A linear logic approach to the semantics of probabilistic programs

joint work with T. Ehrhard and M. Pagani

Christine Tasson
Christine.Tasson@irif.fr

IRIF - University Paris Diderot
Study the *implementation* of probabilistic algorithms with *formal methods*: correctness, termination, contextual behaviour,…

### Bibliography

<table>
<thead>
<tr>
<th>Year</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>Kozen</td>
<td>- Semantics for probabilistic programs</td>
</tr>
<tr>
<td>1989</td>
<td>Jones et al.</td>
<td>- A probabilistic powerdomain of evaluation</td>
</tr>
<tr>
<td>1999</td>
<td>Panangaden</td>
<td>- The category of markov kernel</td>
</tr>
<tr>
<td>2008</td>
<td>Danos et al.</td>
<td>- Probabilistic coherent spaces</td>
</tr>
<tr>
<td>2008</td>
<td>Park et al.</td>
<td>- A probabilistic language based on sampling functions</td>
</tr>
<tr>
<td>2016</td>
<td>Staton et al.</td>
<td>- Semantics for probabilistic programming: higher-order functions, con-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tinuous distributions,</td>
</tr>
<tr>
<td>2018</td>
<td>Ehrhard et al.</td>
<td>- Measurable cones and stable, measurable functions: a model for pro-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>babilistic higher-order programming</td>
</tr>
</tbody>
</table>

**Differences:** CBV or CBN evaluation, Discrete or Continuous data, first or higher order programs.
Semantics of Probabilistic Programs

Operational Semantics: how probabilistic programs compute

The evaluation of a program is a Markov process described by the probability of reduction from $M$ to $N$: $\text{Prob}(M, N)$

- *Discrete type*: stochastic matrix
- *Continuous type*: stochastic kernel

Denotational Semantics: invariant of computation

If $M$ is a closed program, $\lbrack M \rbrack$ can represent the results.

- *Discrete type* ($\mathbb{N}$): discrete distributions over integers
- *Continuous type* ($\mathbb{R}$): continuous distributions over reals
Two examples of Probabilistic Programs

We will prove that the correctness of the implementation of two classic probabilistic algorithms in probability.

Conditioning - handling discrete integers

Given an array containing 0/1 cells, find the index of a 0 cell.

1. choose an index k
2. test if the content of the kth cell is 0
3. if yes output k
4. if no start from 1

Prove that LV outputs a correct value with probability 1
Two examples of Probabilistic Programs

We will prove that the correctness of the implementation of two classic probabilistic algorithms in probability.

**Metropolis Hasting - handling continuous reals**

Simulate a markov chain following a probabilistic law that we know only up to a scaling.

1. Start from a well-chosen point
2. Sample the proposal next point from a gaussian
3. Test if it is coherent with the previous one according to the wanted law up to a scaling
4. if yes use the proposal next point and start from 2
5. if no keep the previous point and start from 2

Prove that MH produces a markov chain following the wanted probabilistic law.
## What tools to study this programs

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Describe programs, types and implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>Describe the evaluation of programs using</td>
</tr>
<tr>
<td>semantics</td>
<td>$\text{Prob}(M, N)$ a stochastic matrix or kernel</td>
</tr>
<tr>
<td>Denotational</td>
<td>Interpret types using mathematical spaces</td>
</tr>
<tr>
<td>semantics</td>
<td>Interpret programs using mathematical functions</td>
</tr>
</tbody>
</table>

### Invariance of semantics

- **Discrete:** $\llbracket M \rrbracket = \sum_N \text{Prob}(M, N)\llbracket N \rrbracket$
- **Continuous:** $\llbracket M \rrbracket = \int \text{Prob}(M, dt)\llbracket t \rrbracket$

### Adequacy

- If $\vdash M : \text{nat}$, then $\llbracket M \rrbracket_n = \text{Prob}(M, n)$
- Lemma: If $\vdash M : \text{real}$, then $\llbracket M \rrbracket(U) = \text{Prob}(M, U)$

### Full Abstraction

- $\llbracket P \rrbracket = \llbracket Q \rrbracket$ iff $P \simeq Q$  
  (Discr.✓ / Cont.✓)

---

C. Tasson

Introduction: Discrete ($\text{Pcoh}$)  Continuous ($\text{Cstab}_m$)  Conclusion 5/38
1 Discrete Probability
- Syntax: Discrete Probabilistic PCF
- Semantics: $\text{Pcoh}$ (Probabilistic Coherent Spaces)
- Results: Probabilistic Adequacy & Full Abstraction

2 Continuous Probability
Syntax of PPCF:

Types: \(A, B ::= \text{nat} | A \to B\)

Terms: \(M, N, L ::= x | \lambda x^A.M | (M)N | \text{fix}(M) | n | \text{succ}(M) | \text{ifz}(L, M, N) | \text{let } x = M \text{ in } N | \text{coin}\)

Operational Semantics as a stochastic process: \(M \xrightarrow{P} N\)

\[
\begin{align*}
(\lambda x^A.M)N & \xrightarrow{1} M[N/x] \\
\text{ifz}(0, M, N) & \xrightarrow{1} M \\
\text{ifz}(n + 1, M, N) & \xrightarrow{1} N \\
\text{let } x = n \text{ in } N & \xrightarrow{1} N[n/x]
\end{align*}
\]
Syntax of PPCF:

**Types:**

\[ A, B ::= \text{nat} | A \to B \]

**Terms:**

\[ M, N, L ::= x | \lambda x^A.M | (M)N | \text{fix}(M) | n | \text{succ}(M) | \text{ifz}(L, M, N) | \text{let } x = M \text{ in } N | \text{coin} \]

Operational Semantics as a stochastic process:

\[ M \xrightarrow{p} N \]

- If \( M \xrightarrow{p} M' \) then
  - \( (M)N \xrightarrow{p} (M')N \)
  - \( \text{let } x = M \text{ in } N \xrightarrow{p} \text{let } x = M' \text{ in } N \)
  - \( \text{succ}(M) \xrightarrow{p} \text{succ}(M') \)
  - \( \text{ifz}(M, L, N) \xrightarrow{p} \text{ifz}(M', L, N), \ldots \)
### Syntax of PPCF:

**Types:**  
\[ A, B ::= \text{nat} \mid A \to B \]

**Terms:**  
\[ M, N, L ::= x \mid \lambda^A.M \mid (M)N \mid \text{fix}(M) \mid n \mid \text{succ}(M) \mid \text{ifz}(L, M, N) \mid \text{let } x = M \text{ in } N \mid \text{coin} \]

---

### Operational Semantics as a stochastic matrix \( \text{Prob} (\cdot, \cdot) \)

\[
\text{Prob}((\lambda^A.M)N, M[N/x]) = 1 : (\lambda^A.M)N \xrightarrow{1} M[N/x]
\]

\[
\text{Prob}(\text{coin}, 0) = \text{Prob}(\text{coin}, 1) = \frac{1}{2} : \begin{array}{c}
\text{coin} \\
\frac{1}{2} \quad \frac{1}{2}
\end{array} \begin{array}{c}
0 \\
1
\end{array}
\]

---

C. Tasson  
Introduction  
Discrete (Pcoh)  
Continuous (Cstab_m)  
Conclusion  
7/38
### Syntax of PPCF:

**Types:**

\[ A, B ::= \text{nat} \mid A \to B \]

**Terms:**

\[ M, N, L ::= x \mid \lambda x^A.M \mid (M)N \mid \text{fix}(M) \mid n \mid \text{succ}(M) \mid \text{ifz}(L, M, N) \mid \text{let } x = M \text{ in } N \mid \text{coin} \]

### Operational Semantics as a stochastic matrix \( \text{Prob}(\cdot, \cdot) \)

- \( \text{Prob}(M, N) \): **probability** that \( M \to N \) in **one** step.
- \( \text{Prob}^2(M, N) \): **probability** that \( M \to N \) in **two** steps.
- \( \ldots \)
- \( \text{Prob}^\infty(M, N) \): **probability** that \( M \to N \) in **any** steps (when \( N \) is a normal form)
Syntax of PPCF:

Types: \( A, B ::= \text{nat} \mid A \rightarrow B \)

Terms: \( M, N, L ::= x \mid \lambda x^A. M \mid (M)N \mid \text{fix}(M) \mid n \mid \text{succ}(M) \mid \text{ifz}(L, M, N) \mid \text{let } x = M \text{ in } N \mid \text{coin} \)

Operational Semantics as a stochastic matrix \( \text{Prob}(\cdot, \cdot) \)

\[
\text{Prob}^2(M, N) = \sum_L \text{Prob}(M, L) \text{Prob}(L, N)
\]

If \( \vdash M : \text{nat} \), then \( \text{Prob}^\infty(M, \_ \_ ) \) is the subprobability discrete distribution over \( \mathbb{N} \) of normal forms of \( M \).
How to encode a LasVegas Algorithm?

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

\[
\begin{array}{cccccc}
0 & 1 & 2 & 3 & 4 & 5 \\
0 & 1 & 0 & 1 & 1 & 0 \\
\end{array}
\]

$f : 0, 2, 5 \mapsto 0$

$1, 3, 4 \mapsto 1$

Output: Find the index of a cell containing 0.

Caml:

```caml
let rec LasVegas = let k = random n in
  if (f k = 0) then k
  else LasVegas
```

pPCF:

```ppcf
fix (\LasVegas : nat. \k : nat.
  ifz f k then k
  else LasVegas) (rand n)
```

C. Tasson
How to encode a LasVegas Algorithm?

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

```
0 1 2 3 4 5
0 1 0 1 1 0
```

$f : \begin{align*}
0, 2, 5 & \mapsto 0 \\
1, 3, 4 & \mapsto 1
\end{align*}$

**Output:** Find the index of a cell containing 0.

**Caml:**

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

**pPCF:**

```
fix (\LasVegas : \k : nat .
    \_ :
    ifz f : \k : nat .
        \_ :
            ifz f k then k
        else LasVegas)
    (rand n))
```
How to encode a LasVegas Algorithm?

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

```
0 1 2 3 4 5
0 1 0 1 1 0
```

$f : 0, 2, 5 \mapsto 0$

$1, 3, 4 \mapsto 1$

**Output:** Find the index of a cell containing 0.

**Caml:**

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

**pPCF:**

```
fix (\text{LasVegas}^{nat}. let k = rand n in
    ifz (f k) then k
    else LasVegas)
```
Syntactical proof of correction of LasVegas

\[ \text{LV} = \text{fix}(\lambda \text{LasVegas}^{\text{nat}}. \text{let } k = \text{rand } n \text{ in } \text{ifz } (f \ k) \text{ then } k \text{ else } \text{LasVegas}) \]

What is the probability LV terminates with a success: \( k \) such that \( f(k) = 0 \):

\[ \frac{1}{2} \quad \text{LV} \xrightarrow{\frac{1}{2}} \text{Success} \]

\[ \frac{1}{2} \]

\[ \frac{1}{2} \quad \text{LV} \xrightarrow{\frac{1}{2}} \text{Success} \]

\[ \frac{1}{2} \quad \text{LV} \xrightarrow{\frac{1}{2}} \text{Success} \]

\[ \text{Prob}^\infty(\text{LV, Success}) = \sum_{k=1}^{\infty} \frac{1}{2^n} = 1 \]
1 Discrete Probability

- Syntax: Discrete Probabilistic PCF
- Semantics: \textbf{Pcoh} (Probabilistic Coherent Spaces)
- Results: Probabilistic Adequacy & Full Abstraction

2 Continuous Probability
<table>
<thead>
<tr>
<th>General Framework</th>
<th>Domains Semantics</th>
<th>Quantitative Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types</td>
<td>Continuous <em>dcpo</em> $(X, \leq)$</td>
<td><em>Proba.</em> spaces $(</td>
</tr>
<tr>
<td>Programs</td>
<td>Scott Continuous</td>
<td>Analytic Functions</td>
</tr>
<tr>
<td>Probability</td>
<td>Proba. <em>monad</em></td>
<td>Values as <em>proba.</em> distr.</td>
</tr>
</tbody>
</table>

### Bibliography

- **1976** Plotkin
- **1981** Kozen
- **1989** Plotkin and Jones
- **1998** Jung and Tix
- **2013** Goubault Larrecq and Varraca
- **2013** Mislove
- **1988** Girard
- **1994** Blute, Panangaden and Seely
- **2002** Hasegawa
- **2004** Girard
- **2011** Danos and Ehrhard
- **2014** Ehrhard, Pagani, T.
- **2016** Ehrahrerd, T.
<table>
<thead>
<tr>
<th>General Framework</th>
<th>Domains Semantics</th>
<th>Quantitative Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types</td>
<td>Continuous dcpos $(X, \leq)$</td>
<td>Proba. spaces $(</td>
</tr>
<tr>
<td>Programs</td>
<td>Scott Continuous</td>
<td>Analytic Functions</td>
</tr>
<tr>
<td>Probability</td>
<td>Proba. monad</td>
<td>Values as proba. distr.</td>
</tr>
</tbody>
</table>

**How to interpret a program** \( M : \text{nat} \Rightarrow \text{nat} \)

**Type:**

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

**Prog:**

\[
[M] : \mathbb{N}_\perp \rightarrow \mathcal{V}(\mathbb{N}_\perp),
[\text{let } n=x \text{ in } M] : \mathcal{V}(\mathbb{N}_\perp) \rightarrow \mathcal{V}(\mathbb{N}_\perp)
\]

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

**Type:**

\[
|\text{Nat}| = \mathbb{N}
\]

\( P(\text{Nat}) \) subproba. dist. over \( \mathbb{N} \)

**Prog:**

\[
[M] : P(\text{Nat}) \rightarrow P(\text{Nat})
\]

\[
x \mapsto \left( \sum_{\mu=[n_1,\ldots,n_k]} \left[ M \right]_{\mu,q} \prod_{i=1}^{k} x_{n_i} \right)_q
\]
<table>
<thead>
<tr>
<th>General Framework</th>
<th>Domains Semantics</th>
<th>Quantitative Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types</td>
<td>Continuous dcpos $(X, \leq)$</td>
<td>Proba. spaces $(</td>
</tr>
<tr>
<td>Programs</td>
<td>Scott Continuous</td>
<td>Analytic Functions</td>
</tr>
<tr>
<td>Probability</td>
<td>Proba. monad</td>
<td>Values as proba. distr.</td>
</tr>
</tbody>
</table>

### Problematic in domain

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

### Full Abs.: PCOH/pPCF

\[
\text{Prob}(C[M], n) \\
\forall n, \forall C[[]] \quad \equiv \\
\text{Prob}(C[N], n) \quad \text{iff} \\
[M] = [N].
\]
Types as **Probabilistic Coherent Spaces**: \((|X|, P(X))\)

**Proba. Space**

- \(|X|\): the **web**, a (potentially infinite) set of final states
- \(P(X)\): a set of vectors \(\subseteq (\mathbb{R}^+)^{|X|}\) such that
  
  **closure:** \(P(X)^\perp\perp = P(X)\) with
  
  \[
  \forall u, v \in (\mathbb{R}^+)^{|X|}, \quad \langle u, v \rangle = \sum_{a \in |X|} u_a v_a
  \]
  
  \[
  \forall P \subseteq (\mathbb{R}^+)^{|X|}, \quad P^\perp = \{v \in (\mathbb{R}^+)^{|X|}; \forall u \in P, \langle u, v \rangle \leq 1\}
  \]

  **bounded covering:** \(\forall a \in |X|,\)
  
  \[
  \exists v \in P(X); \quad v_a \neq 0 \quad \text{and} \quad \exists p > 0; \quad \forall v \in P(X), \quad v_a \leq p.
  \]

**Proposition: Proba. spaces as Domains**

\((|X|, P(X))\) is a **Proba. space iff** \(P(X)\) is bounded covering, **Scott Closed** (downwards-closed and dcpo) and **Convex**.
Types as \textbf{Probabilistic Coherent Spaces}: \(|X|, P(X)\)

\begin{align*}
|1| &= \{\ast\} & P(1) &= [0, 1] \\
|\text{Bool}| &= \{t, f\} & P(\text{Bool}) &= \{(x_t, x_f) ; x_t + x_f \leq 1\} \\
|\text{Nat}| &= \{0, 1, 2, \ldots\} & P(\text{Nat}) &= \{x \in [0, 1]^\mathbb{N} ; \sum_n x_n \leq 1\} \\
|\text{Bool} \rightarrow 1| &= \{[t^n, f^m] ; n, m \in \mathbb{N}\}, & P(\text{Bool} \rightarrow 1) &= \{Q \in (\mathbb{R}^+)_{\text{Bool} \rightarrow 1} ; \} \\
& & \forall x_t + x_f \leq 1, & \sum_{n,m=0}^{\infty} Q[t^n,f^m] x^n_t x^m_f \leq 1 \}
\end{align*}

\textbf{Proposition: Proba. spaces as Domains}

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

C. Tasson

Introduction  Discrete (\textit{Pcoh})  Continuous (\textit{Cstab}_m)  Conclusion
A model of Linear Logic

\[ \text{Pcoh : Linear Category} \]

\begin{itemize}
  \item Objects: Proba. Spaces
  \item Morphisms: Linear Functions
\end{itemize}

\[ \text{Call by Name} \quad A \rightarrow B = !A \circ B \]

\[ \text{Pcoh}!: \text{Kleisli Category} \]

\begin{itemize}
  \item Objects: Proba. Spaces
  \item Morphisms: Analytic Functions
\end{itemize}

\[ \bullet \quad \text{Smcc} \ (1, \otimes, \rightarrow) \]
\[ \bullet \quad \text{biproduct} \]
\[ \bullet \quad \text{Comonad} \ (!, \text{der}, \text{dig}) \]
\[ \bullet \quad \text{Com. Comonoid} \ (!A, 1, \otimes) \]

\[ \bullet \quad \text{CCC} \]
\[ \bullet \quad (\text{PCF}+\text{coin}) \]
Linear Category

**Pcoh**($X, Y$)

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

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

**Example**

**Pcoh(Nat, 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**

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

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

**Example**

Let \( \textbf{Bcoin} = (p, 1 - p) \in \mathcal{P}(\text{Bool}) = \{(p, q) ; p + q \leq 1\} \).

\[ \textbf{Bcoin}^1_{[]} = 1, \quad \textbf{Bcoin}^1_{[t,t]} = p^2, \quad \textbf{Bcoin}^1_{[t,f]} = p(1 - p), \ldots \]

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

This exponential computes the free commutative comonoid.
The set of finite multisets $|!X| = \mathcal{M}_{\text{fin}}(|X|)$ is defined as $P(!X) = \{x^! : x \in P(X)\}$ where $x^!_{[a_1, \ldots, a_k]} = \prod_{i=1}^{k} x_{a_i}$.

**Commutative Comonoid**

- **Cocontr.**: $!X \xrightarrow{c^!X} !X \otimes !X$
- **Coweight**: $!X \xrightarrow{w^!X} 1$

**Comonad**

- **Comult.**: $\text{dig}^! X : !!X \to !X$
- **Counit**: $\text{der}^! X : !X \to X$

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

This exponential computes the free commutative comonoid.
Non-Linear Category

\[
P_{\text{coh}}(X, Y) = P_{\text{coh}}(!X, Y)
\]

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

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

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

\[
P_{\text{coh}}(\text{Bool}, 1) = \{ Q \in (\mathbb{R}^+)_{\text{Bool} \rightarrow 1} \mid s.t. \quad Q_{[t^n, f^m]} \leq \frac{(n+m)^{n+m}}{n^n m^m} \}
\]

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

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

C. Tasson

Introduction  Discrete (Pcoh)  Continuous (Cstab_m)  Conclusion
Non-Linear Category

\[ P_{c}h!(X, Y) = P_{c}h(!X, Y) \]

Density

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

\[ \forall x \in P(X), \; \tilde{Q}(x) = \left( \sum_{m \in M_{\text{fin}}(|X|)} Q_{m,b} x^!_m \right) \in P(Y) \]

Non-Linear Morphisms are \textbf{analytic} and \textbf{Scott Continuous}.

\[ P_{c}h!(\text{Bool}, 1) = \{ Q \in (\mathbb{R}^+)^{|\text{Bool} \rightarrow 1|} \mid \text{s.t. } Q[t^n, f^m] \leq \frac{(n+m)^{n+m}}{n^n m^m} \} \]

denotes

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

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

C. Tasson

Introduction Discrete (\textit{Pcoh}) Continuous (\textit{Cstab}_m) Conclusion 16/38
Non-Linear Category

\[ P_{coh}(X, Y) = P_{coh}(!X, Y) \]

Density

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

\[
\forall x \in P(X), \quad \hat{Q}(x) = \left( \sum_{m \in M_{\text{fin}}(|X|)} Q_{m,b} x^!_{m} \right) \in P(Y)
\]

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

\[ P_{coh}(\text{Bool}, 1) = \{ Q \in (\mathbb{R}^+)^{|\text{Bool}| \rightarrow |1|} \mid s.t. \quad Q[t^n, f^m] \leq \frac{(n+m)^{n+m}}{n^n m^m} \} \]

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

**pb of DEFINABILITY**

\[
\sum_{n,m=0}^{\infty} \frac{(n + m)!}{n! m!} x_t^{2n+1} x_f^{2m+1}
\]
Interpretation of terms

If $\Gamma \vdash M : A$, then $[[A]]_{\Gamma} \in \text{Pcoh}_!(\Gamma, A)$

$\vdash n : \text{nat}$, thus $[[n]] \in \text{P}(\text{Nat})$ is a distribution over $\mathbb{N}$:

$$[[n]] = (0, \ldots, 0, 1^{\leftarrow}, 0, \ldots)$$  

$n$th

$\vdash \text{rand } n : \text{nat}$, thus $[[\text{rand } n]]$ is a distribution over $\mathbb{N}$:

$$[[\text{rand } n]] = \left(\frac{1}{n}, \ldots, \frac{1}{n},^{\leftarrow}, 0, \ldots\right)$$  

$(n - 1)$th

If $\vdash N : \text{nat}$ and $\vdash P : A$ and $\vdash Q : A$, then

$$[[\text{ifz}(N, P, Q)]] = [[N]]_0[[P]] + \sum_{k=0}^{\infty} [[N]]_{k+1}[[Q]]$$

$$[[\text{let } x = N \text{ in } P]] = \sum_{k=0}^{\infty} [[N]]_k[[P]](k)$$
1 Discrete Probability
- Syntax: Discrete Probabilistic PCF
- Semantics: Pcoh (Probabilistic Coherent Spaces)
- Results: Probabilistic Adequacy & Full Abstraction

2 Continuous Probability
First results [Danos-Ehrhard 2011]

Syntax \( \text{pPCF} \)

Operational semantics \( \text{Prob}(M, N) = p \) iff \( M \xrightarrow{p} N \)

stochastic matrix vs. stochastic process

Denotational semantics Types as probabilistic spaces: \( [A] = (|A|, P(A)) \)

Programs as analytic functions:
if \( A \vdash M : B \) then \( \hat{[M]} : P(A) \to P(B) \)

\[
\forall x \in P(A), \forall b \in |B|, \hat{[M]}(x)_b = \sum_{m \in M_{\text{fin}}(|A|)} [M]_{m,b} \prod_{a \in m} x_a^{m(a)}
\]

Compositionality \( \hat{([M]N)}_b = \hat{[M]}(\hat{[N]})_b = \sum_{m} [M]_{m,b} \prod_{a \in m} [N]_a^{m(a)} \)

Invariance of sem. \( [M] = \sum_{N} \text{Prob}(M, N)[N] \)

Adequacy Lemma if \( \vdash M : \text{nat} \), then \( \text{Prob}^{\infty}(M, n) = [M]_n \)
Probabilistic Full Abstraction

**Theorem (2014: Ehrhard - Pagani - T.)**

\[ P_{coh} \]

\[ J_M K = J_N K \]

\[ \iff \]

\[ \text{Adequacy} \]

\[ \text{Full Abstraction} \]

\[ \text{pPCF} \]

\[ M \simeq_o N \]

\[ \forall C \forall n \]

\[ \text{Prob}^\infty (C[M], n) = \text{Prob}^\infty (C[N], n) \]

**Adequacy proof:**

If \( [M] = [N] \) then, \( \text{Prob}^\infty ((C)M, n) = \text{Prob}^\infty ((C)N, n) \)

1. Apply Adequacy Lemma:

\[ \text{Prob}^\infty ((C)M, n) = [((C)M) n] \]

2. Apply Compositionality:

\[ [((C)M) n] = \sum_m [C]_{m,n} \prod_{a \in m} [M]_{a}^{m(a)} = \sum_m [C]_{m,n} \prod_{a \in m} [N]_{a}^{m(a)} = [((C)N) n] \]
**Theorem (2014: Ehrhard - Pagani - T.)**

<table>
<thead>
<tr>
<th>Pcoh</th>
<th>pPCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[[M]] = [[N]]$</td>
<td>$M \simeq o N$</td>
</tr>
<tr>
<td>Adequacy</td>
<td>$\forall C \forall n \forall C[[M]] \forall n = \forall C[[N]] \forall n$</td>
</tr>
<tr>
<td>Full Abstraction</td>
<td>$\Prob^\infty (C[M], n) \neq \Prob^\infty (C[N], n)$</td>
</tr>
</tbody>
</table>

**Full Abstraction Proof:**

1. **By contradiction:** $\exists \alpha \in |\sigma|$, $[[M]]_\alpha \neq [[N]]_\alpha$
2. **Find testing context:** $T_\alpha$ such that $[(T_\alpha)M] \neq [(T_\alpha)N]$
   (context only depends on $\alpha$)
3. **Prove definability:** $T_\alpha \in \text{pPCF}$ using coin and regularity of analytic functions
4. **Apply Adequacy Lemma:**
   $\Prob((T_\alpha)M \rightarrow 0) \neq \Prob((T_\alpha)N \rightarrow 0)$. 
Semantical proof of correction of LasVegas

\[ \text{LV} = \mathbf{fix}(\lambda \text{LasVegas}^{\text{nat}}. \quad \text{let } k = \text{rand } n \text{ in} \]
\[ \quad \text{ifz} (f \, k) \text{ then } k \text{ else } \text{LasVegas} \) \]

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

\[
\begin{array}{cccccc}
0 & 1 & 2 & 3 & 4 & 5 \\
0 & 1 & 0 & 1 & 1 & 0 \\
\end{array}
\]

\( f : \) 0, 2, 5 \( \mapsto \) 0
\( 1, 3, 4 \mapsto 1 \)

**Output:** Find the index of a cell containing 0.

We want to prove that \( \text{Prob}^{\infty}(\text{LV}, \text{Success}) = 1 \)
Semantical proof of correction of LasVegas

$$LV = \textbf{fix} (\lambda \text{LasVegas}^{\text{nat}}. \ \text{let} \ k = \text{rand} \ n \ \text{in} \ \text{ifz} \ (f \ k) \ \text{then} \ k \ \text{else} \ \text{LasVegas})$$

By operational semantics:

$$LV \xrightarrow{1} \text{let} \ k = \text{rand} \ n \ \text{in} \ \text{ifz} \ (f \ k) \ \text{then} \ k \ \text{else} \ LV$$
Semantical proof of correction of LasVegas

\[ \text{LV} = \text{fix}(\lambda \text{LasVegas}^{nat}. \text{let } k = \text{rand } n \text{ in } \begin{cases} f(k) \text{ then } k \text{ else LasVegas} \end{cases}) \]

By operational semantics:

\[ \text{LV} \xrightarrow{1} \text{let } k = \text{rand } n \text{ in } \begin{cases} f(k) \text{ then } k \text{ else LV} \end{cases} \]

By invariance of the semantics and interpretation of let and ifz:

\[ \llbracket \text{LV} \rrbracket_p = \sum_{k=0}^{\infty} \llbracket \text{rand } n \rrbracket_k \llbracket \begin{cases} f(k) \text{ then } k \text{ else LV} \end{cases} \rrbracket_p \]

\[ = \frac{1}{n} \cdot \left( \sum_{f(k) = 0 k < n} \llbracket k \rrbracket_p + \sum_{f(k) \neq 0 k < n} \llbracket \text{LV} \rrbracket_p \right) \]

If \( p < n \) \& \( f(p) = 0 \), then \( \llbracket \text{LV} \rrbracket_p = \frac{1}{n} + \frac{1}{n} \cdot \frac{n}{2} \cdot \llbracket \text{LV} \rrbracket_p \), so \( \llbracket \text{LV} \rrbracket_p = \frac{2}{n} \).

If \( p \geq n \) or \( f(p) \neq 0 \), then \( \llbracket \text{LV} \rrbracket_p = \frac{1}{n} \cdot \frac{n}{2} \cdot \llbracket \text{LV} \rrbracket_p \), so \( \llbracket \text{LV} \rrbracket_p = 0 \).
Semantical proof of correction of LasVegas

\[
LV = \text{fix}(\lambda\text{LasVegas}^{\text{nat}}. \text{let } k = \text{rand } n \text{ in } \text{ifz } (f \ k) \text{ then } k \text{ else } \text{LasVegas})
\]

If \( p < n \) and \( f(p) = 0 \), then \([LV]_p = \frac{2}{n}\), otherwise \([LV]_p = 0\).
Semantical proof of correction of LasVegas

\[ \text{LV} = \textbf{fix}(\lambda \text{LasVegas}^{\text{nat}}. \; \text{let } k = \text{rand } n \ \text{in} \]
\[ \text{ifz } (f \ k) \ \text{then } k \ \text{else} \ \text{LasVegas}) \]

If \( p < n \) and \( f(p) = 0 \), then \( [\text{LV}]_p = \frac{2}{n} \), otherwise \( [\text{LV}]_p = 0 \).

Using Adequacy Lemma, the probability that LV converges:

\[
\text{Prob}^{\infty}(\text{LV}, \text{Success}) = \sum_p \text{Prob}^{\infty}(\text{LV}, p) \\
= \sum_p [\text{LV}]_p \\
= \sum_{p < n} \frac{2}{n} = \frac{n}{2} \cdot \frac{2}{n} \\
= 1
\]
1. **Discrete Probability**

2. **Continuous Probability**
   - Syntax: **Real** Probabilistic PCF
   - Semantics: **Cstab_m** (Cones and Stable measurable functions)
   - Results: **Adequacy**
Nat PPCF

Types: \( A, B ::= \text{nat} | A \to B \)

Terms: \( M, N, L ::= \)
\( x | \lambda x^A.M | (M)N | \text{fix}(M) | \)
\( n | \text{succ}(M) | \)
\( \text{ifz}(L, M, N) | \)
\( \text{coin} | \text{let } x = M \text{ in } N \)

Operational Semantics:
\( \text{Prob}(\text{coin}, 0) = \frac{1}{2} \)

If \( \vdash M : \text{nat} \), \( \text{Prob}^\infty(M, _) \) is
the discrete distribution over \( \mathbb{N} \)
computed by \( M \).
From Discrete to Continuous syntax

<table>
<thead>
<tr>
<th>Nat PPCF</th>
<th>Real PPCF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Types:</strong> $A, B ::= \text{nat}</td>
<td>A \rightarrow B$</td>
</tr>
<tr>
<td><strong>Terms:</strong> $M, N, L ::= x</td>
<td>\lambda x^A.M</td>
</tr>
</tbody>
</table>

**Operational Semantics:**

$\text{Prob}(\text{coin}, 0) = \frac{1}{2}$

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

If $\vdash M : \text{real}$, $\text{Prob}^\infty(M, _) \text{ is the continuous distribution over } \mathbb{R}$ computed by $M$. 
The probability to observe $U$ after at most one reduction step applied to $M$ is $\text{Prob}(M, U)$.

$\text{Prob} : \Lambda^{\Gamma \vdash A} \times \Sigma^{\Lambda_{\Gamma \vdash A}} \rightarrow \mathbb{R}^+$ is a stochastic Kernel, i.e:

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

$\text{Prob}^\infty(M, U)$ is the probability to observe $U$ after any steps.
Operational Semantics: the kernel of terms

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

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

$\text{Prob} : \Lambda^{\Gamma \vdash A} \times \Sigma_{\Lambda^{\Gamma \vdash A}} \to \mathbb{R}^+$ is a stochastic Kernel, i.e:

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

$\text{Prob}^\infty(M, U)$ is the probability to observe $U$ after any steps.
The probability to observe $U$ after at most one reduction step applied to $M$ is $\text{Prob}(M, U)$.

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

$\sum_{A \Gamma \vdash A}$, i.e. $U$ is measurable:

$\forall n, \forall S, \{ \vec{r} \mid S\vec{r} \in U \}$ meas. in $\mathbb{R}^n$

$\text{Prob} : \Lambda \Gamma \vdash A \times \sum_{A \Gamma \vdash A} \rightarrow \mathbb{R}^+$ is a stochastic Kernel, i.e:

- for all $M \in \Lambda \Gamma \vdash A$, $\text{Prob}(M, \_)$ is a measure;
- for all $U \in \sum_{A \Gamma \vdash A}$, $\text{Prob}(\_, U)$ is a measurable function.

$\text{Prob}^\infty(M, U)$ is the probability to observe $U$ after any steps.
The probability to observe $U$ after at most one reduction step applied to $M$ is $\text{Prob}(M, U)$.

$\Lambda^{\Gamma \vdash A}$: the set of terms $M$ s.t. $\Gamma \vdash M : A$. $\Sigma^{\Lambda^{\Gamma \vdash A}}$, i.e. $U$ is measurable:

$\forall n, \forall S, \{\bar{r} | S\bar{r} \in U\}$ meas. in $\mathbb{R}^n$

$\text{Prob} : \Lambda^{\Gamma \vdash A} \times \Sigma^{\Lambda^{\Gamma \vdash A}} \rightarrow \mathbb{R}^+$ is a stochastic **Kernel**, i.e:

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

**Measurable sets and kernels constitute the category **Kern**.**

$\text{Prob}^\infty(M, U)$ is the probability to observe $U$ after any steps.
The probability to observe $U$ after at most one reduction step applied to $M$ is $\text{Prob}(M, U)$.

$\Lambda^{\Gamma \vdash A}$: 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 $\mathbb{R}^n$

$\text{Prob} : \Lambda^{\Gamma \vdash A} \times \Sigma_{\Lambda^{\Gamma \vdash A}} \to \mathbb{R}^+$ is a stochastic Kernel, i.e:

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

Measurable sets and kernels constitute the category $\text{Kern}$.

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

It is computed by composition and lub.
The Bernoulli distribution takes the value 1 with probability $p$ and the value 0 with probability $1 - p$.

$$p\delta_1 + (1 - p)\delta_0$$

\[
\text{beroulli } p ::= \text{let } x = \text{sample in } x \leq p
\]

tests if sample draws a value within $[0, p]$.

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

\[
\text{exp ::= let } x = \text{sample in } -\log(x)
\]

by the inversion sampling method.

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

\[
\text{gauss ::= let } x = \text{sample in let } y = \text{sample in } \sqrt{-2\log(x)} \cos(2\pi y)
\]

by the Box Muller method.
Conditioning: If $U \subseteq \mathbb{R}$ measurable, then observe$(U)$ of type real $\rightarrow$ real, taking a term $M$ and returning the renormalization of the distribution of $M$ on the only samples that satisfy $U$: conditioning by rejection sampling.

$$\text{observe}(U) = \lambda m. \text{fix}\left(\lambda y. \text{let } x = \min\left(\min\left(\min(\min(x \in U, x, y))\right)\right)\right)$$

Monte Carlo Simulation,...
How to encode Metropolis Hasting

**Input:** $\mu$ a distribution on $\mathbb{R}$ with density $\pi$:

$$\mu(U) = \int_U \pi(x) dx,$$

but we only know $\gamma \pi$.

**Output:** Markov Chain $x_n$ converging to a random variable $x$ with law $\mu$

1. Initialized $x$ with a well-chosen point $x_0$
2. Sample $y$ from a Gaussian $\text{gauss}$
3. Compute $\alpha(x, y) = \min(1, \frac{\pi(y)}{\pi(x)})$
4. With probability $\alpha(x, y)$, update $x := y$
5. With probability $1 - \alpha(x, y)$, keep $x$
How to encode Metropolis Hasting

**Input:** \( \mu \) a distribution on \( \mathbb{R} \) with density \( \pi \):
\[
\mu(U) = \int_U \pi(x)dx,
\]
but we only know \( \gamma \pi \).

**Output:** Markov Chain \( x_n \) converging to
a random variable \( x \) with law \( \mu \)

\[
\text{MH} = \text{fix}(\lambda \text{MetHast}^{\text{nat} \to \text{nat}}. \lambda n^{\text{nat}}. \text{if } n=0 \text{ then } x_0 \text{ else }
\]
\[
\text{let } x = \text{MetHast} (n-1) \text{ in }
\]
\[
\text{let } y = \text{gauss } x \text{ in }
\]
\[
\text{let } z = \text{bernouilli}(\alpha(x,y)) \text{ in }
\]
\[
\text{if } z = 0 \text{ then } x \text{ else } y)
\]
1. **Discrete Probability**

2. **Continuous Probability**
   - Syntax: **Real** Probabilistic PCF
   - Semantics: **Cstab_m** (*Cones and Stable measurable functions*)
   - Results: **Adequacy**
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 dcpo's equipped with a convex structure (basic operations being Scott continuous) with Scott continuous functions  
How to restrict to measurable functions?
Semantical needs

<table>
<thead>
<tr>
<th>Discrete</th>
<th>Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>If ⊢ M : nat, then ⌊M⌋ is a</td>
<td>If ⊢ M : real, then ⌊M⌋ is a</td>
</tr>
<tr>
<td>distribution over ℕ</td>
<td>measure over ℝ</td>
</tr>
</tbody>
</table>

- ⌊real⌋ as Meas(ℝ) the set of measures over ℝ.
- Fixpoint of terms.

Cstab\textsubscript{m} is a CCC based on Selinger’s cones (dcpoś 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\textsubscript{m} which is a subcategory of Kegelspitzzen.
Our purpose is to be able to interpret \( \text{real} \) 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** (\( \infty \)-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. **\( \text{Cstab}_m \) is a CCC with measurability included**!
From Discrete to Continuous semantics

\[ \text{Pcoh}_! \]

- For \( \vdash n : \mathbb{N} \),
  \[ \llbracket n \rrbracket_p = \delta_{p,n} \]

- For \( \vdash \text{coin} : \mathbb{N} \),
  \[ \llbracket \text{coin} \rrbracket_p = \frac{1}{2} \delta_{0,p} + \frac{1}{2} \delta_{1,p} \]

- For \( \vdash N : \mathbb{N}, \vdash P : A, \vdash Q : A \),
  \[ \llbracket \text{ifz}(N, P, Q) \rrbracket_a = \llbracket N \rrbracket_0[\llbracket P \rrbracket_a + \sum_{n \neq 0} \llbracket N \rrbracket_{n+1}[\llbracket Q \rrbracket_a] ] \]

  \[ \llbracket \text{let } x = N \text{ in } P \rrbracket_a = \sum_{n=0}^{\infty} \llbracket N \rrbracket_n [\llbracket P \rrbracket](n)_a \]
From Discrete to Continuous semantics

\[ P_{\text{coh}} \]

- For \( \vdash n : \mathbb{N} \),
  \[ [n]_p = \delta_{p,n} \]

- For \( \vdash \text{coin} : \mathbb{N} \),
  \[ [\text{coin}]_p = \frac{1}{2} \delta_{0,p} + \frac{1}{2} \delta_{1,p} \]

- For \( \vdash N : \mathbb{N}, \vdash P : A, \vdash Q : A \),
  \[ [\text{ifz}(N, P, Q)]_a = [N]_0[P]_a + \sum_{n \neq 0} [N]_n[Q]_a \]
  \[ [\text{let } x = N \text{ in } P]_a = \sum_{n=0}^{\infty} [N]_n[P](n)_a \]

\[ C_{\text{stab}} \]

- For \( \vdash r : \text{real} \),
  \[ [r](U) = \delta_r(U) \]

- For \( \vdash \text{sample} : \text{real} \),
  \[ [\text{sample}] = \lambda_{[0,1]}(U) \]

- For \( \vdash R : \text{real}, \vdash P, Q : A \),
  \[ [\text{ifz}(R, P, Q)](U) = [R](\{0\})[P](U) + [R](\mathbb{R}\setminus\{0\})[Q](U) \]
  \[ [\text{let } x = R \text{ in } P](U) = \int [R](dr)[P](\delta_r)(U) \]
1 Discrete Probability

2 Continuous Probability

- Syntax: **Real** Probabilistic PCF
- Semantics: **Cstab_m** (Cones and Stable measurable functions)
- Results: **Adequacy**
The category $\text{Cstab}_m$ is a CCC and a model of Real PPCF.

Invariance of the semantics

$$\left[ M \right]^{\Gamma \vdash A} = \int_{\Lambda^{\Gamma \vdash A}} \left[ t \right]^{\Gamma \vdash A} \text{Prob}(M, dt)$$

Adequacy

$$\left[ M \right] \vdash \text{real}(U) = \text{Prob}^\infty(M, U)$$

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

\[ p\delta_1 + (1 - p)\delta_0 \]

\[ [\operatorname{bernoulli} p] \dashv \mathbb{R} = p\delta_1 + (1 - p)\delta_0 \]

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

\[ \operatorname{exp} : \mathbb{R} := \text{let } x = \text{sample in } -\log(x) \]

\[ [\operatorname{exp}] \dashv \mathbb{R} (U) = \int_{\mathbb{R}^+} \chi_U(s)e^{-s}\lambda(ds) \]

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

\[ \operatorname{gauss} := \text{let } x = \text{sample in let } y = \text{sample in } \sqrt{-2\log(x)} \cos(2\pi y) \]

\[ [\operatorname{gauss}] \dashv \mathbb{R} (U) = \frac{1}{\sqrt{2\pi}} \int_{U} e^{-\frac{x^2}{2}} \lambda(dx) \]
Conditioning: If $U \subseteq \mathbb{R}$ measurable, then $\text{observe}(U)$ of type \( \text{real} \rightarrow \text{real} \), taking a term $M$ and returning the renormalization of the distribution of $M$ on the only samples that satisfy $U$:

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

conditioning by rejection sampling.

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

$$\left[ \text{observe}(U)M \right](V) = \frac{[M](V \cap U)}{[M](U)}$$
How to encode Metropolis Hasting

**Input:**  
$\mu$ a distribution on $\mathbb{R}$ with density $\pi$:  
$\mu(U) = \int_U \pi(x) dx$, but we only know $\gamma \pi$.

**Output:**  
Markov Chain $x_n$ converging to  
a random variable $x$ with law $\mu$

1. Initialized $x$ with a well-chosen point $x_0$
2. Sample $y$ from a gaussian centered on $x$
3. Compute $\alpha(x, y) = \min(1, \frac{\pi(y)}{\pi(x)})$
4. With probability $\alpha(x, y)$, update $x := y$
5. With probability $1 - \alpha(x, y)$, keep $x$
How to encode Metropolis Hasting

**Input:** μ a distribution on $\mathbb{R}$ with density $\pi$:
$\mu(U) = \int_U \pi(x) \, dx$, but we only know $\gamma \pi$.

**Output:** Markov Chain $x_n$ converging to a random variable $x$ with law $\mu$

$$MH = \text{fix}(\lambda \text{MetHast}^{\text{nat} \to \text{nat}}. \lambda n^{\text{nat}}. \text{if } n=0 \text{ then } x_0 \text{ else}
\text{let } x = \text{MetHast} (n-1) \text{ in}
\text{let } y = \text{gauss } x \text{ in}
\text{let } z = \text{bernouilli}(\alpha(x,y)) \text{ in}
\text{if } z = 0 \text{ then } x \text{ else } y)$$
How to encode Metropolis Hasting

\[
\text{MH} = \text{fix}(\lambda \text{MetHast}^{\text{nat}\to\text{nat}}. \lambda n^{\text{nat}}. \text{if } n=0 \text{ then } x_0 \text{ else let } x = \text{MetHast}(n-1) \text{ in}
\]

\[
\text{let } y = \text{gauss } x \text{ in}
\]

\[
\text{let } z = \text{bernouilli(}\alpha(x, y)) \text{ in}
\]

\[
\text{if } z = 0 \text{ then } x \text{ else } y)
\]

\[
\text{MH(0)} \rightarrow x_0 \text{ thus, } \text{Prob(MH(0), } U) = \delta_{x_0}(U)
\]

\[
\text{MH}(n+1) \rightarrow M = \text{let } x = \text{MH}(n) \text{ in let } y = \text{gauss } x \text{ in}
\]

\[
\text{let } z = \text{bernoulli(}\alpha(x, y)) \text{ in if } z(x, y) \text{ then } x \text{ else } y
\]
How to encode Metropolis Hasting

\[ MH = \text{fix}(\lambda \text{MetHast}^{\text{nat}} \to \text{nat}. \lambda n^{\text{nat}}. \text{if } n=0 \text{ then } x_0 \text{ else let } x = \text{MetHast} (n-1) \text{ in let } y = \text{gauss} x \text{ in let } z = \text{bernoulli}(\alpha(x,y)) \text{ in if } z = 0 \text{ then } x \text{ else } y) \]

MH(0) \rightarrow x_0 \text{ thus, } \text{Prob}(MH(0), U) = \delta_{x_0}(U)

MH(n+1) \rightarrow M = \text{let } x=MH(n) \text{ in let } y=\text{gauss} x \text{ in let } z=\text{bernoulli}(\alpha(x,y)) \text{ in if } z(z, x, y)

\text{Prob}(MH(n+1), U) = [MH(n+1)](U) = [M](U) \text{ (Adequacy/Reduction)}

= \int_R \left[ N \right] (\delta_r)(U) \left[ MH(n) \right](dr) = \int_R P_{MH}(r, U) \text{Prob}(MH(n), dr)

P_{MH}(r, U) = \delta_r(U) \left( 1 - \int_R \alpha(r, t)g(t, r)\lambda(dt) \right) + \int_U \alpha(r, t)g(t, r)\lambda(dt).
How to encode Metropolis Hasting

**Input:** $\mu$ a distribution on $\mathbb{R}$ with density $\pi$:
$\mu(U) = \int_U \pi(x) dx$, but we only know $\gamma \pi$.

**Output:** Markov Chain $x_n$ converging to a random variable $x$ with law $\mu$

$$
\text{Prob}(\text{MH}(n+1), U) = \int_{\mathbb{R}} P_{\text{MH}}(r, U) \text{Prob}(\text{MH}(n), dr),
$$

$$
P_{\text{MH}}(r, U) = \delta_r(U) \left(1 - \int_{\mathbb{R}} \alpha(r, t) g(t, r) \lambda(dt)\right) + \int_U \alpha(r, t) g(t, r) \lambda(dt).
$$

This shows that $x_n$ is a Markov-Chain whose law is defined with respect to the kernel $P_{\text{MH}}(r, U)$. It is standard mathematics to prove that $\mu$ is its invariant measure.
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.
Why can we use CBV in CBN?

### Storage Operator

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

### Integer in Pcoh:

\[ \text{[nat]} = \text{Nat} = (\mathbb{N}, P(\text{Nat}) = \{ (\lambda_n) | \sum_n \lambda_n \leq 1 \}) \]

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

- **Cocontraction**: \( c^{\text{nat}} : \text{nat} \to \text{nat} \otimes \text{nat} \)
- **Coweakening**: \( w^{\text{nat}} : \text{nat} \to 1 \)

### 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.
The Eilenberg Moore Category: $\text{Pcoh}$! 

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

Coalgebras have a comonoid structure: values can be stored.

Types interpreted as coalgebras:

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

Example

- **Stream:** $S_\varphi = \varphi \otimes !S_\varphi$  
  - **List:** $\lambda_0 = 1 \oplus (\varphi \otimes \lambda_0)$
Probabilistic Call By Push Value

**Types:**

\[
A ::= \text{UB} \mid A_1 \oplus A_2 \mid 1 \mid A_1 \otimes A_2 \mid \alpha \mid \text{Fix} \alpha \cdot A
\]

Example of natural numbers:
\[
\text{nat} ::= \text{Fix} \alpha \cdot 1 \oplus \alpha
\]

**Terms:**

\[
V ::= x \mid \text{thunk}(M) \mid \text{in}_i V \mid () \mid (V, W)
\]

\[
M ::= \text{return}(V) \mid \text{force}(M) \mid \lambda x^A M \mid \langle M \rangle V \mid \text{fix}(M) \mid \text{coin} \mid \text{case}(M, x_1 \cdot N_1, x_2 \cdot N_2) \mid n \mid \text{succ}(V) \mid \text{let } x = V \text{ in } M \mid \text{ifz}(V, M, N)
\]
Probabilistic Call By Push Value

Types: \( !B \)

| (Value) | \( A ::= UB \mid A_1 \oplus A_2 \mid 1 \mid A_1 \otimes A_2 \mid \alpha \mid \text{Fix} \alpha \cdot A \) |

Example of natural numbers: \( \text{nat} ::= \text{Fix} \alpha \cdot 1 \oplus \alpha \)

| (Computation) | \( B ::= FA \mid A \multimap B \) |

Terms:

| (Value) | \( V ::= x \mid \text{thunk}(M) \mid \text{in}_{i}V \mid () \mid (V, W) \) |

| (Computation) | \( M ::= \text{return}(V) \mid \text{force}(M) \) |

| \( | \lambda x^A M \mid \langle M \rangle V \mid \text{fix}(M) \) |

| \( | \text{coin} \mid \text{case}(M, x_1 \cdot N_1, x_2 \cdot N_2) \) |

| \( | n \mid \text{succ}(V) \mid \text{let } x = V \text{ in } M \mid \text{ifz}(V, M, N) \) |
### Probabilistic Call By Push Value

**Types:**

\[
!B
\]

\[
A ::= UB | A_1 \oplus A_2 | 1 | A_1 \otimes A_2 | \alpha | \text{Fix } \alpha \cdot A
\]

**Example of natural numbers:**

\[
nat ::= \text{Fix } \alpha \cdot 1 \oplus \alpha
\]

**Terms:**

\[
V ::= x | \text{thunk}(M) | \text{in}_i V | () | (V, W)
\]

\[
M ::= \text{return}(V) | \text{force}(M)
\]

\[
| \lambda x^A M | \langle M \rangle V | \text{fix}(M)
\]

\[
| \text{coin} | \text{case}(M, x_1 \cdot N_1, x_2 \cdot N_2)
\]

\[
| n | \text{succ}(V) | \text{let } x = V \text{ in } M | \text{ifz}(V, M, N)
\]
# Probabilistic Call By Push Value

### Types:

- \( \!B \)

### (Value)

- \( A ::= UB | A_1 \oplus A_2 | 1 | A_1 \otimes A_2 | \alpha | \text{Fix} \alpha \cdot A \)

### Example of natural numbers:

- \( \text{nat} ::= \text{Fix} \alpha \cdot 1 \oplus \alpha \)

### (Computation)

- \( B ::= FA | A \leadsto B \)

### Forget:

- \( A \)

### Terms:

### (Value)

- \( V ::= x | \text{thunk}(M) | \text{in}_i V | () | (V, W) \)

### (Computation)

- \( M ::= \text{return}(V) | \text{force}(M) \)

- \( \lambda x^A M | \langle M \rangle V | \text{fix}(M) \)

- \( \text{coin} | \text{case}(M, x_1 \cdot N_1, x_2 \cdot N_2) \)

- \( n | \text{succ}(V) | \text{let } x = V \text{ in } M | \text{ifz}(V, M, N) \)
Probabilistic Call By Push Value

Types: \( \!B \)

| (Value) | \( A ::= UB | A_1 \oplus A_2 | 1 | A_1 \otimes A_2 | \alpha | \text{Fix} \alpha \cdot A \) |

Example of natural numbers: \( \text{nat} ::= \text{Fix} \alpha \cdot 1 \oplus \alpha \)

| (Computation) | \( B ::= FA | A \rightarrow B \) Forget: \( A \) |

Terms:

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

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

\[ | \lambda x^A M | \langle M \rangle V | \text{fix}(M) \]

\[ | \text{coin} | \text{case}(M, x_1 \cdot N_1, x_2 \cdot N_2) \]

\[ | n | \text{succ}(V) | \text{let } x = V \text{ in } M | \text{ifz}(V, M, N) \]
The Eilenberg Moore category and the Linear Category

Dense coalgebra

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

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

The Eilenberg Moore category \( \text{Pcoh}^! \)

**Value Types** are interpreted as **dense** coalgebras

**Values** are morphisms of coalgebras

The Linear category \( \text{Pcoh} \)

**Computation Types** are interpreted in \( \text{Pcoh} \)

**Computations** are linear morphisms in \( \text{Pcoh} \)
## Probabilistic Full Abstraction

**Theorem (2016: Ehrhard - T.)**

<table>
<thead>
<tr>
<th>Pcoh</th>
<th>pCBPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[M] = [N]$</td>
<td>$M \simeq_o N$</td>
</tr>
<tr>
<td>Adequacy</td>
<td></td>
</tr>
<tr>
<td>Full Abstraction</td>
<td></td>
</tr>
</tbody>
</table>

### 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

Probabilistic Full Abstraction

Theorem (2016: Ehrhard - T.)

\[ [M] = [N] \quad \Rightarrow \quad \text{Adequacy} \quad \quad \text{Full Abstraction} \]

\[ \text{Full Abstraction} \quad \quad \quad \Rightarrow \quad \text{pCBPV} \]

\[ \forall C \quad \text{Prob}(C[M], ()) \equiv \text{Prob}(C[N], ()) \]

Full Abstraction Proof:

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

\[
x + y = 0 \Rightarrow x, y = 0, \quad \|x + x'\|_P \leq \|x\|_P + \|x'\|_P, \quad \|\alpha x\|_P = \alpha \|x\|_P
\]

\[
x + y = x + y' \Rightarrow y = y', \quad \|x\|_P = 0 \Rightarrow x = 0, \quad \|x\|_P \leq \|x + x'\|_P
\]

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

**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 $BP$ 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.
- $\hat{\mathcal{X}} = \{u \in (\mathbb{R}^+) | \exists \varepsilon > 0 \ \varepsilon u \in P\mathcal{X} \}$ if $\mathcal{X} \in \text{Pcoh}$. 
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 $f : B^P \to Q$ is **n-non-decreasing function** if:

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

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

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

**Weak Parallel Or**

$\text{wp} : [0, 1] \times [0, 1] \to [0, 1]$ given as $\text{wp}(s, t) = s + t - st$ is Scott-continuous, but not Stable. Its currying is not Scott-continuous.
Step 3: The Measurability Problem

**Type** \( \text{real} \) is interpreted as \( \llbracket \text{real} \rrbracket = \text{Meas}(\mathbb{R}) \),

**Closed term** \( \vdash M : \text{real} \) as a measure \( \mu \) and

**Term** \( x : \text{real} \vdash N : \text{real} \) as a stable \( f : \text{Meas}(\mathbb{R}) \rightarrow \text{Meas}(\mathbb{R}) \).

**Operational semantics**

\[
\forall r, \text{s.t. } M \rightarrow r, \text{ let } x = M \text{ in } N \rightarrow N\{r/x\}
\]

By **Soundness**

\[
\llbracket \text{let } x = M \text{ in } N \rrbracket = \int_{\mathbb{R}} (f \circ \delta)(r) \mu(dr)
\]
Step 3: The Measurability Problem

Type real is interpreted as \([\text{real}] = \text{Meas}(\mathbb{R})\),
Closed term \(\vdash M : \text{real}\) as a measure \(\mu\) and
Term \(x : \text{real} \vdash N : \text{real}\) as a stable \(f : \text{Meas}(\mathbb{R}) \to \text{Meas}(\mathbb{R})\).

Operational semantics

\[\forall r, \text{ s.t. } M \to r, \ \text{let } x = M \text{ in } N \to N\{r/x\}\]

By Soundness

\[\llbracket \text{let } x = M \text{ in } N \rrbracket = \int_\mathbb{R} (f \circ \delta)(r) \ \mu \ (dr)\]

\[\llbracket N \rrbracket\]
Step 3: The Measurability Problem

**Type** $\text{real}$ is interpreted as $[\text{real}] = \text{Meas}(\mathbb{R})$,

**Closed term** $\vdash M : \text{real}$ as a measure $\mu$ and

**Term** $x : \text{real} \vdash N : \text{real}$ as a stable $f : \text{Meas}(\mathbb{R}) \rightarrow \text{Meas}(\mathbb{R})$.

**Operational semantics**

$$\forall r, \text{ s.t. } M \rightarrow r, \text{ let } x = M \text{ in } N \rightarrow N\{r/x\}$$

By **Soundness**

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

$\llbracket N \rrbracket$ Dirac measure
Step 3: The Measurability Problem

**Type** real is interpreted as \([\text{real}] = \text{Meas}(\mathbb{R})\),
**Closed term** \(\vdash M : \text{real}\) as a measure \(\mu\) and
**Term** \(x : \text{real} \vdash N : \text{real}\) as a stable \(f : \text{Meas}(\mathbb{R}) \rightarrow \text{Meas}(\mathbb{R})\).

**Operational semantics**

\[ \forall r, \text{ s.t. } M \rightarrow r, \text{ let } x = M \text{ in } N \rightarrow N\{r/x\} \]

By **Soundness**

\[ [\text{let } x = M \text{ in } N] = \int_{\mathbb{R}} (f \circ \delta)(r) \mu(dr) \]

\(\int_{\mathbb{R}}\) Dirac measure \([M]\)
Step 3: The Measurability Problem

**Type** real is interpreted as $[\text{real}] = \text{Meas}(\mathbb{R})$,

**Closed term** $\vdash M : \text{real}$ as a measure $\mu$ and

**Term** $x : \text{real} \vdash N : \text{real}$ as a stable $f : \text{Meas}(\mathbb{R}) \rightarrow \text{Meas}(\mathbb{R})$.

**Operational semantics**

$$\forall r, \text{ s.t. } M \rightarrow r, \text{ let } x = M \text{ in } N \rightarrow N\{r/x\}$$

By **Soundness**

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

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 Meas(\(\mathbb{R}\)) are given by measurable sets of \(\mathbb{R}\):

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

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

Measurable Cone

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

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

\ell \in M^n(P) \text{ and } x \in P \Rightarrow \left\{ \begin{array}{c}
\mathbb{R}^n \\
\mathbb{R}^+
\end{array} \right\} \xrightarrow{\ell(\vec{r})} \ell(\vec{r})(x) \text{ measurable.}
**Cstab**}_m \text{ is the category of complete and measurable cones with stable and measurable functions.}

Let $P$ and $Q$ be measurable and complete cones:

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

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

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

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

$$\forall n \in \mathbb{N}, \forall \gamma \in \text{Path}^n_1(P), \quad f \circ \gamma \in \text{Path}^n(Q)$$

If $X$ is a measurable space, then $\text{Meas}(X)$ is equipped with:

$M^n(X) = \{ \varepsilon_U : \mathbb{R}^n \to \text{Meas}(X)' \text{ s.t. } \varepsilon_U(\vec{r})(\mu) = \mu(U), \ U \text{ meas.} \}$

$\text{Path}^n_1(P)$ is the set of stochastic kernels from $\mathbb{R}^n$ to $X$. 