Coalgebras and Kleisli Maps for Probability

Coalgebra Now — Oxford

Bart Jacobs — Radboud University
bart@cs.ru.nl
July 8, 2018

Outline

Introduction

Background on (categorical) probability

States as states

Recent work on destructive and constructive updating

Conclusions

Where we are, so far

Introduction

Background on (categorical) probability

States as states

Recent work on destructive and constructive updating

Conclusions

Coalgebras and states

Coalgebra is about state-based computation

- A coalgebra is a map of the form $X \longrightarrow F(X)$
- X is the state space
- the map captures transitions and/or observations

Starting point:
- the term "state" is rather widely used
- e.g. in "state transformation" — as companion of predicate transformation
- are such occurrences suggestions for connections with coalgebra?

Example: deterministic automaton

$$X \xrightarrow{\langle \delta, \varepsilon \rangle} X^A \times 2$$

- a state of such an automaton is an element $x \in X$
- a bit more abstractly, a map $1 \to X$

Example: non-deterministic automaton

$$X \xrightarrow{c} \mathcal{P}(1 + A \times X)$$

- what is now understood as a state?
- an element of X?
- or a subset of X — representing the reached states at a certain point in a computation
  - state transformation associated with c is a map $\mathcal{P}(X) \to \mathcal{P}(X)$
  - it sends $S \subseteq X$ to $\{y \in X \mid \exists x \in S. \exists a \in A. (a, y) \in c(x)\}$
  - implicitly, Kleisli extension is used
- Aside: a subset of X is a point $1 \to X$ in $\mathcal{K}\ell(\mathcal{P})$

Example: probabilistic automaton

$$X \xrightarrow{c} \mathcal{D}(A \times X)$$

- what is now a state? An element of X?
- or a probability distribution on X?
  - state transformation associated with c has type $\mathcal{D}(X) \to \mathcal{D}(X)$
  - it is obtained from Kleisli extension and marginalisation
- A distribution on X is a map $1 \to X$ in $\mathcal{K}\ell(\mathcal{D})$.

Example/excursion: quantum states

- A state on a Hilbert space $\mathscr{H}$ is a density operator
  - a bounded linear map $\rho \colon \mathscr{H} \to \mathscr{H}$
  - with $\rho \geq 0$ and $\mathrm{tr}(\rho) = 1$
- if $\mathscr{H}$ is e.g. finite-dimensional, such a state corresponds to a completely positive unital map $\mathcal{B}(\mathscr{H}) \to \mathbb{C}$
- equivalently, to a map $\mathcal{B}(\mathscr{H}) \to 0$ in the category **vNA** of von Neumann algebras
- equivalently, to a point $1 \to \mathcal{B}(\mathscr{H})$ in $\mathbf{vNA}^{\mathrm{op}}$.

(This perspective of "states as points" is further developed in effectus theory)

States and states

We seem to find two kinds of states:

(1) elements of state spaces — of coalgebras

(2) stages in computations — used in state transformations
   - points in a suitable category

Question: can we use the second kind of states also in the first form?

- we shall concentrate on the probabilistic case
- thus we seek coalgebras $\mathcal{D}(X) \to F(\mathcal{D}(X))$
- state transformation is a special case, for $F = \mathrm{id}$

Where we are, so far

Introduction

Background on (categorical) probability

States as states

Recent work on destructive and constructive updating

Conclusions

Main ingredients

(1) probability distributions/states $\omega \in \mathcal{D}(X)$ and predicates $p \in [0,1]^X$

(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) combinations of these

We shall elaborare these points in greater detail.

Discrete probability distributions / states

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

Predicates, as fuzzy functions

- 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 $x \in X$
  - sharp predicates are indicator functions $\mathbf{1}_E$ for an "event" $E \subseteq X$
- There are "truth", "falsum", "orthosupplement" predicates
  - e.g. $(p^{\perp})(x) = 1 - p(x)$, so that $p^{\perp\perp} = p$
  - then: $(\mathbf{1}_E)^{\perp} = \mathbf{1}_{\neg E}$
  - the set $[0,1]^X$ of predicates on X forms an effect module
- There is also fuzzy conjunction $p \,\&\, q$ via pointwise multiplication
  - $(p \,\&\, q)(x) = p(x) \cdot q(x)$
  - then $\mathbf{1}_E \,\&\, \mathbf{1}_D = \mathbf{1}_{E \cap D}$
  - this makes $[0,1]^X$ a commutative monoid in the category of effect modules

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: $\omega|_p = \sum_x \frac{\omega(x) \cdot p(x)}{\omega \models p} |x\rangle$

Validity and conditioning example

- Take $X = \{1,2,3,4,5,6\}$ with state dice $\in \mathcal{D}(X)$
  - recall dice $= \frac{1}{6}|1\rangle + \frac{1}{6}|2\rangle + \frac{1}{6}|3\rangle + \frac{1}{6}|4\rangle + \frac{1}{6}|5\rangle + \frac{1}{6}|6\rangle$
- Take even predicate $\mathbf{1}_E \in [0,1]^X$ for $E \subseteq X$; it's sharp, given by:
  - $E(1) = E(3) = E(5) = 0, \;\; E(2) = E(4) = E(6) = 1$
  - define odd via orthosupplement: $O = E^{\perp}$
- $\text{dice} \models \mathbf{1}_E \;\; = \;\; \frac{1}{2}$
- $\text{dice}|_{\mathbf{1}_E} \;\; = \;\; \frac{1/6}{1/2}|2\rangle + \frac{1/6}{1/2}|4\rangle + \frac{1/6}{1/2}|6\rangle = \frac{1}{3}|2\rangle + \frac{1}{3}|4\rangle + \frac{1}{3}|6\rangle$
- $\text{dice}|_{\mathbf{1}_E} \models \mathbf{1}_O \;\; = \;\; 0$

Two basic laws of conditioning

Recall that we write $p \,\&\, q$ for the pointwise product $(p \,\&\, q)(x) = p(x) \cdot q(x)$ of predicates $p, q \in [0,1]^X$.

product rule	$\omega|_p \models q \;=\; \frac{\omega \models p \,\&\, q}{\omega \models p}$
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
(the rules differ in a quantum setting)

State and predicate transformation

A channel $X \to Y$ is a function $X \to \mathcal{D}(Y)$

- thus, such a channel is an X-indexed family of states of Y
- alternatively, it is a stochastic matrix
- For a state $\omega \in \mathcal{D}(X)$ we get $c \gg \omega \in \mathcal{D}(Y)$ via:

$$(c \gg \omega)(y) \quad := \quad \sum_x c(x)(y) \cdot \omega(x).$$

- For a predicate $q \in [0,1]^Y$ we have $c \ll q \in [0,1]^X$ by:

$$(c \ll q)(x) \quad := \quad \sum_y c(x)(y) \cdot q(y).$$

Basic relation

$$\omega \models c \ll q \quad = \quad c \gg \omega \models q.$$

Calculus of channels

Channels can be composed sequentially, and in parallel:

- $(d \bullet c)(x) = d \gg c(x)$
- $(e \otimes f)(x, y) = e(x) \otimes f(y)$

- These $\bullet$ and $\otimes$ interact appropriately — abstractly because $\mathcal{K}\ell(\mathcal{D})$ is a symmetric monoidal category
- They also interact well with state and predicate transformation, eg:

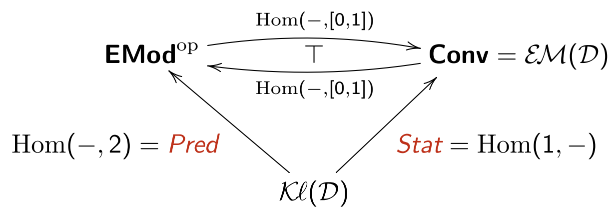
$$(d \bullet c) \gg \omega = d \gg (c \gg \omega) \quad \text{and } (d \bullet c) \ll q = c \ll (d \ll q)$$

Keeping states and predicates apart

- States and predicates look similar and are often confused
  - each state is a predicate: $\mathcal{D}(X) \subseteq [0,1]^X$
  - but not the other way around: predicates may have infinite support, and their probabilities need not add up to one.
- States and predicates have entirely different algebraic structures
  - states on a set X form a convex set
  - predicates on a set X form an effect module
- State transformation preserves convex sums, and predicate transformation preserves the effect module structure.
- Explicitly, for a channel $c \colon X \to \mathcal{D}(Y)$,
  - $c \gg (-) \colon \mathcal{D}(X) \to \mathcal{D}(Y)$ is a map in **Conv** $= \mathcal{EM}(\mathcal{D})$
  - $c \ll (-) \colon [0,1]^Y \to [0,1]^X$ is map in **EMod**

Summary as state-and-effect triangle

Predicates sit on the left, and states on the right in:



Conditioning and transformation

Overview table for joint work with Fabio Zanasi:

notation	action	terminology
$\omega\big\|_{(c \ll q)}$	first do predicate transformation, then update the state	evidential reasoning, or explanation, or backward inference
$c \gg (\omega\big\|_{p})$	first update the state, then do state transformation	causal reasoning, or prediction, or forward inference

Coalgebra example: taxicabs (Kahneman & Tverski 1972)

Consider the following description and question:

- A cab was involved in a hit and run accident at night. Two cab companies, Green and Blue, operate in the city. You are given the following data:
  - 85% of the cabs in the city are Green and 15% are Blue
  - A witness identified the cab as Blue. The court tested the reliability of the witness under the circumstances that existed on the night of the accident, and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.
- What is the probability that the cab involved in the accident was Blue rather than Green?

The answer is 41%, via Bayes. Many people give a higher probability because they do not take the prior cab distribution into account.

Where is the coalgebra in the taxicab example?

- Take as set of taxicab colours $C = \{g, b\}$
  - there is a base rate / a prori distribution: $\tau = 0.85|\,g\,\rangle + 0.15|\,b\,\rangle$
- There is correctness coalgebra / channel $c\colon C \to \mathcal{D}(C)$, namely
  $$c(g) = 0.8|\,g\,\rangle + 0.2|\,b\,\rangle \qquad c(b) = 0.2|\,g\,\rangle + 0.8|\,b\,\rangle$$
- There is the "blue taxi" witness evidence, given as singleton predicate $\mathbf{1}_{\{b\}}$ on C
- Now we can do backward inference:
  $$\tau\big|_{c \ll \mathbf{1}_{\{b\}}} = 0.4138|\,b\,\rangle + 0.5862|\,g\,\rangle$$
  ↑
  such point updates are important, as "daggers"
- The coalgebra-as-test view goes back to Pearl
  - not in terms of $X \to \mathcal{D}(X)$, but as correctness tables

Where we are, so far

Introduction

Background on (categorical) probability

States as states

Recent work on destructive and constructive updating

Conclusions

Conditioning as coalgebra I, via partiality

Let $p \in [0,1]^X$ be a fixed predicate on a set X.

- Consider the conditioning operation $\omega \longmapsto \omega|_p$
- It is partial operation — undefined if $\omega \models p = 0$.
- Hence we can see it as coalgebra:

$$\mathcal{D}(X) \longrightarrow 1 + \mathcal{D}(X)$$

$$\omega \longmapsto \begin{cases} * & \text{if } \omega \models p = 0 \\ \omega|_p & \text{otherwise} \end{cases}$$

- The functor F involved is $F(Y) = 1 + Y$, on **Sets**

Conditioning as coalgebra II, via "hypernormalisation"

- Idea: put both $\omega|_p$ and $\omega|_{p^\perp}$ in the same coalgebra
- moreover, include validities of p and $p^\perp$ as probabilities
  - the "validity=zero" problem can be made to disappear
- We now have:

$$\mathcal{D}(X) \longrightarrow \mathcal{D}\Big(\mathcal{D}(X) + \mathcal{D}(X)\Big)$$

$$\omega \longmapsto (\omega \models p)\big|\, \omega|_p \,\big\rangle + (\omega \models p^\perp)\big|\, \omega|_{p^\perp} \,\big\rangle$$

- E.g. for $X = \{1,2,3,4,5,6\}$ and even predicate $p = \mathbf{1}_E$,

$$\text{dice} \longmapsto \frac{1}{2}\big|\, \tfrac{1}{3}|2\rangle + \tfrac{1}{3}|4\rangle + \tfrac{1}{3}|6\rangle \,\big\rangle \;+\; \frac{1}{2}\big|\, \tfrac{1}{3}|1\rangle + \tfrac{1}{3}|3\rangle + \tfrac{1}{3}|5\rangle \,\big\rangle$$

- The functor is $F(Y) = \mathcal{D}(Y + Y)$ on **Sets**.
  - alternatively, $F(Y) = Y + Y$ on $\mathcal{K}\ell(\mathcal{D})$

A "draw" coalgebra

- Suppose we have vase/urn whose elements are given by a state $\omega \in \mathcal{D}(X)$
- Taking one element out is a state-changing operation; what is it coalgebraically?

We propose $F(Y) = \mathcal{D}(X \times (1 + Y))$ with coalgebra:

$$\mathcal{D}(X) \xrightarrow{\quad d \quad} \mathcal{D}\Big(X \times (1 + \mathcal{D}(X))\Big)$$

$$\omega \longmapsto \begin{cases} 1\big|x, *\big\rangle & \text{if } \omega(x) = 1 \text{ for a unique } x \\ \sum_x \omega(x)\big|\, x, \omega|_{\mathbf{1}_{\{x\}}^\perp} \big\rangle & \text{otherwise} \end{cases}$$

Then: $\mathcal{D}(\pi_1) \circ d = \text{id}$

Draw example

Question

Suppose we have a vase with one red, two black and one green marbles. You draw one marble, and somehow know that there is still a green marble in the vase. What is the probability that you have drawn the red one?

Let's analyse:

- $X = \{r, b_1, b_2, g\}$
- $\omega = \frac{1}{4}|r\rangle + \frac{1}{4}|b_1\rangle + \frac{1}{4}|b_2\rangle + \frac{1}{4}|g\rangle$
- there are obvious'colour' events $R, B, G \subseteq X$

Draw example, solution

$$\begin{aligned} d(\omega) &= \frac{1}{4}\big| r, \omega|_{\mathbf{1}^{\perp}_{\{r\}}} \big\rangle + \frac{1}{4}\big| b_1, \omega|_{\mathbf{1}^{\perp}_{\{b_1\}}} \big\rangle + \frac{1}{4}\big| b_2, \omega|_{\mathbf{1}^{\perp}_{\{b_2\}}} \big\rangle + \frac{1}{4}\big| g, \omega|_{\mathbf{1}^{\perp}_{\{g\}}} \big\rangle \\ &= \frac{1}{4}\big| r, \tfrac{1}{3}| b_1 \rangle + \tfrac{1}{3}| b_2 \rangle + \tfrac{1}{3}| g \rangle \big\rangle + \frac{1}{4}\big| b_1, \tfrac{1}{3}| r \rangle + \tfrac{1}{3}| b_2 \rangle + \tfrac{1}{3}| g \rangle \big\rangle \\ &\quad + \frac{1}{4}\big| b_2, \tfrac{1}{3}| b_1 \rangle + \tfrac{1}{3}| b_1 \rangle + \tfrac{1}{3}| g \rangle \big\rangle + \frac{1}{4}\big| g, \tfrac{1}{3}| r \rangle + \tfrac{1}{3}| b_1 \rangle + \tfrac{1}{3}| b_2 \rangle \big\rangle \end{aligned}$$

We have the predicate "green in the vase" $q(x, \sigma) = \sigma \models \mathbf{1}_G$.

$$d(\omega) \models q = \tfrac{1}{4} \cdot \tfrac{1}{3} + \tfrac{1}{4} \cdot \tfrac{1}{3} + \tfrac{1}{4} \cdot \tfrac{1}{3} + \tfrac{1}{4} \cdot 0 = \tfrac{1}{4}$$

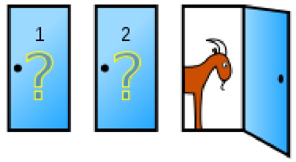
$$\begin{aligned} d(\omega)\big|_q &= \frac{1}{3}\big| r, \tfrac{1}{3}| b_1 \rangle + \tfrac{1}{3}| b_2 \rangle + \tfrac{1}{3}| g \rangle \big\rangle + \frac{1}{3}\big| b_1, \tfrac{1}{3}| r \rangle + \tfrac{1}{3}| b_2 \rangle + \tfrac{1}{3}| g \rangle \big\rangle \\ &\quad + \frac{1}{3}\big| b_2, \tfrac{1}{3}| b_1 \rangle + \tfrac{1}{3}| b_1 \rangle + \tfrac{1}{3}| g \rangle \big\rangle \end{aligned}$$

$$\mathcal{D}(\pi_1)\Big(d(\omega)\big|_q\Big) = \frac{1}{3}| r \rangle + \frac{1}{3}| b_1 \rangle + \frac{1}{3}| b_2 \rangle$$

Monty Hall problem

Problem statement — due to Steve Selvin, Sci. Am. 1975

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?



Monty Hall, solution, part I

There are two state-changing operations: "drawing" and "opening a non-car" door.

We take $X = \{1, 2, 3\}$ and write in $\mathcal{K}\ell(\mathcal{D})$,

$$\mathcal{D}(X) \xrightarrow{\;d \,=\, \text{draw}\;} X \times (1 + \mathcal{D}(X)) \xrightarrow{\;\text{id} \times (\text{id} + \text{condition})\;} X \times (1 + 1 + \mathcal{D}(X))$$

We condition with predicate $\mathbf{1}_G$ for the goat subset $G \subseteq X$.

Monty Hall, solution, part II

We start from the uniform distribution $\omega = \frac{1}{3}|1\rangle + \frac{1}{3}|3\rangle + \frac{1}{3}|3\rangle$ and assume $G = \{1, 2\}$, so the car is behind door 3.

$$\begin{aligned} \omega &\overset{\text{draw}}{\longmapsto} \frac{1}{3}\big|1, \tfrac{1}{2}|2\rangle + \tfrac{1}{2}|3\rangle\big\rangle + \frac{1}{3}\big|2, \tfrac{1}{2}|1\rangle + \tfrac{1}{2}|3\rangle\big\rangle + \frac{1}{3}\big|3, \tfrac{1}{2}|1\rangle + \tfrac{1}{2}|2\rangle\big\rangle \\ &\overset{\text{cond}}{\longmapsto} \frac{1}{3}\big|1, 1|2\rangle\big\rangle + \frac{1}{3}\big|2, 1|1\rangle\big\rangle + \frac{1}{3}\big|3, \tfrac{1}{2}|1\rangle + \tfrac{1}{2}|2\rangle\big\rangle \end{aligned}$$

Conclusion

- If you switch, in the first two cases you will win the car; in the third case you loose it.
- This happens in 2 out of 3 cases. Hence switching is better.

Where we are, so far

Introduction

Background on (categorical) probability

States as states

Recent work on destructive and constructive updating

Conclusions

Problem description

- Typical Bayesian inference (reasoning) proceeds as follows:
  - I have "evidence" $E_1, \ldots, E_n$, used to condition my state
  - I then "observe" A, via marginalisation of conditioned state
- The evidence (and observation) are usually "point" or "singleton" predicates
- What if the evidence is "soft"
  - I saw the object in the dark and believe with 30% certainty that it is red and 70% certainty that it is blue
  - How to handle is called soft evidential update problem (Darwiche)
- There are two approaches, giving different outcomes
  - following Jeffrey, renamed as destructive
  - following Pearl, renamed as constructive

Virus – blood pressure example

We consider patients having a virus or not, and their blood pressure:

virus?	Low	Medium	High
yes (v)	20%	20%	60%
no ($\sim v$)	60%	30%	10%

We know, as base rate, that 1 in 15 patients have the virus.

Mathematical formalisation:

- underlying domains $V = \{v, \sim v\}$ and $B = \{L, M, H\}$
- prior / base rate distribution $\omega = \frac{1}{15}|\, v\,\rangle + \frac{14}{15}|\sim v\,\rangle$
- channel / Kleisli map $c\colon V \to \mathcal{D}(B)$ extracted from table:

$$c(v) = \tfrac{2}{10}|\,L\,\rangle + \tfrac{2}{10}|\,M\,\rangle + \tfrac{6}{10}|\,H\,\rangle \qquad c(\sim v) = \tfrac{6}{10}|\,L\,\rangle + \tfrac{3}{10}|\,M\,\rangle + \tfrac{1}{10}|\,H\,\rangle$$

Point evidence example

- Suppose we have **high** blood pressure evidence
  - what is the updated virus probability (distribution)?
  - typical Bayes' rule problem
- Channel-based solution, with point predicate $\mathbf{1}_{\{H\}}$ on $B = \{L, M, H\}$

$$\omega|_{c \ll \mathbf{1}_{\{H\}}} = 0.3|\,v\,\rangle + 0.7|\sim v\,\rangle$$

  This 30% probability is higher than the base rate $\frac{1}{15} \sim 6.67\%$
- More abstractly, this involves the dagger channel in opposite direction:

$$B \xrightarrow{\;c^{\dagger}_{\omega}\;} \mathcal{D}(V)$$
$$y \longmapsto \omega|_{c \ll \mathbf{1}_{\{y\}}}$$

Soft evidence example

Suppose we have 25% certainty of low blood pressure, 25% of medium 50% of high. What is the updated virus probability?

- **Destructive answer**, after Jeffrey
  - Idea: convex combination of point observations
  - 0.25 · update with L + 0.25 · update with M + 0.5 · update with H

    $= c^{\dagger}_{\omega} \gg \big(0.25|L\rangle + 0.25|M\rangle + 0.5|H\rangle\big)$

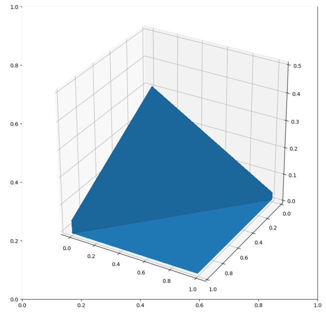
    $= 0.0941|v\rangle + 0.9059|{\sim}v\rangle$
- **Constructive answer**, after Pearl
  - Idea: reason backward with evidence as fuzzy predicate
  - define $p \in [0,1]^B$ as $p(L) = p(M) = 0.25$, $p(H) = 0.5$
  - $\omega|_{c \ll p} = 0.1672|v\rangle + 0.8328|{\sim}v\rangle$

Substantial difference: 9% versus 17%

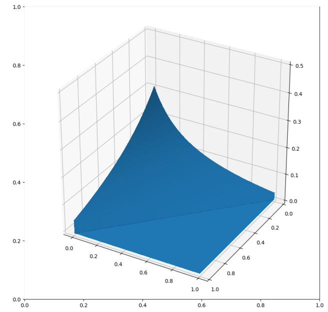
What should decision support systems do — e.g. in medicine?

Plots

We describe the virus probability, given soft evidence $x|L\rangle + y|M\rangle + (1-x-y)|H\rangle$, for $0 \leq x + y \leq 1$ in:



destructive update



constructive update

General observations

Destructive & constructive update coincide on point evidence.

- **Destructive update**
  - interprets soft evidence as state / probability distribution
  - the prior is (largely) overridden by the evidence
  - successive updates do not commute
  - starting from what you can predict you learn nothing: $c^{\dagger}_{\omega} \gg (c \gg \omega) = \omega$
- **Constructive update**
  - interprets soft evidence as fuzzy predicate
  - prior is smoothly combined with the evidence — as inner product
  - successive updates do commute
  - starting from nothing (constant/uniform predicate) you learn nothing: $\omega|_{c \ll (r \cdot \mathbf{1})} = \omega$

It is unclear to me which approach is "the right one" — or even what criterion to use!

Where we are, so far

Introduction

Background on (categorical) probability

States as states

Recent work on destructive and constructive updating

Conclusions

Final remarks

Coalgebras are important in probabilistic reasoning via:

(1) State-transformations, like conditioning and drawing
- where states as stages in computations are used as coalgebraic states

(2) Channels
- actually, Kleisli maps are more useful — with coalgebras as special endomap case