Stochastic Domain Theory

Michael Mislove

Tulane University

Simons Institute on the Theory of Computing
Reunion Workshop on Logical Structures for Computation

December 12, 2017

Supported by US AFOSR

Stochastic Processes and Domains

Scott's Stochastic Lambda Calculus

- Untyped lambda calculus with probabilistic choice
- Probabilistic choice implemented via *random variables* $X\colon ([0,1],\lambda) \to \mathcal{P}(\mathbb{N})$ using Borel sets generated by *Scott topology*.

Barker's Randomized PCF

- PCF (simply typed lambda calculus $+\ \mathbb{N}\ +\ \mathbb{B}\ +$ `rec`) with *randomized choice*
- Models randomized algorithms – reveals speedup in Miller-Rabin Prime Testing Algorithm

Stochastic Lambda Calculus for Probabilistic Programming

This talk: *Applying domain theory to stochastic processes.*

Stochastic Processes and Skorohod's Theorem

A *stochastic process* is a time-indexed family $\{X_t \mid t \in T \subseteq \mathbb{R}_+\}$ of random variables $X_t : \Omega \to S$, where $(\Omega, \Sigma_\Omega, \mu)$ is a probability space, and S is a Polish space.

Note: If S is Polish, then so is $(Prob(S), d_p)$, where d_p is the *Prokhorov metric*. d_P generates the *weak topology* on $Prob(S)$.

Examples:

- Brownian motion, Lévy processes, Markov chains

  MCMC – Markov chain Monte Carlo
  Theme in *Probabilistic Programming Semantics*

Stochastic Processes and Skorohod's Theorem

Let λ denote Lebesgue measure on $[0,1]$.

Skorohod's Theorem

If S is a Polish space, and $\nu \in \mathrm{Prob}\, S$, then there is a random variable $X\colon [0,1] \to S$ with $X_* \lambda = \nu$; i.e., $\nu(A) = \lambda(X^{-1}(A))$ $\forall A$ measurable.

Moreover, if $\nu_n, \nu \in \mathrm{Prob}\, S$ satisfy $\nu_n \to_w \nu$, then there are random variables $X_n, X\colon [0,1] \to S$ with $X_* \lambda = \nu, X_{n*} \lambda = \nu_n$ and $X_n \to X$ λ-a.e.

So: 1) *Every* stochastic process arises from some $X_t\colon [0,1] \to S$.

2) Convergence in $(\mathrm{Prob}\, S, \mathit{weak})$ is equivalent to pointwise convergence a.e. of measurable maps $X\colon [0,1] \to S$.

Goal: Obtain domain-theoretic version of Skorohod's Theorem with Skorohod's Theorem as a Corollary.

Interlude: Some Domain Theory

Domains are partially ordered sets with additional properties:

Directed completeness

$\emptyset \neq A \subseteq D$ *directed* if $x, y \in A \ \Rightarrow \ (\exists z \in A)\, x, y \leq z$.
D *directed complete:* A directed $\ \Rightarrow\ \sup A$ exists.

Approximation

$x \ll y$ iff $y \leq \sup A$ directed $\ \Rightarrow\ (\exists a \in A)\, x \leq a$.

Domain: $\twoheaddownarrow y = \{x \mid x \ll y\}$ directed and $y = \sup \twoheaddownarrow y$

Basis: $B_D \subseteq D$ satisfying $\downarrow y \cap B_D \subseteq \twoheaddownarrow y$ & $y = \sup \downarrow y \cap B_D$ $(\forall y \in D)$

Scott Topology

U *Scott open* if:

- $U = \uparrow U = \{x \in D \mid (\exists u \in U)\, u \leq x\}$ and
- A directed, $\sup A \in U \ \Rightarrow\ A \cap U \neq \emptyset$.

Example: $\twoheaduparrow x = \{y \mid x \ll y\}$ is Scott open $(\forall x \in D)$.

Interlude: Some Domain Theory

Morphisms

$f : D \to E$ is *Scott continuous* if:

- f is monotone, and
- A directed $\Rightarrow$ $f(\sup A) = \sup f(A)$.

Lawson Topology

Basis: $\{\twoheaduparrow x \setminus \uparrow F \mid x \in D,\ F \in \mathcal{P}_{<\omega} D\}$

Hausdorff refinement of Scott topology.

All the domains we discuss are Lawson compact.

Towards a Domain-theoretic Skorohod Theorem

First step: We can use any standard probability space for $([0,1], \lambda)$:

Let $\mathcal{C} = 2^\omega$ denote a countable product of 2-point groups, and let $\mu_\mathcal{C}$ denote Haar measure on $\mathcal{C}$.

Theorem:
If S is a Polish space, and $\nu \in \mathrm{Prob}\, S$, then there is a random variable $X \colon \mathcal{C} \to S$ with $X_* \mu_\mathcal{C} = \nu$.

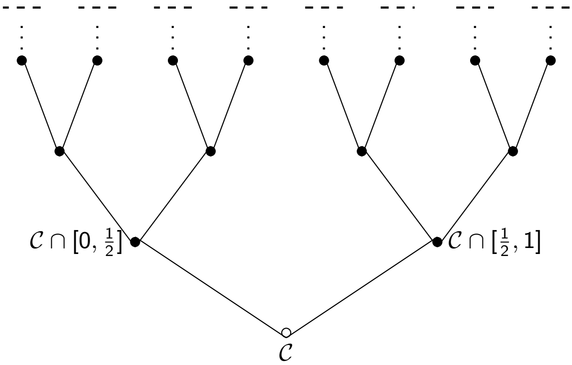
Moreover, if $\nu_n, \nu \in \mathrm{Prob}\, S$ satisfy $\nu_n \to_w \nu$, then there are random variables $X_n, X \colon \mathcal{C} \to S$ with $X_* \mu_\mathcal{C} = \nu, X_{n*} \mu_\mathcal{C} = \nu_n$ and $X_n \to X$ $\mu_\mathcal{C}$-a.e.

Proof: Use $\varphi \colon \mathcal{C} \twoheadrightarrow [0,1]$. $\square$

Towards a Domain-theoretic Skorohod Theorem

Second Step: Embed $\mathcal{C}$ in an appropriate domain:

$\mathbb{CT} = \{0,1\}^\infty$ is a domain in the prefix order.



$$\mathcal{C} \simeq (\{0,1\}^\omega, \Sigma(\mathbb{CT})|_{\{0,1\}^\omega}) = (\mathsf{Max}\,\mathbb{CT}, \Lambda(\mathbb{CT})|_{\mathsf{Max}\,\mathbb{CT}})$$

Third Step: Which domains represent Polish spaces?

BCD_ω – countably based bounded complete domains and Scott continuous maps.

- $D^\infty \simeq [D^\infty \to D^\infty]$ is in BCD_ω.
- $\mathbb{CT} = \{0,1\}^\infty$ is a bounded complete domain.

Theorem: (Lawson; Ciesielski, Flagg & Kopperman)

Each countably-based bounded complete domain D satisfies $\mathrm{Max}\, D$ is a Polish space in the inherited Scott topology. Moreover, $\mathrm{Max}\, D$ is a G_δ in D.

Conversely, every Polish space can be embedded as $\mathrm{Max}\, D$ for a countably based bounded complete domain D.

Examples:

1) $\mathcal{C} \simeq \mathrm{Max}\, \mathbb{CT} \hookrightarrow \mathbb{CT}$.

2) $\mathbb{R} \simeq \mathrm{Max}\, \mathbb{IR} \hookrightarrow \mathbb{IR} = (\{[a,b] \mid a \leq b \in \mathbb{R}\} \cup \{\mathbb{R}\}, \supseteq)$.

Domain-theoretic Skorohod Theorem (cont'd)

Skorohod's Theorem for Domains

If D is a countably based bounded complete domain and $\nu \in \mathrm{Prob}\, D$, then there is a Scott-continuous map $X \colon \mathbb{CT} \to D$ with $X_* \, \mu_{\mathcal{C}} = \nu$.

Moreover, if $\nu_n, \nu \in \mathrm{Prob}\, D$ satisfy $\nu_n \to_w \nu$, then there are Scott-continuous maps $X_n, X \colon \mathbb{CT} \to D$ with $X_* \, \mu_{\mathcal{C}} = \nu$, $X_{n*} \, \mu_{\mathcal{C}} = \nu_n$ and $X_n \to X$ in $[\mathbb{CT} \to D]$.

BCD_ω is Cartesian closed:

- $[D \to E] = \{f \colon D \to E \mid f \text{ Scott continuous}\}$
- $f \leq g$ iff $f(x) \leq g(x)$ $(\forall x \in D)$.

So: $X \mapsto X_* \mu_{\mathcal{C}} \colon [\mathbb{CT} \to D] \twoheadrightarrow (\mathrm{Prob}\, D, \mathit{weak})$ is continuous surjection

Domain-theoretic Skorohod Theorem (cont'd)

Skorohod's Theorem for Domains

If D is a countably based bounded complete domain and $\nu \in \mathrm{Prob}\, D$, then there is a Scott-continuous map $X \colon \mathbb{CT} \to D$ with $X_* \, \mu_{\mathcal{C}} = \nu$.

Moreover, if $\nu_n, \nu \in \mathrm{Prob}\, D$ satisfy $\nu_n \to_w \nu$, then there are Scott-continuous maps $X_n, X \colon \mathbb{CT} \to D$ with $X_* \, \mu_{\mathcal{C}} = \nu$, $X_{n*} \, \mu_{\mathcal{C}} = \nu_n$ and $X_n \to X$ in $[\mathbb{CT} \to D]$.

Corollary: Skorohod's Theorem

Proof: If S is Polish, then $(\mathrm{Prob}\, S, \mathit{weak}) \hookrightarrow (\mathrm{Max}\, \mathrm{Prob}\, D, \mathit{weak})$ for some BCD_ω D. Then $\nu \in \mathrm{Prob}\, S \;\Rightarrow\; (\exists X \colon \mathbb{CT} \to D)\, X_* \mu_{\mathcal{C}} = \nu$.

$X|_{\mathcal{C}} \colon \mathcal{C} \to D$ is measurable is easy argument. □

Note: $(\mathbb{CT}, \mu_{\mathcal{C}})$ is a standard probability space (mod 0), so we get more information about X, X_n: they're all Scott continuous.

Deflations

$\phi\colon D \to D$ is a *deflation* if ϕ is Scott continuous and $\phi(D)$ is finite.

$D \in \mathsf{BCD}_\omega \implies 1_D = \sup_n \phi_n,\ \phi_n \le \phi_{n+1}$, deflations

Example: $\pi_n\colon \mathbb{CT} \to \downarrow\mathcal{C}_n,$ where $\mathcal{C}_n \simeq 2^n$

Prob functorial $\implies 1_{\mathsf{Prob}\,D} = \sup_n \phi_{n*}$

Example: $\mu_{\mathcal{C}} = \sup_n \pi_{n*}\,\mu_{\mathcal{C}} = \sup_n \mu_{\mathcal{C}_n}$

So: If $D \in \mathsf{BCD}_\omega$ and $\mu \in \mathsf{Prob}\,D$, then $\mu = \sup_n \phi_{n*}\,\mu$
with $\phi_{n*}\,\mu = \sum_{x\in F_n} r_x\delta_x$, where F_n finite for all n.

Interlude: Valuations and Probability Measures

The Domain of Valuations

$\mathbb{V}(D) = \{v \colon \Sigma(D) \to [0,1] \mid v \text{ a valuation}\}$:

- $v(\emptyset) = 0$ & $v(D) \leq 1$
- $v(U \cup V) + v(U \cap V) = v(U) + v(V)$ $(\forall U, V \in \Sigma(D))$
- v Scott continuous: $v(\bigcup_i U_i) = \sup_i v(U_i)$ $(\forall \{U_i\}_i$ directed).

Define: $v \leq v'$ iff $v(U) \leq v'(U)$ $(\forall U \in \Sigma(D))$.

Fact: D (Lawson compact) domain $\Longrightarrow$
$\mathbb{V}(D)$ (Lawson compact) domain

Interlude: Valuations and Probability Measures

Prob D is a Domain

1° $\mathbb{V}_1(D) = \{v \in \mathbb{V}(D) \mid v(D) = 1\}$ also a domain in pointwise order.

2° $\mathrm{Prob}\, D \simeq \mathbb{V}(D)$

$\mu \leq \nu$ iff $\int f d\mu \leq \int f d\nu$ ($\forall f : D \to \mathbb{R}_+$ Scott continuous)

3° D Lawson compact $\Rightarrow$ $(\mathrm{Prob}(D), \mathit{weak}) = (\mathrm{Prob}(D), \mathit{Lawson})$.

(Edalat; van Breugel, M., Ouaknine and Worrell
- plays off Portmanteau Theorem).

Finitary Mappings

Fix $\mu \in \mathrm{Prob}\, D$, and fix $\phi_{n*}\, \mu = \sum_{x \in F_n} r_x \delta_x$.

Want: $f_n \colon \mathcal{C}_{m_n} \to D$ with $f_{n*}\, \mu_{\mathcal{C}_{m_n}} = \phi_{n*}\, \mu$ for some $m_n > n$.

Requires: r_x must be dyadic for $x \in F_n$. ↯

So, we *approximate* $\phi_{n*}\, \mu$:

Choose $m_n > n, |F_n|$ with $r_x - s_x < \frac{1}{2^{m_n}}$ $(\forall x \in F_n)$, where

$$s_x = \max \downarrow (r_x \cap Dyad_{m_n}), \text{ with } Dyad_{m_n} = \{\tfrac{k}{2^{m_n}} \mid k \leq 2^{m_n}\}.$$

Choose $y_x \ll x$ for each $x \in F_n$.

Let $\nu_n = \sum_{x \in F_n} s_x \delta_{y_x} + (1 - \sum s_x)\delta_\perp$.

Define $f_n \colon \mathcal{C}_{m_n} \to F_n \cup \{\perp\} \subseteq D$ by

$f_n^{-1}(x) = s_x$ $(\forall x \in F_n)$, and $f_n^{-1}(\perp) = 1 - \sum_{x \in F_n} s_x$.

Then $f_{n*}\, \mu_{\mathcal{C}_{m_n}} = \nu_n \ll \phi_{n*}\, \mu$

Outline of Proof

We can extend f_n to $\widetilde{f}_n\colon {\downarrow}\mathcal{C}_{m_n} \to D$ with $\widetilde{f}_{n*}\,\mu_{\mathcal{C}_{m_n}} = \phi_{n*}\,\mu$.

Problem: $\{\widetilde{f}_n\}_n$ is not a chain.

Proposition: Let $\nu = \sum_{x\in F} r_x\delta_x \le \sum_{y\in G} s_y\delta_y = \nu' \in \mathrm{Prob}\, D$.

Assume r_x, s_y are dyadic rationals for each $x \in F, y \in G$.

Suppose $f_m\colon \mathcal{C}_{k_m} \to D$ satisfies $f_{m*}\,\mu_{\mathcal{C}_{k_m}} = \nu$.

Then there are $n > m$, $k_n > k_m$, and $f_n\colon \mathcal{C}_{k_n} \to D$ satisfying:

- $f_{n*}\,\mu_{\mathcal{C}_{k_n}} = \nu'$, and
- $f_m \circ \pi_{k_m k_n} \le f_n$, where $\pi_{k_m k_n}\colon \mathcal{C}_{k_n} \to \mathcal{C}_{k_m}$ is the canonical projection, which implies $\widetilde{f}_m \le \widetilde{f}_n$.

The proof uses the Splitting Lemma, the fact that if r_x, s_y are dyadic, then the transport numbers $t_{x,y}$ are, too, and a generalization of Hall's Marriage Problem.

The proof of the first part of the Theorem follows by recursively defining the mappings $\widetilde{f}_n$, starting with $\widetilde{f}_0\colon \mathcal{C}_0 \to D$ by $f_0(\mathbf{1}) = \perp_D$, and then letting $X = \sup_n \widetilde{f}_n$.

Questions?