

Foundations for type-driven probabilistic modelling

Ohad Kammar
University of Edinburgh

Logic Summer School
Australian National University
4–16 December, 2023
Canberra, ACT, Australia



THE UNIVERSITY of EDINBURGH
informatics

lfcs

Laboratory for Foundations
of Computer Science



supported by:



THE ROYAL
SOCIETY

The
Alan Turing
Institute

Facebook Research NCSC

Computational golden era of:

logic type rich
computation

Statistical
computation

Computational golden era of:

logic & type rich
computation

expressive type systems:

Haskell, OCaml, Idris

mechanised mathematics:

Agda, Coq, Isabelle/HOL, Lean

Verification:

SMT-powered, realistic
systems

Statistical
computation

generative modelling
+

efficient inference:

Monte-Carlo simulation
or gradient-based
optimisation

"AI"

Computational golden era of:

logic type rich
computation

Statistical
computation

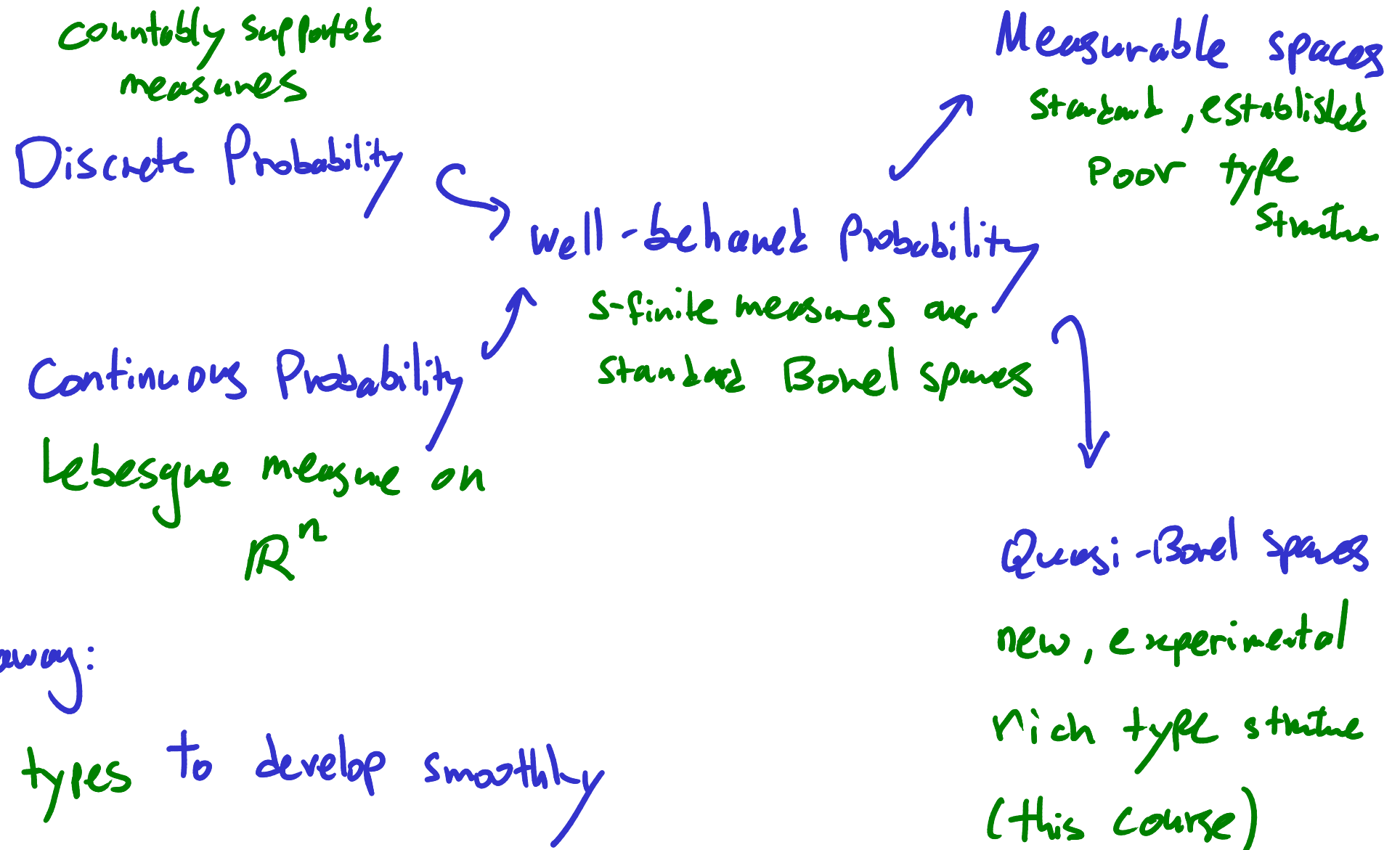
Clear connection to

Foundations:

- Reid's
- John's courses
- Michael's
- Dominik's

- this course

Why foundations?



Takeaway:

use types to develop smoothly

Plan:

- 1) Type-driven probability: discrete case (Mon + Tue(?))
- 2) Borel sets & measurable spaces (Tue)
- 3) Quasi Borel spaces, simple type structure (Wed)
- 4) Dependent type structure & standard Borel spaces (Thu)
- 5) Integration & random variables (Fri)

please ask questions!

smile



Course
web
page

Advertisement

Interested in the ^{Computational} mathematical logical } foundations of Computer Science?

Hedging your bets on a funded PhD offer from AMU?

Check out our PhD Programs:



lfcs

Laboratory for Foundations
of Computer Science



LFCS
PhD
↖



↖ Repentable
AI for
Robotics
CDT

Language of Distribution & Probability

X type (=space) of values/outcomes

DX type of distributions/measures over X

$PX \subseteq DX$ Sub type of probability measures (total measure 1)

BX type of measurable events - subsets of X we wish to measure

W type of weights: $[0, \infty]$

$\mu: DX, E: BX \vdash c_{\mu}[E] : W$

→ type judgement

↳ measure μ assigns to E

Axioms for measures

Empty event: $\emptyset : BX$

Its measure is $0 : W$:

$$\mu : DX \vdash \underset{\mu}{Ce}[\emptyset] = 0 : W$$

Axioms for measures

BX is a Boolean sub-algebra:

$$E : BX \vdash E^c : BX$$

$$E, F : BX \vdash E \cup F, E \cap F : BX$$

$$E, C : BX, \mu : DX \vdash \quad (\text{disjoint additivity})$$

$$C_\mu[E] = C_\mu[E \cap C] + C_\mu[E \cap C^c] : W$$

Axioms for measures

$\omega := (\mathbb{N}, \leq)$ (B, \subseteq) (W, \leq) posets

$$(BX, \subseteq)^\omega := \left\{ (E_n)_{n \in \mathbb{N}} \in (BX)^\mathbb{N} \mid E_0 \subseteq E_1 \subseteq E_2 \subseteq \dots \right\}$$

(BX, \subseteq) and (W, \leq) are ω -chain-closed:

$$E_- : (BX, \subseteq)^\omega \vdash \bigcup_n E_n : BX \qquad a_- : (W, \leq)^\omega \vdash \sup_n a_n : W$$

$$E_- : (BX, \subseteq)^\omega, \mu : DX \vdash \qquad \text{(Scott Continuity)}$$

$$C_{e, \mu} \left[\bigcup_n E_n \right] = \sup_n C_{e, \mu} [E_n] : W$$

Axiom for Probability

$$\text{Cast} : PX \xrightarrow{\varepsilon} DX$$

$$1 : W$$

$$\mu : PX \vdash \underset{\text{Cast } \mu}{Ce}[X] = 1 : W$$

Avoid casting:

$$E : BX, \mu : PX \vdash \underset{\mu}{Pr}[E] := \underset{\text{Cast } \mu}{Ce}[E] : [0,1] \subseteq W$$

Axioms for measures

Integration:

$$\mu : DX, \varphi : \mathbb{W}^X \vdash \int \mu \varphi : \mathbb{W} \quad (\text{Lebesgue integral})$$

Again, avoid casting:

$$\mu : PX, \varphi : \mathbb{W}^X \vdash \underset{\mu}{E}[\varphi] := \int (\text{cast } \mu) \varphi : \mathbb{W} \quad (\text{Expectation})$$

More structure & notation later (...technical...)

Have: language + axioms

Want: model

today: discrete measures

rest of course: discrete + continuous

Discrete model

type X : set

$$DX := \{ \mu: X \rightarrow \mathbb{N} \mid \mu \text{ is countably supported} \}$$

(next slide)

Support

→ Powerset

$$\mu : W^X, S : \mathcal{P}X \vdash S \text{ supports } \mu :=$$

$$\forall x : X. \mu x > 0 \Rightarrow x \in S \quad : \text{Prop}$$

$$\mu : W^X \vdash \text{supp } \mu := \{x \in X \mid \mu x > 0\} : \mathcal{P}X$$

supp μ is the smallest set supporting μ

Discrete model

type X : set

$$DX := \{ \mu: X \rightarrow \mathbb{W} \mid \mu \text{ is countably supported} \}$$

$$:= \{ \mu: X \rightarrow \mathbb{W} \mid \text{supp } \mu \text{ is countable} \}$$

Ex. measures

- X ctbl, Counting measure $\#_X : DX$

$$\#_X := \lambda x:X. 1 \quad (\text{NB: } \text{Supp} \#_X = X \quad \checkmark \text{ ctbl})$$

- Dirac measure:

$$x:X \vdash \delta_x := \lambda x'. \begin{cases} x=x' : 1 \\ \text{o.w.} : 0 \end{cases} : DX$$

$$\text{NB: } \text{Supp} \delta_x = \{x\} \quad \checkmark \text{ ctbl}$$

- Zero measure $\underline{0} := \lambda x. 0 : DX$

$$\text{NB: } \text{Supp} \underline{0} = \emptyset \quad \checkmark \text{ ctbl}$$

Discrete model

type X : set

$DX := \{ \mu: X \rightarrow \mathbb{W} \mid \mu \text{ is countably supported} \}$

$$\mu: DX, E: \mathcal{B}X \vdash \underbrace{Ce[E]}_{\mu} := \sum_{x \in E} \mu x$$

$$:= \sum_{x \in E \cap \text{Supp } \mu} \mu x$$

Lemma: $\mu: DX, S \in \mathcal{P}_{\text{ctbl}} X, S \text{ supports } \mu, E: \mathcal{B}X \vdash$

$$\underbrace{Ce[E]}_{\mu} = \sum_{x \in E \cap S} \mu x$$

Ex:

- $E: B X \vdash \quad C_e[E] = |E| := \begin{cases} E \text{ has } n \text{ elements:} & n \\ E \text{ infinite:} & \infty \end{cases}$
 $\#_x$
- $E: B X, x: X \vdash \quad C_e[E] = \begin{cases} x \in E: & 1 \\ x \notin E: & 0 \end{cases} =: [x \in E] : W$
 δ_x

NB: $E: B X \vdash [- \in E] : X \rightarrow W$

indicator
function

- $E: B X \vdash \quad C_e[E] = 0$
 $\underline{0}$

Validate axioms

$$\mu : DX \vdash C_{e, \mu}[\emptyset] = 0 \quad : \forall W$$

$$E, C : BX, \mu : DX \vdash$$

$$C_{e, \mu}[E] = C_{e, \mu}[E \cap C] + C_{e, \mu}[E \cap C^c] \quad : \forall W$$

$$E_- : (BX, \subseteq)^\omega, \mu : DX \vdash$$

$$C_{e, \mu}[\bigcup_n E_n] = \sup_n C_{e, \mu}[E_n] \quad : \forall W$$

Kernels

κ from Γ to X :

$$\kappa : (DX)^\Gamma$$

kernels are "open/parameterised" measures

Ex: Dirac kernel. $\delta_- : (DX)^X$

Kock Integral

$$\mu : D\Gamma, \kappa : DX^\Gamma \vdash \oint \mu \kappa : DX$$

In discrete model:

$$\oint \mu \kappa := \lambda x : X. \sum_{r \in \Gamma} \mu r \cdot \overbrace{k(r; x)}^{:= k r x}$$

(Weak) disintegration problem:

Input: $\mu: D\Gamma$ $V: DX$

Output: a kernel $k: (DX)^\Gamma$ s.t.

$$\oint \mu k = V$$

Call such k a (weak) disintegration of V

w.r.t. μ .

(non-standard terminology)

Ex disintegration:

$$\underline{n} := \{0, 1, 2, \dots, n-1\}$$

disintegrate $\#_{\underline{z}^{n+1}}$ w.r.t. $\#_{\underline{z}}$

$$k: (D(\underline{z}^{n+1}))^{\underline{z}} \quad k(x; f) := \begin{cases} f(n) = x: & 1 \\ \text{o.w.} & : 0 \end{cases}$$

$$(\oint \#_{\underline{z}} k) f = \sum_{x \in \underline{z}} \overset{1}{\#_{\underline{z}} x} \cdot k(x; f)$$

NB: $\text{Supp}(kx)$
 $\sqrt{c+b}$

$$= k(0; f) + k(1; f) = k(fn; f) = 1 = \#_{\underline{z}^{n+1}}(f)$$

Probability measures

$$PX := \{ \mu : DX \mid C_{\mu}[X] = 1 \} \xhookrightarrow{\subseteq} DX$$

Lemma: $\delta_- : X \rightarrow DX$ and $\oint : D\Gamma \times (DX)^{\Gamma} \rightarrow DX$

lift along the inclusion $\text{cast} : P \xhookrightarrow{\subseteq} D :$

$$\begin{array}{ccc} X & \xrightarrow{\delta_-} & PX \\ & \text{ii} & \downarrow \text{cast} \\ & & DX \\ X & \xrightarrow{\delta_-} & DX \end{array}$$

$$\begin{array}{ccc} D\Gamma \times (PX)^{\Gamma} & \xrightarrow{\oint} & PX \\ \text{cast} \times \text{cast} \downarrow & \text{ii} & \downarrow \text{cast} \\ D\Gamma \times (DX)^{\Gamma} & \xrightarrow{\oint} & DX \end{array}$$

Prop (discrete Giry):

(Michèle Giry '82)

(D, δ_-, \oint) is a monad i.e.

$$\nu: \Gamma, \kappa: (DX)^\Gamma \vdash \oint \delta_{\nu} \kappa = \kappa \nu$$

$$\mu: DX \vdash \oint \mu(\lambda x) \delta_x = \mu \quad : DX$$

$$\mu: D\Gamma, \kappa: (DX)^\Gamma, t: (DY)^X \vdash$$

$$\oint \mu(\lambda x) \left(\oint (\kappa x) t \right) = \oint \left(\oint \mu \kappa \right) (\lambda x) t(x)$$

Corollary: (P, δ_-, \oint) is a monad.

Language of Distribution & Probability

X type (=space) of values/outcomes

DX type of distributions/measures over X

$PX \subseteq DX$ Sub type of probability measures (total measure 1)

BX type of measurable events - subsets of X we wish to measure

W type of weights: $[0, \infty]$

$\mu: DX, E: BX \vdash c_{\mu}[E] : W$

→ type judgement

↳ measure μ assigns to E

Plan:

- 1) Type-driven probability: discrete case (Mon + Tue) ✓
- 2) Borel sets & measurable spaces (Tue)
- 3) Quasi Borel spaces, simple type structure (Wed)
- 4) Dependent type structure & standard Borel spaces (Thu)
- 5) Integration & random variables (Fri)

please ask questions!

snibble



Course
web
page