

A domain theory for statistical probabilistic programming

Matthijs Vákár, Ohad Kammar, and Sam Staton



Laboratoire Spécification et Vérification Seminar
École Normale Supérieure Paris-Saclay
29 January 2019



Statistical probabilistic programming

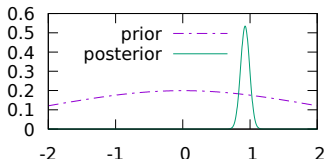
$\llbracket - \rrbracket$: programs \rightarrow unnormalised distributions

- ▶ Bayesian inference: compiler computes normalisation
- ▶ Continuous types: $\mathbb{R}, [0, \infty]$
- ▶ Probabilistic effects:

normally
distributed
sample

sample(μ, σ) : \mathbb{R}

$\llbracket \text{sample}(0, 2) \rrbracket$



scale
distribution
by r

$r : [0, \infty]$
score(r) : 1

conditioning/fitting
to observed data
with likelihood

prior

```
let  $x = \text{sample}(0, 2)$   
in score( $\text{normalPdf}(1.1 \mid x, \frac{1}{4})$ );  
    score( $\text{normalPdf}(1.9 \mid 2x, \frac{1}{4})$ );  
    score( $\text{normalPdf}(2.7 \mid 3x, \frac{1}{4})$ );  
 $x$ 
```

posterior

Statistical probabilistic programming

- ▶ Commutativity/exchangability/Fubini

Exact Bayesian inference
using disintegration
[Shan-Ramsey'17]

$$\left[\begin{array}{l} \text{let } x = K \text{ in} \\ \text{let } y = L \text{ in} \\ f(x, y) \end{array} \right] = \left[\begin{array}{l} \text{let } y = L \text{ in} \\ \text{let } x = K \text{ in} \\ f(x, y) \end{array} \right] \quad \int \llbracket K \rrbracket (dx) \int \llbracket L \rrbracket (dy) f(x, y) = \int \llbracket L \rrbracket (dy) \int \llbracket K \rrbracket (dx) f(x, y)$$

probability
distributions



σ -finite
distributions



not closed under
push-forward

arbitrary
distributions



s-finite
distributions



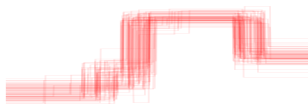
full definability
[Staton'17]

Statistical probabilistic programming

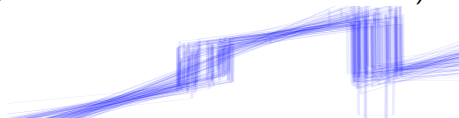
Express continuous distributions using:

- Higher-order functions:

(e.g. generative random function models)



piecewise(random-constant)

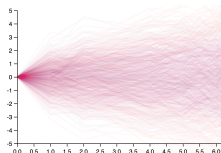


piecewise(random-linear)

- Term recursion:

```
rw(x, σ) = λ().    // thunk  
  let y = sample(x, σ)  
  in (x, rw(y, σ))
```

(e.g. Gaussian random walk)



- Type recursion (à la FPC)

(e.g. dynamic types, IRs)

$$Dynamic = \mu\alpha. \{ \text{Val}(\mathbb{R}) \mid \text{Fun}(\alpha \rightarrow \alpha) \}$$

Application: modular Bayesian inference

Resample-Move Sequential Monte Carlo

[Ścibior et al.'18a+b]

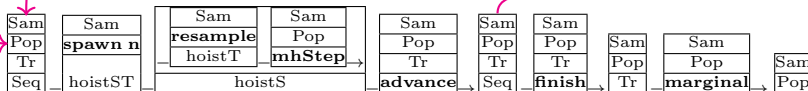
resamples k particles n moves t recursion

```
rmsmc k n t =  
  marginal . finish . compose k (  
    advance . hoistS (  
      compose t mhStep . hoistT resample  
    )  
  ) . hoistST (spawn n >>)
```

inference representation

inference transformer


higher order



inference transformation (invariant preserving)

recursive types

ProbProg: Important Language Features

Church  WebPPL Venture	sample	\mathbb{R}	score	higher order	term rec	type rec	Fubini (commute)
sets + probability	✓	✗	✗	✓	✗	✗	✓
meas space + subprobability	✓	✓	✗	✗	1 st	✗	✓
CPO + subprobability	✓	✓	✗	✓	✓	✓	?
cont domain + subprobability [Jones-Plotkin'89]	✓	✓	✗	✗	1 st	✗	✓
⋮ [Jung-Tix'98]	⋮	⋮	⋮	⋮	⋮	⋮	⋮
meas + s-finite distributions [Staton'17]	✓	✓	✓	✗	✗	✗	✓
qbs + s-finite distributions [Heunen et al'17, Ścibior et al'18]	✓	✓	✓	✓	✗	✗	✓
coh/meas cone + probability [Ehrhard-Pagani-Tasson'18, Ehrhard-Tasson'15-'19]	✓	✓ ✗	✗	✓	✓	? ✓	? ✓
ω qbs + s-finite distributions [This work]	✓	✓	✓	✓	✓	✓	✓



Summary

Contribution

- ▶ $\omega\mathbf{Qbs}$: a category of pre-domain quasi-Borel spaces
- ▶ M : commutative probabilistic powerdomain over $\omega\mathbf{Qbs}$
- ▶ Axiomatic treatment of measure and domain theory in $\omega\mathbf{Qbs}$
- ▶ Adequacy: $(\omega\mathbf{Qbs}, M)$ adequately interprets:
 - ▶ Statistical FPC
 - ▶ Untyped Statistical λ -calculus

This talk

- ▶ $\omega\mathbf{Qbs}$
- ▶ A probabilistic powerdomain
- ▶ Axiomatic treatment

Iso-recursive types: FPC

type variable contexts

$$\Delta = \{\alpha_1, \dots, \alpha_n\}$$

[Fiore-Plotkin'94]

$$\frac{\Delta, \alpha \vdash_k \tau : \text{type}}{\Delta \vdash_k \mu\alpha.\tau : \text{type}}$$

$$\begin{aligned} \text{Lam} = \mu\alpha. \{ & \text{Bool}\{\text{True} \mid \text{False}\} \\ & \mid \text{App}(\alpha * \alpha) \\ & \mid \text{Abs}(\alpha \rightarrow \alpha) \} \end{aligned}$$

type recursion

$$\frac{\Gamma \vdash t : \sigma[\alpha \mapsto \tau]}{\Gamma \vdash \tau.\text{roll}(t) : \tau} \quad \frac{\Gamma \vdash t : \tau \quad \Gamma, x : \sigma[\alpha \mapsto \tau] \vdash s : \rho}{\Gamma \vdash \text{match } t \text{ with roll } x \Rightarrow s : \rho}$$

Iso-recursive types: FPC

type variable contexts

$$\Delta = \{\alpha_1, \dots, \alpha_n\}$$

[Fiore-Plotkin'94]

$$\frac{\Delta, \alpha \vdash_k \tau : \text{type}}{\Delta \vdash_k \mu\alpha.\tau : \text{type}}$$

ω Cpo-enriched
category of
domains

type recursion

$$[\![\Delta \vdash_k \tau : \text{type}]\!] : (\mathcal{C}^{\text{op}})^n \times \mathcal{C}^n \rightarrow \mathcal{C}$$

$$[\![\Delta \vdash_k \mu\alpha.\tau : \text{type}]\!] = \text{minimal invariants}$$

[Freyd'91,92,
Pitts'96]

locally continuous
functor

Challenge

- ▶ probabilistic powerdomain
 - ▶ commutativity/Fubini
 - ▶ domain theory
 - ▶ higher-order functions
- continuous domains [Jones-Plotkin'89]
- open problem [Jung-Tix'98]
-

traditional approach:

domain \mapsto Scott-open sets \mapsto Borel sets \mapsto distributions/valuations

our approach: ^{as in} [Ehrhard-Pagani-Tasson'18]

(domain, quasi-Borel space) \mapsto distributions

separate
but compatible

Rudimentary measure theory

Borel sets

- ▶ $[a, b]$ Borel
- ▶ A Borel $\implies A^c$ Borel
- ▶ $(A_n)_{n \in \mathbb{N}}$ Borel $\implies \bigcup_{n \in \mathbb{N}} A_n$ Borel

Measurable functions $f: \mathbb{R} \rightarrow \mathbb{R}$

$$f^{-1}[A] \text{ Borel} \iff A \text{ Borel}$$

Measures $\mu: \text{Borel} \rightarrow [0, \infty]$

- ▶ monotone:
 $A \subseteq B \implies \mu(A) \leq \mu(B)$
- ▶ Scott-continuous:
 $A_0 \subseteq A_1 \subseteq \dots \implies \mu(\bigcup_n A_n) = \bigvee_n \mu(A_n)$

Rudimentary measure theory

Borel sets

- ▶ $[a, b]$ Borel
- ▶ A Borel $\implies A^c$ Borel
- ▶ $(A_n)_{n \in \mathbb{N}}$ Borel \implies
 $\bigcup_{n \in \mathbb{N}} A_n$ Borel

Measurable functions $f: \mathbb{R} \rightarrow \mathbb{R}$

$$f^{-1}[A] \text{ Borel} \iff A \text{ Borel}$$

Measures $\mu: \text{Borel} \rightarrow [0, \infty]$

- ▶ monotone:
 $A \subseteq B \implies \mu(A) \leq \mu(B)$
- ▶ Scott-continuous:
 $A_0 \subseteq A_1 \subseteq \dots \implies \mu\left(\bigcup_n A_n\right) = \bigvee_n \mu(A_n)$

1 dimensional

Example (Lebesgue measures)

$$\begin{aligned}\lambda[a, b] &= b - a \text{ on } \mathbb{R} \\ (\lambda \otimes \lambda)([a, b] \times [c, d]) &= \\ &= (b - a)(d - c) \text{ on } \mathbb{R}^2\end{aligned}$$

2 dimensional

Example (Push-forward measure)

$$f_*\mu(A) := \mu(f^{-1}[A])$$

Borel set

measure

$f: \mathbb{R} \rightarrow \mathbb{R}$

Quasi-Borel pre-domains

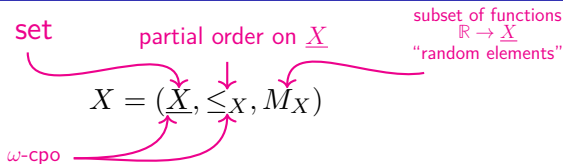
ω -qbs:

set partial order on \underline{X} subset of functions $\mathbb{R} \rightarrow \underline{X}$
"random elements"

$$X = (\underline{X}, \leq_X, M_X)$$

Quasi-Borel pre-domains

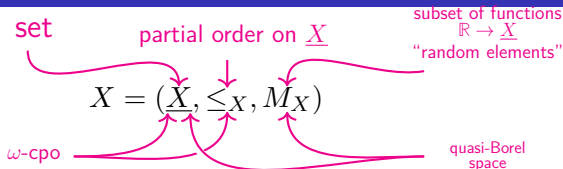
ω -qbs:



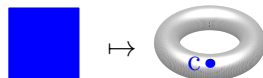
$$\blacksquare x_0 \leq x_1 \leq x_2 \leq \dots \quad \implies \quad \exists \bigvee_n x_n$$

Quasi-Borel pre-domains

ω -qbs:

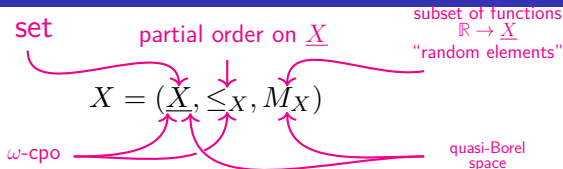


- $\lambda_{_}.x \in M_X$



Quasi-Borel pre-domains

ω -qbs:



- $\lambda_.x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$

$\mathbb{R} \xrightarrow{\varphi}_{\text{Borel}} \mathbb{R}$



$\models \varphi$

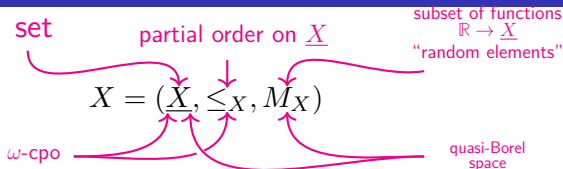


$\models \alpha$



Quasi-Borel pre-domains

ω -qbs:

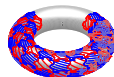


- $\lambda_.x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n. \alpha(r)] \in M_X$

$$\mathbb{R} \xrightarrow{\varphi} \text{Borel } \mathbb{R}$$



$$[S_n. \alpha_n]$$

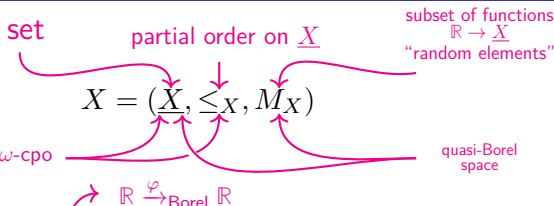


Borel measurable
countable partition

$$\mathbb{R} = \biguplus_{n \in \mathbb{N}} S_n$$

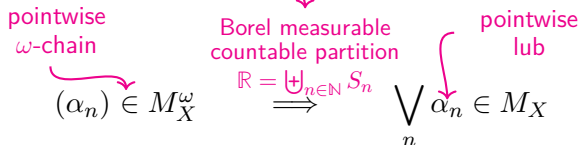
Quasi-Borel pre-domains

ω -qbs:



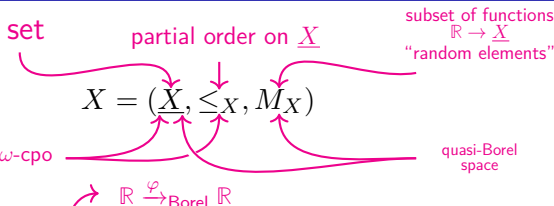
- $\lambda _.x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n. \alpha(r)] \in M_X$

s.t.:



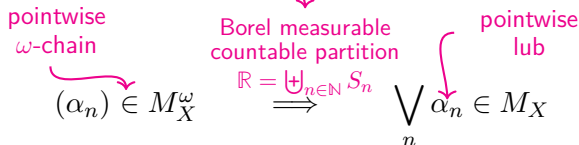
Quasi-Borel pre-domains

ω -qbs:



- $\lambda _ . x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n . \alpha(r)] \in M_X$

s.t.:



Morphisms $f : X \rightarrow Y$: Scott continuous qbs maps

monotone and
 $f \bigvee_n x_n = \bigvee_n f x_n$

$\forall \alpha \in M_X .$
 $f \circ \alpha \in M_Y$



Example

$S = (\underline{S}, \Sigma_S)$ measurable space

$$(\underline{S}, =, \{\alpha : \mathbb{R} \rightarrow \underline{S} \mid \alpha \text{ Borel measurable}\})$$

so $\mathbb{R} \in \omega\mathbf{Qbs}$

Reminder

wqbs: $X = (\underline{X}, \leq_X, M_X)$

- $\lambda_{\underline{\cdot}}.x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n. \alpha(r)] \in M_X$

s.t.:

$$(\alpha_n) \in M_X^\omega \implies \bigvee_n \alpha_n \in M_X$$



Example

$P = (\underline{P}, \leq_P)$ ω -cpo

lubs of
step functions

$$\left(\underline{P}, \leq_P, \left\{ \bigvee_k [_ \in S_n^k \cdot a_n^k] \mid \forall k. \mathbb{R} = \bigcup_n S_n^k \right\} \right)$$

so $\mathbb{L} = ([0, \infty], \leq, \{\alpha : \mathbb{R} \rightarrow [0, \infty] \mid \alpha \text{ Borel measurable}\}) \in \omega\mathbf{Qbs}$

Reminder

wqbs: $X = (\underline{X}, \leq_X, M_X)$

- $\lambda_ . x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n \cdot \alpha(r)] \in M_X$

s.t.:

$$(\alpha_n) \in M_X^\omega \implies \bigvee_n \alpha_n \in M_X$$



Example

X ω -qbs

$$X_{\perp} := \left(\{\perp\} + \underline{X}, \perp \leq \underline{X}, \left\{ [S.\perp, S^{\mathbb{C}}.\alpha] \mid \alpha \in M_X, S \text{ Borel} \right\} \right)$$

Reminder

wqbs: $X = (\underline{X}, \leq_X, M_X)$

- $\lambda_{\perp}.x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n.\alpha(r)] \in M_X$

s.t.:

$$(\alpha_n) \in M_X^{\omega} \implies \bigvee_n \alpha_n \in M_X$$

Quasi-Borel pre-domains

Products

$$\underline{X_1} \times \underline{X_2} = \underline{X_1} \times \underline{X_2} \qquad x \leq y \iff \forall i. x_i \leq y_i$$

$$M_{X_1 \times X_2} = \{(\alpha_1, \alpha_2) : \mathbb{R} \rightarrow \underline{X_1} \times \underline{X_2} \mid \forall i. \alpha_i \in M_{X_i}\}$$



correlated
random elements

Quasi-Borel pre-domains

Products

$$\underline{X_1} \times \underline{X_2} = \underline{X_1} \times \underline{X_2} \quad x \leq y \iff \forall i. x_i \leq y_i$$

$$M_{X_1 \times X_2} = \{(\alpha_1, \alpha_2) : \mathbb{R} \rightarrow \underline{X_1} \times \underline{X_2} \mid \forall i. \alpha_i \in M_{X_i}\}$$

Theorem

$\omega\mathbf{Qbs} \rightarrow \omega\mathbf{Cpo} \times \mathbf{Qbs}$ *creates limits*

correlated
random elements




Quasi-Borel pre-domains

Products

$$\underline{X_1} \times \underline{X_2} = \underline{X_1} \times \underline{X_2} \quad x \leq y \iff \forall i. x_i \leq y_i$$

$$M_{\underline{X_1} \times \underline{X_2}} = \{(\alpha_1, \alpha_2) : \mathbb{R} \rightarrow \underline{X_1} \times \underline{X_2} \mid \forall i. \alpha_i \in M_{X_i}\}$$



correlated
random elements

Exponentials

► $\underline{Y^X} = \{f : \underline{X} \rightarrow \underline{Y} \mid f \text{ Scott continuous qbs morphism}\}$
 $= \mathbf{Qbs}(X, Y)$

► $f \leq g \iff \forall x \in \underline{X}. f(x) \leq g(x)$

► $M_{Y^X} = \left\{ \alpha : \mathbb{R} \rightarrow \underline{Y^X} \mid \begin{array}{l} \text{uncurry } \alpha : \mathbb{R} \times X \rightarrow Y \\ \text{Scott continuous qbs morphism} \end{array} \right\}$
so $\underline{Y^{\mathbb{R}}} = M_Y$

Fundamentals of measure theory

s-finite measures

► μ_n **bounded**:

$$\mu_n(\mathbb{R}) < \infty$$

► μ **s-finite**:

$$\mu = \sum_n \mu_n, \mu_n \text{ bounded}$$

Randomisation Theorem

Every s-finite measure is a push-forward of Lebesgue:

$$\mu \text{ s-finite} \implies \mu = f_*\lambda \text{ for some } f : \mathbb{R} \rightarrow \mathbb{R}_\perp$$

Transfer principle

$$\tau_*\lambda = \lambda \otimes \lambda \text{ for some measurable } \tau : \mathbb{R} \rightarrow (\mathbb{R} \times \mathbb{R})_\perp$$

Randomisation monad structure

▶ $(X_{\perp})^{\mathbb{R}}$

▶ $\text{return}_X(x) : r \in [0, 1] \mapsto x$

▶ $(\alpha \gg= f) : \mathbb{R} \xrightarrow{\tau} \mathbb{R} \times \mathbb{R} \xrightarrow{\mathbb{R} \times \alpha} \mathbb{R} \times X \xrightarrow{\mathbb{R} \times f} \mathbb{R} \times (Y_{\perp})^{\mathbb{R}} \xrightarrow{\text{eval}} Y$

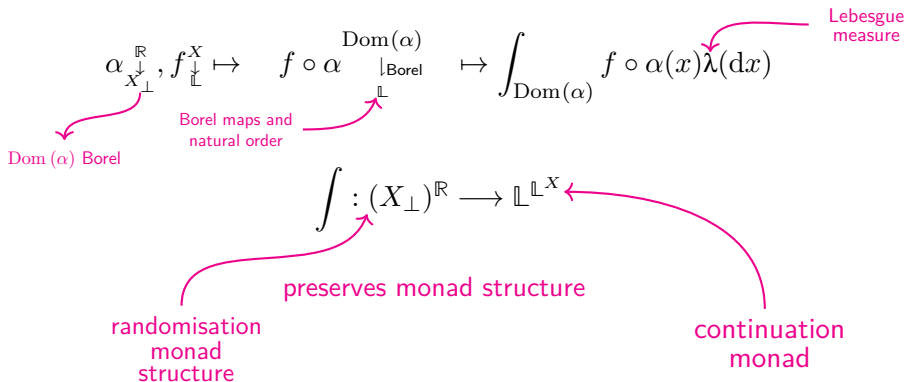
$\mathbb{R} \rightarrow X_{\perp}$ $X \rightarrow (Y_{\perp})^{\mathbb{R}}$

▶ sample from randomisation of normal distribution

▶ $\text{score}(r) : r' \in [0, |r|] \mapsto ()$

monad laws fail
(associativity)

Lebesgue integration

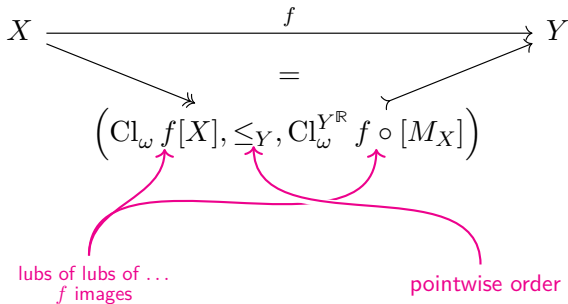


A probabilistic powerdomain

$$\begin{array}{ccc} (X_{\perp})^{\mathbb{R}} & \xrightarrow{\int} & \mathbb{L}^X \\ & \searrow \quad \swarrow & \\ & = & \\ & MX & \end{array}$$

MX : randomisable integration operators

A probabilistic powerdomain



$(\mathcal{E}, \mathcal{M}) := (\text{densely strong epi, full mono})$ factorisation system

A probabilistic powerdomain

\mathcal{E} = densely strong epis closed under:

▶ products:

$$e_1, e_2 \in \mathcal{E} \implies e_1 \times e_2 \in \mathcal{E}$$

▶ lifting:

$$e \in \mathcal{E} \implies e_{\perp} \in \mathcal{E}$$

▶ random elements:

$$e \in \mathcal{E} \implies e^{\mathbb{R}} \in \mathcal{E}$$

$\implies M$ strong monad for sampling + conditioning



[Kammar-McDermott'18]

A probabilistic powerdomain

$$(X_{\perp})^{\mathbb{R}} \begin{array}{c} \xrightarrow{\quad} \mathbb{L}^{\mathbb{L}^X} \\ \searrow \quad \nearrow \\ \quad \quad \quad = \quad \quad \quad \\ \quad \quad \quad MX \quad \quad \quad \end{array}$$

► M locally continuous \implies may appear in domain equations

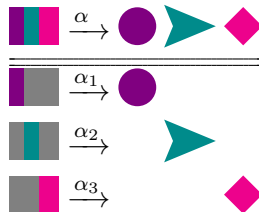
► M commutative

► M models synthetic measure theory

$$M \sum_{n \in \mathbb{N}} X_n \cong \prod_{n \in \mathbb{N}} MX_n$$

[Kock'12,
Ścibior et al.'18]

\implies satisfies Fubini



► $MX \cong \left\{ \mu \mid \text{Scott opens} \mid \mu \text{ is s-finite} \right\}$ generalises valuations

standard Borel space

Axiomatic domain theory

[Fiore-Plotkin'94, Fiore'96]

Structure

- ▶ Total map category: $\omega\mathbf{Qbs}$
- ▶ Admissible monos: **Borel-open** map $m : X \rightrightarrows Y$:

$$\forall \beta \in M_Y. \quad \beta^{-1}[m[X]] \in \mathcal{B}(\mathbb{R})$$

take Borel-Scott open maps as admissible monos

- ▶ **Pos**-enrichment: pointwise order
- ▶ Pointed monad on total maps: the powerdomain

\Rightarrow model axiomatic domain theory

\Rightarrow solve recursive domain equations

Axiomatic domain theory

Structure

\mathfrak{D} total map category
 $\omega\mathbf{Qbs}$
 $f \leq g$ **Pos**-enrichment
 pointwise order
 $\mathcal{M}_{\mathfrak{D}}$ admissible monos
 Borel-Scott opens
 T monad for effects
 power-domain
 m partiality encoding
 $m : -_{\perp} \rightarrow T, \perp \mapsto 0$

Derived axioms/structure

$\mathbf{p}\mathfrak{D}$ partial map category
 $-_{\perp}$ partiality monad
 (\dashv_{\vee}) the adjunction $J \dashv L$
 is locally continuous
 (\mathbf{p}_{\vee}) $\mathbf{p}\mathfrak{D}$ is $\omega\mathbf{Cpo}$ -enriched
 (1_{\leq}) $\mathbf{p}\mathfrak{D}$ has a partial terminal

Axioms

(\dashv) every object has a partial
 map classifier $\downarrow_X : X \rightarrow X_{\perp}$
 (fup) every admissible mono is full
 and upper-closed
 (\dashv_{\leq}) $\lfloor - \rfloor$ is locally monotone
 (\bigvee) \mathfrak{D} is $\omega\mathbf{Cpo}$ -enriched
 (U) ω -colimits behave uniformly
 (1) \mathfrak{D} has a terminal object
 (\rightarrow_{\leq}) \mathfrak{D} has locally monotone
 exponentials
 $(+)$ locally continuous total
 coproducts
 $(?!)$ $0 \rightarrow 1$ is admissible
 (\times_{\vee}) \mathfrak{D} has a locally
 continuous products
 (CL) \mathfrak{D} is cocomplete
 (T_{\vee}) T is locally continuous

(\otimes) $\mathbf{p}\mathfrak{D}$ has partial products
 (\otimes_{\vee}) (\otimes) is locally continuous
 (\rightarrow_{\vee}) \mathfrak{D} has locally continuous
 exponentials
 (\Rightarrow_{\vee}) $\mathbf{p}\mathfrak{D}$ has locally continuous
 partial exponentials
 $(\mathbf{p}CL)$ $\mathbf{p}\mathfrak{D}$ is cocomplete
 $(\mathbf{p}+\vee)$ $\mathbf{p}\mathfrak{D}$ has locally continuous
 partial coproducts
 (BC) $J : \hookrightarrow \mathbf{p}\mathfrak{D}$ is a bilimit
 compact expansion

Summary

Contribution

- ▶ $\omega\mathbf{Qbs}$: a category of pre-domain quasi-Borel spaces
- ▶ M : commutative probabilistic powerdomain over $\omega\mathbf{Qbs}$
- ▶ Axiomatic treatment of measure and domain theory in $\omega\mathbf{Qbs}$
- ▶ Adequacy: $(\omega\mathbf{Qbs}, M)$ adequately interprets:
 - ▶ Statistical FPC
 - ▶ Untyped Statistical λ -calculus

[Fiore-Plotkin'94, Fiore'96]


This talk

- ▶ $\omega\mathbf{Qbs}$
- ▶ A probabilistic powerdomain
- ▶ Axiomatic treatment

Also in the paper

- ▶ Axiomatic domain theory
- ▶ Operational semantics
à la [Borgström et al.'16]
- ▶ Characterising $\omega\mathbf{Qbs}$

ProbProg: Important Language Features

Church  WebPPL Venture	sample	\mathbb{R}	score	higher order	term rec	type rec	Fubini (commute)
sets + probability	✓	✗	✗	✓	✗	✗	✓
meas space + subprobability	✓	✓	✗	✗	1 st	✗	✓
CPO + subprobability	✓	✓	✗	✓	✓	✓	?
cont domain + subprobability [Jones-Plotkin'89]	✓	✓	✗	✗	1 st	✗	✓
⋮ [Jung-Tix'98]	⋮	⋮	⋮	⋮	⋮	⋮	⋮
meas + s-finite distributions [Staton'17]	✓	✓	✓	✗	✗	✗	✓
qbs + s-finite distributions [Heunen et al'17, Ścibior et al'18]	✓	✓	✓	✓	✗	✗	✓
coh/meas cone + probability [Ehrhard-Pagani-Tasson'18, Ehrhard-Tasson'15-'19]	✓	✓ ✗	✗	✓	✓	? ✓	? ✓
ω qbs + s-finite distributions [This work]	✓	✓	✓	✓	✓	✓	✓

