Nebojša Ikodinović*, Zoran Ognjanović**, Miodrag Rašković** and Zoran Marković**

FIRST-ORDER PROBABILISTIC LOGICS AND THEIR APPLICATIONS

Abstract. The paper offers an overview of the authors' results on formalization of uncertain reasoning in first-order framework.

Mathematics Subject Classification (2010): Primary: 03B48, 03B60, 03B70, 03B35, 03B42; Secondary 68T37, 68T15, 68T30.

Keywords: probability logic, conditional probability, approximate probability, non-standard analysis, strong completeness, decidability, default reasoning

*Faculty of Mathematics, University of Belgrade, Belgrade, Serbia
ikodinovic@matf.bg.ac.rs

**Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia
zorano@mi.sanu.ac.rs, zoranm@mi.sanu.ac.rs

Contents

1. Introduction

The paper [92] presented a survey of our work on probabilistic logics [10, 18–22, 43,44,52,67–69,76–88,100–117]. Various variants of probabilistic logics (i.e., finitary or infinitary, with or without iterations of probabilistic operators, with different types of probabilistic operators - both for conditional and unconditional probabilities, propositional or first-order, extending classical, intuitionistic or temporal logics, with different ranges of probability functions, finitely or σ-additive, etc.) were presented there, and the usual model-theoretical and proof-theoretical issues about compactness, axiomatizations, completeness and decidability were discussed. That paper also contained a short historical overview of studies relating logic and probability: from Leibnitz via Jacobus Bernoulli and Boole until the late 1980's, and the work of Keisler, Nilsson and Fagin and Halpern [23, 24, 35, 55–60, 74, 75]. Interested readers can consult an extensive discussion on this historical topic given in [34].

After the survey [92] was published, we continued our investigation in several directions, and presented the obtained results in a number of papers, for example in [14–17, 45–50, 71, 72, 89–92, 96–98, 123]. However, here we will not repeat the approach from [92], and will not present all those results in detail. Rather, we will briefly mention some of them, and devote the rest of the paper to a particular topic - development of a first-order probabilistic logic and applications of probabilistic logic to default reasoning which was briefly addressed in [92, Section 9.1.4].

We will first mention one of our latest papers: in [45] a classification, in a form of a hierarchy, of probability logics without iterations of probability operators from [92] was given. The considered logics have different languages, and we say that the logic L_2 is more expressive than the logic L_1 iff for every L_1-formula α there is a L_2-formula β such that the classes of all models of α and β coincide.

The central point in the hierarchy is the logic LPP_2 which starts from classical propositional logic with real-valued probability functions and does not allow iterations of the probability operators. The upper part of the hierarchy, in the form of a non-modular lattice with the smallest element consists of probability logics with languages richer than the language of LPP_2 (those languages may contain probability operators capable to express that the probability of formula belongs to a particular set). On the other hand, the lower part of the hierarchy, in the form of an atomic, non-modular lattice, is made of probability logics with probabilistic functions with finite ranges.

The paper [17] compares several syntactic and semantic approaches to measuring inconsistency of a set of propositional formulas T. The syntactic notion of n-consistency is based on the number of formulas needed to derive a contradiction: T is said to be *n-consistent* iff each $T' \subseteq T$ of cardinality n is consistent. On the other hand, T is called *n-probable* if there is a probability measure μ which assigns the probability higher than $1 - \frac{1}{n}$ to every formula of T. The main result related to these notions says that n-probability is stronger than n-consistency. Then, a weaker notion, so-called *local n-probability*, which requires n-probability of $(n+1)$-element subsets of T, is introduced and proved to be equivalent to n-consistency. The paper also provides a detailed comparison of introduced notions with Knight's η-consistency [61].

The introduced notions are generalized to measure inconsistency of theories incorporating beliefs of different agents. This is obtained using the notion of n-probability of T modulo a set of formulas $\{\phi_1, \ldots, \phi_k\}$ with the idea to model situations with k agents, each knowing some fact ϕ_i, and each formula in T is believed to be probable by at least one agent. In other words, there exists a probability measure μ such that $\mu(\phi_1 \wedge \cdots \wedge \phi_k) > 0$ and for all $\psi \in T$, there is an $i \in \{1, \ldots, k\}$ such that

$$\mu(\psi|\phi_i) > 1 - \frac{1}{n}.$$

This approach is reasonable if the agent's beliefs are highly compatible, and thus the additional constraint $\mu(\phi_i|\phi_j) > 1 - \frac{1}{n}$ is required. Under those assumptions, we show that there is the following connection between the conditional and unconditional versions of n-probability: if theory T is n-probable modulo $\{\phi_1, \ldots, \phi_k\}$, then T is $(n-k+1)$-probable.

Next, using non-standard probability measures, and the connection between defaults and non-standard conditional probability infinitesimally close to 1, relations between the presented results and default reasoning were investigated. We introduce the notion of strong n-consistency for given rational relation $\mid\!\sim$ and say that a theory T is strongly n-consistent for $\mid\!\sim$, if for any n formulas $\phi_1, \ldots, \phi_n \in T$, $\not\mid\!\sim \neg(\phi_1 \wedge \cdots \wedge \phi_n)$. The strong n-consistency for $\mid\!\sim$ is shown to be preserved under default derivations.

Another application of the previous results concerns so-called *finite approximations of defaults.* Using the notion "ψ is n-probable modulo $\{\phi\}$," the relations $\mid\!\sim_n$ are defined. It is proved that they satisfy a weak version of System **P** [54] (which is considered as the base for default reasoning and will be discussed below):

$$\text{REF}_n : \frac{}{\alpha \mathrel{|\!\sim}_n \alpha}; \qquad \text{LLE}_n : \frac{\vdash \alpha \leftrightarrow \beta, \alpha \mathrel{|\!\sim}_n \gamma}{\beta \mathrel{|\!\sim}_n \gamma};$$

$$\text{RW}_n : \frac{\vdash \alpha \to \beta, \gamma \mathrel{|\!\sim}_n \alpha}{\gamma \mathrel{|\!\sim}_n \beta}; \qquad \text{AND}_n : \frac{\alpha \mathrel{|\!\sim}_{2n} \beta, \alpha \mathrel{|\!\sim}_{2n} \gamma}{\alpha \mathrel{|\!\sim}_n \beta \wedge \gamma};$$

$$\text{OR}_n : \frac{\alpha \mathrel{|\!\sim}_{2n} \gamma, \beta \mathrel{|\!\sim}_{2n} \gamma}{\alpha \vee \beta \mathrel{|\!\sim}_n \gamma}; \qquad \text{CM}_n : \frac{\alpha \mathrel{|\!\sim}_n \beta, \alpha \mathrel{|\!\sim}_n \gamma}{\alpha \wedge \beta \mathrel{|\!\sim}_{n-1} \gamma}.$$

If $\phi \mathrel{|\!\sim}_n \psi$ means "n is the degree of belief that ψ follows from ϕ", the above rules determine degradation of this degree in each step of a possible deduction.

In [15] a first-order temporal logic for reasoning about branching time was introduced. Since the corresponding set of valid formulas is not recursively enumerable and there is no finitary complete axiomatization, a sound and strongly complete infinitary axiomatization was provided.

Another branching time probability logic with probability operators speaking about probabilities on branches, and about probabilities on sets of branches with the same initial state, was discussed in [89].

A sound, strongly complete and decidable probabilistic temporal logic suitable for reasoning about evidence was provided in [90]. That paper offered a solution for a problem proposed by Halpern and Pucella in [36].

Other formal systems that considered probability and time were presented in [14]. Actually, they involved also reasoning about space so that the systems were suitable to model probabilities that objects were in particular locations in particular time instants.

Probabilistic characterizations for some preferential consequence relations that are obtained by adding/subtracting rules (determinacy preservation, fragmented disjunction, conditional excluding middle) to/from System **P** were given in [16].

The papers [71, 72] considered first order conditional probability logics. Beside infinitary axiom systems which were proved to be sound and strongly complete with respect to the corresponding classes of models, some decidable fragments of the logics were determined.

A number of probability logics with not linearly ordered ranges of the corresponding probabilistic functions (i.e., p-adic, complex or monoid-valued probabilistic functions) were given in [46, 48–51]. In [46], some logics that generalize various proposal for capturing uncertainty in propositional framework are introduced. Uncertainty is modeled by $\mathbb{G}$-valued measures, where $\mathbb{G} = (G, \leqslant, *, 0)$ is a partially (not necessarily linearly) ordered commutative monoid such that the neutral is the least element. If $\mathcal{F}$ is a Boolean algebra of events (considered as classical propositional formulas), then $\mathbb{G}$-valued measure is a function $m : \mathcal{F} \to \mathbb{G}$ satisfying the following conditions: $m(\emptyset) = 0$ and $m(A \cup B) = m(A) * m(B)$ if A and B are disjoint.

Examples of structures suitable to be ranges of $\mathbb{G}$-valued measures are:

(1) Additive monoid of nonnegative rational numbers $(Q^+, \leqslant, +, 0)$;
(2) $(\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}, \leqslant, \oplus, 0)$, $x \oplus y = \min\{1, x + y\}$;

(3) $([0,1]_Q, \leqslant, \oplus, 0)$ ($[0,1]_Q$ is the set of rational numbers from unit real interval $[0,1]$);
(4) $(Q^+(\varepsilon), \leqslant, +, 0)$, where $Q^+(\varepsilon)$ is the set of nonnegative elements of the domain of the nonarchimedean field $Q(\varepsilon)$ which is the smallest field obtained by adding a positive infinitesimal ε to rational numbers;
(5) $(Q^{+n}, \leqslant, +, (0,\dots,0))$, $Q^{+n} = \underbrace{Q^+ \times \cdots \times Q^+}_{n}$, $\leqslant$ is lexicographic order;
(6) $(\omega, \leqslant, +, 0)$;
(7) $(\omega, \leqslant, \max, 0)$;
(8) $(Q^+, \leqslant, \max, 0)$.

Starting with a $\mathbb{G}$-valued measure defined on the set of classical propositional formulas, we extended the language of propositional calculus with a list of modal (unary) quantitative operators $M_{=a}$, $a \in G$, and/or (binary) qualitative operators, e.g. $\leqslant$. The intended meanings of $M_{=a}\varphi$ and $\varphi \leqslant \psi$ are 'the measure of φ is a' and 'the measure of φ is smaller or equal to the measure of ψ', respectively. The semantics is based on Kripke-style models with a set of possible worlds W and a $\mathbb{G}$-measure over the algebra of subsets of W. The main results of the paper concern complete axiomatizations of the proposed logics. Also, the paper describes in detail a general method for studying this kind of logics.

An approach, close to ours, is given in [28], where the authors consider plausibility measures mapping sets from $\mathcal{F}$ to elements of some arbitrary partially ordered set that have no algebraic structures. In contrast, the papers [48–51] discuss the measures whose ranges are the field of complex numbers $\mathbb{C}$ or the fields of p-adic numbers $\mathbb{Q}_p$, for any prime number p. It is well-known that these fields cannot be turned into ordered fields, and one has to work with such probabilities without comparing them.

The p-adic probability theory addressed in [48, 50, 51] is an alternative to the real valued probabilities which gives the opportunity to work, for example, with negative probabilities appearing in different settings in physics, or when one is not able to compare probabilities, as was proposed by J. M. Keynes in A Treatise on Probability.

In our approach, classical propositional logic is extended with operators of the form $K_{r,\rho}$, with the intended meaning "the probability belongs to the p-adic ball with the center r and the radius ρ". Since it is not possible to compare the p-adic probabilities of two events, it is not possible to say that "the probability of α is greater than the probability of β", but we can measure how close are p-adic probabilities. Namely, the p-adic probabilities of two events α and β are close enough, if their p-adic difference is close enough to 0. In the mentioned papers strongly complete axiomatizations for several p-adic logics are provided, and the corresponding decidability theorems are proved. It is also shown how logics of this sort can be used to model processes of thinking.

As it was pointed above, the main part of this paper will be used to present probability logics suitable to model default reasoning. So, the rest of of the article is organized as follows.

In Section 2 we give a propositional background on the connections between default reasoning in System **P** and probabilistic reasoning with approximate conditional probabilities. In Section 3, we describe an extension of ordinary first-order logic which enable us to generalize considerations from the previous section. That logic was introduced in [47]. We use $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ to denote this logic and to emphasize that it is an extension of the ordinary first order classical logic $L_{\omega\omega}$, while P and $\mathbb{I}$ indicate that the values of the probabilistic functions will belong to $\mathbb{I}$. The logic $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ is completely developed: syntax and semantics are specified, and sound and complete deductive system is provided. Section 4 contains proofs of decidability of two fragments of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$. In Section 5 we describe how our system can be used to model default reasoning and analyze some properties of the corresponding default consequence relation. In Section 6 we discuss how one could straightforwardly develop some extensions of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ by adding new (qualitative and/or quantitative) quantifiers which might be very interesting for further applications. We conclude in Section 7.

2. Default reasoning

In this section, we give an overview of the relationship between default reasoning in System **P** and probabilistic reasoning with approximate conditional probabilities. For simplicity, in both cases we assume that the underlying language is the classical propositional language constructed from a set Var of propositional letters (denoted by lower case latters p, q, r, $\dots$) and the usual connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$. The set of classical propositional formulas For_C is defined as usual.

Default reasoning in System P. Roughly speaking, nonmonotonic reasoning is a formalization of reasoning when information is incomplete. If we, human beings, are forced to make a decision under incomplete information, we use commonsense to supplement lack of information. Such reasoning does not satisfy the monotonicity property: we can draw sensible conclusions from what we know, but, faced with new information, we often have to take back previous conclusions, even when the new information we gathered in no way made us want to take back our previous assumptions. For example, we may hold the assumption that most birds fly, but that penguins are birds that do not fly and learning that Tweety is a bird, we infer that it flies. Learning that Tweety is a penguin will in no way make us change our mind about the fact that most birds fly and that penguins are birds that do not fly, or about the fact that Tweety is a bird. It should make us abandon our conclusion about its flying capabilities, though. Thus, as the set of assumptions grows, the set of conclusions (theorems) may shrink. This reasoning is called nonmonotonic in contrast to standard logic, which is monotonic: as one's set of assumptions grows, one's set of theorems grows as well.

A comprehensive treatment of nonmonotonic reasoning systems is far beyond the scope of this article. Various researchers have proposed and studied a large number of nonmonotonic systems which have been mainly suggested by various problems in computer science and AI. We list just a few early ones: Hintikka's theory of various multiple believers [40], Doyle's truth maintenance system [70], Reither's default logic [119], and Moor's autoepistemic logic [73] as well as various versions of negation as failure in extensions of Prolog by Apt, Clark and others [13]. Gabbay [29] was probably the first to suggest to focus the study of nonmonotonic logics on their consequence relations, defined in the style of Gentzen. Today there is an extensive work on desired properties of nonmonotonic consequence relations, but the core of these properties are the postulates proposed by Kraus, Lehmann, and Magidor [54]. System **P** is a deductive system studied in that paper and occupies a central position in the hierarchy of nonmonotonic logical systems ([93]). It turns out that System **P** is a fragment of a conditional logic studied by J. Burgess [12]. Interestingly, this very fragment had been also considered by E. Adams in the paper [1] whose purpose was not to study nonmonotonic logic, but to propose probabilistic semantics for indicative conditionals. Below, we briefly describe System **P**.

If α and β are classical propositional formulas, then the pair (α, β), denoted[1] $\alpha \rightarrowtail \beta$, is called a *default rule* (default in short)[2]. A *default base* is a set $\Delta = \{\alpha_i \rightarrowtail \beta_i \mid i \in I\}$ of default rules. A default base is expected to specify defeasible information. Default reasoning is described in terms of a consequence relation $\mid\!\sim_{\mathbf{P}}$ which determines the set of defaults that are $\mid\!\sim_{\mathbf{P}}$-consequences of a default base. In [54, 62], the relation $\mid\!\sim_{\mathbf{P}}$ is defined by a set of properties, called System **P**, and commonly regarded as a core of default reasoning ($\models$ denotes classical validity):

- REF $\alpha \rightarrowtail \alpha$ (reflexivity);
- LLE if $\models \alpha \leftrightarrow \beta$, from $\alpha \rightarrowtail \gamma$ infer $\beta \rightarrowtail \gamma$; (left logical equivalence)
- RW if $\models \alpha \to \beta$, from $\gamma \rightarrowtail \alpha$ infer $\gamma \rightarrowtail \beta$; (right weakening)
- CUT from $\alpha \wedge \beta \rightarrowtail \gamma$ and $\alpha \rightarrowtail \beta$ infer $\alpha \rightarrowtail \gamma$;
- CM from $\alpha \rightarrowtail \beta$ and $\alpha \rightarrowtail \gamma$ infer $\alpha \wedge \gamma \rightarrowtail \beta$ (caution monotonicity)
- OR from $\alpha \rightarrowtail \gamma$ and $\beta \rightarrowtail \gamma$ infer $\alpha \vee \beta \rightarrowtail \gamma$;

Given a default basis $\Delta = \{\alpha_i \rightarrowtail \beta_i \mid i \in I\}$, the notation $\Delta \mid\!\sim_{\mathbf{P}} \alpha \rightarrowtail \beta$ denotes that $\alpha \rightarrowtail \beta$ can be deduced from Δ using System **P**.

Example 2.1. Let $\Delta = \{b \rightarrowtail f, p \rightarrowtail b, p \rightarrowtail \neg f\}$ where b stands for 'birds', f for 'flies', and p for 'penguin'. Then, $\Delta \mid\!\sim_{\mathbf{P}} b \rightarrowtail \neg p$ can be proved in System **P**.

In practice, in addition to a default base, some knowledge described by classical formulas is often present. A default knowledge base KB contains a finite set of propositional formulas and a finite set of defaults.

Example 2.2. Let us consider the knowledge base $KB = \{b \rightarrowtail f, p \to b, p \to \neg f\}$, where $\to$ is the material implication, and the intended meaning of b, p and f is as in the previous example. Note that $p \to b$, $p \to \neg f$ are classical formulas, and $b \rightarrowtail f$ is a default. It is easy to see that $KB \mid\!\sim_{\mathbf{P}} b \rightarrowtail \neg p$.

---

[1]Note that the other authors use different symbols ($\to$, $\mid\!\sim$, for example) to denote the 'default implication'. In the present setting those symbols may cause confusion, so we prefer to introduce a new symbol here.

[2]Here are some widespread intuitive interpretations of $\alpha \rightarrowtail \beta$: 'from α sensibly conclude β'; 'generally, if α then β (possibly having some exception)'; 'α is a good enough reason to believe β'; 'β is plausible consequence of α'; 'if α, normally β'; 'if α is true, I am willing to (defeasibly) jump to the conclusion that β is true'; etc.

ε-semantics for defaults. Semantics for defaults based on infinitesimal probabilities are discussed at length and shown capable of serving as universal core for a variety of nonmonotonic logics. A probabilistic interpretation of defaults is suggested by Adams [1], and later Pearl [94] and Lehmann and Magidor [62]: a default $\alpha \rightarrowtail \beta$ denotes that 'the probability of β given α is *very close* to 1'. The meaning of *very close* involves probabilities whose range is a non-Archimedean field, i.e., an ordered field containing infinitesimals: an element x is infinitesimal if $|x| < 1/n$, for every positive integer n. Two elements x and y are *infinitely close*, denoted $x \approx y$, if $x - y$ is infinitesimal. Thus, x is infinitesimal iff $x \approx 0$.

Using a proper elementary extension $^*\mathbb{R}$ of the standard real numbers[3], i.e., relying on the fundamental results of Robinson's nonstandard analysis, an $^*\mathbb{R}$-probabilistic model can be defined as a triple $(W, \mathcal{F}, \mu)$, where W is a set of possible words (truth assignments to propositional letters), $\mathcal{F}$ is a field of subsets of W containing all sets definable by propositional formulas, and $\mu : \mathcal{F} \to {}^*\mathbb{R}$ is a finitely additive $^*\mathbb{R}$-valued probability measure. A default $\alpha \rightarrowtail \beta$ holds in an $^*\mathbb{R}$-probabilistic model if:

- either the probability of α is 0,
- or the conditional probability of β given α is infinitely close to 1.

A default base Δ ε-entails a default $\alpha \rightarrowtail \beta$, denoted by $\Delta \vdash_\varepsilon \alpha \rightarrowtail \beta$, if we can ensure that $P(\beta \mid \alpha)$ is almost 1, by taking the probabilities of defaults in Δ to be almost 1.

It turns out that $^*\mathbb{R}$ is a complicated space for many practical purposes and can be replaced with a simpler non-Archimedean field. The search for a minimal non-Archimedean range of probabilities led many authors [37, 39] to consider the ordered field $\mathbb{R}(\varepsilon)$. This field is the smallest field generated by adding to the reals a single infinitesimal[4] ε. $\mathbb{Q}(\varepsilon)$ is another example of non-Archimedean field extending the field of rational numbers.

Hardy field $\mathbb{Q}(\varepsilon)$. Elements of $\mathbb{Q}(\varepsilon)$ are rational expressions of the form $\frac{p(\varepsilon)}{q(\varepsilon)}$, where $p(\varepsilon)$ and $q(\varepsilon)$ are polynomials in ε over $\mathbb{Q}$, and $q(\varepsilon)$ is not identically equal to zero. Two rational expressions $\frac{p(\varepsilon)}{q(\varepsilon)}$ and $\frac{p_1(\varepsilon)}{q_1(\varepsilon)}$ are equal if polynomials $p(\varepsilon)q_1(\varepsilon)$ and $p_1(\varepsilon)q(\varepsilon)$ have the same non-zero coefficients. Addition and multiplication are defined in the usual way. With these definition $\mathbb{Q}(\varepsilon)$ is a field. Each element η of $\mathbb{Q}(\varepsilon)$ can be transformed into the normalized form:

$$(*) \qquad \eta = \frac{a\varepsilon^k + \sum_{i=k+1}^{n} a_i \varepsilon^i}{1 + \sum_{j=1}^{m} b_j \varepsilon^j}, k < n, 0 < m$$

for some unique integer k and some unique leading coefficient a such that $a \neq 0$ unless $\eta = 0$. Define the ordering $<$ on $\mathbb{Q}(\varepsilon)$ so that $\eta > 0$ iff $a > 0$. This makes $\mathbb{Q}(\varepsilon)$ a non-Archimedean ordered field since ε is an infinitesimal. Note that $\mathbb{Q} \subsetneqq \mathbb{Q}(\varepsilon)$. The elements of $\mathbb{Q}(\varepsilon) \setminus \mathbb{Q}$ are called 'non-standard rational numbers'.

[3] Such extension exists by basic results in model theory; for example, we could take $^*\mathbb{R}$ to be the ultrapower of $\mathbb{R}$ with respect to a nonprincipal ultrafilter on $\mathbb{N}$.

[4] One can regard ε as a positive infinitesimal from $^*\mathbb{R}$.

Given the normalized form $(*)$ of η, the unique integer k is the *order* of η, written $\mathrm{ord}(\eta)$, with $\mathrm{ord}(0) = \infty$. If $k = 0$, then η is infinitesimaly different from a non-zero rational number called the *standard part* of η and denoted $\mathrm{st}(\eta) = a$; if $k > 0$, then $\mathrm{st}(\eta) = 0$.

$\mathbb{Q}(\varepsilon)$ has a number of nice properties:

- $\mathbb{Q}(\varepsilon)$ is countable and recursive, i.e., its operations are computable and its ordering is decidable;
- $\mathbb{Q}(\varepsilon)$ is (isomorphic to) a dense subfield of ${}^*\mathbb{R}$;
- the monad (halo) of a rational $q \in [0, 1]$, defined by

$$\mathrm{monad}(x) = \{y \mid y \approx x\},$$

can be characterized by:

$$\mathrm{monad}(q) = \bigcap_{n\in\mathbb{N}^+} \Big[\max\Big\{0, q - \frac{1}{n}\Big\}, \min\Big\{1, q + \frac{1}{n}\Big\}\Big].$$

In the sequel, we focus on the $\mathbb{Q}(\varepsilon)$-probabilistic spaces $(W, \mathcal{F}, \mu)$ with a finitely additive probability measure $\mu : \mathcal{F} \to \mathbb{I}$, where $\mathbb{I}$ is the unit interval of $\mathbb{Q}(\varepsilon)$.

A probabilistic propositional logic with approximate conditional probabilities. In [117], we developed a probabilistic propositional logic, denoted LPP, by adding new binary probabilistic operators to build basic probabilistic formulas of the form $\mathrm{CP}_{\geqslant r}(\alpha \mid \beta)$, $\mathrm{CP}_{\leqslant r}(\alpha \mid \beta)$, $r \in \mathbb{I}$ and $\mathrm{CP}_{\approx q}(\alpha \mid \beta)$, $q \in \mathbb{Q} \cap \mathbb{I}$, where α and β are classical propositional formulas. The intended meaning of these formulas are ‘the conditional probability of α given β is at least r’, ‘at most r’ and ‘approximately q’, respectively. The set For_P of probabilistic formulas is the set of all Boolean combinations of the basic probabilistic formulas. Relying on $\mathbb{Q}(\varepsilon)$-probabilistic spaces, the corresponding strong completeness theorem is proved for the proposed axiomatic system denoted Ax_{LPP}. In that logic, a default $\alpha \rightarrowtail \beta$ is represented by the formula $\mathrm{CP}_{\approx 1}(\beta \mid \alpha)$. Of course, a finite default base $\Delta = \{\alpha_1 \rightarrowtail \beta_1, \ldots, \alpha_m \rightarrowtail \beta_m\}$ is represented by the set $\overline{\Delta} = \{\mathrm{CP}_{\approx 1}(\beta_1 \mid \alpha_1), \ldots, \mathrm{CP}_{\approx 1}(\beta_1 \mid \alpha_1)\}$. It is proved that the proposed approach gives a characterization of System **P**.

Theorem 2.1. *For every finite default base Δ and for every default $\alpha \rightarrowtail \beta$:*

$$\Delta \mathrel{|\!\!\sim_{\mathbf{P}}} \alpha \rightarrowtail \beta \textit{ iff } \Delta \vdash_{\mathrm{Ax}_{LPP}} \alpha \rightarrowtail \beta.$$

A finite default base $\Delta = \{\alpha_1 \rightarrowtail \beta_1, \ldots, \alpha_m \rightarrowtail \beta_m\}$ can be regarded as the formula $\bigwedge \overline{\Delta} = \bigwedge_{i=1}^{m} \mathrm{CP}_{\approx 1}(\beta_i \mid \alpha_i)$. Using Deduction and Completeness theorems for LPP [117], we have:

$$\begin{aligned}
&\Delta \mathrel{|\!\!\sim_{\mathbf{P}}} \alpha \rightarrowtail \beta\\
\text{iff}\quad &\vdash_{\mathrm{Ax}_{LPP}} \textstyle\bigwedge \overline{\Delta} \Rightarrow \mathrm{CP}_{\approx 1}(\beta \mid \alpha)\\
\text{iff}\quad &\textstyle\bigwedge \overline{\Delta} \Rightarrow \mathrm{CP}_{\approx 1}(\beta \mid \alpha) \text{ is valid}\\
&\qquad\qquad \text{with respect to the class of } \mathbb{Q}(\varepsilon)\text{-probabilistic spaces}\\
\text{iff}\quad &\neg\big(\textstyle\bigwedge \overline{\Delta} \Rightarrow \mathrm{CP}_{\approx 1}(\beta \mid \alpha)\big) \text{ is not satisfied in any } \mathbb{Q}(\varepsilon)\text{-probabilistic space}\\
\text{iff}\quad &\textstyle\bigwedge \overline{\Delta} \wedge \neg\, \mathrm{CP}_{\approx 1}(\beta \mid \alpha) \text{ is not satisfied in any } \mathbb{Q}(\varepsilon)\text{-probabilistic space}
\end{aligned}$$

For applications, the following result is very important.

Theorem 2.2. *The satisfiability problem for LPP-formulas is decidable.*

The proof of the theorem is rather long and it is based on a reduction of satisfiability problem to linear programming problem. We illustrate the method on formulas of the form

$$\delta = \bigwedge_{i=1}^{m} \mathrm{CP}_{\approx 1}(\beta_i \mid \alpha_i) \wedge \neg\, \mathrm{CP}_{\approx 1}(\beta \mid \alpha).$$

For these considerations the set Var is finite, for definiteness let $\mathrm{Var} = \{p_1, \ldots, p_n\}$.

For $p \in \mathrm{Var}$, let $p^1 = p$ and $p^0 = \neg p$. Let $\xi_1, \ldots, \xi_N$, where $N = 2^n$, run through the 2^n classical formulas of the form $p_1^{e_1} \wedge \cdots \wedge p_n^{e_n}$, where $e_1, \ldots, e_n \in \{0, 1\}$. We call ξ_i's atoms, and denote the set of these atoms by At. It is obvious that for a given atom ξ there is a unique valuation $v_\xi : \mathrm{Var} \to \{0, 1\}$ such that $v_\xi(\xi) = 1$, and vice versa. Moreover, for each $\alpha \in For_C$, there is $S_\alpha \subseteq \mathrm{At}$ such that α is classically equivalent to $\bigvee S_\alpha$:

$$S_\alpha = \{\xi \in \mathrm{At} \mid \xi \models \alpha\} = \{\xi \in \mathrm{At} \mid v_\xi(\alpha) = 1\}.$$

Thus, when the set of propositional letters is finite, we can only consider $\mathbb{Q}(\varepsilon)$-probabilistic spaces of the form $(\mathrm{At}, \mathcal{P}(\mathrm{At}), P)$, where $P : \mathcal{P}(\mathrm{At}) \to \mathbb{I}$ is a finitely additive probability measure. Each probability measure of this form is completely determined by its values on the atoms, i.e., by the vector

$$(P(\xi_1), \ldots, P(\xi_N)) \in \mathbb{D} = \left\{ \vec{x} \in \mathbb{Q}(\varepsilon)^N \,\middle|\, x_i \geqslant 0, 1 \leqslant i \leqslant N, \sum_{i=1}^{N} x_i = 1 \right\}.$$

Conversely, given $\vec{x} \in \mathbb{D}$ determines a unique probability measure P' satisfying

$$(P'(\xi_1), \ldots, P'(\xi_N)) = \vec{x}.$$

The function $P' : For_C \to \mathbb{I}$ is defined by $P'(\alpha) = \sum_{\xi_i \in S_\alpha} x_i$. This 1-1 correspondence between probability measures and points in $\mathbb{D}$ is very useful in what follows. Namely, $\delta = \bigwedge_{i=1}^{m} \mathrm{CP}_{\approx 1}(\beta_i \mid \alpha_i) \wedge \neg\, \mathrm{CP}_{\approx 1}(\beta \mid \alpha)$ is satisfiable iff the system

$$(*) \qquad \begin{cases} \sum_{i=1}^{N} x_i = 1, \\ x_1 \geqslant 0, \ldots, x_N \geqslant 0, \\ \sum_{\xi_i \in S_{\beta_1}} x_i > 0, \ldots, \sum_{\xi_i \in S_{\beta_m}} x_i > 0, \sum_{\xi_i \in S_\beta} x_i > 0, \\ \dfrac{\sum_{\xi_i \in S_{\alpha_1 \wedge \beta_1}} x_i}{\sum_{\xi_i \in S_{\beta_1}} x_i} \approx 1, \ldots, \dfrac{\sum_{\xi_i \in S_{\alpha_m \wedge \beta_m}} x_i}{\sum_{\xi_i \in S_{\beta_m}} x_i} \approx 1, \\ \dfrac{\sum_{\xi_i \in S_{\alpha \wedge \beta}} x_i}{\sum_{\xi_i \in S_\beta} x_i} \not\approx 1, \end{cases}$$

has a solution in $\mathbb{Q}(\varepsilon)$.

As shown in the paper [117], the system $(*)$ has a solution in $\mathbb{Q}(\varepsilon)$ iff the system

$$(\star)\qquad \begin{cases} \sum_{i=1}^{N} x_i = 1, \\ x_1 \geqslant 0, \dots, x_N \geqslant 0, \\ \sum_{\xi_i \in S_{\beta_1}} x_i > 0, \dots, \sum_{\xi_i \in S_{\beta_m}} x_i > 0, \sum_{\xi_i \in S_{\beta}} x_i > 0, \\ 0 \leqslant 1 - \frac{\sum_{\xi_i \in S_{\alpha_1 \wedge \beta_1}} x_i}{\sum_{\xi_i \in S_{\beta_1}} x_i} < K\varepsilon, \dots, 0 \leqslant 1 - \frac{\sum_{\xi_i \in S_{\alpha_m \wedge \beta_m}} x_i}{\sum_{\xi_i \in S_{\beta_m}} x_i} < K\varepsilon, \\ 1 - \frac{\sum_{\xi_i \in S_{\alpha \wedge \beta}} x_i}{\sum_{\xi_i \in S_{\beta}} x_i} > \frac{1}{K}, \end{cases}$$

has a solution in ${}^*\mathbb{R}$, where ε and K are suitably chosen elements of ${}^*\mathbb{R}$: $\varepsilon > 0$ is an infinitesimal, $K > 0$ is infinitely large and $K^k \cdot \varepsilon \approx 0$, for every $k \in \mathbb{N}$.[5] More specific, we solve the system $(\star)$ in an ordered field whose elements are rational expressions in ε and K with rational coefficients. We add and multiply in this field in the usual way. The ordering is generated by the suitable ordering on the set of polynomials in ε and K. Note that each polynomial in $Q(\varepsilon, K)$ can be expressed in the form $Q_0(K)\varepsilon^0 + Q_1(K)\varepsilon^1 + \dots + Q_n(K)\varepsilon^n$, where $Q_i(K)$'s are polynomials in K with rational coefficients. Comparison of polynomials $Q_1(\varepsilon, K)$ and $Q_2(\varepsilon, K)$ is carried out as follows:

$$\begin{aligned} & Q_1(\varepsilon, K) < Q_2(\varepsilon, K) \\ \text{iff}\quad & Q_{1,0}(K)\varepsilon^0 + \dots + Q_{1,n_1}(K)\varepsilon^{n_1} < Q_{2,0}(K)\varepsilon^0 + \dots + Q_{2,n_2}(K)\varepsilon^{n_2} \\ \text{iff}\quad & Q_{1,i}(K) < Q_{2,i}(K), \text{ for some } i \leqslant \max\{n_1, n_2\}, \text{ and} \\ & Q_{1,j}(K) = Q_{2,j}(K), j < i\, (Q_{1,i}(K) = 0, i > n_1 \text{ and } Q_{2,i}(K) = 0, i > n_2) \end{aligned}$$

and

$$\begin{aligned} & Q_{1,i}(K) < Q_{2,i}(K) \\ \text{iff}\quad & q_{1,i,m_1} K^{m_1} + \dots + q_{1,i,1} K + q_{1,i,0} < q_{1,i,m_2} K^{m_2} + \dots + q_{2,i,1} K + q_{2,i,0} \\ \text{iff}\quad & q_{1,i,r} < q_{2,i,r} \text{ for some } r \leqslant \max\{m_1, m_2\} \text{ and } q_{1,i,t} = q_{2,i,t} \text{ for } t < r \\ & (q_{1,i,r} = 0 \text{ for } r > m_1 \text{ and } q_{2,i,r} = 0 \text{ for } r > m_2) \end{aligned}$$

For example,

$$\Big(\frac{1}{2}K^2 + K\Big) + \Big(\frac{1}{3}K^3 + \frac{1}{2}K^2 + K\Big)\varepsilon + \varepsilon^2 < \Big(\frac{1}{2}K^2 + K\Big) + \Big(\frac{1}{3}K^3 + \frac{2}{3}K^2 + 1\Big)\varepsilon$$

since $\frac{1}{2}K^2 + K = \frac{1}{2}K^2 + K$ ($\frac{1}{2} = \frac{1}{2}$, $1 = 1$) and $\frac{1}{3}K^3 + \frac{1}{2}K^2 + K < \frac{1}{3}K^3 + \frac{2}{3}K^2 + 1$ ($\frac{1}{3} = \frac{1}{3}$, $\frac{1}{2} < \frac{2}{3}$).

If $Q_1'(\varepsilon, K) > 0$ and $Q_2'(\varepsilon, K) > 0$, then

$$\frac{Q_1(\varepsilon, K)}{Q_1'(\varepsilon, K)} < \frac{Q_2(\varepsilon, K)}{Q_2'(\varepsilon, K)} \text{ iff } Q_1(\varepsilon, K) \cdot Q_2'(\varepsilon, K) < Q_2(\varepsilon, K) \cdot Q_1'(\varepsilon, K).$$

---

[5] For example, if we consider the usual construction of a set of hyperreals ${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}}/\mathcal{U}$, where $\mathcal{U}$ is a nonprincipal ultrafilter on $\mathbb{N}$, then we can choose ε and K to be the following $\mathcal{U}$-equivalence classes: $\varepsilon = \langle 1, \frac{1}{2^2}, \frac{1}{3^3}, \dots, \frac{1}{n^n}, \dots \rangle$ and $K = \langle 1, 2, 3, \dots, n, \dots \rangle$.

It is easy to see that the system ($\star$) is equivalent to a system of linear inequalities of the form

$$s_1 x_{i_1} + \cdots + s_\ell x_{i_\ell} \ \rho\ 0,$$

where $\ell \leqslant N$, $1 \leqslant i_1, \ldots, i_\ell \leqslant N$, s_j's are rational functions in ε and K, and $\rho \in \{\leqslant, <, >, \geqslant\}$. Now, we can perform Fourier–Motzkin elimination, which iteratively rewrites the starting system into a new system without a variable x_i such that two systems are equisatisfiable. During the procedure, numerators and denominators of coefficients in inequalities remain polynomials in ϵ and K. When no variables are left, we have to verify relations between polynomials in ϵ and K.

Example 2.3. In Example 2.1, we prove $\{b \rightarrowtail f, p \rightarrowtail b, p \rightarrowtail \neg f\} \mid\!\sim_{\mathbf{P}} b \rightarrowtail \neg p$. Now we will prove this in a different way in order to illustrate the method described above.

There are eight atoms that should be considered:

$$\begin{array}{llll} \xi_1 = b \wedge p \wedge f & \xi_2 = b \wedge p \wedge \neg f & \xi_3 = b \wedge \neg p \wedge f & \xi_4 = b \wedge \neg p \wedge \neg f \\ \xi_5 = \neg b \wedge p \wedge f & \xi_6 = \neg b \wedge p \wedge \neg f & \xi_7 = \neg b \wedge \neg p \wedge f & \xi_8 = \neg b \wedge \neg p \wedge \neg f \end{array}$$

We have to prove that the formula

$$\mathrm{CP}_{\approx 1}(f \mid b) \wedge \mathrm{CP}_{\approx 1}(b \mid p) \wedge \mathrm{CP}_{\approx 1}(\neg f \mid p) \wedge \mathrm{CP}_{\not\approx 1}(\neg p \mid b)$$

is not satisfied in any $\mathbb{Q}(\varepsilon)$-probabilistic space, i.e., that the system:

$$(*) \qquad \begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1, \\ x_1 \geqslant 0,\ x_2 \geqslant 0,\ x_3 \geqslant 0,\ x_4 \geqslant 0,\ x_5 \geqslant 0,\ x_6 \geqslant 0,\ x_7 \geqslant 0,\ x_8 \geqslant 0, \\ x_1 + x_2 + x_3 + x_4 > 0,\ x_1 + x_2 + x_5 + x_6 > 0, \\ \frac{x_1+x_3}{x_1+x_2+x_3+x_4} \approx 1,\ \frac{x_1+x_2}{x_1+x_2+x_5+x_6} \approx 1,\ \frac{x_2+x_6}{x_1+x_2+x_5+x_6} \approx 1, \\ \frac{x_3+x_4}{x_1+x_2+x_3+x_4} \not\approx 1, \end{cases}$$

has no solution in $\mathbb{Q}(\varepsilon)$. The first step of the described procedure is the elimination of the sign $\approx$, i.e., the transformation of the system into the following one:

$$(\star) \qquad \begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1, \\ x_1 \geqslant 0,\ x_2 \geqslant 0,\ x_3 \geqslant 0,\ x_4 \geqslant 0,\ x_5 \geqslant 0,\ x_6 \geqslant 0,\ x_7 \geqslant 0,\ x_8 \geqslant 0, \\ x_1 + x_2 + x_3 + x_4 > 0,\ x_1 + x_2 + x_5 + x_6 > 0, \\ 1 - \frac{x_1+x_3}{x_1+x_2+x_3+x_4} \geqslant 0,\ 1 - \frac{x_1+x_2}{x_1+x_2+x_5+x_6} \geqslant 0, 1 - \frac{x_2+x_6}{x_1+x_2+x_5+x_6} \geqslant 0, \\ 1 - \frac{x_1+x_3}{x_1+x_2+x_3+x_4} < K\varepsilon,\ 1 - \frac{x_1+x_2}{x_1+x_2+x_5+x_6} < K\varepsilon,\ 1 - \frac{x_x+x_6}{x_1+x_2+x_5+x_6} < K\varepsilon, \\ 1 - \frac{x_3+x_4}{x_1+x_2+x_3+x_4} > \frac{1}{K}, \end{cases}$$

which is equivalent to the system of linear inequalities:

$$(\star)\quad \begin{cases} x_1+x_2+x_3+x_4+x_5+x_6+x_7+x_8=1,\\ x_1\geqslant 0,\ x_2\geqslant 0,\ x_3\geqslant 0,\ x_4\geqslant 0,\ x_5\geqslant 0,\ x_6\geqslant 0,\ x_7\geqslant 0,\ x_8\geqslant 0,\\ x_1+x_2+x_3+x_4>0,\ x_1+x_2+x_5+x_6>0,\\ x_2+x_4\geqslant 0,\ x_5+x_6\geqslant 0,\ x_1+x_5\geqslant 0,\\ x_2+x_4<K\varepsilon(x_1+x_2+x_3+x_4),\\ x_5+x_6<K\varepsilon(x_1+x_2+x_5+x_6),\\ x_1+x_5<K\varepsilon(x_1+x_2+x_5+x_6),\\ x_1+x_2>\frac{1}{K}(x_1+x_2+x_3+x_4). \end{cases}$$

Fourier–Motzkin elimination can be performed in the standard way, and the procedure will finish with a false condition.

3. Probabilistic first-order logic $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$

In this section, we describe an extension of ordinary first-order logic, introduced in [47], which will enable us to generalize considerations from the previous section. The language of the logic is rich enough so we can represent in it simultaneously statistical knowledge, imprecise probabilities, beliefs and defaults. We follow the standard steps in the development of first-order logic.

Syntax. Let L be a first-order language, consisting of predicate and function symbols. As usual, constant symbols are 0-ary function symbols. The logic $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ is an extension of $L_{\omega\omega}$, where $L_{\omega\omega}$ denotes the classical first-order logic. In addition to logical symbols of $L_{\omega\omega}$, $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ has new probabilistic quantifiers $(\mathrm{CP}\ \vec{x}\geqslant r)$ and $(\mathrm{CP}\ \vec{x}\leqslant r)$, for every $r\in\mathbb{I}$, and $(\mathrm{CP}\ \vec{x}\approx q)$, for every $q\in\mathbb{I}\cap\mathbb{Q}$, where $\vec{x}=(x_1,\dots,x_n)$ is a tuple of pairwise distinct variables.

The set of terms and the set of atomic formulas of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ are the same as in $L_{\omega\omega}$.

Definition 3.1. The set of formulas of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ is the least set such that:

- each atomic formula of first-order logic is a formula of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$;
- if α is a formula of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$, then $\neg\alpha$ is a formula of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$;
- if α and β are formulas of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$, then so is $\alpha\wedge\beta$;
- if α is a formula of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$, and x is a variable, then $\forall x\alpha$ is a formula of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$;
- if α and β are formulas of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$, and $(\mathrm{CP}\ \vec{x}\diamond r)$ is a quantifier of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ ($\diamond$ is a placeholder for $\leqslant,\geqslant,\approx$), then $(\mathrm{CP}\,\vec{x}\diamond r)(\alpha\mid\beta)$ is a formula of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$.

The set of formulas is countable and recursive, since the field $\mathbb{Q}(\varepsilon)$ is countable and recursive.

The connectives $\vee$, $\rightarrow$, and $\leftrightarrow$, and the quantifier $\exists$ are defined as usual. We abbreviate $\alpha\wedge\neg\alpha$ by $\bot$, and $\neg\bot$ by $\top$. Also, it is convenient to use the following abbreviations in $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$:

- $(\mathrm{CP}\,\vec{x}<r)(\alpha\mid\beta)$ for $\neg(\mathrm{CP}\,\vec{x}\geqslant r)(\alpha\mid\beta)$;
- $(\mathrm{CP}\,\vec{x}>r)(\alpha\mid\beta)$ for $\neg(\mathrm{CP}\,\vec{x}\leqslant r)(\alpha\mid\beta)$;
- $(\mathrm{CP}\,\vec{x}=r)(\alpha\mid\beta)$ for $(\mathrm{CP}\,\vec{x}\geqslant r)(\alpha\mid\beta)\wedge(\mathrm{CP}\,\vec{x}\leqslant r)(\alpha\mid\beta)$;
- $(\mathrm{P}\vec{x}\diamond r)\alpha$ for $(\mathrm{CP}\,\vec{x}\diamond r)(\alpha\mid\top)$, $\diamond\in\{\leqslant,=,\geqslant,\approx,<,>\}$.

Note that we allow interleaving of first order quantifiers and probability quantifiers.

The notions of free and bound variables are defined as usual, with the quantifier $(\mathrm{CP}\ \vec{x} \diamond r)$ binding all the variables in the tuple $\vec{x}$, $\diamond \in \{\leqslant, \geqslant, \approx\}$. If $\vec{x}$ is a tuple of pairwise distinct variables that contains all the free variables of α, we write $\alpha(\vec{x})$. A *sentence* is a formula having no free variables. A *theory* is a set of sentences.

Example 3.1. Let $L = \{U, B\}$, where U and B are relation symbols, $\mathrm{ar}(U) = 1$ and $\mathrm{ar}(B) = 2$. The following strings are $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-formulas:

- $(\mathrm{CP}\, y \approx 1)(B(x,y) \to U(y) \mid B(x,y))$ (x occurs free in this formula);
- $(\mathrm{CP}\, xy = \varepsilon)(\exists z(B(x,z) \wedge B(z,y)) \mid x = y)$ (this formula is a sentence);
- $(\mathrm{P}z \approx 1)(\mathrm{CP}\, y = 1)(B(y,z) \vee \neg B(z,z) \mid (\mathrm{CP}\, u \geqslant 0.99)(U(u) \mid B(z,u)))$, etc.

The rules for renaming quantified variables in formulas expand the standard first order rules. Any formula with bound variables has a number of variants obtained by renaming of its bound variables. Note that if α' is a variant of α, then α and α' have the same free variables. As we will see later, variants preserve the meaning of the original formula. Now, we can specify how a term t may be substituted for a variable x in any formula α: transform α to a variant α' which does not have any variable in common with t, and then substitute t for all free occurrences of x in α'. The new formula which is formed by the substitution process is denoted by $\alpha(x := t)$. In a similar manner, we specify the procedure for simultaneously replacing several variables: with a given formula α, pairwise distinct variables $\vec{x} = (x_1, \dots, x_n)$ and arbitrary terms $\vec{t} = (t_1, \dots, t_n)$, we form a formula $\alpha(\vec{x} := \vec{t})$, which is obtained by simultaneously substituting $t_1, \dots, t_n$ for $x_1, \dots, x_n$ in a suitable variant of α.

Semantics. The structures to be considered are of the form $\langle \mathfrak{A}, \mathcal{F}_n, \mu_n \rangle_{n \in \mathbb{N}}$, where μ_n is a finitely-additive probability on the n-fold product of the domain of a first order structure $\mathfrak{A}$.

Definition 3.2. A model for $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ is a structure $\overline{\mathfrak{A}} = \langle \mathfrak{A}, \mathcal{F}_n, \mu_n \rangle_{n \in \mathbb{N}}$ such that

- $\mathfrak{A} = (A, S^{\mathfrak{A}})_{S \in L}$ is a classical $L_{\omega\omega}$-model for L;
- for all $n \geqslant 1$, $(A^n, \mathcal{F}_n, \mu_n)$ is a finitely additive probability space, where $\mathcal{F}_n$ is a field of subsets of A^n and μ_n is a finitely additive probability measure whose domain is $\mathcal{F}_n$, and whose range is $\mathbb{I}$ (i.e., $\mu_n(X) \geqslant 0$, $\mu_n(X \cup Y) = \mu_n(X) + \mu_n(Y)$ if $X \cap Y = \emptyset$, and $\mu_n(A^n) = 1$); moreover,
  - for each n-ary function symbol f of L, the graph of $f^{\mathfrak{A}}$ is in $\mathcal{F}_{n+1}$;
  - for each n-ary relation symbol R of L, $R^{\mathfrak{A}} \in \mathcal{F}_n$;
  - for all $i, j \leqslant n$, $\{(x_1, \dots, x_n) \in A^n \mid x_i = x_j\} \in \mathcal{F}_n$;
  - if $X \in \mathcal{F}_n$, then $A \times X \in \mathcal{F}_{n+1}$;
  - if $X \in \mathcal{F}_{n+1}$ and $\Pi : A^{n+1} \to A^n$ is the projection map
  $$\Pi(x_1, \dots, x_n, x_{n+1}) = (x_1, \dots, x_n),$$
  then $\Pi(X) \in \mathcal{F}_n$;
  - if $X \in \mathcal{F}_{n+m}$ and $\vec{b} \in A^m$, then $\{\vec{a} \in A^n \mid (\vec{a}, \vec{b}) \in X\} \in \mathcal{F}_n$;

– if $X \in \mathcal{F}_{n+m}$, then $\{\vec{a} \in A^n \mid \mu_m\{\vec{b} \in A^m \mid (\vec{a}, \vec{b}) \in X\} \diamond r\} \in \mathcal{F}_n$, where $\diamond \in \{\leqslant, \geqslant\}$ and $r \in \mathbb{I}$, or $\diamond = \approx$ and $r \in \mathbb{I} \cap \mathbb{Q}$;
– each μ_n is invariant under permutation: for every permutation π of $\{1, 2, \dots, n\}$ and $X \in \mathcal{F}_n$, if

$$X^\pi = \{(a_{\pi(1)}, \dots, a_{\pi(n)}) \in A^n \mid (a_1, \dots, a_n) \in X\},$$

then $X^\pi \in \mathcal{F}_n$ and $\mu_n(X^\pi) = \mu_n(X)$;
– if $X \in \mathcal{F}_n$, then $\mu_{n+1}(A \times X) = \mu_n(X)$.

Note that we do not assume that the sequence of probabilities $(\mu_n\ :\ n = 1, 2, \dots)$ is a sequence of product measures. More specific, instead of the conditions:

(i) $\mathcal{F}_{m+n}$ is generated by the set $\{X \times Y \mid X \in \mathcal{F}_m, Y \in \mathcal{F}_n\}$, and
(ii) the measure μ_{m+n} is the product probability measure of μ_m and μ_n,

$$\mu_{m+n}(X \times Y) = \mu_m(X) \cdot \mu_n(Y).$$

we take their weakening:

(1) if $X \in \mathcal{F}_n$, then $A \times X \in \mathcal{F}_{n+1}$,
(2) if $X \in \mathcal{F}_n$, then $\mu_{n+1}(A \times X) = \mu_n(X)$.

Nevertheless, we can restrict the class of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-models to those structures whose sequence of probabilities is a sequence of product measures and give a complete axiomatization for them (Theorem 3.4).

Example 3.2. The weak conditions (1) and (2) are satisfied in the trivial spaces $(A^n, \{\emptyset, A^n\}, \mu_n)$, where $\mu_n(A^n) = 1$, $\mu_n(\emptyset) = 0$. Note also that for any n, μ_n is invariant under permutations: for any permutation $\pi : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$, $(A^n)^\pi = \{(a_{\pi(1)}, \dots, a_{\pi(n)}) \mid (a_1, \dots, a_n) \in A^n\} = A^n$ and $\emptyset^\pi = \emptyset$.

Here are more illustrative examples of probability spaces over the finite set $A = \{1, 2, 3, 4\}$. Let us consider the space $(A, \mathcal{P}(A), \mu_1)$, where $\mu_1\ : \mathcal{P}(A) \to [0, 1]$ is probability measure defined by a nonuniform distribution on the singletons:

$$\mu_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \varepsilon & \frac{1-4\varepsilon}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}, \ \mu_1(X) = \sum_{x \in X} \mu_1(\{x\}), X \subseteq A.$$

There are many possibilities to define a probability measure on $\mathcal{P}(A^2)$ that the requirements (1) and (2) are fulfilled. Here are two examples defined by distributions on singletons:

μ_2	1	2	3	4
1	ε	0	0	0
2	0	$\frac{1-4\varepsilon}{4}$	0	0
3	0	0	$\frac{1}{4}$	0
4	0	0	0	$\frac{1}{2}$

and

μ'_2	1	2	3	4
1	ε^2	$\frac{\varepsilon-\varepsilon^2}{4}$	$\frac{\varepsilon}{4}$	$\frac{\varepsilon}{2}$
2	$\frac{\varepsilon-\varepsilon^2}{4}$	$\frac{(1-\varepsilon)^2}{16}$	$\frac{1-\varepsilon}{16}$	$\frac{1-\varepsilon}{8}$
3	$\frac{\varepsilon}{4}$	$\frac{(1-\varepsilon)^2}{16}$	$\frac{1}{16}$	$\frac{1}{8}$
4	$\frac{\varepsilon}{2}$	$\frac{1-\varepsilon}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

Note that μ'_2 is the product probability measure: $\mu'_2(X \times Y) = \mu_1(X) \cdot \mu_1(Y)$, $X, Y \subseteq A$. Analogously, we can define $(A^n, \mathcal{P}(A^n),\ \mu_n)$ and $(A^n, \mathcal{P}(A^n),\ \mu'_n)$, for $n \geqslant 3$.

An assignment in $\mathfrak{A}$ is a function v whose domain is the set of variables Var and the range is a subset of A. We use $v(\vec{x} := \vec{a})$, where $\vec{x} = (x_1, \dots, x_n)$ is a tuple of pairwise distinct variables and $\vec{a} \in A^n$, for the assignment v' which is the same as v except that $v'(x_i) = a_i$, $i = 1, k$.

If $\overline{\mathfrak{A}}$ is an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model, then every assignment $v : \mathrm{Var} \to A$ is extended to the set of all terms of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ (i.e., of $L_{\omega\omega}$) in the standard recursive way. For a term t of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$, we denote by $t^{\overline{\mathfrak{A}},v}$ the associated element from the domain A.

Definition 3.3. Let $\overline{\mathfrak{A}}$ be an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model and v an assignment. We define the relation $\overline{\mathfrak{A}}, v \models \varphi$ (read: the assignment v satisfies the formula φ in $\overline{\mathfrak{A}}$) for all assignments $v : \mathrm{Var} \to A$ and all formulas in the same way as it is defined for $L_{\omega\omega}$ with the new clauses:

- $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \leqslant r)(\alpha \mid \beta)$ iff $\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\} = 0$ and $r = 1$ or

$\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\} > 0$ and $\dfrac{\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \alpha \wedge \beta\}}{\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\}} \leqslant r$;

- $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \geqslant r)(\alpha \mid \beta)$ iff $\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\} = 0$ or

$\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\} > 0$ and $\dfrac{\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \alpha \wedge \beta\}}{\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\}} \geqslant r$;

- $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$ iff $\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\} = 0$ and $q = 1$or

$\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\} > 0$ and

$$\frac{\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \alpha \wedge \beta\}}{\mu_n\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \beta\}} \in \mathrm{monad}(q).$$

If Γ is a set of formulas, we write $\overline{\mathfrak{A}}, v \models \Gamma$ to mean that $\overline{\mathfrak{A}}, v \models \gamma$ for every formula γ in Γ.

The clause for $(\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$, in the above definition, means that the conditional probability equals either $q - \varepsilon'$ or $q + \varepsilon'$ for some infinitesimal $\varepsilon' \in \mathbb{I}$. Note that all clauses are formulated on the assumption that the conditional probability is 1, whenever the condition has the probability 0. If we had kept the standard definition according to which conditional probability is undefined when the condition has probability 0, then many formulas would not have truth-values, i.e., for some formulas the question whether they are satisfied in a given model would make no sense. Moreover, this approach is fully in line with probabilistic interpretation of defaults.

As in first order logic, the following technical lemma is important.

Lemma 3.1. *Let $\overline{\mathfrak{A}}$ be an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model, and α an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-formula.*

1. *If u and v are assignments in $\overline{\mathfrak{A}}$ which agree on each free variable of α, then*

$$\overline{\mathfrak{A}}, u \models \alpha \ \textit{ iff } \ \overline{\mathfrak{A}}, v \models \alpha.$$

2. *If α' is a variant of α then for every assignment $v : \mathrm{Var} \to A$,*

$$\overline{\mathfrak{A}}, v \models \alpha \ \textit{ iff } \ \overline{\mathfrak{A}}, v \models \alpha'.$$

3. *If t is a term, and x a variable, then for every assignment* $v : \mathrm{Var} \to A$

$$\overline{\mathfrak{A}}, v \models \alpha(x := t) \ \textit{ iff } \ \overline{\mathfrak{A}}, v(x := t^{\overline{\mathfrak{A}},v}) \models \alpha.$$

We omit the proofs since they are similar to their counterparts for $L_{\omega\omega}$.

The first statement of the previous lemma says that only free variables of formulas might alter their meanings. Thus, if $\alpha(\vec{x})$ is a formula, and $\vec{a} \in A^n$, then for every two assignments u and v: $\overline{\mathfrak{A}}, u(\vec{x} := \vec{a}) \models \alpha$ iff $\overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \alpha$; this fact allows us to adopt the common practice and write $\overline{\mathfrak{A}}, \vec{a} \models \alpha(\vec{x})$ or $\overline{\mathfrak{A}} \models \alpha[\vec{a}]$. In particular, if α is a sentence, we write $\overline{\mathfrak{A}} \models \alpha$, since the truth value of a sentence depends only on the underlying structure $\overline{\mathfrak{A}}$.

The second statement means that the renaming of quantified variables in formulas preserves the meaning.

The third statement shows that substitutions in $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$ behave semantically in the same way as in the first order logic.

If $\overline{\mathfrak{A}}$ is an $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-model, $v : \mathrm{Var} \to A$ an assignment, $\vec{x}$ a tuple of distinct variables, and $\varphi(\vec{x})$ an $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-formula, we write $[\varphi]^{\vec{x}}_{\overline{\mathfrak{A}},v}$ for the set $\{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v(\vec{x} := \vec{a}) \models \varphi\}$.

Example 3.3. The language $L = \{U, B\}$, from Example (3.1), is interpreted on $A = \{1, 2, 3, 4\}$: $U^{\mathfrak{A}} = \{2, 3\}$ and $R^{\mathfrak{A}} = \{(1, 1), (1, 3), (2, 4)\}$. We check whether some sentences are satisfiable in $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-models

$$\overline{\mathfrak{A}} = (A, U^{\mathfrak{A}}, R^{\mathfrak{A}}, \mathcal{P}(A^n), \mu_n)_{n \geqslant 1} \ \text{ and } \ \overline{\mathfrak{A}}' = (A, U^{\mathfrak{A}}, R^{\mathfrak{A}}, \mathcal{P}(A^n), \mu'_n)_{n \geqslant 1},$$

where μ_n and μ'_n are measures defined in Example 3.2. It is not hard to prove that:

- $\overline{\mathfrak{A}} \models \exists x(\mathrm{CP}\, y \approx 1)(B(x, y) \to U(y) \mid B(x, y))$, and
  $\overline{\mathfrak{A}}' \models \exists x(\mathrm{CP}\, y \approx 1)(B(x, y) \to U(y) \mid B(x, y))$;
- $\overline{\mathfrak{A}} \models (\mathrm{CP}\, xy = \varepsilon)(\exists z(B(x, z) \wedge B(z, y)) \mid x = y)$, and
  $\overline{\mathfrak{A}}' \not\models (\mathrm{CP}\, xy = \varepsilon)(\exists z(B(x, z) \wedge B(z, y)) \mid x = y)$;
- $\overline{\mathfrak{A}} \models (\mathrm{P}z \approx 1)(\mathrm{CP}\, y = 1)(B(y, z) \vee \neg B(z, z) \mid (\mathrm{CP}\, u \geqslant 0.9)(U(u) \mid B(z, u)))$.

Example 3.4. Now, we present some examples of what can be represented by $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-formulas choosing the suitable language.

- 'Most birds fly': $(\mathrm{CP}\, x > 0.5)(\texttt{fly}(x) \mid \texttt{bird}(x))$, where '$> 0.5$' corresponds to 'Most',
- '90% of birds can fly': $(\mathrm{CP}\, x = 0.9)(\texttt{fly}(x) \mid \texttt{bird}(x))$,
- 'Approximately 90% of birds fly': $(\mathrm{CP}\, x \approx 0.9)(\texttt{fly}(x) \mid \texttt{bird}(x))$,
- 'More than 90% of birds can fly': $(\mathrm{CP}\, x > 0.9)(\texttt{fly}(x) \mid \texttt{bird}(x))$,
- 'Almost all birds fly': $(\mathrm{CP}\, x \approx 1)(\texttt{fly}(x) \mid \texttt{bird}(x))$, etc.

Example 3.5. In medicine, two commonly used statistics, when considering some diagnostic test as an indicator that a patient has some disease, are so called sensitivity and specificity. Sensitivity is defined as the percentage of true positive cases relative to the sum of true positives and false negatives (i.e., the total number of tested patients having the disease). Specificity is defined as the percentage of true negative cases relative to the sum of true negatives and false positives (i.e., the total number of tasted patients who do not have the disease). Let the formulas $\texttt{tested}(x)$, $\texttt{positive}(x)$, $\texttt{disease}(x)$ have the following meaning:

$\texttt{tested}(x)$: the patient x has been tested,
$\texttt{positive}(x)$: the patient x was positive on the test,
$\texttt{disease}(x)$: the patient x has the disease.

We can express the requirement that the sensitivity of the test is at least 95% by

$$(\mathrm{CP}\, x \geqslant 0.95)(\texttt{positive}(x) \mid \texttt{tested}(x) \wedge \texttt{disease}(x)).$$

If we know that the specificity of the test is higher than 90%, this can be stated as

$$(\mathrm{CP}\, x > 0.9)(\neg\texttt{positive}(x) \mid \texttt{tested}(x) \wedge \neg\texttt{disease}(x)).$$

Definition 3.4. Suppose $\overline{\mathfrak{A}}$ is an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model, and α a formula. Then $\overline{\mathfrak{A}} \models \alpha$ (read: α is true in $\overline{\mathfrak{A}}$) if $\overline{\mathfrak{A}}, v \models \alpha$ for every assignment $v : \mathrm{Var} \to A$.

A formula is valid if it is true in every $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model.

Note that in the case some free variables occur in a valid formula we treat them as universally quantified.

Example 3.6. The formula

$$(\mathrm{P}x > 0)\alpha \wedge (\mathrm{CP}\, x \geqslant r)(\beta \mid \alpha) \to (\mathrm{CP}\, x \leqslant 1 - r)(\neg\beta \mid \alpha)$$

is $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-valid, but

$$(\mathrm{CP}\, x \geqslant r)(\beta \mid \alpha) \to (\mathrm{CP}\, x \leqslant 1 - r)(\neg\beta \mid \alpha)$$

is not.

Axiomatization. As we shall prove, the set of all valid formulas can be characterized by the following set of axiom schemata:

(**FO**) all $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-instances of the axioms for $L_{\omega\omega}$;
(**CP1**) $(\mathrm{CP}\,\vec{x} \geqslant 0)(\alpha \mid \beta)$;
(**CP2**) $(\mathrm{CP}\,\vec{x} \leqslant r_1)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} < r_2)(\alpha \mid \beta)$,$r_1 < r_2$;
(**CP3**) $(\mathrm{CP}\,\vec{x} < r)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \leqslant r)(\alpha \mid \beta)$;
(**CP4**) $(\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \geqslant q - 1/n)(\alpha \mid \beta)$, for every positive integer n such that $0 \leqslant q - \frac{1}{n}$;
(**CP5**) $(\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \leqslant q + 1/n)(\alpha \mid \beta)$, for every positive integer n such that $q + \frac{1}{n} \leqslant 1$;
(**CP6**) $(\mathrm{P}\vec{x} = 0)\beta \to (\mathrm{CP}\,\vec{x} = 1)(\alpha \mid \beta)$;
(**CP7**) $((\mathrm{P}\vec{x} = r_1)\beta \wedge (\mathrm{P}\vec{x} = r_2)(\alpha \wedge \beta)) \to (\mathrm{CP}\,\vec{x} = \frac{r_2}{r_1})(\alpha \mid \beta)$, $r_1 \neq 0$;
(**P1**) $(\mathrm{P}\vec{x} \geqslant 1)(\alpha \leftrightarrow \beta) \to ((\mathrm{P}\vec{x} = r)\alpha \to (\mathrm{P}\vec{x} = r)\beta)$;
(**P2**) $(\mathrm{P}\vec{x} \leqslant r)\alpha \leftrightarrow (\mathrm{P}\vec{x} \geqslant 1 - r)\neg\alpha$;
(**P3**) $((\mathrm{P}\vec{x} = r_1)\alpha \wedge (\mathrm{P}\vec{x} = r_2)\beta \wedge (\mathrm{P}\vec{x} = 0)(\alpha \wedge \beta)) \to$
$(\mathrm{P}\vec{x} = \min\{1, r_1 + r_2\})(\alpha \vee \beta)$;
(**P4**) $(\mathrm{P}x_1 \cdots x_i \cdots x_n {\diamond} r)\alpha \leftrightarrow (\mathrm{P}x_1 \cdots y \cdots x_n {\diamond} r)\alpha(x_i := y)$, where y is a variable that does not occur in α, and $\diamond \in \{\leqslant, \geqslant, \approx\}$;
(**P5**) $(\mathrm{P}x_1 \cdots x_n \diamond r)\alpha \leftrightarrow (\mathrm{P}x_{\pi(1)} \cdots x_{\pi(n)} \diamond r)\alpha$, where π is a permutation of $\{1, \ldots, n\}$, and $\diamond \in \{\leqslant, \geqslant, \approx\}$;
(**P6**) $(\mathrm{P}\vec{x} {\diamond} r)\alpha(\vec{x}) \leftrightarrow (\mathrm{P}\vec{x}\vec{y} {\diamond} r)\alpha(\vec{x})$, where $\diamond \in \{\leqslant, \geqslant, \approx\}$, and variables $\vec{y}$ are not free in α

(**Gen**) (generalization) plus the following:

(**Nec**) $\dfrac{(\forall x)\alpha}{(\mathrm{P}x = 1)\alpha}$;

(**Ran**) $\dfrac{\alpha \to (\mathrm{P}\vec{x} \neq r)\beta(\vec{x}), r \in \mathbb{I}}{\alpha \to \bot}$;

(**Approx**) For every $q \in \mathbb{I}_{\mathbb{Q}} \setminus \{0, 1\}$

$$\frac{\gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta), n \geqslant \frac{1}{q} \quad \gamma \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta), n \geqslant \frac{1}{1-q}}{\gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)},$$

$$\frac{\gamma \to (\mathrm{CP}\,\vec{x} \geqslant 1 - \frac{1}{n})(\alpha \mid \beta), n \geqslant 1}{\gamma \to (\mathrm{CP}\,\vec{x} \approx 1)(\alpha \mid \beta)},$$

and

$$\frac{\gamma \to (\mathrm{CP}\,\vec{x} \leqslant \frac{1}{n})(\alpha \mid \beta), n \geqslant 1}{\gamma \to (\mathrm{CP}\,\vec{x} \approx 0)(\alpha \mid \beta)},$$

where the range of the parameter n is the set of naturals $\mathbb{N}$.

Beside the axioms which are analogous to the corresponding ones from [117], we add several new ones which support the coherent behavior of the sequence of measures of an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-structure: Axiom (**P4**) captures the fact that variants of a formula are equivalent, Axiom (**P5**) enforces the constraint that measures are invariant under permutations. Note that the axioms (**CP4**) and (**CP5**), and the rule (**Approx**) describe the relationship between the standard conditional probability and the conditional probability infinitesimally close to some rational r. By Rule (**Ran**), at the syntax level, we define the range of probability functions to be the set $\mathbb{I}$. The rules (**Ran**) and (**Approx**) are infinitary in the sense that they have an infinite number of premises (the premises of these rules are not infinite conjunctions).

The intuition behind the rule (**Ran**) is the following. Since we want the set $\mathbb{I}$ to be the range of probability functions, the infinitely long formula $\bigvee_{r\in\mathbb{I}}(\mathrm{P}\vec{x} = r)\beta(\vec{x})$ must be valid, i.e. the (meta-)equivalence $\bigwedge_{r\in\mathbb{I}}(\mathrm{P}\vec{x} \neq r)\beta(\vec{x}) \Leftrightarrow \bot$ must be valid. Our language does not allow infinitary formulas, so the above equivalence may be replaced by an infinitary rule of inference: given the set of premises $(\mathrm{P}\vec{x} \neq r)\beta(\vec{x})$, $r \in \mathbb{I}$, one may infer $\bot$. In order to be able to prove Deduction theorem (see the proof of Theorem 3.2), we modify this rule by adding a prefix '$\alpha \to$' to the premises and to the conclusion: given the set of premises $\alpha \to (\mathrm{P}\vec{x} \neq r)\beta(\vec{x})$, $r \in \mathbb{I}$, one may infer $\alpha \to \bot$. Note that in the rule (**Ran**), the formula α can have free variables in common with β. For an example we use the language $L = \{U, B\}$ from Example 1. Let $\overline{\mathfrak{A}}$ be an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model such that $\overline{\mathfrak{A}} \models U(x) \to (\mathrm{P}y \neq r)B(x, y)$. If $U^{\mathfrak{A}} \neq \emptyset$, choose any $a \in U^{\mathfrak{A}}$. Then $\mu_1(\{b \in A \mid (a, b) \in B^{\mathfrak{A}}\}) \neq r$, for each $r \in \mathbb{I}$, which contradicts the fact $\{b \in A \mid (a, b) \in B^{\mathfrak{A}}\} \in \mathcal{F}_1$. Similarly, the (meta-)equivalence

$$\bigwedge_{n\geqslant\frac{1}{q}} \Big(\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n}\Big)(\alpha \mid \beta) \wedge \bigwedge_{n\leqslant\frac{1}{1-q}} \Big(\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n}\Big)(\alpha \mid \beta) \Leftrightarrow (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$$

leads to the axioms (**CP4**) and (**CP5**), and to the rule (**Approx**) (taking into account the mentioned modification needed for Deduction theorem).

The infinitary rules of inferences are indispensable to the proof of the strong completeness theorem of our logic. It is easy to see that the logic $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ is not compact, i.e., that there is a set of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-sentences T such that every finite subset of T is satisfiable (has an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model), but the set T itself is not. Consider the set $T = \{(\mathrm{P}x \neq r)U(x) \mid r \in \mathbb{I}\}$, where U is a unary relation symbol. The lack of compactness forces us to consider infinitary inference rules if we want to solve one of the main proof-theoretical problem: providing an axiomatic system which would be strongly complete.

Definition 3.5. A proof of a formula α from a set Γ of formulas is a countable sequence of formulas α_κ indexed by countable ordinal numbers such that the last formula is α, and each formula in the sequence is an axiom, or a formula in Γ or it is derived from the preceding formulas by a rule of inference with no application of (**Gen**) to a formula when the variable is free in formulas of Γ. A formula α is deducible from Γ ($\Gamma \vdash \alpha$) if there is a proof of α from Γ. A formula α is a theorem ($\vdash \alpha$) if it is deducible from the empty set.

A set Γ of formulas is consistent if there is at least one $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ formula that is not deducible from Γ, otherwise Γ is inconsistent.

A set Γ is maximal consistent iff Γ is consistent and for every formula α, either $\alpha \in \Gamma$ or $\neg\alpha \in \Gamma$. Note that if Γ is maximal consistent and $\Gamma \vdash \alpha$ then $\alpha \in \Gamma$.

Many general meta-theorems about deductions can be proven as in $L_{\omega\omega}$ including the so called 'generalization on constants': if c is a constant symbol, Γ is a set of formulas in which c does not occur, and $\alpha(x := c)$ is a formula such that $\Gamma \vdash \alpha(x := c)$, then there is a deduction of $(\forall x)\alpha(x)$ from Γ in which c does not occur.

Soundness of our system follows from the soundness of $L_{\omega\omega}$, and from the properties of probabilistic measures.

Theorem 3.1 (Soundness theorem)**.** *The axiomatic system for $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ is sound with respect to the class of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-models, i.e., each theorem is valid.*

Proof. We can show that every instance of an axiom schema is true in every $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model, while the inference rules preserve truth in a model (if the premise(s) of a rule are true in a model, then the conclusion of that rule is true in the model). The proof is straightforward, and we consider only a few cases, as an illustration.

For example, let us consider Axiom (**P6**) in the case $\diamond = \geqslant$.

Let $\overline{\mathfrak{A}}$ be an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model. We have to prove that for every formula $\alpha(x_1, \ldots, x_n)$, which does not have free variables among $y_1, \ldots, y_m$,

$$\overline{\mathfrak{A}} \models (\mathrm{P}\vec{x} \geqslant r)\alpha(\vec{x}) \leftrightarrow (\mathrm{P}\vec{x}\vec{y} \geqslant r)\alpha(\vec{x}),$$

i.e., by Definition 3.4, $\overline{\mathfrak{A}}, v \models (\mathrm{P}\vec{x} \geqslant r)\alpha(\vec{x}) \leftrightarrow (\mathrm{P}\vec{x}\vec{y} \geqslant r)\alpha(\vec{x})$, for every assignment $v : \mathrm{Var} \to A$. So, choose and fix an assignment v_0. Suppose that $\overline{\mathfrak{A}}, v_0 \models (\mathrm{P}\vec{x} \geqslant r)\alpha(\vec{x})$. Then, by Definition 3.3, we have $\mu_n[\alpha]^{\vec{x}}_{\overline{\mathfrak{A}},v_0} \geqslant r$. (Remember that $(\mathrm{P}\vec{x} \geqslant r)\alpha(\vec{x})$ abbreviates $(\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x}) \mid \top)$ and $\mu_n[\top]^{\vec{x}}_{\overline{\mathfrak{A}},v_0} = \mu_n(A^n) = 1$.) Since

$$\{(\vec{a}, \vec{b}) \in A^{n+m} \mid \overline{\mathfrak{A}}, v_0(\vec{x} := \vec{a}, \vec{y} := \vec{b}) \models \alpha\} = \{\vec{a} \in A^n \mid \overline{\mathfrak{A}}, v_0(\vec{x} := \vec{a}) \models \alpha\} \times A^m,$$

because the variables $\vec{y}$ are not free in α, it follows (by Lemma 3.1.1) that for all $\vec{b} \in A^m$: $\overline{\mathfrak{A}}, v_0(\vec{x} := \vec{a}) \models \alpha$ iff $\overline{\mathfrak{A}}, v_0(\vec{x} := \vec{a}, \vec{y} := \vec{b}) \models \alpha$. From the last clause of Definition 3.2, we obtain

$$\mu_{n+m}[\alpha]^{\vec{x},\vec{y}}_{\overline{\mathfrak{A}},v_0} = \mu_n[\alpha]^{\vec{x}}_{\overline{\mathfrak{A}},v_0} \geqslant r.$$

This proves $\overline{\mathfrak{A}}, v_0 \models (\mathrm{P}\vec{x}\vec{y} \geqslant r)\alpha(\vec{x})$. By the similar argumentation, we obtain the other direction: $\overline{\mathfrak{A}}, v_0 \models (\mathrm{P}\vec{x}\vec{y} \geqslant r)\alpha(\vec{x}) \to (\mathrm{P}\vec{x} \geqslant r)\alpha(\vec{x})$.

We now turn to the inference rules and give the proofs only for two of them.

(**Gen**) Suppose $\overline{\mathfrak{A}} \models \alpha$, for any formula α. For any variable x, we prove that $\overline{\mathfrak{A}} \models \forall x\alpha$, i.e., $\overline{\mathfrak{A}}, v \models \forall x\alpha$, for every assignment $v : \mathrm{Var} \to A$ (Definition 3.4). Choose and fix an assignment v_0. By the assumption $\overline{\mathfrak{A}} \models \alpha$, we have $\overline{\mathfrak{A}}, v \models \alpha$, for every assignment v. In particular, $\overline{\mathfrak{A}}, v_0(x := a) \models \alpha$, for all $a \in A$, and this gives $\overline{\mathfrak{A}}, v_0 \models \forall x\alpha$.

(**Approx**) We consider only the case $q \in (\mathbb{I} \cap \mathbb{Q}) \smallsetminus \{0,1\}$. Assume $\overline{\mathfrak{A}}$ satisfy the formulas $\gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{q}$ and $\gamma \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{1-q}$. We prove $\overline{\mathfrak{A}} \models \gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$, i.e., $\overline{\mathfrak{A}}, v \models \gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$, for every assignment v. Choose and fix an assignment v_0, and suppose that $\overline{\mathfrak{A}}, v_0 \models \gamma$. Then, by the above assumption, we have $\overline{\mathfrak{A}}, v_0 \models (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta)$, for each $n \geqslant \frac{1}{q}$, and $\overline{\mathfrak{A}}, v_0 \models (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$, for each $n \geqslant \frac{1}{1-q}$. It is not possible that $\mu_n[\beta]^{\vec{x}}_{\overline{\mathfrak{A}},v_0} = 0$, because $\overline{\mathfrak{A}}, v_0 \models (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$ and $q + \frac{1}{n} < 1$, for $n > \frac{1}{1-q}$ (see Definition 3.3). Hence, $\mu_n[\beta]^{\vec{x}}_{\overline{\mathfrak{A}},v_0} > 0$. For abbreviation, let $X = [\alpha \wedge \beta]^{\vec{x}}_{\overline{\mathfrak{A}},v_0}$ and $Y = [\beta]^{\vec{x}}_{\overline{\mathfrak{A}},v_0}$. Then, $\frac{\mu_n(X)}{\mu_n(Y)} \geqslant q - \frac{1}{n}$, for each $n \geqslant \frac{1}{q}$, and $\frac{\mu_n(X)}{\mu_n(Y)} \leqslant q + \frac{1}{n}$, for each $n \geqslant \frac{1}{1-q}$. We thus get $\frac{\mu_n(X)}{\mu_n(Y)} \in \bigcap_{n \in \mathbb{N}^+} \left[\max\left\{0, q - \frac{1}{n}\right\}, \min\left\{1, q + \frac{1}{n}\right\}\right] = \mathrm{monad}(q)$, i.e., $\frac{\mu_n(X)}{\mu_n(Y)} \approx q$, and so $\overline{\mathfrak{A}}, v_0 \models (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$. □

Theorem 3.2 (Deduction theorem)**.** *If* Γ *is a set of formulas and* $\Gamma, \varphi \vdash \psi$*, then* $\Gamma \vdash \varphi \to \psi$.

Proof. We use the transfinite induction on the length of the proof of ψ from Γ, φ. The cases when ψ is an axiom, or φ coincides with ψ, or $\psi \in \Gamma$ are standard.

Suppose that $\psi = \forall x\theta$ is derived from Γ, φ by (**Gen**) with premise θ. This means that x is not free in Γ, φ.

$\Gamma, \varphi \vdash \theta$

$\Gamma \vdash \varphi \to \theta$ [Induction hypothesis (IH)]

$\Gamma \vdash \forall x(\varphi \to \theta)$ [(**Gen**); x is not free in Γ]

$\Gamma \vdash \forall x(\varphi \to \theta) \to (\varphi \to \forall x\theta)$ [x is not free in φ]

$\Gamma \vdash (\varphi \to \forall x\theta)$ [Modus Ponens]

Suppose that $\psi = \gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$ is derived from Γ, φ by (**Approx**) with premises $\gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{q}$, $\gamma \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{1-q}$.

1_n. $\Gamma, \varphi \vdash \gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta)$, for each $n \geqslant \frac{1}{q}$,

2_n. $\Gamma, \varphi \vdash \gamma \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$, for each $n \geqslant \frac{1}{1-q}$,

3_n. $\Gamma \vdash \varphi \to (\gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta))$, for each $n \geqslant \frac{1}{q}$, [IH]
4_n. $\Gamma \vdash \varphi \to (\gamma \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta))$, for each $n \geqslant \frac{1}{1-q}$, [IH]
5_n. $\Gamma \vdash (\varphi \wedge \gamma) \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta)$, for each $n \geqslant \frac{1}{q}$,
6_n. $\Gamma \vdash (\varphi \wedge \gamma) \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$, for each $n \geqslant \frac{1}{1-q}$,
7. $\Gamma \vdash (\varphi \wedge \gamma) \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$ [(**Approx**)]
8. $\Gamma \vdash \varphi \to (\gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta))$
Other cases are similar. □

The next lemma gives some auxiliary statements which are needed for the proof of the completeness theorem.

Lemma 3.2. *Let* $\alpha, \beta \in L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$. *Then:*

(1) $\vdash (\mathrm{CP}\,\vec{x} \geqslant r_1)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \geqslant r_2)(\alpha \mid \beta)$, $r_1 > r_2$
(2) $\vdash (\mathrm{CP}\,\vec{x} \leqslant r_1)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \leqslant r_2)(\alpha \mid \beta)$, $r_1 < r_2$
(3) $\vdash (\mathrm{CP}\,\vec{x} = r_1)(\alpha \mid \beta) \to \neg(\mathrm{CP}\,\vec{x} = r_2)(\alpha \mid \beta)$, $r_1 \neq r_2$
(4) $\vdash (\mathrm{CP}\,\vec{x} = r_1)(\alpha \mid \beta) \to \neg(\mathrm{CP}\,\vec{x} \geqslant r_2)(\alpha \mid \beta)$, $r_1 < r_2$
(5) $\vdash (\mathrm{CP}\,\vec{x} = r_1)(\alpha \mid \beta) \to \neg(\mathrm{CP}\,\vec{x} \leqslant r_2)(\alpha \mid \beta)$, $r_1 > r_2$
(6) $\vdash (\mathrm{CP}\,\vec{x} = q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$, $q \in \mathbb{I}_{\mathbb{Q}}$
(7) $\vdash (\mathrm{CP}\,\vec{x} \approx q_1)(\alpha \mid \beta) \to \neg(\mathrm{CP}\,\vec{x} \approx q_2)(\alpha \mid \beta)$, $q_1, q_2 \in \mathbb{I}_{\mathbb{Q}}$, $q_1 \neq q_2$
(8) $\vdash (\mathrm{P}\vec{x} = 0)\beta \to \neg(\mathrm{CP}\,\vec{x} \leqslant r)(\alpha \mid \beta)$, $r < 1$
(9) $\vdash (\mathrm{P}\vec{x} \leqslant 1)\alpha$

Proof. As an illustration, we prove the statement (6). For shortness, we omit details related to the obvious arguments.

1. $(\mathrm{CP}\,\vec{x} = q)(\alpha \mid \beta) \leftrightarrow (\mathrm{CP}\,\vec{x} \geqslant q)(\alpha \mid \beta) \wedge (\mathrm{CP}\,\vec{x} \leqslant q)(\alpha \mid \beta)$
2_n. $(\mathrm{CP}\,\vec{x} \geqslant q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{q}$ [(1) of this lemma]
3_n. $(\mathrm{CP}\,\vec{x} \leqslant q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{1-q}$ [(2) of this lemma]
4_n. $(\mathrm{CP}\,\vec{x} = q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{q}$
5_n. $(\mathrm{CP}\,\vec{x} = q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{n})(\alpha \mid \beta)$, $n \geqslant \frac{1}{1-q}$
6. $(\mathrm{CP}\,\vec{x} = q)(\alpha \mid \beta) \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$
[from 4_n, $n \geqslant \frac{1}{q}$ and 5_n, $n \geqslant \frac{1}{1-q}$ by (**Approx**)] □

Note that, by restricting β to $\top$, we obtain the analogous statements for unconditional probabilities (except the statements 8 and 9, of course).

Example 3.7. From 'Most birds fly'

$$(\mathrm{CP}\,x > 0.5)(\texttt{fly}(x) \mid \texttt{bird}(x))$$

and 'Penguins do not fly'

$$\forall x(\texttt{penguin}(x) \to \neg\texttt{fly}(x)),$$

we can deduce 'Most birds are not penguins'

$$(\mathrm{CP}\,x > 0.5)(\neg\texttt{penguin}(x) \mid \texttt{bird}(x)).$$

In order to prove the completeness theorem for $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$, we follow the Henkin style procedure and the methodology used in our previous works [92].

Theorem 3.3 (Completness theorem)**.** *If T is a consistent set of formulas, then T has a model.*

Proof. First, we extend L to a new language $L \cup C$ by adding a denumerable set of new constant symbols $C = \{c_n \mid n = 0, 1, \ldots\}$. Let $(\alpha_n : n = 0, 1, 2, \ldots)$ be an enumeration of all $(L \cup C)^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-formulas.

Next, we extend T to a maximal consistent set of formulas T^* which has witnesses, i.e., if $\exists x\alpha \in T^*$ then for some constant symbol $c \in C$, $\alpha(x := c) \in T^*$. This will be done by defining a sequence $(T_n : n = 0, 1, \ldots)$ of sets of formulas such that:

- $T_0 \subseteq T_1 \subseteq T_2 \subseteq \cdots$;
- T_n is consistent for each n;
- only finitely many new constants from C occur in T_n for each n.

The sequence T_n, $n = 0, 1, \ldots$, is defined as follows:
Let $T_0 = T$.
For every $n \geqslant 0$,

(1) If α_n is a sentence, and
  (a) If $T_n \cup \{\alpha_n\}$ is consistent, we have:
    (i) if α_n is of the form $\exists x\alpha(x)$, then $T_{n+1} = T_n \cup \{\alpha_n, \alpha(x := c)\}$, for some new constant symbol $c \in C$ not occurring in any formula of $T_n \cup \{\alpha_n\}$, say the first one under the given well-ordering of all the new constant symbols,
    (ii) otherwise, $T_{n+1} = T_n \cup \{\alpha_n\}$,
  (b) otherwise, if $T_n \cup \{\alpha_n\}$ is inconsistent, we have:
    (i) if α_n is of the form $\gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha \mid \beta)$, then $T_{n+1} = T_n \cup \{\neg\alpha_n, \gamma \to \neg(\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{k})(\alpha \mid \beta)\}$ or $T_{n+1} = T_n \cup \{\neg\alpha_n, \gamma \to \neg(\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{k})(\alpha \mid \beta)\}$, where k is chosen so that T_{n+1} is consistent (we will prove that this is possible below),
    (ii) otherwise, $T_{n+1} = T_n \cup \{\neg\alpha_n\}$
(2) otherwise, if α_n is an open formula, then $T_{n+1} = T_n \cup \{(\mathrm{P}\vec{x} = r)\alpha_n\}$, where $\vec{x}$ is the tuple of all free variables of α_n and r is chosen to be an element of $\mathbb{I}$ such that T_{n+1} is consistent (we will prove that this is possible below),

Note that at each stage we extend the previous set of formulas by finitely many formulas.

Let $T^* = \bigcup_{n=0}^{\infty} T_n$.

Claim 1. T_n is consistent for each n.

Proof of Claim 1. We prove this by considering the steps of the previous construction.

T_0 is consistent by the assumption of the theorem.

The sets obtained by the steps 1(a)(ii) and 1(b)(ii) are obviously consistent. The step 1(a)(i) produces a consistent set by 'the generalization on constants'.

Let us consider the step 1(b)(i).

Suppose that α_n is of the form $\gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x}) \mid \beta(\vec{x}))$, and that $T_n \cup \{\alpha_n\}$ is inconsistent. Also, let all the sets

$$T_n \cup \Big\{\neg\alpha_n, \gamma \to \neg\Big(\,\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{k}\Big)\big(\alpha(\vec{x}) \mid \beta(\vec{x})\big)\Big\},$$

for every positive integer k such that $q - \frac{1}{k} \geqslant 0$, and

$$T_n \cup \Big\{\neg\alpha_n, \gamma \to \neg\Big(\,\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{k}\Big)\big(\alpha(\vec{x}) \mid \beta(\vec{x})\big)\Big\},$$

for every positive integer k such that $q + \frac{1}{k} \leqslant 1$, be inconsistent. Then the following contradicts consistency of T_n:
$T_n, \neg\alpha_n, \gamma \to \neg(\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{k})(\alpha(\vec{x}) \mid \beta(\vec{x})) \vdash \bot$, for every $k, q - \frac{1}{k} \geqslant 0$,
$T_n, \neg\alpha_n, \gamma \to \neg(\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{k})(\alpha(\vec{x}) \mid \beta(\vec{x})) \vdash \bot$, for every $k, q + \frac{1}{k} \leqslant 1$,
$T_n, \neg\alpha_n \vdash \gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - \frac{1}{k})(\alpha(\vec{x}) \mid \beta(\vec{x}))$, for every $k, q - \frac{1}{k} \geqslant 0$,
$T_n, \neg\alpha_n \vdash \gamma \to (\mathrm{CP}\,\vec{x} \leqslant q + \frac{1}{k})(\alpha(\vec{x}) \mid \beta(\vec{x}))$, for every $k, q + \frac{1}{k} \leqslant 1$,
$T_n, \neg\alpha_n \vdash \gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x}) \mid \beta(\vec{x}))$ by Rule (**Approx**).

Finally, we consider the step 2. Suppose that for every $r \in \mathbb{I}$, $T_n \cup \{(\mathrm{P}\vec{x} = r)\alpha_n(\vec{x})\}$ is inconsistent. Then the following contradicts consistency of T_n:
$T_n, (\mathrm{P}\vec{x} = r)\alpha_n(\vec{x}) \vdash \bot$, for every $r \in \mathbb{I}$, by the hypothesis,
$T_n \vdash \top \to \neg(\mathrm{P}\vec{x} = r)\alpha_n(\vec{x})$, for every $r \in \mathbb{I}$,
$T_n \vdash \top \to \bot$, by Rule (**Ran**),
$T_n \vdash \bot$.

Thus, the proof of the Claim 1 is completed.

Claim 2. T^* is deductively closed.

Proof of Claim 2.

Let φ be an $L^{\mathrm{P};\mathbb{I}}_{\omega\omega}$-formula. It can be proved by induction on the length of the inference that if $T^* \vdash \varphi$, then $\varphi \in T^*$. Note that if $T_m \vdash \varphi$ and $\varphi = \alpha_n$, it must be $\varphi \in T^*$ because $T_{\max\{m,n+1\}}$ is consistent.

Suppose that the sequence $\varphi_1, \varphi_2, \dots, \varphi$ forms the proof of φ from T^*.

If the sequence is finite, there must be a set T_n such that $T_n \vdash \varphi$, and $\varphi \in T^*$. Thus, suppose that the sequence is countable infinite. We can show that for every i, if φ_i is obtained by an application of an inference rule, and all the premises belong to T^*, then it must be $\varphi_i \in T^*$.

If the rule is a finitary one, then we conclude $\varphi_i \in T^*$ by reasoning as above. Next we consider infinitary rules. Let $\varphi_i = \gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x}) \mid \beta(\vec{x}))$ be obtained by Rule (**Approx**) from the set of premises φ_{i_n} of the form $\gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - 1/n)(\alpha(\vec{x}) \mid \beta(\vec{x}))$, $n \geqslant 1$, $q - 1/n \geqslant 0$, and φ_{j_n} of the form $\gamma \to (\mathrm{CP}\,\vec{x} \leqslant q + 1/n)(\alpha(\vec{x}) \mid \beta(\vec{x}))$, $n \geqslant 1$, $q + 1/n \leqslant 1$. Suppose that $\varphi_i \notin T^*$. By the step 1(b)(i) of the construction, there are some k and m such that $\gamma \to \neg(\mathrm{CP}\,\vec{x} \geqslant q - 1/k)(\alpha(\vec{x}) \mid \beta(\vec{x}))$ or $\gamma \to \neg(\mathrm{CP}\,\vec{x} \leqslant q + 1/k)(\alpha(\vec{x}) \mid \beta(\vec{x}))$ belongs to T_m. Let us suppose the former case, while the latter one follows similarly. It means that there is some ℓ such that $\gamma \to (\mathrm{CP}\,\vec{x} \geqslant q - 1/k)(\alpha(\vec{x}) \mid \beta(\vec{x})), \gamma \to \neg(\mathrm{CP}\,\vec{x} \geqslant q - 1/k)(\alpha(\vec{x}) \mid \beta(\vec{x})) \in T_\ell$. Then, $T_\ell \vdash \gamma \to \bot$, and $T_\ell \vdash \gamma \to (\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x}) \mid \beta(\vec{x}))$. It follows that $\varphi_i \in T^*$, a contradiction. The case concerning formulas obtained by Rule (**Ran**) can be proved in the same way.

Claim 3. T^* is maximal consistent.

Proof of Claim 3. Let us first observe that T^* does not contain all formulas. For a formula φ, either $\varphi \in T^*$ or $\neg\varphi \in T^*$, and the set T^* does not contain both. Thus, T^* is consistent. The construction guarantees that it is maximal.

Claim 4. T^* has the following properties:

(1) T^* contains all theorems;
(2) If $\varphi \in T^*$, then $\neg\varphi \notin T^*$;
(3) $\varphi \wedge \psi \in T^*$ iff $\varphi \in T^*$ and $\psi \in T^*$;
(4) $\varphi \vee \psi \in T^*$ iff $\varphi \in T^*$ or $\psi \in T^*$;
(5) If $\varphi, \varphi \to \psi \in T^*$, then $\psi \in T^*$;
(6) If $\exists x\alpha(x) \in T^*$, then there is a constant symbol c such that $\alpha(x := c) \in T^*$;
(7) There exists exactly one $r \in \mathbb{I}$, such that $(\mathrm{P}\vec{x} = r)\alpha(\vec{x}) \in T^*$;
(8) There exists exactly one $r \in \mathbb{I}$, such that $(\mathrm{CP}\,\vec{x} = r)(\alpha(\vec{x}) \mid \beta(\vec{x})) \in T^*$;
(9) If $(\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x}) \mid \beta(\vec{x})) \in T^*$, there is $r' \in \mathbb{I}$ such that $r' \geqslant r$ and $(\mathrm{CP}\,\vec{x} = r')(\alpha(\vec{x}) \mid \beta(\vec{x})) \in T^*$;
(10) If $(\mathrm{CP}\,\vec{x} \leqslant r)(\alpha(\vec{x}) \mid \beta(\vec{x})) \in T^*$, there is $r' \in \mathbb{I}$ such that $r' \leqslant r$ and $(\mathrm{CP}\,\vec{x} = r')(\alpha(\vec{x}) \mid \beta(\vec{x})) \in T^*$;
(11) If $(\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x}) \mid \beta(\vec{x})) \in T^*$, and $q' \in (\mathbb{I} \cap \mathbb{Q}) \setminus \{q\}$, then $(\mathrm{CP}\,\vec{x} \approx q')(\alpha(\vec{x}) \mid \beta(\vec{x})) \notin T^*$.

We omit the proof of this claim.

Using T^*, we define a model for T.

The construction of the classical model $\mathfrak{A}$ from the constants $c \in C$ by taking the equivalence classes $[c]$ is standard. If $\vec{c} = (c_1, \ldots, c_n)$ is a tuple of constant symbols, we write $[\vec{c}]$ for the tuple $([c_1], \ldots, [c_n])$. For every formula $\varphi(x_1, \ldots, x_n)$, let

$$\langle\varphi(\vec{x})\rangle = \Big\{[\vec{c}] \in A^n \Big| \varphi(\vec{x} := \vec{c}) \in T^*\Big\}.$$

Let $\mathcal{F}_n$ be the collection of subsets of A^n of the form $\langle\varphi(\vec{x})\rangle$, for some formula $\varphi(\vec{x})$ with n free variables. It is easy to prove that each $\mathcal{F}_n$ is a field of subsets of A^n and that all the clauses of Definition 3.2 are fulfilled. Now, we define the probabilities $\mu_n : \mathcal{F}_n \to \mathbb{I}$:

$$\mu_n(\langle\varphi(\vec{x})\rangle) = r \ \text{ iff } \ (\mathrm{P}\vec{x} = r)\varphi(\vec{x}) \in T^*.$$

For each n, μ_n is well-defined. Suppose that $\langle\varphi(\vec{x})\rangle = \langle\psi(\vec{x})\rangle$. Then, for every tuple $\vec{c} \in C^n$,

$$\varphi(\vec{x} := \vec{c}) \in T^* \text{ iff } \ \psi(\vec{x} := \vec{c}) \in T^*.$$

This equivalence implies $(\forall\vec{x})(\varphi \leftrightarrow \psi) \in T^*$. If not, then $(\exists\vec{x})\neg(\varphi \leftrightarrow \psi) \in T^*$, and there would be a tuple $\vec{c} \in C^n$ such that $(\varphi(\vec{c}) \wedge \neg\psi(\vec{c})) \vee (\neg\varphi(\vec{c}) \wedge \psi(\vec{c})) \in T^*$, i.e., $(\varphi(\vec{c}) \wedge \neg\psi(\vec{c})) \in T^*$ or $(\neg\varphi(\vec{c}) \wedge \psi(\vec{c})) \in T^*$, which contradicts the above equivalence. Thus, we have $(\mathrm{P}\vec{x} = 1)(\varphi \leftrightarrow \psi) \in T^*$. Axiom (**P1**) and the definition of μ_n give the equality $\mu_n(\langle\varphi(\vec{x})\rangle) = \mu_n(\langle\psi(\vec{x})\rangle)$.

For every n, μ_n is a finitely additive probability measure.

Let $\overline{\mathfrak{A}} = \langle\mathfrak{A}, \mathcal{F}_n, \mu_n\rangle_{n\in\omega}$.

Claim 5. $\overline{\mathfrak{A}}$ is an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model.

It is clear from the construction of $\overline{\mathfrak{A}}$ that μ_n is defined on each singleton set of A^n, since the formula $x_1 = c_1 \wedge \cdots \wedge x_n = c_n$ defines the singleton set $\{([c_1], \ldots, [c_n])\}$.

Claim 6. For every formula $\varphi(\vec{y})$ and every assignment $v : \mathrm{Var} \to A$, $v(y_i) = [c_i]$, we have

$$\overline{\mathfrak{A}}, v \models \varphi \text{ iff } \varphi(\vec{y} := \vec{c}) \in T^*.$$

We consider only the case when φ begins with a probability quantifier.

Let φ be a formula of the form $(\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$.

Suppose that $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$. If $\mu_n(\langle\beta(\vec{x},\vec{c})\rangle) = 0$, then $(\mathrm{P}\vec{x} = 0)\beta(\vec{x},\vec{c}) \in T^*$, and by Axiom (**CP6**) and Lemma 2.1 $(\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$. So, let $\mu_n(\langle\alpha(\vec{x},\vec{c}) \wedge \beta(\vec{x},\vec{c})\rangle) = s$, $\mu_n(\langle\beta(\vec{x},\vec{c})\rangle) = t$, $t > 0$, and by Claim 4, $s/t \geqslant r$. By Axiom (**CP7**) it must be $(\mathrm{CP}\,\vec{x} = s/t)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$, and again by the monotonicity of the conditional probability, $(\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$.

For the other direction, suppose that $(\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$. From Claim 4 there are unique $s, t \in \mathbb{I}$ such that $(\mathrm{P}\vec{x} = s)(\alpha(\vec{x},\vec{c}) \wedge \beta(\vec{x},\vec{c})) \in T^*$ and $(\mathrm{P}\vec{x} = t)\beta(\vec{x},\vec{c}) \in T^*$, i.e., $\mu_n(\langle\alpha(\vec{x},\vec{c}) \wedge \beta(\vec{x},\vec{c})\rangle) = s$ and $\mu_n(\langle\beta(\vec{x},\vec{c})\rangle) = t$. If $t = 0$, then $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$ by Definition 4. If $t \neq 0$, by Axiom (**CP7**), $(\mathrm{CP}\,\vec{x} = s/t)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$. It follows from Claim 4 that $s/t \geqslant r$, so $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \geqslant r)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$.

The case $\varphi = (\mathrm{CP}\,\vec{x} \leqslant r)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$ follows similarly.

Finally, let $\varphi = (\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$.

Suppose that $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$. Then, for all positive integer n, m such that $0 \leqslant q - 1/n < q < q + 1/m \leqslant 1$, we have $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \geqslant q - 1/n)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$ and $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \leqslant q + 1/m)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$. It follows that $(\mathrm{CP}\,\vec{x} \geqslant q - 1/n)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y})), (\mathrm{CP}\,\vec{x} \leqslant q + 1/m)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y})) \in T^*$. If $(\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \notin T^*$, the step 2(b)(i) of the construction of the set T^* guarantees that for some positive integer k, either $\neg(\mathrm{CP}\,\vec{x} \geqslant q - 1/k)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$ or $\neg(\mathrm{CP}\,\vec{x} \leqslant q + 1/k)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$ which contradicts consistency of T^*. Thus, $(\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$.

For the other direction, suppose that $(\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$. From Claim 4 there are unique $t \in \mathbb{I}$ such that $(\mathrm{P}\vec{x} = t)\beta(\vec{x},\vec{c}) \in T^*$. If $t = 0$, by Axiom (**CP6**), it must be $q = 1$, and then $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \approx 1)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{y}))$. If $t > 0$, by Axioms (**CP4**) and (**CP5**), $(\mathrm{CP}\,\vec{x} \geqslant q + 1/m)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})), (\mathrm{CP}\,\vec{x} \leqslant q - 1/n)(\alpha(\vec{x},\vec{c}) \mid \beta(\vec{x},\vec{c})) \in T^*$ for all positive integers n, m such that $0 \leqslant q - 1/n < q < q + 1/m \leqslant 1$. Then reasoning as above, we have $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \geqslant q - 1/n)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{c}))$ and $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \leqslant q + 1/m)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{c}))$, which means that $\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x} \approx q)(\alpha(\vec{x},\vec{y}) \mid \beta(\vec{x},\vec{c}))$. □

If we extend the list of axioms and rules of $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$ with the following axiom

(**Prod**) $\quad (\mathrm{P}\vec{x} = r)\alpha(\vec{x}) \wedge (\mathrm{P}\vec{y} = s)\beta(\vec{y}) \to (\mathrm{P}\vec{x}\,\vec{y} = rs)(\alpha(\vec{x}) \wedge \beta(\vec{y}))$,

provided all variables in $\vec{x}$, $\vec{y}$ are distinct, we are able to prove the completeness theorem for the class of product $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-models.

Definition 3.6. A product model is an $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-model $\overline{\mathfrak{A}} = \langle \mathfrak{A}, \mathcal{F}_n, \mu_n \rangle_{n \in \mathbb{N}}$ such that the sequence of probabilities $(\mu_n : n = 1, 2, \ldots)$ is a sequence of product measures: for any two sets $X \subseteq A^m$ and $Y \subseteq A^n$ $(m, n \in \mathbb{N})$, and their Cartesian product $X \times Y \subseteq A^{m+n}$, if $X \in \mathcal{F}_m$ and $Y \in \mathcal{F}_n$, then $X \times Y \in \mathcal{F}_{m+n}$ and $\mu_{m+n}(X \times Y) = \mu_m(X) \cdot \mu_n(Y)$.

Theorem 3.4. *If $T \cup \{\mathbf{Prod}\}$ is consistent, then T has a product $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-model.*

Proof. The proof of the theorem is almost the same as the proof of Theorem 3.3. Of course, the definition of the corresponding sequence of theories $(T_n : n = 0, 1, \ldots)$ begins with $T_0 = T \cup \{\mathbf{Prod}\}$. The only new fact that should be proven in this case is that the sequence of probabilities of the canonical $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$-model is a sequence of product measures. But this is evident since every instance of the axiom (**Prod**) is true in the canonical model. Let $\mu_n(\langle \varphi(x_1, \ldots, x_n) \rangle) = r$ and $\mu_m(\langle \psi(y_1, \ldots, y_m) \rangle) = s$, provided all variables in $\vec{x}$, $\vec{y}$ are distinct. It is easy to see that the set $\langle \varphi(\vec{x}) \wedge \psi(\vec{y}) \rangle$ represents the Cartesian product $\langle \varphi(x_1, \ldots, x_n) \rangle \times \langle \psi(y_1, \ldots, y_m) \rangle$. From $(\mathrm{P}\vec{x} = r)\varphi(\vec{x}) \in T^*$, $(\mathrm{P}\vec{y} = s)\psi(\vec{y}) \in T^*$, using the the axiom (**Prod**), we have $(\mathrm{P}\vec{x}\,\vec{y} = r \cdot s)(\varphi(\vec{x}) \wedge \psi(\vec{y})) \in T^*$. □

Example 3.8. Note that even if we consider the axiomatic system with the axiom (**Prod**), no two of the sentences

$$(\mathrm{P}x \geqslant 1/2)(\mathrm{P}y \geqslant 1/2)R(x, y),$$
$$(\mathrm{P}y \geqslant 1/2)(\mathrm{P}x \geqslant 1/2)R(x, y),$$
$$(\mathrm{P}xy \geqslant 1/4)R(x, y),$$

are equivalent. These facts can be checked by considering structures with three elements, each singleton having the measure 1/3.

4. Decidable fragments of $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$

Since first-order classical logic is undecidable, the same holds for the logic $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$. If we want our logic to be applicable, a compromise has to be made: we should find a fragment which is decidable but still has significant expressive power. Numerous fragments of $L_{\omega\omega}$ were proved decidable (for validity or satisfiability). One of the best general reference sources here is [11]. Among these fragments we consider some of the most traditional ones, i.e. collections of prenex formulas defined by restrictions on the quantifiers prefix and language. A prenex formula is a formula with all its quantifiers up front. Recall that there is a simple algorithm for transforming an arbitrary first-order formula to an equivalent one in the prenex form.

In the search for decidable fragments of $L_{\omega\omega}^{\mathrm{P},\mathbb{I}}$, the idea was to forbid the nesting of probability quantifiers and try to find a decidable fragment of classical first-order logic to which a technique for proving decidability, developed in [117], could be applied. It turned out that the key were the closure conditions from Lemma 4.1 below, i.e., closure under negation, conjunction and existential quantification. Based on this we defined a subfragment of the fragment defined by Gödel, Kalmár, and Schütte, for which decidability can be proved. Gödel, Kalmár, and Schütte discovered a decision procedure for the satisfiability of sentences from the class

$[\exists^*\forall^2\exists^*, \text{all}]$ which contains the sentences of the form $\exists x_1 \dots \exists x_k \forall y_1 \forall y_2 \exists z_1 \dots \exists z_\ell \varphi$ where φ is quantifier free, contains any number of relation symbols of any arity and no function symbols of arity $\geqslant 1$, and does not contain equality symbol.

Since the equality may be important for some applications, we also define a subfragment of the fragment introduced by Shelah, which allows equality. The proof of decidability is practically the same for both fragments.

Note that both fragments may include individual constants.

Theorem 4.1 (Gödel 1932, Kalmár 1933, Schütte 1934)**.** *The satisfiability and the finite satisfiability problems are decidable for the class* $\left[\exists^*\forall^2\exists^*, \text{all}\right]$.

We choose a class $\mathcal{G}_L$ of first-order formulas (in a language L) in order to find a class of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-sentences for which the satisfiability and validity problems are decidable.

Definition 4.1. Let L be a first-order language. By $\mathcal{G}_L$ we denote the class of all first-order formulas whose prenex form has the quantifier prefix from the set

$$\{\exists, \forall, \exists\exists, \exists\forall, \forall\exists, \forall\forall, \exists\exists\forall, \exists\forall\forall, \forall\exists\exists, \forall\forall\exists, \exists\exists\forall\forall, \forall\forall\exists\exists\}.$$

Let us fix two variables x and y. Note that we can always transform a formula from $\mathcal{G}_L$ in the prenex form so that variables x and y go only with universal quantifiers, and the other variables with existential or no quantifiers. So, we suppose that only formulas in the prenex form containing the quantifiers $\{\forall x, \forall y\} \cup \{\exists v : v \in \text{Var} \setminus \{x, y\}\}$ belong to $\mathcal{G}_L$.

Lemma 4.1. 1) *The class $\mathcal{G}_L$ is closed under negation.*

2) *The existential closure of a finite conjunction of formulas from $\mathcal{G}_L$ belongs to the class* $\left[\exists^*\forall^2\exists^*, \text{all}\right]$.

Proof. 1) The key idea is to rename variables in a suitable way, as in the following examples:

$$\neg\exists u \exists v \forall x \forall y \varphi \equiv \forall x \forall y \exists u \exists v \neg\varphi[u := x, v := y, x := u, y := v],$$
$$\neg\forall x \forall y \exists u \exists v \varphi \equiv \exists u \exists v \forall x \forall y \neg\varphi[u := x, v := y, x := u, y := v].$$

2) For shortness, we give only an example. Let $\varphi_1, \varphi_2, \varphi_3$ have free variables not appearing in the scope of any given quantifier.

$$\exists u_1 \exists v_1 \forall x \forall y \varphi_1 \wedge \forall x \forall y \exists u_2 \exists v_2 \varphi_2 \wedge \exists u_3 \exists v_3 \forall x \forall y \varphi_3 \equiv$$
$$\exists u_1 \exists v_1 \exists u_3 \exists v_3 \forall x \forall y \exists u_2 \exists v_2 (\varphi_1 \wedge \varphi_2 \wedge \varphi_3) \quad \square$$

Definition 4.2. Let $\mathcal{G}^{\text{prob}}_L$ be the class of all $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-sentences which are boolean combinations of probabilistic formulas of the form $(\text{CP}\,\vec{x} \diamond r)(\varphi(\vec{x}) \mid \psi(\vec{x}))$, $\varphi(\vec{x}), \psi(\vec{x}) \in \mathcal{G}_L$.

Theorem 4.2. *Let L be a first-order language without function symbols and equality sign. The satisfiability problem for $\mathcal{G}^{\text{prob}}_L$ is decidable.*

Proof. Let β be a sentence from $\mathcal{G}_L^{\text{prob}}$. According to Axiom (**P6**), without loss of generality, we assume that all probabilistic quantifiers appearing in β bound the same variables. We consider the simplest case when probabilistic quantifiers bound one variable. The same procedure can be performed when probabilistic quantifiers in β bound a sequence of variables. In that case we obtain the measure μ_n where n is the length of the sequence of variables.

Using propositional reasoning it is easy to show that β is equivalent to a formula: $\text{DNF}(\beta) = \bigvee_{j=1}^{d} \bigwedge_{i=1}^{k_j} \pm(\text{CP } x \diamond_{ij} r_{ij})(\varphi_{ij}(x) \mid \psi_{ij}(x))$, where $\varphi_{ij}(x), \psi_{ij}(x) \in \mathcal{G}_L$, and the prefix $\pm$ denotes presence $(-)$ or absence $(+)$ of $\neg$.

Obviously, to prove decidability of our class of probabilistic sentences, it is enough to show that satisfiability of probabilistic formulas of the form

$$\bigwedge_{i=1}^{k} \pm(\text{CP } x \diamond_i r_i)(\varphi_i(x) \mid \psi_i(x)),\ \diamond_i \in \{\leqslant, \geqslant, \approx\}$$

is decidable. Moreover, an easy verification (see [117]) shows that it is enough to prove decidability of satisfiability of formulas of the form

$$\beta = \bigwedge_{i=1}^{k} \big[\pm (\text{CP } x \diamond_i r_i)(\varphi_i(x) \mid \psi_i(x)) \wedge (\text{P } x > 0)\psi_i(x)\big],\ \diamond_i \in \{\leqslant, \geqslant, \approx\}.$$

Let $\mathcal{C}_\beta = \{\varphi_1(x), \ldots, \varphi_k(x), \psi_1(x), \ldots, \psi_k(x)\}$. For every $\ell : \{1, \ldots, 2k\} \to \{0, 1\}$, let

$$\alpha_\ell(x_\ell) = \bigwedge_{i=1}^{k} \varphi_i(x_\ell)^{\ell(i)} \wedge \bigwedge_{i=1}^{k} \psi_i(x_\ell)^{\ell(k+i)},$$

where x_ℓ, $\ell = 1, \ldots, 2^{2k}$ are variables variables not appearing in formulas from $\mathcal{C}_\beta$, and φ^0 denotes $\neg\varphi$, and φ^1 denotes φ.

For every nonempty subset $I = \{i_1, \ldots, i_n\} \subseteq \{1, 2, \ldots, 2^{2k}\}$, by the previous lemma,

$$\beta_I = \exists x_{i_1} \cdots \exists x_{i_n} \left(\alpha_{i_1}(x_{i_1}) \wedge \cdots \wedge \alpha_{i_n}(x_{i_n})\right)$$

belongs to the class $\big[\exists^* \forall^2 \exists^*, \text{all}\big]$. Thus, it is decidable if β_I is satisfiable or not.

Let $\mathcal{B}$ be the set of all formulas β_I having models. If $I' \subseteq I$, then $\beta_I \to \beta_{I'}$. Thus, we can suppose that $\mathcal{B}$ contains only those β_I indexed by subsets of $\{1, \ldots, 2^{2k}\}$ which are as maximal as possible.

Next, we reduce the satisfiability problem to linear programming problem. More precisely, for every $\beta_I \in \mathcal{B}$, we solve a system $\mathcal{S}_I$ with coefficients in $\mathbb{Q}(\varepsilon)$ and unknowns $y_1, \ldots, y_h$, where $h = 2^{2k}$. Note that the same approach is used to prove decidability in [117].

For every $\gamma(x) \in \mathcal{C}_\beta$, $I_{\gamma(x)}$ is the set of all $\ell \in \{1, 2, \ldots, 2^{2k}\}$ such that the exponent of $\gamma(x)$ in $\alpha_\ell(x)$ is 1. We denote $\sum_{\ell \in I_{\gamma(x)}} y_\ell$ briefly by $\Sigma(\gamma(x))$, and $\dfrac{\Sigma(\gamma(x) \wedge \delta(x))}{\Sigma(\delta(x))}$ by $\Sigma(\gamma(x), \delta(x))$.

Fix β_I. $\mathcal{S}_I$ is the following system:

$$\begin{cases} ly_1 + \cdots + y_h = 1, \\ y_\ell \geqslant 0, \ \ \ell \in I, \\ y_\ell = 0, \ \ \ell \notin I, \\ \Sigma(\psi_i(x)) > 0, \ \ i = 1, \ldots, k, \\ \Sigma(\varphi(x), \psi(x)) \geqslant r, \ \text{if } (\mathrm{CP}\ x \geqslant r)(\varphi(x) \mid \psi(x)) \ \ \text{is a conjunct of } \beta, \\ \Sigma(\varphi(x), \psi(x)) < r, \ \text{if } (\neg\,\mathrm{CP}\ w \geqslant r)(\varphi(x) \mid \psi(x)) \ \ \text{is a conjunct of } \beta, \\ \Sigma(\varphi(x), \psi(x)) \leqslant r, \ \text{if } (\mathrm{CP}\ w \leqslant r)(\varphi(x) \mid \psi(x)) \ \ \text{is a conjunct of } \beta, \\ \Sigma(\varphi(x), \psi(x)) > r, \ \text{if } \neg(\mathrm{CP}\ w \leqslant r)(\varphi(x) \mid \psi(x)) \ \ \text{is a conjunct of } \beta, \\ \Sigma(\varphi(x), \psi(x)) \approx r, \ \text{if } (\mathrm{CP}\ w \approx r)(\varphi(x) \mid \psi(x)) \ \ \text{is a conjunct of } \beta, \\ \Sigma(\varphi(x), \psi(x)) \not\approx r, \ \text{if } \neg(\mathrm{CP}\ w \approx r)(\varphi(x) \mid \psi(x)) \ \ \text{is a conjunct of } \beta. \end{cases}$$

Systems of this kind have been considered in detail in [117]. The decision procedure on whether such a system is solvable in $\mathbb{Q}(\varepsilon)$ is described in Section 2, and here we omit it.

It remains to prove that β is satisfiable iff there is $I \subseteq \{1, 2, \ldots, 2^{2k}\}$ such that $\beta_I \in \mathcal{B}$ and $\mathcal{S}_I$ has a solution. Note that if $\mathcal{B} = \emptyset$, then it is senseless to look for a model satisfying β.

If $\beta_I \in \mathcal{B}$ and $\mathcal{S}_I$ has a solution, then we define a model $\langle \mathfrak{A}, \mu_1 \rangle$ such that $\mathfrak{A} \models \beta_I$, μ_1 is defined according to the solution of $\mathcal{S}_I$. It is enough to choose an arbitrary element a_ℓ such that $\mathfrak{A} \models \alpha_\ell[a_\ell]$, and set $\mu_1(\{a_\ell\}) = y_\ell$, and $\mu_1(\{a\}) = 0$, for $a \neq a_\ell$ and $\mathfrak{A} \models \alpha_\ell[a]$. Obviously, $\langle \mathfrak{A}, \mu_1 \rangle \models \beta$.

For the other direction, suppose that $\langle \mathfrak{A}, \mu \rangle \models \beta$. For every $\ell = 1, 2, \ldots, 2^{2k}$ let

$$A_\ell = \{a \in A \mid \mathfrak{A} \models \alpha_\ell[a]\}.$$

Let $I = \{\ell : A_\ell \neq \emptyset\}$. Then $\mathfrak{A} \models \beta_I$. Is is easy to verify that $\mathcal{S}_I$ has a solution: $y_\ell = \mu_1(A_\ell)$, $\ell \in I$ and $y_\ell = 0$, $\ell \notin I$. □

As we have mentioned, it is possible to find other decidable fragment of our logic. We define here one, based on a theorem of Shelah, which allows less quantifiers but admits one unary function symbol and equality, which might make it more suitable for some applications.

Theorem 4.3 (Shelah 1977)**.** *The satisfiability and the finite satisfiability problems are decidable for the classes* $[\exists^*\forall\exists^*, \mathrm{all}, (1)]_=$*, i.e., for the class of formulas whose prenex form has the quantifier prefix of the form* $\exists^*\forall\exists^*$ *and whose language contains equality sign, arbitrary relation and constant symbols, and at most one unary function symbol.*

Theorem 4.4. *Let L be a language with at most one unary function symbol and no function symbols of arity $\geqslant 2$ (there are no restrictions on relation and constant symbols). Let $\mathcal{S}_L$ denote the class of all $L_{\omega\omega}$-formulas possibly with equality whose prenex form has the quantifier prefix from the set $\{\exists, \forall, \exists\forall, \forall\exists\}$, and let $\mathcal{S}_L^{\mathrm{prob}}$ denote* *the class of all $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-sentences which are boolean combinations of probabilistic formulas of the form* $(\mathrm{CP}\,\vec{x} \diamond r)(\varphi(\vec{x}) \mid \psi(\vec{x}))$*, where* $\varphi(\vec{x}), \psi(\vec{x}) \in \mathcal{S}_L$*. The satisfiability problem for* $\mathcal{S}_L^{\mathrm{prob}}$ *is decidable.*

The proof is completely analogous to the proof of Theorem 4.2, starting with the analogue of Lemma 4.1, so we omit it.

5. Applications

The question might be raised whether the fragments $\mathcal{G}_L^{\mathrm{prob}}$ and $\mathcal{S}_L^{\mathrm{prob}}$ are too weak for applications. We shall argue here that they are rich enough for most practical purposes. First, we may claim that iterated probabilities ('the probability of probability ...') rarely occur in practical considerations, so one probabilistic quantifier should suffice in most situations. On the other hand, the famous 'empirical theorem' of logic that three quantifiers are the limit of human understanding is only partly a joke. It is only in the most difficult mathematical theorems that we may see three quantifiers (e.g., $\forall x \exists y \forall z \ldots$). While many mathematical statements really involve many quantifiers, they are always rephrased by replacing existential quantifiers with new constants and blocks $\forall x \exists y$ with predefined functions, so that the final statement has two or at most three (blocks of same) quantifiers.

In this section we give a few applications of our decidable fragments to problems of nonmonotonic reasoning.

In [117] defaults of the form $\alpha \mid\!\sim \beta$, where α and β are propositional formulas, are syntactically represented in a propositional probabilistic framework.

It is disputable whether the propositional language properly captures the idea behind defaults.

For instance, the classical example 'by default birds fly' is usually represented by $\mathtt{bird} \mid\!\sim \mathtt{fly}$, but a better reading might be 'if x is a bird, then - by default - x flies' which can be written as $\mathtt{bird}(x) \mid\!\sim \mathtt{flies}(x)$, i.e., it is a part of monadic first-order logic.

Proper formal analysis should be: $\forall x(\mathtt{bird}(x) \to \mathtt{flies}(x))$, but here we may note two facts:

(1) the entailment $\to$ is not the usual material implication,
(2) the quantifier $\forall x$ is not the usual universal quantifier (it is rather a kind of 'almost all' quantifier).

The first fact has been extensively analyzed and it is generally accepted that the best description of this entailment is System **P** introduced by Kraus, Lehmann and Magidor in [54, 62]. The second fact has been neglected and we believe that this is the main reason why there is a number of problems with defaults as defined by P. For a list of 6 such problems see [7].

In [47], we proposed to represent the default $\mathtt{bird}(x) \mid\!\sim \mathtt{flies}(x)$ by a formula of our system: $(\mathrm{CP}\,x \approx 1)(\mathtt{flies}(x) \mid \mathtt{bird}(x))$, where $(\mathrm{CP}\,x \approx 1)$ plays the role of 'almost all' quantifier, and conditional probability replaces the material implication. This constitutes one of many probabilistic approaches to nonmonotonic (and, in particular, default) reasoning. In the context of our decidable fragments, it is possible to apply the same analysis also to more general first-order formulas.

We may introduce what might be called binary (or n-ary) defaults, namely the statements of the type $\mathtt{a}(x, y, ...) \mid\!\sim \mathtt{b}(x, y, ...)$. For example, consider the following statement 'Students generally respect their teachers'. This may be formalized as: $(\mathrm{CP}\, xy \approx 1)(\mathtt{Respect}(x, y) \mid \mathtt{Student}(y) \wedge \mathtt{Teacher}(x, y))$.

Let us denote $(\mathrm{CP}\, \vec{x} \approx 1)(\beta \mid \alpha)$ by $\alpha \rightsquigarrow_{\vec{x}} \beta$ (we will omit the subscript $\vec{x}$ when it is clear from the context). In this new notation, the previous statement would read $\mathtt{Student}(y) \wedge \mathtt{Teacher}(x, y) \rightsquigarrow_{x,y} \mathtt{Respect}(x, y)$.

Example 5.1. If $\overline{\mathfrak{A}}$ and $\overline{\mathfrak{A}}'$ are $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-structures from Example 3.2, we have:

$$\overline{\mathfrak{A}} \not\models B(x, y) \rightsquigarrow_{x,y} U(x) \vee U(y) \text{ and } \overline{\mathfrak{A}}' \models B(x, y) \rightsquigarrow_{x,y} U(x) \vee U(y),$$

where $B(x, y) \rightsquigarrow_{x,y} U(x) \vee U(y)$ can be read 'almost all pairs (x, y) satisfying $B(x, y)$, also satisfy $U(x) \vee U(y)$'. But, note that formulas

$$(\mathrm{P}y \approx 1)(B(x, y) \rightsquigarrow_x U(x) \vee U(y)) \text{ and } \forall x(B(x, y) \rightsquigarrow_y U(x) \vee U(y))$$

hold in both models $\overline{\mathfrak{A}}$ and $\overline{\mathfrak{A}}'$.

Example 5.2. For an example of a ternary default we may use street basketball: 'Three people playing on the same street basket team are usually friends'. This might be formalized as

$$(\mathrm{CP}\, xyz \approx 1)(\mathtt{Friends}(x, y) \wedge \mathtt{Friends}(y, z) \wedge \mathtt{Friends}(z, x) \mid \mathtt{Team}(x, y, z)),$$

or $\mathtt{Team}(x, y, z) \rightsquigarrow_{x,y,z} \mathtt{Friends}(x, y) \wedge \mathtt{Friends}(y, z) \wedge \mathtt{Friends}(z, x)$.

It turns out that all the rules of System **P** are satisfied for this translation.

Theorem 5.1. *(We assume that all the free variables of the formulas α, β, γ are contained in the fixed tuple $\vec{x}$.)*

(1) $\alpha \to \gamma, \beta \rightsquigarrow \alpha \vdash \beta \rightsquigarrow \gamma$ [Right Weakening]
(2) $\vdash \beta \rightsquigarrow \beta$ [Reflexivity]
(3) $\beta \leftrightarrow \beta', \beta \rightsquigarrow \alpha \vdash \beta' \rightsquigarrow \alpha$ [Left Logical Equivalence]
(4) $\beta \rightsquigarrow \alpha, \alpha \wedge \beta \rightsquigarrow \gamma \vdash \beta \rightsquigarrow \gamma$ [Cut]
(5) $\beta \rightsquigarrow \alpha, \beta \rightsquigarrow \gamma \vdash \beta \wedge \alpha \rightsquigarrow \gamma$ [Cautions Monotonicity]

Proof. The statements (1)–(3) are obvious.

We use the completeness theorem to get properties (4) and (5). Let $\overline{\mathfrak{A}}$ be an $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-model, and v a valuation. Set $A = [\alpha]^{\vec{x}}_{\overline{\mathfrak{A}},v}$, $B = [\beta]^{\vec{x}}_{\overline{\mathfrak{A}},v}$, $C = [\gamma]^{\vec{x}}_{\overline{\mathfrak{A}},v}$. For shortness we omit the subscript in the notation of the measure. The claims (4) and (5) are obviously true if $\mu(B) = 0$. So, suppose that $\mu(B) = b > 0$.

(4) Assume that $\overline{\mathfrak{A}} \models (\mathrm{CP}\, \vec{x} \approx 1)(\alpha(\vec{x}) \mid \beta(\vec{x})), (\mathrm{CP}\, \vec{x} \approx 1)(\gamma(\vec{x}) \mid \alpha(\vec{x}) \wedge \beta(\vec{x}))$. From $\mu(A \mid B) \approx 1$, we have $\mu(A \mid B) = 1 - \varepsilon'$, where ε' is an infinitesimal. Thus, $\mu(A \cap B) = b - \varepsilon' b$. Reasoning in the same manner, we obtain that $\mu(A \cap B \cap C) = b - \varepsilon' b - \varepsilon''(b - \varepsilon' b)$, for some infinitesimal ε''. Now, by the monotonicity of the probability, we have

$$b - (\varepsilon' - \varepsilon'' + \varepsilon'\varepsilon'')b = \mu(A \cap B \cap C) \leqslant \mu(B \cap C) \leqslant \mu(B) = b,$$

Thus, $\mu(B \cap C) \approx b$, and hence $\mu(C \mid B) \approx 1$.

(5) Assume that $\overline{\mathfrak{A}} \models (\mathrm{CP}\,\vec{x} \approx 1)(\alpha(\vec{x}) \mid \beta(\vec{x})), (\mathrm{CP}\,\vec{x} \approx 1)(\gamma(\vec{x}) \mid \beta(\vec{x}))$. Note that $\mu(A \cap B) \neq 0$ since $\mu(A \mid B) \approx 1$ and $\mu(B) > 0$. From $\mu(A \mid B) \approx 1$, as in the proof of the previous statement, we have $\mu(A \cap B) = b - \varepsilon' b$, where ε' is an infinitesimal, and consequently $\mu(A^c \cap B) = \varepsilon' b$. Also, there is an infinitesimal ε'' such that $\mu(B \cap C) = b - \varepsilon'' b$. It is easy to see that $b - \varepsilon' b - \varepsilon'' b \leqslant \mu(A \cap B \cap C) \leqslant b - \varepsilon' b$. Therefore,

$$1 - \frac{\varepsilon'' b}{b - \varepsilon' b} \leqslant \frac{\mu(A \cap B \cap C)}{\mu(A \cap B)} \leqslant 1. \qquad \square$$

One of the problems with System **P** mentioned above is the so called "inheritance blocking". If the elements of some set A have a certain property R, we expect that the elements of a subset $A \cap C$ will "inherit" the property R. However, this would be in conflict with the essence of nonmonotonicity. Namely, if we were allowed to infer $\alpha \wedge \gamma \mid\!\sim \beta$ from $\alpha \mid\!\sim \beta$, for arbitrary γ, the system would be monotonous. Therefore, as expected, System **P** does not allow such derivations. Still, we would like to be able to make such derivations in some special cases, under some restrictions, but such restrictions cannot be formulated in a propositional language. Our first-order language offers un opportunity to formulate some possible solutions.

Example 5.3. Let us consider the set Δ consisting of the following two defaults: "the Swedes are blond" and "the Swedes are tall", i.e., $\mathtt{Swede}(x) \rightsquigarrow_x \mathtt{blond}(x)$ and $\mathtt{Swede}(x) \rightsquigarrow_x \mathtt{tall}(x)$. Because of the inheritance blocking problem, in P it is not possible to conclude that Swedes who are not tall are blond ($\mathtt{Swede}(x) \wedge \neg\mathtt{tall}(x) \rightsquigarrow_x \mathtt{blond}(x)$). In this particular case, it may turn out that the short Swedes are exactly the ones which are not blond. A solution might be to add a clause which excludes such possibility, for example: $(\mathrm{CP}\,x \approx 0)(\neg\mathtt{blond}(x) \mid \mathtt{Swede}(x) \wedge \neg\mathtt{tall}(x))$.

Example 5.4. We can express in our language things like: 'Married people with children do not get divorced' (a default 100 years ago). We may formalize this as

$$(\mathrm{CP}\,xy \approx 1)(\neg\mathtt{Devorced}(x,y) \mid \mathtt{Married}(x,y) \wedge \exists z \mathtt{Child}(x,y,z))$$

(or $\mathtt{Married}(x,y) \wedge \exists z\mathtt{Child}(x,y,z) \rightsquigarrow_{x,y} \neg\mathtt{Devorced}(x,y)$).

More complicated example would be: 'Generally, parents hate drug dealers'. This default may be formalized as

$$(\mathrm{CP}\,x \approx 1)(\forall y(\mathtt{DrugDealer}(y) \to \mathtt{Hate}(x,y)) \mid \exists y \mathtt{Parent}(x,y))$$

or

$$(\mathrm{CP}\,xy \approx 1)(\mathtt{DrugDealer}(y) \to \mathtt{Hate}(x,y) \mid \exists z \mathtt{Parent}(x,z)).$$

Note the slight difference in meaning of the last two formulas. While the first says that generally, a parent hates every drug dealer, the second is slightly weaker, saying that that if x is a parent and y is a drug dealer, generally, x will hate y, i.e., while the first allows exceptions only among parents, the second allows also exceptions among drug dealers.

Example 5.5. Let $L = \{M, R, P\}$ consists of three relation symbols, $\mathrm{ar}(M) = \mathrm{ar}(R) = 2$ and $\mathrm{ar}(P) = 1$. The intended meanings of these symbols might be:

$M(x,y)$ – 'x is married to y', $R(x,y)$ – 'x is a relative of y', and $T(x)$ – 'x is a member of a tribe'. Let us consider the set Δ of the universal closures of the following formulas: $M(x,y) \to M(y,x)$, $R(x,y) \to R(y,x)$, $T(x) \wedge M(x,y) \to T(y)$, $T(x) \wedge R(x,y) \to T(y)$, $\neg T(x) \wedge M(x,y) \to \neg T(y)$, $\neg T(x) \wedge R(x,y) \to \neg T(y)$, $(\exists y)M(x,y) \rightsquigarrow_x (\forall y)(M(x,y) \to \neg R(x,y))$, $T(x) \rightsquigarrow_x (\exists y)M(x,y)$, $T(x) \rightsquigarrow_x (\exists y)(M(x,y) \wedge R(x,y))$. The classical part of the conjunction is obviously satisfiable. We give a model for this part and use it in the further considerations. The model $\mathfrak{A}$ is defined over the set $A = \{\underline{1}, \underline{2}, \overline{2}, \underline{3}, \overline{3}, \underline{4}, \underline{5}, \overline{5}, \underline{6}, \overline{6}\}$. Let

$$M^{\mathfrak{A}} = \{(\underline{2}, \overline{2}), (\overline{2}, \underline{2}), (\underline{3}, \overline{3}), (\overline{3}, \underline{3}), (\underline{5}, \overline{5}), (\overline{5}, \underline{5}), (\underline{6}, \overline{6}), (\overline{6}, \underline{6})\}$$
$$R^{\mathfrak{A}} = \{(\underline{2}, \overline{2}), (\overline{2}, \underline{2}), (\underline{5}, \overline{5}), (\overline{5}, \underline{5})\},\ P^{\mathfrak{A}} = \{\underline{4}, \underline{5}, \overline{5}, \underline{6}, \overline{6}\}$$

For shortness, let $\varphi(x) = (\forall y)(M(x,y) \to \neg R(x,y))$, $\psi(x) = (\exists y)M(x,y)$, $\theta(x) = T(x)$. Note that $\neg\varphi(x) \leftrightarrow (\exists y)(M(x,y) \wedge R(x,y))$.

Now, the probabilistic part of Δ consists of the formulas:

$$(\mathrm{CP}\, x \approx 1)(\varphi(x) \mid \psi(x)), (\mathrm{CP}\, x \approx 1)(\psi(x) \mid \theta(x)), (\mathrm{CP}\, x \approx 1)(\neg\varphi(x) \mid \theta(x)),$$
$$(\mathrm{P}x > 0)\psi(x), (\mathrm{P}x > 0)\theta(x).$$

Since $\neg\psi(x) \to \varphi(x)$, there are six 'atoms' that should be considered:

$$\begin{aligned}
\alpha_1(x) &= \neg\theta(x) \wedge \neg\psi(x) \wedge \varphi(x),\\
\alpha_2(x) &= \neg\theta(x) \wedge \psi(x) \wedge \neg\varphi(x),\\
\alpha_3(x) &= \neg\theta(x) \wedge \psi(x) \wedge \varphi(x),\\
\alpha_4(x) &= \theta(x) \wedge \neg\psi(x) \wedge \varphi(x),\\
\alpha_5(x) &= \theta(x) \wedge \psi(x) \wedge \neg\varphi(x),\\
\alpha_6(x) &= \theta(x) \wedge \psi(x) \wedge \varphi(x).
\end{aligned}$$

It is easy to see that $\mathfrak{A} \models (\exists x_1) \cdots (\exists x_6) \bigwedge_{i=1}^{6} \alpha_i(x_i)$ ($\underline{i}$ is the witness for $\alpha_i(x_i)$, $i = 1, \ldots, 6$). The corresponding system is:

$$\begin{aligned}
& y_i \geqslant 0, i = 1, \ldots, 6,\ y_7 = y_8 = 0,\\
& \sum_{i=1}^{8} y_i = 1,\ y_2 + y_3 + y_5 + y_6 > 0,\ y_4 + y_5 + y_6 > 0,\\
& \frac{y_3 + y_6}{y_2 + y_3 + y_5 + y_6} \approx 1,\ \frac{y_5 + y_6}{y_4 + y_5 + y_6} \approx 1,\ \frac{y_5}{y_4 + y_5 + y_6} \approx 1.
\end{aligned}$$

The next step is the elimination of the sign $\approx$, and then the Fourier-Motzkin procedure which finishes with the true condition. Thus, the probabilistic part of Δ is satisfiable. Of course, using the results obtained during the solving of the system, we can define a measure on A. For example, the structure $(\mathfrak{A}, \mu)$ is a model for Δ, with the measure defined by: $\mu(\{\underline{1}\}) = \frac{1}{2}\varepsilon + \frac{3}{4}\varepsilon^2$, $\mu(\{\underline{2}\}) = \mu(\{\overline{2}\}) = \frac{1}{8}\varepsilon^2$, $\mu(\{\underline{3}\}) = \mu(\{\overline{3}\}) = \frac{1}{2} - \frac{3}{4}\varepsilon - \frac{1}{4}\varepsilon^2$, $\mu(\{\underline{4}\}) = \varepsilon^2$, $\mu(\{\underline{5}\}) = \mu(\{\overline{5}\}) = \frac{1}{2}\varepsilon - \frac{1}{2}\varepsilon^2$, $\mu(\{\underline{6}\}) = \mu(\{\overline{6}\}) = \frac{1}{2}\varepsilon^2$.

Example 5.6. Now we consider the set Δ containing the same classical formulas as in the previous example and the following probabilistic formulas:
$M(x,y) \rightsquigarrow_{xy} \neg R(x,y)$, and $M(x,y) \wedge T(x) \wedge T(y) \rightsquigarrow_{xy} R(x,y)$.

The following six 'atoms' should be now considered:

$$\alpha_1(x,y) = \neg(M(x,y) \wedge T(x) \wedge T(y)) \wedge \neg M(x,y) \wedge \neg R(x,y),$$
$$\alpha_2(x,y) = \neg(M(x,y) \wedge T(x) \wedge T(y)) \wedge \neg M(x,y) \wedge R(x,y),$$
$$\alpha_3(x,y) = \neg(M(x,y) \wedge T(x) \wedge T(y)) \wedge M(x,y) \wedge \neg R(x,y),$$
$$\alpha_4(x,y) = \neg(M(x,y) \wedge T(x) \wedge T(y)) \wedge M(x,y) \wedge R(x,y),$$
$$\alpha_5(x,y) = (M(x,y) \wedge T(x) \wedge T(y)) \wedge M(x,y) \wedge \neg R(x,y),$$
$$\alpha_6(x,y) = (M(x,y) \wedge T(x) \wedge T(y)) \wedge M(x,y) \wedge R(x,y).$$

If we perform the same procedure as in the previous example, we obtain the desired measure μ_2 on A^2 by choosing one solution of the corresponding system, for example, $y_8 = y_7 = 0$, $y_6 = \frac{\varepsilon}{2}$, $y_5 = \frac{\varepsilon^2}{4(1-\varepsilon)}$, $y_4 = \frac{\varepsilon}{4}$, $y_3 = \frac{1}{2}$, $y_2 = \frac{1}{4}$, $y_1 = \frac{1-4\varepsilon+2\varepsilon^2}{4(1-\varepsilon)}$.

Note that for each i, the value y_i can be understood as the basic belief mass given to α_i, i.e., as an amount of belief that supports the fact that the actual pair (c,d) 'belongs' to α_i.

6. Some extensions of the logic $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$

Adding qualitative quantifiers is crucial for AI applications. The list of logical symbols of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ can be extended with a range of new binary (qualitative) quantifiers. Some of them are $\leqslant_{\vec{x}}$, and $\trianglelefteq_{\vec{x}}$, with the intended meanings:

- $\mathfrak{A}, v \models \alpha \leqslant_{\vec{x}} \beta$ iff $\mu_n([\alpha]^{\vec{x}}_{\mathfrak{A},v}) \leqslant \mu_n([\beta]^{\vec{x}}_{\mathfrak{A},v})$;
- $\mathfrak{A}, v \models \alpha \trianglelefteq_{\vec{x}} \beta$ iff $\mu_n([\alpha \wedge \neg\beta]^{\vec{x}}_{\mathfrak{A},v}) \leqslant \mu_n([\alpha \wedge \beta]^{\vec{x}}_{\mathfrak{A},v})$, etc.

The corresponding axiomatic system can be obtained by using the ideas from Section 3. All new quantifiers can be described by infinitarly $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-formulas. For example,

$$\alpha \leqslant_{\vec{x}} \beta \Leftrightarrow \bigwedge_{q\in\mathbb{I}} (\mathrm{P}\vec{x} = q)\alpha \to (\mathrm{P}\vec{x} \geqslant q)\beta.$$

Having in mind this meta-equivalence, we add the following axiom schema:

$$\alpha \leqslant_{\vec{x}} \beta \to ((\mathrm{P}\vec{x} = q)\alpha \to (\mathrm{P}\vec{x} \geqslant q)\beta), q \in \mathbb{I}$$

and the rule:

$$\frac{\varphi \to ((\mathrm{P}\vec{x} = q)\alpha \to (\mathrm{P}\vec{x} \geqslant q)\beta),\ q \in \mathbb{I}}{\varphi \to \alpha \leqslant_{\vec{x}} \beta}$$

In a similar manner, we deal with the other quantifiers.

The axiomatic systems, obtained in this way, are sound with the class of $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$-models, and the proofs of the completeness theorems are analogous to the proof of the theorem 3.3.

As it can be seen from the previous section, the quantifiers $(\mathrm{CP}\cdot \approx 1)$ and $(\mathrm{CP}\cdot \approx 0)$ play a central role in modeling nonmonotonic reasoning. We can introduce new quantifiers of finer meanings then $(\mathrm{CP}\cdot \approx 0)$, and significantly increase the expressive power of the logic $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$. The idea is to add new quantifiers that will be able to express the order of an infinitesimal. We extend the logic $L^{\mathrm{P},\mathbb{I}}_{\omega\omega}$ with new quantifiers $(\mathrm{CP}\cdot\approx_k 0)$, $k\in\mathbb{N}$, whose meaning is given by the following clause:

$$\overline{\mathfrak{A}}, v \models (\mathrm{CP}\,\vec{x}\approx_k 0)(\alpha\mid\beta) \text{ iff } \mu_n([\beta]^{\vec{x}}_{\mathfrak{A},v})>0 \text{ and } \operatorname{ord}\left(\frac{\mu_n([\alpha]^{\vec{x}}_{\mathfrak{A},v}\wedge[\beta]^{\vec{x}}_{\mathfrak{A},v})}{\mu_n([\beta]^{\vec{x}}_{\mathfrak{A},v})}\right)\geqslant k.$$

Note that $(\mathrm{CP}\cdot\approx_0 0)$ and $(\mathrm{CP}\cdot\approx_1 0)$ are the quantifiers $(\mathrm{CP}\cdot\not\approx 0)$ and $(\mathrm{CP}\cdot\approx 0)$, respectively. It is not difficult to see that each $(\mathrm{CP}\cdot\approx_k 0)$, $k\geqslant 2$, can be introduced by an infinitary formula:

$$(\mathrm{CP}\,x\approx_k 0)(\alpha\mid\beta)\Leftrightarrow\bigvee_{r\in\mathbb{I}}(\mathrm{CP}\,x=r\varepsilon^k)(\alpha\mid\beta),$$

i.e.,

$$\neg(\mathrm{CP}\,x\approx_k 0)(\alpha\mid\beta)\Leftrightarrow\bigwedge_{r\in\mathbb{I}}(\mathrm{CP}\,x\neq r\varepsilon^k)(\alpha\mid\beta),$$

since the set of infinitesimals whose order is at least k can be defined by:

$$\{x\in\mathbb{Q}(\varepsilon)\mid(\exists r\in\mathbb{Q}(\varepsilon))\,x=r\varepsilon^k\}.$$

Having in mind this meta equivalence, we add the following axiom schema:

$$(\mathrm{CP}\,\vec{x}=r\varepsilon^k)(\alpha\mid\beta)\to(\mathrm{CP}\approx_k\vec{x})(\alpha\mid\beta),\ r\in\mathbb{I},$$

and the rule:

$$\frac{\gamma\to(\mathrm{CP}\,\vec{x}\neq r\varepsilon^k)(\alpha\mid\beta), r\in\mathbb{I}}{\gamma\to\neg(\mathrm{CP}\approx_k\vec{x})(\alpha\mid\beta)}.$$

We are now able to introduce new default relations between classical first-order formulas. Namely, for each $k\geqslant 1$ we have $\beta\rightsquigarrow^{\geqslant k}_{\vec{x}}\alpha$ and $\beta\rightsquigarrow^{=k}_{\vec{x}}\alpha$ denoting respectively

$$(\mathrm{CP}\,\vec{x}\approx_k 0)(\neg\alpha\mid\beta) \text{ and } (\mathrm{CP}\,\vec{x}=k)(\neg\alpha\mid\beta).$$

It is easy to prove that the relations $\rightsquigarrow^{\geqslant k}_{\vec{x}}$ and $\rightsquigarrow^{=k}_{\vec{x}}$ for each $n\geqslant 1$, satisfy all the rules of System **P**. Studying properties of these two relations shows that they may be useful in characterizations of various types of weakened monotonicity that are significant in nonmonotonic reasoning. We believe that more detailed research in this direction will be very fruitful.

Example 6.1. Let us consider again the set Δ introduced in Example 5.3, consisting of the following two defaults: $\texttt{Swede}(x)\rightsquigarrow_x\texttt{blond}(x)$ and $\texttt{Swede}(x)\rightsquigarrow_x\texttt{tall}(x)$. If we assume that the starting defaults are of different strength, we can avoid the inheritance blocking and obtain the desired conclusion. For example, from $\texttt{Swede}(x)\rightsquigarrow^{=1}_x\texttt{tall}(x)$ and $\texttt{Swede}(x)\rightsquigarrow^{=2}_x\texttt{blond}(x)$, it follows that $\texttt{Swede}(x)\wedge\neg\texttt{tall}(x)\rightsquigarrow_x\texttt{blond}(x)$.

Let us introduce one more binary quantifier $\prec_{\vec{x}}$ that may be interesting in formal treatment of Pearl's κ-calculus [94] with the meaning:

$$\mathfrak{A}, v\models\alpha\prec_{\vec{x}}\beta \text{ iff } \operatorname{ord}(\mu_n([\alpha]^{\vec{x}}_{\mathfrak{A},v}))<\operatorname{ord}(\mu_n([\beta]^{\vec{x}}_{\mathfrak{A},v})).$$

To facilitate the reading in what follows, we use the abbreviation $(\mathrm{O}\vec{x}=k)(\alpha \mid \beta)$ for $(\mathrm{CP}\, x \approx_k 0)(\alpha \mid \beta) \wedge \neg(\mathrm{CP}\, x \approx_{k+1} 0)(\alpha \mid \beta)$, and $(\mathrm{O}\vec{x}=k)\alpha$ for $(\mathrm{O}\vec{x}=k)$ $(\alpha \mid \top)$. The meta equivalence

$$\alpha \prec_{\vec{x}} \beta \Leftrightarrow \bigwedge_{k\in\mathbb{N}} ((\mathrm{O}\vec{x}=k)\alpha \rightarrow (\mathrm{P}\vec{x} \approx_{k+1} 0)\beta)$$

gives the following axiom schema:

$$\alpha \prec_{\vec{x}} \beta \rightarrow ((\mathrm{O}\vec{x}=k)\alpha \rightarrow (\mathrm{P}\vec{x} \approx_{k+1} 0)\beta),\ k\in\mathbb{N}$$

and the rule:

(κ) from the premises $\gamma \rightarrow ((\mathrm{O}\vec{x}=k)\alpha \rightarrow (\mathrm{P}\vec{x} \approx_{k+1} 0)\beta)$, $k\in\mathbb{N}$, infer $\gamma \rightarrow \alpha \prec_{\vec{x}} \beta$.

It is not hard to prove that the axiomatic system extended in this way is sound and complete with respect to the class of $L^{\mathrm{P};\mathbb{I}}_{\omega\omega}$-models. The proofs are straightforward modifications of the corresponding proofs for the logic $L^{\mathrm{P};\mathbb{I}}_{\omega\omega}$.

7. Conclusions

In this paper we described some first order logics with probabilistic quantifiers, provided a strongly complete axiomatization and proved the decidability of its two fragments which are expressive enough to enable different practical applications.

There are a lot of papers that are closely related to the subject of this paper [3–6,8,27,32,63,65,94,121,124]. General and detailed comparison of the corresponding logical systems would be very interesting and useful, but it is left for a future work.

Acknowledgement. The work was partially supported by the Serbian Ministry of Education and Science (projects 174026 and III044006).

References

[1] E. W. Adams, *The logic of Conditional*, Dordrecht, Reidel, 1975.

[2] M. Abadi, J. Y. Halpern, *Decidability and expressiveness for first-order logics of probability*, Inform. Comput. **112** (1994), 1–36.

[3] N. Alechina, *Logic with probabilistic operators*, In Proc. of the ACCOLADE '94 (1995), 121–138.

[4] F. Bacchus, *Lp, A Logic for Representing and Reasoning with Statistical Knowledge*, Comput. Intell. **6** (1990), 209–231.

[5] F. Bacchus, A. J. Grove, J. Y. Halpern, D. Koller, *From Statistical Knowledge Bases to Degrees of Belief*, Artif. Intell. **87**(1-2) (1996), 75–143.

[6] S. Benferhat, D. Dubois, H. Prade, *Nonmonotonic reasoning, conditional objects and possibility theory*, Artif. Intell. **92** (1997), 259–276.

[7] S. Benferhat, A. Saffiotti, P. Smets, *Belief functions and default reasoning*, Artif. Intell. **122** (2000), 1–69.

[8] V. Biazzo, A. Gilio, T. Lukasiewicz, G. Sanfilippo, *Probabilistic logic under coherence, model-theoretic probabilistic logic, and default reasoning in System P*, J. Appl. Non-Class. Log. **12**(2) (2002), 189–213.

[9] ———, *Probabilistic logic under coherence: Complexity and algorithms*, Ann. Math. Artif. Intell. **45**(1-2) (2005), 35–81.

[10] B. Boričić, M. Rašković, *A probabilistic validity measure in intuitionistic propositional logic*, Math. Balk. **10**(4) (1996), 365–372.

[11] E. Börger, E. Grädel, Y. Gurevich, *The Classical Decision Problem*, Springer, Berlin, Heidelberg, 1997.

[12] J. P. Burgess, *Quick completeness proofs for some logics of conditionals*, Notre Dame J. Formal Logic **22** (1981), 76–84.

[13] K. L. Clark, *Negation as failure*, in: H. Gallaire and J. Minker (eds.) *Logics and Databases*, Plenum Press, 1978, 293–322.

[14] D. Doder, J. Grant, Z. Ognjanović, *Probabilistic logics for objects located in space and time*, J. Log. Comput. **23**(3) (2013), 487–515.

[15] D. Doder, Z. Ognjanović, Z. Marković, *An Axiomatization of a First-order Branching Time Temporal Logic*, J. UCS **16**(11) (2010), 1439–1451.

[16] D. Doder, A. Perović, Z. Ognjanović, *Probabilistic approach to nonmonotonic consequence relations*, in: Weiru Liu (ed.), *The 11th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty ECSQARU 2011*, 29th June – 1st July 2011, Belfast, North Ireland, UK, LNCS 6717 (2011), 459–471.

[17] D. Doder, M. Rašković, Z. Marković, Z. Ognjanović, *Measures of inconsistency and defaults*, Int. J. Approx. Reasoning **51** (2010), 832–845.

[18] R. Đrđvić, *Barwise completeness theorems for logics with integrals*, Publ. Inst. Math., Nouv. Sér. **49**(63) (1991) 1–5.

[19] ———, *Analytic completeness theorem for absolutely continuous biprobability models*, Z. Math. Logik Grundlag. Math. **38** (1992), 241–246.

[20] ———, *Analytic completeness theorem for singular biprobability logic*, Math. Logic Quarterly **39** (1993), 228–230.

[21] ———, *Logics with two types of integral operators*, Publ. Inst. Math., Nouv. Sér. **54**(68) (1993), 18–24.

[22] R. Đrđvić, M. Rašković, Z. Ognjanović, *Completeness theorem for propositional probabilistic models whose measures have only finite ranges*, Arch. Math. Logic **43** (2004), 557–563.

[23] R. Fagin, J. Halpern, N. Megido, *A Logic For Reasoning About Probabilities*, Inf. Comput. **87**(1-2) (1990), 78–128.

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

[25] B. de Finetti, *La logique de la probabilié*, in: Actes du Congrès International de Philosophie Scientifique, Paris, Hermann: IV, 1936, 1–9.

[26] T. Flaminio, F. Montagna, *A logical and algebraic treatment of conditional probability*, in: L.A. Zadeh (ed.) *Proceedings of IPMU '04*, 2004, 493–500.

[27] N. Friedman, D. Koller, J. Halpern, *First-order conditional logic for default reasoning revisited*, ACM Trans. Comput. Log. **1**(2) (2000), 175–207.

[28] N. Friedman, J. Halpern, *Plausibility measures and default reasoning*, J. ACM **48**(6) (2001), 648–685.

[29] D. M. Gabbay, *Theoretical foundations for non-monotonic reasoning in expert systems*, in: K. R. Apt (eds.) *Proc. of the NATO Advanced Study Institute on Logics and Models of Concurrent Systems* La Colle-sur-Loup, France, 1985, 439–457.

[30] H. Gaifman, *Concerning measures in first order calculi*, Israel J. Math. **2** (1964), 1–18.

[31] P. Gillett, R. Scherl, G. Shafer, *Basing Probabilistic Logic on Gambles*, Fourth international symposium on imprecise probabilities and their applications, Carnegie Mellon University Pittsburgh, PA, USA, 2005.

[32] A. Gilio, *Probabilistic reasoning under coherence in System P*, Ann. Math. Artif. Intell. **34** (2002), 5–34.

[33] M. Goldszmidt, J. Pearl, *Qualitative probabilities for default reasoning, belief revision and causal modeling*, Artif. Intell. **84** (1996), 57–112.

[34] T. Hailperin, *Sentential Probability Logic, Origins, Development, Current Status, and Technical Applications*, Lehigh University Press, 1996.

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

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

[37] J. Halpern, *Lexicographic probability, conditional probability, and nonstandard probability*, Games Econ. Behav. **68**(1) (2010), 155–179.

[38] ———, *A Nonstandard Characterization of Sequential Equiliblium, Perfect Equilibrioum, and Proper Equilibrium*, Int. J. Game Theory **38**(1) (2009) 37–49

[39] P. Hammond, *Non-Archimedean Subjective probabilities in Decision Theory and Games*, Math. Soc. Sci. **38**(2) (1999) 139–156.

[40] J. Hintikka, *Knowledge and Belief - An Introduction to the Logic of the Two Notions*, Cornell University Press, 1962.

[41] W. van der Hoeck. *Some consideration on the logics P_FD*, J. Appl. Non-Class. Log. **7**(3) (1997), 287–307.

[42] D. N. Hoover, *Probability logic*, Ann. Math. Logic **14** (1978), 287–313.

[43] N. Ikodinović, *Craig interpolation theorem for classical propositional logic with some probability operators*, Publ. Inst. Math., Nouv. Sér. **69**(83) (2001), 27–33.

[44] N. Ikodinović, Z. Ognjanović, *A logic with coherent conditional probabilities*, in: L. Godo (ed.) *Proceedings of the 8th European Conference Symbolic and Quantitative Approaches to Reasoning with Uncertainty ECSQARU 2005*, Barcelona, Spain, July 6-8, 2005, Lecture Notes in Computer Science (Subseries: Lecture Notes in Artif. Intell.) 3571, Springer-Verlag, 2005, 726–736.

[45] N. Ikodinović, Z. Ognjanović, M. Rašković, A. Perović, *Hierarchies of probabilistic logics*, Int. J. Approx. Reasoning **55**(9) (2014), 1830–1842.

[46] N. Ikodinović, M. Rašković, Z. Marković, Z. Ognjanović, *Logics with Generalized Measure Operators*, J. Mult.-Val. Log. Soft Comput. **20**(5-6) (2013), 527–555.

[47] ———, *A first-order probabilistic logic with approximate conditional probabilities*, Log. J. IGPL **22**(4) (2014), 539–564.

[48] A. Ilić-Stepić, Z. Ognjanović, N. Ikodinović, A. Perović, *A p-adic probability logic*, Math. Log. Q. **58**(4-5) (2012), 263–280.

[49] A. Ilić-Stepić, Z. Ognjanović, *Complex valued probability logics*, Publ. Inst. Math., Nouv. Sér. **95**(109) (2014), 73–86.

[50] A. Ilić-Stepić, Z. Ognjanović, N. Ikodinović, *Conditional p-adic probability logic*, Int. J. Approx. Reasoning **55**(9) (2014), 1843–1865.

[51] A. Ilić-Stepić, Z. Ognjanović, *Logics for Reasoning About Processes of Thinking with Information Coded by p-adic Numbers*, Stud. Log. **103** (2015), 145–174.

[52] D. Jovanović, N. Mladenović, Z. Ognjanović, *Variable Neighborhood Search for the Probabilistic Satisfiability Problem*, in K. F. Doerner, M. Gendreau, P. Greistorfer, W. Gutjahr, R. F. Hartl, M. Reimann (eds.) *Metaheuristics: Progress in Complex Systems Optimization*, Springer Series: Operations Research/Computer Science Interfaces Series vol. 39, Springer, Berlin-NY, 2007, 173–188.

[53] J. Keisler. *Elementary calculus. An infinitesimal approach*, 2nd ed., Boston, Massachusetts: Prindle, Weber & Schmidt, 1986.

[54] S. Kraus, D. Lehmann, M. Magidor, *Nonmonotonic reasoning, preferential models and cumulative logics*, Artif. Intell. **44** (1990), 167–207.

[55] H. J. Keisler, *Model theory for infinitary logic*, North-Holland, Amsterdam, 1971.

[56] ———, *Hyperfinite model theory*, in: R. O. Gandy, J. M. E. Hyland (eds.) *Logic Colloquim 76*, North-Holland, 1977, 5–110.

[57] ———, *Probability quantifiers*, in: J. Barwise and S. Feferman (eds.) *Model Theoretic Logics*, Springer-Verlag, Berlin, 1985, 509–556.

[58] J. Keisler, *Elementary calculus. An infinitesimal approach*, 2nd ed., Boston, Massachusetts: Prindle, Weber & Schmidt, 1986.

[59] ———, *A completeness proof for adapted probability logic*, Ann. Pure Appl. Logic **31** (1986), 61–70.

[60] ———, *Hyperfinite models of adapted probability logic*, Ann. Pure Appl. Logic **31** (1986), 71–86.
[61] K. Knight, *Measuring inconsistency*, J. Philos. Log. **31** (2001), pp. 77–91.
[62] D. Lehmann, M. Magidor, *What does a conditional knowledge base entail?*, Artif. Intell. **55** (1992), 1–60.
[63] T. Lukasiewicz, *Probabilistic Default Reasoning with Conditional Constraints*, Ann. Math. Artif. Intell. **34** (2002), 35–88.
[64] ———, *Nonmonotonic probabilistic logics under variable-strength inheritance with overriding: Complexity, algorithms, and implementation*, Int. J. Approx. Reasoning **44**(3) (2007), 301–321.
[65] E. Marchioni, L. Godo, *A Logic for Reasoning about Coherent Conditional Probability: A Modal Fuzzy Logic Approach*, Proc. of the JELIA'04, Lecture notes in artificial intelligence 3229, Springer-Verlag, 2004, 213–225.
[66] Z. Marković, M. Rašković, Z. Ognjanović, *Completeness theorem for a logic with imprecise and conditional probabilities*, Publ. Inst. Math., Nouv. Sér. **78**(92) (2005), 35–49.
[67] Z. Marković, Z. Ognjanović, M. Rašković, *A probabilistic extension of intuitionistic logic*, Math. Log. Q. **49** (2003), 415–424.
[68] ———, *An intuitionistic logic with probabilistic operators*, Publ. Inst. Math., Nouv. Sér. **73**(87) (2003), 31 – 38.
[69] ———, *What is the Proper Propositional Base for Probabilistic Logic?* in: L.A. Zadeh (ed.) *Proceedings of the Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference IPMU 2004*, Perugia, Italy, July, 4-9, 2004, 443–450.
[70] D. McDermott, Jon Doyle, *Non-monotonic logic*, Artif. Intell. **13** (1980), 41–72.
[71] M. Milošević, Z. Ognjanović, *A first-order conditional probability logic*, Log. J. IGPL **20**(1) (2012), 235–253.
[72] ———, *A First-Order Conditional Probability Logic With Iterations*, Publ. Inst. Math., Nouv. Sér. **93**(107) (2013), 19–27.
[73] R. C. Moore, *Possible-world semantics for autoepistemic logic*, in: *Proceedings of AAAI Workshop on Nomonotonic Reasoning*, New Paltz, 1984, 369–401.
[74] N. Nilsson, *Probabilistic logic*, Artif. Intell., **28** (1986), 71–87.
[75] ———, *Probabilistic logic revisited*, Artif. Intell. **59** (1993), 39–42.
[76] Z. Ognjanović, M. Rašković, *A logic with higher order probabilities*, Publ. Inst. Math., Nouv. Sér. **60**(74) (1996), 1–4.
[77] Z. Ognjanović, *A logic for temporal and probabilistic reasoning*, Workshop on Probabilistic Logic and Randomised Computation ESSLLI '98, Saarbruecken, Germany, 1998.
[78] ———, *Neke verovatnosne logike i njihove primene u računarstvu*, PhD thesis, PMF Kragujevac, 1999.
[79] Z. Ognjanović, M. Rašković, *Some probability logics with new types of probability operators*, J. Log. Comput. **9**(2) (1999), 181–195.
[80] ———, *Some first-order probability logics*, Theor. Comput. Sci. **247**(1-2) (2000), 191–212.
[81] Z. Ognjanović, J. Kratica, M. Milovanović, *A genetic algorithm for satisfiability problem in a probabilistic logic: A first report*, Lecture notes in computer science (LNCS/LNAI), 2143, Springer-Verlag, 2001, 805–816.
[82] Z. Ognjanović, U. Midić, J. Kratica, *A genetic algorithm for probabilistic SAT problem*, in: L. Rutkowski, J. Siekmann, R. Tadeusiewicz, L.A. Zadeh (eds.) *Artif. Intell. and Soft Computing ICAISC 2004*, Zakopane, Poland, June 7 - 11, 2004, Lecture notes in artificial intelligence (LNCS/LNAI), 3070, Springer-Verlag, 2004, 462–467,
[83] Z. Ognjanović, Z. Marković, M. Rašković, *Completeness theorem for a Logic with imprecise and conditional probabilities*, Publ. Inst. Math., Nouv. Sér. **78**(92) (2005), 35–49.
[84] Z. Ognjanović, U. Midić, N. Mladenović, *A Hybrid Genetic and Variable Neighborhood Descent for Probabilistic SAT Problem*, in: M. J. Blesa (ed.) *Proceedings of the Second International Workshop on Hybrid Metaheuristics HM2005*, Barcelona, Spain, August 29-31, 2005, Lecture Notes in Computer Science 3636, Springer-Verlag, 2005, 42–53.

[85] Z. Ognjanović, N. Ikodinović, Z. Marković, *A logic with Kolmogorov style conditional probabilities*, in: *Proceedings of the 5th Panhellenic logic symposium*, Athens, Greece, July 25-28, 2005, 111–116.

[86] Z. Ognjanović, *Discrete Linear-time Probabilistic Logics: Completeness, Decidability and Complexity*, J. Log. Comput. **16**(2) (2006), 257–285.

[87] Z. Ognjanović, Z. Marković, M. Rašković, *Completeness Theorem for a Logic with Imprecise and Conditional Probabilities*, in: *Logic Colloquium '05, Athens, Greece, July 28 – August 3, 2005*, Bulletin of Symbolic Logic **12**(2) (2006), 357.

[88] Z. Ognjanović, A. Perović, M. Rašković, *Logics with the Qualitative Probability Operator*, Log. J. IGPL **16**(2) (2008), 105–120.

[89] Z. Ognjanović, D. Doder, Z. Marković, *A Branching Time Logic with Two Types of Probability Operators*, in: S. Benferhat and J. Grant (eds.) *Fifth International Conference on Scalable Uncertainty Management SUM-2011* October 10 - 12, 2011, Dayton, Ohio, USA, LNCS 6929 (2011), 219–232,

[90] Z. Ognjanović, Z. Marković, M. Rašković, D. Doder, A. Perović, *A Probabilistic Temporal Logic That Can Model Reasoning about Evidence*, Ann. Math. Artif. Intell. **65**(2-3) (2012), 217–243.

[91] Z. Ognjanović, A. Perović, D. Doder, *A First-Order Dynamic Probability Logic*, in: Linda C. van der Gaag (ed.) *The 12th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty ECSQARU 2013* Utrecht, The Netherlands, LNCS 7958 (2013), 461–472.

[92] Z. Ognjanović, M. Rašković, Z. Marković, *Probability logics*, Zb. Rad., Nova Ser. **12**(20) 2009, 35–111.

[93] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan & Kaufmann, San Mateo, CA, 1988.

[94] ———, *System Z: A natural ordering of defaults with tractable application to default reasoning*, Proc. of Theoretical Aspects of Reasoning about Knowledge (1990), 121–135.

[95] R. P. Perez, C. Uzcategui, *On representation theorems for nonmonotonic consequence relations*, J. Symb. Log. **65**(3) (2000), 1321–1337.

[96] A. Perović, D. Doder, Z. Ognjanović, *On Real-Valued Evaluation of Propositional Formulas*, in: T. Lukasiewicz and A. Sali (eds.) *Proceedings of the Seventh International Symposium on Foundations of Information and Knowledge Systems (FoIKS)* March 5-9, 2012 - Kiel, Germany, Lecture Notes in Computer Science 7153 (2012), 264–277.

[97] A. Perović, Z. Ognjanović, M. Rašković, Z. Marković, *Qualitative Possibilities and Necessities*, in: C. Sossai and G. Chemello (eds.) *The 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU 2009*. LNCS/LNAI volume 5590 (2009), 651 – 662.

[98] A. Perović, Z. Ognjanović, M. Rašković, D. Radojević, *Finitely additive probability measures on classical propositional formulas definable by Godel's t-norm and product t-norm*, Fuzzy Sets Syst. **169** (2011), 65–90.

[99] K. P. S. B. Rao, M. B. Rao, *Theory of Charges*, Academic Press, London, 1983.

[100] M. Rašković, *Model theory for L_{AM} logic*, Publ. Inst. Math., Nouv. Sér. **37**(51) (1985), 17–22.

[101] ———, *Completeness theorem for biprobability models*, J. Symb. Log. **51**(3) (1986), 586–590.

[102] M. Rašković, R. Živaljević, *Barwise completeness theorem for some biprobability logics* Z. Math. Logik Grundlagen Math. **31** (1986), 133–135.

[103] M. Rašković, *Weak completeness theorem for $L_{AP\forall}$*, Zb. Rad., Prir.-Mat. Fak., Univ. Kragujevcu **8** (1987), 69–72.

[104] ———, *Completeness theorem for singular biprobability models*, Proc. Am. Math. Soc. **102** (1988), 389–392.

[105] ———, *Completeness theorem for a monadic logic with both ordinary first-order and probability quantifiers*, Publ. Inst. Math., Nouv. Sér. **47**(61) (1990), 1–4.

[106] ______, *A completeness theorem for an infinitary intuitionistic logic with both ordinary and probability quantifiers*, Publ. Inst. Math., Nouv. Sér. **50**(64) (1991), 7–13.
[107] M. Rašković, R. Đrđvić, *Finite completeness theorem for biprobability logics*, Math. Balkanica **5** (1991), 12–14.
[108] ______, *Second order probability logic*, Math. Balkanica **6** (1992), 105–108.
[109] M. Rašković, *Classical logic with some probability operators*, Publications de l'Institut Mathématique, Ns. **53**(67) (1993), 1–3.
[110] M. Rašković, R. Đrđvić, *Continuous time probability logic*, Publ. Inst. Math., Nouv. Sér. **57**(71) (1995), 143–146.
[111] M. Rašković, R. Đrđvić, *Probability Quantifiers and Operators*, VESTA, Beograd, 1996.
[112] M. Rašković, R. Đrđvić, M. Bradić, *Weak cylindric probability algebras*, Publ. Inst. Math., Nouv. Sér. **61**(75) (1997), 6–16.
[113] M. Rašković, Z. Ognjanović, *A first order probability logic, LP_Q*, Publ. Inst. Math., Nouv. Sér. **65**(79) (1999), 1–7.
[114] M. Rašković, R. Đrđvić, *Cylindric probability algebras*, Publ. Inst. Math., Nouv. Sér. **68**(82) (2000), 20–36.
[115] M. Rašković, Z. Ognjanović, Z. Marković, *A Probabilistic Approach to Default Reasoning* in: *Proc. of the NMR '04* (2004), 335–341.
[116] M. Rašković, Z. Ognjanović, Z. Marković, *A Logic with Conditional Probabilities*, Proc. of the JELIA '04, LNCS 3229 (2004), 226–238.
[117] M. Rašković, Z. Marković, Z. Ognjanović, *A logic with approximate conditional probabilities that can model default reasoning*, Int. J. Approx. Reasoning **49**(1) (2008), 52–66.
[118] H. Reichenbach, *The Theory of Probability*, University of California Press, Berkeley-Los Angeles, 1949.
[119] R. Reiter, *A logic for default reasoning*, Artif. Intell. **13** (1980), 81–132.
[120] A. Robinson, *Non-standard analysis*, North-Holland, Amsterdam, 1966.
[121] K. Satoh, *A probabilistic interpretation for lazy nonmonotonic reasoning*, Proc. of the 8th American Conference on Artificial Intelligence 1990, 659–664.
[122] D. Scott, P. Krauss, *Assigning probabilities to logical formulas*, in: J. Hintikka, P. Suppes (eds), *Aspects of Inductive logic*, North-Holland Publishing Company, Amsterdam, 1966, 219–264.
[123] T. Stojanović, T. Davidović, Z. Ognjanović, *Bee Colony Optimization for the satisfiability problem in probabilistic logic*, Appl. Soft Comput. **31** (2015), 339–347.
[124] N. Wilson, *A Logic of Extended Probability*, First international symposium on imprecise probabilities and their applications, Ghent, Belgium, 1999.