Higher-order building blocks for statistical modelling

Ohad Kammar
University of Edinburgh

PPS-PIHOC-DIAPASoN Workshop
17 February 2021

THE UNIVERSITY *of* EDINBURGH informatics lfcs Laboratory for Foundations of Computer Science BayesCentre

supported by: THE ROYAL SOCIETY The Alan Turing Institute Facebook Research

$S = 2, \mathbb{N}, \mathbb{R}, \ldots$

Theorem (Aumann)

No σ*-algebra on* $\mathbf{Meas}(\mathbb{R}, S)$ *makes evaluation measurable:*

$$\mathrm{eval} : \mathbf{Meas}(\mathbb{R}, S) \times \mathbb{R} \to S$$

$\Longrightarrow$ bad fit for higher-order programming semantics:

$\mathbf{Meas}$ is not Cartesian closed

higher-order measure theory

Bacci-et-al'18
Dahlqvist-Kozen'20
Ehrhard-Pagani-Tasson'18

dependent
(co)products

Recent breakthroughs:

$\times, \to$

alongside

$\mathrm{Pr}, \int$

This talk:

$\coprod, \prod$

$\mathbf{Type}, \mathbf{Prop}, \mathbf{Dec}$

higher-order measure theory
with **quasi-Borel spaces**

Bacci-et-al'18
Dahlqvist-Kozen'20
Ehrhard-Pagani-Tasson'18

dependent
(co)products

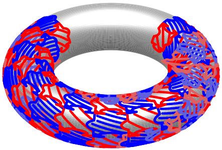
Recent breakthroughs:
$\times$, $\to$
alongside
Pr, $\int$

This talk:
$\coprod$, $\prod$
$\mathbf{Type}$, $\mathbf{Prop}$, $\mathbf{Dec}$

Quasi-Borel spaces [Heunen-Kammar-Staton-Yang'17]

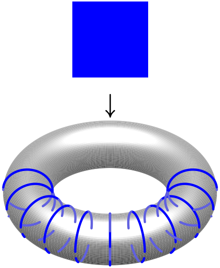
Intuition

measurable space



axiomatise
measurable events

quasi Borel space

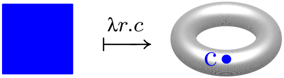


axiomatise
random elements

quasi-Borel space $X = (\llcorner X \lrcorner, \mathcal{R}_X)$

random element $\alpha \in \mathcal{R}_X \subseteq \llcorner X \lrcorner^{\mathbb{R}}$ axioms:

determinism. elements are random elements:

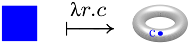


Quasi-Borel spaces [Heunen-Kammar-Staton-Yang'17]

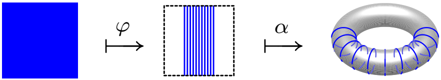
quasi-Borel space $X = (\llcorner X \lrcorner, \mathcal{R}_X)$

random element $\alpha \in \mathcal{R}_X \subseteq \llcorner X \lrcorner^{\mathbb{R}}$ axioms:

determinism.

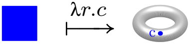


precomposition. $(\varphi \in \mathbf{Meas}(\mathbb{R}, \mathbb{R}))$

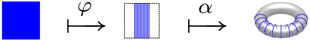


quasi-Borel space $X = (\llcorner X \lrcorner, \mathcal{R}_X)$

random element $\alpha \in \mathcal{R}_X \subseteq \llcorner X \lrcorner^{\mathbb{R}}$ axioms:

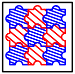
determinism. 

$\xmapsto{\lambda r.c}$

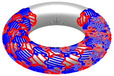
precomposition. $(\varphi \in \mathbf{Meas}(\mathbb{R}, \mathbb{R}))$ 

$\xmapsto{\varphi} \quad \xmapsto{\alpha}$

recombination $(\mathbb{R} = \biguplus_{i=0}^{\infty} S_n)$



$$\xmapsto{\lambda r.\begin{cases} \vdots \\ r \in S_n : & \alpha_n(r) \\ \vdots \end{cases}}$$



Quasi-Borel spaces [Heunen-Kammar-Staton-Yang'17]

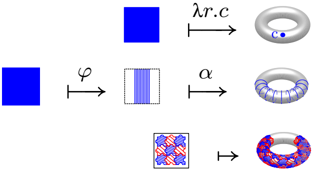
quasi-Borel space $X = (\llcorner X \lrcorner, \mathcal{R}_X)$

random element $\alpha \in \mathcal{R}_X \subseteq \llcorner X \lrcorner^{\mathbb{R}}$ axioms:

determinism.

precomposition. $(\varphi \in \mathbf{Meas}(\mathbb{R}, \mathbb{R}))$

recombination $(\mathbb{R} = \biguplus_{i=0}^{\infty} S_n)$



Morphisms $f : X \to Y$

functions $\llcorner f \lrcorner : \llcorner X \lrcorner \to \llcorner Y \lrcorner$:

$$\alpha \in \mathcal{R}_X \implies f \circ \alpha \in \mathcal{R}_Y$$

Free and cofree qbs

set A:

- free: $\mathcal{R}_{\ulcorner A \urcorner^{\mathbf{Qbs}}} = \sigma$-simple functions:

$$\lambda r. \begin{cases} \vdots \\ r \in S_n : \quad a_n \\ \vdots \end{cases}$$

- cofree: $\mathcal{R}_{\llcorner A \lrcorner_{\mathbf{Qbs}}} =$ all functions

$$\begin{tikzcd}
 & \mathbf{Qbs} \arrow[dd, "\llcorner - \lrcorner_{\mathbf{Set}}"] & \\
\arrow[r, phantom, "\dashv"] & & \arrow[l, phantom, "\dashv"] \\
 & \mathbf{Set} \arrow[uu, bend left=80, "\ulcorner - \urcorner^{\mathbf{Qbs}}"] \arrow[uu, bend right=80, "\llcorner - \lrcorner_{\mathbf{Qbs}}"'] &
\end{tikzcd}$$

qbses for:

$$2, \mathbb{N}, \mathbb{Q}$$

set-theoretic escape hatch

$\llcorner - \lrcorner_{\mathbf{Set}} : \mathbf{Qbs} \to \mathbf{Set}$ generates limits and colimits

preserves
lifts
~~reflects~~

Using measure theory

Measurable space are carried by qbses:

$$
\begin{tikzcd}
\mathbf{Sbs} \arrow[r, hook] \arrow[d, bend left, shift left] \arrow[d, phantom, "\cong"] & \mathbf{Meas} \arrow[d, bend left, "\llcorner - \lrcorner_{\mathbf{Qbs}}"] \arrow[d, phantom, "\dashv"] \\
\mathbf{Sbs} \arrow[r, hook] \arrow[u, bend left, shift left] & \mathbf{Qbs} \arrow[u, bend left, "\ulcorner - \urcorner"]
\end{tikzcd}
$$

$$
X \overset{\ulcorner - \urcorner}{\longmapsto} \left(\llcorner X \lrcorner, \begin{array}{l} U \in \mathcal{B} \iff \\ \forall \alpha. \alpha^{-1}[U] \in \mathcal{B}_{\mathbb{R}} \end{array}\right)
\qquad
S \overset{\llcorner - \lrcorner}{\longmapsto} (\llcorner S \lrcorner, \mathbf{Meas}(\mathbb{R}, S))
$$

Recover qbses for:

$$\mathbb{R}, \mathbb{W} := [0, \infty], \mathbb{I} := [0, 1]$$

Conservative extension for standard Borel spaces
Measure-theoretic escape hatch

Simple types

Simple products

Correlated random elements:

$$\mathcal{R}_{X \times Y} \xleftarrow[\cong]{(-,-)} \mathcal{R}_X \times \mathcal{R}_Y$$

Simple types

Simple products

Correlated random elements:

$$\mathcal{R}_{X \times Y} \xleftarrow[\cong]{(-,-)} \mathcal{R}_X \times \mathcal{R}_Y$$

Simple coproducts

Recombinations:

$$\alpha \in \mathcal{R}_{\coprod_{i \in \mathcal{I}} X_i} \iff \alpha = \lambda r. \begin{cases} \vdots \\ r \in S_n : & (i_n, \alpha_n r) \\ \vdots \end{cases} \qquad (\mathbb{R} = \biguplus_{n=0}^{\infty}, \alpha_n \in X_{i_n})$$

Simple types

Simple products

Correlated random elements:

$$\mathcal{R}_{X\times Y} \xleftarrow[\cong]{(-,-)} \mathcal{R}_X \times \mathcal{R}_Y$$

Simple coproducts

Recombinations:

$$\alpha \in \mathcal{R}_{\coprod_{i\in\mathcal{I}} X_i} \iff \alpha = \lambda r. \begin{cases} \vdots \\ r \in S_n : \quad (i_n, \alpha_n r) \\ \vdots \end{cases} \qquad \left(\mathbb{R} = \biguplus_{n=0}^{\infty}, \alpha_n \in X_{i_n}\right)$$

Simple function spaces

$$\llcorner Y^X \lrcorner = \mathbf{Qbs}(X, Y) \qquad \mathcal{R}_{Y^X} \xrightarrow[\cong]{\text{uncurry}} \mathbf{Qbs}(\mathbb{R} \times X, Y)$$

Random element space: $\mathcal{R}_X := X^{\mathbb{R}}$

Subspaces

$$\begin{tikzcd} X \arrow[rr, hook] & & Y \\ & \arrow[u, phantom, "="] & \\ & \mathbb{R} \arrow[uul, dashed, "\alpha'"] \arrow[uur, "\alpha"'] & \end{tikzcd}$$

m injective and $\mathcal{R}_X = (m\circ)^{-1}[\mathcal{R}_Y]$

$\Omega := \llcorner \mathbb{2} \lrcorner_{\mathbf{Qbs}}$ subspace classifier

Subspaces

$$\begin{tikzcd} X \arrow[rr, hook] & \arrow[d, phantom, "="] & Y \\ & \mathbb{R} \arrow[ul, dashed, "\alpha'"] \arrow[ur, "\alpha"'] & \end{tikzcd}$$

m injective and $\mathcal{R}_X = (m\circ)^{-1}[\mathcal{R}_Y]$

$\mathbb{\Omega} := \llcorner \underset{\mathbf{Qbs}}{2} \lrcorner$ subspace classifier

Example

- Use $[\![\mathbf{Prop}]\!] := \mathbb{\Omega}$ for reasoning/axiomatics [Sato et al.'19].

Subspaces

$$\begin{tikzcd} X \arrow[rr, hook] & & Y \\ & \mathbb{R} \arrow[ul, dashed, "\alpha'"] \arrow[ur, "\alpha"'] \arrow[u, phantom, "="] & \end{tikzcd}$$

m injective and $\mathcal{R}_X = (m\circ)^{-1}[\mathcal{R}_Y]$

$\Omega := \llcorner 2_{\mathbf{Qbs}} \lrcorner$ subspace classifier

Example

- Use $[\![\mathbf{Prop}]\!] := \Omega$ for reasoning/axiomatics [Sato et al.'19].
- Differentiation:

$$D_1\mathbb{R} := \{f : \mathbb{R} \to \mathbb{R} | f \text{ differentiable everywhere}\} \hookrightarrow \mathbb{R}^{\mathbb{R}}$$

$$\frac{\mathrm{d}}{\mathrm{d}} : D_1 \to \mathbb{R}^{\mathbb{R}}$$

Borel subspaces

$m : X \hookrightarrow Y$ when $m : X \hookrightarrow Y$ and $[- \in X] \in \mathbb{\Omega}^Y$ factors through $\mathbb{2}^Y \to \mathbb{\Omega}^Y$

Borel subspaces

$$m : X \rightarrowtail Y \qquad \text{when} \qquad \begin{array}{l} m : X \hookrightarrow Y \text{ and } [- \in X] \in \Omega^Y \\ \text{factors through } 2^Y \to \Omega^Y \end{array}$$

Example

- higher-order $\mathbf{Qbs}$-internal σ-algebra:

$$-^{\complement} : \mathcal{B}_X \to \mathcal{B}_X \qquad \bigcap_{n=0}^{\infty} : \mathcal{B}_X^{\mathbb{N}} \to \mathcal{B}_X \qquad -^{-1}[-] : Y^X \times \mathcal{B}_Y \to \mathcal{B}_X$$

Borel subspaces

$$m : X \leftarrowtail\!\!\!\!\to Y \qquad \text{when} \qquad \begin{array}{l} m : X \hookrightarrow Y \text{ and } [- \in X] \in \Omega^Y \\ \text{factors through } 2^Y \to \Omega^Y \end{array}$$

Example

- higher-order $\mathbf{Qbs}$-internal σ-algebra:

$$-^{\complement} : \mathcal{B}_X \to \mathcal{B}_X \qquad \bigcap_{n=0}^{\infty} : \mathcal{B}_X^{\mathbb{N}} \to \mathcal{B}_X \qquad -^{-1}[-] : Y^X \times \mathcal{B}_Y \to \mathcal{B}_X$$

- Non-$\mathbf{Qbs}$-morphisms:

$$\exists : \mathcal{B}_{X \times Y} \to \mathcal{B}_X \qquad [- = \varnothing] : \mathcal{B}_X \to 2 \qquad [- \subseteq -] : \mathcal{B}_X^2 \to 2$$

Borel subspaces

$$m : X \hookleftarrow\!\!\rightarrow Y \qquad \text{when} \qquad \begin{array}{l} m : X \hookrightarrow Y \text{ and } [- \in X] \in \mathbb{\Omega}^Y \\ \text{factors through } 2^Y \to \mathbb{\Omega}^Y \end{array}$$

Example

- higher-order $\mathbf{Qbs}$-internal σ-algebra:

$$-^{\complement} : \mathcal{B}_X \to \mathcal{B}_X \qquad \bigcap_{n=0}^{\infty} : \mathcal{B}_X^{\mathbb{N}} \to \mathcal{B}_X \qquad -^{-1}[-] : Y^X \times \mathcal{B}_Y \to \mathcal{B}_X$$

- Non-$\mathbf{Qbs}$-morphisms:

$$\exists : \mathcal{B}_{X \times Y} \to \mathcal{B}_X \qquad [- = \varnothing] : \mathcal{B}_X \to 2 \qquad [- \subseteq -] : \mathcal{B}_X^2 \to 2$$

- $\mathcal{B}_{\mathcal{B}_X}$: Borel-on-Borel sets [Sabok-Staton-Stein-Wolman'21]

Measures

- Unrestricted Giry:

$$\llcorner G_1 X \lrcorner := \left\{ \mu : \ulcorner \overset{\mathbf{Qbs}}{\llcorner \mathcal{B}_X \lrcorner} \urcorner \to \mathbb{W} \;\middle|\; \text{measure on } \ulcorner \overset{\mathbf{Meas}}{X} \urcorner \right\}$$

$$\mathcal{R}_{G_1 X} := \left\{ k : \mathbb{R} \times \ulcorner \overset{\mathbf{Qbs}}{\llcorner \mathcal{B}_X \lrcorner} \urcorner \to \mathbb{W} \;\middle|\; k \text{ is a kernel} \right\}$$

Analogous to the measure-theoretic Giry
But: careful to evaluate only σ-simple random Borel sets

Measures

- Unrestricted Giry:

$$\llcorner G_1 X \lrcorner := \left\{ \mu : \ulcorner \overset{\mathbf{Qbs}}{\llcorner \mathcal{B}_X \lrcorner} \urcorner \to \mathbb{W} \,\middle|\, \text{measure on } \ulcorner \overset{\mathbf{Meas}}{X} \urcorner \right\}$$

$$\mathcal{R}_{G_1 X} := \left\{ k : \mathbb{R} \times \ulcorner \overset{\mathbf{Qbs}}{\llcorner \mathcal{B}_X \lrcorner} \urcorner \to \mathbb{W} \,\middle|\, k \text{ is a kernel} \right\}$$

Analogous to the measure-theoretic Giry
But: careful to evaluate only σ-simple random Borel sets

- Following can evaluate any random Borel set:

$$\llcorner G_2 X \lrcorner := \left\{ \mu : \mathcal{B}_X \to \mathbb{W} \,\middle|\, \text{measure on } \ulcorner \overset{\mathbf{Meas}}{X} \urcorner \right\}$$

$$\mathcal{R}_{G_2 X} := \{ k : \mathbb{R} \times \mathcal{B}_X \to \mathbb{W} | k \text{ is a kernel} \}$$

$$\llcorner \mathbf{M}X \lrcorner \coloneqq \{v_* \boldsymbol{\lambda}_\Omega | v : \Omega \to X, \Omega \hookrightarrow \mathbb{R} \text{ a } \sigma\text{-finite standard measure space}\}$$

$$\mathcal{R}_{\mathbf{M}X} \coloneqq \{v_* \circ k | k : \mathbb{R} \times \Omega \to X, \Omega \hookrightarrow \mathbb{R} \ldots \text{ ditto } \ldots\}$$

Randomisable measures [Ścibior et al.'18a]

$$\llcorner \mathbf{M}X \lrcorner := \{v_* \boldsymbol{\lambda}_\Omega | v : \Omega \to X, \Omega \hookrightarrow \mathbb{R} \text{ a } \sigma\text{-finite standard measure space}\}$$

$$\mathcal{R}_{\mathbf{M}X} := \{v_* \circ k | k : \mathbb{R} \times \Omega \to X, \Omega \hookrightarrow \mathbb{R} \ldots \text{ ditto } \ldots\}$$

- For standard Borel spaces S,T:
  - $\mathbf{M}S$ are the s-finite measures
  - $(\mathbf{M}S)^T$ are the s-finite kernels

  Fully-definable semantic domain for first-order Monte Carlo models [Staton'17]

Randomisable measures [Ścibior et al.'18a]

$$\llcorner \mathbf{M}X \lrcorner := \{v_* \boldsymbol{\lambda}_\Omega | v : \Omega \to X, \Omega \hookrightarrow \mathbb{R} \text{ a } \sigma\text{-finite standard measure space}\}$$

$$\mathcal{R}_{\mathbf{M}X} := \{v_* \circ k | k : \mathbb{R} \times \Omega \to X, \Omega \hookrightarrow \mathbb{R} \ldots \text{ ditto } \ldots\}$$

- For standard Borel spaces S,T:
  - $\mathbf{M}S$ are the s-finite measures
  - $(\mathbf{M}S)^T$ are the s-finite kernels

  Fully-definable semantic domain for first-order Monte Carlo models [Staton'17]
- Integration is commutative
- Models synthetic measure theory [Kock'12]
- Probabilistic fragment $\mathbf{P}X := \{\mu | \mu X = 1\} \hookrightarrow \mathbf{M}X$ with de Finetti's theorem [Heunen et al'17]
- $\mathbf{P}X$ also models name generation [Sabok et al.'21]

Syntactic spaces and recursive domains

$\mathbf{Qbs}$ is locally presentable
$\implies$ initial algebra semantics for inductive types

Syntactic spaces and recursive domains

$\mathbf{Qbs}$ is locally presentable
$\implies$ initial algebra semantics for inductive types

Example

- Syntactic spaces for operational semantics
- Meta-programming data structures for Monte Carlo inference [Ścibior et al.'18, Lew et al.'20]
- Extends to **recursive** types with domain theory [Vákár-Kammar-Staton'19]
- Opportunity: abstract syntax with binding [Fiore-Plotkin-Turi'99]

Monadic operational semantics [Dal Lago et al.'17, Gavazzo'19, Vákár et al.'19]

$$\frac{k_1(t)\, w_1 \quad k_2(t, w_1)\, w_2 \quad \ldots \quad k_n(t, w_1, \ldots, w_n)\, v}{l(t)\, f(t, w_1, \ldots, w_n, v)}$$

means

$$\begin{aligned} l(t) \coloneqq k_1\,(t) \quad & \gg\!= \lambda w_1. \\ k_2\,(t, w_1) \quad & \gg\!= \lambda w_2.\ \ldots \\ k_n(t, w_1, \ldots, w_{n-1}) & \gg\!= \lambda v. \\ \boldsymbol{\delta}_{f(t, w_1, \ldots, w_n, v)} & \end{aligned}$$

and $l(t) \coloneqq k_1(t) \quad l(t) \coloneqq k_2(t)$ means $l(t) \coloneqq k_1(t) + k_2(t)$

Monadic operational semantics [Dal Lago et al.'17, Gavazzo'19, Vákár et al.'19]

$$\frac{k_1(t)\,w_1 \quad k_2(t,w_1)\,w_2 \quad \ldots \quad k_n(t,w_1,\ldots,w_n)\,v}{l(t)\,f(t,w_1,\ldots,w_n,v)}$$

means

$$\begin{aligned} l(t) \coloneqq k_1(t) & \gg\!\!= \lambda w_1. \\ k_2(t,w_1) & \gg\!\!= \lambda w_2.\ \ldots \\ k_n(t,w_1,\ldots,w_{n-1}) & \gg\!\!= \lambda v. \\ \boldsymbol{\delta}_{f(t,w_1,\ldots,w_n,v)} & \end{aligned}$$

and $l(t) \coloneqq k_1(t) \quad l(t) \coloneqq k_2(t)$ means $l(t) \coloneqq k_1(t) + k_2(t)$

Example

$$\frac{t \Downarrow_n \underline{0} \quad s_1 \Downarrow_n v}{\mathbf{match}\, t\, \mathbf{with}\, \{\underline{0} \to s_1 \mid \_ \to s_2\} \Downarrow_n v}$$

$$\frac{t \Downarrow_n \underline{r} \quad s_2 \Downarrow_n v}{\mathbf{match}\, t\, \mathbf{with}\, \{\underline{0} \to s_1 \mid \_ \to s_2\} \Downarrow_n v}(r \neq 0)$$

Random variable spaces

Refinement types:

$$\frac{B : X \to \mathbb{\Omega}^{Y}}{\prod_{x:X} Bx, \coprod_{x:X} Bx} \qquad \prod_{x:X} Bx \coloneqq \{f : X \to Y | \forall x \in X. f x \in Bx\} \hookrightarrow Y^{X}$$

$$\coprod_{x:X} Bx \coloneqq \{(x, y) \in X \times Y | y \in Bx\} \hookrightarrow X \times Y$$

Example

Lebesgue spaces and modes of convergence:

$$\mathcal{L}^{-}_{-} \in \prod_{\substack{\boldsymbol{\lambda} \in \mathbf{P}\Omega \\ p \in \mathbb{R}}} \left\{ f : \Omega \to [-\infty, \infty] \middle| \int \mathrm{d}\omega \, |f \, \omega|^{p} < \infty \right\}$$

- Observation: $\mathbf{Sbs}$ closed under dependent pairs
- To get dependent types, want “good” universe $\mathbf{Type}$

Monte Carlo inference

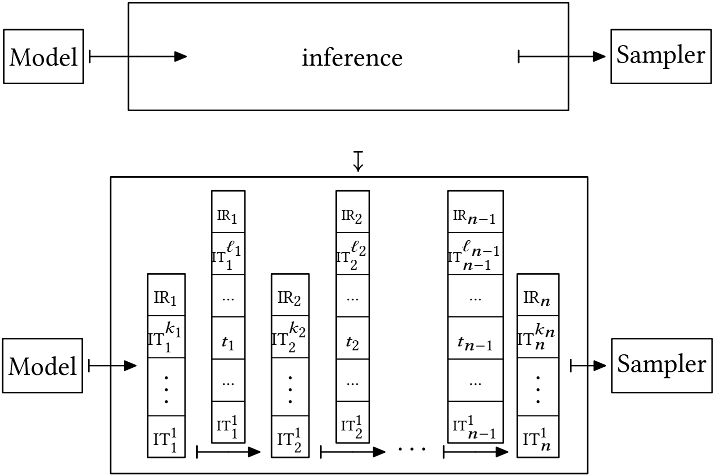
$$\text{model} = \overset{\text{likelihood}}{\rho} \odot \overset{\text{prior}}{\boldsymbol{\lambda}} \coloneqq \lambda\varphi. \int \boldsymbol{\lambda}(\mathrm{d}\omega)\rho(\omega) \cdot \varphi(\omega)$$

Defined programmatically:

$$\mathrm{sample} : \mathbf{M}\mathbb{R} \qquad \text{\#Uniform distribution on } [0, 1]$$

$$\mathrm{score} : \mathbb{R} \to \mathbf{M}\mathbb{1}$$

Modular inference [Ścibior et al.'18a+b, Lew et al.'20]



cf. inference with handlers [Bingham et al.'19] and ongoing

Thank you!

Exact conditioning