Some Probabilistic Riddles and Some Logical Solutions

Radboud University Nijmegen
Dedicated to Herman Geuvers, May 29, 2024

Bart Jacobs
bart@cs.ru.nl

Outline

Where we are, so far

Geuvers on the future of maths

> *"We believe that these systems will become the future of mathematics, where definitions, statements, computations and proofs are all available in a computerized form."*

From: Herman Geuvers. Proof assistants: History, ideas and future, 2009

- Geuvers' work concentrates on the formalisation techniques
- less on actual formalisations

An area that urgently requires such formalisation is probability theory
- much decision making is now left to AI, based on correlations
- whether you get a mortgage, or medical treatment
- clear need for explainable AI (XAI), with a "logic"

Challenges in probabilistic logic (from Pearl'89)

> *To those trained in traditional logics, symbolic reasoning is the standard, and nonmonotonicity a novelty. To students of probability, on the other hand, it is symbolic reasoning that is novel, not nonmonotonicity. Dealing with new facts that cause probabilities to change abruptly from very high values to very low values is a commonplace phenomenon in almost every probabilistic exercise and, naturally, has attracted special attention among probabilists. The new challenge for probabilists is to find ways of abstracting out the numerical character of high and low probabilities, and cast them in linguistic terms that reflect the natural process of accepting and retracting beliefs.*

Embarrassingly, there is still **no** probabilistic logic for symbolic reasoning. Probabilistic updating will have to be crucial, non-trivial ingredient.

Own Festschrift paper

- Six riddles in (discrete) probability theory, as reasoning tests
  - copied from website briddles.com, many with updating
- Not in the language of a symbolic probabilistic logic (yet)
  - but in an interpreted mathematical language
  - with explicit operations on distributions: products, updating
- Inspiration comes from categorical probability theory
  - new, refreshing, more structured approach, with string diagrams
  - traditional probabilistic language is confusing (and untyped!)
  - it is a hindrance to progress
- Own opinion: a new formalism is badly needed for probability theory
  - Wittgenstein: "The limits of my language mean the limits of my world"
  - See also recent artile: *Getting Wiser from Multiple Data: Probabilistic Updating according to Jeffrey and Pearl*, on ArXiv
  - Big differences in answers of 17 experts on standard question

Aside about Wittgenstein

In 1992 Herman and I tried went looking for Wittgenstein's grave in Cambridge

It took us a whole afternoon to find it. Nowadays there is a website with a nice video: britishwittgensteinsociety.org/wittgensteins-grave

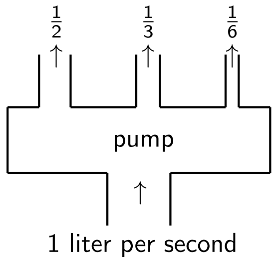
Where we are, so far

Coins and dices

- I write a fair coin as a distribution $\frac{1}{2}|H\rangle + \frac{1}{2}|T\rangle$
- Similarly, a dice is: $\frac{1}{6}|1\rangle + \frac{1}{6}|2\rangle + \frac{1}{6}|3\rangle + \frac{1}{6}|4\rangle + \frac{1}{6}|5\rangle + \frac{1}{6}|6\rangle$
- Suppose I roll this dice behind my hand and ask you the probability of each number / pip ...
  - of course, you will say: $\frac{1}{6}$
- Suppose I tell you that the outcome of the roll is even
  - what probability will you assign to each of the pips $1, 2, 3, 4, 5, 6$?
  - that is, what is the resulting, updated distribution?
  - It is: $\frac{1}{3}|2\rangle + \frac{1}{3}|4\rangle + \frac{1}{3}|6\rangle$.

A physical model of probabilistic updating I

Consider the water pump with different diameters of the outgoing pipes:

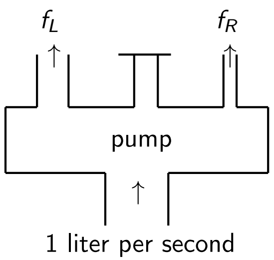


It represents a distribution, written as:

$$\omega = \frac{1}{2}|L\rangle + \frac{1}{3}|M\rangle + \frac{1}{6}|R\rangle.$$

A physical model of probabilistic updating II

Suppose now that we have evidence that the middle pipe is blocked:



The pump keeps on pushing 1 liter per second

What are the resulting left and right flows f_L and f_R ?

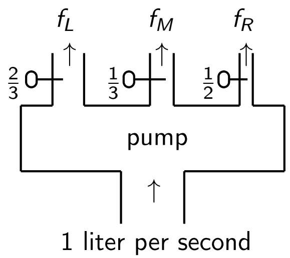
Recall the diameters: $\frac{1}{2}, \frac{1}{3}, \frac{1}{6}$

- The diameter ratio $L : R$ is $\frac{1}{2} : \frac{1}{6}$, which is $3 : 1$.
- Hence the updated distributions is: $\omega' = \frac{3}{4}|L\rangle + \frac{1}{4}|R\rangle$.
- More formally, we renormalise as in:

$$\omega' = \frac{1/2}{1/2 + 1/6}|L\rangle + \frac{1/6}{1/2 + 1/6}|R\rangle = \frac{3}{4}|L\rangle + \frac{1}{4}|R\rangle.$$

A physical model of probabilistic updating III

To make it more interesting, suppose the pipes can be closed by taps:



The numbers left of the taps are the fractions of openness

- The normalisation factor is: $\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{6} \cdot \frac{1}{2} = \frac{19}{36}$
- The updated distribution is:

$$\omega'' = \frac{1/2 \cdot 2/3}{19/36}|L\rangle + \frac{1/3 \cdot 1/3}{19/36}|M\rangle + \frac{1/6 \cdot 1/2}{19/36}|R\rangle = \frac{12}{19}|L\rangle + \frac{4}{19}|M\rangle + \frac{3}{19}|R\rangle$$

General formulations

- A distribution over a set X is a finite sum $\sum_i r_i | x_i \rangle$
  - where $x_i \in X$ and $r_i \in [0, 1]$ with $\sum_i r_i = 1$
- A (fuzzy) predicate is a function $p \colon X \to [0, 1]$
  - think of it as a collection of taps
- The validity (or expected value) $\omega \models p$ of predicate p in distribution ω is $\sum_i r_i \cdot p(x_i)$.
- The updated distribution $\omega|_p$ is the renormalisation:

$$\omega|_p = \sum_i \frac{r_i \cdot p(x_i)}{\omega \models p} | x_i \rangle.$$

- Basic fact: $(\omega|_p \models p) \geq (\omega \models p)$
  - Slogan: updating (learning) makes you wiser
  - The inequality does not occur in the literature — since distributions are left implicit in the traditional language

Where we are, so far

Life or death (from briddles.com)

You are a prisoner sentenced to death. The Emperor offers you a chance to live by playing a simple game. He gives you 50 black marbles, 50 white marbles and 2 empty bowls. He then says, 'Divide these 100 marbles into these 2 bowls. You can divide them any way you like as long as you use all the marbles. Then I will blindfold you and mix the bowls around. You then can choose one bowl and remove ONE *marble. If the marble is* WHITE *you will live, but if the marble is* BLACK *. . . you will die.'*
How do you divide the marbles up so that you have the greatest probability of choosing a WHITE *marble?*

Solution: Put one white marble in the first bowl and all other marbles in the second bowl

- Chance of survival: $\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{49}{99} = \frac{148}{198} = \frac{74}{99} \approx 0.747$

Life or Death, formalisation I

- We write $(w, b) \in \mathbb{N}^2$ for the white, black marbles in the first bowl
  - The second bowl is then contains $(50 - w, 50 - b)$.
- We use as space $\mathcal{MD}$ of marble divisions the set:

$$\mathcal{MD} \coloneqq \{(w, b) \mid 0 \leq w, b \leq 50 \text{ with } 0 < w + b < 100\}.$$

- This set $\mathcal{MD}$ has $51^2 - 2 = 2599$ elements
- We use it with uniform distribution:

$$\omega = \tfrac{1}{2599}\big|0,1\big\rangle + \tfrac{1}{2599}\big|0,2\big\rangle + \cdots + \tfrac{1}{2599}\big|50,48\big\rangle + \tfrac{1}{2599}\big|50,49\big\rangle$$

Life or Death, formalisation II

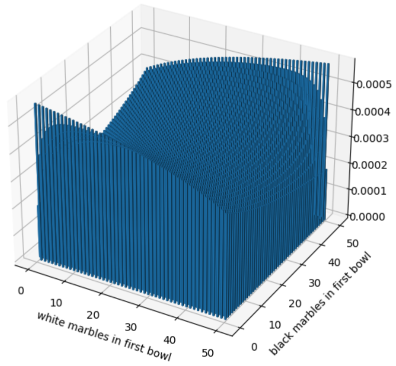
- The predicates $p_1, p_2 \colon MD \to [0,1]$ give the probability of drawing white from the first, resp. second bowl
- Explicitly, for $(w,b) \in MD$,

$$p_1(w,b) := \frac{w}{w+b} \qquad p_2(w,b) := \frac{50-w}{(50-w)+(50-b)} = \frac{50-w}{100-w-b}$$

- There is $\frac{1}{2}$ to choose each bow
  - So we use predicate $p := \frac{1}{2} \cdot p_1 + \frac{1}{2} \cdot p_2$
- We then form the update $\omega|_p$
  - Interpretation: given that we have drawn white (and survived), what are the probabilities?

Life or Death, formalisation III

Plot of updated distribution $\omega|_p$ on MD



Highest probabilities for white, black in first bowl: $(1,0)$ and $(49,50)$.

Where we are, so far

Suggestions and wishes

Developing a formal, symbolic logic and type theory for probability theory has scientific and societal urgency

- it should include non-trivial rules for updating

Happy Anniversary to Herman!