Probabilistic Coherent Spaces - a tutorial

DILL - CIRM, Marseille, May 2024

Christine Tasson

Institut Supérieur de l'Aéronautique et de l'Espace
SUPAERO

Introduction

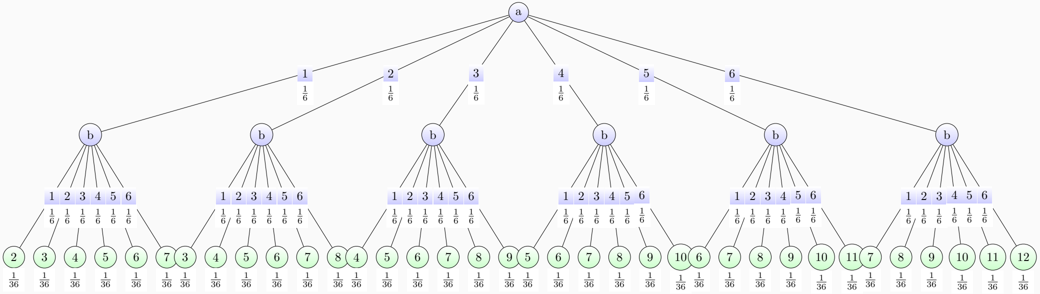
Probabilistic Programming

Sum of two Dice

Baudart, *muPPL, https://github.com/gbdrt/mu-ppl, a PPL prototype in Python*

```
def dice() → int:
    a = sample(RandInt(1, 6), name="a")
    b = sample(RandInt(1, 6), name="b")
    return a + b
```

dice is a Random Variable whose semantics is a distribution over $\{2, \ldots, 12\}$.

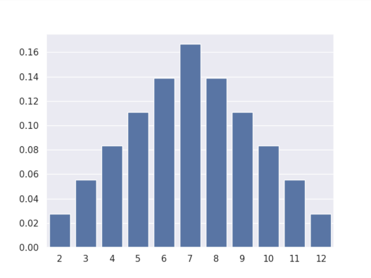


Sum of two Dice

Baudart, muPPL, https://github.com/gbdrt/mu-ppl, a PPL prototype in Python

```
def dice() → int:
    a = sample(RandInt(1, 6), name="a")
    b = sample(RandInt(1, 6), name="b")
    return a + b

with Enumeration():
    dist: Categorical[float] = infer(dice)
```

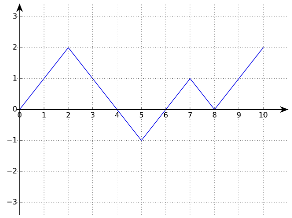


dice is a Random Variable whose semantics is a distribution over $\{2, \ldots, 12\}$.

$$
\begin{aligned}
[\![\mathrm{dice}]\!] : \quad \mathbb{N} &\rightarrow \mathbb{R}^+ \\
k &\mapsto \sum_{a=1}^{6} \sum_{b=1}^{6} \frac{1}{36} \mathbb{1}_{\{a+b=k\}}
\end{aligned}
$$

Random Walk

```
def RandomWalk(n:int, p:float) → int:
    if n == 0 :
        return 0
    else:
        step = 1 if sample(Bernoulli(p)) else -1
        return RandomWalk(n - 1, p) + step
```



Let RW be the Random Variable associated to $\mathrm{RandomWalk}(100, 0.3)$.
Its semantics is a distribution over $\mathbb{N}$, with finite support.

$$\begin{array}{rccc} [\![\mathrm{RW}]\!] : & |\mathrm{int}| & \to & \mathbb{R}^+ \\ & k & \mapsto & \mathbb{P}(\mathrm{RW} = k) \end{array} \qquad \mathrm{supp}([\![\mathrm{RW}]\!]) = \{k \mid -100 \leq k \leq 100\}.$$

Monte-Carlo Simulation

The mean and the mass function of the distribution is approximated by 10 000 executions.

```
def RandomWalk(n:int, p:float) → int:
    if n == 0 :
        return 0
    else:
        step = 1 if sample(Bernoulli(p)) else -1
        return RandomWalk(n - 1, p) + step

with ImportanceSampling(num_particles=10000):
    RW: Categorical[float] = infer(RandomWalk, 100, 0.3)
```

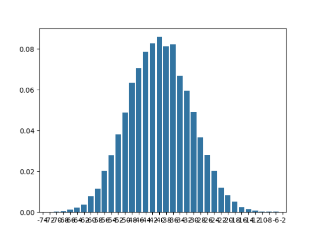
Galton Board

Monte-Carlo Simulation

The mean and the mass function of the distribution is approximated by 10 000 executions.

```
def RandomWalk(n:int, p:float) → int:
    if n == 0 :
        return 0
    else:
        step = 1 if sample(Bernoulli(p)) else -1
        return RandomWalk(n - 1, p) + step

with ImportanceSampling(num_particles=10000):
    RW: Categorical[float] = infer(RandomWalk, 100, 0.3)
```



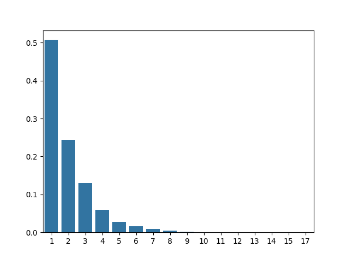
Approximated Mass Function

What is the exact distribution $[\![\mathrm{RW}]\!] \in (\mathbb{R}^+)^{\mathbb{N}}$?

Stopping Time

```
def StoppingTime(p:float) → int:
    time = 1
    while sample(Bernoulli(p)):
        time = time +1
    return time

with ImportanceSampling(num_particles=10000):
    ST: Categorical[float] = infer(StoppingTime, 0.5)
```



Approximated Mass Function

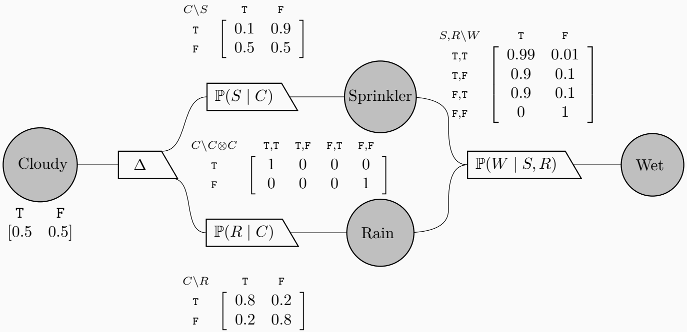
$\mathrm{ST} = \mathrm{StoppingTime}(0.5)$ is a Random Variable whose semantics is a distribution over $\mathbb{N}$, with infinite support.

$$\begin{aligned} [\![\mathrm{ST}]\!] : \quad |\mathrm{int}| &\to \mathbb{R}^+ \\ k &\mapsto \mathbb{P}(\mathrm{ST} = k) \end{aligned} \qquad \mathrm{supp}([\![\mathrm{ST}]\!]) = \mathbb{N}.$$

What is its exact distribution $[\![\mathrm{ST}]\!] \in (\mathbb{R}^+)^{\mathbb{N}}$?

Bayesian Network

Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference



$\boxed{\mathbb{P}(C) : |\mathrm{bool}| = \{\mathrm{T}, \mathrm{F}\} \to \mathbb{R}^+}$

$\boxed{\mathbb{P}(S \mid C) : |\mathrm{bool}| \times |\mathrm{bool}| \to \mathbb{R}^+}$

$$\mathbb{P}(W) = \mathbb{P}(C) \; ; \; \Delta \; ; \; (\; \mathbb{P}(S \mid C) \otimes \mathbb{P}(R \mid C) \;) \; ; \; \mathbb{P}(W \mid S, R)$$

Introduction

Origins of Probabilistic Coherent Spaces

Origins of Probabilistic Coherent Spaces

Girard, "Normal functors, power series and λ-calculus" define Coherent Spaces, where values of type A are given by a carrier $|X|$ and a closed program $\vdash t : A$ as a part $[\![t]\!] \in \mathcal{P}(|X|)$.
An environment $x : A \vdash s : 1$ interacts deterministically with any closed program $\vdash t : A$ when:

$$[\![s]\!] \perp [\![t]\!] \iff \# [\![t]\!] \cap [\![s]\!] \leq 1.$$

Girard, "Between Logic and Quantic: a Tract" introduces Probabilistic Coherent Spaces, as a generalization of Coherent Spaces where subsets are replaced by factors: $[\![t]\!] : |X| \to \mathbb{R}^+$. An environment $x : A \vdash s : 1$ interacts probabilistically with any closed program $\vdash t : A$ when:

$$[\![s]\!] \perp [\![t]\!] \iff \sum_{x \in |X|} [\![t]\!]_x \cdot [\![s]\!]_x \leq 1.$$

Danos and Ehrhard, "Probabilistic coherence spaces as a model of higher-order probabilistic computation" gives explicit definition of all Linear Logic connectors and fixpoint of types. It studies mathematical properties of Probabilistic Coherent Spaces. It gives a model in Probabilistic Coherent Spaces of pure lambda-calculus and PCF with binary choice and proves adequacy.

Probabilistic Coherent Spaces

Linear Category

Linear Category Pcoh (objects)

Orthogonality

Let $\mathbb{I}$ be countable and $u, u' \in (\mathbb{R}^+)^{\mathbb{I}}$.

$$u \perp u' \iff \sum_{a \in \mathbb{I}} u_a u'_a \leq 1$$

Definition

A Probabilistic Coherent Space (PCS) is a pair $X = (|X|, \mathrm{P}X)$ where $|X|$ is a countable set and $\mathrm{P}X \subseteq (\mathbb{R}^+)^{|X|}$ satisfies

- $\mathrm{P}X^{\perp\perp} = \mathrm{P}X$ (equivalently, $\mathrm{P}X^{\perp\perp} \subseteq \mathrm{P}X$) (Biorthogonality),
- for each $a \in |X|$ there exists $u \in \mathrm{P}X$ such that $u_a > 0$ (Coverage),
- for each $a \in |X|$ there exists $A > 0$ such that $\forall u \in \mathrm{P}X \ u_a \leq A$ (Boundedness).

Examples of Base Type

Product

$$\left|\prod_{i\in I} X_i\right| = \bigcup_{i\in I}\{i\}\times|X_i| \quad \text{and} \quad \mathrm{P}\left(\prod_{i\in I} X_i\right) = \{u \mid \forall i\in I\ u(i)\in \mathrm{P}X_i, \text{ where } \forall a\in|X_i|\ u(i)_a = u_{(i,a)}\}$$

unit is interpreted as distributions over the singleton $|1| = \{*\}$ and $\mathrm{P}(1) = [0,1] \subseteq \mathbb{R}^+$.

Enumeration

$$\left|\bigoplus_{i\in I} X_i\right| = \bigcup_{i\in I}\{i\}\times|X_i| \quad \text{and} \quad \mathrm{P}\left(\bigoplus_{i\in I} X_i\right) = \left\{u \mid \forall i\in I\ u(i)\in \mathrm{P}X_i \text{ and } \sum_{i\in I}\|u(i)\|_{X_i}\leq 1\right\}$$

$\mathrm{bool} = 1\oplus 1$. Booleans are interpreted as sub-probability distributions over booleans

$$|\mathrm{bool}| = \{\mathrm{T},\mathrm{F}\} \qquad \text{and} \qquad \mathrm{P}(\mathrm{bool}) = \{a\delta_T + b\delta_F \in \left(\mathbb{R}^+\right)^{\{\mathrm{T},\mathrm{F}\}} \mid a+b\leq 1\}.$$

Type Fixpoint

PCSs is a CPO with least element 0 with $|0| = \emptyset$ when ordered by

$$X \subseteq Y \qquad \text{iff} \qquad |X| \subseteq |Y| \qquad \text{and} \qquad \mathrm{P}X = \{v_{||X|} \mid v \in \mathrm{P}Y\}$$

$\mathrm{int} = 1 \oplus \mathrm{int}$. Integers are interpreted as sub-probability distributions over integers

$$|\mathrm{int}| = \mathbb{N} \qquad \text{and} \qquad \mathrm{P}(\mathrm{int}) = \left\{ \sum_{n \in \mathbb{N}} x_n \delta_n \mid \sum_{n \in \mathbb{N}} x_n \leq 1 \right\}.$$

Properties of $\mathrm{P}X$

$\mathrm{P}X$ is unitary $\|u\|_X \in [0,1]$ for all $u \in \mathrm{P}X$ where the norm is defined as

$$\|u\|_X = \sup\left\{ \sum_{a \in |X|} u_a u'_a \;\middle|\; u' \in \mathrm{P}X^{\perp} \right\}$$

$\mathrm{P}X$ is a cone

$$\forall u, v \in \mathrm{P}X \, \forall \alpha, \beta \in \mathbb{R}^+ \quad \alpha + \beta \leq 1 \Rightarrow \alpha\, u + \beta\, v \in \mathrm{P}X.$$

$\mathrm{P}X$ is an ω-continuous domain where the partial order is defined as

$$u \leq v \qquad \text{iff} \qquad \forall a \in |X| \;\; u_a \leq v_a \in \mathbb{R}^+.$$

Back to the StoppingTime Example

```
def StoppingTime(p:float) → int:
    time = 1
    while sample(Bernoulli(p)):
        time = time +1
    return time
```

After k iterations,

$$\llbracket \mathrm{ST} \rrbracket^1 = (1-p)\delta_1$$
$$\llbracket \mathrm{ST} \rrbracket^{k+1}_t = p \llbracket \mathrm{ST} \rrbracket^k_{t-1} + (1-p) \llbracket \mathrm{ST} \rrbracket^k_t$$

$\llbracket \mathrm{ST} \rrbracket$ is the lub of the increasing sequence $\llbracket \mathrm{ST} \rrbracket^k$ in $\mathrm{P}(\mathrm{int}) \subseteq (\mathbb{R}^+)^{\mathbb{N}}$

Linear Category Pcoh (morphisms)

A morphism of PCSs from X to Y is a matrix $t \in (\mathbb{R}^+)^{|X| \times |Y|}$ which maps $\mathrm{P}X$ to $\mathrm{P}Y$.

$$\forall u \in \mathrm{P}X \; t\, u \in \mathrm{P}Y \iff \forall v' \in \mathrm{P}Y^\perp, \qquad \sum_{(a,b) \in |X| \times |Y|} t_{a,b} u_a v'_b \leq 1 \,.$$

Identity

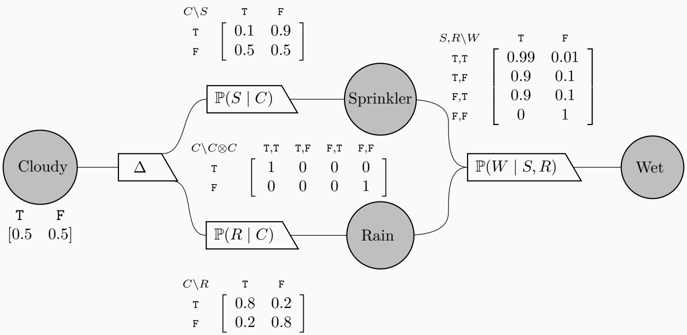
The diagonal matrix $\mathrm{id} \in (\mathbb{R}^+)^{|X| \times |X|}$, given by $\mathrm{id}_{a,b} = 1$ if $a = b$ and $\mathrm{id}_{a,b} = 0$ otherwise

Composition

Matrix multiplication, let $s \in \mathbf{Pcoh}(X, Y)$ and $t \in \mathbf{Pcoh}(Y, Z)$

$$(t\, s)_{a,c} = \sum_{b \in |Y|} s_{a,b} t_{b,c}$$

Back to a Bayesian Network example



$\boxed{\mathbb{P}(C) \in \mathrm{P}(\mathrm{bool}) \subseteq (\mathbb{R}^+)^{|\mathrm{bool}|}}$

$\boxed{\mathbb{P}(S \mid C) \in \mathbf{Pcoh}(\llbracket \mathrm{bool} \rrbracket, \llbracket \mathrm{bool} \rrbracket) : (\mathbb{R}^+)^{|\mathrm{bool}| \times |\mathrm{bool}|}}$

$$\mathbb{P}(W) = \mathbb{P}(C)\ ;\ \Delta\ ;\ (\ \mathbb{P}(S \mid C) \otimes \mathbb{P}(R \mid C)\ )\ ;\ \mathbb{P}(W \mid S, R)$$

Probabilistic Coherent Spaces

Pcoh is Symmetric Monoidal, Closed

Pcoh is Symmetric Monoidal Closed

Joint Distributions

if U is a random variable of distribution $u \in (\mathbb{R}^+)^{|X|}$ and V is a random variable of distribution $v \in (\mathbb{R}^+)^{|Y|}$, then the joint distribution $u \otimes v \in (\mathbb{R}^+)^{|X| \times |Y|}$ is given by:

$$(u \otimes v)_{(a,b)} = \mathbb{P}(U \otimes V = (a,b)) = \mathbb{P}(U = a \wedge V = b) = u_a\, v_b$$

Tensor Product

$$|X \otimes Y| = |X| \times |Y| \qquad \text{and} \qquad \mathrm{P}(X \otimes Y) = \{u \otimes v \;\mid\; u \in \mathrm{P}X \quad \text{and} \quad v \in \mathrm{P}Y\}^{\perp\perp}$$

Associativity, Symmetry, unitor are morphisms of PCSs

Closed Structure

$$|X \multimap Y| = |X| \times |Y| \quad \text{and} \quad \mathrm{P}(X \multimap Y) = \mathbf{Pcoh}(X,Y) = \left\{ t \in \left(\mathbb{R}^+\right)^{|X| \times |Y|} \;\mid\; \forall u \in \mathrm{P}X\; t\, u \in \mathrm{P}Y \right\}$$

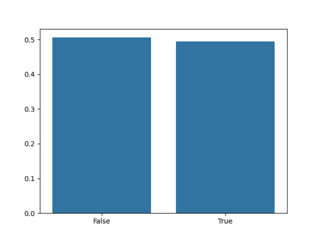
Evaluation and Curryfication are morphisms of PCSs

Fair Coin Example

```
def FairCoin(p:float) → bool:
    a = sample(Bernoulli(p))
    b = sample(Bernoulli(p))
    if (a and not b):
        return True
    elif (b and not a):
        return False
    else:
        return FairCoin(p)

with ImportanceSampling(num_particles=1000):
    FC: Categorical[bool] = infer(FairCoin, 0.3)
```

FC is a fair coin !



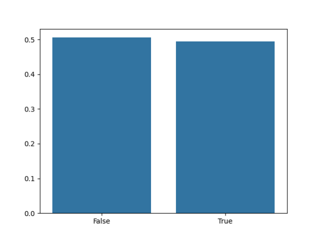
Pcoh *is CPO enriched indeed,* $\mathrm{P}(X,Y) = \mathrm{P}(X \multimap Y)$ *is a CPO with* 0 *as least element.*

Fair Coin Example

```
def FairCoin(p:float) → bool:
    a = sample(Bernoulli(p))
    b = sample(Bernoulli(p))
    if (a and not b):
        return True
    elif (b and not a):
        return False
    else:
        return FairCoin(p)

with ImportanceSampling(num_particles=1000):
    FC: Categorical[bool] = infer(FairCoin, 0.3)
```

FC is a fair coin !



Recursive equation (if $p \notin \{0,1\}$):

$$\mathscr{F}(t) = p(1-p)(\delta_{\mathrm{T}} + \delta_{\mathrm{F}}) + (1 - 2(p(1-p)))\, t$$

$$[\![\mathrm{FC}]\!] = \tfrac{1}{2}(\delta_{\mathrm{T}} + \delta_{\mathrm{F}})$$

Pcoh *is CPO enriched indeed,* $\mathrm{P}(X, Y) = \mathrm{P}(X \multimap Y)$ *is a CPO with* 0 *as least element.*

Probabilistic Coherent Spaces

Exponential

The exponential comonad

Finite Multiset. If $u \in (\mathbb{R}^+)^{|X|}$ is a distribution, then $u^! \in (\mathbb{R}^+)^{\mathfrak{M}_{\mathrm{fin}}|X|}$ is defined by if $\mu = [a_1, \ldots, a_k]$ and $U_1, \ldots, U_k$ are *i.i.d.* of law u,

$$u^!_\mu = \mathbb{P}(U_1 = a_1 \wedge \cdots \wedge U_k = a_k) = \prod_{a \in |X|} u_a^{\mu(a)}$$

Definition

$$|!X| = \mathfrak{M}_{\mathrm{fin}}|X| \quad \text{and} \quad \mathrm{P}(!X) = \left\{u^! \mid u \in \mathrm{P}X\right\}^{\perp\perp}$$

For $t \in \mathrm{P}(X, Y)$, the promotion $t^! \in \mathrm{P}(!X, !Y)$ is defined such that

$$t^! u^! = (t\, u)^!$$

Comonad: ! with counit $\mathrm{der}^X \in \mathrm{P}(!X \multimap X)$ and comultiplication $\mathrm{dig}^X \in \mathrm{P}(!X \multimap !!X)$.

Strong symmetric monoidal structure given by isomorphisms from $(\mathbf{Pcoh}, \&)$ to $(\mathbf{Pcoh}, \otimes)$,

$$m^0 \in \mathbf{Pcoh}(!\top, 1) \qquad \text{and} \qquad m^2 \in \mathbf{Pcoh}(!(X \& Y), !X \otimes !Y)$$

Duplication and Eilenberg Moore Category

$!X$ is the free Commutative Comonoid over X.

$$\mathrm{contr}^{!X} \in \mathrm{P}(!X \multimap !X \otimes !X) \quad \text{and} \quad \mathrm{weak}^{!X} \in \mathrm{P}(!X \multimap 1)$$

Crubillé et al., "The Free Exponential Modality of Probabilistic Coherence Spaces"

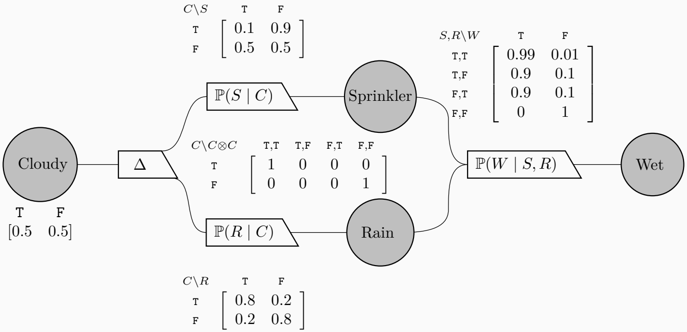
A Coalgebra is a PCS P with $h \in \mathrm{P}(P \multimap !P)$ compatible with the !-comonad structure.
Every coalgebra P comes with marginalization, duplication and erasure

$$\pi_P \in \mathrm{P}(P \otimes Q \multimap P) \quad \text{and} \quad \Delta_P \in \mathrm{P}(P \multimap P \otimes P) \quad \text{and} \quad \epsilon_P \in \mathrm{P}(P \multimap 1)$$

Value types are interpreted as coalgebras

$$\varphi, \psi := \mathrm{unit} \mid !\sigma \mid \varphi \otimes \psi \mid \varphi \oplus \psi \mid \zeta \mid \mathit{Rec}\, \zeta \varphi$$

Back to a Bayesian Network example



$\mathbb{P}(C) \in \mathrm{P}(\mathrm{bool}) \subseteq (\mathbb{R}^+)^{|\mathrm{bool}|}$

$\mathbb{P}(S \mid C) \in \mathbf{Pcoh}(\llbracket \mathrm{bool} \rrbracket, \llbracket \mathrm{bool} \rrbracket) : (\mathbb{R}^+)^{|\mathrm{bool}| \times |\mathrm{bool}|}$

$$\mathbb{P}(W) = \mathbb{P}(C) \; ; \; \Delta \; ; \; (\; \mathbb{P}(S \mid C) \otimes \mathbb{P}(R \mid C) \;) \; ; \; \mathbb{P}(W \mid S, R)$$

Probabilistic Coherent Spaces

Semantics of Probabilistic Programming

Pcoh models probabilistic programming

Syntaxes

Probabilistic Call-By-Push-Value.

General types $\sigma, \tau := \varphi \mid \varphi \multimap \sigma$

Value types $\varphi, \psi := \mathrm{unit} \mid !\sigma \mid \varphi \otimes \psi \mid \varphi \oplus \psi \mid \zeta \mid \mathit{Rec}\, \zeta \varphi$

Programs are typed in value contexts $x_1 : \varphi_1, \ldots, x_k : \varphi_k \vdash M : \sigma$

$[\![t]\!] \in \mathrm{P}(P_1 \otimes \cdots \otimes P_k \multimap X)$ where $[\![\varphi_i]\!] = P_i$ are coalgebras and $[\![\sigma]\!] = X$

Probabilistic Call-By-Value. $A \to B$ encoded by $!(A \multimap B)$.

Probabilistic Call-By-Name. $A \Rightarrow B$ encoded by $(!A) \multimap B$.

Probabilistic PCF. Fixpoint operator: $\mathcal{Y} \in \mathbf{Pcoh}_!(X \Rightarrow X, X)$

Probabilistic Untyped Calculus. The reflexive object satisfies $D = \left(!D^{\mathbb{N}}\right)^{\perp}$.

Soundness

$$[\![M]\!] = \sum_{M'} \mathbb{P}(M \to M')\, [\![M']\!]$$

Observational equivalence and Adequacy

Observational Distance For every $\vdash M : \sigma$ and $\vdash N : \sigma$,

$$\mathrm{d}_{\mathrm{obs}}(M, N) = \sup\{|\mathbb{P}(CM\downarrow) - \mathbb{P}(CN\downarrow)| \;\mid \vdash C :!\sigma \multimap \mathrm{unit}\}$$

Semantics Distance

$$\mathrm{d}_X(x, y) = \|x - (x \wedge y)\| + \|(y - (x \wedge y))\|$$

Adequacy

If $[\![M]\!] = [\![N]\!]$, that is $\mathrm{d}_{[\![\sigma]\!]}([\![M]\!], [\![N]\!]) = 0$ then $\mathrm{d}_{\mathrm{obs}}(M, N) = 0$

The converse is also true !

$$\mathrm{d}_{[\![\sigma]\!]}([\![M]\!], [\![N]\!]) = 0 \quad \text{iff} \quad \mathrm{d}_{\mathrm{obs}}(M, N) = 0$$

Ehrhard, Pagani, and Tasson, "Full Abstraction for Probabilistic PCF"

Metric Adequacy

Ehrhard, "Differentials and Distances in Probabilistic Coherence Spaces"

Amplification of Probability Take $C = \mathit{fix}(\lambda f.\, \lambda x.\, \mathrm{if}(\mathrm{Bernoulli}(r), \mathrm{T}, (f)x))$

$$\forall \epsilon \in [0,1] \;\; \mathrm{d_{obs}}(\mathrm{Bernoulli}(0), \mathrm{Bernoulli}(\epsilon)) = 1$$

***p*-tamed observational distance** $C^{\langle p \rangle} = \mathit{fix}(\lambda x.\, (C)\mathrm{if}(\mathrm{Bernoulli}(p), x, \Omega))$

$$\mathrm{d}^{\langle \mathsf{p} \rangle}_{\mathrm{obs}}(M, N) = \sup \left\{ \left| \mathbb{P}(C^{\langle p \rangle} M \downarrow) - \mathbb{P}(C^{\langle p \rangle} N \downarrow) \right| \;\middle|\; \vdash C : !\sigma \multimap \mathrm{unit} \right\}$$

Metric Adequacy $\mathrm{d}^{\langle \mathsf{p} \rangle}_{\mathrm{obs}}(M, N) \leq \frac{p}{1-p} \, \mathrm{d}_{[\![\sigma]\!]}([\![M]\!], [\![N]\!])$

Thus

$$\mathrm{d}^{\langle \mathsf{p} \rangle}_{\mathrm{obs}}(\mathrm{Bernoulli}(0), \mathrm{Bernoulli}(\epsilon)) \leq \frac{p\epsilon}{1-p}$$

Probabilistic Coherent Spaces

Non-Linear Category

Non-Linear Category

Kleisli Category $\mathbf{Pcoh}_!$ with PCSs as objects and morphisms $\mathbf{Pcoh}_!(X,Y) = \mathbf{Pcoh}(!X,Y)$.

Taylor Expansion. $t \in \mathbf{Pcoh}_!(X,Y)$ iff $t \in (\mathbb{R}^+)^{\mathfrak{M}_{\mathrm{fin}}(X)\times Y}$ and

$$\forall u \in \mathrm{P}(X) \quad t(u) = t\, u^! = \left(\sum_{\mu \in \mathfrak{M}_{\mathrm{fin}}(X)} t_{\mu,b} \prod_{a\in|X|} u_a^{\mu(a)}\right)_{b\in|Y|} \in \mathrm{P}(Y)$$

Non Definable morphisms

$$\mathbf{Pcoh}(!\mathrm{bool},1) = \left\{Q \in (\mathbb{R}^+)^{\mathfrak{M}_{\mathrm{fin}}\mathrm{T},\mathrm{F}} \;\middle|\; Q_{\mathrm{T}^n,\mathrm{F}^m} \leq \frac{(n+m)^{n+m}}{n^n m^m}\right\} \quad \text{and} \quad \text{max. coeff.} \frac{(n+m)!}{n!m!}$$

Fully Abstract.

$$\mathbf{Pcoh}(!1,1) = \left\{q \in (\mathbb{R}^+)^{\mathbb{N}} \;\middle|\; \forall x \in [0,1] \sum q_n x^n \in [0,1]\right\}$$

if $[\![M]\!] \neq [\![N]\!]$, build testing terms such that $[\![\lambda x.C(x)M : !1 \multimap 1]\!]$ and $[\![\lambda x.C(x)N : !1 \multimap 1]\!]$ are power series of with different coefficients, then there is $p \in [0,1]$ on which they differ.

Yet, Pcoh is not a model of DILL

Cocontraction is not a PCSs morphism If $f \in \mathbf{Pcoh}_!(X, Y)$, and $u, v \in \mathrm{P}(X)$, we cannot ensure that $f(u + v) \in \mathrm{P}(Y)$

No global derivative: $\mathbf{Pcoh}_!(1, 1)$ are entire series defined on $[0, 1]$ and not necessarily derivable at 1.

Example: $M_r = \mathit{fix}(\lambda f.\, \lambda x.\, \mathrm{if}(\mathrm{Bernoulli}(r), (f)x; (f)x, x; x))$.

$$\varphi(r)(u) = \llbracket M_r \rrbracket (u) = \sum_n a_n(r) u^n$$

$\varphi(r)(1) = \mathbb{P}(M \Downarrow)$ is the termination probability by adequacy.

The termination time conditioned to convergence can be computed by $\frac{\varphi'(r)(1)}{\varphi(r)(1)}$

But if $r = 0.5$, then $\varphi(r)$ is not derivable at $u = 1$, as

$$\forall u \in [0, 1]\, \varphi(0.5)(u) = 1 - \sqrt{1 - u^2}$$

Fortunately, a notion of Derivative

Local PCSs let $x \in \mathrm{P}X$, the local PCS X_x is given by

$$|X_x| = \{a \in |X| \mid \exists\epsilon > 0 x + \epsilon\delta_a \in \mathrm{P}X\} \qquad \text{and} \qquad \mathrm{P}(X_x) = \{u \in \mathrm{P}X \mid x + u \in \mathrm{P}X\}$$

Local Derivative if $x + u \in \mathrm{P}X$ and $t \in \mathbf{Pcoh}_!(X, Y)$ then $t(x + u) \in \mathrm{P}Y$

$$t(x+u) = \sum_{\mu} t_{\mu\in\mathfrak{M}_{\mathrm{fin}}|X|}(x+u)^{\mu} \in \mathrm{P}Y = t(x) + \sum_{a\in|X|} u_a \sum_{\nu\in\mathfrak{M}_{\mathrm{fin}}|X|} (\nu(a)+1)t_{\nu+[a],b}x^{\nu} + \dots$$

Local Derivative define $t' \in \mathbf{Pcoh}_!(X_x, Y_{t(x)})$

$$t'(x)_{a,b} = \sum_{\nu\in\mathfrak{M}_{\mathrm{fin}}|X|} (\nu(a)+1)t_{\nu+[a],b}x^{\nu}$$

Chain Rule $(t \circ s)'(x)(u) = t'(s(x))\, s'(x)u$

Ehrhard, "Differentials and Distances in Probabilistic Coherence Spaces"

Take Home

A model of **higher order probabilistic programming** with discrete probabilities.

A fully abstract model of **Linear Logic** with **Taylor expansion**, but not of DILL.

A model of **local differentiation**

At the origin of Coherent Differentiation

More on this subject is coming this week.