Stable Semantics of Probabilistic Higher-Order Programs

Raphaëlle Crubillé

IMDEA, IRIF, Univ. Paris-Diderot

TLLA 2019
Higher-order Probabilistic Programming

Higher-order languages extended with:

- **Discrete randomized** algorithms (e.g. randomized sorting);
- **continuous** probability distributions (e.g. to model physical systems);
- **Bayesian reasoning** (e.g. machine learning algorithms.)

\[
M \in \text{PCF}_\oplus ::= x \mid \lambda x^A \cdot M \mid (MN) \mid (YN) \\
| \text{ifz} (M, N, L) \mid \text{let}(x, M, N) \\
| M \oplus N \mid n \mid \text{succ} (M) \mid \text{pred} (M) \mid n \in \mathbb{N} \\
| \text{sample} \mid r \mid f \mid r \in \mathbb{R}, f : \mathbb{R} \to \mathbb{R} \text{ measurable} \\
| \text{score} \mid \text{normalize}
\]
Examples

Discrete randomized programs

\[ M \downarrow \mathcal{D} \text{ discrete sub-distribution} \quad \mathcal{D} : \{ \text{normal forms} \} \rightarrow [0, 1]. \]

\[ M = 0 \oplus 1 \]

\[ N = Y(\lambda x. N \rightarrow N \lambda y. N(y \oplus x(\text{succ}(y)))) \]

\[ M \downarrow \frac{1}{2}\{0^1\} + \frac{1}{2}\{1^1\} \]

\[ N \downarrow \sum_{n \in \mathbb{N}} \frac{1}{2^{n+1}} \{n^1\} \]

Program sampling from a continuous distribution

\[ M \downarrow \mathcal{D} \text{ continuous sub-distribution} \quad \mathcal{D} \in \sum\{ \text{normal forms} \} \rightarrow [0, 1]. \]

- \((\lambda x. \text{sample} + x) \downarrow \mu[1,2]\)
  with \(\mu[1,2]\) Lebesgue Measure on \([1,2]\);

- Exponential Distribution: \(M := \text{let}(x, \text{sample}, -\log(x))\)

\[ M \downarrow (E \mapsto \int_E e^{-x}) \]

\[ \text{Density Function} \]

\[ y \]

\[ x \]
Two stable Models of Higher-Order Probabilistic Computations

PCF + discrete probabilities $\to$ PCoh\textsuperscript{!}  
\hspace{2cm} $\llbracket \cdot \rrbracket$ fully abstract

PCF + continuous probabilities $\to$ Cstab\textsubscript{m}  
\hspace{2cm} $\llbracket \cdot \rrbracket$

- probabilistic coherence spaces (Danos, Ehrhard 2011) built as a model of LL
- measurable complete cones measurable stable functions (Ehrhard, Pagani, Tasson 2018)
Two stable Models of Higher-Order Probabilistic Computations

PCF + discrete probabilities $\rightarrow$ $\llbracket \cdot \rrbracket$ fully abstract $\rightarrow$ PCoh!

PCF + continuous probabilities $\rightarrow$ $\llbracket \cdot \rrbracket$ ? $\rightarrow$ Cstab_m

probabilistic coherence spaces (Danos, Ehrhard 2011) built as a model of LL

measurable complete cones measurable stable functions (Ehrhard, Pagani, Tasson 2018)

Probability Coherence Spaces

Proving operational properties on programs denotationally.
Outline

1. The Discrete Models of PCSs
   - Probabilistic Coherence Spaces
   - Proving operational properties on programs denotationally.

2. The Continuous Stable Model
   - Construction of the Continuous Model: An Overview
   - Probabilistic Stability
   - Adding measurability constraints.
The Discrete Models of PCSs

- Probabilistic Coherence Spaces
- Proving operational properties on programs denotationally.

The Continuous Stable Model

- Construction of the Continuous Model: An Overview
- Probabilistic Stability
- Adding measurability constraints.
**PCoh**: A Linear Logic model for discrete probabilistic computations (Danos, Ehrhard 2011).

Probabilistic coherence spaces (PCS)

pair $(|A|, P(A))$ where:
- $|A|$: *countable* web;
- $P(A) \subseteq \mathbb{R}^{|A|}$: quantitative cliques.

Orthogonality Relation

- For $u, v \in \mathbb{R}^{|A|}$, $\langle u, v \rangle = \sum_{a \in |A|} u_a \cdot v_a$;
- $A^\perp = \{ u \in \mathbb{R}^{|A|} \mid \forall v \in P(A), \langle u, v \rangle \leq 1 \}$.

Bi-Orthogonality Condition on PCSs

$A = A^\perp \perp$

Example (Booleans)

$|\text{Bool}| = \{t, f\}$

$P(\text{Bool}) = \{(p, q) ; p + q \leq 1\} = \{(1, 0), (0, 1)\}^\perp \perp$

$\Rightarrow$ sub-probability distributions on booleans.
**Linear Morphisms in \( \text{PCoh} \)**

**Morphisms of PCS**

\( \text{PCoh}(\mathcal{A}, \mathcal{B}) \): matrices \( \phi \in \mathbb{R}^{\lvert \mathcal{A} \rvert} \times \lvert \mathcal{B} \rvert \)

\( = \) linear functions \( \mathbb{R}^{\lvert \mathcal{A} \rvert} \rightarrow \mathbb{R}^{\lvert \mathcal{B} \rvert} \).

such that:

\( \forall x \in \mathbb{P}(\mathcal{A}), \phi(x) \in \mathbb{P}(\mathcal{B}) \).

**Example (\( \text{Bool} \rightarrow \text{Bool} \))**

\( \lvert \text{Bool} \rightarrow \text{Bool} \rvert = \lvert \text{Bool} \rvert \times \lvert \text{Bool} \rvert \).

\( \mathbb{P}(\text{Bool} \rightarrow \text{Bool}) = \{ u \mid u_{t,t} + u_{t,f} \leq 1 \land u_{f,t} + u_{f,f} \leq 1 \} \)

\[ \begin{array}{c}
0.6 \\
1 \\
0.4 \\
\end{array} \]

\( \Rightarrow \) Markov transitions.
PCoh: Exponential Modality

Exponential Comonad

- web: $|!A| = \mathcal{M}_f(|A|)$
- cliques:
  \[ P(!A) = \{ x^! \mid x \in P(A) \} \]

Promotion of $x$:
\[
\begin{align*}
x^! & \in \mathbb{R}^{\mathcal{M}_f(|A|)}; \\
x^!_{[a_1,\ldots,a_n]} & := \prod_i x_{a_i}.
\end{align*}
\]

Theorem (C., Ehrhard, Pagani, Tasson)

$\text{PCoh}$ is a Lafont model of Linear Logic
\textit{i.e.} $\forall A \text{ PCS, } !A$ is the free commutative comonoid generated by $A$. 
\textbf{PCoh}_! \text{ morphisms}

Functional meaning for \( f \in \mathbb{R}^{\mathcal{M}_f(|A|) \times |B|} \):

\[
\hat{f} : \mathcal{P}(A) \rightarrow (\mathbb{R} \cup +\infty)^{|B|}
\]
\[
x \mapsto f \cdot x^! 
\]

Lemma

\( f \in \text{PCoh}_!(A, B) \iff \hat{f} \text{ preserves cliques.} \)

Structure of Functional Interpretations for \( \text{PCoh}_!(A, B) \)

\textit{Power series} from \( \mathcal{P}(A) \) to \( \mathcal{P}(B) \) with:

- non-negative coefficients;
- a countable number of variables.
**PCoh**! morphisms

Functional meaning for $f \in \mathcal{R}^{\mathcal{M}_f(|A|) \times |B|}$:

$$\hat{f} : \mathcal{P}(A) \rightarrow \mathcal{P}(B)$$

$$x \mapsto f \cdot x^!$$

**Lemma**

$f \in \textbf{PCoh}_!(A, B)$ iff $\hat{f}$ preserves cliques.

**Structure of Functional Interpretations for** $\textbf{PCoh}_!(A, B)$

*Power series* from $\mathcal{P}(A)$ to $\mathcal{P}(B)$ with:

- non-negative coefficients;
- a countable number of variables.
1. The Discrete Models of PCSs
   - Probabilistic Coherence Spaces
   - Proving operational properties on programs denotationally.

2. The Continuous Stable Model
   - Construction of the Continuous Model: An Overview
   - Probabilistic Stability
   - Adding measurability constraints.
PCoh! as model of PCF⊕.

Example ($M : \text{Bool} \rightarrow 1$)

$$M = \text{let } b = \text{true} \oplus \text{false} \text{ in }$$
\[\text{fix } (\lambda f. \lambda x. \text{if } x = b \text{ then } \star \text{ else } fx).\]

The program $M$:
- chooses randomly a boolean;
- calls its argument until the result coincides with the boolean chosen.

$$\widehat{M} : \mathbb{P (\text{Bool})} \rightarrow \mathbb{P (1)} = [0, 1]$$
\[(x_t, x_\bar{t}) \mapsto \frac{1}{2} \sum_{n \geq 1} x_\bar{t}^{n-1} \cdot x_t + \frac{1}{2} \sum_{n \geq 1} x_t^{n-1} \cdot x_\bar{t}\]
Equivalence on Programs

Definition (Context Equivalence)

\[ M \equiv_{\text{ctx}} N \text{ when: } \forall \text{ context } C, \, \text{Obs}(C[M]) = \text{Obs}(C[N]). \]

In a probabilistic setting, \( \text{Obs}(\cdot) \) can be:
- the probability of termination of a program;
- the probability of returning 0 (for ground types program) . . .

Theorem

- \( \text{PCoh}^1 \): full abstraction for CBN probabilistic PCF. [EPT POPL '2014] i.e.:
  \[ M \equiv_{\text{ctx}} N \iff \llbracket M \rrbracket = \llbracket N \rrbracket. \]

- \( \text{PCoh}^1 \): full abstraction for a probabilistic version of Levy’s CBPV [ET ’2016].

Crucial component in the proof:
A power series is entirely characterized by its coefficients.
Toward a quantitative generalization of context equivalence (1)

From there:
- Can we also express **quantitative properties** in the model?
- What kind of quantitative operational properties do we want to model?

**Context Distance**

A quantitative generalization of Context Equivalence:

\[ \delta_{\text{ctx}}(M, N) = \sup_{C \text{ a context}} |C[M] - C[N]|. \]

**Problem with Context Distance:**

Contexts may be *too powerful*
i.e. amplify too much the distance between programs.
Toward a quantitative generalization of context equivalence (2)

Example (Two very similar programs at context distance 1)

\[ M = \text{true} \]
\[ N_\epsilon = \text{true} \oplus \epsilon \text{ false} \quad \text{with } \epsilon \ll 1 \]

We can show that \( \forall \epsilon > 0: \delta^{\text{ctx}}(M, N_\epsilon) = 1 \).

Proof.

We can construct a sequence of amplification contexts \( C_n \) such that:

\[ \text{Obs}(C_n[\text{true}]) = 1 \quad \text{Obs}(C_n[\text{true} \oplus \epsilon \text{ false}]) = (1 - \epsilon)^n \]

\[ C_n = (\lambda x. \text{if } (x \land \ldots \land x) \text{ then } l \text{ else } \Omega)[\cdot]. \]
Proposition (C., Dal Lago 2017)

*In a probabilistic \( \lambda \)-calculus where programs almost surely terminate, all programs are at distance either 0 or 1.*

Remark

- \( \text{PCF} \oplus \) has non-deterministic programs,
- but context distance may nonetheless be a too strong notion.
Tamed Context Distances

Idea

Each time a context uses its argument, it must pay a price in its contribution to context distance.

Definition (Tamed context distances (Ehrhard 2019))

For $p$ a dyadic number in $[0, 1]$:

$$\delta^\text{ctx}_p(M, N) = \sup_{C \text{ a context}} |C[\Omega \oplus^p M] - C[\Omega \oplus^p N]|.$$  

Theorem (Metric Adequacy of PCSs (Ehrhard 2019))

$$\delta^\text{ctx}_p(M, N) \leq \frac{p}{1 - p} d(\llbracket M \rrbracket, \llbracket N \rrbracket)$$  

where:

- $d$ is defined using the norm in $\text{PCoh}$;
  $$\| x \|_A = \sup_{y \in A^\perp} \langle x, y \rangle.$$

- For every $t \in P(A)$, $t'$ represents the derivative of $t$ (seen as a power series).
1. The Discrete Models of PCSs
   - Probabilistic Coherence Spaces
   - Proving operational properties on programs denotationally.

2. The Continuous Stable Model
   - Construction of the Continuous Model: An Overview
   - Probabilistic Stability
   - Adding measurability constraints.
Question:
Can we generalize this fully-abstract model of discrete probabilistic PCF to continuous probabilistic PCF?

Problem
Distribution on a continuous data-type (e.g. $\mathbb{R}$) cannot be seen as vectors over a web.
⇒ Does not allow to model continuous computations.

Stable semantics
A generalization of PCSs to a continuous setting.
Kozen’s semantics of First-Order Language with Continuous Probabilities

Kozen’s Language
A first-order while language with a random number generator.

Kozen’s semantics
- Possible configurations of the memory: **Measurable Spaces**; equipped with a cone structure to manage recursion.
- Program interpretation: **Stochastic Kernels** between measurable spaces.

Stochastic Kernels
$(X, \Sigma_X), (Y, \Sigma_Y)$ two measurable spaces $k : X \times \Sigma_Y \rightarrow [0, 1]$ such that:
- $\forall B \in \Sigma_Y, (x \in X \mapsto k(x, B))$ is $X$-measurable.
- $\forall x \in X, (B \in \Sigma_Y \mapsto k(x, B))$ is a probability measure on $Y$. 
Semantics for Probabilistic Higher-Order Languages

Fact
Neither Kern nor Meas are cartesian closed categories.

Staton et al’s Quasi Borel spaces

Idea: considering a cartesian closed category of presheaves embedding Meas.
⇒ Replace measurable spaces with space of the form \((X, V(X))\), where:

- \(X\) is any set
- \(V \subseteq (\mathbb{R} \to X)\) is a set of random variables.

The measurability constraints on the space are replaced by constraints on the set of random variables.

Example (Quasi-Borel spaces)

A continuous data-type: 
**The \(n\)-uples of reals**
\((\mathbb{R}^n, \{f : \mathbb{R} \to \mathbb{R}^n \mid f \text{ measurable}\})\)

A discrete data-type: 
**The booleans**
\((\{0, 1\}, \{f : \mathbb{R} \to \{0, 1\} \mid f \text{ characteristic function of Borel set}\})\)
Construction of the Continuous Stable Model [Ehrhard, Pagani, Tasson]

Two steps:

- **Cstab**: The category of complete cones and stable functions on cones;
  ⇒ Generalizing PCS to spaces without an underlying countable web

- **Cstab**\textsubscript{m}: The category obtained from **Cstab** by adding measurability constraints
  ⇒ using random variables in a similar way as in Staton et al’s QBS.

The CCC \( \text{Cstab} \)

Forget

The CCC \( \text{Cstab}\textsubscript{m} \)

interpretation of PCF\textsubscript{sample}
The Discrete Models of PCSs
- Probabilistic Coherence Spaces
- Proving operational properties on programs denotationally.

The Continuous Stable Model
- Construction of the Continuous Model: An Overview
- Probabilistic Stability
- Adding measurability constraints.
From PCSs to complete cones

**Definition (Order on a PCS)**

\[ A = (|A|, \mathcal{P}(A)) \] a PCS, \( x, y \in \mathcal{P}(A) \):

\[ x \leq_A y \quad \text{when} \quad \forall a \in |A|, x_a \leq y_a. \]

**Properties of the order \( \leq_A \)**

- \( \mathcal{P}(A) \) is an \( \omega \)-cpo;
  \( \Rightarrow \) allows to interpret fixpoints of programs.

- \( \forall x, y \in \mathcal{P}(A) : \ (x \leq_A y \iff \exists z \in \mathcal{P}(A), y = x + z. ) \)
  \( \Rightarrow \) it gives a cone structure to \( \mathcal{P}(A) \).

**Illustration on function space**

\( f, g \in \text{PCoh}_!(A, B) \) such that \( f \leq!_{A \rightarrow B} g. \)

- \( \hat{f} \leq \hat{g} \) coefficient-wise;

- \( \hat{f} – \hat{g} \) is still a power series with non-negative coefficients.
The CCC $\text{Cstab}_m$ [EPT(2018)] : (1)- Cones

Definition (Cones)

$C$: $\mathbb{R}$-semimodule, $\parallel \cdot \parallel_C : C \to \mathbb{R}$ with equational constraints on $+$, $\parallel \cdot \parallel_C$.

\[
\begin{align*}
(x + y = x + y') \Rightarrow y &= y' \\
\parallel x + x' \parallel_C &\leq \parallel x \parallel_C + \parallel x' \parallel_C \\
\parallel x \parallel_C &\leq \parallel x + x' \parallel_C \\
\parallel \alpha x \parallel_C &= \alpha \parallel x \parallel_C \\
\parallel x \parallel_C &= 0 \Rightarrow x = 0
\end{align*}
\]

Definitions:

- **Closed Unit Ball** $\mathcal{B}_C$;
  
  $\mathcal{B}_C ::= \{x | \parallel x \parallel_C \leq 1\}$

- **partial order** $\preceq_C$.
  
  $x \preceq_C y ::= \exists z \in C, y = x + z$.

- If $x \preceq_C y$, $x - y$ is the **unique** element $z$ of $C$ s.t. $x + z = y$.

**Additional Requirement:**

Order-Completeness of the Unit Ball
The CCC $\text{Cstab}$ [EPT(2018)] : (2)

Example (Some Complete Cones)

- $\mathbf{1} = (\mathbb{R}, |.|)$

- Non-negative cone of Lebesgues spaces:
  $L_1^+ (\mathbb{R}^n) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}_+$
  $\|f\| = \int_{x \in \mathbb{R}^n} f \cdot dx < \infty.$

- $\text{Meas}(X)$: finite measures over a measurable space $X$,
  $\|\mu\|_{\text{Meas}(X)} = \mu(X).
  \| [R] \| = \text{Meas}(\mathbb{R})$
Building a CCC from cones?

Morphisms in \( \text{Cstab} \): First Idea

Scott-Continuous Functions \( f : B C \rightarrow B D \).

However (Ehrhard Pagani Tasson)

It does not give a CCC
\( \Rightarrow \) We need stronger requirements on morphisms.

Example (Parallel Or)

\[
\text{or} : B_1 \times B_1 \rightarrow B_1 \\
\text{curr(or)} : B_1 \rightarrow (B_1 \rightarrow B_1) \\
x, y \mapsto x + y - xy \\
x \mapsto (y \mapsto x + y - xy)
\]

- \( \text{or} \) is Scott-continuous \( \Rightarrow \) it should be a morphism
- For the category to be CCC, it will require \( \text{curr(or)} \) also a morphism.
  It cannot be so: \( \text{curr(or)} \) is not non-decreasing.
Building a CCC from cones?

Morphisms in **Cstab**: First Idea

Scott-Continuous Functions $f : B C \rightarrow B D$. Since

$$\text{curr(or)}(0) \not \preceq_1 \text{curr(or)}(1) : B 1 \rightarrow (B 1 \rightarrow B 1)$$

**Proof:**

$$\text{curr(or)}(0) = (y \mapsto y)$$

$$\text{curr(or)}(1) = (y \mapsto 1)$$

$\text{curr(or)}(1) - \text{curr(or)}(0) = 1 - y \quad x \mapsto (y \mapsto x + y - xy)$

*not* non-decreasing $\Rightarrow$ *not* in $1 \rightarrow 1$.

- For the category to be CCC, it will require $\text{curr(or)}$ also a morphism.
  - *It cannot be so:* $\text{curr(or)}$ is *not* non-decreasing.
Building a CCC from cones?

**Morphisms in Cstab: First Idea**

Scott-Continuous Functions $f : BC \to BD$.

However (Ehrhard Pagani Tasson)

It does not give a CCC

$\implies$ We need stronger requirements on morphisms.

**Example (Parallel Or)**

$$\text{or} : B1 \times B1 \to B1$$

\[ x, y \mapsto x + y - xy \]

$$\text{curr}(\text{or}) : B1 \to (B1 \to B1)$$

\[ x \mapsto (y \mapsto x + y - xy) \]

- or is Scott-continuous $\implies$ it should be a morphism
- For the category to be CCC, it will require $\text{curr}(\text{or})$ also a morphism.
  - **It cannot be so**: $\text{curr}(\text{or})$ is not non-decreasing.
Girard-Berry Stability

Coherent space

\( X = (|X|, \text{Cl}(X)) \):

- The elements of \( \text{Cl}(X) \) are subset of \(|X|\), that moreover are cliques of an underlying graph.
- The elements of \( \text{Cl}(X) \) are ordered by inclusion.

Definition (Stable function)

\( X, Y \) coherent spaces. \( f : \text{Cl}(X) \to \text{Cl}(Y) \) is stable when:

- \( f \) is not decreasing and Scott-continuous
- Stability condition: if \( x \cup y \in \text{Cl}(X) \), then \( f(x \cap y) = f(x) \cap f(y) \).

Facts

- The interpretation of parallel or is not stable;
- A function \( \text{Cl}(X) \to \text{Cl}(Y) \) is stable if and only if \( \text{Tr}(f) \) is a morphism in \( \text{Coh}_1(X, Y) \).
A reformulation of Girard-Berry’s stability [Ehrhard, Pagani, Tasson]

Definition (Local coherent space)

\( X \) a coherent space, \( u \in \text{Cl}(X) \). \( X_u \): the local PCS at \( u \):
- \( |X_u| = \{ a \in |X| \mid \{a\} \prec X u \} \)
- \( x \prec_{X_u} y \) when \( x \prec X y \).

Proposition (Characterisation of stable functions)

\( f : P(X) \rightarrow P(Y) \) is stable if and only if:

\[
\forall u \in \text{Cl}(X), \text{ } \Delta f_u \text{ is non-decreasing},
\]

where \( \Delta f_u \) is the local difference of \( f \) at \( u \):

\[
\Delta f_u : \text{Cl}(E_u) \rightarrow \text{Cl}(F) \\
 v \mapsto f(u \cup v) - f(v).
\]
Pre-Stable Functions

**Definition (n-pre-stable functions for \( n \in \mathbb{N} \))**

\( f : \mathcal{B}C \rightarrow \mathcal{B}D \) with:

- \( n = 0 \): \( f \) is non-decreasing
- \( n > 0 \): \( \forall u \in \mathcal{B}C, \Delta f_{\mid u} \) is \((n - 1)\) pre-stable.

**Notations:**
- Local Cone at \( u \).
- Local Differences.

**Example (3-pre-stability)**

- \( f(x) \preceq_D f(x + u) \);
- \( f(x + u_1) + f(x + u_2) \preceq_D f(x + u_1 + u_2) + f(x) \);
- \( f(x + u_1 + u_2) + f(x + u_1 + u_3) + f(x + u_2 + u_3) + f(x) \preceq_D f(x + u_1 + u_2 + u_3) + f(x + u_1) + f(x + u_2) + f(x + u_3) \).
The CCC \textbf{Cstab}.

Example (\(\infty\) pre-stable functions)

- \textit{linear} functions;
- every class of functions preserved by \(f \mapsto \Delta f|_u\): e.g. in \(\mathbb{R}^n \to \mathbb{R}\):
  - polynomials with non-negative coefficients
  - power series with non-negative coefficients

Definition (Probabilistic Stable functions \(\mathcal{B}C \to \mathcal{B}D\))

- Scott-continuous.
- \(\infty\) pre-stable

Definition (The CCC \textbf{Cstab})

- Objects: complete cones
- Morphisms: stable functions.
The CCC $\mathbf{Cstab}$.

**Example ($\infty$ pre-stable functions)**
- *Linear* functions;
- every class of functions preserved by $f \mapsto \Delta f|_u$:
  - e.g. in $\mathbb{R}^n \to \mathbb{R}$:
    - polynomials with non-negative coefficients
    - power series with non-negative coefficients

**Definition (Probabilistic Stable functions $\mathcal{B}C \to \mathcal{B}D$)**
- Scott-continuous.
- $\infty$ pre-stable $\Rightarrow$ excludes the Scott-continuous function denoting the parallel or

**Definition (The CCC $\mathbf{Cstab}$)**
- Objects: complete cones
- Morphisms: stable functions.
Connection between $\text{PCoh}_!$ and $\text{Cstab}$. 

**Theorem (C 2018)**

*There exists a functor $F: \text{PCoh}_! \rightarrow \text{Cstab}_m$:*

- *which is full and faithful*
- *and respects the cartesian closed structure.*

**Proof sketch**

- Every PCS can be seen as a complete cone;
- Stable functions on PCSs coincide with power series with non-negative coefficients:
  - uses a result due to McMillan on absolutely monotonous functions on partitions systems
  - from a stable function, build *generalized derivatives* that allows to recover the power series coefficients.
Stable Semantics of Probabilistic Higher-Order Programs

Raphaëlle Crubillé

The Discrete Models of PCSs
- Probabilistic Coherence Spaces
- Proving operational properties on programs denotationally.

The Continuous Stable Model
- Construction of the Continuous Model: An Overview
- Probabilistic Stability
- Adding measurability constraints.
Why do we need additional constraints on stable functions?

Goal:
Sample first, then do a computation.

Example

\[ M := \text{let } x = L \text{ in } N. \]

\[ x : \mathbb{R} \vdash N : \mathbb{R} \Rightarrow [N] : \text{Meas}(\mathbb{R}) \mapsto \text{Meas}(\mathbb{R}). \]

\[ L : \mathbb{R} \Rightarrow [L] : \text{Meas}(\mathbb{R}) \]

We would like the interpretation of \( M \) to be:

\[
[M] \in [\mathbb{R}] = \text{Meas}(\mathbb{R})
= U \in \Sigma_\mathbb{R} \mapsto \int f.d[L]
\]

where \( f : r \in \mathbb{R} \mapsto [N](\delta_r)(U) \in \mathbb{R}. \)

\[\Rightarrow \text{ We need to add constraints to guarantee that this integral makes sense} \]
Measurable Cones

Definition (Measurable cone)

$C$ equipped of a family of random variables $\text{Paths}^n(C) \subseteq \mathbb{R}^n \rightarrow C$, such that:

- For every $\gamma \in \text{Paths}^n(C)$, $\gamma(\mathbb{R}^n)$ is bounded in $C$;
- $\forall \gamma \in \text{Paths}^n(C)$, $f : \mathbb{R}^m \mapsto \mathbb{R}^n$ measurable, $\gamma \circ f \in \text{Paths}^m(C)$;
- $\forall x \in C$, $n \in \mathbb{N}$, $(\vec{r} \in \mathbb{R}^n \mapsto x) \in \text{Paths}^n(C)$.

Example (The Cone $\text{Meas}(X)$ of Bounded Measures on $X$)

Unitary measurable paths: Stochastic Kernels from $\mathbb{R}$ into the measurable space $X$.

Definition (Measurable Functions $f : \mathcal{B}C \rightarrow D$.)

$\forall \gamma \in \text{Paths}^n(C)$ with $\gamma(\mathbb{R}^n) \subseteq \mathcal{B}C$, $f \circ \gamma \in \text{Paths}^n(D)$. i.e. it preserves measurable paths.
Stable Semantics

The $\textbf{Cstab}_m$ category

- Objects: Measurable Cones
- Morphisms: stable measurable functions $\mathcal{B}C \rightarrow \mathcal{B}D$.

Theorem

Ehrhard, Pagani, Tasson $\textbf{Cstab}_m$ is an adequate model of $\text{PCF}_{\text{sample}}$. 
Cstab\textsubscript{m} is a conservative extension of PCoh\textsubscript{!}.

\begin{align*}
\text{PCF} & \quad \text{PCF} \\
\Rightarrow & \quad \Rightarrow \\
\text{Cstab}\textsubscript{m} & \quad \text{PCoh}\textsubscript{!} \\
\downarrow F_m & \quad \downarrow [\cdot] \\
\text{Cstab} & \quad \text{Cstab}
\end{align*}

Theorem (C 2018)

We can extend $F$ into a functor $F_m : \text{PCoh}\textsubscript{!} \rightarrow \text{Cstab}\textsubscript{m}$ that

- is full and faithful,
- preserves the cartesian closed structure.
Conclusion

Probabilistic coherence space models
- full abstraction results
- Express quantitative property on programs (Ehrhard tamed distance)
- genericity of its exponential structure (i.e. it is a Lafont model)

Stable continuous model
A generalization of PCSs where coexist:
- Wild data structures (e.g. bounded measures over any measurable set)
- Very regular morphisms (that can be understood using our result as generalization of analytic functions).
Perspectives

- Find a model of linear logic with $\text{Cstab}_m$ as Kleisli category.
- Extension of $\text{PCoh}_!$ full abstraction proof for PCF$_\oplus$ to $\text{Cstab}_m$ for PCF$_\text{sample}$.
Perspectives

- Find a model of linear logic with $\text{Cstab}_m$ as Kleisli category.
- Extension of $\text{PCoh}_!$ full abstraction proof for $\text{PCF} \oplus$ to $\text{Cstab}_m$ for $\text{PCF}_{\text{sample}}$.

\[
\text{C} \xrightarrow{f \text{ stable function}} 1
\]
Perspectives

- Find a model of linear logic with $\text{Cstab}_m$ as Kleisli category.
- Extension of $\text{PCoh}!$ full abstraction proof for $\text{PCF} \oplus$ to $\text{Cstab}_m$ for $\text{PCF}_{\text{sample}}$.

\[
\begin{array}{ccc}
g \text{ stable function} & f \text{ stable function} \\
\mathbb{N} & \rightarrow & C & \rightarrow & 1
\end{array}
\]
Perspectives

- Find a model of linear logic with $\text{Cstab}_m$ as Kleisli category.
- Extension of $\text{PCoh}_!$ full abstraction proof for $\text{PCF}_\oplus$ to $\text{Cstab}_m$ for $\text{PCF}_{\text{sample}}$. 

```
\begin{tikzpicture}
    \node (N) at (0,0) {$\mathbb{N}$};
    \node (C) at (2,0) {$C$};
    \node (1) at (4,0) {$1$};
    \node (PCoh) at (2,-2) {$\text{PCoh}_!$ -morphism};
    \draw[->] (N) -- (C);
    \draw[->] (C) -- (1);
    \draw[->, dashed] (N) -- (PCoh);
    \draw[->, dashed] (PCoh) -- (C);
    \draw[->, dashed] (PCoh) -- (1);
    \draw[->, dashed] (C) -- (1);
    \node at (1,-1) {g stable function};
    \node at (3,-1) {f stable function};
\end{tikzpicture}
```