Frequentist Probability Logic

Alessandro Aldini[0000−0002−7250−5011], Pierluigi Graziani[0000−0002−8828−8920], and Mirko Tagliaferri[0000−0003−3875−0512]

University of Urbino Carlo Bo, Italy

Abstract. In this paper, we present a logic **FPL** to reason about probabilities with a relative frequency interpretation. We show that it is possible to interpret the language of **FPL** with the standard semantics for propositional logic. **FPL** can give a peculiar frequentist interpretation of a probability operator. We then give a proof system for the language, prove that the traditional theorems of probability hold, and prove soundness and completeness.

Keywords: Probability Logic · Relative Frequency · Reasoning about Uncertainty.

1 Introduction

Probability is important to every human endeavour which involves quantitative data analysis. Moreover, it is safe to claim that probability theory can be considered the best formal tool researchers have to represent uncertainty. From a mathematical standpoint, probability theory is a rigorous and thoroughly studied theory with precise axioms and proven theorems. Thus, it makes sense to try and combine logic and probability, the former providing a qualitative perspective over human reasoning and the latter providing the quantitative analysis [4].

One potential issue with all the formalizations of probability logic is that interpretations of probability are often left tacit; they provide the formal framework and the user applies its preferred interpretation. Although it is true that, as far as an interpretation is consistent with the axioms of probability theory, applying such interpretation to whatever probability logic does not change the characteristics of the logic, it is also true that some interpretations allow for a more fine-grained analysis of why a specific event has a given probability in the first place. For instance, saying that the probability of a fair coin landing heads is $\frac{1}{2}$ is not affected by the interpretation you give to the fraction $\frac{1}{2}$. However, showing that such probability follows from the fact that a coin landed heads twice out of four tosses adds some depth to our understanding of *why* the probability is $\frac{1}{2}$. This last example is an instance of a frequentist interpretation of probability [9] and is extensively employed in scientific settings to establish the probability of different events.

Moreover, a logic that can explicitly capture the idea of relative frequency and that can allow reasoning about probability under said interpretation can become extremely useful in context of statistical and approximate model checking [10, 8] and of probabilistic verification of machine learning [7, 11, 12].

Statistical model checking refers to techniques which are employed in computer science to check properties of stochastic systems. Differently from Boolean model checkers, which establish whether a property is reached or not in a model of a system, statistical model checkers start from the assumption that exhaustively exploring all paths of a model might not always be feasible and, thus, quantitative questions should be asked rather than qualitative one. For example, a statistical model checker would look for answers to questions such as: is the probability that the system satisfies a certain property greater than a certain threshold?

In a similar way, the logic here proposed could be employed to construct sophisticated reinforcement learning strategies for machine learning. Reinforcement learning is based on the idea that whenever a classifier gets an answer wrong, it get punished, while it gets a prize whenever the classification is correct. However, we might want to cluster sets of answers together and punish only classifications that fall below certain thresholds that we have chosen as appropriate. This might thus allow those systems to create appropriate strategies that, although not perfect, might still be useful in contexts in which fuzzy evaluations are sufficient to make the best decision.

The logic presented in this paper allows computer scientists to represent and reason in those contexts with ease and thus represent valuable tools in the hands of the right modeller.

The aim of the paper is to enrich the semantics of standard probability logic by capturing explicitly a frequentists interpretation of probability. In particular, we will start from Fagin's et al. [5] Probability Logic and apply to it a frequentist semantics. Moreover, we will enrich the language by allowing combinations of propositional and probabilistic formulas, thus allowing an agent to reason about the relationship between local and global results (Section 3). We will then show that all meta-theoretic results from [5] hold for our semantic structures (Section 4) and we will discuss some future works that make our semantics an interesting evolution which might be of interest to computer scientists (Section 5).

2 Related Works

To the best of our knowledge, the works that mostly resemble our approach are those on counting propositions [3] and those on graded modal logics [6]. The former family of works present formal semantics to reason about how many times a specific proposition is true in a model. The general idea presented in those works is really similar to our, with the only distinction being the lack of connections to probability and its theorems. In this paper, we extend the analysis carried out by those authors by separating the truth of a proposition from its probability (thus capturing an intrinsic notion of uncertainty) and by constructing all the bridges between such formalism and probability logic. On the other hand, graded modal logic offers a way to quantify the uncertainty about the application of specific modal operators to formulas. When such modal operators are interpreted to stand from probability, the connection with our proposal becomes obvious. However, in graded modal logics a relational structure must be assumed, which is not strictly necessary in our semantics. In particular, we claim that graded modal logics might offer a subjective interpretation (due to the accessibility relations) of an objective interpretation of probability (due to the counting of true instances of a proposition in the structure). We offer a completely objective interpretation, creating a stronger tie with probability theory and the reasoning thereof.

3 Frequentist Probability Logic

3.1 Syntax

Take a countable set of primitive atomic formulas At ranging over p_i, with $i \in \mathbb{N}$. We will also assume that a, b range over the set $\mathbb{Z}$.

Definition 1 (Language of FPL). *The language $\mathcal{L}_{\mathbf{FPL}}$ ranging over χ_i, with $i \in \mathbb{N}$ is generated recursively by the following three-level grammar, with $\odot \in \{\geq, \leq\}$ and $\circ \in \{+, -\}$:*

$$\chi \coloneqq \varphi \mid \psi \mid \neg\chi \mid \chi \wedge \chi$$
$$\psi \coloneqq P(\varphi) \odot b \mid P(\varphi) \circ P(\varphi) \odot b$$
$$\varphi \coloneqq \top \mid p_i \mid \neg\varphi \mid \varphi \wedge \varphi$$

The elements of $\mathcal{L}_{\mathbf{FPL}}$ are called formulas. Specifically, the φ_is will be called *propositional formulas*, while the ψ_is will be called *probabilistic formulas*. Finally, the χ_is will be simply called *formulas*.

The following definitions will be employed throughout the paper:

Definition 2 (Defined formulas).

- $\bot \coloneqq \neg\top$;
- $\varphi_i \vee \varphi_j \coloneqq \neg(\neg\varphi_i \wedge \neg\varphi_j)$;
- $\varphi_i \rightarrow \varphi_j \coloneqq \neg\varphi_i \vee \varphi_j$;
- $\varphi_i \leftrightarrow \varphi_j \coloneqq (\varphi_i \rightarrow \varphi_j) \wedge (\varphi_j \rightarrow \varphi_i)$;[1]
- $P(\varphi_i) \odot P(\varphi_j) \coloneqq (P(\varphi_i) - P(\varphi_j)) \odot 0$;
- $P(\varphi) = b \coloneqq P(\varphi) \geq b \wedge P(\varphi) \leq b$;
- $P(\varphi) > b \coloneqq \neg(P(\varphi) \leq b)$;
- $P(\varphi) < b \coloneqq \neg(P(\varphi) \geq b)$.

Example 1. Imagine that Alice wants to draw some conclusions about a Monte Carlo simulation she ran over throwing two dice. For the purpose of this example, she will only focus on the outcomes of the dice that correspond to three (it is not hard to expand the example, but to keep the example simple, we will limit

[1] The same abbreviations will also be employed for the χ_i formulas.

ourselves to the case mentioned). Thus, the important propositions that Alice is interested in are p_{13} which stands for "the outcome of the throw of die 1 is 3", and p_{23} which stands for "the outcome of the throw of die 2 is 3". Inside **FPL**, Alice could then reason about formulas such as "both dice landed on the value 3" ($p_{13} \wedge p_{23}$); "whenever die one landed on three, die two did not" ($p_{13} \rightarrow \neg p_{23}$); "the relative frequency of die one landing three is $\frac{1}{6}$" ($P(p_{13}) = \frac{1}{6}$); "the relative frequency of die one landing three is lower than $\frac{1}{10}$" ($P(p_{13}) < \frac{1}{10}$).

3.2 Semantics

To define the semantics of **FPL**, we first introduce Frequentist Models.

Definition 3 (Frequentist Models). *A frequentist model is a couple $\mathfrak{M} = (O, v)$, where O is a non-empty finite set of possible outcomes ranging over o_i with $i \in \mathbb{N}$, and v is a valuation function $v : O \times At \rightarrow \{0, 1\}$.*

Intuitively, an outcome $o_i \in O$ can be seen as a valuation of all the atomic formulas of the language. Those outcomes could also be interpreted as experiments (or Montecarlo simulations) in which the atomic formulas are tested in order to verify whether they hold or not. The valuation function, on the other hand, indicates the results of the evaluations/experiment, indicating which formulas turned out to be true.

Example 2. We continue from Example 1. In order to evaluate the formulas Alice is interested in, she can perform a Monte Carlo simulation, throwing the two dice repeatedly and keeping track of the results. Imagine that Alice threw the two dice 30 times. The 30 throws performed by Alice would constitute the set O of possible outcomes. Moreover, the valuation function v would be equivalent to the notes that Alice made about the specific outcomes of the throws. Imagine that out of the 30 throws, die one landed on three five times (e.g., on throws one, three, eight, fourteen, twenty-seven), while die two landed on three on ten different occasions (e.g., on throws two, four, six, eleven, fourteen, sixteen, twenty-two, twenty-three, twenty-eight, and thirty). Formally, this would be captured by a model $\mathfrak{M}$ where $O = \{o_i \mid 1 \leq i \leq 30\}$, and v would be defined as follows (we will only specify the cases in which the formulas are true):

- $v(o_1, p_{13}) = v(o_3, p_{13}) = v(o_8, p_{13}) = v(o_{14}, p_{13}) = v(o_{27}, p_{13}) = 1$;
- $v(o_2, p_{23}) = v(o_4, p_{23}) = v(o_6, p_{23}) = v(o_{11}, p_{23}) = v(o_{14}, p_{23}) = v(o_{16}, p_{23}) = v(o_{22}, p_{23}) = v(o_{23}, p_{23}) = v(o_{28}, p_{23}) = v(o_{30}, p_{23}) = 1$.

It is fairly easy to extend the valuation function to all propositional formulas φ through induction over the structure of the formulas:

Definition 4 (Extended Valuation). *Given a frequentist model $\mathfrak{M}$, the valuation function v is extended to all the propositional formulas recursively as follows (we use v^e to indicate the extension of v):*

- $v^e(o_i, p_j) = v(o_i, p_j)$;

– $v^e(o_i, \neg\varphi) = 1$ *iff* $v^e(o_i, \varphi) = 0$;
– $v^e(o_i, \varphi_j \wedge \varphi_m) = 1$ *iff* $v^e(o_i, \varphi_j) = 1$ *and* $v^e(o_i, \varphi_m) = 1$.

Valuation functions do not apply to probability formulas. Since we are not allowing the nesting of probability operators, this will not constitute a problem.

Definition 5 (Validating Sets). *Given a frequentist model $\mathfrak{M}$, the validating set of a propositional formula φ, indicated with $[\![\varphi]\!]$, is a subset of O ($[\![\varphi]\!] \subseteq O$) s.t. $o_i \in [\![\varphi]\!]$ iff $v^e(o_i, \varphi) = 1$.*

Example 3. Given our example, the validating sets for propositions p_{13} and p_{23} would be, respectively:

– $[\![p_{13}]\!] = (o_1, o_3, o_8, o_{14}, o_{27})$;
– $[\![p_{23}]\!] = (o_2, o_4, o_6, o_{11}, o_{14}, o_{16}, o_{22}, o_{23}, o_{28}, o_{30})$.

The validating set of a formula, sometimes referred to as the truth set of the formula, simply indicates the set of outcomes where the formula is true.

Proposition 1 (Properties of Validating Sets). *Given a validating set $[\![\varphi]\!]$, the following properties follow:*

– $[\![\neg\varphi]\!] = O \setminus [\![\varphi]\!]$;
– $[\![\varphi_j \wedge \varphi_m]\!] = [\![\varphi_j]\!] \cap [\![\varphi_m]\!]$.

Proof. The proof follows directly from definitions 4 and 5.

Definition 6 (Validating Space). *The **validating space** $\mathcal{V}_{\mathfrak{M}}$ of a model $\mathfrak{M}$, is the set of all validating sets of propositional formulas of the language $\mathcal{L}_{\mathbf{FPL}}$:*

$$\mathcal{V}_{\mathfrak{M}} = \{[\![\varphi]\!] \mid \varphi \in \mathcal{L}_{\mathbf{FPL}}\} \tag{1}$$

Proposition 2. *Given a frequentist model $\mathfrak{M}$, the validating space $\mathcal{V}_{\mathfrak{M}}$ forms a σ-algebra over $O \in \mathfrak{M}$.*

Proof. Take an arbitrary model $\mathfrak{M} = (O, v)$. First note that by definitions 6 and 5, it follows that $\mathcal{V}_{\mathfrak{M}}$ is a set of subsets of O. Since $\top \in \mathcal{L}_{\mathbf{FPL}}$ and $[\![\top]\!] = O$, it follows that $O \in \mathcal{V}_{\mathfrak{M}}$.

Assume that $[\![\varphi]\!] \in \mathcal{V}_{\mathfrak{M}}$. Since $\mathcal{L}_{\mathbf{FPL}}$ is closed under negation, i.e., if $\varphi \in \mathcal{L}_{\mathbf{FPL}}$, then $\neg\varphi \in \mathcal{L}_{\mathbf{FPL}}$, it follows that $[\![\neg\varphi]\!] \in \mathcal{V}_{\mathfrak{M}}$. By proposition 1, $[\![\neg\varphi]\!] = O \setminus [\![\varphi]\!]$, which is the complement of $[\![\varphi]\!]$. Therefore, $\mathcal{V}_{\mathfrak{M}}$ is closed under complement.

Assume that $[\![\varphi_i]\!] \in \mathcal{V}_{\mathfrak{M}}$ and that $[\![\varphi_j]\!] \in \mathcal{V}_{\mathfrak{M}}$. Since $\mathcal{L}_{\mathbf{FPL}}$ is closed under disjunction (through definition 2), i.e., if $\varphi_i \in \mathcal{L}_{\mathbf{FPL}}$ and $\varphi_j \in \mathcal{L}_{\mathbf{FPL}}$, then $\varphi_i \vee \varphi_j \in \mathcal{L}_{\mathbf{FPL}}$, it follows that $[\![\varphi_i \vee \varphi_j]\!] \in \mathcal{V}_{\mathfrak{M}}$. Through set operations and proposition 1, $[\![\varphi_i \vee \varphi_j]\!] = [\![\varphi_i]\!] \cup [\![\varphi_j]\!]$. Therefore, $\mathcal{V}_{\mathfrak{M}}$ is closed under finite union.

From the previous four properties, it follows that $\mathcal{V}_{\mathfrak{M}}$ forms an algebra over $O \in \mathfrak{M}$.

Moreover, since O is finite, $\mathcal{V}_{\mathfrak{M}}$ is also a σ-algebra over O, since countable union collapses onto finite union.

In the following definitions, we will indicate with $|O|$, and $|[\![\varphi]\!]|$ the cardinality of, respectively, the outcomes set and the validating set.

Definition 7 (Relative Frequency Function). *Given a model $\mathfrak{M}$, it is possible to define a **relative frequency function** τ that assigns to every propositional formula of the language its relative frequency. The assignment procedure for τ is defined as follows: $\tau(\varphi) = \frac{|[\![\varphi]\!]|}{|O|}$.*

Definition 8 (Probability Measure). *Given a set S of states and a σ-algebra $\mathcal{X}$ of measurable sets X_i, a **probability measure** is a function μ which takes as arguments the measurable sets and returns as values a number from $[0,1] \in \mathbb{Q}$. Moreover, a probability measure satisfies the following properties:*

- $\mu(X_i) \geq 0$, *for all* $X_i \in \mathcal{X}$;
- $\mu(S) = 1$;
- *If* $X_i \cap X_j = \emptyset$, *then* $\mu(X_i \cup X_j) = \mu(X_i) + \mu(X_j)$.

*The triple $(S, \mathcal{X}, \mu)$ is called a **probability space**.*

Remark 1. Given a frequentist model $\mathfrak{M}$, the triple $(O, \mathcal{V}_{\mathfrak{M}}, \tau)$ is a probability space.

In the semantics of **FPL**, formulas are interpreted over frequentist models. Propositional formulas will be interpreted locally, while probability formulas will be interpreted globally over the whole model. In particular, the semantics of the operator $P(\cdot)$ is given in terms of the relative frequency function $\tau(\cdot)$.

Definition 9 (Truth of propositional formulas). *Let $\varphi \in \mathcal{L}_{\mathbf{FPL}}$ be a propositional formula and $\mathfrak{M} = (O, v)$ be a frequentist model. We inductively define the notion of φ being verified (or satisfied) by an outcome $o_i \in O$ in $\mathfrak{M}$, written $o_i \models_{\mathfrak{M}} \varphi$, as follows:*

1. $o_i \models_{\mathfrak{M}} \top$, *always*;
2. $o_i \models_{\mathfrak{M}} p_j$ *iff* $v^e(o_i, p_j) = 1$;
3. $o_i \models_{\mathfrak{M}} \neg\varphi$ *iff* $o_i \not\models_{\mathfrak{M}} \varphi$;
4. $o_i \models_{\mathfrak{M}} \varphi_j \wedge \varphi_m$ *iff* $o_i \models_{\mathfrak{M}} \varphi_j$ *and* $o_i \models_{\mathfrak{M}} \varphi_m$.

Definition 10 (Truth of probability formulas). *Let $\psi \in \mathcal{L}_{\mathbf{FPL}}$ be a probability formula, $\chi \in \mathcal{L}_{\mathbf{FPL}}$ a general formula, and $\mathfrak{M} = (O, v)$ be a frequentist model. We inductively define the notion of ψ (or χ) being verified (or satisfied) by an outcome $o_i \in O$ in $\mathfrak{M}$, written $o_i \models_{\mathfrak{M}} \psi$ (the same notation applies to χ), as follows:*

5. $o_i \models_{\mathfrak{M}} P(\varphi) \odot b$ *iff* $\tau(\varphi) \odot b$;
6. $o_i \models_{\mathfrak{M}} P(\varphi_j) \circ P(\varphi_m) \odot b$ *iff* $\tau(\varphi_j) \circ \tau(\varphi_m) \odot b$;
7. $o_i \models_{\mathfrak{M}} \neg\chi$ *iff* $o_i \not\models_{\mathfrak{M}} \chi$;
8. $o_i \models_{\mathfrak{M}} \chi_j \wedge \chi_m$ *iff* $o_i \models_{\mathfrak{M}} \chi_j$ *and* $o_i \models_{\mathfrak{M}} \chi_m$.[2]

[2] Note that if χ has no occurrences of probability formulas inside of it, then the truth definition collapses onto that of propositional formulas.

The semantic interpretation of the operation signs + (addition) and − (subtraction) is the standard one from arithmetic as is that of the inequality/equality signs. It is easy to notice that in the truth definition of probability formulas, the specific outcome chosen plays no role (i.e., there is no difference in evaluating a probability formula ψ in o_1 or in o_2, whatever the outcomes say). This is perfectly reasonable, since we already said that the probability formulas are evaluated globally, rather than locally. Therefore, whenever a probability formula is satisfied at a pointed frequentist model, it is satisfied in all the model. We kept the notation of pointed models for probability formulas for coherence of notation and to more easily transition to general formulas χs. Note that for a χ formula, the valuation must be taken at a pointed model, since they might contain propositional formulas as their elements, which would then have to be evaluated locally.

Remark 2. Note that by simply dropping the probabilistic formulas from the language, the resulting language is that of propositional logic. Indeed, the semantics is also that of propositional logic, where the couples (O, v) can be seen as truth assignment to atomic formulas. This implies that, strictly speaking, **FPL** expands the expressivity of propositional logic maintaining the same semantics elements of it.

Example 4. We are now in a position to show how Alice could evaluate her formulas. Recall that Alice was interested in those propositions: "both dice landed on the value 3" ($p_{13} \wedge p_{2_3}$); "whenever die one landed on three, dice two did not" ($p_{13} \to \neg p_{23}$); "the relative frequency of die one landing three is $\frac{1}{6}$" ($P(p_{13}) = \frac{1}{6}$); "the relative frequency of die one landing three is lower than $\frac{1}{10}$" ($P(p_{13}) < \frac{1}{10}$). As should be expected, for the propositional propositions that Alice is asking, it must be specified which specific trial she has in mind. However, for the probability propositions, her evaluations should be global, again, as expected. In particular, in **FPL**, it is easy to show that $p_{13} \wedge p_{23}$ is satisfied only by outcome o_{14}, i.e., $o_{14} \models_{\mathfrak{M}} p_{13} \wedge p_{23}$. At the same time, it is always false that $p_{13} \to \neg p_{23}$. Moreover, it happens to be true in the model we constructed that $P(p_{13}) = \frac{1}{6}$, i.e., $\models_{\mathfrak{M}} P(p_{13}) = \frac{1}{6}$, while it is false that $P(p_{13}) < \frac{1}{10}$.

Definition 11 (Validity of propositional and probability formulas). *A propositional formula φ is said to be* ***satisfiable in a frequentist model*** $\mathfrak{M} = (O, v)$ *iff $\exists o_i \in O$ such that $o_i \models_{\mathfrak{M}} \varphi$. A propositional formula φ is said to be* ***valid in a frequentist model*** $\mathfrak{M} = (O, v)$ *(in symbols, $\models_{\mathfrak{M}} \varphi$) iff $\forall o_i \in O \in \mathfrak{M}, o_i \models_{\mathfrak{M}} \varphi$. For a probability formula, as already mentioned, satisfiability and validity in a model collapse, since the valuation of a probability formula ψ is global on the model. Thus, a probability formula ψ is said to be* ***valid in a frequentist model*** $\mathfrak{M} = (O, v)$ *iff $\models_{\mathfrak{M}} \psi$. A formula χ retains the same definitions of a propositional formula, with the distinction that whenever χ only contains probability formulas, then satisfiability collapses onto validity in a model.*

Proposition 3 (Relation between validity and satisfiability). *A propositional formula φ is valid iff its negation is not satisfiable in a frequentist model.*

Moreover, a probability formula ψ is valid iff its negation is not valid in a frequentist model.

Proof. Assume that a propositional formula φ is valid. This is equivalent to the fact that the formula is satisfied in every outcome o_i of every model $\mathfrak{M}$. Take an arbitrary model $\mathfrak{M}$ and an arbitrary outcome $o_i \in O \in \mathfrak{M}$. Take the negation of φ, i.e., $\neg\varphi$. To satisfy such formula, it must hold that $o_i \models_{\mathfrak{M}} \neg\varphi$, which, by definition, means that $o_i \not\models_{\mathfrak{M}} \varphi$. However, by assumption, we know that there is no o_i s.t. $o_i \not\models_{\mathfrak{M}} \varphi$, thus there is no o_i s.t. $o_i \models_{\mathfrak{M}} \neg\varphi$.

For probability formula the argument is similar, with the difference being that validity in a model is taken into consideration instead of satisfiability. This is due to the fact that probability formulas are evaluated globally on the model and not at a specific outcome.

An example of a valid formula in **FPL** would be $P(\varphi) \geq \frac{1}{2} \to P(\varphi) \geq \frac{1}{4}$.

Proposition 4. *The following two sentences are equivalent, where φ is a propositional formula:*

(a) $\models \varphi$;
(b) $\models P(\varphi) = 1$.

Proof. Take an arbitrary model $\mathfrak{M} = (O, v)$.

(From a to b). Assume that $\models \varphi$. It therefore follows, by definition 11, that $\forall o_i \in O, o_i \models_{\mathfrak{M}} \varphi$. By definition 9, that $\forall o_i \in O, v^e(o_i, \varphi) = 1$. The validating set $[\![\varphi]\!]$ is therefore equivalent to O, from which it follows that $\frac{|[\![\varphi]\!]|}{|O|} = 1$, which in turn implies $\frac{|[\![\varphi]\!]|}{|O|} = 1$. This is the definition of $\models_{\mathfrak{M}} P(\varphi) = 1$, via definition 10. Since the model $\mathfrak{M}$ was taken arbitrarily, this holds for all possible models chosen.[3]

(From b to a). Assume that $\models_{\mathfrak{M}} P(\varphi) = 1$. It follows, by definition 10, that $\frac{|[\![\varphi]\!]|}{|O|} = 1$, i.e., $|[\![\varphi]\!]| = |O|$. By definition 5, $[\![\varphi]\!] \subseteq O$. Since O is finite, the only possible case in which the two conditions $[\![\varphi]\!] \subseteq O$ and $|[\![\varphi]\!]| = |O|$ are both true at the same time is when $[\![\varphi]\!] = O$. From such equality and definition 5, it follows that $\forall o_i \in O, v^e(o_i, \varphi) = 1$, which, by definition 9, means that $\forall o_i \in O, o_i \models_{\mathfrak{M}} \varphi$. Finally, by definition 11, this means that $\models_{\mathfrak{M}} \varphi$. Since the model $\mathfrak{M}$ was taken arbitrarily, this holds for all possible models chosen.[4]

3.3 Axiomatic system for FPL

We will now provide an axiomatic system for **FPL**. We will then show that such axiomatic system is sound and complete with respect to frequentist models.

[3] Technically, both $P(\varphi)$ and 1 should be qualified with a coefficient a. However, multiplying both for the same number, would not change the result of the valuation, thus the formula would be still valid for any a chosen.

[4] Note that it is assumed in **FPL** that O is finite. While using an infinite O would not change the a to b it might indeed make the b to a direction false.

Definition 12 (Frequentist Probability Logic). *A* ***frequentist probability logic*** *(FPL) is a set $\Lambda \subseteq \mathcal{L}_{\mathbf{FPL}}$ containing (i) all propositional tautologies, (ii) all the substitution instances of valid formulas about linear inequalities, (iii) all the substitution instances of the following axiom schemata, for $\varphi, \psi \in \mathcal{L}_{\mathbf{FPL}}$:*

Ax. 1 $P(\varphi) \geq 0$;
Ax. 2 $P(\top) = 1$;
Ax. 3 If $\neg(\varphi_i \wedge \varphi_j)$, *then* $P(\varphi_i \vee \varphi_j) = P(\varphi_i) + P(\varphi_j)$;[5]
Ax. 4 If $(\varphi_i \leftrightarrow \varphi_j)$, *then* $P(\varphi_i) = P(\varphi_j)$.

Λ is also closed under Modus Ponens (MP).

Instances of (i) formalize reasoning about propositional formulas and Boolean combinations between formulas (either propositional or probability formulas or a mix of the two); instances of (ii) formalize reasoning about formulas of the form $P(\varphi_i) \circ P(\varphi_j) \odot b$, where the $P(\varphi)$s can be seen as variables in the inequalities. A formula about linear inequality is valid only if every numerical assignment to the variables in the inequality. i.e., the $\tau(\varphi)$ of the $P(\varphi)$s appearing in the inequality, satisfy the inequality. Ax. 1 states that probability formulas can never receive a negative value. Ax. 2 states that the probability of a tautology is always 1 (see also proposition 4). Ax. 3 captures the idea of finite additivity of probability theory; this axiom is also what creates the connection between $P(\varphi) \circ P(\varphi) \odot b$ formulas and $P(\varphi) \odot b$ formulas (i.e., it is an interaction axiom between simple probability formulas and linear combinations of them). Ax. 4 captures the idea that equivalent formulas should have the same probability.

Definition 13. *We say that a formula χ is* ***provable*** *in Λ ($\vdash_\Lambda \chi$) iff the formula is an instance of an axiom schema in Λ or is obtained through the application of MP to formulas already proven. A formula χ is* ***derivable*** *in Λ from a set of premises Γ ($\Gamma \vdash_\Lambda \chi$) iff the formula is an instance of an axiom schema in Λ, is an instance of a formula contained in Γ, or is obtained through the application of MP to formulas already derived or proven.*

We say that a formula χ is ***inconsistent*** *if its negation $\neg\chi$ is provable ($\vdash_\Lambda \neg\chi$). A formula is* ***consistent*** *otherwise.*

3.4 Probability Theorems

We will now show that the main theorems from probability theory can be proven in **FPL**. We will provide semantic proofs rather than syntactic ones. Our choice is justified by the fact that the major novel contribution to the literature of **FPL** is semantic in nature, rather than syntactic. For convenience, we also provide syntactic proofs in the appendix of this paper.

Theorem 1 (Probability of $\bot$).

[5] Finite additivity could also be expressed by the formula $P(\varphi_i) = P(\varphi_i \wedge \varphi_j) + P(\varphi_i \wedge \neg\varphi_j)$. This fact is important since we will use this formulation in our proof of completeness.

$$\models P(\bot) = 0$$

Proof. Take an arbitrary model $\mathfrak{M}$ and an arbitrary outcome $o_i \in O \in \mathfrak{M}$. By definition 2, $\bot \coloneqq \neg\top$, and by definition 9, $\top$ holds for every $o_i \in O \in \mathfrak{M}$. By definition 5, it follows that $[\![\top]\!] = O$. By definition 1, it follows that $[\![\bot]\!] = O \backslash [\![\top]\!]$. By substitution, it follows that $[\![\bot]\!] = \emptyset$. Therefore, $\tau(\bot) = \frac{|\emptyset|}{|O|}$. Whatever the cardinality of O is, the cardinality of $\emptyset$ is zero. It follows that $\tau(\bot) = 0$, which means that $P(\bot) = 0$ holds. Since the outcome and the model were chosen arbitrarily, the results holds for any outcome and any model.

Theorem 2 (Finite Additivity).

$$\textit{If} \models \neg(\varphi_i \wedge \varphi_j),\ \textit{then} \models P(\varphi_i \vee \varphi_j) = P(\varphi_i) + P(\varphi_j)$$

Proof. Take an arbitrary model $\mathfrak{M}$. Assume that $\models \neg(\varphi_i \wedge \varphi_j)$, thus, in particular, $\models_{\mathfrak{M}} \neg(\varphi_i \wedge \varphi_j)$. This means that for each $o_i \in O \in \mathfrak{M}$, either $o_i \notin [\![\varphi_i]\!]$ or $o_i \notin [\![\varphi_j]\!]$. By definition 10, $P(\varphi_i \vee \varphi_j) = b$ iff $\tau(\varphi_i \vee \varphi_j) = b$. By definition 7, $\tau(\varphi_i \vee \varphi_j) = \frac{|[\![\varphi_i \vee \varphi_j]\!]|}{|O|}$. By proposition 1, it follows that $[\![\varphi_i \vee \varphi_j]\!] = [\![\varphi_i]\!] \cup [\![\varphi_j]\!]$. This implies that $|[\![\varphi_i \vee \varphi_j]\!]| = |[\![\varphi_i]\!] \cup [\![\varphi_j]\!]|$. By the initial assumption, it follows that $[\![\varphi_i]\!] \cap [\![\varphi_j]\!] = \emptyset$. Thus, every $o_i \in [\![\varphi_i]\!] \cup [\![\varphi_j]\!]$ contributes only once to the cardinality of $|[\![\varphi_i]\!] \cup [\![\varphi_j]\!]|$, and coming from just one of the sets $[\![\varphi_i]\!]$ or $[\![\varphi_j]\!]$. This implies that $|[\![\varphi_i]\!] \cup [\![\varphi_j]\!]| = |[\![\varphi_i]\!]| + |[\![\varphi_j]\!]|$. By substitution, it follows that $\tau(\varphi_i \vee \varphi_j) = \frac{|[\![\varphi_i \vee \varphi_j]\!]|}{|O|} = \frac{|[\![\varphi_i]\!]| + |[\![\varphi_j]\!]|}{|O|}$. By splitting the fraction, it follows that $\tau(\varphi_i \vee \varphi_j) = \frac{|[\![\varphi_i]\!]|}{|O|} + \frac{|[\![\varphi_j]\!]|}{|O|}$. By definition 10, it follows that $P(\varphi_i \vee \varphi_j) = P(\varphi_i) + P(\varphi_j)$. Since the model $\mathfrak{M}$ was chosen arbitrarily, this holds for every model.

Theorem 3 (Probability of Negation).

$$\models P(\neg\varphi) = 1 - P(\varphi)$$

Proof. Take an arbitrary model $\mathfrak{M}$ and an arbitrary outcome $o_i \in O \in \mathfrak{M}$. Take the tautology $\varphi \vee \neg\varphi$. Given the fact that it is a tautology, it follows that $\models (\varphi \vee \neg\varphi)$. By proposition 4, it follows that $\models P(\varphi \vee \neg\varphi) = 1$. It also holds that $\models_{\mathfrak{M}} \neg(\varphi \vee \neg\varphi)$. By theorem 2, it follows that $P(\varphi \vee \neg\varphi) = P(\varphi) + P(\neg\varphi)$. By substitution, it follows that $P(\varphi) + P(\neg\varphi) = 1$. By algebra, it follows that $P(\neg\varphi) = 1 - P(\varphi)$. Since the model and the outcome were chosen arbitrarily, this holds for any outcome and any model.

Theorem 4 (Probability of Equivalence).

$$\textit{If} \models \varphi_i \leftrightarrow \varphi_j,\ \textit{then} \models P(\varphi_i) = P(\varphi_j)$$

Proof. Take an arbitrary model $\mathfrak{M}$. Assume that $\models_{\mathfrak{M}} \varphi_i \leftrightarrow \varphi_j$. From the assumption, it follows that $[\![\varphi_i]\!] = [\![\varphi_j]\!]$ (if the formulas are satisfied in the same outcomes of the model, they must have the same validating set). Therefore, the conclusion follows directly by definitions 7 and 10, since $\frac{|[\![\varphi_i]\!]|}{|O|} = \frac{|[\![\varphi_j]\!]|}{|O|}$ (the equivalence follows easily from the fact that the denominators are equivalent and so are the numerators of the fractions). The previous fact directly implies that $P(\varphi_i) = P(\varphi_j)$. Since the model was chosen arbitrarily, the same holds in every model.

Theorem 5 (Strong Additivity).

$$\models P(\varphi_i \vee \varphi_j) + P(\varphi_i \wedge \varphi_j) = P(\varphi_i) + P(\varphi_j)$$

Proof. Take the two tautologies $\varphi_i \leftrightarrow (\varphi_i \wedge \varphi_j) \vee (\varphi_i \wedge \neg\varphi_j)$ and $\varphi_j \leftrightarrow (\varphi_j \wedge \varphi_i) \vee (\varphi_j \wedge \neg\varphi_i)$. For simplicity, name $(\varphi_i \wedge \varphi_j)$ as α_1, $(\varphi_i \wedge \neg\varphi_j)$ as α_2, $(\varphi_j \wedge \varphi_i)$ as β_1 and $(\varphi_j \wedge \neg\varphi_i)$ as β_2.

Since the formulas taken are tautologies, it follows that $\models \varphi_i \leftrightarrow (\alpha_1 \vee \alpha_2)$ and that $\models \varphi_j \leftrightarrow (\beta_1 \vee \beta_2)$ By theorem 4, we derive that $\models P(\varphi_i) = P(\alpha_1 \vee \alpha_2)$ and $\models P(\varphi_j) = P(\beta_1 \vee \beta_2)$. Note that $\models \neg(\alpha_1 \wedge \alpha_2)$ and that $\models \neg(\beta_1 \wedge \beta_2)$. By theorem 2, it follows that $\models P(\alpha_1 \vee \alpha_2) = P(\alpha_1) + P(\alpha_2)$ and that $\models P(\beta_1 \vee \beta_2) = P(\beta_1) + P(\beta_2)$. By algebra and substitution of equivalents, it follows that $\models P(\varphi_i) + P(\varphi_j) = P(\alpha_1) + P(\alpha_2) + P(\beta_1) + P(\beta_2)$. Name this *fact one*.

Now, take the tautology $(\varphi_i \vee \varphi_j) \leftrightarrow (\alpha_1) \vee (\alpha_2) \vee (\beta_2)$. Since this is a tautology, it follows that $\models (\varphi_i \vee \varphi_j) \leftrightarrow (\alpha_1) \vee (\alpha_2) \vee (\beta_2)$. By theorem 4, $\models P(\varphi_i \vee \varphi_j) = P(\alpha_1 \vee \alpha_2 \vee \beta_2)$. Also note that $\models \neg(\alpha_1 \wedge (\alpha_2 \vee \beta_2))$. By theorem 2, this means that $\models P(\alpha_1 \vee \alpha_2 \vee \beta_2) = P(\alpha_1) + P(\alpha_2 \vee \beta_2)$. Note also that $\models \neg(\alpha_2 \wedge \beta_2)$. By theorem 2, it follows that $\models P(\alpha_2 \vee \beta_2) = P(\alpha_2) + P(\beta_2)$. Therefore, $\models P(\alpha_1 \vee \alpha_2 \vee \beta_2) = P(\alpha_1) + P(\alpha_2) + P(\beta_2)$. By equivalence, this means that $\models P(\varphi_i \vee \varphi_j) = P(\alpha_1) + P(\alpha_2) + P(\beta_2)$ Name this *fact two*.

By putting together *fact one* and *fact two*, it follows that $\models P(\varphi_i \vee \varphi_j) = P(\varphi_i) + P(\varphi_j) - P(\beta_1)$, which, by algebra, is equivalent to $\models P(\varphi_i \vee \varphi_j) + P(\beta_1) = P(\varphi_i) + P(\varphi_j)$. By using the definition of β_1 and the commutativity of conjunction, it follows that $\models P(\varphi_i \vee \varphi_j) + P(\varphi_i \wedge \varphi_j) = P(\varphi_i) + P(\varphi_j)$.

4 Soundness and Completeness

We will now show that Λ is sound with respect to the class of all frequentist models.

4.1 Soundness

Theorem 6 (Soundness). *If* $\vdash_\Lambda \chi$ *then* $\models \chi$

Proof. In order to prove soundness, we will show that all axiom schemata are valid with respect to the class of all frequentist models and then show that MP preserves validity.

For instances of (i) and (ii) of definition 12, the proof is straightforward and follows from the definition of propositional tautology and from the algebraic properties of the mathematical operations, which were assumed in frequentists models.

For instances of Ax. 1, note that the value of $P(\psi)$ is obtained by taking $\frac{|[\![\varphi]\!]|}{|O|}$. By definition 5, $\emptyset \subseteq [\![\varphi]\!] \subseteq O$. This implies that $|\emptyset| \leq |[\![\varphi]\!]| \leq |O|$. In addition, by definition 3, it follows that $\emptyset \subset O$, which implies that $|\emptyset| < |O|$. Since $|\emptyset| = 0$, it follows that O is strictly positive. From those premises, it follows directly that $\frac{|[\![\varphi]\!]|}{|O|}$ is both defined ($|O|$ is different from zero) and is bigger or equal to zero, which proves the validity of Ax. 1.

For instances of Ax. 2, note that $[\![\top]\!] = O$. Since the two sets are equivalent, so is their cardinality, i.e., $|[\![\top]\!]| = |O|$. Moreover, the value of $P(\top)$ is equivalent to $\frac{|[\![\top]\!]|}{|O|}$. Given the fact that $|O|$ is finite and strictly bigger than 0, the fact that the numerator and denominator of $\frac{|[\![\top]\!]|}{|O|}$ are equivalent, implies that their ratio is equal to 1, which means that $P(\top) = 1$.

Instances of Ax. 3 are just instances of theorem 2.

Instances of Ax. 4 are just instances of theorem 4.

Finally, we must show that MP preserves validity. For the propositional formulas, this follows straightforwardly from propositional logic and the way the valuation function works. For probability formulas, it must be shown that if $P(\varphi_i) \odot b_m$ and $P(\varphi_i) \odot b_m \rightarrow P(\varphi_j) \odot b_t$, it follows that $P(\varphi_j) \odot b_t$.

Assume that $P(\varphi_i) \odot b_m$ and that $P(\varphi_i) \odot b_m \rightarrow P(\varphi_j) \odot b_t$. By definition 10, the first assumption asserts that $\frac{|[\![\varphi_i]\!]|}{|O|} \odot b_m$. By the definition of $\rightarrow$, the second assumption asserts that it must hold that either it is not the case that $\frac{|[\![\varphi_i]\!]|}{|O|} \odot b_m$ or $\frac{|[\![\varphi_j]\!]|}{|O|} \odot b_t$. By the first assumption, we know that it cannot be the first disjunct, therefore it must be the second one. However, the second one is just the definition of $P(\varphi_j) \odot b_t$. Thus, MP preserves validity.

4.2 Weak Completeness

To proof weak completeness we will mimic the proof given in [5]. In order to do so, some fact must be established.

Since **FPL** is equivalent to propositional logic with the addition of probability formulas, and it is known that propositional logic is complete, we must only show that the added probability formulas also preserve completeness. This means, in particular, that we can focus only on the ψ formulas of the language. Thus, we must show that if a probability formula is valid, then it is provable.

Definition 14 (Literals and Literal Formula). *We define as **literal** any atomic formula p_i or its negation $\neg p_i$. A **literal formula** δ is a formula $\varphi_1 \wedge \cdots \wedge \varphi_n$, where each φ_i is a literal.*

Definition 15. *We use the notation $At(\chi)$ to indicate the set of all atomic formulas contained in χ.*

Lemma 1. *Let φ be a propositional formula. Call $Lit(\varphi)$ the set of all literal formulas δ obtained from $At(\varphi)$ such that $\delta \rightarrow \varphi$ is a propositional tautology. It is provable that $P(\varphi) = \sum_{\delta \in Lit(\varphi)} P(\delta)$.*

Proof. The proof will be given by induction on the size of $Lit(\varphi)$. Note that the main claim of lemma 1 can be expanded as follows (where the γ formulas are literals taken from $Lit(\varphi)$):

$$P(\varphi) = P(\varphi \wedge \gamma_1) + \cdots + P(\varphi \wedge \gamma_{2^i})$$

Base case. Assume that $i = 1$.

It must be shown that $P(\varphi) = P(\varphi \wedge \gamma_1) + P(\varphi \wedge \gamma_2)$ is provable. By the construction processes mentioned above, $\gamma_2 := \neg\gamma_1$ (γ must be taken from the set of literal formulas, but since in the base of the induction, the set $Lit(\varphi)$ contains only one element, it is possible only to construct the two literal formulas γ_1 and $\neg\gamma_1$). Therefore, it must be proven that $P(\varphi) = P(\varphi \wedge \gamma_1) + P(\varphi \wedge \neg\gamma_1)$. This follows from Ax. 3 of definition 12. Thus, the formula is provable.

Inductive base. Assume that the following is provable:

$$P(\varphi) = P(\varphi \wedge \gamma_1) + \cdots + P(\varphi \wedge \gamma_{2^i})$$

By Ax. 3, it is provable that $P(\varphi \wedge \gamma_1) = P(\varphi \wedge \gamma_1 \wedge p_{i+1}) + P(\varphi \wedge \gamma_1 \wedge \neg p_{i+1})$. Using propositional reasoning (clause (i) in definition 12) and instances of valid inequality formulas (clause (ii) in definition 12), it is possible to substitute each formula $P(\varphi \wedge \gamma_r)$ in the inductive base with a formula $P(\varphi \wedge \gamma_r \wedge p_{i+1}) + P(\varphi \wedge \gamma_r \wedge \neg p_{i+1})$. This implies that the inductive step is provable:

$$P(\varphi) = P(\varphi \wedge \gamma_1) + \cdots + P(\varphi \wedge \gamma_{2^{i+1}}).$$

As a particular case, the following formula follows:

$$P(\varphi) = P(\varphi \wedge \delta_1) + \cdots + P(\varphi \wedge \delta_{2^n}) \tag{2}$$

Now, take the set $At(\varphi)$, i.e., the set of all atomic formulas contained in φ. By propositional reasoning, if $\delta_r \in At(\varphi)$, then $(\varphi \wedge \delta_r) \leftrightarrow \delta_r$, which, by Ax. 4, means that $P(\varphi \wedge \delta_r) = P(\delta_r)$. If $\delta_r \notin At(\varphi)$, then $(\varphi \wedge \delta_r) = \bot$, which, by Ax. 4 and Theorem 1, means that $P(\varphi \wedge \delta_r) = 0$.

Given those facts, in formula 2, each $P(\varphi \wedge \delta_r)$ can be either substituted with $P(\delta_r)$ or 0, depending on whether δ_r is included in $At(\varphi)$ or not. Therefore, lemma 1 follows.

Theorem 7 (Weak Completeness). *If* $\models \chi$ *then* $\vdash_\Lambda \chi$.

Proof. As previously stated, we can limit ourselves to probability formulas and their Boolean combinations, since the completeness of the portion of the language with propositional formulas follows from the completeness of propositional logic (note that our semantics does not differ from that of propositional logic). Moreover, due to proposition 4, it is possible to reduce the problem of the provability of combinations of propositional and probability formulas to that of combinations of just probability formulas (which we will prove now).

We therefore prove completeness by showing that any consistent formula is satisfiable. For the proof we will use χ to refer to Boolean combinations of probability formulas, thus excluding the portion of χ containing propositional formulas. A simple probability formula $P(\varphi) \odot b$ could always be read as a Boolean combination between $P(\varphi) \odot b \wedge P(\top) = 1$.

Assume that χ is consistent. First, transform χ into its disjunctive normal form (DNF), i.e., a disjunction of conjunctions (call each one of the disjunctive clauses χ_{clause_i}). Call this formula χ_{DNF} (χ_{DNF} will look like this: $\chi_{clause_1} \vee \dots \vee \chi_{clause_r}$). By propositional reasoning it is possible to show that χ and χ_{DNF} are provably equivalent. Thus, since χ is consistent, so is χ_{DNF}. This implies that at least one of the clauses χ_{clause_i} must be consistent. Imagine that this is not the case, i.e., $\neg\chi_{clause_i}$ was provable for every i, then it is easy to see that $\neg(\chi_{clause_1} \vee \dots \vee \chi_{clause_r})$ was also provable, which would make χ_{DNF} inconsistent, which is a contradiction. In addition to knowing that there is at least one χ_{clause_i} which is consistent, every model that satisfies χ_{clause_i} must also satisfies χ. We can therefore limit ourselves to the evaluation of a χ_{clause_i}. Call this formula f. Recall that all the elements of f are probability formulas ψ.

By lemma 1, we know that we can construct a formula f' starting from f by substituting every conjunct in f, with a formula $P(\delta 1) + \dots + P(\delta_{2^n})$, where $At(f)$ includes all the atomic propositions that appear in f and where the $\delta_1, \dots, \delta_{2^n}$ are the $Lit(f)$.

Now, construct a formula f'' from f' by adding as conjuncts all the formulas $P(\delta_j) \geq 0$, with $1 \leq j \leq 2^n$, and the formula $P(\delta_1) + \dots + P(\delta_{2^n}) = 1$. The first set of additions follows from Ax. 1, while the last addition is an instance of lemma 1, where $\varphi = \top$, and Ax. 2.

This new formula f'' is provably equivalent to f' and therefore to f. To prove completeness, we therefore must show that f'' is satisfiable. Note that f'' is a conjunction of $2^n + r + s + 1$ formulas of the form:

$$\begin{aligned} P(\delta_1) + \dots + P(\delta_{2^n}) = 1 \\ P(\delta_1) \geq 0 \\ \dots \\ P(\delta_{2^n}) \geq 0 \\ P(\delta_1)_r + \dots + P(\delta_{2^n})_r \geq b_r \\ P(\delta_1)_s + \dots + P(\delta_{2^n})_s < b_s \end{aligned} \tag{3}$$

where each $P(\delta_1)_r + \dots + P(\delta_{2^n})_r \geq b_r$ captures the positive conjuncts of f', while each $P(\delta_1)_s + \dots + P(\delta_{2^n})_s < b_s$ captures the negative conjuncts of f'.

Equation 3 can be easily transformed into a system of linear inequalities by substituting all instances of formulas with a variable (maintaining uniform substitution). Thus, the equation would become:

$$\begin{aligned} x_1 + \dots + x_{2^n} = 1 \\ x_1 \geq 0 \\ \dots \\ x_{2^n} \geq 0 \\ x_{1_r} + \dots x_{2^n_r} \geq b_r \\ x_{1_s} + \dots x_{2^n_s} \geq b_s \end{aligned} \tag{4}$$

If equation 4 is satisfiable, so is f'', and, in turn, f', f and χ. Now, assume that f'' is unsatisfiable. This implies that equation 4 is unsatisfiable. Therefore, $\neg(f'')$ becomes an instance of (ii) from definition 12. That would mean, however, that f and therefore χ is inconsistent, which is a contradiction. Thus, f'' must be satisfiable.

5 Conclusion and Future Works

In the paper, we presented a logic **FPL** to reason about probabilities with a relative frequency interpretation. We showed that it is possible to interpret the language of **FPL** with the standard semantics for propositional logic. **FPL** can give a peculiar frequentist interpretation of the probability operator as presented in [5]. We then gave a proof system for the language, proved that the traditional theorems of probability hold in our language, and established that the techniques employed in [5] to prove soundness and completeness work also for our interpretation.

In the future, we have two evolving plans. The first plan is to increase the expressiveness of **FPL** by evolving the logic in various directions. The second plan is to apply the logic to various contexts related to computer science and reasoning under uncertainty. As far as the first plan goes, we would like: (i) to add a third truth value to the codomain of v, in order to express irrelevance of a proposition to a given outcome. This would allow us to construct examples where only specific propositions (and not all of them) are tested during an experiment. (ii) to add multiple agents to the logic and to add communication channels between them. This would allow us to model scenarios in which different agents ran different experiments and then communicated their results to each other. (iii) to add dynamism to the language. This is the most interesting addition to the logic, since it would allow us to provide updating techniques which do not reduce to Bayesian updating. Specifically, we would like to construct a logic which allows a frequentist updating possibility.

Obviously, also combinations of (i), (ii), and (iii) would be interesting additions to **FPL**. As far as the second plan goes, we would like: (i) to employ **FPL** to reason about uncertainty in logics with trust operators, which could then measure the ratio of positive recommendations over all recommendations and of positive direct experiences (e.g., it would be interesting to evolve the propositional components of, e.g., [14, 13, 1, 2] to add a probabilistic part). (ii) to employ **FPL** to reason about probabilistic verification of machine learning, and statistical and approximate model checking.

References

1. Aldini, A., Curzi, G., Graziani, P., Tagliaferri, M.: Trust evidence logic. In: Vejnarová, J., Wilson, N. (eds.) Symbolic and Quantitative Approaches to Reasoning with Uncertainty: 16th European Conference (ECSQARU 2021). LNAI, vol. 12897, pp. 575–589. Springer (2021)

2. Aldini, A., Tagliaferri, M.: Logics to reason formally about trust computation and manipulation. In: Saracino, A., Mori, P. (eds.) Emerging Technologies for Authorization and Authentication. LNCS, vol. 11967, pp. 1–15. Springer (2020)
3. Antonelli, M., Lago, U.D., Pistone, P.: On counting propositional logic and wagner's hierarchy. In: Coen, C.S., Salvo, I. (eds.) Proceedings of the 22nd Italian Conference on Theoretical Computer Science. CEUR Workshop Proceedings, vol. 3072. Technical University of Aachen (2021)
4. Demey, L., Kooi, B., Sack, J.: Logic and Probability. In: Zalta, E.N., Nodelman, U. (eds.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Fall 2023 edn. (2023)
5. Fagin, R., Halpern, J.Y., Megiddo, N.: A logic for reasoning about probabilities. Information and Computation **87**(1-2), 78–128 (1990)
6. Fattorosi-Barnaba, M., Caro, F.D.: Graded modalities. I. Studia Logica **44**(2), 197–221 (1985)
7. Fazlyab, M., Morari, M., Pappas, G.J.: Probabilistic verification and reachability analysis of neural networks via semidefinite programming. In: 2019 IEEE 58th Conference on Decision and Control (CDC). pp. 2726–2731 (2019)
8. Hérault, T., Lassaigne, R., Magniette, F., Peyronnet, S.: Approximate probabilistic model checking. In: Steffen, B., Levi, G. (eds.) Verification, Model Checking, and Abstract Interpretation. pp. 73–84. Springer (2004)
9. Hájek, A.: Interpretations of Probability. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Fall 2019 edn. (2019)
10. Legay, A., Delahaye, B., Bensalem, S.: Statistical model checking: An overview. In: Barringer, H., Falcone, Y., Finkbeiner, B., Havelund, K., Lee, I., Pace, G., Roşu, G., Sokolsky, O., Tillmann, N. (eds.) Runtime Verification. pp. 122–135. Springer (2010)
11. O'Donoghue, B., Osband, I., Ionescu, C.: Making sense of reinforcement learning and probabilistic inference (2020)
12. Pilipovsky, J., Sivaramakrishnan, V., Oishi, M., Tsiotras, P.: Probabilistic verification of ReLU neural networks via characteristic functions. In: Matni, N., Morari, M., Pappas, G.J. (eds.) Proceedings of The 5th Annual Learning for Dynamics and Control Conference. vol. 211, pp. 966–979. PMLR (2023)
13. Tagliaferri, M., Aldini, A.: A trust logic for the varieties of trust. In: Camara, J., Steffen, M. (eds.) Software Engineering and Formal Methods. LNCS, vol. 12226, pp. 119–136. Springer (2020)
14. Tagliaferri, M., Aldini, A.: From belief to trust: A quantitative framework based on modal logic. Journal of Logic and Computation **32**(6), 1017–1047 (2022)

A Syntactic Proofs

Theorem 8 (Probability of $\bot$).

$$\vdash P(\bot) = 0 \tag{5}$$

Proof. Direct Proof.

1. $\neg(\top \wedge \bot)$	Taut.
2. $P(\top \vee \bot) = P(\top) + P(\bot)$	Ax. 3 from (1)
3. $P(\top) = 1$	Ax. 2
4. $(\top \vee \bot) \leftrightarrow \top$	Taut.
5. $P(\top \vee \bot) = P(\top)$	Ax. 4 from (4)
6. $P(\top \vee \bot) = 1$	Transitivity of =
7. $1 = 1 + P(\bot)$	Substitution from (2), (3), and (6)
8. $P(\bot) = 0$	By algebra over (7).

Theorem 9 (Probability of Negation).

$$\vdash P(\neg\varphi) = 1 - P(\varphi) \tag{6}$$

Proof. Direct Proof.

1. $P(\top) = 1$	Ax.2
2. $(\varphi \vee \neg\varphi) \leftrightarrow \top$	Taut.
3. $P(\varphi \vee \neg\varphi) = P(\top)$	Ax. 4 from (2)
4. $P(\varphi \vee \neg\varphi) = 1$	Substitution from (1) and (3)
5. $\neg(\varphi \wedge \neg\varphi)$	Taut.
6. $P(\varphi \vee \neg\varphi) = P(\varphi) + P(\neg\varphi)$	Ax. 3 from (5)
7. $P(\top) = P(\varphi) + P(\neg\varphi)$	Transitivity of =
8. $1 = P(\varphi) + P(\neg\varphi)$	Substitution over (1) and (7)
9. $P(\neg\varphi) = 1 - P(\varphi)$	By algebra over (8).

Theorem 10 (Strong Additivity).

$$\vdash P(\varphi_i \vee \varphi_j) + P(\varphi_i \wedge \varphi_j) = P(\varphi_i) + P(\varphi_j) \tag{7}$$

Proof. Direct Proof.

1. $\varphi_i \leftrightarrow (\varphi_i \wedge \varphi_j) \vee (\varphi_i \wedge \neg\varphi_j)$ — Taut.
2. $\varphi_j \leftrightarrow (\varphi_j \wedge \varphi_i) \vee (\varphi_j \wedge \neg\varphi_i)$ — Taut.
3. $P(\varphi_i) = P((\varphi_i \wedge \varphi_j) \vee (\varphi_i \wedge \neg\varphi_j))$ — Ax. 4 from (1)
4. $P(\varphi_j) = P((\varphi_j \wedge \varphi_i) \vee (\varphi_j \wedge \neg\varphi_i))$ — Ax. 4 from (2)
5. $\neg((\varphi_i \wedge \varphi_j) \wedge (\varphi_i \wedge \neg\varphi_j))$ — Taut.
6. $P((\varphi_i \wedge \varphi_j) \vee (\varphi_i \wedge \neg\varphi_j)) =$ $= P(\varphi_i \wedge \varphi_j) + P(\varphi_i \wedge \neg\varphi_j)$ — Ax. 3 from (5)
7. $\neg((\varphi_j \wedge \varphi_i) \wedge (\varphi_j \wedge \neg\varphi_i))$ — Taut.
8. $P((\varphi_j \wedge \varphi_i) \vee (\varphi_j \wedge \neg\varphi_i)) =$ $= P(\varphi_j \wedge \varphi_i) + P(\varphi_j \wedge \neg\varphi_i)$ — Ax. 3 from (7)
9. $P(\varphi_i) + P(\varphi_j) = P(\varphi_i \wedge \varphi_j) + P(\varphi_i \wedge \neg\varphi_j) +$ $+ P(\varphi_j \wedge \varphi_i) + P(\varphi_j \wedge \neg\varphi_i)$ — Algebra and substitution from (3), (4), (6) and (8)
10. $(\varphi_i \vee \varphi_j) \leftrightarrow (\varphi_i \wedge \neg\varphi_j) \vee (\varphi_j \wedge \neg\varphi_i) \vee (\varphi_i \wedge \varphi_j)$ — Taut.
11. $P(\varphi_i \vee \varphi_j) =$ $= P((\varphi_i \wedge \neg\varphi_j) \vee (\varphi_j \wedge \neg\varphi_i) \vee (\varphi_i \wedge \varphi_j))$ — Ax. 4 from (10)
12. $\neg((\varphi_i \wedge \varphi_j) \wedge ((\varphi_i \wedge \neg\varphi_j) \vee (\varphi_i \wedge \varphi_j)))$ — Taut.
13. $P((\varphi_i \wedge \neg\varphi_j) \vee (\varphi_j \wedge \neg\varphi_i) \vee (\varphi_i \wedge \varphi_j)) =$ $= P(\varphi_i \wedge \neg\varphi_j) + P((\varphi_j \wedge \neg\varphi_i) \vee (\varphi_i \wedge \varphi_j))$ — Ax. 3 from (12)
14. $\neg((\varphi_j \wedge \neg\varphi_i) \wedge (\varphi_i \wedge \varphi_j))$ — Taut.
15. $P((\varphi_j \wedge \neg\varphi_i) \vee (\varphi_i \wedge \varphi_j)) =$ $= P(\varphi_j \wedge \neg\varphi_i) + P(\varphi_i \wedge \varphi_j)$ — Ax. 3 from (14)
16. $P((\varphi_i \wedge \neg\varphi_j) \vee (\varphi_j \wedge \neg\varphi_i) \vee (\varphi_i \wedge \varphi_j)) =$ $= P(\varphi_i \wedge \neg\varphi_j) + P(\varphi_j \wedge \neg\varphi_i) + P(\varphi_i \wedge \varphi_j)$ — From (13) and (15)
17. $P(\varphi_i) + P(\varphi_j) = P(\varphi_i \vee \varphi_j) + P(\varphi_j \wedge \varphi_i)$ — From (9), (11) and (16)

Note that the proofs of *Finite Additivity* and of *Equivalence of Probabilities* follow directly from the axioms of our proof system. You just have to assume the condition of the axioms and then get the conclusion directly by applying the relevant axiom.