Lower and Upper Conditioning in Quantum Bayesian Theory

Simons Institute Berkeley

Bart Jacobs — Radboud University Nijmegen
bart@cs.ru.nl
Dec. 15, 2017

iCIS | Digital Security
Radboud University

Where we are, so far

Introduction

State updates and disintegration in classical probability

State updates and disintegration in quantum probability

Conclusions

Overview of the talk

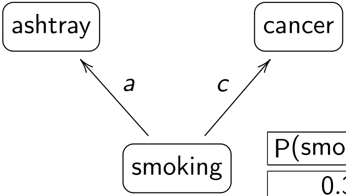
- Two key ingredients of classical Bayesian network theory
  (1) translations back-and-forth between joint states and Bayesian networks
  (2) equivalence of inference via updating of joint states and forward/backward reasoning via the network graph

  Point 1 is completely standard; point 2 less so

- Our aim is to investigate quantum analogues of these two points, towards "quantum Bayesian theory"
  (1) the above point 1 exists in [Leifer Spekkens 2013]
  (2) point 2 is unchartered territory: it will lead to two forms of quantum conditioning — which coincide classically.

A Bayesian network example

smoking	P(ashtray)
t	0.95
f	0.25



smoking	P(cancer)
t	0.4
f	0.05

P(smoking)
0.3

Two questions:

- what is the a prior probability of cancer? **Easy:** 0.155
- what the probability of cancer given an ashtray?
  - two solutions: via back/for-ward inference, or via the joint state

The two solutions in some detail

(1) **Via backward & forward inference** (see [Zanasi & Jacobs 2016])

- transform the ashtray predicate p along the ashtray channel a to a predicate $p' = a \ll p$ on the initial smoking state σ
- update σ with this predicate $\sigma' = \sigma|_{p'}$
- transform this state along cancer channel to $c \gg \sigma'$

(2) **Via the joint state**

- compute the joint state $\omega \in \mathcal{D}(A \times S \times C)$
- update with the (weakened) ashtray predicate p, to $\omega' = \omega|_{p \otimes \mathbf{1} \otimes \mathbf{1}}$
- take third ('cancer') marginal of ω'

The two solutions computed, in EfProb

The inference formulation:

```
c >> (smoking / (a << tt))
0.267|t> + 0.733|f>
```

The joint version:

```
joint
0.114|t,t,t> + 0.171|t,t,f> + 0.00875|t,f,t> +
   0.166|t,f,f> + 0.006|f,t,t> + 0.009|f,t,f> +
   0.0263|f,f,t> + 0.499|f,f,f>

(joint / (tt @ 1 @ 1)) % [0,0,1]
0.267|t> + 0.733|f>
```

Main ingredients

(1) states ω and predicates p

(2) state tranformation $c \gg \omega$ and predicate transformation $c \ll p$ along a channel c

(3) update $\omega|_p$ of a state ω with a predicate p

(4) translations between joint states and channels (“disintegration”)

- points 1 and 2 are sufficiently familiar — also in the quantum case
- we elaborate on 3 and 4 — first in classical discrete probability

Where we are, so far

Discrete probability distributions

Notation

- Fair coin: $\frac{1}{2}|H\rangle + \frac{1}{2}|T\rangle$
- Fair dice: $\frac{1}{6}|1\rangle + \frac{1}{6}|2\rangle + \frac{1}{6}|3\rangle + \frac{1}{6}|4\rangle + \frac{1}{6}|5\rangle + \frac{1}{6}|6\rangle$

ket notation

- $|-\rangle$ is pure syntactic sugar — stemming from quantum
- more confusing to omit them, as in: $\frac{1}{6}1 + \frac{1}{6}2 + \frac{1}{6}3 + \frac{1}{6}4 + \frac{1}{6}5 + \frac{1}{6}6$

- Write $\mathcal{D}(X)$ for the set of such probability distributions $\sum_i r_i|x_i\rangle$ where $x_i \in X$, $r_i \in [0,1]$ with $\sum_i r_i = 1$
- Distributions $\omega \in \mathcal{D}(X)$ will often be called states of X
- A predicate on a set X is a function $p\colon X \to [0,1]$
  - It is called sharp (non-fuzzy) if $p(x) \in \{0,1\}$ for each $x \in X$

Combining states and predicates

Let $\omega \in \mathcal{D}(X)$ be state/distribution, $p \in [0,1]^X$ a predicate, both on X.

- **Validity** $\omega \models p$, in $[0,1]$
  - defined as $\sum_x \omega(x) \cdot p(x)$
  - also known as expected value of p in state ω
- **Conditioning** $\omega|_p$, in $\mathcal{D}(X)$
  - assuming validity $\omega \models p$ is non-zero
  - defined as: $\omega|_p = \sum_x \frac{\omega(x) \cdot p(x)}{\omega \models p} \big| x \big\rangle$

Two basic laws of conditioning

We write $p \mathbin{\&} q$ for the pointwise product $(p \mathbin{\&} q)(x) = p(x) \cdot q(x)$ of predicates $p, q \in [0,1]^X$.

product rule

$$\omega|_p \models q \;=\; \frac{\omega \models p \mathbin{\&} q}{\omega \models p}$$

Bayes' rule

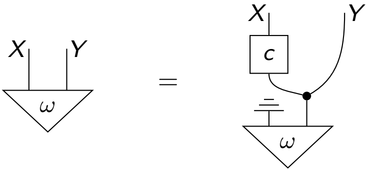
$$\omega|_p \models q \;=\; \frac{(\omega|_q \models p) \cdot (\omega \models q)}{\omega \models p}$$

Easy but important observation:

These rules are equivalent, using that & is commutative

Disintegration: extraction of channel

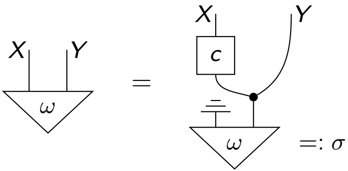
- Assume a joint state (distribution) ω on X, Y as depicted below
- A disintegration of ω is a channel $c\colon Y \to X$ such that:



- Equationally, $\omega(x, y) = \omega(x \mid y) \cdot \omega(y)$
- Disintegration is a fundamental concept, esp. in conditional probability theory
  - to construct a Bayesian network from a joint state
  - also to define conditional independence abstractly

Inference Theorem

Assume a joint state ω and a channel c such that:



We write σ for the first marginal ω % $[1, 0]$

Theorem

For predicates p on X and q on Y,

$$\left(\omega|_{p\otimes\mathbf{1}}\right) \% [0,1] = \sigma|_{c \ll p}$$

$$\left(\omega|_{\mathbf{1}\otimes q}\right) \% [1,0] = c \gg \left(\sigma|_{q}\right)$$

This explains the same outcomes in the ashtray-cancer example.

Where we are, so far

Quantum setting

- Let $\mathscr{H}, \mathscr{K}$ be (finite-dimensional) Hilbert spaces
- A state is a density matric $\varrho\colon \mathscr{H} \to \mathscr{H}$
  - this means: $\varrho \geq 0$ and $\mathrm{tr}\big(\varrho\big) = 1$
- A predicate is an effect $p\colon \mathscr{H} \to \mathscr{H}$
  - this means: $0 \leq p \leq 1$
  - sequential conjunction $p \mathbin{\&} q \coloneqq \sqrt{p}\, q \sqrt{p}$ is not commutative
- Validity is given by Born's rule: $\varrho \models p \;\coloneqq\; \mathrm{tr}\big(\varrho\, p\big) \;\in\; [0,1]$
- A channel $\mathscr{H} \to \mathscr{K}$ is a (completely) positive unit map $c\colon \mathcal{B}(\mathscr{K}) \to \mathcal{B}(\mathscr{H})$.
  - it comes with state transformation $c \gg \varrho$
  - and predicate transformation $c \ll p$

Two forms of quantum conditioning

$$\textbf{lower:} \quad \sigma|_p := \frac{\sqrt{p}\,\sigma\sqrt{p}}{\sigma \models p} \qquad \textbf{upper:} \quad \sigma|^p := \frac{\sqrt{\sigma}\,p\sqrt{\sigma}}{\sigma \models p}$$

- The 'lower' one comes from effectus theory, the 'upper' one from Leifer-Spekkens
- Classically they coincide

Theorem

Lower satisfies the product rule, upper satisfies Bayes' rule:

$$\sigma|_p \models q \;=\; \frac{\sigma \models p \,\&\, q}{\sigma \models p} \qquad \sigma|^p \models q \;=\; \frac{(\sigma|^q \models p)\cdot(\sigma \models q)}{\sigma \models p}$$

Intermezzo: assert channels

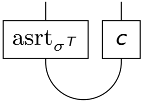
Given a predicate p on a Hilbert space $\mathscr{H}$ there is a (sub)channel $\mathscr{H} \to \mathscr{H}$ given by:

$$\mathrm{asrt}_p(A) := \sqrt{p}\,A\sqrt{p}$$

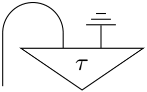
- These assert maps play an important role in effectus theory, for measurement
- Here we also apply them to states σ — which are special predicates

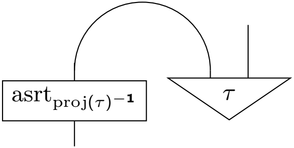
Channel extraction à la Leifer-Spekkens

- pairing of a state σ and a channel c to form a joint state:

$\mathrm{pair}(\sigma, c) \;:=$ 

- extraction of a channel from a joint state τ

$\mathrm{proj}(\tau) \;:=$ 

$\mathrm{extr}(t) \;:=$ 

They satisfy:

$$\mathrm{proj}(\mathrm{pair}(\sigma, c)) = \sigma \qquad\qquad \tau = \mathrm{pair}(\mathrm{proj}(\tau), \mathrm{extr}(\tau))$$
$$\mathrm{extr}\big(\mathrm{pair}(\sigma, c)\big) = c$$

(These constructions are basis-dependent via the cups & caps)

Main result about quantum inference

Let τ be a joint state on $\mathscr{H} \otimes \mathscr{K}$, and p, q predicates on $\mathscr{H}$ and $\mathscr{K}$.

Theorem

$$\big(\tau|_{p\otimes\mathbf{1}}\big) \,\%\, [0,1] = \mathrm{extr}(\tau) \gg \big(\mathrm{proj}(\tau)|^{p^T}\big)$$

$$\big(\tau|_{\mathbf{1}\otimes q}\big) \,\%\, [1,0] = \big(\mathrm{proj}(\tau)|^{\mathrm{extr}(\tau)\ll q}\big)^T.$$

Informally:

- on the left-hand-side of the equations there is crossover influence, via conditioning in one coordinate and marginalising in the other
  - this crossover happens through entanglement
- on the right-hand-side, this crossover is equivalently obtained via:
  - upper and lower conditioning
  - state- and predicate-transformation with the extracted channel $\mathrm{extr}(\tau)$

Where we are, so far

Concluding remarks

- Quantum Bayesian theory is still in its infancy
  - guiding examples are missing, so we're "air gitar playing"
  - analogies with classical Bayesian theory are a good starting point
- Essential points of the classical approach:
  - back-and-forth between joint and conditional probabilities (using disintegration)
  - analogues for conditioning, equivalently via crossover on joint states and transformations along edges of the network graph
- classical conditioning falls apart in two quantum forms:
  - the product rule and Bayes' rule are no longer equivalent
  - product holds for 'lower', Bayes for 'upper'
- Equivalence of conditioning via crossover and transformation still holds in the quantum case, but only by using the appropriate form of conditioning at the appropriate place
- Results are mathematically elegant, but intuition is lacking