability approaches). Consequently, a sentence of the form

(*) "probability of α is at least s"

has a fixed truth value (true or false) in each model, depending on the measure of the semantical set that corresponds to α.

In this paper, we consider the situations in which uncertainty is represented by a set of probability measures. In the cases where are different measures in a single semantical model M, we cannot always assign a truth value to the sentence $(*)$. Indeed, consider sentence $(\exists x)\text{Head\_Crash}(x)$ from Example 9 and denote the sentence by β. The model M from Example 9 contains the measures μ_1 and μ_2 such that $\mu_1([\beta]) = 0.8$ and $\mu_2([\beta]) = 0.6$. In that case, we cannot assign the value true or false to the sentence

(**) "probability of β is at least 0.7"

in the model M. Obviously, the syntax with classical probability operators from the papers discussed above is not appropriate for the models with multiple measures. On the other hand, we can state some simple probability boundaries in our semantical frameworks, which speak that the probability of a formula being true is at least s for all/some probability measure from a model, i.e.,

"upper/lower probability of α is at least s."

If we consider Example 9 again, instead of statement $(**)$, we can consider the sentences

"upper (lower) probability of β is at least 0.7,"

which has the value true (false, respectively) in the model M (see Example 9).

The only logic for reasoning about upper probability measures is introduced by Halpern and Pucella (2002). Following the technique by Fagin et al. (1990), they developed a propositional logic which allows reasoning about linear inequalities involving upper probabilities, and they proved weak completeness. On the other hand, we followed the technique of Ognjanovic and Raskovic (1999, 2000), and we proved strong completeness using infinitary rules of inference. We introduced the syntax with unary operators for lower and upper probabilities, which is a simpler syntax than the syntax from Halpern and Pucella (2002), since we don't have the arithmetical operations built into formulas. Finally, while Halpern and Pucella (2002) introduced the propositional logic only, we also developed a first-order logic for reasoning about lower and upper probabilities.

10. Conclusion

In this paper, we introduced the logics $LUPP$ and $LUPFO$, whose languages are obtained by adding the operators for upper and lower probabilities to propositional and first-order logic, respectively. We proposed the axiomatizations for the logics and proved strong completeness. Since the logics are not compact, the axiomatizations contain infinitary rules of inference. Then we simplified the semantics and we achieved compactness using finite sets of probability values for logics $LUPP^{FR(n)}$ and $LUPFO^{FR(n)}$. For those logics we provide finitary axiomatizations.

The logics presented in this work extend several existing results. First, $LUPP$ is similar in spirit to the logic by Halpern and Pucella (2002) (the former provides strongly complete axiomatization, while the latter provides a finitary axiomatization which is weakly complete and incorporate arithmetical operations into syntax). $LUPFO$ can be seen as the first-order extension of those logics. The logics LPP_2 and $LFOP_2$ by Ognjanovic and Raskovic (2000) can be seen as special cases of the logics $LUPP$ and $LUPFO$, respectively, if the set of probability measures is a singleton (thus we are reduced to standard probabilistic logics). Syntactically, that restriction corresponds to an additional axiom: $U_{\geq r}\alpha \rightarrow L_{\geq r}\alpha$. Similarly, the logic by van der Hoek (1997) can be seen as special case of $LUPP^{FR(n)}$.

Finally, we recall that our logics do not allow mixing of probabilistic and non-probabilistic knowledge in a single formula or higher-order lower/upper probabilities. In the future work, we plan to extend the approach from this paper in the way that we allow both types of formulas. We believe that a more modal approach, in which every world is equipped with a set of probability measures, is the obvious choice for the semantics of such logic.

Acknowledgements

This work was supported by the National Research Fund (FNR) of Luxembourg through project PRIMAT, Swiss National Science Foundation grant 200021_165549 and by the Serbian Ministry of Education and Science.

We wish to thank the reviewers, whose careful reading, comments and corrections greatly improved the paper.

References

Anger, B., Lembcke, J., 1985. Infinitely subadditive capacities as upper envelopes of measures. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 68 (3), 403–414.

Bacchus, F., 1990. Lp, a logic for representing and reasoning with statistical knowledge. Computational Intelligence 6 (4), 209–231.

Cintula, P., Noguera, C., 2014. Modal logics of uncertainty with two-layer syntax: A general completeness theorem. In: Kohlenbach et al. (2014), pp. 124–136.

de Cooman, G., Hermans, F., 2008. Imprecise probability trees: Bridging two theories of imprecise probability. Artif. Intell. 172 (11), 1400–1427.

Dubois, D., Prade, H., 1988. Possibility theory. Plenum Press, New-York.

Fagin, R., Halpern, J. Y., 1994. Reasoning about knowledge and probability. J. ACM 41 (2), 340–367.

Fagin, R., Halpern, J. Y., Megiddo, N., 1990. A logic for reasoning about probabilities. Inf. Comput. 87 (1/2), 78–128.

Fattorosi-Barnaba, M., Amati, G., 1995. Modal operators with probabilistic interpretations, i. Studia Logica 46 (4), 383–393.

Flaminio, T., Godo, L., 2007. A logic for reasoning about the probability of fuzzy events. Fuzzy Sets and Systems 158 (6), 625–638.

Frisch, A. M., Haddawy, P., 1994. Anytime deduction for probabilistic logic. Artif. Intell. 69 (1-2), 93–122.

Grove, A. J., Halpern, J. Y., Koller, D., 1992. Asymptotic conditional probabilities for first-order logic. In: Kosaraju, S. R., Fellows, M., Wigderson, A., Ellis, J. A. (Eds.), Proceedings of the 24th Annual ACM Symposium on Theory of Computing, May 4-6, 1992, Victoria, British Columbia, Canada. ACM, pp. 294–305.

Hájek, P., Godo, L., Esteva, F., 1995. Fuzzy logic and probability. In: Besnard, P., Hanks, S. (Eds.), UAI '95: Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence, Montreal, Quebec, Canada, August 18-20, 1995. Morgan Kaufmann, pp. 237–244.

Halpern, J. Y., 1990. An analysis of first-order logics of probability. Artif. Intell. 46 (3), 311–350.

Halpern, J. Y., Pucella, R., 2002. A logic for reasoning about upper probabilities. J. Artif. Intell. Res. (JAIR) 17, 57–81.

Halpern, J. Y., Pucella, R., 2006. A logic for reasoning about evidence. J. Artif. Intell. Res. (JAIR) 26, 1–34.

Heifetz, A., Mongin, P., 2001. Probability logic for type spaces. Games and Economic Behavior 35 (1-2), 31–53.

Hoover, D. N., 1978. Probability logic. Annals of Mathematical Logic 14 (3), 287 – 313.

Huber, P., 1981. Robust statistics. Wiley New York.

Ikodinovic, N., Ognjanovic, Z., Perovic, A., Raskovic, M., 2014. Hierarchies of probabilistic logics. Int. J. Approx. Reasoning 55 (9), 1830–1842.

Ilic-Stepic, A., Ognjanovic, Z., Ikodinovic, N., 2014. Conditional p-adic probability logic. Int. J. Approx. Reasoning 55 (9), 1843–1865.

Jaeger, M., 2005. A logic for inductive probabilistic reasoning. Synthese 144 (2), 181–248.

Keisler, H. J., 1977. Hyperpinite model theory. In: Gandy, R., Hyland, J. (Eds.), Logic Colloquium 76Proceedings of a conference. Vol. 87 of Studies in Logic and the Foundations of Mathematics. Elsevier, pp. 5 – 110.

Keisler, H. J., 1985. Chapter XIV: Probability Quantifiers. Vol. Volume 8 of Perspectives in Mathematical Logic. Springer-Verlag, New York, pp. 507–556.

Kohlenbach, U., Barceló, P., de Queiroz, R. J. G. B. (Eds.), 2014. Logic, Language, Information, and Computation - 21st International Workshop, WoLLIC 2014, Valparaíso, Chile, September 1-4, 2014. Proceedings. Vol. 8652 of Lecture Notes in Computer Science. Springer.

Kyburg, H., 1961. Probability and the Logic of Rational Belief. Wesleyan University Press.

Levi, I., 1980. The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance. The MIT Press.

Lorentz, G. G., 1952. Multiply subadditive functions. Synthese 4 (4), 455–462.

Meier, M., 2012. An infinitary probability logic for type spaces. Israel Journal of Mathematics 192 (1), 1–58.

Miranda, E., 2008. A survey of the theory of coherent lower previsions. International Journal of Approximate Reasoning 48 (2), 628 – 658, in Memory of Philippe Smets (19382005).

Nilsson, N. J., 1986. Probabilistic logic. Artif. Intell. 28 (1), 71–87.

Ognjanovic, Z., Raskovic, M., 1999. Some probability logics with new types of probability operators. J. Log. Comput. 9 (2), 181–195.

Ognjanovic, Z., Raskovic, M., 2000. Some first-order probability logics. Theor. Comput. Sci. 247 (1-2), 191–212.

Ognjanovic, Z., Raskovic, M., Markovic, Z., 2016. Probability Logics - Probability-Based Formalization of Uncertain Reasoning. Springer.

Poole, D., 2003. First-order probabilistic inference. In: Gottlob, G., Walsh, T. (Eds.), IJCAI-03, Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, Mexico, August 9-15, 2003. Morgan Kaufmann, pp. 985–991.

Raskovic, M., Markovic, Z., Ognjanovic, Z., 2008. A logic with approximate conditional probabilities that can model default reasoning. Int. J. Approx. Reasoning 49 (1), 52–66.

Savic, N., Doder, D., Ognjanovic, Z., 2015. A logic with upper and lower probability operators. In: Kohlenbach et al. (2014), pp. 267–276.

Shafer, G., et al., 1976. A mathematical theory of evidence. Vol. 1. Princeton university press Princeton.

van der Hoek, W., 1997. Some considerations on the logic pfd˜. Journal of Applied Non-Classical Logics 7 (3).

Walley, P., 1991. Statistical reasoning with imprecise probabilities.

Walley, P., 2000. Towards a unified theory of imprecise probability. Int. J. Approx. Reasoning 24 (2-3), 125–148.

Zadeh, L., 1978. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1 (1), 3 – 28.