A domain theory for statistical probabilistic programming

Matthijs Vákár, Ohad Kammar, and Sam Staton



Second Workshop on Probabilistic Interactive and Higher-Order Computation 6-7 February 2019

























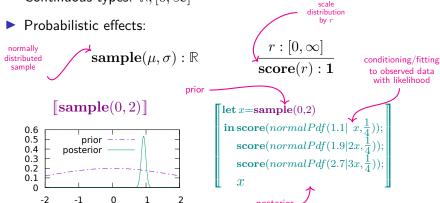






Statistical probabilistic programming

- $\llbracket \rrbracket : programs \rightarrow unnormalised distributions$
- ▶ Bayesian inference: compiler computes normalisation
- ▶ Continuous types: $\mathbb{R}, [0, \infty]$



Statistical probabilistic programming

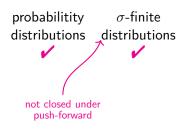
Commutativity/exchangability/Fubini-

$$\begin{bmatrix} \mathbf{let} \ x = K \mathbf{in} \\ \mathbf{let} \ y = L \mathbf{in} \end{bmatrix} = \begin{bmatrix} \mathbf{let} \ y = L \mathbf{in} \\ \mathbf{let} \ x = K \mathbf{in} \end{bmatrix}$$

 $\begin{bmatrix} \mathbf{let} \ x = K \ \mathbf{in} \\ \mathbf{let} \ y = L \ \mathbf{in} \\ f(x,y) \end{bmatrix} = \begin{bmatrix} \mathbf{let} \ y = L \ \mathbf{in} \\ \mathbf{let} \ x = K \ \mathbf{in} \\ f(x,y) \end{bmatrix} \quad \int \begin{bmatrix} K \end{bmatrix} (\mathrm{d}x) \int \begin{bmatrix} L \end{bmatrix} (\mathrm{d}y) f(x,y) \\ = \int \begin{bmatrix} L \end{bmatrix} (\mathrm{d}y) \int \begin{bmatrix} K \end{bmatrix} (\mathrm{d}x) f(x,y)$

arbitrary

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



distributions s-finite distributions full definability [Staton'17]

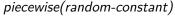
Statistical probabilistic programming

Express continuous distributions using:

► Higher-order functions:

(e.g. generative random function models)



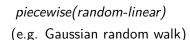


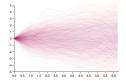
► Term recursion:

$$\begin{split} rw(x,\sigma) &= \lambda(). \qquad \text{// thunk} \\ &\mathbf{let} \ y = \mathbf{sample}(x,\sigma) \\ &\mathbf{in} \ (x,rw \ (y,\sigma)) \end{split}$$

Type recursion (à la FPC)

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





(e.g. dynamic types, IRs)

Application: modular Bayesian inference

Resample-Move Sequential Monte Carlo [Scibior et al.'18a+b] resamples particles moves recursion rmsmc k marginal . finish . compose k (advance . hoistS (compose t mhStep . hoistT resample inference representation . hoistST (spawn n >>) higher order inference transformer Sam Sam Sam Sam Sam Sam $_{\text{Sam}}$ resample Pop Pop Pop Sam Pop Pop Sam spawn n hoistT mhStep Tr Tr Tr TrPop Pop Sam Seq hoistST hoistS advance Seq finish Tr margina Pop inference recursive (invariant types preserving)

ProbProg: Important Language Features

Church RebPPL Venture	sampl	e ℝ	score	higher		٠.	
Church Venture				order	rec	rec	(commute)
sets + probability	✓	X	X	✓	X	X	✓
meas space + subprobability	✓	✓	X	X	1^{st}	X	✓
CPO + subprobability	✓	1	Х	✓	√	√	?
cont domain + subprobability	✓	✓	X	X	1^{st}	X	✓
[Jones-Plotkin'89]							
: [Jung-Tix'98]	:	:	:	÷	:	:	:
meas + s-finite distributions	✓	√	√	X	1 st	X	✓
[Staton'17]							
qbs + s-finite distributions	✓	√	√	√	1 st	X	✓
[Heunen et al'17, Ścibior et al'18]							
coh/meas cone + probability		,				2	2
[Ehrhard-Pagani-Tasson'18,	✓	V	X	✓	✓	!	!
Ehrhard-Tasson'15-'19]		^				•	•
ω qbs + s-finite distributions	✓	✓	✓	✓	✓	✓	✓
. [This work]							

Summary

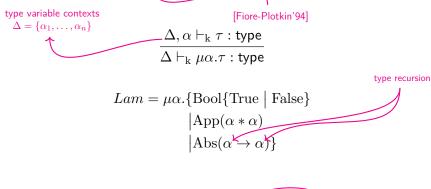
Contribution

- ightharpoonup M: commutative probabilistic powerdomain over $\omega \mathbf{Qbs}$
- lacktriangle 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

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

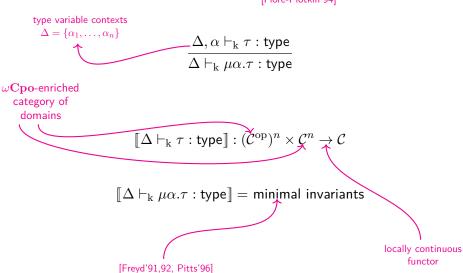
Iso-recursive types: FPC



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

Iso-recursive types: FPC

[Fiore-Plotkin'94]



Challenge

- probabilistic powerdomain
- commutativity/Fubini
- continuous domains [Jones-Plotkin'89]

open problem [Jung-Tix'98]

- domain theory
- ▶ higher-order functions

traditional approach:

 $\mathsf{domain} \mapsto \mathsf{Scott}\text{-}\mathsf{open} \ \mathsf{sets} \mapsto \mathsf{Borel} \ \mathsf{sets} \mapsto \mathsf{distributions}/\mathsf{valuations}$

our approach:

as in [Ehrhard-Pagani-Tasson'18]

 $(domain, quasi-Borel \ space) \mapsto distributions$ separatebut compatible

Rudimentary measure theory

Borel sets

- ightharpoonup [a,b] Borel
- ightharpoonup A Borel $\implies A^{\complement}$ Borel
- $(A_n)_{n \in \mathbb{N}} \text{ Borel } \Longrightarrow$ $\bigcup_{n \in \mathbb{N}} A_n \text{ Borel }$

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

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

Measures $\mu : \mathsf{Borel} \to [0, \infty]$

monotone:

$$A \subseteq B \implies \mu(A) \le \mu(B)$$

Scott-continuous:

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

▶ strict $(\mu\emptyset = 0)$ and additive $(\mu(A \uplus B) = \mu A + \mu B)$

Rudimentary measure theory

1 dimensional

Borel sets

- ightharpoonup [a,b] Borel
- ightharpoonup A Borel $\implies A^{\complement}$ Borel
- $\begin{array}{c} \blacktriangleright \ (A_n)_{n\in \mathbb{N}} \ \mathrm{Borel} \implies \\ \bigcup_{n\in \mathbb{N}} A_n \ \mathrm{Borel} \end{array}$

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

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

Measures $\mu : \mathsf{Borel} \to [0, \infty]$

- monotone: $A \subseteq B \implies \mu(A) < \mu(B)$
- Scott-continuous: $A_0 \subseteq A_1 \subseteq \ldots \implies \mu(\bigcup_n A_n) = \bigvee_n \mu(A_n)$
- strict $(\mu \emptyset = 0)$ and additive $(\mu(A \uplus B) = \mu A + \mu B)$

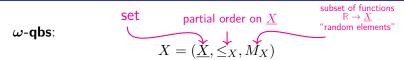
Example (Lebesgue measures)

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

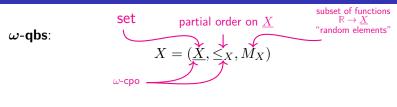
2 dimensional

Example (Push-forward measure)

$$f_*\mu(A) := \mu\left(f^{-1}[A]\right)$$
Borel set
measure
$$f: \mathbb{R} \to \mathbb{R}$$

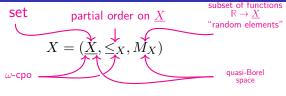


 $x_0 \le x_1 \le x_2 \le \dots$

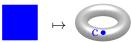


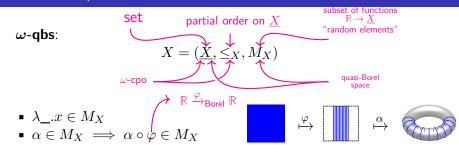
 $\exists \bigvee_{n} x_{n}$

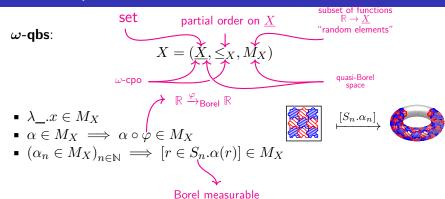




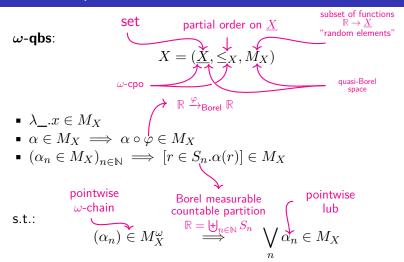
 $\lambda_x : X \in M_X$

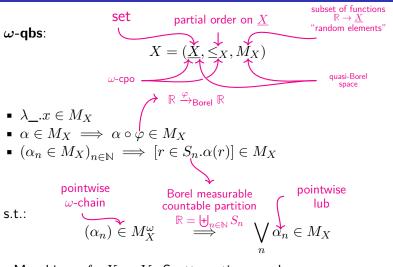






countable partition $\mathbb{R} = \left\{ + \right\}_{n \in \mathbb{N}} S_n$





Morphisms $f: X \to Y$: 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

$$\omega \mathbf{Qbs}$$
 \top \mathbf{Qbs}

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

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

Reminder

wqbs:
$$X = (X, \leq_X, M_X)$$

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

s.t.:

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

Example step functions
$$\omega \mathbf{Qbs}$$
 \top $\omega \mathbf{Cpo}$ $P = (\underline{P}, \leq_P) \omega\text{-cpo}$
$$\left(\underline{P}, \leq_P, \left\{\bigvee_k [\underline{\ } \in S_n^k.a_n^k] \middle| \forall k.\mathbb{R} = \biguplus_n S_n^k\right\}\right)$$

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

Reminder

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

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

s.t.:

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

Example

X ω -qbs

$$\omega \mathbf{Qbs}$$

$$X_{\perp} := \Big(\{\bot\} + \underline{X}, \bot \leq \underline{X}, \Big\{[S.\bot, S^{\complement}.\alpha] \Big| \alpha \in M_X, S \text{ Borel} \Big\}\Big)$$

Reminder

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

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

s.t.:

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

Products

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

$$M_{X_1 \times X_2} = \{(\alpha_1, \alpha_2) : \mathbb{R} \to \underline{X_1} \times \underline{X_2} | \forall i.\alpha_i \in M_{X_i} \}$$
 correlated random elements

Products

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

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

random elements

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

Products

$$\underline{X_1 \times X_2} = \underline{X}_1 \times \underline{X}_2 \qquad \qquad x \leq y \iff \forall i.x_i \leq y_i$$

$$M_{X_1 \times X_2} = \{(\alpha_1, \alpha_2) : \mathbb{R} \to \underline{X}_1 \times \underline{X}_2 | \forall i.\alpha_i \in M_{X_i} \}$$
 correlated random elements

Exponentials

- $\underline{Y}^X = \{f: \underline{X} \to \underline{Y} | f \text{ Scott continuous qbs morphism} \}$ $= \mathbf{Qbs}(X,Y)$
- $f \le g \iff \forall x \in \underline{X}.f(x) \le g(x)$

Fundamentals of measure theory

s-finite measures

 $\blacktriangleright \mu_n$ bounded:

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

 \blacktriangleright μ s-finite:

$$\mu = \sum_n \mu_n$$
, μ_n bounded

Randomisation Theorem

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

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

Transfer principle

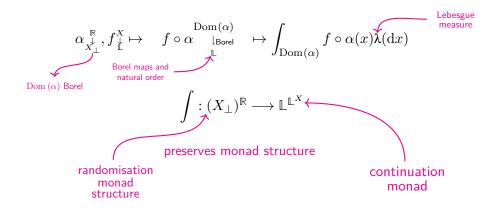
$$au_*\lambda=\lambda\otimes\lambda$$
 for some measurable $au:\mathbb{R} o(\mathbb{R} imes\mathbb{R})_\perp$

Randomisation monad structure

- $(X_{\perp})^{\mathbb{R}}$
- ightharpoonup return $_X(x): r \in [0,1] \mapsto x$
- $(\alpha \gg f) : \mathbb{R} \xrightarrow{\tau} \mathbb{R} \times \mathbb{R} \xrightarrow{\mathbb{R} \times \alpha} \mathbb{R} \times X \xrightarrow{\mathbb{R} \times f} \mathbb{R} \times (Y_{\perp})^{\mathbb{R}} \xrightarrow{\text{eval}} Y$ $X \to (Y_{\perp})^{\mathbb{R}}$
 - ▶ sample from randomisation of normal distribution
 - ightharpoonup score $(r): r' \in [0, |r|] \mapsto ()$

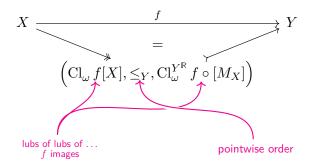
monad laws fail (associativity)

Lebesgue integration



$$(X_{\perp})^{\mathbb{R}} \xrightarrow{\int \int \mathbb{L}^{\mathbb{L}^X}} MX$$

MX: randomisable integration operators



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

 $\mathcal{E} =$ 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:

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

 \Longrightarrow M strong monad for sampling + conditioning

[Kammar-McDermott'18]

$$(X_{\perp})^{\mathbb{R}} \xrightarrow{= \longrightarrow \mathbb{L}^{\mathbb{L}^{X}}} MX$$

- ightharpoonup M locally continuous \implies may appear in domain equations
- M commutative

⇒ satisfies Fubini

[Kock'12,

▶ M models synthetic measure theory $M \sum_{n \in \mathbb{N}} X_n \cong \prod_{n \in \mathbb{N}} M X_n$

Ścibior et al.'18]

 $MX \cong \left\{ \mu |_{\text{Scott opens}} \middle| \mu \text{ is s-finite} \right\}$ generalises valuations

standard Borel space

Axiomatic domain theory

[Fiore-Plotkin'94, Fiore'96]

Structure

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

onos: **Borel-open** map
$$m: X \xrightarrow{\varphi} Y$$
: $\forall \beta \in M_Y. \qquad \beta^{-1}[m[X]] \in \mathcal{B}(\mathbb{R})$

take Borel-Scott open maps as admissible monos

- ▶ Pos-enrichment: pointwise order
- Pointed monad on total maps: the powerdomain
- → model axiomatic domain theory
- ⇒ solve recursive domain equations

Axiomatic domain theory

Structure

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

Derived axioms/structure

- partial map category partiality monad
- $(\dashv_{\mathcal{M}})$ the adjunction $J \dashv L$ is locally continuous

- $(\mathbb{1}_{<})$ $\mathbf{p}\mathfrak{D}$ has a partial terminal

Axioms

- every object has a partial $(\rightarrow_{<}) \mathfrak{D}$ has locally monotone map classifier $\downarrow_X: X \to X_\perp$
- (fup) every admissible mono is full (+) and upper-closed
- $(\dashv_{<})$ |-| is locally monotone
- \mathfrak{D} is $\omega \mathbf{Cpo}$ -enriched
- ω -colimits behave uniformly
- D has a terminal object (1)

exponentials

locally continuous total

- coproducts $\mathbb{O} \to \mathbb{1}$ is admissible
- $(\times_{\mathcal{M}})$ \mathfrak{D} has a locally
- continuous products
- (CL) \mathfrak{D} is cocomplete
- T is locally continuous
- (\otimes) pD has partial products (\otimes_V) (\otimes) is locally continuous
- D has locally continuous exponentials
- (\mathbf{p}_V) $\mathbf{p}\mathfrak{D}$ is $\omega \mathbf{Cpo}$ -enriched (\Longrightarrow_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

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

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

Also in the paper

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

ProbProg: Important Language Features

Church RebPPL Venture	sampl	e ℝ	score	higher		٠.	
Church Venture				order	rec	rec	(commute)
sets + probability	✓	X	X	✓	X	X	✓
meas space + subprobability	✓	✓	X	X	1^{st}	X	✓
CPO + subprobability	✓	1	Х	✓	√	√	?
cont domain + subprobability	✓	✓	X	X	1^{st}	X	✓
[Jones-Plotkin'89]							
: [Jung-Tix'98]	:	:	:	÷	:	:	:
meas + s-finite distributions	✓	√	√	X	1 st	X	✓
[Staton'17]							
qbs + s-finite distributions	✓	√	√	√	1 st	X	✓
[Heunen et al'17, Ścibior et al'18]							
coh/meas cone + probability		,				2	2
[Ehrhard-Pagani-Tasson'18,	✓	V	X	✓	✓	!	!
Ehrhard-Tasson'15-'19]		^				•	•
ω qbs + s-finite distributions	✓	✓	✓	✓	✓	✓	✓
. [This work]							