COMPUTATIONAL HARDNESS OF VALIDITY IN PROBABILITY LOGIC

RUTGER KUYPER

ABSTRACT. We consider the complexity of validity in ε-logic, a probability logic introduced by Terwijn. We prove that the set of valid formulas is Π^1_1-hard, improving a previous undecidability result by Terwijn.

1. INTRODUCTION

Over the years, there have been many attempts at combining logic and probability through so-called *probability logics*. We study the computational aspects of a probability logic introduced by Terwijn in [4]. This logic has two main characteristics whose combination sets it apart from earlier attempts: first, the logic is closely related to probabilistic induction and Valiant's pac-model; and second, it is a probabilistic interpretation of first-order logic instead of a probability logic with an entirely new syntax. Terwijn's probability logic depends on a fixed error parameter ε and is hence called *ε-logic*.

Valiant [6] also introduced a probability logic related to his pac-model. Nonetheless, the logic most closely related to ε-logic is the logic $\mathcal{L}_{\omega P}$ introduced by Keisler, surveyed in Keisler [2]. This logic contains quantifiers of the form $(Px \geq r)$ which should be read as "holds for at least measure r many x". However, Keisler's logic does not contain the classical universal and existential quantifiers and does not attempt to model probabilistic induction in any way. Nevertheless, it turns out we can adapt some of the ideas used to prove results about $\mathcal{L}_{\omega P}$ to obtain similar results for ε-logic. For a discussion of more probability logics related to ours, we refer to the introduction of Kuyper and Terwijn [3].

Unfortunately, it turns out that ε-logic is computationally quite hard. Previously, Terwijn [5] has shown that the set of ε-tautologies is undecidable. This should not be too surprising, because this is of course also the case for classical first-order logic. For $\mathcal{L}_{\omega P}$, Hoover [1] has shown that validity is Π^1_1-complete (i.e. of the same complexity as first-order arithmetic with second-order universal quantifiers), which is computationally much harder than first-order logic. In this paper, we will

Research supported by the Dutch mathematics research cluster DIAMANT and by the John Templeton Foundation Grant ID#15619: 'Mind, Mechanism and Mathematics: Turing Centenary Research Project'.

combine some of his ideas with our own to show that ε-logic is Π^1_1-hard. This shows that ε-logic is computationally much harder than first-order logic and that we cannot hope to find an effective calculus for it.

In the next section we will briefly recall the definition of ε-logic and some elementary facts. After that, we will prove in section 3 that there exist many-one reductions between ε_0-logic and ε_1-logic for $\varepsilon_0 \neq \varepsilon_1$. Finally, in section 4 we will prove that ε-logic is Π^1_1-hard by utilising these reductions.

2. ε-Logic

As mentioned above, ε-logic was introduced in Terwijn [4]. Afterwards, the definition has gone through a few minor modifications. We use the definition from Kuyper and Terwijn [3]. For a discussion of this definition and more information about ε-logic, we refer to the same paper.

Definition 2.1. Let $\mathcal{L}$ be a first-order language, possibly containing equality, of a countable signature. Let $\varphi = \varphi(x_1, \dots, x_n)$ be a first-order formula in the language $\mathcal{L}$, and let $\varepsilon \in [0,1]$. Furthermore, let $\mathcal{M}$ be a classical first-order model for $\mathcal{M}$ and let $\mathcal{D}$ be a probability measure on $\mathcal{M}$. Then we inductively define the notion of *ε-truth*, denoted by $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$, as follows (where we leave the parameters implicit).

(i) For every atomic formula φ:

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi \text{ if } \mathcal{M} \models \varphi.$$

(ii) We treat the logical connectives $\wedge$ and $\vee$ classically, e.g.

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi \wedge \psi \text{ if } (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi \text{ and } (\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi.$$

(iii) The existential quantifier is treated classically as well:

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \exists x \varphi(x)$$

if there exists an $a \in \mathcal{M}$ such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a)$.

(iv) The case of negation is split into sub-cases as follows:

(a) For φ atomic, $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg\varphi$ if $(\mathcal{M}, \mathcal{D}) \not\models_\varepsilon \varphi$.

(b) $\neg$ distributes in the classical way over $\wedge$ and $\vee$, e.g.

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg(\varphi \wedge \psi) \text{ if } (\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg\varphi \vee \neg\psi.$$

(c) $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg\neg\varphi$ if $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$.

(d) $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg(\varphi \to \psi)$ if $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi \wedge \neg\psi$.

(e) $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg\exists x \varphi(x)$ if $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \forall x \neg\varphi(x)$.

(f) $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg\forall x \varphi(x)$ if $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \exists x \neg\varphi(x)$.

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

(vi) Finally, we define $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \forall x \varphi(x)$ if

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a)\big] \geq 1 - \varepsilon.$$

Thus, the crucial change is that the universal quantifier is not treated classically: instead of saying that we have $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a)$ for *all* elements $a \in \mathcal{M}$, we merely say that it holds for "many" of the elements, where "many" depends on the error parameter ε.

The main reason for this change is that we want our logic to be *learnable*, in the sense defined in Terwijn [4] (whose definition of learning is closely related to Valiant's pac-model). We do not want to add the classical universal quantifier to our logic, since it is impossible to decide if a universal quantifier holds from just a finite amount of information. Therefore we take special care in defining our negation: we do not want $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg\exists x \varphi(x)$ to mean $(\mathcal{M}, \mathcal{D}) \not\models_\varepsilon \exists x \varphi(x)$, because the latter is equivalent to saying that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(x)$ holds *classically* for all $x \in \mathcal{M}$, which is exactly what we wanted to avoid. We define our negation in such a way that it still behaves in a classical way on the propositional level, while it interchanges the existential and universal quantifiers.

One other important point is that both $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \forall x \varphi(x)$ and $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \exists x \neg\varphi(x)$ may hold simultaneously, i.e. both a formula and its negation might hold at the same time. Thus, ε-logic is *paraconsistent*. For our current work this has one important implication: it is no longer the case that φ is satisfiable if and only if its negation $\neg\varphi$ is not a tautology, as demonstrated in Example 2.4 below.

To make sure that all necessary sets are measurable, we need to restrict ourselves to the right set of models.

Definition 2.2. Let $\mathcal{L}$ be a first-order language of a countable signature, possibly containing equality, and let $\varepsilon \in [0, 1]$. Then an *ε-model* $(\mathcal{M}, \mathcal{D})$ for the language $\mathcal{L}$ consists of a classical first-order $\mathcal{L}$-model $\mathcal{M}$ together with a probability distribution $\mathcal{D}$ over $\mathcal{M}$ such that:

(1) For all formulas $\varphi = \varphi(x_1, \dots, x_n)$ and all $a_1, \dots, a_{n-1} \in \mathcal{M}$, the set

$$\{a_n \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \dots, a_n)\}$$

is $\mathcal{D}$-measurable (i.e. all definable sets of dimension 1 are measurable).

(2) All relations of arity n are $\mathcal{D}^n$-measurable (including equality, if it is in $\mathcal{L}$) and all functions of arity n are measurable as functions from $(\mathcal{M}^n, \mathcal{D}^n)$ to $(\mathcal{M}, \mathcal{D})$ (where $\mathcal{D}^n$ denotes the n-fold product measure). In particular, constants are $\mathcal{D}$-measurable.

A *probability model* is a pair $(\mathcal{M}, \mathcal{D})$ that is an ε-model for every $\varepsilon \in [0, 1]$.

Definition 2.3. A formula $\varphi(x_1, \dots, x_n)$ is an *ε-tautology* or is *ε-valid* (notation: $\models_\varepsilon \varphi$) if for all probability models $(\mathcal{M}, \mathcal{D})$ and all $a_1, \dots, a_n \in \mathcal{M}$ it holds that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \dots, a_n)$. Similarly, we say that φ is *ε-satisfiable* if there exists a probability model $(\mathcal{M}, \mathcal{D})$ and there exist $a_1, \dots, a_n \in \mathcal{M}$ such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$.

We remark that, for satisfiability, it would be equivalent to use the more satisfying definition of letting φ be ε-satisfiable if there exists just an ε-model $(\mathcal{M}, \mathcal{D})$ such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$, as follows directly from Kuyper and Terwijn [3, Proposition 5.1 and Theorem 5.9]. Unfortunately, we do not know of a similar result for validity. Our proof below needs our models to be probability models, hence our choice for this definition of an ε-tautology.

Example 2.4. Let Q be a unary predicate. Then $\varphi = \forall x Q(x) \vee \forall x \neg Q(x)$ is a $\frac{1}{2}$-tautology. Namely, in every probability model, either the set on which Q holds or its complement has measure at least $\frac{1}{2}$. However, φ is not an ε-tautology for $\varepsilon < \frac{1}{2}$. Furthermore, both φ and $\neg\varphi$ are classically satisfiable and hence ε-satisfiable for every ε; in particular we see that φ can be an ε-tautology while simultaneously $\neg\varphi$ is ε-satisfiable.

The next result will be used below.

Proposition 2.5. (Terwijn [4]) *Every formula φ is semantically equivalent to a formula φ' in prenex normal form; i.e. $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi \Leftrightarrow (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi'$ for all $\varepsilon \in [0, 1]$ and all ε-models $(\mathcal{M}, \mathcal{D})$.*

Note that 1-logic is fairly trivial: every formula in prenex normal form containing a universal quantifier is trivially true, so the only interesting fragment is the existential fragment, which is just the classical fragment. It turns out that 0-validity also does not contain much interesting information.

Proposition 2.6. (Terwijn [5, Proposition 3.2]) *The 0-valid formulas coincide with the classically valid formulas.*

That the 0-tautologies and classical tautologies coincide does not mean that 0-logic is the same as classical logic, because in a fixed 0-model $(\mathcal{M}, \mathcal{D})$ it might be the case that some statement $\varphi(x)$ holds for almost all elements of the model, but not for all; hence $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \forall x \varphi(x)$ but not $\mathcal{M} \models \forall x \varphi(x)$.

Thus, 0-logic and 1-logic have been taken care of from a computational point of view (the first is computably enumerable, while the second is decidable). This paper will deal with rational $\varepsilon \in (0, 1)$.

3. Many-One Reductions Between Different ε

In this section we will show that for rational $\varepsilon_0, \varepsilon_1 \in (0, 1)$, the ε_0-tautologies many-one reduce to the ε_1-tautologies. Not only does this show that we only need to consider one fixed ε for our hardness results (we will take $\varepsilon = \frac{1}{2}$ below), but in our proof of the Π^1_1-hardness of ε-validity these reductions will also turn out to be useful in a different way.

We will begin with reducing to bigger ε_1. To do this, we refine the argument by Terwijn [5], where it is shown that the 0-tautologies many-one reduce to the ε-tautologies for $\varepsilon \in [0, 1)$. Our argument is similar to the one given in Kuyper and Terwijn [3], where we discuss reductions for satisfiability instead of for validity.

Theorem 3.1. *Let $\mathcal{L}$ be a countable first-order language not containing equality. Then, for all rationals $0 \leq \varepsilon_0 \leq \varepsilon_1 < 1$, the ε_0-tautologies many-one reduce to the ε_1-tautologies.*

Proof. We can choose integers n and $0 < m \leq n$ so that $\frac{m}{n} = \frac{1-\varepsilon_1}{1-\varepsilon_0}$. Let $\varphi(y_1, \ldots, y_k)$ be a formula in prenex normal form (see Proposition 2.5). For simplicity we write $\vec{y} = y_1, \ldots, y_k$. Also, for a function π we let $\pi(\vec{y})$ denote the vector $\pi(y_1), \ldots, \pi(y_k)$. We use formula-induction to define a computable function f such that for every formula φ,

(3) φ is an ε_0-tautology if and only if $f(\varphi)$ is an ε_1-tautology.

For propositional formulas and existential quantifiers, there is nothing to be done and we use the identity map. Next, we consider the universal quantifiers. Let $\varphi = \forall x \psi(\vec{y}, x)$. The idea is to introduce new unary predicates, that can be used to vary the strength of the universal quantifier. We will make these predicates split the model into disjoint parts. If we split it into just the right number of parts (in this case n), then we can choose m of these parts to get just the right strength.

So, we introduce new unary predicates $X_1, \ldots, X_n$. We define the sentence n-split by:

$$\forall x \left((X_1(x) \vee \ldots \vee X_n(x)) \wedge \bigwedge_{1 \leq i < j \leq n} \neg \left(X_i(x) \wedge X_j(x) \right) \right).$$

Then one can verify that in any model, $\neg n$-split does *not* hold if and only if the sets X_i disjointly cover the entire model.

Now define $f(\varphi)$ to be the formula

$$\neg n\text{-split} \vee \bigvee_{i_1, \ldots, i_m} \forall x \big((X_{i_1}(x) \vee \cdots \vee X_{i_m}(x)) \wedge f(\psi)(\vec{y}, x) \big)$$

where the disjunction is over all subsets of size m from $\{1, \ldots, n\}$. (It will be clear from the construction that $f(\psi)$ has the same arity as ψ.) Thus, $f(\varphi)$ expresses that for some choice of m of the n parts, $f(\psi)(x)$ holds often enough when restricted to the resulting part of the model.

We will now prove claim (3) above. For the implication from right to left, we will prove the following strengthening:

For every formula $\varphi(\vec{y})$ and every probability model $(\mathcal{M},\mathcal{D})$ there exists a probability model $(\mathcal{N},\mathcal{E})$ together with a measure-preserving surjective measurable function $\pi:\mathcal{N}\to\mathcal{M}$ (i.e. for all $\mathcal{D}$-measurable A we have that $\mathcal{E}(\pi^{-1}(A))=\mathcal{D}(A)$) such that for all $\vec{y}\in\mathcal{N}$ we have that

$$(\mathcal{N},\mathcal{E}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y}) \text{ if and only if } (\mathcal{M},\mathcal{D}) \not\models_{\varepsilon_0} \varphi(\pi(\vec{y})).$$

In particular, if $f(\varphi)$ is an ε_1-tautology, then φ is an ε_0-tautology. We prove this by formula-induction over the formulas in prenex normal form. For propositional formulas, there is nothing to be done (we can simply take the models to be equal and π the identity). For the existential quantifier, let $\varphi=\forall x\psi(x)$ and apply the induction hypothesis to ψ to find a model $(\mathcal{N},\mathcal{E})$ and a mapping π. Then we can take the same model and mapping for φ, as easily follows from the fact that π is surjective.

Next, we consider the universal quantifier. Suppose $\varphi=\forall x\psi(\vec{y},x)$ and let $(\mathcal{M},\mathcal{D})$ be a probability model. Use the induction hypothesis to find a model $(\mathcal{N},\mathcal{E})$ and a measure-preserving surjective measurable function $\pi:\mathcal{N}\to\mathcal{M}$ such that for all $\vec{y},x\in\mathcal{N}$ we have that

$$(\mathcal{N},\mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\vec{y},x) \text{ if and only if } (\mathcal{M},\mathcal{D}) \not\models_{\varepsilon_0} \varphi(\pi(\vec{y}),\pi(x)).$$

Now form the probability model $(\mathcal{N}',\mathcal{E}')$ which consists of n disjoint copies $\mathcal{N}_1,\dots,\mathcal{N}_n$ of $(\mathcal{N},\mathcal{E})$, each with weight $\frac{1}{n}$. That is, $\mathcal{E}'$ is the sum of n copies of $\frac{1}{n}\mathcal{E}$. Let $\pi':\mathcal{N}'\to\mathcal{M}$ be the composition of the projection map $\sigma:\mathcal{N}'\to\mathcal{N}$ with π. Relations in $\mathcal{N}'$ are defined just as on $\mathcal{N}$, that is, for a t-ary relation R we define $R^{\mathcal{N}'}(x_1,\dots,x_t)$ by $R^{\mathcal{N}}(\sigma(x_1),\dots,\sigma(x_t))$. Observe that this is the same as defining $R^{\mathcal{N}'}(x_1,\dots,x_t)$ by $R^{\mathcal{M}}(\pi'(x_1),\dots,\pi'(x_t))$. We interpret constants $c^{\mathcal{N}'}$ by embedding $c^{\mathcal{N}}$ into the first copy $\mathcal{N}_1$. For functions f of arity t, first note that we can see $f^{\mathcal{N}}$ as a function from $\mathcal{N}^t\to\mathcal{N}'$ by embedding its codomain $\mathcal{N}$ into the first copy $\mathcal{N}_1$. We now interpret $f^{\mathcal{N}'}$ as the composition of this $f^{\mathcal{N}}$ with π'. Finally, we let each X_i be true exactly on the copy $\mathcal{N}_i$.

Then π' is clearly surjective. To show that it is measure-preserving, it is enough to show that σ is measure-preserving. If A is $\mathcal{E}$-measurable, then $\sigma^{-1}(A)$ consists of n disjoint copies of A, each having measure $\frac{1}{n}\mathcal{E}(A)$, so $\pi^{-1}(A)$ has $\mathcal{E}'$-measure exactly $\mathcal{E}(A)$.

Now, since $(\mathcal{N}',\mathcal{E}')$ does not satisfy $\neg n$-split (because the X_i disjointly cover $\mathcal{N}'$), we see that

$$(\mathcal{N}',\mathcal{E}') \not\models_{\varepsilon_1} f(\varphi)(\vec{y}) \tag{4}$$

is equivalent to the statement that for all $1\le i_1<\dots<i_m\le n$ we have

$$\Pr_{\mathcal{E}'}\big[x\in\mathcal{N}' \mid (\mathcal{N}',\mathcal{E}') \not\models_{\varepsilon_1} (X_{i_1}(x)\vee\dots\vee X_{i_m}(x))\wedge f(\psi)(\vec{y},x)\big] > \varepsilon_1. \tag{5}$$

By Lemma 3.2 below we have that $(\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} f(\psi)(\vec{y}, x)$ holds if and only if $(\mathcal{N}, \mathcal{E}) \models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), \sigma(x))$ holds. In particular, we see for every $1 \leq i \leq n$ that

$$\begin{aligned}&\Pr_{\mathcal{E}'}\big[x \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} X_i(x) \text{ and } (\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} f(\psi)(\vec{y}, x)\big]\\ &= \frac{1}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big].\end{aligned}$$

It follows that (5) is equivalent to

$$\frac{n-m}{n} + \frac{m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big] > \varepsilon_1.$$

The induction hypothesis tells us that this is equivalent to

$$\frac{n-m}{n} + \frac{m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})), \pi(x))\big] > \varepsilon_1$$

and since π is surjective and measure-preserving, this is the same as

$$\begin{aligned}\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})), x)\big] &> \frac{n}{m}\left(\varepsilon_1 - \frac{n-m}{n}\right)\\ &= \frac{n}{m}(\varepsilon_1 - 1) + 1 = \varepsilon_0.\end{aligned}$$

This proves that we have $(\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$ if and only if $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \varphi(\pi'(\vec{y}))$.

To prove the left to right direction of (3) we will use induction to prove the following stronger statement:

If $(\mathcal{M}, \mathcal{D})$ is a probability model and $\vec{y} \in \mathcal{M}$ are such that $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$, then we also have $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \varphi(\vec{y})$.

In particular, if φ is an ε_0-tautology, then $f(\varphi)$ is an ε_1-tautology. The only interesting case is the universal case, so let $\varphi = \forall x \psi(\vec{y}, x)$. Let $\vec{y} \in \mathcal{M}$ be such that $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$. Assume, towards a contradiction, that $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \varphi(\vec{y})$. Then

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \psi(\vec{y}, x)\big] \geq 1 - \varepsilon_0$$

and by the induction hypothesis we have

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} f(\psi)(\vec{y}, x)\big] \geq 1 - \varepsilon_0. \tag{6}$$

Because $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_1} \neg n$-split, the X_i disjointly cover $\mathcal{M}$, as discussed above. Now, by taking those m of the X_i (say $X_{i_1}, \ldots, X_{i_m}$) which have the largest intersection with this set we find that

$$\begin{aligned}\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} (X_{i_1} \vee \cdots \vee X_{i_m}) \wedge f(\psi)(\vec{y}, x)\big] &\geq \frac{m}{n}(1 - \varepsilon_0)\\ &= 1 - \varepsilon_1\end{aligned}$$

which contradicts our choice of $(\mathcal{M}, \mathcal{D})$. □

Lemma 3.2. (Kuyper and Terwijn [3]) *Let* $(\mathcal{N}', \mathcal{E}')$ *and* $(\mathcal{N}, \mathcal{E})$ *be as in the proof of Theorem 3.1 above. Then for every formula* $\zeta(x_1, \dots, x_t)$ *in the language of* $\mathcal{M}$*, for every* $\varepsilon \in [0,1]$ *and all* $x_1, \dots, x_t \in \mathcal{N}'$*:* $(\mathcal{N}', \mathcal{E}') \models_\varepsilon \zeta(x_1, \dots, x_t)$ *if and only if* $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta(\sigma(x_1), \dots, \sigma(x_t))$*.*

Proof. By induction on the structure of the formulas in prenex normal form. The base case holds by definition of the relations in $\mathcal{N}'$. The only interesting induction step is the one for the universal quantifier. So, let $\zeta = \forall x_0 \zeta'(x_0, \dots, x_t)$ and let $x_1, \dots, x_t \in \mathcal{N}'$. Using the induction hypothesis, we find that the set $A = \{x_0 \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \models_\varepsilon \zeta'(x_0, \dots, x_t)\}$ is equal to the set $\{x_0 \in \mathcal{N}' \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta'(\sigma(x_0), \dots, \sigma(x_t))\}$, which consists of n disjoint copies of the set $B = \{x_0 \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta'(x_0, \sigma(x_1), \dots, \sigma(x_t))\}$; denote the copy of B living inside $\mathcal{N}_i$ by B_i. Then

$$\mathcal{D}(A) = \sum_{i=1}^{n} \mathcal{E}'(B_i) = \sum_{i=1}^{n} \tfrac{1}{n}\mathcal{E}(B) = \mathcal{E}(B)$$

from which we directly see that $(\mathcal{N}', \mathcal{E}') \models_\varepsilon \zeta(x_1, \dots, x_t)$ if and only if $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta(\sigma(x_1), \dots, \sigma(x_t))$. □

Theorem 3.3. *Let* $\mathcal{L}$ *be a countable first-order language not containing equality. Then, for all rationals* $0 < \varepsilon_1 \leq \varepsilon_0 \leq 1$*, the* ε_0*-tautologies many-one reduce to the* ε_1*-tautologies.*

Proof. We can choose integers n and $m < n$ such that $\frac{m}{n} = \frac{\varepsilon_0 - \varepsilon_1}{\varepsilon_0}$. We construct a many-one reduction f such that for all formulas φ,

$$\varphi \text{ is an } \varepsilon_0\text{-tautology if and only if } f(\varphi) \text{ is an } \varepsilon_1\text{-tautology.}$$

Again, we only consider the nontrivial case where φ is a universal formula $\forall x \psi(\vec{y}, x)$. We define $f(\varphi)$ to be the formula

$$\neg\text{-}n\text{-split} \wedge \bigvee_{i_1, \dots, i_m} \forall x \big(X_{i_1}(x) \vee \dots \vee X_{i_m}(x) \vee f(\psi)(\vec{y}, x)\big)$$

where the disjunction is over all subsets of size m from $\{1, \dots, n\}$.

The proof is almost the same as for Theorem 3.1. In the proof for the implication from right to left, follow the proof up to (4), i.e.

$$(\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} f(\varphi)(\vec{y}).$$

This is equivalent to the statement that for all $1 \leq i_1 < \dots < i_m \leq n$ we have

$$\Pr_{\mathcal{E}'}\big[x \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} X_{i_1}(x) \vee \dots \vee X_{i_m}(x) \vee f(\psi)(\vec{y}, x)\big] > \varepsilon_1.$$

Similar as before, using Lemma 3.2, we find that this is equivalent to

$$\frac{n-m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big] > \varepsilon_1.$$

Again, using the induction hypothesis and the fact that π is measure-preserving we find that this is equivalent to

$$\begin{aligned}&\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \not\models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})),x)\big]\\ &> \frac{n}{n-m}\varepsilon_1 = \frac{\varepsilon_0}{\varepsilon_1}\varepsilon_1 = \varepsilon_0.\end{aligned}$$

This proves that $(\mathcal{N}',\mathcal{E}') \models_{\varepsilon_1} f(\varphi)(\vec{y})$ if and only if $(\mathcal{M},\mathcal{D}) \models_{\varepsilon_0} \varphi(\vec{y})$.

For the converse implication, we also need to slightly alter the proof of Theorem 3.1. Assuming that $(\mathcal{M},\mathcal{D}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$, follow the proof up to (6), where we obtain

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\varepsilon_1} f(\psi)(\vec{y},x)\big] \geq 1-\varepsilon_0. \tag{7}$$

Define

$$\eta = \Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\varepsilon_1} f(\psi)(\vec{y},x)\big]$$

and take those m of the X_i (say $X_{i_1},\dots,X_{i_m}$) which have the smallest intersection with this set. Note that by (7) we have $\eta \geq 1-\varepsilon_0$. Then we find that

$$\begin{aligned}&\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\varepsilon_1} X_{i_1} \vee \cdots \vee X_{i_m} \vee f(\psi)(\vec{y},x)\big]\\ &\geq \frac{m}{n} + \left(1-\frac{m}{n}\right)\eta = \frac{\varepsilon_0-\varepsilon_1}{\varepsilon_0} + \frac{\varepsilon_1}{\varepsilon_0}\eta\\ &\geq \frac{\varepsilon_0-\varepsilon_1}{\varepsilon_0} + \frac{\varepsilon_1}{\varepsilon_0}(1-\varepsilon_0) = 1-\varepsilon_1.\end{aligned}$$

which contradicts our choice of $(\mathcal{M},\mathcal{D})$. □

Observe that, because of the inductive nature of the reductions above, we can perform these reductions per quantifier. In particular, we can talk about what it means for a formula with variable ε (that is, a separate ε for each quantifier) to be a tautology. This way, we get something like Keisler's probability logic mentioned in the introduction; however, remember that we still have our non-classical negation (unlike Keisler). This idea will be crucial in our hardness proof.

4. Validity Is Π^1_1-hard

To show that the set of ε-tautologies is indeed Π^1_1-hard, we adapt a proof by Hoover [1] which shows that $\mathcal{L}_{\omega P}$ is Π^1_1-complete. We will show that, to a certain extent, we can define the natural numbers within probability logic.

Definition 4.1. Let φ be a formula in prenex normal form and N a unary predicate. Then φ^N, or φ **relativised to** N, is defined as the formula where each $\forall x\psi(x)$ is replaced by $\forall x(N(x) \to \psi(x))$ and each $\exists x\psi(x)$ is replaced by $\exists x(N(x) \wedge \psi(x))$.

Theorem 4.2. *Let $\mathcal{L}$ be the language consisting of a constant symbol* 0*, a unary relation* $N(x)$*, binary relations* $x = y$*,*[1] $S(x) = y$ *and* $R(x, y)$*, and ternary relations* $x + y = z$ *and* $x \cdot y = z$*. Furthermore, let* f *be the reduction from* 0*-tautologies to* $\frac{1}{2}$*-tautologies from Proposition 3.1. Then there exists finite theories* T, T' *in the language* $\mathcal{L}$ *such that, for every first-order sentence* φ *containing a new predicate symbol* Q*, the following are equivalent:*

(i) $\models_{\frac{1}{2}} f(\neg(\bigwedge T)) \vee \neg(\bigwedge T') \vee f\left(\neg\varphi^N\right)$*;*
(ii) $\mathbb{N} \models \forall Q \neg\varphi(Q)$.[2]

Proof. We will prove the contrapositives of the implications (i) $\to$ (ii) and (ii) $\to$ (i). During the entire proof, one should mainly think about what it means for a formula ψ that its negation $\neg\psi$ does not hold. Note that we have that $(\mathcal{M}, \mathcal{D}) \not\models_0 \neg\psi$ if and only if all universal quantifiers hold classically and all existential quantifiers hold on a set of strictly positive measure. Likewise, $(\mathcal{M}, \mathcal{D}) \not\models_{\frac{1}{2}} \neg\psi$ holds if and only if all universal quantifiers hold classically and all existential quantifiers hold on a set of measure strictly greater than $\frac{1}{2}$.

Inspired by this, we form the theories T and T'. T consists of Robinson's Q relativised to N, axioms specifying that the arithmetical relations only hold on N, and some special axioms for N and R. That is, we put the following axioms in T (keeping in mind that we are mostly interested in what happens when the negation of these formulas does not hold, i.e. one should read the $\forall$ as a classical universal quantifier and the $\exists$ as saying that the statement holds on a set of strictly positive measure):

All equality axioms. For example:

$$\forall x(x = x)$$
$$\forall x \forall y((N(x) \wedge x = y) \to N(y))$$

We should guarantee that 0 is in N:

$$N(0)$$

[1]Here we do not mean true equality, but rather a binary relation that we will use to represent equality.

[2]We denote by $\forall Q \neg\varphi(Q)$ the second-order formula $\forall X \neg\varphi(X/Q)$, where $\varphi(X/Q)$ is the formula where the predicate symbol Q is replaced by a second-order variable X.

[3]We do not really need this last axiom, but we have added it anyway so that all axioms of Robinson's Q are in T.

We now give the axioms for the successor function:

$$\forall x \forall y (S(x) = y \to (N(x) \land N(y)))$$
$$(\forall x \exists y S(x) = y)^N$$
$$(\forall x \forall y \forall u \forall v ((S(x) = y \land S(u) = v \land x = u) \to y = v))^N$$
$$(\forall x \neg S(x) = 0)^N$$
$$(\forall x (x = 0 \lor \exists y S(y) = x))^N.^3$$

In the axioms below, we will leisurely denote by $\psi(S(x))$ the formula $\forall y (S(x) = y \to \psi(y))$ and similarly for $x + y$ and $x \cdot y$. We proceed with the inductive definitions of $+$ and $\cdot$:

$$(\forall x \forall y \forall z (x + y = z \to (N(x) \land N(y) \land N(z))))$$
$$(\forall x (x + 0 = x))^N$$
$$(\forall x \forall y (x + S(y) = S(x + y)))^N$$
$$(\forall x \forall y \forall z (x \cdot y = z \to (N(x) \land N(y) \land N(z))))^N$$
$$(\forall x (x \cdot 0 = 0))^N$$
$$(\forall x \forall y (x \cdot S(y) = (x \cdot y) + x))^N.$$

Finally, we introduce a predicate R. This predicate is meant to function as a sort of 'padding'. The goal of this predicate is to force the measure of a point $S^n(0)$ to be larger than the measure of $\{x \mid N(x) \land x > S^n(0)\}$ (the precise use will be made clear in the proof below).

$$(\forall x \forall y \neg R(x, y))^N$$

The last two axioms will be in T' instead of in T, because these need to be evaluated for $\varepsilon = \frac{1}{2}$ while the rest will be evaluated for $\varepsilon = 0$. So, because we will be looking at when the negation does not hold, the existential quantifier should be read as "strictly more than measure $\frac{1}{2}$ many".

$$\forall x (N(x) \to \exists y (R(x, y) \lor x = y))$$
$$\forall x (N(x) \to \exists y \neg (R(x, y) \lor x < y))$$

Here, $x < y$ is short for $f(\exists z (x + S(z) = y))$, i.e. the usual definition of $x < y$ evaluated for $\varepsilon = 0$.

Note that for universal formulas it does not matter if they are in T or T' because in both cases the negation of the formula does not hold if and only if the formula holds classically.

We will now show that these axioms indeed do what we promised. First, we show that (i) implies (ii). So, assume $\mathbb{N} \not\models \forall Q(\neg\varphi(Q))$. Fix a predicate $Q^{\mathbb{N}}$ such that $\mathbb{N} \not\models \neg\varphi(Q)$. Now take the model $\mathcal{M} = \omega \times \{0, 1\}$ to be the disjoint union of two copies of ω, where we define

$S, +, \cdot, \leq, 0$ on the first copy $\omega \times \{0\}$ of ω as usual, and let these be undefined elsewhere. Let

$$N := \omega \times \{0\} \text{ and } R := \left\{((a,0),(b,1)) \mid \mu k\left[2^{k+1} > 3^{a+1}\right] \neq b\right\}.$$

We let $Q^{\mathcal{M}}(a,0)$ hold if $Q^{\mathbb{N}}(a)$ and we never let it hold on the second copy of ω. Finally, define $\mathcal{D}$ by

$$\mathcal{D}(a,0) = \mathcal{D}(a,1) := \frac{1}{3^{a+1}}.$$

Then it is directly verified that

$$(\mathcal{M},\mathcal{D}) \not\models_0 \neg\left(\bigwedge T\right) \vee \neg\varphi^N,$$

i.e. all formulas in $T \cup \left\{\varphi^N\right\}$ hold in $(\mathcal{M},\mathcal{D})$ if universal quantifiers are interpreted classically and existential quantifiers as expressing that there exists a set of positive measure. Note that because all points have positive measure this is equivalent to the classical existential quantifier, so all we are really saying is that T and φ^N hold classically in $\mathcal{M}$.

Furthermore, if we let $a \in \omega$ and denote b for $\mu k[2^{k+1} > 3^{a+1}]$ then we have that

$$\begin{aligned}
&\Pr_{\mathcal{D}}\left[y \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\frac{1}{2}} R((a,0),y) \vee (a,0) = y\right] \\
&= \frac{1}{2} - \frac{1}{2^{b+1}} + \frac{1}{3^{a+1}} \\
&> \frac{1}{2}
\end{aligned}$$

while we also have that

$$\begin{aligned}
&\Pr_{\mathcal{D}}\left[y \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\frac{1}{2}} \neg(R((a,0),y) \vee (a,0) < y)\right] \\
&= 1 - \left(\frac{1}{2} - \frac{1}{2^{b+1}} + \sum_{i=a+2}^{\infty} 3^{-i}\right) \\
&= 1 - \frac{1}{2}\left(1 - \frac{1}{2^b} + \frac{1}{3^{a+1}}\right) \\
&> \frac{1}{2}
\end{aligned}$$

where the last inequality follows from the fact that b is the smallest $k \in \omega$ such that $2^{k+1} > 3^{a+1}$, so that $2^b \leq 3^{a+1}$. Thus, we see that $(\mathcal{M},\mathcal{D}) \not\models_{\frac{1}{2}} \neg(\bigwedge T')$. But then we see from (the proof of) Theorem 3.1, together with the remark below Theorem 3.3 that there is a probability model $(\mathcal{N},\mathcal{E})$ such that

$$(\mathcal{N},\mathcal{E}) \not\models_{\frac{1}{2}} f\left(\neg\left(\bigwedge T\right)\right) \vee \neg\left(\bigwedge T'\right) \vee f\left(\neg\varphi^N\right),$$

i.e. (i) does not hold.

Conversely, assume that statement (i) does not hold. Without loss of generality, we may assume the equality relation on $\mathcal{M}$ to be true equality; otherwise, because (i) does not hold and all equality axioms are in T we could look at $\mathcal{M}/=$ instead.

Again, from (the proof of) Theorem 3.1 we see that

$$(\mathcal{M},\mathcal{D}) \not\models_0 \neg\Big(\bigwedge T\Big) \vee \neg\varphi^N \text{ and } (\mathcal{M},\mathcal{D}) \not\models_{\frac{1}{2}} \neg\Big(\bigwedge T'\Big).$$

We will now use the three axioms involving R. Let $m \in \mathcal{M}$ with $\mathcal{M} \models N(m)$. Then $\{a \in \mathcal{M} \mid \mathcal{M} \models a = m\} \subseteq N^{\mathcal{M}}$ by the equality axioms, and similarly $\{a \in \mathcal{M} \mid m < a\} \subseteq N^{\mathcal{M}}$. So the axiom $(\forall x \forall y \neg R(x,y))^N$ tells us that these two sets are disjoint from $\{a \in \mathcal{M} \mid \mathcal{M} \models R(m,a)\}$. Therefore, from the two axioms in T' it now follows that

$$\begin{aligned}\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models m = a\big] &> \frac{1}{2} - \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models R(m,a)\big] \\ &> \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models m < a\big].\end{aligned}$$

Thus,

$$\Pr_{\mathcal{D}}\big[\mathcal{M}\big] > \frac{1}{2}\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid m \leq a\big] \tag{8}$$

We now claim that, if we denote $S(x)$ for the unique y such that $S(x) = y$ (as guaranteed to exist and be unique by T):

$$\Pr_{\mathcal{D}}\Big[\big\{0,\ldots,S^k(0)\big\}\Big] > \left(1 - \frac{1}{2^{k+1}}\right)\Pr_{\mathcal{D}}\big[N\big].$$

For $k = 0$ this is clear: from the axioms in T it follows that for all elements $a \in N$ different from 0 we have $a > 0$, and therefore $\Pr_{\mathcal{D}}\big[\{0\}\big] > \frac{1}{2}\Pr_{\mathcal{D}}\big[N\big]$ by (8). Similarly, assume this holds for $k \in \omega$. Then we have by (8):

$$\begin{aligned}&\Pr_{\mathcal{D}}\Big[\big\{0,\ldots,S^{k+1}(0)\big\}\Big] \\ &> \Pr_{\mathcal{D}}\Big[\big\{0,\ldots,S^k(0)\big\}\Big] + \frac{1}{2}\left(\Pr_{\mathcal{D}}\big[N\big] - \Pr_{\mathcal{D}}\Big[\big\{0,\ldots,S^k(0)\big\}\Big]\right)\end{aligned}$$

so from the induction hypothesis we obtain

$$\begin{aligned}&> \left(1 - \frac{1}{2^{k+1}}\right)\Pr_{\mathcal{D}}\big[N\big] + \frac{1}{2^{k+2}}\Pr_{\mathcal{D}}\big[N\big] \\ &= \left(1 - \frac{1}{2^{k+2}}\right)\Pr_{\mathcal{D}}\big[N\big].\end{aligned}$$

Because this converges to $\Pr_{\mathcal{D}}\big[N\big]$ if k goes to infinity, we see that all weight of N rests on $X := \{S^n(0) \mid n \in \omega\} \subseteq N$. Now, if some universal quantifier holds when relativised to N, it certainly holds when restricted to X. Furthermore, if some existential quantifier holds with positive measure in N, then it also has to hold with positive measure in X because $X \subseteq N$ has the same measure as N. Therefore, we see that $(\mathcal{M},\mathcal{D}) \not\models_0 \neg\varphi^N$ implies that also $(\mathcal{M}{\restriction}X,\mathcal{D}{\restriction}X) \not\models_0 \neg\varphi^X$ (see the discussion at the beginning of the proof about what it means for the negation of a formula to not hold).

However, we can directly verify that $\mathcal{M}{\restriction}X$ is isomorphic to the standard natural numbers $\mathbb{N} = (\omega, S, +, \cdot, 0)$. So, by transferring the predicate Q from $\mathcal{M}$ to $\mathbb{N}$ (i.e. letting $Q^{\mathbb{N}}(k)$ hold if $Q^{\mathcal{M}}(S^k(0))$ holds) we find that indeed $\mathbb{N} \not\models \forall Q \neg\varphi(Q)$. □

Putting this together, we reach our conclusion.

Theorem 4.3. *For rational $\varepsilon \in (0,1)$, the set of ε-tautologies is Π^1_1-hard.*

Proof. From Theorem 3.1, Theorem 3.3 and Theorem 4.2. □

In fact, we have shown that even for languages not containing function symbols or equality, ε-validity is already Π^1_1-hard. Our proof above uses one constant: 0. However, we could also replace 0 by a unary relation representing $0 = x$ and modify the proof to show that the relational fragment of ε-validity is Π^1_1-hard.

We do not yet know of an upper bound for the complexity of ε-validity. While we have developed methods for proving upper bounds for ε-satisfiability, which will be discussed in a future paper, these methods do not seem to work for proving any results about ε-validity. Thus, the exact complexity of ε-validity is still an open problem.

References

1. D. N. Hoover, *Probability logic*, Annals of Mathematical Logic **14** (1978), 287–313.
2. H. J. Keisler, *Probability quantifiers*, Model-Theoretic Logics (J. Barwise and S. Feferman, eds.), Perspectives in Mathematical Logic, vol. 8, Springer-Verlag, 1985.
3. R. Kuyper and S. A. Terwijn, *Model theory of measure spaces and probability logic*, The Review of Symbolic Logic **6** (2013), no. 3, 367–393.
4. S. A. Terwijn, *Probabilistic logic and induction*, Journal of Logic and Computation **15** (2005), no. 4, 507–515.
5. ———, *Decidability and undecidability in probability logic*, Proceedings of Logical Foundations of Computer Science (S. Artemov and A. Nerode, eds.), Lecture Notes in Computer Science, vol. 5407, Springer, 2009, pp. 441–450.
6. L. G. Valiant, *Robust logics*, Artificial Intelligence **117** (2000), 231–253.

(Rutger Kuyper) Radboud University Nijmegen, Department of Mathematics, P.O. Box 9010, 6500 GL Nijmegen, the Netherlands.

E-mail address: `r.kuyper@math.ru.nl`