Quantifying over Bernoulli random variables in probability logic

Stanislav O. Speranski

Sobolev Institute of Mathematics

Novosibirsk State University

Novosibirsk 2013

Let A and B be sets of natural numbers.

Say that A is *m-reducible* to B (denoted $A \leqslant_m B$) iff there exists a computable function $f : \mathbb{N} \to \mathbb{N}$ satisfying

$$n \in A \iff f(n) \in B;$$

A and B are called *m-equivalent* (denoted $A \equiv_m B$) iff $A \leqslant_m B$ and $B \leqslant_m A$. Now we define $[A] := \{B \mid A \equiv_m B\}$.

Further, identify each problem specified by a question of the type

Whether a given input has the desired property?

with the set of inputs for which the answer is affirmative, and view, in turn, this set as a collection of natural numbers.

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 the standard model $\mathfrak{N}$, and $\mathscr{P}_\infty$ to be the full second-order theory of $\mathfrak{N}$.

The analytical hierarchy includes the following milestones: $\Pi^1_n := [\mathscr{P}_n]$ and $\Sigma^1_n := [\mathscr{S}_n]$ for all $n \in \mathbb{N}$. Define $\Pi^1_\infty := [\mathscr{P}_\infty]$.

A portion of the related terminology: for $\lambda \in \mathbb{N} \cup \{\infty\}$,

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

and similarly for Σ^1_λ (in place of Π^1_λ) with $\lambda \in \mathbb{N}$.

We present a bunch of quantified probability logics each of which *has the complexity of* $\mathscr{P}^1_\infty$ and, in addition, obeys the conditions:

- the validity problem for its quantifier-free fragment is decidable;
- only two quantifiers, $\forall$ and $\exists$, are available in the logic, both ranging over the unique sort of objects;
- no quantifiers may occur in the scope of the probability symbol, i. e., the formulas cannot contain $\mu\,(\ldots\forall\ldots)$ or $\mu\,(\ldots\exists\ldots)$;
- the quantification employed must be intuitively attractive from the viewpoint of probability theory, and the syntax/semantics of the logic should be easily describable.

Let $\mathscr{X} = \{x_i \mid i \in \mathbb{N}\}$ and $C = \{c_i \mid i \in I\}$,
where I is a non-empty computable subset of $\mathbb{N}$.

The collection of e-terms is the smallest set containing $\mathscr{X} \cup C$, and s. t. if t_1 and t_2 are e-terms, then $\overline{t_1}$ and $t_1 \cap t_2$ are also e-terms.

Definition

By a QPL^C-atom we mean an expression of the sort

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

where f and g are polynomials with coefficients in $\mathbb{Q}$, μ is a fixed special symbol, and $t_1, \ldots, t_{n+m}$ are e-terms.

The QPL^C-formulas are obtained from the QPL^C-atoms by closing under $\neg$, $\wedge$ and applications of $\forall x$, with $x \in \mathscr{X}$. As usual, $\exists x\, \Phi := \neg \forall x\, \neg \Phi$.

A QPL^C-formula belongs to Π_n (Σ_n) iff it has the form

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

with $\{\overline{x}_1, \overline{x}_2, \ldots\}$ a set of tuples from $\mathscr{X}$ and Ψ quantifier-free.

A QPL^C-structure is a discrete probability space $\langle \Omega, \mathscr{A}, \mathrm{P} \rangle$ augmeted by a valuation $v : \mathscr{X} \cup C \to \mathscr{A}$. So Ω is an at most countable set, $\mathscr{A} = \{S \mid S \subseteq \Omega\}$, and P is a *discrete probability measure* on $\mathscr{A}$ determined by a distribution $\mathrm{p} : \Omega \to [0, 1]$ s. t.

$$\sum_{\omega \in \Omega} \mathrm{p}(\omega) = 1, \quad \text{and} \quad \mathrm{P}(S) = \sum_{\omega \in S} \mathrm{p}(\omega) \quad \text{for all } S \subseteq \Omega.$$

$$\mathcal{M} \;=\; (\langle \Omega, \mathscr{A}, \mathrm{P} \rangle, v : \mathscr{X} \cup C \to \mathscr{A})$$

Let's expand v from $\mathscr{X} \cup C$ to the e-terms by interpreting $\overline{t_1}$ as the complement of t_1, $t_1 \cap t_2$ as the intersection of t_1 and t_2. And for every quantifier-free formula Φ of QPL^C, naturally define

$$\mathcal{M} \Vdash \Phi \quad \Longleftrightarrow \quad \begin{array}{l} \text{the result of replacing each } \mu(t) \\ \text{in } \Phi \text{ with } \mathrm{P}(v(t)) \text{ is true in } \mathbb{R} \end{array}$$

(which is, essentially, a variation on the quantifier-free probability logic of Fagin–Halpern–Megiddo). We extend $\Vdash$ to all QPL^C-formulas by:

- *treating the connectives $\neg$ and $\wedge$ clasically*;
- *viewing the quantifier $\forall$ as ranging over all events of $\mathscr{A}$*.

Call a QPL^C-sentence valid if it holds in any QPL^C-structure.

Along with the problem of testing validity for all QPL^C-sentences comes the hierarchy of validity problems for QPL^C containing

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

— hence we have $\Pi_n\text{-}Val^C \leqslant_m \Pi_{n+1}\text{-}Val^C,\ \Sigma_{n+1}\text{-}Val^C$ and $\Sigma_n\text{-}Val^C \leqslant_m \Sigma_{n+1}\text{-}Val^C,\ \Pi_{n+1}\text{-}Val^C$. Such a hierarchy collapses if there exists n fulfilling the condition:

$$\text{for each } k \geqslant n,\ \Pi_k\text{-}Val^C \equiv_m \Pi_n\text{-}Val^C$$

(clearly, one may switch from Π to Σ here).

Before proceeding, it is helpful to list some observations.

Since every event is uniquely determined by its characteristic function, quantifiers over events correspond to quantifiers over Bernoulli random variables — so the quantification employed in QPL^C is very appealing from the viewpoint of probability theory.

In addition, the logics QPL^C are closely related to the logic with quantifiers over propositions, and hence are indirectly connected with formalisms introduced by H. J. Keisler, J. B. Paris, etc.

The validity problem for quantifier-free QPL^C-sentences is easily shown to be decidable by an argument of Fagin–Halpern–Megiddo, via m-reduction to determining membership in $\mathrm{Th}\,(\langle \mathbb{R}, +, \times, \leqslant \rangle)$ along with implementation of the Tarski's decision procedure.

Each logic QPL^C has the same complexity as elementary analysis:

Theorem

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

And there are infinitely many pairwise non-m-equivalent elements of the nondecreasing sequence $\Pi_0\text{-}Val^C \leqslant_m \Pi_1\text{-}Val^C \leqslant_m \Pi_2\text{-}Val^C \leqslant_m \ldots$

Theorem

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

Notice: both proofs exploit an alternative description of the analytical hierarchy without $\times$ (which generalises the result of Halpern about the Π^1_1-completeness of the theory of $\langle \mathbb{N}, + \rangle$ with a free unary predicate).

We now turn to the investigation of the decision problems for QPL^C, viz. the characterisation of all maximum prefix fragments of QPL^C among those for which the validity problem is decidable.

Theorem

The validity problem for Π_2-QPL^C-sentences is decidable, while the validity problem for Σ_2-QPL^C-sentences is undecidable.

Notice: the proof employs the technique of first-order elementary definability and some related results. (And one may see the parallel with formulating the Skolem–Bernays–Shönfinkel classification of decision problems for pure first-order predicate logic.)

Let's present the probability logic $\mathrm{QPL}_\circ$ with quantifiers over propositions in the following way:

- the $\mathrm{QPL}_\circ$-formulas and the $\mathrm{QPL}_\circ$-structures are the same as for $\mathrm{QPL}^{\mathscr{C}}$ with $\mathscr{C} = \{c_i\}_{i \in \mathbb{N}}$;
- in $\mathrm{QPL}_\circ$ the atoms, $\wedge$ and $\neg$ are viewed semantically as in $\mathrm{QPL}^{\mathscr{C}}$;
- the significant distinction concerns the treatment of quantifiers in $\mathrm{QPL}_\circ$: for every $\mathcal{M} = (\langle \Omega, \mathscr{A}, \mathrm{P} \rangle, v : \mathscr{X} \cup \mathscr{C} \to \mathscr{A})$,

  $\forall$ *ranges over all events of* $\{v(t) \mid t \textit{ is a ground e-term}\}$

  (and thus the domain of quantification is at most countable).

Even though the maximum decidable prefix fragments of $\mathrm{QPL}_\circ$ turn out to be the same as of $\mathrm{QPL}_{\mathscr{C}}$, the two logics differ strikingly from the perspective of expressibility.

Namely, as was proved earlier, we have:

- the validity problem for $\mathsf{QPL}_\circ$ is Π^1_1-complete;
- the hierarchy of validity problems for $\mathsf{QPL}_\circ$ collapses — because the Π^1_1-completeness result already holds for the Σ_4-sentences.

In sharp contrast to this, as we have already found out, 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' — and, in effect, the analytical hierarchy behaves in a similar manner.