Math. Struct. in Comp. Science (2017), *vol.* 27, *pp.* 1581–1600. © Cambridge University Press 2016
doi:10.1017/S0960129516000189 First published online 29 June 2016

Quantifying over events in probability logic: an introduction

STANISLAV O. SPERANSKI

Sobolev Institute of Mathematics, 4 Koptyug ave., 630090 Novosibirsk, Russia
Email: katze.tail@gmail.com

Received 25 February 2014; revised 9 October 2014

In this article we describe a bunch of *probability logics with quantifiers over events*, and develop primary techniques for proving computational complexity results (in terms of *m*-degrees) about these logics, mainly over discrete probability spaces. Also the article contains a comparison with some other probability logics and a discussion of interesting analogies with research in the metamathematics of Boolean algebras, demonstrating a number of attractive features and intuitive advantages of the present proposal.

1. Quantified probability logics

1.1. *Preliminary motivation*

Roughly speaking, by a *probability logic* we mean a triple $\langle \mathscr{L}, \mathcal{K}, \Vdash \rangle$ where:

- $\mathscr{L}$ is a formal language, intended for building $\mathscr{L}$*-formulas*;
- $\mathcal{K}$ is a class of $\mathscr{L}$*-structures*, whose descriptions involve probability spaces;
- $\Vdash$ is a *satisfiability relation* between $\mathscr{L}$-structures and $\mathscr{L}$-formulas.

If $\mathscr{L}$ contains quantifiers, then the collection of all valid (with respect to $\Vdash$) $\mathscr{L}$-sentences often turns out to be of very high computational complexity, far from being recursively enumerable, and hence there exists no complete deductive consequence relation for $\mathscr{L}$. E.g. the validity problems for various logics in Abadi and Halpern (1994); Hoover (1978); Speranski (2011); Terwijn (2005) are more complex than the first-order true arithmetic of the natural numbers, so they do not fit into the *arithmetical hierarchy* of *m*-degrees. In effect, some languages in Abadi and Halpern (1994) produce more complexity than the second-order true arithmetic, and thus cannot be characterised even within the *analytical hierarchy*. See Leitgeb (2016) for references to other formalisms.

The main goal of this paper is to describe new quantified probability logics, give the reader some feeling of their behaviour and develop primary techniques which will be of importance in further investigation, rather than to establish unexpected results. As we shall explicitly point out, these logics have nice algebraic features, and moreover allow us to introduce natural analogues of known notions from Boolean algebras — leading to an appropriate classification of probability spaces. The latter reveals interesting connections with achievements in elementary theories of Boolean algebras, whose role seems to be underestimated or even neglected in many quantified probabilistic formalisms.

More precisely we propose a bunch of probability logics, each of which is denoted by QPL^C with C a suitable set of constants. Actually, the quantifier-free fragment of QPL^C

is a variant of the well-known quantifier-free formalism studied in Fagin *et al.* (1990) — namely the one which deals directly with events. So QPL^C may be viewed as extending this variant by adding *quantifiers over events*. The idea turns out to be fruitful for many reasons. Clearly there exists a trade-off between computability and expressibility. If we have a quantified probabilistic logic of high computational complexity (as is usually the case), we wish some natural properties of probability spaces and classifications to be expressible in the corresponding language. These definability issues strongly depend on the chosen formulation of the logic and its features, such as available sorts of objects, prefix normal forms, etc. Among other things, QPL^C allows us to:

- safely take quotients of probability spaces modulo sets of measure zero (this natural property is often not available in various other probabilistic formalisms);
- define the concepts of finiteness and discreteness, modulo events of measure zero, for probability spaces;
- introduce prefix normal forms and the associated hierarchy of validity problems;
- propose a natural classification of probability spaces reminiscent of the famous elementary classification of Boolean algebras, cf. Koppelberg (1989).

E.g. no analogues of the second or third item are known for the other languages mentioned above, except for a version of 'finiteness' in Speranski (2013a). Thus, in particular, the approach looks attractive from an algebraic perspective.

Since every event is uniquely specified by its characteristic function, quantification over events directly corresponds to quantification over Bernoulli random variables. Hence the quantifiers used in QPL^C are very appealing from the viewpoint of probability theory as well (the reader might consult (Suppes *et al.* 1998) for a number of theorems involving quantification over Bernoulli random variables).

One can say that the role of QPL^C (which is natural both logically and algebraically) for probability spaces is similar to that of the languages studied in Solovay *et al.* (2012) for vector spaces, normed spaces, etc. Further, the framework of QPL^C can be exploited to formalise certain proposals in philosophical logic; see the remark on Leitgeb's probabilistic belief theory at the end of the article. Also, investigating various computability and expressibility issues for QPL^C is interesting in its own right.

Sections 2–5 employ, among other things, earlier results of Tarski (1951) and Nies (1996), as well as some recent contributions of Speranski (2011; 2013b) — where those of the latter paper generalise Halpern's Π^1_1-completeness theorem for Presburger arithmetic with a free unary predicate. In this way the work also gives a brief overview of the corresponding methods.

Subsection 1.2 proceeds by providing some necessary background material and mathematically defining QPL^C, along with the associated hierarchy of validity problems. In addition, it contains further remarks which are mainly intended for those who have experience in probabilistic logics and related issues. As we shall see, QPL^C has the same complexity as the second-order true arithmetic — although it does not refer to natural numbers or their subsets — and the corresponding hierarchy of sets of probabilistic formulas does not collapse. Intuitively, to derive complexity lower bounds for a logic $\mathscr{L}$ speaking directly about objects of analysis, it may be practically convenient to interpret

QPL^C in $\mathscr{L}$. In such cases one avoids the need to code $\langle \mathbb{N}, 2^{\mathbb{N}}; +, \times \rangle$ within $\mathscr{L}$ (directly) — so here undecidability and complexity arguments will be concerned.

1.2. *Mathematical formulation*

One of the most popular formal languages in the foundations of mathematics is that of second-order arithmetic: employing certain coding techniques, many classical objects of analysis may be defined in its terms (the reader might consult (Simpson 2009) for more details). For these reasons, the second-order theory of the standard model $\mathfrak{N} = \langle \mathbb{N}, +, \times \rangle$ is often referred to as *elementary analysis*. This language and its prefix fragments also take an important part in characterising so-called *m-degrees* — since they are used for describing the computational complexity of various problems within the *analytical hierarchy* (see e.g. Rogers (1967)).

Here it is helpful to briefly recall a few concepts from classical computability theory. Let $A \subseteq \mathbb{N}$ and $B \subseteq \mathbb{N}$. We say A is *m-reducible* to B (denoted $A \leqslant_m B$) iff there exists a computable function $f : \mathbb{N} \to \mathbb{N}$ such that for every $k \in \mathbb{N}$,

$$k \in A \iff f(k) \in B.$$

A and B are *m-equivalent* (denoted $A \equiv_m B$) iff $A \leqslant_m B$ and $B \leqslant_m A$. Further — we identify each problem specified by a question of the form

Does a given input have the desired property?

with the set of all inputs for which the answer is 'yes,' and view, in turn, this set as a collection of natural numbers, up to a suitable Gödel numbering of the possible inputs. So the m-reducibility extends to such problems as well. Now by the *computational complexity* of A we mean the equivalence class of A under $\equiv_m$, i.e. what is generally known as the *m-degree* of A. Take $\mathscr{P}_n$ (respectively $\mathscr{S}_n$) to be the set of $\Pi^1_n(\Sigma^1_n)$-sentences of second-order arithmetic true in $\mathfrak{N}$ and $\mathscr{P}_\infty$ to be elementary analysis. Consequently the analytical hierarchy includes the following major degrees:

$$\Pi^1_0,\ \Pi^1_1,\ \Pi^1_2,\ \ldots \quad \text{and} \quad \Sigma^1_0,\ \Sigma^1_1,\ \Sigma^1_2,\ \ldots$$

where Π^1_n (Σ^1_n) is the computational complexity of $\mathscr{P}_n$ ($\mathscr{S}_n$). Let Π^1_∞ denote the limiting degree, viz. the complexity of $\mathscr{P}_\infty$. A portion of the related terminology will be exploited below: for $\lambda \in \mathbb{N} \cup \{\infty\}$,

$$\begin{array}{rcl} A \text{ is } \Pi^1_\lambda\textit{-hard} & \text{iff} & \mathscr{P}_\lambda \leqslant_m A, \\ A \text{ is } \Pi^1_\lambda\textit{-bounded} & \text{iff} & A \leqslant_m \mathscr{P}_\lambda, \\ A \text{ is } \Pi^1_\lambda\textit{-complete} & \text{iff} & \mathscr{P}_\lambda \equiv_m A; \end{array}$$

and similarly for Σ^1_λ with $\lambda \in \mathbb{N}$. In the context of the analytical hierarchy, the best way of estimating the complexity of A consists in finding $B \in \{\mathscr{P}_n, \mathscr{S}_n\}_{n\in\mathbb{N}} \cup \{\mathscr{P}_\infty\}$ which satisfies $A \equiv_m B$.

As has been mentioned above, coding is required to express notions of analysis in the language of second-order arithmetic — clearly the latter was not designed to explicitly talk about them. And if one is interested in establishing complexity lower bounds for a language $\mathscr{L}$ speaking directly about objects of analysis, it may be practically more convenient to interpret another such language (or its fragments) in $\mathscr{L}$ than to use second-order arithmetic. The task naturally arises:

develop formal languages that will enable us to reason directly about objects of analysis, having virtually the same computational complexity as elementary analysis itself.

More precisely, since probability theory plays one of the leading roles in the contemporary mathematical philosophy — and yet, according to Kolmogorov, can be treated as a branch of functional analysis — we are aiming to search for such a language among those for reasoning about probabilities, meeting the substantial conditions:

- the quantifier-free fragment of the language is simple enough from the viewpoint of computability theory, viz. the validity problem for it is algorithmically decidable;
- only one type of quantifiers is available in the language (that is, $\forall$ and $\exists$ ranging over the unique sort of objects) — this allows us to introduce the prefix classification in traditional fashion, and to avoid going deep into the investigation of possible relationships between different sorts of objects;
- no quantifiers may occur within the scope of the probability symbol in the formulas of the language;
- the quantification employed must be natural algebraically and intuitively attractive from the perspective of probability theory (or statistics), and the syntax/semantics of the language should be easily describable.

For instance, some logics of Abadi and Halpern (1994) are Π^1_∞-complete, but neither of them fulfils the second or third condition; further discussion of the fourth condition will appear in Section 5 — in analysing advantages of the present proposal.

Next, we provide a bunch of examples of probabilistic languages of the desired kind, and briefly sketch the principal results to be proved in the remaining sections.

Let $\mathscr{X} = \{x_i \mid i \in \mathbb{N}\}$ be the collection of *variables* and $C = \{c_i \mid i \in I\}$ the collection of *constants*, where I is a non-empty initial segment of $\mathbb{N}$. Define the set of all *e-terms* to be the smallest set having the properties:

- $\varnothing$ is an e-term, and every element of $\mathscr{X} \cup C$ is an e-term;
- if t_1 and t_2 are e-terms, then $\overline{t_1}$ and $t_1 \cap t_2$ are also e-terms.

The special unary function symbol μ plays the role of the intended probability measure and applies to e-terms. By a QPL^C-*atom* we mean an expression of the form

$$f\,(\mu\,(t_1), \ldots, \mu\,(t_n)) \;\leqslant\; g\,(\mu\,(t_{n+1}), \ldots, \mu\,(t_{n+k}))$$

where f and g are polynomials with coefficients in $\mathbb{Q}$, and $t_1, \ldots, t_{n+k}$ are e-terms. The QPL^C-*formulas* are obtained from the QPL^C-atoms by closing under $\neg$, $\wedge$ and the applications of $\forall x$, with $x \in \mathscr{X}$. As usual, we abbreviate

$$T_1 = T_2 \;:=\; T_1 \leqslant T_2 \wedge T_2 \leqslant T_1, \quad T_1 \neq T_2 \;:=\; \neg T_1 = T_2,$$
$$\Phi_1 \vee \Phi_2 \;:=\; \neg\,(\neg\Phi_1 \wedge \neg\Phi_2), \quad \Phi_1 \to \Phi_2 \;:=\; \neg\Phi_1 \vee \Phi_2 \quad \text{and} \quad \exists x\,\Phi_1 := \neg\forall x\,\neg\Phi_1$$

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

(here $T_1 \leqslant T_2$ and $T_2 \leqslant T_1$ are QPL^C-atoms, Φ_1 and Φ_2 are QPL^C-formulas). Denoting by $\mathscr{X}^*$ the set of all tuples of elements of $\mathscr{X}$, for any $(v_1, \dots, v_k) \in \mathscr{X}^*$ and $\mathsf{Q} \in \{\forall, \exists\}$, we write $\mathsf{Q}(v_1, \dots, v_k)$ as shorthand for $\mathsf{Q}v_1 \dots \mathsf{Q}v_k$. A QPL^C-formula is in Π_n (Σ_n) iff it has the form

$$\underbrace{\forall \overline{x}_1 \, \exists \overline{x}_2 \dots}_{n-1 \text{ alternations}} \; \Psi \quad (\underbrace{\exists \overline{x}_1 \, \forall \overline{x}_2 \dots}_{n-1 \text{ alternations}} \; \Psi)$$

with $\{\overline{x}_1, \overline{x}_2, \dots\} \subseteq \mathscr{X}^*$ and Ψ quantifier-free. Given a QPL^C-formula Φ, take

$$C(\Phi) \;:=\; \text{the set of constants which occur in } \Phi,$$
$$FV(\Phi) \;:=\; \text{the set of variables that occur free in } \Phi.$$

The languages under consideration will be interpreted over the class of *discrete probability spaces* (cf. Billingsley (1995)), each of which is represented by a triple

$$\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$$

where Ω is an at most countable set of *possible worlds*, $\mathscr{A} = \{S \mid S \subseteq \Omega\}$ (obviously $\mathscr{A}$ is a sigma-algebra), and P is a *discrete probability measure* on $\mathscr{A}$ — so there should be a distribution $\mathsf{p} : \Omega \to [0, 1]$ with $\sum_{\omega \in \Omega} \mathsf{p}(\omega) = 1$ that determines P via:

$$\mathsf{P}(S) \;=\; \sum\nolimits_{\omega \in S} \mathsf{p}(\omega) \quad \text{for every} \quad S \subseteq \Omega.$$

A QPL^C-*structure* is a discrete probability space augmented by a valuation from $\mathscr{X} \cup C$ into $\mathscr{A}$. Now we turn to describing the relation $\Vdash$ for QPL^C. Given a QPL^C-structure

$$\mathcal{M} \;=\; (\mathscr{P}, v) \text{ with } v : \mathscr{X} \cup C \to \mathscr{A},$$

expand v to the e-terms by interpreting, inductively, $\overline{t_1}$ as the complement of t_1, $t_1 \cap t_2$ as the intersection of t_1 and t_2; for a quantifier-free QPL^C-formula Φ, define

$$\mathcal{M} \Vdash \Phi \quad \Longleftrightarrow \quad \begin{array}{l} \text{the result of replacing each } \mu(t) \text{ in } \Phi \text{ by } \mathsf{P}(v(t)) \text{ is} \\ \text{true in the ordered field } \mathfrak{R} = \langle \mathbb{R}, +, \times, \leqslant \rangle \text{ of reals} \end{array}$$

— this is, essentially, a variation on the quantifier-free probability logic from Fagin *et al.* (1990, Section 5). The above relation $\Vdash$ is then extended to arbitrary QPL^C-formulas (in the style of Tarski) by the instructions:

- the connectives $\neg$ and $\wedge$ are treated clasically;
- the quantifier $\forall$ is viewed as ranging over all events in $\mathscr{A}$.

We refer to such languages as *probability logics with quantifiers over events*. Notice: the meaning of Φ in $\mathcal{M}$ does not depend on the values assigned by v to the elements of

$$(\mathscr{X} \cup C) \setminus (FV(\Phi) \cup C(\Phi))$$

— so these may be omitted. Further, for a space $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ and a *partial* valuation v from a subset of $\mathscr{X} \cup C$ into $\mathscr{A}$, define

$$\Phi \textit{ is valid over } (\mathscr{P}, v) \quad \Longleftrightarrow \quad (\mathscr{P}, v') \Vdash \Phi \text{ for any extension } v' : \mathscr{X} \cup C \to \mathscr{A} \text{ of } v.$$

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

In application to classes of spaces,

$$\Phi \textit{ is valid over } K \quad \Longleftrightarrow \quad \Phi \text{ is valid over each } \mathscr{P} \in K.$$

Actually, it all easily generalises to *arbitrary* probability spaces as well, which should be kept in mind. However, the present exposition — except for discussion in Section 5 — is primarily concerned with discrete ones.

Two QPL^C-formulas Φ_1 and Φ_2 are *semantically equivalent* (in symbols, $\Phi_1 \sim \Phi_2$) iff for any QPL^C-structure $\mathcal{M}$,

$$\mathcal{M} \Vdash \Phi_1 \quad \Longleftrightarrow \quad \mathcal{M} \Vdash \Phi_2.$$

The reader may easily verify that every QPL^C-formula is semantically equivalent to one in Π_n or Σ_n, for a suitable $n \in \mathbb{N}$. Call a QPL^C-sentence (i.e. a QPL^C-formula with no free variables) *valid* if it is semantically equivalent to $0 \leqslant 0$ — in other words, it holds in all QPL^C-structures. Along with the entire problem of testing validity for QPL^C-sentences comes the *hierarchy of validity problems for* QPL^C containing, for each $n \in \mathbb{N}$,

$$\begin{aligned} \Pi_n\text{-}Val^C &:= \text{ the set of valid } \Pi_n\text{-}\mathsf{QPL}^C\text{-sentences,} \\ \Sigma_n\text{-}Val^C &:= \text{ the set of valid } \Sigma_n\text{-}\mathsf{QPL}^C\text{-sentences.} \end{aligned}$$

Hence (modulo an appropriate Gödel numbering) we have

$$\Pi_n\text{-}Val^C \leqslant_m \Pi_{n+1}\text{-}Val^C,\ \Sigma_{n+1}\text{-}Val^C \quad \text{and} \quad \Sigma_n\text{-}Val^C \leqslant_m \Sigma_{n+1}\text{-}Val^C,\ \Pi_{n+1}\text{-}Val^C.$$

Such a hierarchy *collapses* in case there exists n with the property that for every $k \geqslant n$,

$$\Pi_k\text{-}Val^C \leqslant_m \Pi_n\text{-}Val^C \quad (\text{which implies } \Pi_k\text{-}Val^C \equiv_m \Pi_n\text{-}Val^C).$$

Clearly one may switch from Π to Σ here.

Remark: the validity problem for the quantifier-free fragment of QPL^C is easily shown to be decidable by an argument exploited earlier in Fagin *et al.* (1990), via m-reduction to determining membership in the first-order theory of $\mathfrak{R}$, along with implementation of Tarski's decision procedure (Tarski 1951).

The remainder of this article is organised as follows. In Section 2 we prove that each logic QPL^C has the desired complexity of elementary analysis (this can be carried over to several enrichments and reducts of QPL^C). Section 3 characterises all the maximum decidable prefix fragments: Π_2-Val^C is computable, while Σ_2-Val^C is not. In Section 4 we show that the corresponding hierarchies do not collapse. Section 5 discusses certain properties of QPL^C's and their generalisations, paying attention to the aforementioned classification of probability spaces. Remark: the proofs in Sections 2 and 4 exploit some observations from Speranski (2013b) concerning the monadic second-order definability in the standard model $\mathfrak{N}^+ = \langle \mathbb{N}, + \rangle$ of Presburger arithmetic; while in Section 3 we use a valuable result of Nies (1996), combined with a fact from Speranski (2011).

2. Computational complexity

The argument for the complexity result is quite explicit, and makes use of the following three basic observations:

- the concept of being an atom of the sigma-algebra of a given probability space, modulo events of measure zero, is definable in our logics (atoms will serve as 'bricks' in building a copy of $\mathbb{N}$);
- obviously $\langle \mathbb{N}, + \rangle$ is isomorphic to $\langle \{1/2^{n+1} \mid n \in \mathbb{N}\}, \times \rangle$, and hence we want to view $n \in \mathbb{N}$ as a distinguished event with probability $1/2^{n+1}$;
- treating atoms as 'first-order objects,' we can deal with quantifiers over events (in a given space) as with monadic second-order quantifiers.

Note: analogues of the second observation were employed, e.g. in Hoover (1978); Abadi and Halpern (1994) for probability logics which are very different from QPL^C. However, even the first two observations together provide no computational complexity bound — they only say that a fragment of the first-order decidable theory of $\mathfrak{N}^+$ is interpretable. As a matter of fact, the choice of $\mathfrak{N}^+$ is not very important — because the technique in Speranski (2013b) demonstrates how some other reducts of $\mathfrak{N}$ can be taken instead. So the crucial part of the reasoning exploits the conjunction of the first and third observations. This should not be overlooked, for in Section 5 we apply the Π^1_∞-completeness in a natural context irrelevant for other quantified probabilistic logics.

Theorem 2.1. The validity problem for QPL^C is Π^1_∞-complete.

Proof. The argument naturally falls into two parts. We begin by showing that determining whether a given sentence of second-order arithmetic is true in $\mathfrak{N}$ can be computably reduced to testing validity for $\mathsf{QPL}^{\{c\}}$-sentences, where $c = c_0$. The main task here consists in familiarising ourselves with atoms and their applications.

Lemma 2.1. $\bigcup_{n \in \mathbb{N}} \Pi_n\text{-}Val^C$ is Π^1_∞-hard.

Proof. A substantial role in our reasoning will be played by the following formulas:

$$
\begin{aligned}
x_1 \preccurlyeq x_2 &:= \mu(\overline{x_1} \cup x_2) = 1; \\
x_1 \sim x_2 &:= x_1 \preccurlyeq x_2 \wedge x_2 \preccurlyeq x_1; \\
At(x_1) &:= \mu(x_1) > 0 \wedge \forall x_2\,((\mu(x_2) > 0 \wedge x_2 \preccurlyeq x_1) \to x_2 \sim x_1); \\
Nat &:= \exists x_1 \big(At(x_1) \wedge \mu(x_1) = 1/2\big) \wedge \forall x_1 \,\big(At(x_1) \to \exists x_2 \big(At(x_2) \wedge \mu(x_2) = \mu(x_1)/2\big)\big).
\end{aligned}
$$

Let $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ be a discrete probability space, and take $\Lambda := \{E \in \mathscr{A} \mid \mathsf{P}(E) = 0\}$, i.e. the events of measure zero. Obviously, for any $\{E, E'\} \subseteq \mathscr{A}$, we have:

$$
\begin{aligned}
\mathscr{P} \Vdash E \preccurlyeq E' &\iff E \text{ is a subset of } E', \text{ modulo } \Lambda; \\
\mathscr{P} \Vdash E \sim E' &\iff E \text{ equals } E', \text{ modulo } \Lambda; \\
\mathscr{P} \Vdash At(E) &\iff \text{there is no event strictly between } \varnothing \text{ and } E, \text{ modulo } \Lambda.
\end{aligned}
$$

In effect, since $x_1 \sim x_2$ defines a congruence relation on the Boolean algebra $\mathscr{A}$, $At(E)$ means that the equivalence class $[E]_\sim$ is an atom of the corresponding quotient-algebra $\mathscr{A}_\sim$. Hence Nat asserts the existence of a sequence of events $E_0, E_1, \ldots$ in $\mathscr{A}$ such that for every $n \in \mathbb{N}$,

$$E_n \text{ is an atom, modulo } \Lambda, \text{ with } \mathsf{P}(E_n) = 1/2^{n+1}.$$

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

In addition, $\sum_{n\in\mathbb{N}} \mathsf{P}(E_n) = 1$, thus $\{E_n\}_{n\in\mathbb{N}}$ is a partition of Ω, modulo Λ. Consequently it ensures that

$$\mathscr{P} \Vdash At(E) \iff \mathscr{P} \Vdash E \sim E_n \text{ for some } n \in \mathbb{N}$$

— this part allows us to identify each E_n (or rather, $[E_n]_\sim$) with the natural number n (compare (Speranski 2013a, Section 4)), and then interpret $\mathfrak{N}^+$ using

$$n + m = k \iff 1/2^{n+1} \times 1/2^{m+1} = 1/2^{k+1} \times 1/2.$$

Remark: although the last equivalence and a formula analogous to *Nat* were employed e.g. in certain proofs of Abadi and Halpern (1994), the notion of an atom did not — and could not — play any role there (as well as quantifiers over events).

Another important observation is that for every $S \subseteq \mathbb{N}$, its union $E_S := \bigcup_{n\in S} E_n$ belongs to the sigma-algebra $\mathscr{A}$, so

$$n \in S \iff \mathscr{P} \Vdash E_n \preccurlyeq E_S.$$

Intuitively, we treat each $n \in \mathbb{N}$ as the singleton $\{n\}$. The monadic second-order theory of $\mathfrak{N}^+$ is Π^1_∞-complete — see an alternative characterisation of the analytical hierarchy obtained in Speranski (2013b, Section 3). Obviously, we want to m-reduce this special theory to the valid $\mathsf{QPL}^{\{c\}}$-sentences. As can be readily checked, in the language of the former every formula is effectively converted into a logically equivalent one all of whose atomic subformulas have the form

$$x_n + x_m = x_k \quad \text{and} \quad x_n \in X_m,$$

i.e. to what is called an L_+-*formula* below. Next the translation τ from L_+ to $\mathsf{QPL}^{\{c\}}$ is described by recursion:

$$\begin{aligned}
\tau(x_n + x_m = x_k) &:= \mu(x_{2n}) \times \mu(x_{2m}) = \mu(x_{2k}) \times 1/2;\\
\tau(x_n \in X_m) &:= x_{2n} \preccurlyeq x_{2m+1};\\
\tau(\neg\varphi) &:= \neg\tau(\varphi);\\
\tau(\varphi \wedge \psi) &:= \tau(\varphi) \wedge \tau(\psi);\\
\tau(\forall x_n\, \varphi) &:= \forall x_{2n}\, (At(x_{2n}) \to \tau(\varphi));\\
\tau(\forall X_m\, \varphi) &:= \forall x_{2m+1}\, \tau(\varphi).
\end{aligned}$$

In a straightforward way, the above considerations imply that for any L_+-sentence φ,

$$\varphi \text{ is true in } \mathfrak{N} \iff Nat \to \tau(\varphi) \text{ is valid in } \mathsf{QPL}^{\{c\}} \qquad (\dagger)$$

(remark: discrete probability spaces satisfying *Nat* certainly do exist — the reader may easily construct such a space if needed). Hence the validity problem for $\mathsf{QPL}^{\{c\}}$, as well as for QPL^C, is at least Π^1_∞-hard. □

The same proof works when we pass from discrete to arbitrary probability spaces — because the former may be identified with 'spaces consisting of atoms.'

Lemma 2.2. $\bigcup_{n\in\mathbb{N}} \Pi_n$-$Val^C$ is Π^1_∞-bounded.

Proof. Now formulas of QPL^C should be encoded as those of second-order arithmetic in an appropriate manner. And since each discrete space $\langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ is uniquely determined by $\mathsf{p} : \Omega \to [0, 1]$ with

$$\mathsf{p}(\omega) \;=\; \mathsf{P}\left(\{\omega\}\right) \quad \text{for every} \quad \omega \in \Omega$$

and, further, Ω may be viewed as the initial segment S of $\mathbb{N}$ with $|S| = |\Omega|$, this will be almost immediate. The reals and the sequences of reals are well-known to be representable in second-order arithmetic (consult (Simpson 2009, Chapter I)), so we can take X_0 to range over all functions f of the sort

$$f : S \to [0, 1] \;\text{ for some initial segment } S \text{ of } \mathbb{N}, \quad \text{with } \sum\nolimits_{n \in S} f(n) \;=\; 1$$

(or rather, subsets of $\mathbb{N}$ representing such functions), while the variables from $\{X_n\}_{n=1}^{\infty}$ continue to range over all subsets of $\mathbb{N}$. For any QPL^C-sentence Φ, let $\rho(\Phi)$ be obtained from Φ by replacing

- each $\mu(t)$ by $\lim_{n\to\infty} \sum_{m \in S \cap \{0,\dots,n\} \cap t} X_0(m)$,
- each occurrence of c_n by X_{2n+1},
- each occurrence of x_m by X_{2m+2}.

(of course, usual set-theoretic operations and the notions of bounded sum and limit are definable in second-order arithmetic). It is then straightforward to verify that

$$\Phi \text{ is valid in } \mathsf{QPL}^C \quad \Longleftrightarrow \quad \overline{\forall X}\, \rho(\Phi) \text{ is true in } \mathfrak{N} \qquad (\ddagger)$$

where $\overline{\forall X}$ abbreviates the list of universal quantifiers binding all free variables of $\rho(\Phi)$, including X_0. Thus the validity problem for QPL^C is at worst Π^1_∞. □

The rest is trivial. □

There are two simple corollaries worth mentioning:

1. the validity problem for QPL^C without $+$ is Π^1_∞-complete;
2. the validity problem for QPL^C augmented by quantifiers over $\mathbb{R}$ is Π^1_∞-complete.

The former is justified by the observation that $\tau(\varphi)$ in the proof of Lemma 2.1 doesn't contain $+$; and the latter is because the proof of Lemma 2.2 can be readily adapted to take quantifiers over $\mathbb{R}$ into account — so the updated ρ will again produce formulas of second-order arithmetic. Note: although $\times$ can be dropped as well, the argument would then become far more complicated and it falls beyond the scope of the paper — involving another version of the analytical hierarchy from Speranski (2013b), stated in terms of the monadic second-order theory of $\langle \mathbb{N}, s \rangle$ with one free binary predicate.

Of course, in the proof of Theorem 2.1 the class of discrete probability spaces (determining the validity in QPL^C) may be replaced by a single $\mathscr{P}$ — like that with

$$\Omega \;=\; \mathbb{N}, \quad \mathscr{A} \;=\; \{S \mid S \subseteq \mathbb{N}\}, \quad \text{and} \quad \mathsf{P}(S) \;=\; \sum\nolimits_{n \in S} 1/2^{n+1} \quad \text{for all} \quad S \subseteq \mathbb{N}.$$

Further, whenever we pass from discrete to arbitrary spaces, the lower bound argument survives, while the upper bound argument fails — so we get the Π^1_∞-hardness and may be uncertain about the Π^1_∞-boundedness. Actually, the latter will hold, but the demonstration

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

exploits a special technique related to 'atomless spaces,' while these deserve an independent investigation. For the opposite effect: if we restrict attention to a reasonable collection of spaces excluding

distinguished sequences of events whose probabilities form an appropriate geometric progression

(given $k \geqslant 2$, the construction can be easily modified to deal with $1/k$ instead of $1/2$ by working in a portion of Ω of probability $\sum_{n=1}^{\infty} 1/k^n$), then the lower bound argument fails. E.g. the QPL^C-theory of the class of all finite spaces is co-recursively enumerable, see Proposition 5.1 below. Yet the next section reveals how to obtain the undecidability in many such situations. At the same time there are very natural classes of non-discrete spaces, whose theories turn out to be decidable: here one finds a lot of suitable geometric progressions but cannot distinguish the corresponding sequences of events. Actually, any class of 'atomless spaces' serves as an example, but the proof of this fact falls beyond the scope of our presentation, being based on completely different techniques.

To emphasise that the first two items from the beginning of this section are not crucial for deriving computational complexity bounds, and to illustrate the importance of the third item, let us consider a version of QPL^C in which

quantifiers over events are replaced by those over atoms,

i.e. relativise every $\forall x$ ($\exists x$) by $At(x)$. Then Nat is again satisfiable, but now τ only reduces the decidable $\mathrm{Th}\left(\mathfrak{N}^+\right)$ to the corresponding validity problem — and hence provides no complexity bound. (Note: even though an essentially different kind of translation can be offered here — which, in contrast, will eventually lead to the Π^1_1-completeness — the demonstration would take us too far aside, and falls beyond the scope of the paper.)

Another curious observation about Nat is that for $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ we have

$$\mathscr{P} \Vdash Nat \quad \Longrightarrow \quad \{\mathsf{P}(E) \mid E \in \mathscr{A}\} \; = \; [0,1], \tag{$\sharp$}$$

because for any $r \in [0,1]$, there exists $S \subseteq \mathbb{N}$ with $r = \sum_{n \in S} 1/2^n$. But the above proof does not exploit ($\sharp$). Instead, we employ the underlying property that

quantifiers range over all elements of $\mathscr{A}$ which is closed under the formation of countable unions,

and hence no union of atoms (modulo events of measure zero) of $\mathscr{A}$ is avoided. And of course, the conclusion of ($\sharp$), together with its premise, can easily be falsified, by any $\mathscr{P}$ in which $[0,1]$ is unavailable. For instance, this is very different from the situation with two-sorted logics of Abadi and Halpern (1994) where one has, in particular, quantifiers over reals — so the whole of $\langle \mathbb{R}, +, \times, \leqslant \rangle$ is built into each language model (even over finite probability spaces). While for a given QPL^C-structure $\mathcal{M}$, we only deal with reals expressible as probabilities in $\mathcal{M}$, or their arithmetical combinations.

3. Maximum decidable prefix fragments

Although Theorem 2.1 establishes the Π^1_∞-complexity, its proof has a rather small range of applications — as was already mentioned in the previous section. Indeed, even if the

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

class of discrete probability spaces is concerned, one can hardly expect that it may give us the minimum undecidable prefix fragments (however, the Π^1_n-completeness of some non-optimal levels can be obtained using the argument of Theorem 4.1 below). Now we shift attention from certain special-purpose machinery employed in proving complexity results to more general undecidability techniques.

So we turn to the investigation of the decision problem for the probabilistic formulae classification of Section 1, viz. for the prefix fragments of QPL^C (the reader may note the parallel with stating the Skolem–Bernays–Shönfinkel classification of decision problems for pure first-order predicate logic, cf. Börger *et al.* (1997)). And a method of studying elementary theories will help us.

Let σ be a signature (for first-order logic) and Val_σ the collection of σ-sentences true in all σ-structures. Traditionally, we call a set Γ of σ-sentences *hereditarily undecidable* (*h.u.*) iff for every Δ,

$$Val_\sigma \cap \Gamma \subseteq \Delta \subseteq \Gamma \implies \Delta \text{ is undecidable}$$

(see Nies (1996) and references therein). A useful example is given by Nies (1996, Theorem 4.2), which establishes the hereditary undecidability of the first-order $\exists\forall$-theory of the class $\mathcal{K}_\circ$ of all finite symmetric irreflexive graphs. An immediate yet helpful observation: for any two classes $\mathcal{K}_1$ and $\mathcal{K}_2$ of σ-structures, and each $n \in \mathbb{N}$,

$$\mathcal{K}_1 \subseteq \mathcal{K}_2, \text{ the } \Sigma_n\text{-theory of } \mathcal{K}_1 \text{ is h.u.} \implies \text{the } \Sigma_n\text{-theory of } \mathcal{K}_2 \text{ is h.u.} \qquad (\natural)$$

This will be exploited below for $n = 2$, i.e. $\Sigma_n = \exists\forall$.

Theorem 3.1. The validity problem for $\forall\exists$-QPL^C-sentences is decidable, while the validity for $\exists\forall$-QPL^C-sentences is undecidable.

Proof. Once again, the argument naturally falls into two parts.

Lemma 3.1. Π_2-Val^C is computable.

Proof. We aim to reduce the corresponding problem to testing validity for quantifier-free $\mathsf{QPL}^{\mathscr{C}}$-sentences, where $\mathscr{C} = \{c_n\}_{n\in\mathbb{N}}$ — this involves a variation on the 'grounding idea' (whose standard application is the decidability of the $\forall\exists$-fragment of Val_σ in case σ does not contain function symbols (Bernays and Schönfinkel 1928)). Consider a space $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P}\rangle$. One easily checks that for any QPL^C-sentence of the form $\forall x_1 \ldots \forall x_n \Phi$ and $v : C(\Phi) \to \mathscr{A}$,

$$\forall x_1 \ldots \forall x_n \Phi \text{ is true in } (\mathscr{P}, v) \iff \Phi\left[x_1/c_{k+1}, \ldots, x_n/c_{k+n}\right] \text{ is valid over } (\mathscr{P}, v)$$

where $k := \sup\{i \mid c_i \text{ occurs in } \Phi\}$ (indeed, the new constants $c_{k+1}, \ldots, c_{k+n}$ can be assigned arbitrary elements of $\mathscr{A}$ and the variables $x_1, \ldots, x_n$ range over $\mathscr{A}$). Further, for every $v : S \to \mathscr{A}$ with $S \subset \mathscr{C}$ finite, we define

$$T_S := \text{the set of ground } e\text{-terms built up from } S \text{ (including } \varnothing),$$
$$\mathscr{A}_S := \{v(t) \mid t \in T_S\} \quad \text{and} \quad \mathsf{P}_S := \text{the restriction of } \mathsf{P} \text{ to } \mathscr{A}_S.$$

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

Obviously $\mathscr{P}_S = \langle \Omega, \mathscr{A}_S, \mathsf{P}_S \rangle$ is a finite subspace of $\mathscr{P}$. Next, for any QPL^C-sentence of the form $\exists x_1 \ldots \exists x_n \Psi$ with Ψ quantifier-free and $v : S \to \mathscr{A}$ with $S = C(\Psi)$,

$$\exists x_1 \ldots \exists x_n \Psi \text{ is true in } (\mathscr{P}_S, v) \iff \bigvee_{t_1 \in D_S} \cdots \bigvee_{t_n \in D_S} \Psi\left[x_1/t_1, \ldots, x_n/t_n\right] \text{ is true in } (\mathscr{P}, v)$$

where D_S denotes the finite collection of all full disjunctive normal forms built up from S (subject to the replacement of $\wedge$ by $\cap$, and so on) plus $\varnothing$ — the required equivalence is a direct consequence of the two simple facts:

- $\mathscr{A}_S$ coincides with $\{v(t) \mid t \in D_S\}$;
- for a quantifier-free QPL^C-sentence Θ with $C(\Theta) \subseteq S$,

$$\Theta \text{ is true in } (\mathscr{P}, v) \iff \Theta \text{ is true in } (\mathscr{P}_S, v).$$

Hence, in the present notation, we get

$$\forall x_1 \ldots \forall x_n \Phi \text{ is valid} \iff \Phi\left[x_1/c_{k+1}, \ldots, x_n/c_{k+n}\right] \text{ is valid};$$
$$\exists x_1 \ldots \exists x_n \Psi \text{ is valid} \iff \bigvee_{t_1 \in D_S} \cdots \bigvee_{t_n \in D_S} \Psi\left[x_1/t_1, \ldots, x_n/t_n\right] \text{ is valid}.$$

The rest is straightforward (remember the decidability of Π_0-Val^C). □

The above ideas appear to be helpful in providing decision procedures for $\forall\exists$-theories of various classes of probability spaces, especially when finite subspaces are included.

Lemma 3.2. Σ_2-Val^C is incomputable.

Proof. Consider $\sigma := \{G^2\}$ (G is a predicate symbol of arity 2), i.e. the signature of graphs. Every $\mathsf{QPL}^{\{c\}}$-structure $\mathcal{M} = (\langle \Omega, \mathscr{A}, \mathsf{P} \rangle, c \mapsto E_c)$ with E_c a fixed element of $\mathscr{A}$ induces the σ-structure $\mathcal{M}_\sigma$ with domain $\{E \in \mathscr{A} \mid \mathsf{P}(E) = \mathsf{P}(E_c)\}$ such that

$$\mathcal{M}_\sigma \Vdash G(E_1, E_2) \iff \mathcal{M} \Vdash \Phi_G(E_1, E_2)$$

where $\Phi_G(x_1, x_2) := x_1 \nsim x_2 \wedge \mu(x_1 \wedge x_2) \neq \mu(c)$. Noting that the binary relation

$$\left\{(E_1, E_2) \in \mathscr{A}^2 \mid \mathcal{M} \Vdash E_1 \sim E_2\right\}$$

is a congruence for $\mathcal{M}_\sigma$, we denote the corresponding quotient graph by $(\mathcal{M}_\sigma)_\sim$. Given a σ-sentence φ, let $\alpha(\varphi)$ be the result of replacing each $G(t_1, t_2)$ in φ by $\Phi_G(t_1, t_2)$. In particular, we immediately have

$$\mathcal{M}_\sigma \Vdash \varphi \iff (\mathcal{M}_\sigma)_\sim \Vdash \varphi \iff \mathcal{M} \Vdash \alpha(\varphi),$$

hence the first-order $\exists\forall$-theory of the class

$$\mathcal{K} := \left\{(\mathcal{M}_\sigma)_\sim \mid \mathcal{M} \text{ is a } \mathsf{QPL}^{\{c\}}\text{-structure}\right\},$$

denoted $\exists\forall$-Th$(\mathcal{K})$, is m-reducible to Σ_2-$Val^{\{c\}}$ via the translation α. What remains is to show that the problem of testing membership in $\exists\forall$-Th$(\mathcal{K})$ is undecidable. And this, in turn, is an immediate consequence of ($\natural$) and the following known facts:

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

- the first-order $\exists\forall$-theory of $\mathcal{K}_\circ$ (the finite symmetric irreflexive graphs) is h.u. (Nies 1996, Theorem 4.2);
- $\mathcal{K}_\circ$ coincides with

$$\mathcal{K}_\bullet \;:=\; \left\{ (\mathcal{M}_\sigma)_\sim \mid \mathcal{M} \text{ is a } \mathsf{QPL}^{\{c\}}\text{-structure with } |\Omega| < \infty \text{ and } \mathsf{P}(\mathscr{A}) \subseteq \mathbb{Q} \right\}$$

— see (Speranski 2011, Section 3).

Thus $\mathcal{K}_\circ \subseteq \mathcal{K}$ and $\exists\forall$-Th$(\mathcal{K})$ is even h.u. □

The rest is trivial. □

Further, adopting the above notation, we get

Corollary 3.1. Let $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ be a space and K a class of QPL^C-structures.

1. For any $\forall\exists$-QPL^C-sentence Φ and $v : C(\Phi) \to \mathscr{A}$,

$$\begin{array}{c}\Phi \text{ is true in all } (\mathscr{P}', v) \text{ with} \\ \mathscr{P}' \text{ a finite subspace of } \mathscr{P}\end{array} \quad \Longrightarrow \quad \Phi \text{ is true in } (\mathscr{P}, v);$$

2. If $\mathcal{K}_\bullet \subseteq \{(\mathcal{M}_\sigma)_\sim \mid \mathcal{M} \in \mathsf{K}\}$, then the $\exists\forall$-QPL^C-theory of K is undecidable.

The first item (being a sort of 'quantifier elimination' for $\forall\exists$-formulas) ensures that a $\forall\exists$-QPL^C-sentence is valid iff it is true in all finite QPL^C-structures.
An example of the application of the second item:

Corollary 3.2. Given a subfield $\mathbb{F}$ of $\mathbb{R}$ (thus $\mathbb{Q} \subseteq \mathbb{F}$), the $\exists\forall$-QPL^C-theory of the class of all finite probability spaces satisfying $\mathsf{P}(\mathscr{A}) \subseteq \mathbb{F}$ is undecidable.

At the same time, there are at least two directions for generalisation:

— pass from discrete to arbitrary probability spaces in the QPL^C-semantics;
— replace '$\exists\forall$-QPL^C-sentences' by '$\exists\forall$-QPL^C-sentences without $+$ and $\times$'

(the latter is justified by the observation that $\alpha(\varphi)$ in the proof of Lemma 3.2 does not contain $+$ and $\times$). Note: the condition '$\mathcal{K}_\bullet \subseteq \{(\mathcal{M}_\sigma)_\sim \mid \mathcal{M} \in \mathsf{K}\}$' can also be weakened in many cases, since there is plenty of freedom in interpreting graphs, as the experience of computability theory shows — however, both the formulation and the demonstration would then become rather technical.

Let me end up with a curious example. Consider

$$Fin \;:=\; \exists x_1 \left(\mu(x_1) > 0 \wedge \forall x_2 \left(\mu(x_2) > 0 \to \mu(x_1) \leqslant \mu(x_2)\right)\right).$$

Clearly Fin is semantically equivalent to a $\exists\forall$-$\mathsf{QPL}^{\{c\}}$-sentence. It is easy to check that for every probability space $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ we have

$$Fin \text{ holds in } \mathscr{P} \quad \Longleftrightarrow \quad \{[A]_\sim \in \mathscr{A}_\sim \mid \mathsf{P}(A) > 0\} \text{ is finite},$$

i.e. iff $\mathscr{P}$ is finite modulo events of measure zero. Thus, the finiteness is expressible by means of Fin. On the contrary, no $\forall\exists$-QPL^C-sentence can do this job.

4. Non-collapsing hierarchies

Finally, we need to prove that there are infinitely many pairwise non-m-equivalent elements of the non-decreasing sequence

$$\Pi_0\text{-}Val^C \;\leqslant_m\; \Pi_1\text{-}Val^C \;\leqslant_m\; \Pi_2\text{-}Val^C \;\leqslant_m\; \dots$$

And the proof will make an essential use of the alternative characterisation of the analytical hierarchy from Speranski (2013b).

Clearly, we want to employ the translation τ from Section 2. Still, here is one delicate point to deal with:

$\Pi^1_n(\Sigma^1_n)$-sentences of second-order arithmetic may contain arbitrary many alternations of first-order quantifiers.

Thus, since τ replaces quantifiers over $\mathbb{N}$ and those over subsets of $\mathbb{N}$ by quantifiers over events, but every element of $\Pi_n\text{-}Val^C$ ($\Sigma_n\text{-}Val^C$) contains only $n-1$ alternations in its prefix, we need a suitable tool for restricting the number of first-order quantifier alternations in the monadic second-order theory of $\mathfrak{N}^+$.

Let $\sigma\,[n,k]$ denote the collection of all monadic second-order σ-sentences of the form

$$\underbrace{\forall X_1\,\exists X_2\,\dots}_{n-1 \text{ alternations}}\;\underbrace{\exists \overline{x}_1\,\forall \overline{x}_2\,\dots}_{k-1 \text{ alternations}}\;\psi,$$

with ψ quantifier-free; $\sigma\,[n] := \bigcup_{k\in\mathbb{N}} \sigma\,[n,k]$. Take σ_* and σ_+ to be the signatures of $\mathfrak{N}$ and $\mathfrak{N}^+$ respectively. It is known that for each $n > 0$,

$$\{\varphi \in \sigma_*\,[n,2] \mid \varphi \text{ is true in } \mathfrak{N}\}$$

is Π^1_n-complete. Further, as was shown in Speranski (2013b, Section 3), there exists an effective translation ζ with the properties:

$$\begin{aligned} \varphi \in \sigma_*\,[n,2] \quad &\Longleftrightarrow \quad \zeta(\varphi) \in \sigma_+\,[n,4]\,; \\ \varphi \text{ is true in } \mathfrak{N} \quad &\Longleftrightarrow \quad \zeta(\varphi) \text{ is true in } \mathfrak{N}^+ \end{aligned}$$

— which implies the Π^1_n-completeness of the set

$$\left\{\varphi \in \sigma_+\,[n,4] \mid \varphi \text{ is true in } \mathfrak{N}^+\right\}.$$

We exploit this technical fact below.

Theorem 4.1. The hierarchy of validity problems for QPL^C does not collapse.

Proof. We begin by demonstrating the Π^1_n-hardness (where $n \in \{1,2,\dots\}$) of certain validity problems for prefix fragments of QPL^C. Assume φ belongs to $\sigma_+\,[n,4]$, and so has the form

$$\underbrace{\forall X_1\,\exists X_2\,\dots}_{n-1 \text{ alternations}}\;\exists \overline{x}_1\,\forall \overline{x}_2\,\exists \overline{x}_3\,\forall \overline{x}_4\,\psi$$

with ψ quantifier-free. By the definition of τ (from the proof of Lemma 2.1),

$$\tau(\varphi) \;=\; \underbrace{\forall x_3\,\exists x_5\,\dots}_{n-1 \text{ alternations}}\;\tau\left(\exists \overline{x}_1\,\forall \overline{x}_2\,\exists \overline{x}_3\,\forall \overline{x}_4\,\psi\right)$$

and also

$$\tau(\exists\overline{x}_1\,\forall\overline{x}_2\,\exists\overline{x}_3\,\forall\overline{x}_4\,\psi) \sim \exists\overline{y}_1\,(At\,(\overline{y}_1)\wedge\tau\,(\forall\overline{x}_2\,\exists\overline{x}_3\,\forall\overline{x}_4\,\psi)) \sim \ldots \sim$$
$$\exists\overline{y}_1\,(At\,(\overline{y}_1)\wedge\forall\overline{y}_2\,(At\,(\overline{y}_2)\rightarrow\exists\overline{y}_3\,(At\,(\overline{y}_3)\wedge\forall\overline{y}_4\,(At\,(\overline{y}_4)\rightarrow\tau\,(\psi)))))$$

where for any $i \in \{1,2,3,4\}$, the tuple $\overline{y}_i$ is obtained from $\overline{x}_i$ by replacing each x_m with x_{2m}, and the expression $At\,(\overline{y}_i)$ abbreviates

$$\bigwedge_{x_m \text{ occurs in } \overline{y}_i} At\,(x_m).$$

Clearly $\tau\,(\psi)$ is quantifier-free and $At\,(x_1)$ is semantically equivalent to a Π_1-QPL^C-formula. Thus

$$\begin{aligned}
\tau(\exists\overline{x}_1\,\forall\overline{x}_2\,\exists\overline{x}_3\,\forall\overline{x}_4\,\psi) &\sim \text{a } \Sigma_5\text{-}\mathsf{QPL}^C\text{-formula},\\
\tau\,(\varphi) &\sim \text{a } \Pi_{n+5}\text{-}\mathsf{QPL}^C\text{-sentence},\\
Nat &\sim \text{a } \Pi_3\text{-}\mathsf{QPL}^C\text{-sentence},\\
Nat \rightarrow \tau\,(\varphi) &\sim \text{a } \Pi_{n+5}\text{-}\mathsf{QPL}^C\text{-sentence}.
\end{aligned}$$

Remark: all these reductions can be performed effectively (and uniformly in n). Then, in view of the property (†) of τ, we conclude that determining membership in

$$\{\varphi \in \sigma_+\,[n,4] \mid \varphi \text{ is true in } \mathfrak{N}^+\}$$

is m-reducible to testing validity for Π_{n+5}-QPL^C-sentences, so Π_{n+5}-Val^C turns out to be at least Π^1_n-hard.

On the other hand, remembering the description of ρ (from the proof of Lemma 2.2), it is straightforward to derive the existence of N such that for any Π^1_n-QPL^C-sentence Φ,

$$\overline{\forall X}\,\rho\,(\Phi) \text{ is equivalent to a sentence in } \sigma_*\,[n+N] \text{ with respect to } \mathfrak{N}$$

(the value of N is not essential to our presentation, although it may depend slightly on the arithmetical coding employed). Since this reduction can be exhibited in an effective way, the corresponding property (‡) of ρ ensures that the computational complexity of Π_n-Val^C is bounded by Π^1_{n+N}.

The rest is almost trivial. Indeed, suppose the hierarchy for QPL^C does collapse, i.e. there is n satisfying the condition that for every $k \geqslant n$, Π_k-$Val^C \leqslant_m \Pi_n$-Val^C. Letting $k := n+N+6$, we would have that

$$\Pi_k\text{-}Val^C \text{ is } \Pi^1_{n+N+1}\text{-hard} \quad\text{and}\quad \Pi_k\text{-}Val^C \text{ is at worst } \Pi^1_{n+N}$$

— which is obviously a contradiction, as desired. □

Note that the reasons why we can eliminate $\times$ from (the description of) the monadic second-order arithmetic are, in fact, quite natural — see the definability results in Speranski (2013b, Section 3). Thus the proofs of Theorems 2.1 and 4.1 demonstrate explicitly how the analytical hierarchy emerges in our probabilistic versions of elementary analysis. In addition, the latter gives us a bound on the size of the gap between Π^1_n and the complexity of Π_n-Val^C.

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

5. Further discussion

It is worth comparing the present proposal with some other approaches in the area. For instance, the probability logic with quantifiers over propositional formulas — which we denote $\mathsf{QPL}_\circ$ — was introduced in Speranski (2011), with essentially the same formulas and probability structures as for $\mathsf{QPL}^{\mathscr{C}}$. The significant distinction concerns the semantical treatment of quantifiers in $\mathsf{QPL}_\circ$: for every $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$,

$\forall$ is viewed as ranging over all the events of $\mathscr{A}$ definable by ground e-terms (i.e. the events having the form $v(t)$ for an e-term t with no variables)

and thus the domain of quantification is always at most countable. Indeed, this feature turns out to be crucial for the issues of expressibility, namely we have:

- both the full validity problem for $\mathsf{QPL}_\circ$ and its restriction to Σ_4-sentences are Π^1_1-complete — so the hierarchy of validity problems for $\mathsf{QPL}_\circ$ collapses, capturing the complexity of the universal fragment of Π^1_∞ (see (Speranski 2013c, Theorem 2)).

In sharp contrast to this, Theorems 2.1 and 4.1 together show that, in a precise sense, the m-degrees corresponding to the members of the sequence

$$\Sigma_0\text{-}Val^C \ \leqslant_m \ \Sigma_1\text{-}Val^C \ \leqslant_m \ \Sigma_2\text{-}Val^C \ \leqslant_m \ \ldots$$

(or of its companion with Π in place of Σ) come infinitely close to Π^1_∞ — which is never actually attained but appears as the 'limit.' In effect, the analytical hierarchy behaves in a similar manner. Note: $\mathsf{QPL}_\circ$ can be interpreted in probabilistic formalisms augmented by countable conjunctions or disjunctions, like those of Keisler (1985), Paris (2011), etc. — see (Speranski 2013a, Section 2.3); while an analogous argument fails for each QPL^C, because the range of quantifiers may easily have the cardinality of the continuum — e.g. it is true of every QPL^C-structure satisfying the sentence *Nat* (from the proof of Lemma 2.1).

Further, it may be instructive to mention several observations concerning the striking differences between the present proposal and the approach of Abadi and Halpern (1994), where certain Π^1_∞-completeness results were proved. Let $\mathscr{L}$ be a probability logic from the article by Abadi and Halpern. Then:

1. unlike QPL^C, $\mathscr{L}$ is always two-sorted, explicitly including *quantifiers over reals*;
2. unlike QPL^C, $\mathscr{L}$ is not based on the quantifier-free language from Fagin *et al.* (1990) and the underlying propositional classical logic — in fact, events are now expressed in terms of two-sorted first-order formulas, so in particular, both types of quantifiers may occur in the scope of μ;
3. unlike for QPL^C, we cannot safely switch to quotients of $\mathscr{L}$-structures — the set of true $\mathscr{L}$-sentences may be different (actually, some of the main arguments in Abadi and Halpern (1994) fail over discrete spaces, but go through for 'discrete spaces enriched with uncountably many elements of measure zero');
4. unlike in QPL^C, to interpret $\mathscr{L}$-formulas over $\langle \Omega, \mathscr{A}, \mathsf{P} \rangle$, we refer only to events definable via two-sorted first-order formulas with parameters from $\mathbb{R}$ and either Ω or a special 'domain' U, not to all elements of $\mathscr{A}$, and hence $\mathscr{L}$ is completely insensible to adding to $\mathscr{A}$ or removing from $\mathscr{A}$ 'undefinable events';

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

5. unlike in QPL^C, neither the notion of an atom nor the finiteness property for underlying probability spaces (both modulo events of measure zero) is expressible in $\mathscr{L}$;
6. unlike for QPL^C, extending the semantics of $\mathscr{L}$ to deal with arbitrary spaces meets the well-known measurability problem for sets definable by first-order formulas:

a projection of a measurable set is not necessarily measurable.

In particular — because of Items 1–2, every prefix classification for the $\mathscr{L}$-formulas has to be rather involved. For these reasons, the issues raised in Sections 3 and 4, as well as various other claims about probabilistic elementary analysis or its hierarchy, do not look relevant here (moreover, the decision problem directly reduces to those in Börger *et al.* (1997) in many cases). These and other features of $\mathscr{L}$, such as those we have been discussing above, play key roles in the Π^1_∞-hardness arguments of Abadi and Halpern (1994), but they may conflict with some desirable conditions (in particular, algebraic and logical ones) on quantified probabilistic logics. At the same time, neither of the aforementioned problems occurs in QPL^C. We proceed with a natural classification of probability spaces closely related to Boolean algebras (which is irrelevant, however, for many other quantified probabilistic logics — e.g. because of analogues of Item 4).

Consider an arbitrary probability space $\mathscr{P} = \langle \Omega, \mathscr{A}, \mathsf{P} \rangle$. Adopting the notation of the proof of Theorem 2.1, we take

$$S \;:=\; \{[E]_\sim \mid At(E) \text{ holds in } \mathscr{P}\},$$

i.e. the collection of all atoms of $\mathscr{A}_\sim$. One can easily check that S is at most countable, and hence there exists the supremum of S in $\mathscr{A}_\sim$ — let A denote it. Define

$$\mathtt{ch}_1(\mathscr{P}) \;:=\; \begin{cases} n & \text{if } |S| = n, \ n \in \mathbb{N} \\ \infty & \text{when } S \text{ is infinite} \end{cases} \quad \text{and} \quad \mathtt{ch}_2(\mathscr{P}) \;:=\; \begin{cases} 0 & \text{if } \mathsf{P}^*(A) = 1 \\ 1 & \text{otherwise} \end{cases}$$

where $\mathsf{P}^*(A) := \mathsf{P}(E)$ for $E \in A$. Hence, we obtain a straightforward adaptation of the important notion of an elementary invariant; cf. Koppelberg (1989).

Note: each $\mathscr{P}$ with $\mathtt{ch}_2(\mathscr{P}) = 0$ can be identified with a suitable discrete probability space, modulo events of measure zero; and if $\mathscr{P}$ is discrete, then $\mathtt{ch}_2(\mathscr{P}) = 0$. Clearly

$$\mathtt{ch}_2(\mathscr{P}) \;=\; 0 \quad \Longleftrightarrow \quad \mathscr{A}_\sim \text{ is an atomic Boolean algebra.}$$

These interactions between probability spaces and Boolean algebras appear naturaly in probability logics with quantifiers over events, while in many other languages they are almost completely abandoned. For instance, the above items for $\mathscr{L}$ make the distance between

what is available in a space and *what is expressible formally*

even bigger. Returning to our logics, we observe that for each $\zeta \in (\mathbb{N} \cup \{\infty\}) \times \{0, 1\}$,

$$K_\zeta \;:=\; \{\mathscr{P} \mid \mathscr{P} \text{ is a probability space with } (\mathtt{ch}_1(\mathscr{P}), \mathtt{ch}_2(\mathscr{P})) = \zeta\}$$

is a definable class in QPL^C, viz. there exists a QPL^C-sentence Φ_ζ for which

$$\Phi_\zeta \text{ holds in } \mathscr{P} \quad \Longleftrightarrow \quad \mathscr{P} \in K_\zeta$$

— e.g. when $\zeta = (\infty, 0)$, we take

$$\Phi_\zeta \;:=\; \neg Fin \wedge \forall x_1 \, \exists x_2 \, (\mu(x_1) > 0 \rightarrow At(x_2) \wedge x_2 \preccurlyeq x_1).$$

Another two important definable classes are

$$K^0_{\mathsf{fin}} \;:=\; \bigcup_{n \in \mathbb{N}} K_{(n,0)} \quad \text{and} \quad K^1_{\mathsf{fin}} \;:=\; \bigcup_{n \in \mathbb{N}} K_{(n,1)}.$$

For an arbitrary class K of probability spaces (not necessarily discrete), let

$$\mathrm{Th}^C(K) \;:=\; \text{the set of } \mathsf{QPL}^C\text{-sentences valid over all spaces in } K.$$

A simple modification of the proof of Theorem 2.1 yields

Theorem 5.1. $\mathrm{Th}^C\left(K_{(\infty,0)}\right)$ is Π^1_∞-complete. $\mathrm{Th}^C\left(K_{(\infty,1)}\right)$ is at least Π^1_∞-hard.

Although the latter theory turns out to be Π^1_∞-bounded as well, more advanced tools are needed to prove this: spaces in $K_{(\infty,1)}$ contain 'atomless counterparts' — so coding from Section 2 is no longer available. Further, similarly to Section 4, for each $n \in \mathbb{N}$ we get the Π^1_n-hardness of certain prefix fragments of $\mathrm{Th}^C\left(K_{(\infty,0)}\right)$ and $\mathrm{Th}^C\left(K_{(\infty,1)}\right)$.

Another simple application of the above technique is

Proposition 5.1. For any $n \in \mathbb{N}$, $\mathrm{Th}^C\left(K_{(n,0)}\right)$ is decidable.

Proof. The case $n = 0$ is trivial, because $K_{(0,0)} = \varnothing$.

Suppose $n > 0$. Letting $U = \{1, \ldots, n\}$ we define Φ_n to be

$$At(x_1) \wedge \cdots \wedge At(x_n) \wedge \mu(x_1 \vee \cdots \vee x_n) = 1 \wedge \bigwedge_{\{i,k\} \subseteq U,\, i \neq k} \mu(x_i \wedge x_k) = 0$$

— thus $\exists x_1 \ldots x_n \, \Phi_n$ characterises $K_{(n,0)}$. If Φ_n holds in $\langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ for some $x_1, \ldots, x_n$, then every element of $\mathscr{A}$ can be identified with a disjunction in

$$D \;:=\; \left\{ \bigvee\nolimits_{i \in I} x_i \mid I \subseteq U \right\},$$

modulo events of measure zero. Consequently, for a QPL^C-sentence Φ,

$$\Phi \text{ holds over } K_{(n,0)} \quad \Longleftrightarrow \quad \forall x_1 \ldots x_n \left(\Phi_n \rightarrow \Phi^*\right) \text{ is valid,}$$

where Φ^* is the result of replacing each $\forall x$ ($\exists x$) in Φ by $\bigwedge_{x \in D}$ (respectively $\bigvee_{x \in D}$). In addition, $\forall x_1 \ldots x_n \left(\Phi_n \rightarrow \Phi^*\right)$ is easily reduced to a semantically equivalent $\forall\exists$-form. It only remains to apply Lemma 3.1. □

As was already mentioned in Section 3, $\exists\forall$-$\mathrm{Th}^C\left(K^0_{\mathsf{fin}}\right)$ is undecidable (even in case $+$ and $\times$ are excluded from the language). We also have

Proposition 5.2. For $K \in \left\{K^1_{\mathsf{fin}}, K_{(\infty,0)}, K_{(\infty,1)}\right\}$, $\exists\forall$-$\mathrm{Th}^C(K)$ is undecidable.

Proof. Suppose $K \in \left\{K^1_{\mathsf{fin}}, K_{(\infty,0)}, K_{(\infty,1)}\right\}$.

Fix some $\langle \Omega_\star, \mathscr{A}_\star, \mathsf{P}_\star \rangle \in K$. For a finite probability space $\langle \Omega, \mathscr{A}, \mathsf{P} \rangle$ with $\Omega \cap \Omega_\star = \varnothing$, we take

$$r \;:=\; \min\left\{ \mathsf{P}(E_1) - \mathsf{P}(E_2) \mid E_1 \in \mathscr{A},\; E_2 \in \mathscr{A} \text{ and } \mathsf{P}(E_1) > \mathsf{P}(E_2) \right\}$$

https://doi.org/10.1017/S0960129516000189 Published online by Cambridge University Press

and consider the composite $\mathscr{P}' = \langle \Omega', \mathscr{A}', \mathsf{P}' \rangle$ where

$$\Omega' = \Omega \cup \Omega_\star, \quad \mathscr{A}' = \{E \cup E_\star \mid E \in \mathscr{A},\ E_\star \in \mathscr{A}_\star\},$$
$$\mathsf{P}'(E \cup E_\star) = \frac{3-r}{3} \times \mathsf{P}(E) + \frac{r}{3} \times \mathsf{P}_\star(E_\star) \quad \text{for all} \quad E \cup E_\star \in \mathscr{A}'.$$

As is easily verified, $\mathscr{P}'$ belongs to K (since $\mathscr{A}$ contains only finitely many events). Let $E_c \in \mathscr{A}$ be a potential value of the constant symbol c. For every $E' \in \mathscr{A}'$,

$$\mathsf{P}'\left(E'\right) = \mathsf{P}'(E_c) \quad \Longrightarrow \quad \mathsf{P}\left(E' \cap \Omega_\star\right) = 0.$$

Indeed, assuming $\mathsf{P}'\left(E'\right) = \mathsf{P}'(E_c)$, if $\mathsf{P}'\left(E' \cap \Omega_\star\right) > 0$, then

$$\frac{r}{3} \geqslant \mathsf{P}'\left(E' \cap \Omega_\star\right) = \mathsf{P}'\left(E'\right) - \mathsf{P}'\left(E' \cap \Omega\right) = \mathsf{P}'(E_c) - \mathsf{P}'\left(E' \cap \Omega\right) \geqslant r \times \frac{3-r}{3}$$

(by the construction of r and $\mathscr{P}'$), a contradiction to $0 \leqslant r \leqslant 1$ — thus the probability of $E' \cap \Omega_\star$ is zero, as desired. Consequently, adopting the notation of Section 3, for the two $\mathsf{QPL}^{\{c\}}$-structures

$$\mathcal{M} = \left(\langle \Omega, \mathscr{A}, \mathsf{P} \rangle, c \mapsto E_c\right) \quad \text{and} \quad \mathcal{M}' = \left(\langle \Omega', \mathscr{A}', \mathsf{P}' \rangle, c \mapsto E_c\right),$$

the corresponding graphs $(\mathcal{M}_\sigma)_\sim$ and $\left(\mathcal{M}'_\sigma\right)_\sim$ turn out to be isomorphic. It remains to apply the second corollary from the same section. □

Note that the argument will not go through for, say, $K_{(0,1)}$, even though every finite probability space can be easily extended to an atomless space: the problem is that we cannot distinguish the initial 'finite counterpart' in such spaces. Indeed, the atomless spaces deserve an independent investigation, which is beyond the scope of the paper. To sum up, there is an interesting analogy between the properties of the suggested probability logics and those of Boolean algebras. At the same time, the ordered field of reals continues to play an important role in the investigation. The two directions are naturally combined in probability logics with quantifiers over events.

Actually, the foregoing classification may turn out to be useful in philosophical logic, too. For instance, the most important role in Leitgeb (2013) is played by the concept of a *probabilistically stable set* in a given space. One easily checks that each such set must consist of finitely many atoms and, modulo events of measure zero, the concept itself is definable in QPL^C, while the axioms for Leitgeb's probabilistic belief theory are expressible in QPL^C with an additional predicate symbol *Bel*.

The writing of this article was supported by a Marie Curie International Research Staff Exchange Scheme Fellowship within the Seventh European Community Framework Programme (project 'Computable analysis — theoretical and applied aspects'), and also by the Russian Foundation for Basic Research (project RFBR-13-01-00015-a).

References

Abadi, M. and Halpern, J.Y. (1994). Decidability and expressiveness for first-order logics of probability. *Information and Computation* **112** (1) 1–36.

Bernays, P. and Schönfinkel, M. (1928). Zum Entscheidungsproblem der mathematischen Logik. *Mathematische Annalen* **99** (1) 342–372.

Billingsley, P. (1995). *Probability and Measure*, John Wiley & Sons.

Börger, E., Grädel E. and Gurevich, Y. (1997). *The Classical Decision Problem*, Springer.

Fagin, R., Halpern, J.Y. and Megiddo, N. (1990). A logic for reasoning about probabilities. *Information and Computation* **87** (1–2) 78–128.

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

Keisler, H.J. (1985). Probability quantifiers. In: Barwise, J. and Feferman, S. (eds.) *Model-Theoretic Logics*, Springer 509–556.

Koppelberg, S. (1989). General theory of Boolean algebras. In: Monk, J.D. and Bonnet, R. (eds.) *Handbook of Boolean Algebras*, vol. 1, North-Holland 1–311.

Leitgeb, H. (2013). Reducing belief simpliciter to degrees of belief. *Annals of Pure and Applied Logic* **164** (12) 1338–1389.

Leitgeb, H. (2016). Probability in logic. In: Hájek, A. and Hitchcock, C. (eds.) *The Oxford Handbook of Probability and Philosophy*, Oxford University Press. To appear.

Nies, A. (1996). Undecidable fragments of elementary theories. *Algebra Universalis* **35** (1) 8–33.

Paris, J.B. (2011). Pure inductive logic. In: Horsten, L. and Pettigrew, R. (eds.) *The Continuum Companion to Philosophical Logic*, Continuum 428–449.

Rogers, H. (1967). *Theory of Recursive Functions and Effective Computability*, McGraw-Hill.

Simpson, S.G. (2009). *Subsystems of Second Order Arithmetic*, Cambridge University Press.

Solovay, R.M., Arthan, R.D. and Harrison, J. (2012). Some new results on decidability for elementary algebra and geometry. *Annals of Pure and Applied Logic* **163** (12) 1765–1802.

Speranski, S.O. (2011). Quantification over propositional formulas in probability logic: decidability issues. *Algebra and Logic* **50** (4) 365–374.

Speranski, S.O. (2013a). Complexity for probability logic with quantifiers over propositions. *Journal of Logic and Computation* **23** (5) 1035–1055.

Speranski, S.O. (2013b). A note on definability in fragments of arithmetic with free unary predicates. *Archive for Mathematical Logic* **52** (5–6) 507–516.

Speranski, S.O. (2013c). Collapsing probabilistic hierarchies. I, *Algebra and Logic* **52** (2) 159–171.

Suppes, P., de Barros, J.A. and Oas, G. (1998). A collection of probabilistic hidden-variable theorems and counterexamples. In: Pratesi, R. and Ronchi, L. (eds.) *Conference Proceedings Vol. 60, Waves, Information and Foundations of Physics*, Bologna 267–291.

Tarski, A. (1951). *A Decision Method for Elementary Algebra and Geometry*, University of California Press.

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