A(nother) Random Variable Monad

Preliminary Report

Michael W. Mislove

Tulane University
New Orleans, LA

Domains XII
Cork, August 26, 2015

Work sponsored by US AFOSR & NSF

A Disclaimer

- This is *NOT* a talk about stochastic processes or Skorohod's Theorem.

  Apologies – a funny thing happened on the way to the talk...

- Instead, I'll talk about some preliminary results on computational models for probabilistic choice.

  Active theme over last several years
  Other talks at the meeting: Plotkin, Barker, Jung

Some History

- The *probabilistic powerdomain* is the traditional domain model for probabilistic choice.
  - Realized on DCPO as $\mathbb{V}(D)$ – the space of Scott-continuous valuations of $\mathcal{O}(D)$
  - Defines a monad on DCPO
  - Leaves Coh invariant, but *no CCC of domains is known to be invariant*
  - No distributive law wrt power domains
- Varacca (2003) devised a monad that overcomes these problems
  - Weakens the probabilistic law: $p +_r p \neq p$.
- Based on Varacca's work, M. (2007) devised a monad of *finite random variables* that leaves RB and FS invariant, and satisfies a distributive law.

Some History

- The *probabilistic powerdomain* is the traditional domain model for probabilistic choice.
- Varacca (2003) devised a monad that overcomes these problems
  - Weakens the probabilistic law: $p +_r p \neq p$.
- Based on Varacca's work, M. (2007) devised a monad of *finite random variables* that leaves RB and FS invariant, and satisfies a distributive law.
- Goubault-Larrecq and Varacca (2011) proposed a monad of *continuous random variables*, but the Kleisli lift failed to be Scott continuous.
- Their work inspired this work – and that by Tyler Barker you'll hear after this talk.

Some Background

- *Random variables versus* Prob
  - Order on $\mathit{Prob}(D)$ is more complicated than that on D:

    $\mu \sqsubseteq \nu$ iff $\mu(U) \leq \nu(U)$ ($\forall U$ open)

    $d \mapsto \delta_d : D \to \mathit{Prob}(D)$ unit of monad
  - *Idea of Random Variables:*

    Choose fixed domain $\mathcal{C}$ for which $\mathit{Prob}(\mathcal{C})$ is well-behaved,

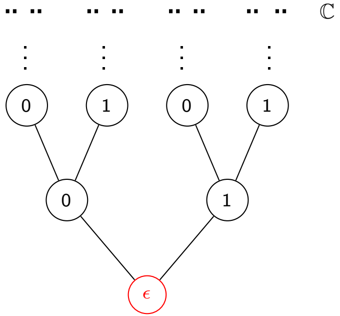
    and use

    $f : \mathcal{C} \to D$ to define outcomes of probabilistic choices in $\mathcal{C}$

Some Background

- *Random variables versus* Prob
  - *Which domain $\mathcal{C}$ to choose?*

    Use the Cantor tree – the full binary tree together with the Cantor set at the top:



  - So, the model looks like $Prob(\mathcal{C}) \times [\mathcal{C} \rightarrow D]$ for a domain D

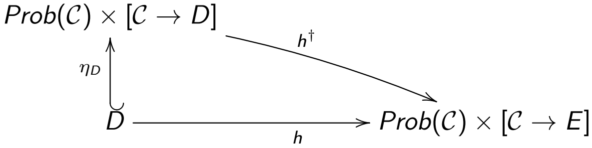
Sanity Check

$RanV(D) \equiv Prob(\mathcal{C}) \times [\mathcal{C} \to D]$ defines a functor:

$$g\colon D \to E \quad \Longrightarrow \quad RanV(g)(\mu, f) = \langle \mu, g \circ f \rangle.$$

Sanity Check

But to be a monad, we need a Keisli lift:



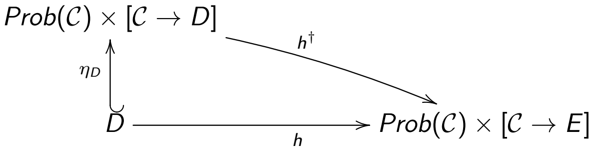
where $\eta_D(d) = \langle \delta_\epsilon, const_d \rangle$.

If $h^\dagger(\mu, f) = \langle \nu, g \rangle$,

g is easy: $f\colon \mathcal{C} \to D$ & $\pi_2 \circ h\colon D \to E \;\Rightarrow\; g = \pi_2 \circ h \circ f\colon \mathcal{C} \to E$.

Sanity Check

But to be a monad, we need a Keisli lift:



where $\eta_D(d) = \langle \delta_\epsilon, const_d \rangle$.

If $h^\dagger(\mu, f) = \langle \nu, g \rangle$,

For ν, we should combine μ and the family $\{\pi_1 \circ h \circ f(d)\}_{d \in \mathsf{supp}\, \mu}$.

Suppose $\mu = \sum_{d \in F} r_d \delta_d$ is simple. Then we expect:

$$h^\dagger \left(\sum_{d \in F} r_d \delta_d, f \right) = \left\langle \sum_{d \in F} r_d (\delta_d \otimes \pi_1 \circ h \circ f(d)), \pi_2 \circ h \circ f \right\rangle$$

where $\otimes$ is a composition operation on measures.

$\otimes$ also should also be monotone!

More Background

- Prob forms a monad on:
  - Comp – category of compact T_2 spaces and continuous maps

    Algebras are compact affine spaces
  - CompMon – category of compact monoids and continuous monoid morphisms

    Algebras are compact affine monoids
  - CompGrp – category of compact groups and continuous group maps

    Algebras again compact affine monoids – right adjoint changes to $M \mapsto H(1_M)$
- On CompMon, the monoid operation on $Prob(S, \cdot, e)$ is *convolution:*

  $*\colon Prob(S) \times Prob(S) \to Prob(S)$ by

$$\mu * \nu(A) = \cdot\,(\mu \times \nu)(A) = \mu \times \nu(\cdot^{-1}(A)) = \mu \times \nu(\{\langle x, y\rangle \mid x \cdot y \in A\}).$$

More Background

- An *ordered monoid* is a monoid $(S, \cdot, e)$ which has a partial order relative to which $\cdot \colon S \times S \to S$ is monotone.
  - $\mathcal{C}$ is a monoid under concatenation, but it is *not* an ordered monoid.
  - $\mathcal{C}_{\surd} \equiv \{0,1\}^\infty \cup \{0,1\}^* \surd$ with the prefix order, and let

    $$s \cdot t = \begin{cases} s' t & \text{if } s = s' \surd \\ s & \text{otherwise} \end{cases}$$

  - $(\mathcal{C}_{\surd}, \cdot, \surd)$ is a compact ordered monoid.
- Prob is a monad on CompOrdMon – category of compact, ordered monoid and continuous, monotone monoid morphisms.

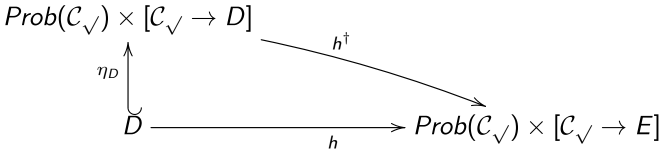
  The convolution operation on $Prob(S, \cdot, e, \leq)$ is monotone:

  $$\mu \sqsubseteq \mu', \nu \sqsubseteq \nu' \quad \Longrightarrow \quad \mu * \nu \sqsubseteq \mu' * \nu'.$$

  - This applies to *any* bounded complete domain in the Lawson topology – in paricular to $\mathcal{C}_{\surd}$.

Towards a Monad

The monad should be: $RanV(D) = Prob(\mathcal{C}_{\surd}) \times [\mathcal{C}_{\surd} \to D]$ where

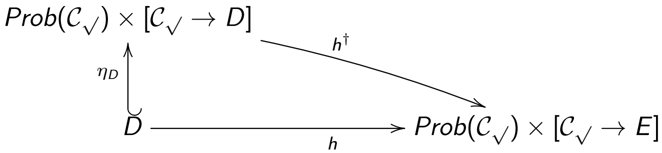


where:

$$\eta_D = \langle \delta_{\surd}, const_d \rangle \quad \text{and for } \mu = \sum_{d \in F} r_d \delta_d$$

$$h^{\dagger}(\mu, f) = \left\langle \sum_{d \in F} r_d(\delta_d * (\pi_1 \circ h \circ f(d)), \pi_2 \circ h \circ f \right\rangle$$

Towards a Monad



where:

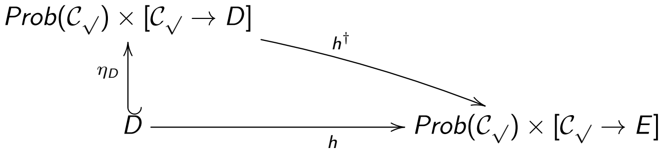
$$\eta_D = \langle \delta_{\surd}, const_d \rangle \text{ and for } \mu = \sum_{d \in F} r_d \delta_d$$

$$h^{\dagger}(\mu, f) = \left\langle \sum_{d \in F} r_d (\delta_d * (\pi_1 \circ h \circ f(d)), \pi_2 \circ h \circ f \right\rangle$$

ToDo's:

- Validate monad structure – May require restricting μ to be "thin"
- Work out equational laws

Towards a Monad



where:

$$\eta_D = \langle \delta_{\surd}, const_d \rangle \text{ and for } \mu = \sum_{d \in F} r_d \delta_d$$

$$h^{\dagger}(\mu, f) = \left\langle \sum_{d \in F} r_d(\delta_d * (\pi_1 \circ h \circ f(d)), \pi_2 \circ h \circ f \right\rangle$$

ToDo's:

- Devise closed form for probability component with μ non-discrete

$$\mu = \lim_n \sum_{x \in F_n} r_x \delta_x \implies \mu \text{``} * \text{''} \nu = \lim_n \sum_{x \in F_n} r_x \left(\delta_x * (\pi_1 \circ h \circ f(x)\right)$$