Diagrammatic Reasoning about Probability and Nondeterminism

1 Effectful programming

Programs with nondeterministic choice:

$$(\mathit{tea} \Box \mathit{coffee}) \Box \mathit{beer}$$

with axioms:

$$\begin{aligned} m \Box m &= m \\ m \Box n &= n \Box m \\ m \Box (n \Box p) &= (m \Box n) \Box p \end{aligned}$$

Also, composition $\gg\!=$ distributes leftwards over choice:

$$(m \Box n) \gg\!= k = (m \gg\!= k) \Box (n \gg\!= k)$$

A semantic model in terms of finite sets.

2 Probabilistic programming

Similarly, programs with probabilistic choice:

$$\mathit{six} = \mathit{True} \lhd {}^1/_6 \rhd \mathit{False}$$

with axioms (where $w \in [0,1]$ and $\bar{w} = 1 - w$):

$$\begin{aligned} m \lhd 0 \rhd n &= n \\ m \lhd w \rhd m &= m \\ m \lhd w \rhd n &= n \lhd \bar{w} \rhd m \\ m \lhd w \rhd (n \lhd x \rhd p) &= (m \lhd y \rhd n) \lhd z \rhd p \\ &\Leftarrow w = y{\times}z \wedge \bar{z} = \bar{w}{\times}\bar{x} \\ (m \lhd w \rhd n) \gg\!= k &= (m \gg\!= k) \lhd w \rhd (n \gg\!= k) \end{aligned}$$

A semantic model as finite probability distributions.

3 Combining effects

Union of signatures and equations, plus interaction:

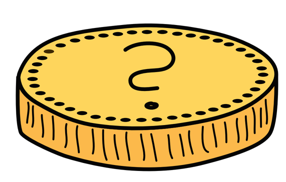
$$m \lhd w \rhd (n \Box p) = (m \lhd w \rhd n) \Box (m \lhd w \rhd p)$$

Say you win when your coin matches die roll:

$$\begin{aligned} \mathit{sixcoin} &= \mathit{six} \gg\!= \lambda c.\,(\mathit{coin} \gg\!= \lambda a.\,(a = c)) \\ \mathit{coinsix} &= \mathit{coin} \gg\!= \lambda a.\,(\mathit{six} \gg\!= \lambda c.\,(a = c)) \end{aligned}$$

where you choose $\mathit{coin} = \mathit{True} \Box \mathit{False}$.

Order of play makes a difference!



A model as finite convex-closed sets of distributions.

4 Distributivity is tricky

$\lhd\,\rhd$ distributes over $\Box$, but not vice-versa:

$$m \Box (n \lhd w \rhd p) \neq (m \Box n) \lhd w \rhd (m \Box p)$$

Asserting both distributivities collapses the theory.

$\gg\!=$ distributes leftwards over $\lhd\,\rhd$, but rightwards

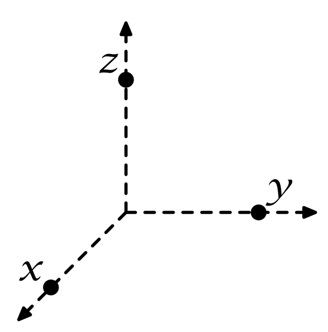
$$\begin{aligned} & m \gg\!= (\lambda a.\,(k\ a) \lhd w \rhd (k'\ a)) \\ &\quad = (m \gg\!= k) \lhd w \rhd (m \gg\!= k') \end{aligned}$$

entails the unwanted distributivity above.
(This law holds only for deterministic m.)

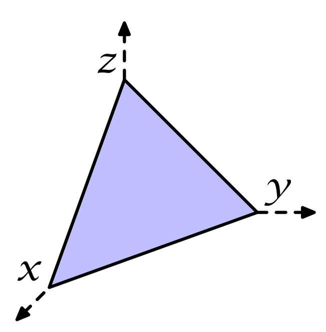
How can we get some intuition for these properties?

5 A geometric model

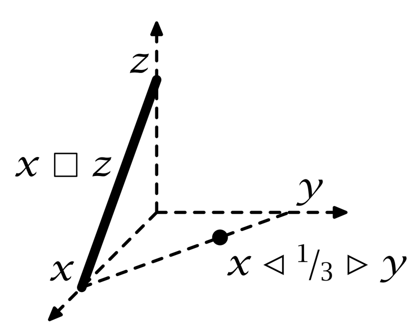
...as finite convex sets of points (polygons) in hyperspace.



Consider programs over three outcomes x, y, z,



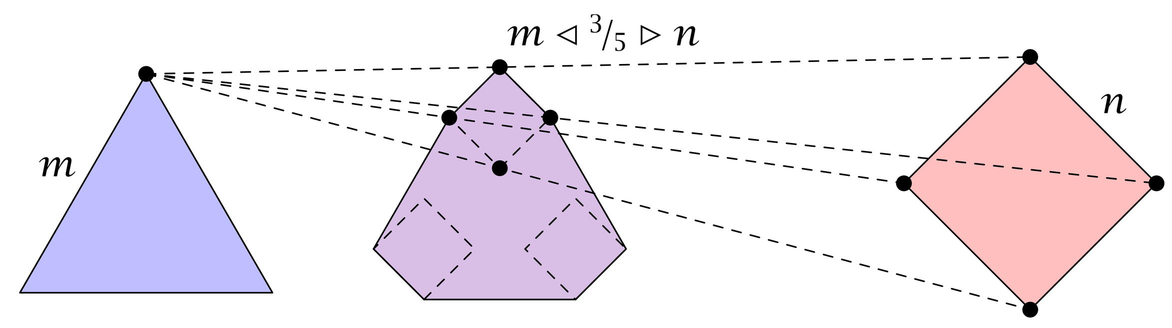
modelled as convex polygons in triangle xyz,



with $\lhd\,\rhd$ as weighted sum, $\Box$ as convex union.

6 Morphing

Generally, $\lhd\,\rhd$ as pointwise weighted sum of polygons:

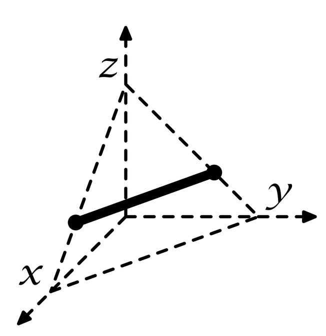


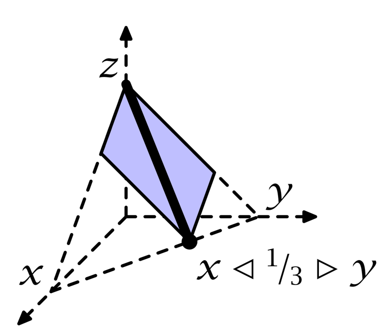
It suffices to take closure of projections of n from each vertex of m—or vice versa.

7 Distributivities, geometrically

The desirable distributivity holds:

$$(x \Box y) \lhd {}^2/_3 \rhd z = (x \lhd {}^2/_3 \rhd z) \Box (y \lhd {}^2/_3 \rhd z)$$



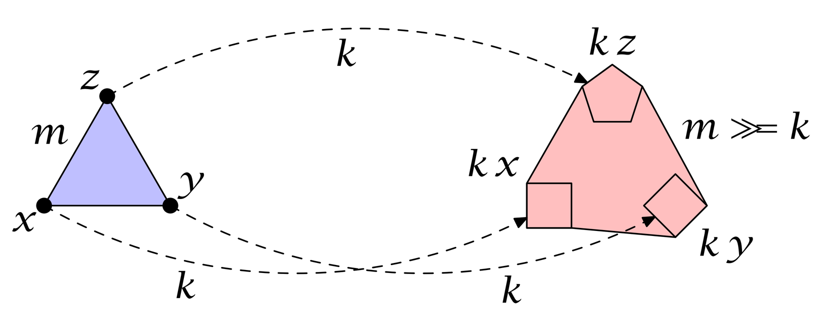


...and the undesirable one does not:

$$(x \lhd {}^1/_3 \rhd y) \Box z \neq (x \Box z) \lhd {}^1/_3 \rhd (y \Box z)$$

8 Barycentric algebra

Composition is convex union of pointwise images:



A barycentric algebra $(A, \lhd\,\rhd)$ satisfies the four axioms of Panel 2. Homomorphisms are affine functions:

$$h\ (m \lhd w \rhd n) = (h\ m) \lhd w \rhd (h\ n)$$

Convex polygons form a barycentric algebra.

Workshop on Probabilistic Programming Semantics, January 2016
Faris Abou-Saleh, Kwok-Ho Cheung, Jeremy Gibbons
Department of Computer Science, University of Oxford
http://www.cs.ox.ac.uk/activities/algprog/

UNIVERSITY OF OXFORD