DECIDABILITY AND UNDECIDABILITY IN PROBABILITY LOGIC

Sebastiaan A. Terwijn *

October 13, 2008

Abstract

We study computational aspects of a probabilistic logic based on a well-known model of induction by Valiant. We prove that for this paraconsistent logic the set of valid formulas is undecidable.

1 Introduction

The probabilistic interpretation of quantifiers has a long tradition and has been studied in many forms, often motivated by the difficulties of obtaining a complete picture of the world outside the realm of mathematical formalisms. We will not attempt to give an historical overview of the various approaches in the restricted context of this brief paper, but instead confine the discussion to those sources that are of direct relevance to it. More references to papers concerning probability logic can be found in [6].

Valiant [10] and Terwijn [7] gave probabilistic interpretations of first-order predicate logic based on Valiants model of pac-learning. In these interpretations universal quantification in a model $\mathcal{M}$ is interpreted as "many", where "many" refers to a given probability distribution $\mathcal{D}$ on $\mathcal{M}$ and to a given error parameter ε. These probabilistic interpretations were partly motivated by considerations from computational learning theory. In this paper our concern is not the induction of formulas but the study of probabilistic truth itself. Both [10] and [7] are predated by Keisler [5] (that also surveys many results of other researchers, notably Hoover), in which a logic is studied with essentially the same probabilistic interpretation of universal quantification, but with no other quantifiers, and with a negation that is different from the one below. Our different interpretation of negation allows for having the classical existential quantifier $\exists$ in the logic, something that Keislers logic does not have. A logic with a measure quantifier was also introduced by H. Friedman

*Institute of Logic, Language and Computation, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, the Netherlands, terwijn@logic.at. This research was supported by the Austrian Science Fund FWF under project P20346-N18.

(cf. Steinhorn in [1]), but this logic is even less related to ours than Keislers. A recent study of a probabilistic logic extending classical predicate logic that is motivated by inductive probabilistic reasoning is Jaeger [4].

We start by repeating the definition of probabilistic truth from [7]. For unexplained measure-theoretic terminology we refer to Doob [3]. Fix a first-order language with predicates and constants, possibly with equality, but no function symbols. Let $\mathcal{M}$ be a classical model (consisting of a universe with interpretations of the predicates in the language) and let $\mathcal{D}$ be a probability measure on $\mathcal{M}$, i.e. a measure such that $\mathcal{D}(\mathcal{M}) = 1$. For a $\mathcal{D}$-measurable subset $X \subseteq \mathcal{M}$ we will sometimes write $\Pr_{\mathcal{D}}\big[X\big]$ instead of $\mathcal{D}(X)$, to emphasize that we think of these measures as probabilities. For the moment we just assume that $\mathcal{D}$ is a probability measure. We will discuss an additional property that one can impose on $\mathcal{D}$ in Section 2.

Given a property $\varphi(x)$ for elements in a model $\mathcal{M}$, and given an error parameter ε, one can calculate (using Chernoff bounds, cf. [7]) how large a sample of x's from $\mathcal{M}$ should be to be able to assert with a certain confidence $1-\delta$ that at least $1-\varepsilon$ of the x's in $\mathcal{M}$ (in terms of the unknown measure $\mathcal{D}$) satisfy $\varphi(x)$. In the context of such large samples, we want $\forall x\varphi(x)$ to mean that this is the case, i.e. that at least $1-\varepsilon$ of the x's in $\mathcal{M}$ satisfy $\varphi(x)$. In contrast, we want $\exists x\varphi(x)$ to mean that an x was found in the sample that satisfies $\varphi(x)$. $\neg\exists x\varphi(x)$ should mean that no such x was found, which is the same as saying that $\forall x\neg\varphi(x)$. $\neg\forall x\varphi(x)$ means that not all sampled x's satisfy $\varphi(x)$, that is, the sample contains an x with $\neg\varphi(x)$, i.e. $\exists x\neg\varphi(x)$. These considerations are reflected in the following definition. Note that we do not model induction of formulas here; at this point we are solely interested in probabilistic truth.

Definition 1.1. (Truth definition) Given $\varepsilon \in [0,1]$, we inductively define the relation $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$ as follows.

1. For every prime formula φ (i.e. φ atomic or the negation of an atomic formula), $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$ if $\mathcal{M} \models \varphi$.
2. The logical connectives $\wedge$ and $\vee$ are treated classically, e.g. $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi \wedge \psi$ if it holds that $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$ and $\mathcal{M} \models_{\mathcal{D},\varepsilon} \psi$.
3. $\mathcal{M} \models_{\mathcal{D},\varepsilon} \exists x\varphi(x)$ if there exists $x \in \mathcal{M}$ such that $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(x)$.
4. The case of negation is split into subcases as follows:
   - 4.1. For φ atomic, $\mathcal{M} \models_{\mathcal{D},\varepsilon} \neg\neg\varphi$ if $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$. Furthermore, $\neg$ distributes in the classical way over $\vee$ and $\wedge$, e.g. $\mathcal{M} \models_{\mathcal{D},\varepsilon} \neg(\varphi \wedge \psi)$ if $\mathcal{M} \models_{\mathcal{D},\varepsilon} \neg\varphi \vee \neg\psi$.
   - 4.2. $\mathcal{M} \models_{\mathcal{D},\varepsilon} \neg\exists x\varphi(x)$ if $\mathcal{M} \models_{\mathcal{D},\varepsilon} \forall x\neg\varphi(x)$.
   - 4.3. $\mathcal{M} \models_{\mathcal{D},\varepsilon} \neg\forall x\varphi(x)$ if $\mathcal{M} \models_{\mathcal{D},\varepsilon} \exists x\neg\varphi(x)$.

5. $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi \to \psi$ if $\mathcal{M} \models_{\mathcal{D},\varepsilon} \neg\varphi \vee \psi$.

6. $\mathcal{M} \models_{\mathcal{D},\varepsilon} \forall x\varphi(x)$ if $\Pr_{\mathcal{D}}\big[x \in \mathcal{M} : \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(x)\big] \geqslant 1 - \varepsilon$.

Note that in the above definition everything is treated classically, except the interpretation of $\forall x\varphi(x)$ in Case 6. The treatment of negation requires some care, since we no longer have that $\mathcal{M} \models_{\mathcal{D},\varepsilon} \neg\varphi$ implies that $\mathcal{M} \not\models_{\mathcal{D},\varepsilon} \varphi$ (though the converse still holds).

Note that both $\mathcal{M} \models_{\mathcal{D},\varepsilon} \exists x\varphi(x)$ and $\mathcal{M} \models_{\mathcal{D},\varepsilon} \forall x\neg\varphi(x)$ may hold, since the interpretation of the first is the classical one, but the interpretation of the second is that *most* x's satisfy $\neg\varphi(x)$. That is, the logic of $\models_{\mathcal{D},\varepsilon}$ is *paraconsistent*. In [7] the above definition is motivated. In particular, the asymmetry in the interpretation of $\exists$ and $\forall$ is motivated by an interpretation in which the truth of first-order statements in an unknown model is established with a given degree of confidence by taking samples from the model.

Case 5 defines $A \to B$ as $\neg A \vee B$. We note that this is weaker than the classical implication. Namely, the classical definition would say that B holds in any model where A holds. Taking $B = \perp$, where $\perp$ is an inconsistency such as $\exists x(R(x)\wedge\neg R(x))$, we would thus obtain the classical negation of A. Taking for A an existential statement, then since $\exists$ expresses classical existence we would thus also obtain the classical universal quantifier $\forall$, and our logic would become a strong extension of classical predicate logic, which is not what we are after. Instead, $\exists x\varphi(x) \to \perp$ by definition means $\neg\exists x\varphi(x) \vee \perp$, which is the same as $\forall x\neg\varphi(x)$. Thus the above definition of implication takes on a probabilistic interpretation: If we interpret $\neg A$ by saying that A is *unlikely*, then $A \to B$ holds if whenever A holds it is *likely* that B holds.

Note that for $\varepsilon = 0$ the truth definition does not coincide with the classical one: If $\mathcal{M} \models_{\mathcal{D},0} \forall xR(x)$ there can still be a set of $\mathcal{D}$-measure zero of x's with $\neg R(x)$. *In the following we will exclude the pathological case of* $\varepsilon = 1$. Note that for $\varepsilon = 1$ all universal statements are always true, for example.

Proposition 1.2. (Prenex normal form) *Every formula φ is semantically equivalent to a formula φ' in prenex normal form, that is, φ' satisfies $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi \Leftrightarrow \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi'$ for all models $\mathcal{M}$, $\mathcal{D}$, ε.*

Proof. By Case 5 in Definition 1.1 we may assume that the formula is free of implications. Case 4 in the definition allows us to rewrite all formulas by pushing the negations inside, so that all negations occur only directly in front of an atomic formula. We then pull all quantifiers outside: Clearly we can pull $\exists$ outside over the connectives $\wedge$ and $\vee$ since $\exists$ has the classical meaning. For $\forall$ we have to check that

$$\varphi \wedge \forall x\psi(x) \equiv \forall x(\varphi \wedge \psi(x)) \tag{1}$$

and

$$\varphi \vee \forall x\psi(x) \equiv \forall x(\varphi \vee \psi(x)) \tag{2}$$

where $\equiv$ denotes semantic equivalence, and under the usual proviso about x not occurring free in φ. Indeed, for (1) we have

$$\begin{aligned}\mathcal{M} \models_{\mathcal{D},\varepsilon} \forall x\big(\varphi \wedge \psi(x)\big) &\iff \Pr_x\big[\varphi \wedge \psi(x)\big] \geqslant 1-\varepsilon\\ &\iff \varphi \wedge \Pr_x\big[\psi(x)\big] \geqslant 1-\varepsilon\\ &\iff \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi \wedge \forall x\psi(x).\end{aligned}$$

The second statement is proved in exactly the same way, replacing $\wedge$ by $\vee$. □

Definition 1.3. We will use the following terminology: By a *probabilistic model* we will mean a triple $\mathcal{M}, \mathcal{D}, \varepsilon$ as above, where $\varepsilon \in [0,1)$. In this case we also call the pair $\mathcal{M}$, $\mathcal{D}$ an *ε-model.* For a given ε, a sentence φ is *ε-satisfiable* if there are $\mathcal{M}$ and $\mathcal{D}$ such that $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$, and φ is *ε-valid* if $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$ for every $\mathcal{M}$ and $\mathcal{D}$. Furthermore, φ is *probabilistically satisfiable* if φ is ε-satisfiable for some $\varepsilon < 1$.

Note that all models are necessarily nonempty since they are measure spaces. From Proposition 1.2 it is easy to see that for $\varepsilon \leqslant \varepsilon'$, every ε-valid formula is ε'-valid. See also Proposition 3.3 below. In Section 2 we will amend Definition 1.3 by imposing an extra restriction on the probabilistic models.

Definition 1.4. Below we will use the shorthand notation $\vec{z}$ for a series of variables $z_1, \dots, z_n$. Let us adopt here the convention that for a formula $\varphi(\vec{z})$ with free variables $\vec{z}$ it holds that $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(\vec{z})$ whenever there are $\vec{z} \in \mathcal{M}$ such that $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(\vec{z})$. So we think of unbound variables as being existentially quantified.

2 The measurability of predicates

In Case 6 of Definition 1.1 we require in particular that the set

$$\big\{x \in \mathcal{M} : \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(x)\big\}$$

is $\mathcal{D}$-measurable. One can argue that it is natural to require a bit more than this, namely that

for every k-ary predicate R occurring in φ the set of k-tuples satisfying R is $\mathcal{D}^k$-measurable, (3)

where $\mathcal{D}^k$ denotes the product measure on $\mathcal{M}^k$. This is a natural assumption: When we are talking about probabilities over certain predicates we may as well require that all such probabilities exist, even if in some cases this would not be necessary. The property (3) and its consequences are discussed more extensively in [8]. *Henceforth, we will assume property (3).*

3 The set of 0-valid formulas

In this section we make some preliminary remarks about the set of 0-valid formulas. We start by repeating an easy preliminary result from [7].

Lemma 3.1. *Let $\mathcal{D}$ be a probability distribution on $\mathcal{M}$ such that for all $x \in \mathcal{M}$, $\mathcal{D}(\{x\}) \neq 0$. Then for every formula φ, $\mathcal{M} \models \varphi \iff \mathcal{M} \models_{\mathcal{D},0} \varphi$.*

Proof. One direction follows from the fact that classical validity implies probabilistic validity, since the only difference is that the probabilistic interpretation of $\forall$ is weaker. For the converse direction, if $\mathcal{D}$ is as in the lemma and $\Pr_{\mathcal{D}}\big[x \in \mathcal{M} : \mathcal{M} \models_{\mathcal{D},0} \varphi(x)\big] = 1$ then in fact $(\forall x \in \mathcal{M})\big[\mathcal{M} \models_{\mathcal{D},0} \varphi\big]$. So the interpretation of $\forall$ is in fact the classical one, and hence every formula is interpreted classically. □

Proposition 3.2. *The* 0*-valid formulas coincide with the classically valid formulas.*

Proof. That every classically valid formula is also probabilistically valid was already noted above. For the converse, suppose that φ is not classically valid. Then there is a countable model $\mathcal{M}$ such that $\mathcal{M} \not\models \varphi$. Since $\mathcal{M}$ is countable, there is a distribution $\mathcal{D}$ on $\mathcal{M}$ such that for all $x \in \mathcal{M}$, $\mathcal{D}(\{x\}) \neq 0$. But then by Lemma 3.1, $\mathcal{M} \not\models_{\mathcal{D},0} \varphi$. Hence φ is not 0-valid. □

Note however that we do not have that every 0-satisfiable sentence is classically satisfiable; a counterexample is $\exists x R(x) \wedge \forall x \neg R(x)$.

Proposition 3.3. (Terwijn [7]) *For all $\varepsilon < \varepsilon'$, the set of ε-valid formulas is strictly included in the set of ε'-valid ones.*[1]

Proposition 3.4. *Let $\varphi(\vec{x})$ be a formula with free variables $\vec{x}$ such that for every probabilistic model $\mathcal{M}$, $\mathcal{D}$ and every $\vec{x} \in \mathcal{M}$*

$$\forall \varepsilon > 0 \big(\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(\vec{x})\big) \Longrightarrow \mathcal{M} \models_{\mathcal{D},0} \varphi(\vec{x}). \tag{4}$$

If furthermore $\forall \vec{x} \varphi(\vec{x})$ is ε-valid for every $\varepsilon > 0$, then $\forall \vec{x} \varphi(\vec{x})$ is 0*-valid.*

Proof. By induction on the number of $\forall$-quantifiers it suffices to prove this

[1] The proof in [7] actually does not take the extra measurability condition (3) into consideration. However an alternative proof using similar ideas of this result can be given that also respects (3).

for $\forall x\varphi(x)$, where $\varphi(x)$ satisfies (4). So suppose $\varphi(x)$ satisfies (4). Then

$$\begin{aligned}
\forall\varepsilon > 0 \ \ \mathcal{M} \models_{\mathcal{D},\varepsilon} \forall x\varphi(x) \ &\Longrightarrow \ \forall\varepsilon > 0 \ \ \Pr_{\mathcal{D}}\big[x : \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(x)\big] \geqslant 1-\varepsilon \\
&\Longrightarrow \ \forall\varepsilon > 0 \ \ \Pr_{\mathcal{D}}\big[x : \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi(x)\big] = 1 \\
&\Longrightarrow \ \Pr_{\mathcal{D}}\Big(\bigcap_{n=2}^{\infty}\big\{x : \mathcal{M} \models_{\mathcal{D},\frac{1}{n}} \varphi(x)\big\}\Big) = 1 \\
&\Longrightarrow \ \Pr_{\mathcal{D}}\big[x : \mathcal{M} \models_{\mathcal{D},0} \varphi(x)\big] = 1 \\
&\Longrightarrow \ \mathcal{M} \models_{\mathcal{D},0} \forall x\varphi(x).
\end{aligned}$$

Here the second to last implication follows because φ satisfies (4). So if for every $\varepsilon > 0$ the sentence $\forall x\varphi(x)$ is ε-valid then it is 0-valid. □

Following standard notation, let $\forall^n\exists^m$ denote the class of $\mathcal{L}$-sentences in prenex form with at most n $\forall$-quantifiers followed by at most m $\exists$-quantifiers. Similarly, let $\exists^*$ and $\forall^*$ denote the fragments defined by any finite number of $\exists$ or $\forall$ quantifiers. Note that in contrast to the classical case, under the probabilistic interpretation we do *not* have that for example the $\forall^2$-fragment is closed under conjunctions. To see this, notice that the pair of formulas $\varphi_0 = \forall x\forall yRxy \wedge \forall x\forall y\neg Ryx$ and $\varphi_1 = \forall x\forall y(Rxy \wedge \neg Ryx)$ are not semantically equivalent. One can for example prove that whereas both formulas are $\frac{1}{3}$-satisfiable, φ_0 has a finite $\frac{1}{3}$-model, but φ_1 has not, cf. [8]. Hence, to put φ_0 in prenex form we have to rename variables and put more than two quantifiers in the prefix.

Corollary 3.5. *For every $\varphi \in \forall^*\exists^*$, if φ is not 0-valid then there is an $\varepsilon > 0$ such that φ is not ε-valid.*

Proof. It suffices to note that every formula φ of the form $\exists\vec{y}P(\vec{x},\vec{y})$, where P is a propositional combination of atomic predicates, satisfies (4). This is because if $\mathcal{M} \models_{\mathcal{D},\varepsilon} P(\vec{x},\vec{y})$ then $\mathcal{M} \models_{\mathcal{D},0} P(\vec{x},\vec{y})$. □

At this point we do not know whether Corollary 3.5 holds for *all* sentences φ, i.e. whether $\bigcap_{\varepsilon>0}\varepsilon$-valid = 0-valid.[2]

[2]Note however that the same proof will not work, since in general (4) fails for $\exists\forall$-formulas. Corollary 3.5 was used in an earlier version of the proof of Theorem 4.2, in which a reduction was built from the formulas in the $\forall^3\exists$-fragment. This fragment is undecidable by a result of Surányi [2, Theorem 3.1.16]

4 The undecidability of the ε-valid formulas

In this section we prove that the set of ε-valid formulas is undecidable for every ε. Note that for $\varepsilon = 0$ the set of ε-valid formulas coincides with the classically valid formulas by Proposition 3.2, and hence is Σ^0_1-complete.

Definition 4.1. Given a probabilistic model $\mathcal{M}, \mathcal{D}$ and a subset $X \subseteq \mathcal{M}$ with $\mathcal{D}(X) > 0$, we define the *restriction* of $\mathcal{M}$ to X, denoted by $\mathcal{M}\restriction X$, as the model with universe X obtained from $\mathcal{M}$ by restricting all relations to X, and with the probability distribution on X defined by multiplying $\mathcal{D}$ with a factor $1/\mathcal{D}(X)$.

Theorem 4.2. *For every rational $\varepsilon \in [0, 1)$, the set of ε-valid formulas is Σ^0_1-hard.*

Proof. Suppose that $0 < m < n$ and that $\varepsilon = 1 - \frac{m}{n}$. We build a many-one reduction from the 0-valid formulas to the ε-valid ones, i.e. we show that there is a computable function f such that φ is 0-valid if and only if $f(\varphi)$ is ε-valid. Note that this suffices since by Proposition 3.2 the 0-valid formulas coincide with the classically valid ones, and these are of course undecidable.

The idea of the proof is to introduce new parts $X_0, \dots X_{n-1}$ into a given model to "dilute" the meaning of the $\forall$-quantifiers in φ. We consider suitably relativized versions $\varphi^{X_{i_1} \cdots X_{i_m}}$ of φ relative to fractions $X_{i_1}, \dots, X_{i_m}$ of m out of the n X_i's. In $\varphi^{X_{i_1} \cdots X_{i_m}}$ the existential quantifiers range over X_{i_1} and $X_{i_2}, \dots, X_{i_m}$ are used to dilute the $\forall$ quantifiers in such a way that φ holds in X_{i_1} with error 0 if and only if $\varphi^{X_{i_1} \cdots X_{i_m}}$ holds in $X_{i_1} \cup \ldots \cup X_{i_m}$ with error 0. If $X_{i_1} \cup \ldots \cup X_{i_m}$ has weight $\geqslant \frac{m}{n}$ then the latter holds if and only if $\varphi^{X_{i_1} \cdots X_{i_m}}$ holds in $X_0 \cup \ldots \cup X_{n-1}$ with error at most ε. The main problem is to express all this correctly in such a way that a 0-countermodel to φ can be transformed to ε-countermodel of the relativized version. The proof below would be would be considerably simpler if we could express $\mathcal{D}(X) > \frac{m}{n}$. To circumvent this technical difficulty, we resort to considering all possible combinations $i_1 \dots i_m$.

Formally, given a first-order formula φ, let $X_0, \dots, X_{n-1}$ be fresh unary predicates, i.e. predicates not occurring in φ. Define the sentence

$$n\text{-}\mathsf{split} = \forall x\Big(\big(X_0(x) \vee \ldots \vee X_{n-1}(x)\big) \wedge \bigwedge_{i<n} \Big[X_i(x) \leftrightarrow \bigwedge_{j \neq i} \neg X_j(x)\Big]\Big)$$

that says that $\mathcal{M}$ splits into n parts. Note that since n-split is purely universal we have that $\mathcal{M}$ probabilistically satisfies $\neg n$-split if and only if it satisfies it classically, hence if $\mathcal{M} \not\models_{\mathcal{D},\varepsilon} \neg n$-$\mathsf{split}$ then really $\mathcal{M} \models n$-split classically. In the following we use set-theoretic notation such as $x \in X_0 \cup \ldots \cup X_{n-1}$ as a shorthand for the formula $X_0(x) \vee \ldots \vee X_{n-1}(x)$. We also write $\mathcal{D}(X) \geqslant \frac{m}{n}$ for the sentence $\forall x(X(x))$. (Note that since $\varepsilon = 1 - \frac{m}{n}$ this last sentence expresses precisely this fact.) Given a sentence φ and a choice $i_1, \dots, i_m$ of

m different numbers from the set $\{0,\ldots,n-1\}$, define a relativized version $\varphi^{X_{i_1}\cdots X_{i_m}}$ of φ by recursively replacing $\exists x$ everywhere by $\exists x(X_{i_1}(x)\wedge\ldots)$ and all occurrences of $\forall x$ by $\forall x\big(x\in X_{i_2}\cup\ldots\cup X_{i_m}\vee(x\in X_{i_1}\wedge\ldots)\big)$.[3] For every φ define

$$f(\varphi)=\neg n\text{-}\mathsf{split}\vee\bigvee_{i_1\ldots i_m}\Big(\mathcal{D}(X_{i_1}\cup\ldots\cup X_{i_m})\geqslant\tfrac{m}{n}\wedge\varphi^{X_{i_1}\cdots X_{i_m}}\Big).$$

Here the disjunction is over all choices $i_1\ldots i_m$ of m different numbers from the set $\{0,\ldots,n-1\}$. Now if φ is 0-valid and $\mathcal{M}\not\models_{\mathcal{D},\varepsilon}\neg n\text{-}\mathsf{split}$ then $\mathcal{M}$ splits into $X_0\ldots X_{n-1}$. By Lemma 4.3 there is always a choice of $i_1\ldots i_m$ such that $\mathcal{D}(X_{i_1}\cup\ldots\cup X_{i_m})\geqslant\frac{m}{n}$. Without loss of generality $\mathcal{D}(X_{i_1})>0$, for if this does not hold we can permute $i_1\ldots i_m$. But then by Lemma 4.4 we have that $\mathcal{M}\models_{\mathcal{D},\varepsilon}\varphi^{X_{i_1}\cdots X_{i_m}}$. Hence $f(\varphi)$ is ε-valid.

Conversely, suppose that φ is not 0-valid, say that $\mathcal{M}\not\models_{\mathcal{D},0}\varphi$. We show that there is a model $\mathcal{M}'$,$\mathcal{D}'$ such that $\mathcal{M}'\not\models_{\mathcal{D}',\varepsilon}f(\varphi)$. Let $\mathcal{M}'$ consist of the n disjoint parts $X_0,\ldots,X_{n-1}$, where each X_i is a copy of $\mathcal{M}$ where in addition every element satisfies the unary predicate X_i. The predicates on $\mathcal{M}'$ are defined exactly as in $\mathcal{M}$ within each given copy X_i, and are defined arbitrarily across different copies. Under $\mathcal{D}'$ we give each of $X_0\ldots X_{n-1}$ weight $\frac{1}{n}$. The structure of $\mathcal{D}'$ on each X_i is like $\mathcal{D}$ on $\mathcal{M}$, multiplied with the factor $\frac{1}{n}$, that is, $\mathcal{D}'$ is the sum of n copies of $\frac{1}{n}\cdot\mathcal{D}$. Notice that by definition $\mathcal{M}'$ does not ε-satisfy $\neg n\text{-}\mathsf{split}$, and that it ε-satisfies $\mathcal{D}(X_{i_1}\cup\ldots\cup X_{i_m})\geqslant\frac{m}{n}$ for any choice $i_1\ldots i_m$ of m different numbers from $\{0,\ldots,n-1\}$. Given any such choice $i_1\ldots i_m$, let $\mathcal{M}'{\restriction}X_{i_1}\cup\ldots\cup X_{i_m}$, $\mathcal{D}''$ be the restriction of $\mathcal{M}'$ to $X_{i_1}\cup\ldots\cup X_{i_m}$. (Cf. Definition 4.1.) So $\mathcal{D}''$ on $X_{i_1}\cup\ldots\cup X_{i_m}$ is $\mathcal{D}'$ multiplied with $1/\mathcal{D}'(X_{i_1}\cup\ldots\cup X_{i_m})=\frac{n}{m}$. Now suppose that $\mathcal{M}'\models_{\mathcal{D}',\varepsilon}\varphi^{X_{i_1}\cdots X_{i_m}}$. Then by Lemma 4.5, $\mathcal{M}'{\restriction}X_{i_1}\cup\ldots\cup X_{i_m}\models_{\mathcal{D}'',0}\varphi^{X_{i_1}\cdots X_{i_m}}$. But since X_{i_1} is a copy of $\mathcal{M}$, this easily implies $\mathcal{M}\models_{\mathcal{D},0}\varphi$: By definition of $\varphi^{X_{i_1}\cdots X_{i_m}}$, witnesses for existential quantifiers can be found in X_{i_1}, and universal quantifiers hold with $\mathcal{D}''$-measure 1 in $\mathcal{M}'{\restriction}X_{i_1}\cup\ldots\cup X_{i_m}$, hence also with $\mathcal{D}$-measure 1 in X_{i_1}. Thus we have $\mathcal{M}\models_{\mathcal{D},0}\varphi$, contrary to the assumption. Hence $\mathcal{M}'$ also does not ε-satisfy $\varphi^{X_{i_1}\cdots X_{i_m}}$, and thus $\mathcal{M}'$ witnesses that $f(\varphi)$ is not ε-valid. □

Lemma 4.3. *Suppose that $n\geqslant m\geqslant 1$ and that $a_i\in\mathbb{R}$ are such that*

$$\sum_{i=0}^{n-1}a_i=1.$$

Then there are m a_i's such that their sum is greater than or equal to $\frac{m}{n}$.

Proof. The average over the a_i is $\frac{1}{n}$, so the m largest of them sum up to at least $\frac{m}{n}$. □

[3] If $m=1$ then we replace $\forall x$ by $\forall x(x\in X_{i_1}\wedge\ldots)$.

Lemma 4.4. *In the proof of Theorem 4.2, suppose that $\mathcal{M} \not\models_{\mathcal{D},\varepsilon} \neg n$-*$\mathsf{split}$*, that $\mathcal{D}(X_{i_1} \cup \ldots \cup X_{i_m}) \geqslant \frac{m}{n}$, and that $\mathcal{D}(X_{i_1}) > 0$. Then $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi^{X_{i_1}\ldots X_{i_m}}$.*

Proof. This follows because φ is 0-valid, hence it holds with error 0 in $\mathcal{M}\restriction X_{i_1}$, which is defined since $\mathcal{D}(X_{i_1}) > 0$ (cf. Definition 4.1). Loosely speaking, $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi^{X_{i_1}\ldots X_{i_m}}$ holds because witnesses for the existential quantifiers can be found in X_{i_1} since φ holds in $\mathcal{M}\restriction X_{i_1}$, and the universal quantifiers in $\varphi^{X_{i_1}\ldots X_{i_m}}$ are satisfied since $\mathcal{D}(X_{i_1} \cup \ldots \cup X_{i_m}) \geqslant \frac{m}{n}$ and the error on X_{i_1} is 0. More formally, the lemma follows from the following claim. Let $\mathcal{D}'$ denote the distribution on $\mathcal{M}\restriction X_{i_1}$.
Claim. If $\mathcal{M}\restriction X_{i_1} \models_{\mathcal{D}',0} \varphi$ then $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi^{X_{i_1}\ldots X_{i_m}}$. The claim is proved by formula induction. By Proposition 1.2 we may assume that φ is in prenex normal form and that all negations occur directly in front of atomic predicates. The induction step for $\exists$ is trivial by definition of $\varphi^{X_{i_1}\ldots X_{i_m}}$, so the only case that requires attention is the induction step for $\forall$. So suppose that $\varphi = \forall x\psi(x)$. Then

$$\varphi^{X_{i_1}\ldots X_{i_m}} = \forall x\big(x \in X_{i_2} \cup \ldots \cup X_{i_m} \vee \big(x \in X_{i_1} \wedge \psi(x)^{X_{i_1}\ldots X_{i_m}}\big)\big). \tag{5}$$

If $\mathcal{M}\restriction X_{i_1} \models_{\mathcal{D}',0} \varphi$ then

$$\Pr_{\mathcal{D}'}\left[x \in X_{i_1} : \mathcal{M}\restriction X_{i_1} \models_{\mathcal{D}',0} \psi(x)\right] \geqslant 1.$$

By induction hypothesis, for every $x \in X_{i_1}$ with $\mathcal{M}\restriction X_{i_1} \models_{\mathcal{D}',0} \psi(x)$ we have $\mathcal{M} \models_{\mathcal{D},\varepsilon} \psi(x)^{X_{i_1}\ldots X_{i_m}}$. It then follows from $\mathcal{D}(X_{i_1} \cup \ldots \cup X_{i_m}) \geqslant \frac{m}{n}$ that

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} : x \in X_{i_2} \cup \ldots \cup X_{i_m} \vee \big(x \in X_{i_1} \wedge \psi(x)^{X_{i_1}\ldots X_{i_m}}\big)\big] \geqslant \tfrac{m}{n} = 1 - \varepsilon$$

hence $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi^{X_{i_1}\ldots X_{i_m}}$. □

Lemma 4.5. *In the proof of Theorem 4.2 we have that*

$$\mathcal{M}' \models_{\mathcal{D}',\varepsilon} \varphi^{X_{i_1}\ldots X_{i_m}} \Longrightarrow \mathcal{M}'\restriction X_{i_1} \cup \ldots \cup X_{i_m} \models_{\mathcal{D}'',0} \varphi^{X_{i_1}\ldots X_{i_m}}.$$

Proof. Again we prove this by induction on φ. By Proposition 1.2 we may assume that φ is in prenex normal form and that all negations occur directly in front of atomic predicates. Then all the steps in the induction are trivial, except the case of universal quantification. So suppose that $\varphi = \forall x\psi(x)$. Then $\varphi^{X_{i_1}\ldots X_{i_m}}$ is of the form (5). Denoting $\psi(x)^{X_{i_1}\ldots X_{i_m}}$ by $\hat{\psi}(x)$ we then have

$$
\begin{aligned}
&\mathcal{M}' \models_{\mathcal{D}',\varepsilon} \varphi^{X_{i_1}\ldots X_{i_m}} \quad \Longrightarrow \\
&\Pr_{\mathcal{D}}\big[x \in \mathcal{M}' : x \in X_{i_2} \cup \ldots \cup X_{i_m} \vee \big(x \in X_{i_1} \wedge \mathcal{M}' \models_{\mathcal{D}',\varepsilon} \hat{\psi}(x)\big)\big] \geqslant \tfrac{m}{n} \quad \Longrightarrow \\
&\Pr_{\mathcal{D}}\big[x \in \mathcal{M}' : x \in X_{i_2} \cup \ldots \cup X_{i_m} \vee \\
&\qquad \big(x \in X_{i_1} \wedge \mathcal{M}'{\restriction}X_{i_1} \cup \ldots \cup X_{i_m} \models_{\mathcal{D}'',0} \hat{\psi}(x)\big)\big] \geqslant \tfrac{m}{n} \quad \Longrightarrow \\
&\Pr_{\mathcal{D}''}\big[x \in X_{i_1} \cup \ldots \cup X_{i_m} : x \in X_{i_2} \cup \ldots \cup X_{i_m} \vee \\
&\qquad \big(x \in X_{i_1} \wedge \mathcal{M}'{\restriction}X_{i_1} \cup \ldots \cup X_{i_m} \models_{\mathcal{D}'',0} \hat{\psi}(x)\big)\big] \geqslant \\
&\qquad\qquad \frac{\frac{m}{n}}{\mathcal{D}(X_{i_1} \cup \ldots \cup X_{i_m})} = 1 \\
&\Longrightarrow \quad \mathcal{M}'{\restriction}X_{i_1} \cup \ldots \cup X_{i_m} \models_{\mathcal{D}'',0} \varphi^{X_{i_1}\ldots X_{i_m}}.
\end{aligned}
$$

Here the second implication follows by the induction hypothesis. □

5 Finite models and decidability

It is shown in [8] that the downward Löwenheim-Skolem theorem fails for ε-logic: Not every infinitely ε-satisfiable sentence has a countable model. The next result shows that countable probabilistic models are in a way analogous to classical finite models:

Theorem 5.1. *Let φ be a sentence. Then*

$$\forall \mathcal{M} \text{ finite } \mathcal{M} \models \varphi \iff \forall \mathcal{M} \text{ countable} \, \forall \mathcal{D} \, \forall \varepsilon > 0 \;\, \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi.$$

Proof. (If) If $\mathcal{M}$ is finite and $\forall \varepsilon > 0 \;\, \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$ then classically $\mathcal{M} \models \varphi$: If $\mathcal{M}$ has n elements then take $\mathcal{D}$ the uniform distribution on $\mathcal{M}$ assigning to every element probability $\frac{1}{n}$ and take $\varepsilon < \frac{1}{n}$. Then there can be no exceptions to $\forall$-statements.

(Only if) The idea is simply that if $\mathcal{M}$ is countable then most of the weight under $\mathcal{D}$ is concentrated on *finitely many* elements of $\mathcal{M}$. If φ holds classically in all finite models, φ also holds on these finitely many elements. More precisely; Fix $\varepsilon > 0$ and a countable probabilistic model $\mathcal{M}$, $\mathcal{D}$, and suppose that φ is classically valid on all finite models. Let $\mathcal{M}' \subseteq \mathcal{M}$ be finite such that $\mathcal{M}'$ has weight at least $1 - \varepsilon$ under $\mathcal{D}$. Since $\mathcal{M}'$ is finite we have $\mathcal{M}' \models \varphi$. But then also $\mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi$, since clearly all existential quantifications from φ are satisfied within $\mathcal{M}$, and all universal quantifications have at most ε exceptions in weight. □

Notice that it is essential that in Theorem 5.1 we exclude the case $\varepsilon = 0$, since otherwise by Lemma 3.1 we would obtain all classical validities instead of only the finitely valid sentences.

Corollary 5.2. *The set*

$$\{\varphi : \forall \mathcal{M} \, \text{countable} \, \forall \mathcal{D} \forall \varepsilon > 0 \;\; \mathcal{M} \models_{\mathcal{D},\varepsilon} \varphi\}$$

is Π^0_1*-complete.*

Proof. By Trakhtenbrot's theorem [9] (a result that was independently obtained by Craig) the set $\{\varphi : \forall \mathcal{M} \, \text{finite} \;\; \mathcal{M} \models \varphi\}$ of finitely valid first-order sentences is Π^0_1-complete. □

In Terwijn [8] it is proven that for fixed ε we do not have the finite model property: There are ε-satisfiable sentences without a finite ε-model. (Cf. the examples quoted on page 6.) Nevertheless, we make the following

Conjecture 5.3. *For rational* $\varepsilon \in [0, 1)$*, it is decidable whether* φ *is* ε*-satisfiable.*

Note that a positive answer to Conjecture 5.3 does not contradict the undecidability from Theorem 4.2 because, in contrast to the classical case, under the probabilistic interpretation we do not have that φ is valid if and only if $\neg\varphi$ is not satisfiable, even if $\varepsilon = 0$. For example the sentence $\varphi = \exists x R(x) \wedge \forall x \neg R(x)$ is probabilistically satisfiable, but its negation $\neg\varphi$ is $\forall x \neg R(x) \vee \exists x R(x)$, which is even classically valid.

References

[1] J. Barwise and S. Feferman (eds.), *Model-theoretic logics*, Springer-Verlag 1985.

[2] E. Börger, E. Grädel, Y. Gurevich, *The classical decision problem*, Springer, 1997.

[3] J. L. Doob, *Measure theory*, Springer, 1994.

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

[5] H. J. Keisler, *Probability quantifiers*, in: [1], 509–556.

[6] Probability logic papers, database, problog.mi.sanu.ac.yu/index.html.

[7] S. A. Terwijn, *Probabilistic logic and induction*, Journal of Logic and Computation 15(4) (2005) 507–515.

[8] S. A. Terwijn, *Model-theoretic aspects of probability logic*, in preparation.

[9] B. Trakhtenbrot, *The impossibility of an algorithm for the decision problem for finite models*, Dokl. Akad. Nauk SSSR 70 (1950) 596–572. English translation in AMS Transl. Ser. 2 (1963) 1–6.

[10] L. G. Valiant, *Robust logics*, Artificial Intelligence 117 (2000) 231–253.