Stability in a probabilistic setting

TLT 2017, Torino

Thomas Ehrhard, IRIF, CNRS and Univ Paris Diderot

September 25, 2017

Joint work with Michele Pagani and Christine Tasson.

Denotational semantics and intersection types are deeply related.

For instance, the points of the model **Rel** (sets and relations) of Linear Logic can be interpreted as (non-idempotent) intersection types.

This remains true of all models "based on **Rel**": coherence spaces, hypercoherences, Scott semantics in prime-algebraic complete lattices.

Typically

$$\frac{\Gamma_0 \vdash M : [a_1, \ldots, a_n] \multimap b \qquad (\Gamma_i \vdash M : a_i)_{i=1}^{n}}{\sum_{i=0}^{n} \Gamma_i \vdash M\,N : b}$$

where contexts assign multisets of types to variables, and $(\Gamma + \Delta)(x) = \Gamma(x) + \Delta(x)$ (sum of multisets).

This can be extended to probabilistic programming languages, using the model of Probabilistic Coherence Spaces (PCS) (Girard, Danos and E.).

$$\Gamma \vdash_\alpha M : a \quad \alpha \in \mathbb{R}^+$$

Typically, if M is a probabilistic PCF closed term of ground type (integers, booleans), then α is the probability of M to reduce to value a.

NB: this simple intuition is lost at higher type, where we can have $\alpha > 1$.

PCS duality

Let I be a finite or countable set.

If $u, u' \in (\mathbb{R}^+)^I$, set

$$\langle u, u' \rangle = \sum_{i \in I} u_i u'_i \in \overline{\mathbb{R}^+}$$

and if $\mathcal{F} \subseteq (\mathbb{R}^+)^I$ then

$$\mathcal{F}^\perp = \{u' \in (\mathbb{R}^+)^I \mid \forall u \in \mathcal{F}\ \langle u, u' \rangle \leq 1\}\,.$$

PCS: $X = (|X|, \mathrm{P}X)$ where $|X|$ countable and $\mathrm{P}X \subseteq (\mathbb{R}^+)^{|X|}$ such that $\mathrm{P}X = \mathrm{P}X^{\perp\perp}$.

Additional conditions to avoid ∞ coefficients:

- $\forall a \in |X| \, \exists x \in \mathrm{P}X \; x_a \neq 0$
- $\forall a \in |X| \; \{x_a \mid x \in \mathrm{P}X\}$ bounded

Just as ordinary coherence spaces, PCS (with suitable linear morphisms generalizing stochastic matrices)

- are a model of full classical LL
- with fixpoints (hence a model of PCF etc)
- with fixpoints of types (hence contain various models of the pure lambda-calculus, of FPC, of CBPV with recursive types etc)
- 2nd order LL, polymorphism? (never explored)

All these languages extended with probabilistic primitives, for instance: random integers in a given range, (fair) coin etc.

Examples of type constructions:

- $|1| = \{*\}$, $\mathsf{P}1 = [0,1]$
- $|X \mathbin{\&} Y| = |X| + |Y|$ and
  $\mathsf{P}(X \mathbin{\&} Y) = \{x \oplus y \mid x \in \mathsf{P}X \text{ and } y \in \mathsf{P}Y\} \simeq \mathsf{P}X \times \mathsf{P}Y$
- $|X \oplus Y| = |X| + |Y|$ and
  $\mathsf{P}(X \oplus Y) = \{px \oplus (1-p)y \mid x \in \mathsf{P}X,\ y \in \mathsf{P}Y \text{ and } p \in [0,1]\}$

So $|1 \oplus 1| = \{\mathbf{t}, \mathbf{f}\}$ and
$\mathsf{P}(1 \oplus 1) = \{p\mathbf{t} + q\mathbf{f} \mid p, q \in \mathbb{R}^+, p + q \leq 1\}$

$\mathsf{N} = 1 \oplus \mathsf{N}$ ("least" fixpoint) so that $|\mathsf{N}| = \mathbb{N}$ and PN is the set of sub-probability distributions on $\mathbb{N}$.

Similarly one defines probabilistic types of lists, trees, streams etc.

Morphisms in the CCC of PCS

By standard LL/categorical considerations we know that we have a CCC: the "Kleisli category" of the "!" comonad that we have not described.

One defines $X \Rightarrow Y$ by

- $|X \Rightarrow Y| = \mathcal{M}_{\mathrm{fin}}(|X|) \times |Y|$
- If $u \in (\mathbb{R}^+)^{|X|}$ and $m \in \mathcal{M}_{\mathrm{fin}}(|X|)$, set $u^m = \prod_{a \in |X|} u_a^{m(a)}$
- Then $t \in (\mathbb{R}^+)^{|X \Rightarrow Y|}$ is in $\mathrm{P}(X \Rightarrow Y)$ if for all $u \in \mathrm{P}X$,

$$\widehat{t}(u) = \left(\sum_{m \in \mathcal{M}_{\mathrm{fin}}(|X|)} t_{m,b} u^m \right)_{b \in |Y|} \in \mathrm{P}Y$$

NB: such a t defines therefore a function $\widehat{t} : \mathrm{P}X \to \mathrm{P}Y$ and t is easily seen to be determined by this function ($\widehat{s} = \widehat{t} \Rightarrow s = t$).

These morphisms are stable under composition: this defines the CCC $\mathbf{Pcoh}_!$. In other words, given $s \in \mathsf{P}(X \Rightarrow Y)$ and $t \in \mathsf{P}(Y \Rightarrow Z)$, there is $t \circ s \in \mathsf{P}(X \Rightarrow Z)$ such that

$$\widehat{t \circ s} = \widehat{t} \circ \widehat{s}$$

and there is an $\mathrm{Id} \in \mathsf{P}(X \Rightarrow X)$, given by $\mathrm{Id}_{[a],a} = 1$ and $\mathrm{Id}_{m,a} = 0$ if $m \neq [a]$. Of course $\widehat{\mathrm{Id}}(x) = x$.

- The cartesian product of X and Y is $X \mathbin{\&} Y$.
- The object of morphisms from X to Y is $X \Rightarrow Y$.

PX is naturally ordered by: $x \leq y$ if $\forall a \in |X|\ x_a \leq y_a$. Then P$X$ is directed-complete (it is actually an ω-continuous cpo).

For any $s \in \mathrm{P}(X \Rightarrow Y)$, the function $\widehat{s}$ is Scott continuous, and hence we can interpret (*e.g.* PCF) fixpoint operators.

Example of morphisms

A $s \in \mathrm{P}(1 \Rightarrow 1)$ is a family $(s_n)_{n\in\mathbb{N}}$ of elements of $\mathbb{R}^+$ such that $\sum_{n=0}^{\infty} s_n \leq 1$, and the associated function is $\widehat{s} : [0,1] \to [0,1]$ given by $\widehat{s}(x) = \sum_{n=0}^{\infty} s_n x^n$.

Functions of this kind are

- very smooth (analytic)
- very monotonic (all derivatives are ≥ 0)

Examples of such functions $[0,1] \to [0,1]$: $f(x) = \frac{1}{3} + \frac{x^2}{2}$, $f(x) = 1 - \sqrt{1-x}$ (with an ∞ derivative at $x = 1$).

Completely similar characterization of $s \in \mathsf{P}((1 \mathbin{\&} 1) \Rightarrow 1)$.

The "weak parallel or" $\mathsf{wpor} : [0,1]^2 \to [0,1]$ defined by $\mathsf{wpor}(x,y) = x + y - xy$ is not a morphism (there is a negative coefficient).

NB: adding such a morphism to the model is incompatible with the fact that all morphisms are analytic and the presence of least fixpoints. Indeed we would be able to define a function $f : [0,1] \to [0,1]$ by $f(x) = \mathsf{wpor}(x, f(x)) = x + f(x) - xf(x)$ as a least fixpoint. But then $f(0) = 0$ and $f(x) = 1$ if $x > 0$ so f is not analytic.

A $s \in \mathrm{P}((1 \oplus 1) \Rightarrow 1)$ is a family $(s_{n,m})_{n,m \in \mathbb{N}}$ such that

$$\forall p \in [0,1] \quad \sum_{n,m \in \mathbb{N}} s_{n,m} p^n (1-p)^m \leq 1$$

For instance: $f : \mathrm{P}(1 \oplus 1) \to [0,1]$ given by $f(x) = 4 x_{\mathbf{t}} x_{\mathbf{f}}$ is such a morphism because $p \in [0,1] \Rightarrow p(1-p) \leq \frac{1}{4}$.

NB: this function is not definable in "PCF" (but the function $f(x) = 2 x_{\mathbf{t}} x_{\mathbf{f}}$ is).

NB: we have no "parallel or" function in the model (requires negative coefficients), but we have an analogue of Gustave's function $g \in \mathrm{P}((1 \oplus 1) \,\&\, (1 \oplus 1) \,\&\, (1 \oplus 1) \Rightarrow 1)$ given by $g(x, y, z) = x_{\mathbf{t}} y_{\mathbf{f}} + y_{\mathbf{t}} z_{\mathbf{f}} + z_{\mathbf{t}} x_{\mathbf{f}}$. But $\frac{1}{2} g$ is definable whereas no εwpor (for $\varepsilon > 0$) is!

Main properties of this model

For probabilistic PCF (and its extensions with recursive types etc):

- Adequacy: if $\vdash M : \iota$, then the interpretation $[M]$ of M, which is an element of PN (ι is interpreted as the PCS N) satisfies:

  For all $n \in \mathbb{N}$, $[M]_n$ is the probability of M to reduce to $\underline{n}$.
- Equational full abstraction: semantical equality $\Rightarrow$ observational equivalence (same probability to reduce to $\underline{0}$ in any context of ground type ι).

Inequational full abstraction fails for the standard order of the model: $x, y \in \mathrm{P}X$ satisfy $x \leq y$ if $\forall a \in |X|\ x_a \leq y_a$.

Major limitation of PCS: no "continuous" types

Main limitation of this model: apparently, does not allow to interpret "continuous types" like the real line $\mathbb{R}$ (very important for the semantics of Machine Learning oriented languages).

Idea to overcome it: our morphisms are functions acting on the cpos $\mathsf{P}X$. Introduce more general such cpos and find a notion of morphisms generalizing those of PCS.

The cone generated by a PCS

Given a PCS X we can consider its associated "cone" $\mathsf{C}(X)$: the set of all $x \in (\mathbb{R}^+)^{|X|}$ such that $\varepsilon x \in \mathsf{P}X$ for some $\varepsilon > 0$. Then

- $\mathsf{C}(X)$ is an $\mathbb{R}^+$-semi module: if $x, y \in \mathsf{C}(X)$ and $\alpha, \beta \in \mathbb{R}^+$ then $\alpha x + \beta y = (\alpha x_a + \beta y_a)_{a \in |X|} \in \mathsf{C}(X)$ (with the usual algebraic properties).
- The canonical ordre relation can be defined by: $x \leq y$ if there is $z \in \mathsf{C}(X)$ such that $y = x + z$.
- There is a "norm" $\|\_\| : \mathsf{C}(X) \to \mathbb{R}^+$ defined by $\|x\| = \sup_{x' \in \mathsf{P}X^\perp} \langle x, x' \rangle$ which satisfies $\|\alpha x\| = \alpha \|x\|$, $\|x + y\| \leq \|x\| + \|y\|$ and $x \leq y \Rightarrow \|x\| \leq \|y\|$.
- Any monotonic bounded sequence of elements of $\mathsf{C}(X)$ has a least upper bound.

$\mathsf{C}(X)$ is a kind of "order complete positive Banach space".

Positive cone

We generalize this situation. A *cone* (many similar notions can be found in the literature) is an $\mathbb{R}^+$-semi module P (there are operations $+$ and $\mathbb{R}^+$ scalar multiplication satisfying the usual laws) equiped with a function $\|\_\|_P : P \to \mathbb{R}^+$ such that:

- $x + y = x' + y \Rightarrow x = x'$ (simplifiability).
- $\|\alpha x\|_P = \alpha\|x\|_P$, $\|x + y\|_P \leq \|x\|_P + \|y\|_P$ and $\|x\|_P = 0 \Rightarrow x = 0$ ($\|\_\|_P$ is a norm).
- Defining $x \leq y$ by $\exists z\ y = x + z$, we have $\|x\|_P \leq \|y\|_P$ (monotonicity of the norm).
- Any sequence $(x_n)_{n\in\mathbb{N}}$ of elements of P such that $\forall n \in \mathbb{N}\ x_n \leq x_{n+1}$ and $\exists\alpha \in \mathbb{R}^+\ \forall n \in \mathbb{N}\ \|x_n\|_P \leq \alpha$ (the sequence is bounded) has a lub $\sup_{n\in\mathbb{N}} x_n \in P$.

Subtraction

If $x \leq y$, there is a unique z such that $x + z = y$, by simplifiability.

This z is denoted as $y - x$.

This subtraction, when defined, satisfies all the usual algebraic laws.

Main motivating examples of cones

- For any PCS X, $\mathsf{C}(X) = \{x \in (\mathbb{R}^+)^{|X|} \mid \exists \varepsilon > 0\ \varepsilon x \in \mathsf{P}X\}$ is a cone with $\|x\|_{\mathsf{C}(X)} = \sup_{x' \in \mathsf{P}X^\perp} \langle x, x' \rangle$.
- Let $(\mathcal{X}, \Sigma_{\mathcal{X}})$ be a measurable space. We define a cone $\mathsf{Meas}(\mathcal{X})$ as the set of all $\mathbb{R}^+$-valued measures μ on $\Sigma_{\mathcal{X}}$ (in particular $\mu(\mathcal{X}) < \infty$). Algebraic operations defined in the obvious way. $\|\mu\|_{\mathsf{Meas}(\mathcal{X})} = \mu(\mathcal{X})$.

What morphisms?

Unit ball of P: $\mathcal{B}P = \{x \in P \mid \|x\|_P \leq 1\}$.

It is a poset where all monotone sequences have a lub (we do not consider uncoutable directed subsets because we will have to use the Monotone Convergence Theorem at some point).

By analogy with PCS, a morphism from P to Q should be a Scott continuous function $\mathcal{B}P \to \mathcal{B}Q$.

But this category is not a CCC.

The problem is that the curryfied version wpor' of $\mathrm{wpor} : [0,1] \times [0,1] \to [0,1]$ should be a monotonic function $[0,1] \to P$ where P is a cone of Scott continuous functions $[0,1] \to [0,1]$. Remember $\mathrm{wpor}(x,y) = x + y - xy$.

We should have $\mathrm{wpor}'(0) \leq \mathrm{wpor}'(1)$ in this cone P where the operations are defined pointwise.

But $f = \mathrm{wpor}'(1) - \mathrm{wpor}'(0) : [0,1] \to [0,1]$ is the function defined by $f(y) = 1 - y$ which is not monotonic.

Local cone

Let P be a cone and $u \in \mathcal{B}P$.

We define a new cone P_u as follows

- $P_u = \{x \in P \mid \exists\varepsilon > 0 \ \varepsilon x + u \in \mathcal{B}P\}$
- algebraic operations defined as in P.
- $\|x\|_{P_u} = \inf\{1/\varepsilon \mid \varepsilon > 0 \text{ and } \varepsilon x + u \in \mathcal{B}P\}$

Fact: P_u is a cone.

Observe that $\mathcal{B}(P_u) = \{x \in P \mid x + u \in \mathcal{B}P\} \subseteq \mathcal{B}P$.

Let P and Q be cones and $f : \mathcal{B}P \to Q$.

Assume that f is monotonic. Then, given $u \in \mathcal{B}P$, we have $\forall x \in \mathcal{B}(P_u)\ f(x) \leq f(x+u)$.

So we can define a function $\Delta f(\_;u) : \mathcal{B}(P_u) \to Q$ by $\Delta f(x;u) = f(x+u) - f(x)$. We require this function to be also monotonic.

Given $v \in \mathcal{B}P$ such that $u + v \in \mathcal{B}P$, we can consider the function $\Delta f(\_;u,v) : \mathcal{B}(P_{u,v}) \to Q$ given by $\Delta f(x;u,v) = \Delta(\Delta f(\_;u))(x;v) = f(x+u+v) - f(x+u) - f(x+v) + f(x)$. We require this function to be also monotonic.

And the same for all functions $\Delta f(\_;u_1,\dots,u_n)$ for $u_1,\dots,u_n \in \mathcal{B}P$ such that $\sum_{i=1}^{n} u_i \in \mathcal{B}P$.

Third iterated difference:
$\Delta f(x; u_1, u_2, u_3) = f(x + u_1 + u_2 + u_3) - f(x + u_1 + u_2) - f(x + u_1 + u_3) - f(x + u_2 + u_3) + f(x + u_1) + f(x + u_2) + f(x + u_3) - f(x)$.

More generally, if $\vec{u} \in \mathcal{B}P^n$ with $\sum_{i=1}^{n} u_i \in \mathcal{B}P$ and $x \in \mathcal{B}(P_{\vec{u}})$ and $\varepsilon \in \{+, -\}$, one defines

$$\Delta^{\varepsilon} f(x; \vec{u}) = \sum_{I \in \mathcal{P}_{\varepsilon}(n)} f(x + \sum_{i \in I} u_i) \in Q$$

where $\mathcal{P}_{+}(n)$ (resp. $\mathcal{P}_{-}(n)$) is the set of all $I \subseteq \{1, \ldots, n\}$ such that $n - \#I$ is even (resp odd).

Definition

The function $f : \mathcal{B}P \to Q$ is *absolutely monotonic* if, for all $\vec{u} \in \mathcal{B}P^n$ with $\sum_{i=1}^{n} u_i \in \mathcal{B}P$ and $x \in \mathcal{B}P_{\vec{u}}$, one has

$$\Delta^- f(x; \vec{u}) \leq \Delta^+ f(x; \vec{u})$$

Then we set $\Delta f(x; \vec{u}) = \Delta^+ f(x; \vec{u}) - \Delta^- f(x; \vec{u})$. This generalizes our previous examples, and the function $\Delta f(\_; \vec{u})$ is monotonic.

Definition

The function $f : \mathcal{B}P \to Q$ is *stable* if it is absolutely monotonic, bounded (that is $\exists\alpha\,\forall x \in \mathcal{B}P\ \|f(x)\|_Q \leq \alpha$) and Scott continuous, that is:
For all monotonic sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{B}P$, one has $\sup_{n\in\mathbb{N}} f(x_n) = f(\sup_{n\in\mathbb{N}} x_n)$.

We define a cone $P \Rightarrow Q$ as follows.

- The elements of $P \Rightarrow Q$ are the stable functions $\mathcal{B}P \to Q$.
- Addition and scalar multiplication are defined pointwise ($(f+g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha f(x)$; these functions are stable by linearity of the operators $\Delta^{\varepsilon}\_(x; \vec{u})$).
- $\|f\|_{P \Rightarrow Q} = \sup_{x \in \mathcal{B}P} \|f(x)\|_Q$.

Then $f \leq g$ is equivalent to the following condition: for all $n \in \mathbb{N}$, all $\vec{u} \in \mathcal{B}P^n$ such that $\sum u_i \in \mathcal{B}P$ and all $x \in \mathcal{B}(P_{\vec{u}})$, one has

$$\Delta f(x; \vec{u}) \leq \Delta g(x; \vec{u})$$

that is

$$\Delta^+ f(x; \vec{u}) + \Delta^- g(x; \vec{u}) \leq \Delta^- f(x; \vec{u}) + \Delta^+ g(x; \vec{u})$$

Composing stable functions

If $f \in \mathcal{B}(P \Rightarrow Q)$ then $f : \mathcal{B}P \to \mathcal{B}Q$.

Theorem

Let $f \in \mathcal{B}(P \Rightarrow Q)$ and $g \in \mathcal{B}(Q \Rightarrow R)$. Then $g \circ f \in \mathcal{B}(P \Rightarrow R)$.

The proof is not straightforward because morphisms are not defined by a preservation property (like Scott continuity of Berry stability).

So we have a category **Cstab** whose objects are the cones and morphisms, the stable functions.

Reminder: a coherence space is a structure $E = (|E|, \coh_E)$ where $|E|$ is a countable set and $\coh_E$ is a binary, reflexive and symmetric relation on $|E|$.

Cliques of E: $\mathrm{Cl}(E) = \{x \subseteq |E| \mid \forall a, a' \in x\ a \coh_E a'\}$.

$(\mathrm{Cl}(E), \subseteq)$ is a cpo, any subset of a clique is a clique.

A function $f : \mathrm{Cl}(E) \to \mathrm{Cl}(F)$ is stable if it is monotonic, Scott continuous and "conditionally multiplicative", that is:

$$\forall x, x' \in \mathrm{Cl}(E) \quad x \cup x' \in \mathrm{Cl}(E) \Rightarrow f(x \cap x') = f(x) \cap f(x')$$

Let $f, g : \mathrm{Cl}(E) \to \mathrm{Cl}(F)$ be stable.
f is stably less than g ($f \leq g$) if

$$\forall x, x' \in \mathrm{Cl}(E) \quad x \subseteq x' \Rightarrow f(x) = f(x') \cap g(x)$$

Reformulating stability in coherence spaces

Let E be a coherence space. If $u \in \mathrm{Cl}(E)$, define a "local" coherence space E_u as follows:
$|E_u| = \{a \in |E| \setminus u \mid \forall a' \in u\ a \frown_E a'\}$. So if $x \in \mathrm{Cl}(E_u)$, $x + u$ (disjoint union) is in $\mathrm{Cl}(E)$.

Let $f : \mathrm{Cl}(E) \to \mathrm{Cl}(F)$ be a function and $u \in \mathrm{Cl}(E)$. We define $\Delta f(\_; u) : \mathrm{Cl}(E_u) \to \mathrm{Cl}(F)$ by
$\Delta f(x; u) = f(x + u) \setminus f(x) \in \mathrm{Cl}(F)$.

Theorem

A Scott continuous function $f : \mathit{Cl}(E) \to \mathit{Cl}(F)$ is stable iff for all $u \in \mathit{Cl}(E)$, the function $\Delta f(\_\,; u) : \mathit{Cl}(E_u) \to \mathit{Cl}(F)$ is monotonic.

If f is stable, then $\Delta f(\_\,; u) : \mathit{Cl}(E_u) \to \mathit{Cl}(F)$ is also stable.

So there is no need to consider $\Delta f(x; u_1, \ldots, u_n)$ for $n \geq 2$: the corresponding conditions are redundant (this due to the idempotency of union).

Theorem

Let $f, g : \mathit{Cl}(E) \to \mathit{Cl}(F)$ *be stable functions.*

One has $f \leq g$ *(for the stable order) iff*

- $\forall x \in \mathit{Cl}(E)\ f(x) \subseteq g(x)$
- $\forall u \in \mathit{Cl}(E) \forall x \in \mathit{Cl}(E_u)\ \Delta f(x; u) \subseteq \Delta g(x; u)$.

The cartesian product is defined in the obvious way: $P \times Q$ with norm defined by $\|(x,y)\|_{P\times Q} = \max(\|x\|_P, \|y\|_Q)$.

We have already defined the cone $P \Rightarrow Q$. The evaluation map $\mathsf{Ev} : (P \Rightarrow Q) \times P \to Q$ is defined by $\mathsf{Ev}(f,x) = f(x)$. It is stable.

If $f : \mathcal{B}R \times \mathcal{B}P \to \mathcal{B}Q$ is stable, it is a very nice exercise to prove that the function $\Lambda(f) : \mathcal{B}R \to \mathcal{B}(P \Rightarrow Q)$ defined as usual by $\Lambda(f)(z)(x) = f(z,x)$ takes actually its values in $\mathcal{B}(P \Rightarrow Q)$ and is stable.

Types of our target language are interpreted as cones. There is a type ρ of real numbers, and $[\rho] = \mathrm{Meas}(\mathbb{R})$ (with respect to the standard Lebesgue σ-algebra). For simplicity, ρ is our unique ground type.

A closed term M such that $\vdash M : \rho$ will be interpreted as an element $[M]$ of $\mathcal{B}(\mathrm{Meas}(\mathbb{R}))$, that is, as a sub-probability measure.

For each $r \in \mathbb{R}$, there is a constant $\underline{r}$ of our language $\vdash \underline{r} : \rho$. We set $[\underline{r}] = \delta_r \in \mathcal{B}(\mathrm{Meas}(\mathbb{R}))$.

There is also a constant $\vdash \mathsf{sample} : \rho$. Intuitively, sample draws a real number in $[0, 1]$ with uniform probability.

The language is CBN but has a "let" construct *restricted to the ground type of real numbers* (we omit contexts for readability):

$$\frac{\vdash M : \rho \qquad x : \rho \vdash N : \sigma}{\vdash \mathsf{let}\, x = M \,\mathsf{in}\, N : \sigma}$$

This construction is crucial: it draws a real number r according to the sub-probability measure defined by M and reduces to $N\,[\underline{r}/x]$.

In our model **Cstab**: $\mu = [M]$ is an element of $\mathcal{B}(\mathsf{Meas}(\mathbb{R}))$. And $[N]_x : \mathcal{B}(\mathsf{Meas}(\mathbb{R})) \to [\sigma]$ is a stable function.

So we have a function $\gamma : [N]_x \circ \delta : \mathbb{R} \to [\sigma]$ where δ is the function $r \mapsto \delta_r$. We have $\gamma(r) = [N\,[\underline{r}/x]]$ (by Substitution Lemma).

So we would like to set
$[\mathsf{let}\, x = M \,\mathsf{in}\, N] = \int_{\mathbb{R}} \gamma(r)\mu(dr) = \int_{\mathbb{R}} [N]_x\,(\delta_r)\mu(dr)$

This integral does not make sense *a priori* for two reasons:

- We don't know how to integrate functions ranging in an arbitrary cone, but this is not a serious issue because $\sigma = \sigma_1 \to \cdots \to \sigma_k \to \rho$ and so we can replace our problem with: given $\gamma : \mathbb{R} \times \mathcal{B}P \to \mathsf{Meas}(\mathbb{R})$, define $\gamma_\mu : \mathcal{B}P \to \mathsf{Meas}(\mathbb{R})$ by

$$\gamma_\mu(x)(U) = \int_{\mathbb{R}} \gamma(r,x)(U)\mu(dr)$$

  for $U \in \Sigma_{\mathbb{R}}$.

- More seriously, given $x \in \mathcal{B}P$ and $U \in \Sigma_{\mathbb{R}}$, there is no reason *a priori* for the function $r \mapsto \gamma(r,x)(U)$ (from $\mathbb{R}$ to $\mathbb{R}$) to be measurable.

Our solution:

Equip all cones with a family of sets of "measurability tests" $(\mathrm{M}^n(P))_{n\in\mathbb{N}}$ where each element l of $\mathrm{M}^n(P)$ is a function $l : \mathbb{R}^n \times P \to \mathbb{R}^+$ with the following properties:

- For each $\vec{r} \in \mathbb{R}^n$, the function $x \mapsto l(\vec{r}, x)$ is linear (commutes with all linear combinations in P) and Scott continuous from P to $\mathbb{R}^+$.
- For each $x \in P$, the function $\vec{r} \mapsto l(\vec{r}, x)$ is measurable.
- For each measurable $h : \mathbb{R}^k \to \mathbb{R}^n$, $l \circ h \in \mathrm{M}^k(P)$.
- $0 \in \mathrm{M}^n(P)$ for all n.

Next we say that a function $\gamma : \mathbb{R}^n \to P$ is a *measurable path* if:

- $\gamma(\mathbb{R}^n)$ is bounded in P
- and for all $l \in \mathrm{M}^k(P)$, the function $l * \gamma : \mathbb{R}^{k+n} \to \mathbb{R}^+$ defined by $(l * \gamma)(\vec{r}, \vec{s}) = l(\vec{r}, \gamma(\vec{s}))$ is measurable.

Last, a morphism $P \to Q$ in our new category $\mathbf{Cstab_m}$ is a stable function $f : \mathcal{B}P \to \mathcal{B}Q$ such that, for any measurable path $\gamma : \mathbb{R}^n \to P$, the function $f \circ \gamma$ is a measurable path.

Such an f will be called a stable measurable function.

Construction of measurability tests

If $\mathcal{X}$ is a measurable space, we equip $\mathsf{Meas}(\mathcal{X})$ with the following measurability tests: $\mathsf{M}^n(\mathsf{Meas}(\mathcal{X})) = \{e_U \mid U \in \Sigma_{\mathcal{X}}\}$ where $e_U(\vec{r}, \mu) = \mu(U)$ for $\vec{r} \in \mathbb{R}^n$ and $\mu \in \mathsf{Meas}(\mathcal{X})$.

$P \Rightarrow Q$ is the cone of all stable and measurable functions $\mathcal{B}P \to Q$, and this cone is equiped with the following measurability tests:

- Given $\gamma : \mathbb{R}^n \to \mathcal{B}P$ a measurable path and $m \in \mathsf{M}^n(Q)$, we define $\gamma \triangleright m : \mathbb{R}^n \times (P \Rightarrow Q) \to \mathbb{R}^+$ by $(\gamma \triangleright m)(\vec{r}, f) = m(\vec{r}, f(\gamma(\vec{r})))$.
- $\mathsf{M}^n(P \Rightarrow Q)$ is the set of all these $\gamma \triangleright m$.

To prove the completeness property of $P \Rightarrow Q$, we need the Monotone Convergence Theorem, so we can consider only lubs of countable families. This is enough for fixpoints!

No surprise in the definition of $P \times Q$.

This defines a CCC $\mathbf{Cstab}_{\mathrm{m}}$ where we can interpret our target language and prove an adequacy theorem.

This solves indeed our integration problem.

$$\frac{\vdash M : \rho \qquad x : \rho \vdash N : \sigma}{\vdash \mathsf{let}\, x = M \,\mathsf{in}\, N : \sigma}$$

We take $\sigma = \rho$ to simplify a bit the notations.

- $\mu = [M] \in \mathcal{B}(\mathsf{Meas}(\mathbb{R}))$,
- $f = [N]_x : \mathcal{B}(\mathsf{Meas}(\mathbb{R})) \to \mathcal{B}(\mathsf{Meas}(\mathbb{R}))$ is stable and mesurable.

The map $\delta : \mathbb{R} \to \mathsf{Meas}(\mathbb{R})$ defined by $\delta(r) = \delta_r$ is a measurable path, because, for any $U \in \Sigma_{\mathbb{R}}$, the test $e_U \in \mathrm{M}^n(\mathsf{Meas}(\mathbb{R}))$ satisfies:

$$(e_U * \delta)(\vec{r}, r) = e_U(\vec{r}, \delta(r)) = \delta_r(U) = \begin{cases} 1 & \text{if } r \in U \\ 0 & \text{otherwise.} \end{cases}$$

so $e_U * \delta$ is measurable since $U \in \Sigma_{\mathbb{R}}$.

Because f is (stable) measurable, it follows that $f \circ \delta : \mathbb{R} \to \mathrm{Meas}(\mathbb{R})$ is a measurable path which means that for all $U \in \Sigma_{\mathbb{R}}$, the function

$$\begin{aligned} \mathbb{R} &\to \mathbb{R}^+ \\ r &\mapsto f(\delta_r)(U) \end{aligned}$$

is measurable.

So we can define $[\mathrm{let}\, x = M \,\mathrm{in}\, N] \in \mathrm{Meas}(\mathbb{R})$ as the measure ν given by

$$\nu(U) = \int_{\mathbb{R}} f(\delta_r)(U)\mu(dr)$$

Conclusion: a few questions

- Conjecture: this is an equationally fully abstract model of our "real probabilistic PCF" target language.
- We have a natural notion of measurable linear maps on cones. Does it give rise to a model of ILL? Probably. Of classical LL? Probably not, but can we find a class of measurable cones for which it is true, and which contains the cones Meas($\mathcal{X}$)?
- Representation theorem for a sufficiently large class of cones (including Meas($\mathcal{X}$)), typically replacing the webs $|X|$ of PCS with more structured spaces? Related to the previous question. This seems a crucial step in the development of an "intersection type systems" adapted to languages with continuous data types like $\mathbb{R}$.
- Probabilistic sequentiality, strong stability?
- Connection with other approaches (in particular: Staton quasi-Borel spaces, Keimel and Plotkin Kegelspitzen)?