FSCD
Oxford
Wednesday, Septembre 6th, 2017

Quantitative Semantics for Probabilistic Programing

joint work with **T. Ehrhard** and **M. Pagani**

Christine Tasson
`Christine.Tasson@irif.fr`

Laboratoire IRIF - Université Paris Diderot

A denotational semantics for probabilistic higher-order functional computation,

(based on **quantitative** semantics of Linear Logic)

Discrete setting:

Probabilistic Coherent Spaces are **fully abstract** for a programming language with **natural numbers** as base types suitable to encode discrete probabilistic programs.

Continuous setting:

A **CCC** of measurable spaces and **stable** maps that soundly denotes a programming language with **reals** as base types suitable to encode continuous probabilistic programs.

1. **Discrete Probability**
    - Syntax: **Discrete** Probabilistic PCF
    - Semantics: **Pcoh** (Probabilistic Coherent Spaces)
    - Results: Probabilistic **Adequacy** & **Full Abstraction**
    - Discrete Probabilistic Call By Push Value
2. **Continuous Probability**

General Framework	**Domains Semantics**	**Quantitative Semantics**
Types	Continuous **dcpos** $(X, \leq)$	**Proba.** spaces $(\lvert X \rvert, \mathrm{P}(X) \subseteq (\mathbb{R}^+)^{\lvert X \rvert})$
Programs	**Scott Continuous**	**Analytic** Functions
Probability	Proba. **monad**	**Values** as proba. distr.

Bibliography

1976	Plotkin
1981	Kozen
1989	Plotkin and Jones
1998	Jung and Tix
2013	Goubault Larrecq and Varraca
2013	Mislove

Bibliography

1988	Girard
1994	Blute, Panangaden and Seely
2002	Hasegawa
2004	Girard
2011	Danos and Ehrhard
2014	Ehrhard, Pagani, T.
2016	Ehrahrd, T.

General Framework	**Domains Semantics**	**Quantitative Semantics**
Types	Continuous **dcpos** $(X, \leq)$	**Proba.** spaces $(\lvert X \rvert, \mathrm{P}(X) \subseteq (\mathbb{R}^+)^{\lvert X \rvert})$
Programs	**Scott Continuous**	**Analytic** Functions
Probability	Proba. **monad**	**Values** as proba. distr.

How to interpret a program $M : \mathcal{N} \Rightarrow \mathcal{N}$

Type:

$\mathbb{N}_\perp$ flat domain,
$\mathcal{V}(\mathbb{N}_\perp)$ proba. distr. over $\mathbb{N}_\perp$,

Prog: $[\![M]\!] : \mathbb{N}_\perp \to \mathcal{V}(\mathbb{N}_\perp)$,
$[\![$`let n=x in M`$]\!]$: $\mathcal{V}(\mathbb{N}_\perp) \to \mathcal{V}(\mathbb{N}_\perp)$

$$x \mapsto \left(\sum_n [\![M]\!]_{n,q} x_n \right)_q$$

Type:

$|\textbf{Nat}| = \mathbb{N}$
$\mathrm{P}(\textbf{Nat})$ subproba. dist. over $\mathbb{N}$

Prog: $[\![M]\!] : \mathrm{P}(\textbf{Nat}) \to \mathrm{P}(\textbf{Nat})$

$$x \mapsto \left(\sum_{\mu = [n_1, \ldots, n_k]} [\![M]\!]_{\mu,q} \prod_{i=1}^{k} x_{n_i} \right)_q$$

General Framework	**Domains Semantics**	**Quantitative Semantics**
Types	Continuous **dcpos** $(X, \leq)$	**Proba.** spaces $(\lvert X\rvert, \mathrm{P}(X) \subseteq (\mathbb{R}^+)^{\lvert X\rvert})$
Programs	**Scott Continuous**	**Analytic** Functions
Probability	Proba. **monad**	**Values** as proba. distr.

Problematic in **domain**

Finding a full subcategory of continuous dcpos that is: **Cartesian Closed** and **closed** under the proba. monad $\mathcal{V}$.

Full Abs.: **PCOH/pPCF**

$$\mathrm{Red}(C[M], \underline{n}) \overset{\forall n,\ \forall C[]}{=} \mathrm{Red}(C[N], \underline{n})$$

iff

$$[\![M]\!] = [\![N]\!].$$

1. **Discrete Probability**
    - Syntax: **Discrete** Probabilistic PCF
    - Semantics: **Pcoh** (Probabilistic Coherent Spaces)
    - Results: Probabilistic **Adequacy** & **Full Abstraction**
    - Discrete Probabilistic Call By Push Value

2. **Continuous Probability**

Syntax of PPCF

Syntax of PPCF:

Types: $A, B ::= \mathcal{N} \mid A \to B$

Terms: $M, N, L ::= x \mid \lambda x^A.M \mid (M)N \mid \mathtt{Y}M \mid$
$\mathsf{coin} \mid \underline{n} \mid \mathsf{succ}(M) \mid \mathtt{ifz}(L, M, N)$

Operational Semantics:

$\mathrm{Red}(M, N)$ is the **probability** that M reduces to N in a step.

$$\mathrm{Red}((\lambda x^A.M)N, M[N/x]) = 1, \text{ as } (\lambda x^A.M)N \xrightarrow{1} M[N/x]$$

$$\mathrm{Red}(\mathsf{coin}, \underline{0}) = \mathrm{Red}(\mathsf{coin}, \underline{1}) = \tfrac{1}{2}, \text{ as } \begin{tikzcd} & 0 \\ \mathsf{coin} \arrow[ru, "\frac{1}{2}"] \arrow[rd] & \\ & 1 \end{tikzcd}$$

If $\vdash M : \mathcal{N}$, then $\mathrm{Red}^\infty(M, \_)$ is the discrete distribution over $\mathbb{N}$ of all normal forms computed by M.

1 **Discrete Probability**

- Syntax: **Discrete** Probabilistic PCF
- Semantics: **Pcoh** (Probabilistic Coherent Spaces)
- Results: Probabilistic **Adequacy** & **Full Abstraction**
- Discrete Probabilistic Call By Push Value

2 **Continuous Probability**

Types as **Probabilistic Coherent Spaces**: $(|X|, P(X))$

Proba. Space

$|X|$: the **web**, a (potentially infinite) set of final states

$\mathrm{P}(X)$: a set of vectors $\subseteq (\mathbb{R}^+)^{|X|}$ such that

closure: $\mathbf{P}(\mathbf{X})^{\perp\perp} = \mathbf{P}(\mathbf{X})$ with

$$\forall u, v \in (\mathbb{R}^+)^{|X|},\ \langle u, v\rangle = \sum_{a\in|X|} u_a v_a$$

$$\forall P \subseteq (\mathbb{R}^+)^{|X|},\ P^\perp = \{v \in (\mathbb{R}^+)^{|X|}\ ;\ \forall u \in P,\ \langle u, v\rangle \leq 1\}$$

bounded covering: $\forall a \in |X|$,

$$\exists v \in P(X)\ ;\ v_a \neq 0 \quad \text{and} \quad \exists p > 0,\ ;\ \forall v \in \mathrm{P}(X),\ v_a \leq p.$$

Proposition: Proba. spaces as Domains

$(|X|, \mathrm{P}(X))$ is a **Proba. space iff** $\mathrm{P}(X)$ is bounded covering, **Scott Closed** (downwards-closed and dcpo) and **Convex**.

Types as **Probabilistic Coherent Spaces**: $(|X|, P(X))$

Example: $P(X) \subseteq (\mathbb{R}^+)^{|X|}$

$$|\mathbf{1}| = \{*\} \qquad \mathrm{P}\,(\mathbf{1}) = [0,1]$$

$$|\mathbf{Bool}| = \{t, f\} \qquad \mathrm{P}\,(\mathbf{Bool}) = \{(x_t, x_f)\ ;\ x_t + x_f \leq 1\}$$

$$|\mathbf{Nat}| = \{0, 1, 2, \dots\} \qquad \mathrm{P}\,(\mathbf{Nat}) = \{x \in [0,1]^{\mathbb{N}}\ ;\ \sum_n x_n \leq 1\}$$

$$|\mathbf{Bool} \to \mathbf{1}| = \{[t^n, f^m]\ ;\ n, m \in \mathbb{N}\},$$

$$\mathrm{P}\,(\mathbf{Bool} \to \mathbf{1}) = \{Q \in (\mathbb{R}^+)^{|\mathbf{Bool} \to \mathbf{1}|}\ ;$$

$$\forall x_t + x_f \leq 1, \quad \sum_{n,m=0}^{\infty} Q_{[t^n, f^m]}\, x_t^n x_f^m \leq 1\}$$

Proposition: Proba. spaces as Domains

$(|X|, \mathrm{P}\,(X))$ is a **Proba. space iff** $\mathrm{P}\,(X)$ is bounded covering, **Scott Closed** (downwards-closed and dcpo) and **Convex**.

A model of Linear Logic

Pcoh : Linear Category

Objects: Proba. Spaces

Morphisms: Linear Functions

- **Smcc** $(\mathbf{1}, \otimes, \multimap)$
- biproduct

Call by Name $A \to B = !A \multimap B$

- **Comonad** $(!, \mathrm{der}, \mathrm{dig})$
- **Com. Comonoid** $(!A, \mathbf{1}, \otimes)$

$\mathbf{Pcoh}_!$: Kleisli Category

Objects: Proba. Spaces

Morphisms: Analytic Functions

- **CCC**
- (PCF+coin)

Linear Category

$\mathbf{Pcoh}(X, Y)$

Matrices $Q \in (\mathbb{R}^+)^{|X| \times |Y|}$ such that:

$$\forall x \in \mathrm{P}(X),\ Q \cdot x = \left(\sum_{a \in |X|} Q_{a,b}\, x_a \right)_b \in \mathrm{P}(Y)$$

Example

$\mathbf{Pcoh}(\mathbf{Nat}, \mathbf{Nat})$: Stochastic Matrices $Q \in (\mathbb{R}^+)^{\mathbb{N} \times \mathbb{N}}$.

$$\forall x \in (\mathbb{R}^+)^{\mathbb{N}} \ ; \ \sum_{n \in \mathbb{N}} x_n \leq 1, \ \sum_{m,n \in \mathbb{N}} Q_{m,n} x_n \leq 1$$

Free Commutative Comonoid and Comonad

Exponential

$$|!X| = \mathcal{M}_{\mathsf{fin}}(|X|) \text{ the set of finite multisets}$$

$$\mathrm{P}(!X) = \{x^! \;;\; x \in \mathrm{P}(X)\}^{\perp\perp} \text{ where } x^!_{[a_1,\ldots,a_k]} = \prod_{i=1}^{k} x_{a_i}$$

Example

Let **Bcoin** $= (p, 1-p) \in \mathrm{P}(\mathbf{Bool}) = \{(p,q) \;;\; p+q \leq 1\}$.

$$\mathbf{Bcoin}^!_{[]} = 1, \qquad \mathbf{Bcoin}^!_{[t,t]} = p^2, \qquad \mathbf{Bcoin}^!_{[t,f]} = p(1-p), \;\ldots$$

Theorem (2017: Crubillé - Ehrhard - Pagani - T.)

This exponential computes the free commutative comonoid.

Free Commutative Comonoid and Comonad

Exponential

$$|!X| = \mathcal{M}_{\text{fin}}(|X|) \text{ the set of finite multisets}$$

$$\mathrm{P}(!X) = \{x^! \; ; \; x \in \mathrm{P}(X)\}^{\perp\perp} \text{ where } x^!_{[a_1,\dots,a_k]} = \prod_{i=1}^{k} x_{a_i}$$

Commutative Comonoid | **Comonad**

Cocontr.: $!X \xrightarrow{c^{!X}} !X \otimes !X$

Coweak.: $!X \xrightarrow{w^{!X}} \mathbf{1}$

Comult.: $\mathrm{dig}_{!X} : !!X \to !X$

Counit: $\mathrm{der}_{!X} : !X \to X$

Theorem (2017: Crubillé - Ehrhard - Pagani - T.)

This exponential computes the free commutative comonoid.

Non Linear Category

$\mathbf{Pcoh}_!(X, Y) = \mathbf{Pcoh}(!X, Y)$

Matrices $Q \in (\mathbb{R}^+)^{\mathcal{M}_{\mathrm{fin}}(|X|)\times|Y|}$ such that

$$\forall U \in \mathrm{P}\,(!X),\ \ Q\cdot U = \left(\sum_{m\in\mathcal{M}_{\mathrm{fin}}(|X|)} Q_{m,b}\, U_m\right)_b \in \mathrm{P}\,(Y)$$

Non-Linear Morphisms are **analytic** and **Scott Continuous**.

$$\mathbf{Pcoh}_!(\mathbf{Bool},\mathbf{1}) = \{Q \in (\mathbb{R}^+)^{|\mathbf{Bool}\to\mathbf{1}|}\ \ s.t.\ \ Q_{[t^n,f^m]} \le \frac{(n+m)^{n+m}}{n^n\, m^m}\}$$

```
let rec f x =
 if x then if x then f x
                else ()
      else if x then ()
                else f x
```

denotes

$$\sum_{n,m=0}^{\infty} \frac{(n+m)!}{n!\, m!} x_t^{2n+1} x_f^{2m+1}$$

Non Linear Category

$\mathbf{Pcoh}_!(X, Y) = \mathbf{Pcoh}(!X, Y)$ **Density**

Matrices $Q \in (\mathbb{R}^+)^{\mathcal{M}_{\text{fin}}(|X|)\times|Y|}$ such that if $x^!_m = \prod_{a\in m} x_a^{m(a)}$

$$\forall x \in \mathrm{P}\,(X),\ \ Q(x) = \left(\sum_{m\in\mathcal{M}_{\text{fin}}(|X|)} Q_{m,b}\, {x^!}_m\right)_b \in \mathrm{P}\,(Y)$$

Non-Linear Morphisms are **analytic** and **Scott Continuous**.

$\mathbf{Pcoh}_!(\mathbf{Bool}, \mathbf{1}) = \{Q \in (\mathbb{R}^+)^{|\mathbf{Bool}\to\mathbf{1}|}\ \ s.t.\ \ Q_{[t^n, f^m]} \leq \frac{(n+m)^{n+m}}{n^n\, m^m}\}$

```
let rec f x =
 if x then if x then f x
                 else ()
      else if x then ()
                 else f x
```

denotes

$$\sum_{n,m=0}^{\infty} \frac{(n+m)!}{n!\,m!} x_t^{2n+1} x_f^{2m+1}$$

Non Linear Category

$\mathbf{Pcoh}_!(X, Y) = \mathbf{Pcoh}(!X, Y)$ **Density**

Matrices $Q \in (\mathbb{R}^+)^{\mathcal{M}_{\text{fin}}(|X|)\times|Y|}$ such that if $x^!_m = \prod_{a\in m} x_a^{m(a)}$

$$\forall x \in \mathrm{P}(X),\ Q(x) = \left(\sum_{m\in\mathcal{M}_{\text{fin}}(|X|)} Q_{m,b}\, {x^!}_m\right)_b \in \mathrm{P}(Y)$$

Non-Linear Morphisms are **analytic** and **Scott Continuous**.

$$\mathbf{Pcoh}_!(\mathbf{Bool},\mathbf{1}) = \{Q \in (\mathbb{R}^+)^{|\mathbf{Bool}\to\mathbf{1}|}\ s.t.\ Q_{[t^n,f^m]} \le \frac{(n+m)^{n+m}}{n^n\, m^m}\}$$

```
let rec f x =
 if x then if x then f x
                else ()
      else if x then ()
                else f x
```

pb of DEFINABILITY

$$\sum_{n,m=0}^{\infty} \frac{(n+m)!}{n!\,m!} x_t^{2n+1} x_f^{2m+1}$$

1. **Discrete Probability**
   - Syntax: **Discrete** Probabilistic PCF
   - Semantics: **Pcoh** (Probabilistic Coherent Spaces)
   - Results: Probabilistic **Adequacy** & **Full Abstraction**
   - Discrete Probabilistic Call By Push Value
2. **Continuous Probability**

Probabilistic Full Abstraction

Theorem (2014: Ehrhard - Pagani - T.)

Pcoh

$[\![M]\!] = [\![N]\!]$

Adequacy

$\Longrightarrow$

$\Longleftarrow$

Full Abstraction

pPCF

$M \simeq_o N$

$\mathrm{Red}^\infty(C[M], n) \overset{\forall C[\,]\forall n}{=} \mathrm{Red}^\infty(C[N], n)$

Adequacy Lemma (2011: Danos - Ehrhard):

If $\vdash M : \mathcal{N}$, then $\forall n \in \mathbb{N}, [\![M]\!]_n = \mathrm{Red}^\infty(M, n)$.

Adequacy proof:

If $[\![M]\!] = [\![N]\!]$ then, $\mathrm{Red}^\infty((C)M, \underline{n}) = \mathrm{Red}^\infty((C)N, \underline{n})$

1. Apply Adequacy Lemma : $\mathrm{Red}^\infty((C)M, \underline{n}) = [\![(C)M]\!]$.
2. Apply Compositionality:

$$[\![(C)M]\!] = \sum_\mu [\![C]\!]_\mu \prod_{\alpha \in \mu} [\![M]\!]_\alpha^{\mu(\alpha)} = \sum_\mu [\![C]\!]_\mu \prod_{\alpha \in \mu} [\![N]\!]_\alpha^{\mu(\alpha)} = [\![(C)N]\!]$$

Probabilistic Full Abstraction

Theorem (2014: Ehrhard - Pagani - T.)

Pcoh

pPCF

$[\![M]\!] = [\![N]\!]$

Adequacy

$\Longrightarrow$

$\Longleftarrow$

Full Abstraction

$M \simeq_o N$

$\mathrm{Red}^\infty(C[M], n) \overset{\forall C[\,]\forall n}{=} \mathrm{Red}^\infty(C[N], n)$

Adequacy Lemma (2011: Danos - Ehrhard):

If $\vdash M : \mathcal{N}$, then $\forall n \in \mathbb{N}, [\![M]\!]_n = \mathrm{Red}^\infty(M, n)$.

Full Abstraction proof:

- Find **testing terms** that depend only on points of the web.
- Use regularity of **analytic functions**.

1. **Discrete Probability**
   - Syntax: **Discrete** Probabilistic PCF
   - Semantics: **Pcoh** (Probabilistic Coherent Spaces)
   - Results: Probabilistic **Adequacy** & **Full Abstraction**
   - Discrete Probabilistic Call By Push Value

2. **Continuous Probability**

How to encode a LasVegas Algorithm?

Input: A $\underline{0}/\underline{1}$ array of length $n \geq 2$ s.t. $\frac{1}{2}$ cells are $\underline{0}$.

0	1	2	3	4	5
$\underline{0}$	$\underline{1}$	$\underline{0}$	$\underline{1}$	$\underline{1}$	$\underline{0}$

$$f : \begin{array}{ccc} 0, 2, 5 & \mapsto & \underline{0} \\ 1, 3, 4 & \mapsto & \underline{1} \end{array}$$

Output: Find the index of a cell containing $\underline{0}$.

Caml:

```
let rec LasVegas (f: nat -> nat) (n:nat) =
                let k = random n in
                    if (f k = 0) then k
                    else LasVegas f n
```

pPCF: CBN

$$\mathbf{Y}\Big(\lambda \mathtt{LasVegas}^{(\mathtt{nat}\Rightarrow\mathtt{nat})\Rightarrow\mathtt{nat}\Rightarrow\mathtt{nat}}\ \lambda \mathtt{f}^{\mathtt{nat}\Rightarrow\mathtt{nat}} \lambda \mathtt{n}^{\mathtt{nat}}$$
$$(\lambda \mathtt{k}^{\mathtt{nat}}\ \mathtt{ifz}\ \mathtt{f}\ \mathtt{k}\ \mathtt{then}\ \mathtt{k}$$
$$\mathtt{else}\ \mathtt{LasVegas}\ \mathtt{f}\ \mathtt{n})\ (\mathtt{rand}\ \mathtt{n})$$

How to encode a LasVegas Algorithm?

Input: A $\underline{0}/\underline{1}$ array of length $n \geq 2$ s.t. $\frac{1}{2}$ cells are $\underline{0}$.

0	1	2	3	4	5
$\underline{0}$	$\underline{1}$	$\underline{0}$	$\underline{1}$	$\underline{1}$	$\underline{0}$

$$f: \begin{array}{lcl} 0,2,5 & \mapsto & \underline{0} \\ 1,3,4 & \mapsto & \underline{1} \end{array}$$

Output: Find the index of a cell containing $\underline{0}$.

Caml:
`let in`
CBV

```
let rec LasVegas (f: nat -> nat) (n:nat) =
                let k = random n in
                    if (f k = 0) then k
                    else LasVegas f n
```

pPCF:
CBN

$$\mathbf{Y}\Big(\lambda \mathtt{LasVegas}^{(\mathtt{nat}\Rightarrow\mathtt{nat})\Rightarrow\mathtt{nat}\Rightarrow\mathtt{nat}}\ \lambda \mathtt{f}^{\mathtt{nat}\Rightarrow\mathtt{nat}}\lambda \mathtt{n}^{\mathtt{nat}}$$
$$(\lambda \mathtt{k}^{\mathtt{nat}}\ \mathtt{ifz\ f\ k\ then\ k}$$
$$\mathtt{else\ LasVegas\ f\ n})\ (\mathtt{rand\ n})\Big)$$

Semantics gives the answer

Storage Operator

```
let k = rand n in if k = 0 then k else 42
```

Integer in Pcoh: $[\![\mathcal{N}]\!] = \mathbf{Nat} = (\mathbb{N}, \mathrm{P}(\mathbf{Nat}) = \{(\lambda_n) \mid \sum_n \lambda_n \leq 1\})$

Equipped with a structure of comonoid in the *linear* Pcoh:

- Cocontraction: $c^{\mathcal{N}} : \mathcal{N} \to \mathcal{N} \otimes \mathcal{N}$
- Coweakening: $w^{\mathcal{N}} : \mathcal{N} \to \mathbf{1}$

Bibliography

1990 Krivine, Opérateurs de mise en mémoire et Traduction.

1999 Levy, Call by Push Value, a subsuming paradigm.

2000 Nour, On Storage operator.

2016 Curien, Fiore, Munch-Maccagnoni, A Theory of Effects and Resources .

What sem. object to encode Storage Operator.

The Eilenberg Moore Category: $\mathbf{Pcoh}^!$

Coalgebras $P = (\underline{P}, h_P)$ with $\underline{P} \in \mathbf{Pcoh}$ and $h_P \in \mathbf{Pcoh}(\underline{P}, !\underline{P})$:

$$\begin{tikzcd}
\underline{P} \arrow[r, "h_P"] \arrow[dr, "\mathrm{Id}"'] & !\underline{P} \arrow[d, "\mathrm{der}_{\underline{P}}"] \\
 & \underline{P}
\end{tikzcd}
\qquad
\begin{tikzcd}
\underline{P} \arrow[r, "h_P"] \arrow[d, "h_P"'] & !\underline{P} \arrow[d, "\mathrm{dig}_{\underline{P}}"] \\
!\underline{P} \arrow[r, "!h_P"] & !!\underline{P}
\end{tikzcd}$$

Coalgebras have a comonoid structure: values can be **stored**.

Types interpreted as coalgebras:

$!X$ by def. of the exp. | $\otimes$, $\oplus$ and Y preserve coalgebras.

Example

Stream: $\mathsf{S}_\varphi = \varphi \otimes !\mathsf{S}_\varphi$ | **List:** $\lambda_0 = \mathbf{1} \oplus (\varphi \otimes \lambda_0)$

Probabilistic Call By Push Value

Types:

(Value) $A ::= U\underline{B} \mid A_1 \oplus A_2 \mid \mathbf{1} \mid A_1 \otimes A_2 \mid \alpha \mid \mathrm{Fix}\,\alpha \cdot A$

Example of natural numbers: $\mathcal{N} ::= \mathrm{Fix}\,\alpha \cdot \mathbf{1} \oplus \alpha$

(Computation) $\underline{B} ::= FA \mid A \multimap \underline{B}$

Terms:

(Value) $V ::= x \mid \mathrm{thunk}(M) \mid \mathrm{in}_i V \mid () \mid (V, W)$

(Computation) $M ::= \mathrm{return}(V) \mid \mathrm{force}(M)$

$\mid \lambda x^A\, M \mid \langle M \rangle V \mid \Upsilon M$

$\mid \mathrm{coin} \mid \mathrm{case}(M, x_1 \cdot N_1, x_2 \cdot N_2)$

$\mid \underline{n} \mid \mathrm{succ}(V) \mid \mathtt{let}(x, V, M) \mid \mathtt{ifz}(V, M, N)$

Probabilistic Call By Push Value

Types: $!\underline{B}$

(Value) $A \ ::= \ U\underline{B} \mid A_1 \oplus A_2 \mid \mathbf{1} \mid A_1 \otimes A_2 \mid \alpha \mid \mathrm{Fix}\,\alpha \cdot A$

Example of natural numbers: $\mathcal{N} ::= \mathrm{Fix}\,\alpha \cdot \mathbf{1} \oplus \alpha$

(Computation) $\underline{B} \ ::= \ FA \mid A \multimap \underline{B}$

Terms:

(Value) $V \ ::= \ x \mid \mathrm{thunk}(M) \mid \mathrm{in}_i V \mid () \mid (V, W)$

(Computation) $M \ ::= \ \mathrm{return}(V) \mid \mathrm{force}(M)$

$\mid \lambda x^A\, M \mid \langle M \rangle V \mid \Upsilon M$

$\mid \mathrm{coin} \mid \mathrm{case}(M, x_1 \cdot N_1, x_2 \cdot N_2)$

$\mid \underline{n} \mid \mathrm{succ}(V) \mid \mathtt{let}(x, V, M) \mid \mathtt{ifz}(V, M, N)$

Probabilistic Call By Push Value

Types: $!\underline{B}$

(Value) $A \ ::= \ U\underline{B} \ | \ A_1 \oplus A_2 \ | \ \mathbf{1} \ | \ A_1 \otimes A_2 \ | \ \alpha \ | \ \mathrm{Fix}\,\alpha \cdot A$

Example of natural numbers: $\mathcal{N} ::= \mathrm{Fix}\,\alpha \cdot \mathbf{1} \oplus \alpha$

(Computation) $\underline{B} \ ::= \ FA \ | \ A \multimap \underline{B}$ Forget: A

Terms:

(Value) $V \ ::= \ x \ | \ \mathrm{thunk}(M) \ | \ \mathrm{in}_i V \ | \ () \ | \ (V, W)$

(Computation) $M \ ::= \ \mathrm{return}(V) \ | \ \mathrm{force}(M)$

$| \ \lambda x^A\, M \ | \ \langle M \rangle V \ | \ \Upsilon M$

$| \ \mathrm{coin} \ | \ \mathrm{case}(M, x_1 \cdot N_1, x_2 \cdot N_2)$

$| \ \underline{n} \ | \ \mathrm{succ}(V) \ | \ \mathtt{let}(x, V, M) \ | \ \mathtt{ifz}(V, M, N)$

Probabilistic Call By Push Value

Types: $!\underline{B}$

(Value) $A ::= U\underline{B} \mid A_1 \oplus A_2 \mid \mathbf{1} \mid A_1 \otimes A_2 \mid \alpha \mid \mathrm{Fix}\,\alpha \cdot A$

Example of natural numbers: $\mathcal{N} ::= \mathrm{Fix}\,\alpha \cdot \mathbf{1} \oplus \alpha$

(Computation) $\underline{B} ::= FA \mid A \multimap \underline{B}$ Forget: A

Terms: $M^!$

(Value) $V ::= x \mid \mathrm{thunk}(M) \mid \mathrm{in}_i V \mid () \mid (V, W)$

(Computation) $M ::= \mathrm{return}(V) \mid \mathrm{force}(M)$

$\mid \lambda x^A M \mid \langle M \rangle V \mid \Upsilon M$

$\mid \mathrm{coin} \mid \mathrm{case}(M, x_1 \cdot N_1, x_2 \cdot N_2)$

$\mid \underline{n} \mid \mathrm{succ}(V) \mid \mathtt{let}(x, V, M) \mid \mathtt{ifz}(V, M, N)$

Probabilistic Call By Push Value

Types: $!\underline{B}$

(Value) $A \ ::= \ U\underline{B} \mid A_1 \oplus A_2 \mid \mathbf{1} \mid A_1 \otimes A_2 \mid \alpha \mid \mathsf{Fix}\,\alpha \cdot A$

Example of natural numbers: $\mathcal{N} ::= \mathsf{Fix}\,\alpha \cdot \mathbf{1} \oplus \alpha$

(Computation) $\underline{B} \ ::= \ FA \mid A \multimap \underline{B}$ Forget: A

Terms: $M^!$ $\mathsf{der}(M)$

(Value) $V \ ::= \ x \mid \mathsf{thunk}(M) \mid \mathsf{in}_i V \mid () \mid (V, W)$

(Computation) $M \ ::= \ \mathsf{return}(V) \mid \mathsf{force}(M)$

$\mid \lambda x^A\, M \mid \langle M \rangle V \mid \Upsilon M$

$\mid \mathsf{coin} \mid \mathsf{case}(M, x_1 \cdot N_1, x_2 \cdot N_2)$

$\mid \underline{n} \mid \mathsf{succ}(V) \mid \mathtt{let}(x, V, M) \mid \mathtt{ifz}(V, M, N)$

The Eilenberg Moore categoy and the Linear Category

Dense coalgebra

$P = (\underline{P}, h_P)$ such that coalgebraic points characterize morphisms:
$\forall Y \in \mathbf{Pcoh}$ and $\forall t, t' \in \mathbf{Pcoh}(\underline{P}, Y)$,
if $\forall v \in \mathbf{Pcoh}^!(1, P),\ t\,v = t'\,v$, then $\forall u \in \mathbf{Pcoh}(1, \underline{P}),\ t\,u = t'\,u$.

Already known for $!X$ as: if $\forall x \in \mathbf{Pcoh}(1, X),\ t\,x^! = t'\,x^!$ then $t = t'$.

The Eilenberg Moore category $\mathbf{Pcoh}^!$

Value Types are interpreted as **dense** coalgebras

Values are morphisms of coalgebras

The Linear category **Pcoh**

Computation Types are interpreted in **Pcoh**

Computations are linear morphisms in **Pcoh**

Probabilistic Full Abstraction

Theorem (2016: Ehrhard - T.)

Pcoh

$[\![M]\!] = [\![N]\!]$

Adequacy

$\Longrightarrow$

$\Longleftarrow$

Full Abstraction

pCBPV

$M \simeq_o N$

$\mathrm{Red}(C[M], ()) \overset{\forall C[\,]}{=} \mathrm{Red}(C[N], ())$

Adequacy Lemma Proof:

- Handle **values** separately
- Logical relations: **fixpoint** of types (hidden step indexing, biorthogonality closure, fixpoints of pairs of logical relations)
- **Density:** Morphisms on positive types are characterized by their action on coalgebraic points.

Probabilistic Full Abstraction

Theorem (2016: Ehrhard - T.)

Pcoh

pCBPV

$[\![M]\!] = [\![N]\!]$ $\overset{\text{Adequacy}}{\Longrightarrow}$ $\underset{\text{Full Abstraction}}{\Longleftarrow}$ $M \simeq_o N$

$\mathrm{Red}(C[M], ()) \overset{\forall C[\,]}{=} \mathrm{Red}(C[N], ())$

Full Abstraction Proof:

1. By **contradiction**: $\exists \alpha \in |\sigma|,\ [\![M]\!]_\alpha \neq [\![N]\!]_\alpha$
2. Find **testing context**: T_α such that $[\![\langle T_\alpha\rangle M^!]\!] \neq [\![\langle T_\alpha\rangle N^!]\!]$ (context only depends on α)
3. Prove **definability**: $T_\alpha \in$ **pCBPV** using coin and regularity of analytic functions and density.
4. Apply **Adequacy Lemma**: $\mathrm{Red}(\langle T_\alpha\rangle M^! \xrightarrow{*} ()) \neq \mathrm{Red}(\langle T_\alpha\rangle N^! \xrightarrow{*} ())$.

A denotational semantics for probabilistic higher-order functional computation,

(based on **quantitative** semantics of Linear Logic)

Discrete setting:

Probabilistic Coherent Spaces are **fully abstract** for a programming language with **natural numbers** as base types suitable to encode discrete probabilistic programs.

Continuous setting:

A **CCC** of measurable spaces and **stable** maps that soundly denotes a programming language with **reals** as base types suitable to encode continuous probabilistic programs.

1. Discrete Probability

2. **Continuous Probability**
   - Syntax: **Real** Probabilistic PCF
   - Semantics: $\mathbf{Cstab}_{\mathrm{m}}$ (Cones and Stable measurable functions)
   - Results: **Adequacy**

From Discrete to Continuous syntax

PPCF

Types: $A, B ::= \mathcal{N} \mid A \to B$

Terms: $M, N, L ::=$

$x \mid \lambda x^{A}.M \mid (M)N \mid \mathtt{Y}M \mid$

$\underline{n} \mid \mathsf{succ}(M) \mid$

$\mathtt{ifz}(L, M, N) \mid$

$\mathsf{coin} \mid \mathtt{let}(x, M, N)$

Operational Semantics:

$\mathrm{Red}(\mathsf{coin}, \underline{0}) = \mathrm{Red}(\mathsf{coin}, \underline{1}) = \frac{1}{2}$

If $\vdash M : \mathcal{N}$, $\mathrm{Red}^{\infty}(M, \_)$ is the discrete distribution over $\mathbb{N}$ computed by M.

From Discrete to Continuous syntax

PPCF

Types: $A, B ::= \mathcal{N} \mid A \to B$

Terms: $M, N, L ::=$
$x \mid \lambda x^A.M \mid (M)N \mid \mathtt{Y}M \mid$
$\underline{n} \mid \mathsf{succ}(M) \mid$
$\mathtt{ifz}(L, M, N) \mid$
$\mathsf{coin} \mid \mathtt{let}(x, M, N)$

Operational Semantics:

$\mathrm{Red}(\mathsf{coin}, \underline{0}) = \mathrm{Red}(\mathsf{coin}, \underline{1}) = \frac{1}{2}$

If $\vdash M : \mathcal{N}$, $\mathrm{Red}^\infty(M, \_)$ is the discrete distribution over $\mathbb{N}$ computed by M.

Real PPCF

Types: $A, B ::= \mathcal{R} \mid A \to B$

Terms: $M, N, L ::=$
$x \mid \lambda x^A.M \mid (M)N \mid \mathtt{Y}M \mid$
$\underline{r} \mid \underline{f}(M_1, \ldots, M_n) \mid$
$\mathtt{ifz}(L, M, N) \mid$
$\mathtt{sample} \mid \mathtt{let}(x, M, N)$

Operational Semantics:

$\mathrm{Red}(\mathtt{sample}, U) = \lambda_{[0,1]}(U)$

If $\vdash M : \mathcal{R}$, $\mathrm{Red}^\infty(M, \_)$ is the continuous distribution over $\mathbb{R}$ computed by M.

Operational Semantics

The probability to observe U after at most one reduction step applied to M is $\mathrm{Red}(\ M\ ,\ U\ )$

$\mathrm{Red} : \Lambda^{\Gamma\vdash A} \times \Sigma_{\Lambda^{\Gamma\vdash A}} \to \mathbb{R}^+$ is a **Kernel**, i.e:

- for all $M \in \Lambda^{\Gamma\vdash A}$, $\mathrm{Red}(M, \_)$ is a measure;
- for all $U \in \Sigma_{\Lambda^{\Gamma\vdash A}}$, $\mathrm{Red}(\_, U)$ is a measurable function.

$\mathrm{Red}^{\infty}(M, U)$ is the probability to observe U after any steps.

Operational Semantics

The probability to observe U after at most one reduction step applied to M is $\mathrm{Red}(\ M\ ,\ U\ )$

$\Lambda^{\Gamma\vdash A}$: the set of terms M
s.t. $\Gamma \vdash M : A$.

$\mathrm{Red} : \Lambda^{\Gamma\vdash A} \times \Sigma_{\Lambda^{\Gamma\vdash A}} \to \mathbb{R}^+$ is a **Kernel**, i.e:

- for all $M \in \Lambda^{\Gamma\vdash A}$, $\mathrm{Red}(M, \_)$ is a measure;
- for all $U \in \Sigma_{\Lambda^{\Gamma\vdash A}}$, $\mathrm{Red}(\_, U)$ is a measurable function.

$\mathrm{Red}^{\infty}(M, U)$ is the probability to observe U after any steps.

Operational Semantics

The probability to observe U after at most one reduction step applied to M is $\mathrm{Red}(\ M\ ,\ U\ )$

$\Lambda^{\Gamma\vdash A}$: the set of terms M s.t. $\Gamma \vdash M : A$.

$\Sigma_{\Lambda^{\Gamma\vdash A}}$, i.e. U is measurable: $\forall n, \forall S,\ \{\vec{r} \mid S\vec{r} \in U\}$ meas. in $\mathbb{R}^n$

$\mathrm{Red} : \Lambda^{\Gamma\vdash A} \times \Sigma_{\Lambda^{\Gamma\vdash A}} \to \mathbb{R}^+$ is a **Kernel**, i.e:

- for all $M \in \Lambda^{\Gamma\vdash A}$, $\mathrm{Red}(M, \_)$ is a measure;
- for all $U \in \Sigma_{\Lambda^{\Gamma\vdash A}}$, $\mathrm{Red}(\_, U)$ is a measurable function.

$\mathrm{Red}^\infty(M, U)$ is the probability to observe U after any steps.

Operational Semantics

The probability to observe U after at most one reduction step applied to M is $\mathrm{Red}(\ M\ ,\ U\ )$

$\Lambda^{\Gamma\vdash A}$: the set of terms M s.t. $\Gamma \vdash M : A$.

$\Sigma_{\Lambda^{\Gamma\vdash A}}$, i.e. U is measurable: $\forall n, \forall S,\ \{\vec{r} \mid S\vec{r} \in U\}$ meas. in $\mathbb{R}^n$

$\mathrm{Red} : \Lambda^{\Gamma\vdash A} \times \Sigma_{\Lambda^{\Gamma\vdash A}} \to \mathbb{R}^+$ is a **Kernel**, i.e:

- for all $M \in \Lambda^{\Gamma\vdash A}$, $\mathrm{Red}(M, \_)$ is a measure;
- for all $U \in \Sigma_{\Lambda^{\Gamma\vdash A}}$, $\mathrm{Red}(\_, U)$ is a measurable function.

Measurable sets and kernels constitute the category Kern.

$\mathrm{Red}^{\infty}(M, U)$ is the probability to observe U after any steps.

Operational Semantics

The probability to observe U after at most one reduction step applied to M is $\mathrm{Red}(\ M\ ,\ U\ )$

$\Lambda^{\Gamma\vdash A}$: the set of terms M s.t. $\Gamma \vdash M : A$.

$\Sigma_{\Lambda^{\Gamma\vdash A}}$, i.e. U is measurable: $\forall n, \forall S,\ \{\vec{r} \mid S\vec{r} \in U\}$ meas. in $\mathbb{R}^n$

$\mathrm{Red} : \Lambda^{\Gamma\vdash A} \times \Sigma_{\Lambda^{\Gamma\vdash A}} \to \mathbb{R}^+$ is a **Kernel**, i.e:

- for all $M \in \Lambda^{\Gamma\vdash A}$, $\mathrm{Red}(M, \_)$ is a measure;
- for all $U \in \Sigma_{\Lambda^{\Gamma\vdash A}}$, $\mathrm{Red}(\_, U)$ is a measurable function.

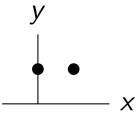
Measurable sets and kernels constitute the category Kern.

$\mathrm{Red}^{\infty}(M, U)$ is the probability to observe U after any steps.

It is computed by composition and lub.

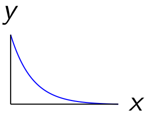
Examples: Distributions

The Bernoulli distribution takes the value 1 with probability p and the value 0 with probability $1-p$.



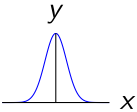
$\texttt{bernoulli}\ p\ ::=\ \texttt{let}(x, \texttt{sample}, x \leq p)$ tests if `sample` draws a value within $[0, p]$.

The exponential distribution is specified by its density e^{-x}.



$\texttt{exp} : \mathcal{R} ::= \texttt{let}(x, \texttt{sample}, -\log(x))$
by the inversion sampling method.

The standard normal distribution defined by its density $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.



$\texttt{normal} ::=$
$\texttt{let}(x, \texttt{sample}, \texttt{let}(y, \texttt{sample}, \sqrt{-2\log(x)}\cos(2\pi y)))$
by the Box Muller method.

Examples:

Conditioning: If $U \subseteq \mathbb{R}$ measurable, then $\texttt{observe}(U)$ of type $\mathcal{R} \to \mathcal{R}$, taking a term M and returning the renormalization of the distribution of M on the only samples that satisfy U:
$\texttt{observe}(U) = \lambda m.\texttt{Y}(\lambda y.\texttt{let}(x, m, \texttt{if}(x \in U, x, y)))$
conditioning by rejection sampling.

Monte Carlo Simulation, Metropolis Hasting,...

1. Discrete Probability

2. **Continuous Probability**
   - Syntax: **Real** Probabilistic PCF
   - Semantics: $\mathbf{Cstab}_{\mathsf{m}}$ (Cones and Stable measurable functions)
   - Results: **Adequacy**

Semantical context

1981, Kozen Memory as measurable space and programs as kernels representing the transformation of the memory.

What is a measurable subset for function space ?

1999, Panangaden

Meas, the category of measurable sets and functions

Kern, the category of measurable sets and kernels

They are **cartesian** but **not closed**.

2017, Heunen, Kammar, Staton, Yang **Quasi-borel spaces**

A **CCC** based on **Meas** embedded into presheaves.

How to interpret recursive types ?

2017, Keimel and Plotkin **Kegelspitzen**

A **CCC** of dcpos equipped with a convex structure (basic operations being scott continous) with scott continuous functions

How to restrict to measurable functions ?

Semantical context

If $\vdash M : \mathcal{N}$, then $[\![M]\!]$ is a discrete distribution over $\mathbb{N}$	If $\vdash M : \mathcal{R}$, then $[\![M]\!]$ is a continuous measure over $\mathbb{R}$

- $[\![\mathcal{R}]\!]$ as Meas($\mathbb{R}$) the set of measures over the measurable space $\mathbb{R}$.
- Fixpoint of terms.

Cstab$_{\mathrm{m}}$ is a **CCC** based on Selinger's **cones** (dcpos with the order induced by addition and a convex structure).

Objects are cones and measurable spaces

Morphisms are stable and measurable functions

Pcoh is a subcategory of **Cstab$_{\mathrm{m}}$** which is a subcategory of Kegelspitzen.

An elegant model in 3 steps

Our purpose is to be able to interpret $\mathcal{R}$ as the set of bounded measures.

1. **Complete cones** (convex dcpos with the order induced by addition) with Scott continuous functions
   However, the category is cartesian but not closed.
2. Complete cones and **Stable functions** (∞-non-decreasing functions) is a CCC.
   However, not every stable function is measurable.
3. **Measurable Cones** (complete cones with **measurable tests**). Measurable paths pass measurable tests and Measurable functions preserve measurable paths.
   $\mathbf{Cstab_m}$ is a CCC with measurability included !

Step 1: Complete Cones

A Cone P is analogous to a real normed vector space, except that scalars are $\mathbb{R}^+$ and the norm $\|\_\|_P : P \to \mathbb{R}^+$ satisfies:

$$x+y=0 \Rightarrow x,y=0, \qquad \|x+x'\|_P \leq \|x\|_P + \|x'\|_P, \qquad \|\alpha x\|_P = \alpha\|x\|_P$$

$$x+y=x+y' \Rightarrow y=y', \qquad \|x\|_P = 0 \Rightarrow x=0, \qquad \|x\|_P \leq \|x+x'\|_P$$

The Unit Ball is the set $\mathcal{B}P = \{x \in P \mid \|x\|_P \leq 1\}$.

Order $x \leq_P x'$ if there is a $y \in P$ such that $x' = x + y$. This unique y is denoted as $y = x' - x$.

A Complete Cone is s.t. any non-decreasing $(x_n)_{n\in\mathbb{N}}$ of $\mathcal{B}P$ has a lub and $\|\sup_{n\in\mathbb{N}} x_n\|_P = \sup_{n\in\mathbb{N}} \|x_n\|_P$.

Example of Complete Cones

- Meas(X) with X a measurable space.
- $\widehat{\mathcal{X}} = \{u \in (\mathbb{R}^+)^{|\mathcal{X}|} \mid \exists \varepsilon > 0\ \varepsilon u \in \mathrm{P}\mathcal{X}\}$ if $\mathcal{X} \in$ **Pcoh**.

Step 2: Stable functions

The category of **complete cones** and **Scott-continuous** functions is not cartesian closed as *currying* fails to be *non-decreasing*.

A function $f : \mathcal{B}P \to Q$ is **n-non-decreasing function** if:

$n = 0$ and f is non-decreasing

$n > 0$ and $\forall u \in \mathcal{B}P$, $\Delta f(x; u) = f(x + u) - f(x)$ is $(n-1)$-non-decreasing in x.

A function is **stable** if it is Scott-continuous and ∞-non-decreasing, i.e. n-non-decreasing for all $n \in \mathbb{N}$.

Complete cones and **stable** functions constitute a **CCC**.

Weak Parallel Or

$\mathrm{wpor} : [0,1] \times [0,1] \to [0,1]$ given as $\mathrm{wpor}(s,t) = s + t - st$ is Scott-continuous, but not Stable. Its currying is not Scott-continuous.

Step 3: The Measurability Problem

Type $\mathcal{R}$ is interpreted as $[\![\mathcal{R}]\!] = \mathrm{Meas}(\mathbb{R})$,
Closed term $\vdash M : \mathcal{R}$ as a measure μ and
Term $x : \mathcal{R} \vdash N : \mathcal{R}$ as a stable $f : \mathrm{Meas}(\mathbb{R}) \to \mathrm{Meas}(\mathbb{R})$.

Operational semantics

$$\forall r,\ \text{s.t. } M \to r,\ \mathtt{let}(x, M, N) \to N\{r/x\}$$

By **Soundness**

$$[\![\mathtt{let}(x, M, N)]\!] = \int_{\mathbb{R}} (\ f\ \circ\ \delta\ )(r)\ \ \mu\ (dr)$$

Step 3: The Measurability Problem

Type $\mathcal{R}$ is interpreted as $[\![\mathcal{R}]\!] = \mathrm{Meas}(\mathbb{R})$,
Closed term $\vdash M : \mathcal{R}$ as a measure μ and
Term $x : \mathcal{R} \vdash N : \mathcal{R}$ as a stable $f : \mathrm{Meas}(\mathbb{R}) \to \mathrm{Meas}(\mathbb{R})$.

Operational semantics

$$\forall r,\ \text{s.t. } M \to r,\ \texttt{let}(x, M, N) \to N\{r/x\}$$

By **Soundness**

$$[\![\texttt{let}(x, M, N)]\!] = \int_{\mathbb{R}} (\underset{\overset{\uparrow}{[\![N]\!]}}{f} \circ \delta)(r)\ \mu\,(dr)$$

Step 3: The Measurability Problem

Type $\mathcal{R}$ is interpreted as $\llbracket \mathcal{R} \rrbracket = \mathrm{Meas}(\mathbb{R})$,
Closed term $\vdash M : \mathcal{R}$ as a measure μ and
Term $x : \mathcal{R} \vdash N : \mathcal{R}$ as a stable $f : \mathrm{Meas}(\mathbb{R}) \to \mathrm{Meas}(\mathbb{R})$.

Operational semantics

$$\forall r, \text{ s.t. } M \to r, \ \texttt{let}(x, M, N) \to N\{r/x\}$$

By **Soundness**

$$\llbracket \texttt{let}(x, M, N) \rrbracket = \int_{\mathbb{R}} (\underbrace{f}_{\llbracket N \rrbracket} \circ \underbrace{\delta}_{\text{Dirac measure}})(r) \ \ \mu\,(dr)$$

Step 3: The Measurability Problem

Type $\mathcal{R}$ is interpreted as $[\![\mathcal{R}]\!] = \mathrm{Meas}(\mathbb{R})$,
Closed term $\vdash M : \mathcal{R}$ as a measure μ and
Term $x : \mathcal{R} \vdash N : \mathcal{R}$ as a stable $f : \mathrm{Meas}(\mathbb{R}) \to \mathrm{Meas}(\mathbb{R})$.

Operational semantics

$$\forall r,\ \text{s.t. } M \to r,\ \texttt{let}(x, M, N) \to N\{r/x\}$$

By **Soundness**

$$[\![\texttt{let}(x, M, N)]\!] = \int_{\mathbb{R}} (\underset{[\![N]\!]}{f} \circ \underset{\text{Dirac measure}}{\delta})(r)\ \underset{[\![M]\!]}{\mu}\,(dr)$$

Step 3: The Measurability Problem

Type $\mathcal{R}$ is interpreted as $[\![\mathcal{R}]\!] = \mathrm{Meas}(\mathbb{R})$,
Closed term $\vdash M : \mathcal{R}$ as a measure μ and
Term $x : \mathcal{R} \vdash N : \mathcal{R}$ as a stable $f : \mathrm{Meas}(\mathbb{R}) \to \mathrm{Meas}(\mathbb{R})$.

Operational semantics

$$\forall r,\ \text{s.t. } M \to r,\ \texttt{let}(x, M, N) \to N\{r/x\}$$

By **Soundness**

$$[\![\texttt{let}(x, M, N)]\!] = \int_{\mathbb{R}} (\ f\ \circ\ \delta\ )(r)\ \ \mu\ (dr)$$

Thus $f \circ \delta$ needs to be measurable.

- There are non measurable stable functions
- We need to equip every cone with a notion of measurability

Step 3: Measurability tests

Measurability tests of $\mathrm{Meas}(\mathbb{R})$ are given by measurable sets of $\mathbb{R}$:

$$\forall U \subseteq \mathbb{R} \text{ measurable, } \varepsilon_U \in \mathrm{Meas}(\mathbb{R})' : \mu \mapsto \mu(U)$$

For needs of CCC, we parameterized measurable tests of a cone:

Measurable Cone

A cone P with a collection $(\mathrm{M}^n(P))_{n\in\mathbb{N}}$ with $\mathrm{M}^n(P) \subseteq (P')^{\mathbb{R}^n}$ s.t.:

$$0 \in \mathrm{M}^n(P), \quad \ell \in \mathrm{M}^n(P) \text{ and } h : \mathbb{R}^p \to \mathbb{R}^n \Rightarrow \ell \circ h \in \mathrm{M}^p(P)$$

$$\ell \in \mathrm{M}^n(P) \text{ and } x \in P \Rightarrow \left\{ \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^+ \\ \vec{r} & \mapsto & \ell(\vec{r})(x) \end{array} \right. \text{ measurable.}$$

Measurable Tests, Paths and Functions

$\mathbf{Cstab}_{\mathsf{m}}$ is the category of complete and measurable cones with stable and measurable functions.

Let P and Q be measurable and complete cones:

Measurable Test: $\mathsf{M}^n(P) \subseteq (P')^{\mathbb{R}^n}$

Measurable Path: $\mathsf{Path}^n(P) \subseteq P^{\mathbb{R}^n}$ the set of bounded $\gamma : \mathbb{R}^n \to P$ such that $\ell * \gamma : \mathbb{R}^{k+n} \to \mathbb{R}^+$ is measurable with

$$\ell * \gamma : (\vec{r}, \vec{s}) \mapsto \ell(\vec{r})(\gamma(\vec{s}))$$

Measurable Functions: Stable functions $f : P \to Q$ such that:

$$\forall n \in \mathbb{N},\ \forall \gamma \in \mathsf{Path}_1^n(P), \quad f \circ \gamma \in \mathsf{Path}^n(Q)$$

If X is a measurable space, then $\mathsf{Meas}(X)$ is equipped with:
$\mathsf{M}^n(X) = \{\varepsilon_U : \mathbb{R}^n \to \mathsf{Meas}(X)' \text{ s.t. } \varepsilon_U(\vec{r})(\mu) = \mu(U),\ U \text{ meas.}\}$
$\mathsf{Path}_1^n(P)$ is the set of stochastic kernels from $\mathbb{R}^n$ to X.

1. Discrete Probability

2. **Continuous Probability**
    - Syntax: **Real** Probabilistic PCF
    - Semantics: $\mathbf{Cstab}_\mathrm{m}$ (Cones and Stable measurable functions)
    - Results: **Adequacy**

Results

The category **Cstab**$_{\mathrm{m}}$ is a CCC and a model of Real PPCF.

Interpretation of some terms:

$$[\![\underline{r}]\!] = \delta_r,\ [\![\texttt{sample}]\!] = \lambda_{[0,1]},\ [\![\texttt{let}(x, M, N)]\!](U) = \int_{\mathbb{R}} [\![N]\!](\delta_r)(U)\,[\![M]\!](dr)$$

Soundness

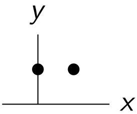
$$[\![M]\!]^{\Gamma\vdash A} = \int_{\Lambda^{\Gamma\vdash A}} [\![t]\!]^{\Gamma\vdash A}\mathrm{Red}(M, dt)$$

Adequacy

$$[\![M]\!]^{\vdash\mathcal{R}}(U) = \mathrm{Red}^{\infty}(M, U)$$

Examples: Distributions

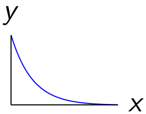
The Bernoulli distribution takes the value 1 with probability p and the value 0 with probability $1 - p$.



$$\texttt{bernoulli}\ p ::= \texttt{let}(x, \texttt{sample}, x \leq p)$$

$$\llbracket \texttt{bernoulli}\ \underline{p} \rrbracket^{\vdash \mathcal{R}} = p\delta_1 + (1-p)\delta_0$$

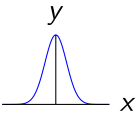
The exponential distribution is specified by its density e^{-x}.



$$\exp : \mathcal{R} ::= \texttt{let}(x, \texttt{sample}, -\log(x))$$

$$\llbracket \exp \rrbracket^{\vdash \mathcal{R}}(U) = \int_{\mathbb{R}^+} \chi_U(s) \mathrm{e}^{-s} \lambda(ds)$$

The standard normal distribution defined by its density $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.



$$\texttt{normal} ::=$$

$$\texttt{let}(x, \texttt{sample}, \texttt{let}(y, \texttt{sample}, \sqrt{-2\log(x)}\cos(2\pi y)))$$

$$\llbracket \texttt{normal} \rrbracket^{\vdash \mathcal{R}}(U) = \frac{1}{\sqrt{2\pi}} \int_U \mathrm{e}^{-\frac{x^2}{2}} \lambda(dx)$$

Examples:

Conditioning: If $U \subseteq \mathbb{R}$ measurable, then $\texttt{observe}(U)$ of type $\mathcal{R} \to \mathcal{R}$, taking a term M and returning the renormalization of the distribution of M on the only samples that satisfy U:

$\texttt{observe}(U) = \lambda m.\texttt{Y}(\lambda y.\texttt{let}(x, m, \texttt{if}(x \in U, x, y)))$

conditioning by rejection sampling.

Whenever M represents a probability distribution, this equation gives the conditional probability:

$$\llbracket \texttt{observe}(U) M \rrbracket (V) = \frac{\llbracket M \rrbracket (V \cap U)}{\llbracket M \rrbracket (U)}$$

Conclusion

Pcoh and $\mathbf{Cstab_m}$ models of probabilistic programming

- For countable data types, **Pcoh** is fully abstract.
- For real data types, $\mathbf{Cstab_m}$ is a sound model that encodes probability measures used in probabilistic programming.

Further directions:

- A model of LL ?
- A model of pCBPV ?
- Full abstraction ?