

A domain theory for statistical probabilistic programming

Ohad Kammar
with
Matthijs Vákár and Sam Staton

7th ACM SIGPLAN Workshop on
Higher-Order Programming with Effects
23 September 2018



UNIVERSITY OF
OXFORD

EPSRC
Engineering and Physical Sciences
Research Council



THE ROYAL
SOCIETY

The
Alan Turing
Institute

Statistical probabilistic programming

$\llbracket - \rrbracket : \text{programs} \rightarrow \text{distributions}$

► Continuous types: $\mathbb{R}, [0, \infty]$

► Probabilistic effects:

normally
distributed
sample

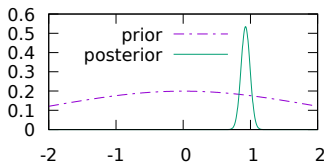
$\text{sample}(\mu, \sigma) : \mathbb{R}$

scale
distribution
by r

$\frac{r : [0, \infty]}{\text{score}(r) : 1}$

conditioning/fitting
to observed data

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



prior

```
 $\llbracket$  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( $\text{normalPdf}(2.7 \mid 3x, \frac{1}{4})$ );  
   $x$   
 $\rrbracket$ 
```

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

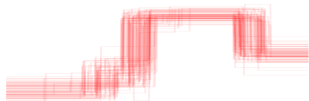


full definability
[Staton'17]

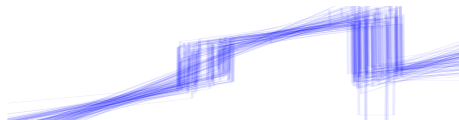
Statistical probabilistic programming

Express continuous distributions using:

- Higher-order functions



piecewise(random-constant)



piecewise(random-linear)

example: generative random function models

measure theory



Theorem (Aumann'61)

measurable cones
and stable
measurable functions



[Heunen et al.'17]

quasi-Borel spaces



[Ehrhard-Pagani-Tasson'18]

No σ -algebra over $\mathbf{Meas}(\mathbb{R}, \mathbb{R})$ with measurable evaluation:

$$\text{eval} : \mathbf{Meas}(\mathbb{R}, \mathbb{R}) \times \mathbb{R} \rightarrow \mathbb{R}$$

Statistical probabilistic programming

Express continuous distributions using:

- Inductive types and bounded iteration

Ścibior et al.'18a+b

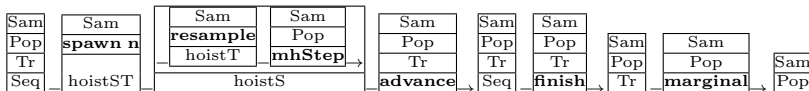
resamples

particles

moves

rmsmc k n t =

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



see Adam Ścibior's talk

Monday 17:02–17:25

Functional Programming for Modular Bayesian Inference

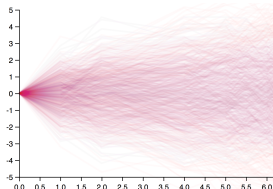
Statistical probabilistic programming

Express continuous distributions using:

[Ehrhard-Pagani-Tasson'18]

- ▶ Term recursion

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



Gaussian random walk

- ▶ Type recursion and dynamic types

Church  WebPPL
Venture

this talk

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

Summary

Contribution

- ▶ $\omega\mathbf{Qbs}$: a category of pre-domain quasi-Borel spaces
- ▶ M : commutative probabilistic powerdomain over $\omega\mathbf{Qbs}$

Theorem (adequacy)

M adequately interprets:

- ▶ *Statistical FPC*
- ▶ *Untyped Statistical λ -calculus*

This talk

- ▶ $\omega\mathbf{Qbs}$
- ▶ a powerdomain over $\omega\mathbf{Qbs}$
- ▶ a domain theory for $\omega\mathbf{Qbs}$

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:
 $A_0 \subseteq A_1 \subseteq \dots \implies \mu(\bigcup_n A_n) = \bigvee_n \mu(A_n)$

1 dimensional Example

Lebesgue measures:

$$\lambda[a, b] = b - a \text{ on } \mathbb{R}$$

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

2 dimensional

Push-forward measure

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

Borel set

measure

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

Quasi-Borel pre-domains

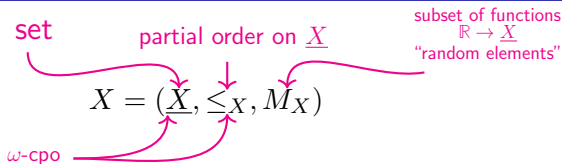
ω -qbs:

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

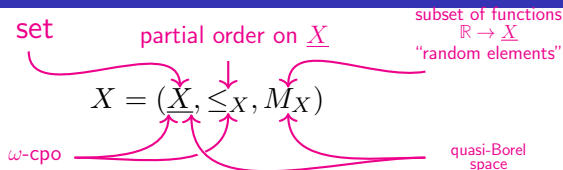
ω -qbs:



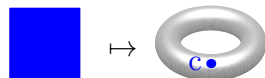
$$\bullet x_0 \leq x_1 \leq x_2 \leq \dots \quad \implies \quad \exists \bigvee_n x_n$$

Quasi-Borel pre-domains

ω -qbs:

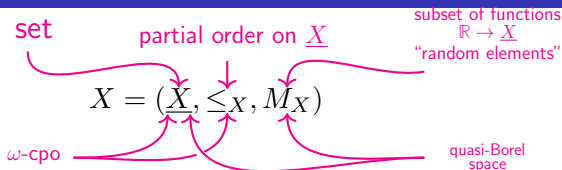


- $\lambda_.x \in M_X$



Quasi-Borel pre-domains

ω -qbs:



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

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



$$\xrightarrow{\varphi}$$

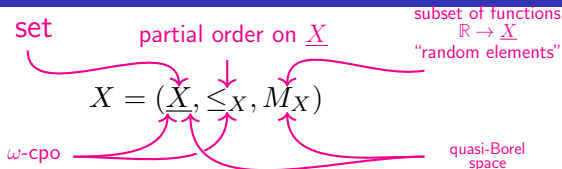


$$\xrightarrow{\alpha}$$



Quasi-Borel pre-domains

ω -qbs:

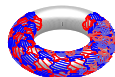


- $\lambda_.x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n \cdot \alpha(r)] \in M_X$

Borel measurable
countable partition
 $\mathbb{R} = \biguplus_{n \in \mathbb{N}} S_n$

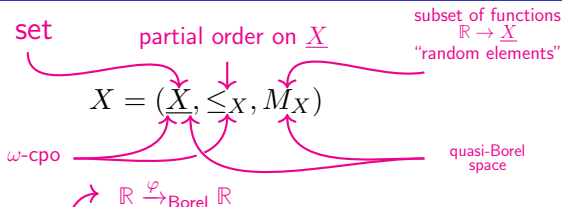


$[S_n \cdot \alpha_n]$



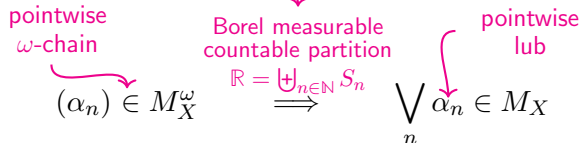
Quasi-Borel pre-domains

ω -qbs:



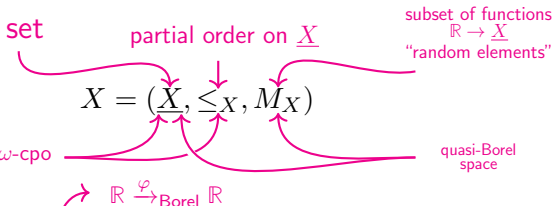
- $\lambda..x \in M_X$
- $\alpha \in M_X \implies \alpha \circ \varphi \in M_X$
- $(\alpha_n \in M_X)_{n \in \mathbb{N}} \implies [r \in S_n. \alpha(r)] \in M_X$

s.t.:

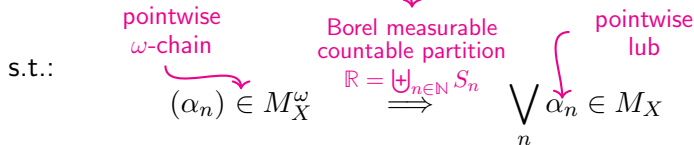


Quasi-Borel pre-domains

ω -qbs:



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



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

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

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

Quasi-Borel pre-domains

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

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

s.t.:

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


Quasi-Borel pre-domains

Example

$P = (\underline{P}, \leq_P)$ ω -cpo

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

lubs of
step functions



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

Reminder

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

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

s.t.:

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

Example

X ω -qbs

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

Reminder

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

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

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

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

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

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

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

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

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

Fundamentals of measure theory

s-finite measures

► μ **bounded**:

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

► 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

$$\tau_*\lambda = \lambda \otimes \lambda \text{ for some measurable } \tau : \mathbb{R} \xrightarrow{\cong} \mathbb{R} \times \mathbb{R}$$

Randomisation monad structure

► $(X_{\perp})^{\mathbb{R}}$

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

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

$\mathbb{R} \rightarrow X_{\perp}$

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

monad laws fail
(associativity)

Lebesgue integration

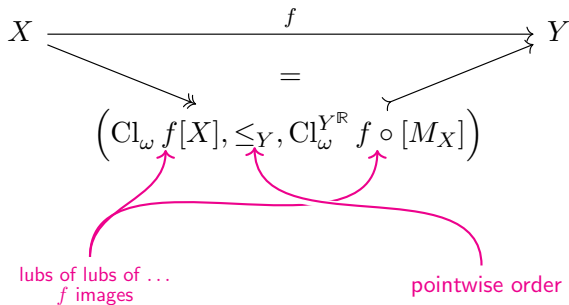
$$\begin{array}{c}
 \alpha_{X_{\perp}}^{\mathbb{R}}, f_{\mathbb{L}}^X \mapsto f \circ \alpha_{\text{Borel}}^{\text{Dom}(\alpha)} \mapsto \int_{\text{Dom}(\alpha)} f \circ \alpha(x) \lambda(dx) \\
 \text{Dom}(\alpha) \text{ Borel} \quad \text{Borel maps and natural order} \quad \text{Lebesgue measure} \\
 \int : (X_{\perp})^{\mathbb{R}} \longrightarrow \mathbb{L}^{\mathbb{L}^X} \\
 \text{preserves monad structure} \quad \text{continuation monad}
 \end{array}$$

A probabilistic powerdomain

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

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



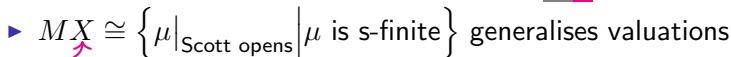
[Kammar-McDermott'18]

A probabilistic powerdomain

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

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

\Rightarrow satisfie Fubini



A domain theory for statistical probabilistic programming

Axiomatic domain theory

[Fiore-Plotkin'94, Fiore'96]

Structure

- ▶ Total map category: $\omega\mathbf{Qbs}$

- ▶ Admissible monos: **Borel-open** map $m : X \rightarrowtail Y$:

strong mono

qbses

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

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_V) the adjunction $J \dashv L$
 is locally continuous
 (\mathbf{p}_V) $\mathbf{p}\mathfrak{D}$ 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_V) \mathfrak{D} has a locally
 continuous products
 (CL) \mathfrak{D} is cocomplete
 (T_V) T is locally continuous

(\otimes) $\mathbf{p}\mathfrak{D}$ has partial products
 (\otimes_V) (\otimes) is locally continuous
 (\rightarrow_V) \mathfrak{D} has locally continuous
 exponentials
 (\Rightarrow_V) $\mathbf{p}\mathfrak{D}$ has locally continuous
 partial exponentials
 $(\mathbf{p}CL)$ $\mathbf{p}\mathfrak{D}$ is cocomplete
 $(\mathbf{p}+_V)$ $\mathbf{p}\mathfrak{D}$ 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 $\omega\mathbf{Qbs}$

Theorem (adequacy)

M adequately interprets:

- ▶ *Statistical FPC*
- ▶ *Untyped Statistical λ -calculus*

This talk

- ▶ $\omega\mathbf{Qbs}$
- ▶ a powerdomain over $\omega\mathbf{Qbs}$
- ▶ a domain theory for $\omega\mathbf{Qbs}$

Not in this talk

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