Probability and Non-Determinism in Domain Theory, Part II

Jean Goubault-Larrecq

Partially supported by the ANR CPP project

LI'12 — Feb 23, 2012

Outline

Outline

The Hoare Powerdomain

Definition

The Hoare powerdomain $\mathcal{H}(X)$ is the set of closed non-empty subsets of X, with inclusion ordering $\subseteq$.

- Idea: if $F \in \mathcal{H}(X)$, pick $x \in F$ non-deterministically
- Can be justified in many ways

The Categorical Explanation

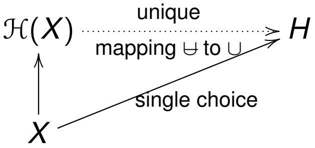
We want a domain H of "choices" such that:

- Can encode single choice in H $\quad$ ("$\{x\}$")
- Binary choice $F \uplus F'$ between two choices F, F' $\quad$ ("$F \cup F'$")

$$(F \uplus F') \uplus F'' = F \uplus (F' \uplus F'') \qquad F \uplus F' = F' \uplus F \qquad F \leq F \uplus F'$$

Theorem

The free domain of choices over the domain X is $\mathcal{H}(X)$.



Note on Free Domains

In continuous dcpos (domains), we can describe free domains through a basis.

- This is what we did for indexed valuations
- For angelic non-determinism, basis given by $\downarrow\{x_1, \ldots, x_n\}$, free choice on $x = \downarrow\{x\}$, $\uplus$ is union
- For probabilistic choice, basis given by simple valuations $a_1\delta_{x_1} + \ldots + a_n\delta_{x_n}$

Note on Free Domains

In continuous dcpos (domains), we can describe free domains through a basis. In general, for an (in)equational theory E:

Building a free domain

Basis of free E-algebra is quotient of terms through all derivable equations.
Then complete this by forming the rounded ideal completion.

- The last step is hard to understand intuitively
- Miraculously, sometimes we have concrete descriptions of it: e.g., $\mathcal{H}(X)$, space of closed subsets.

Outline

Random Variables

- Started in 2010: Daniele wanted a concrete description of indexed valuations (not just the finite ones)
- In a sense, we failed, but not by much [GLV, LICS'11]
- Similar to the probabilist's notion of random variable:

Definition

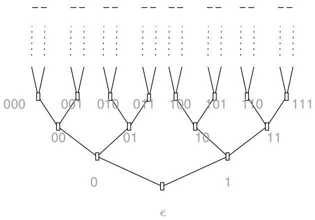
A random variable on X is a triple (Ω, μ, f) of:

- a sample space Ω . . . . draw $\omega \in \Omega$ at random
- a probability measure μ on Ω . . . . using μ
- a measurable map $f : \Omega \to X$ . . . . and obtain $x = f(\omega)$

Note the separation between Ω and X.

The Sample Space Ω

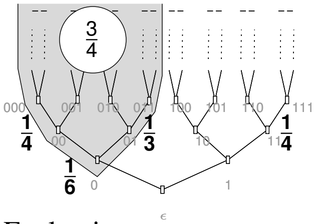
We take
$\Omega = \{0, 1\}^{\leq\omega}$:
the Cantor tree.



- Basic open sets: $\uparrow x$ for finite sequence x

The Sample Space Ω

We take
$\Omega = \{0,1\}^{\leq\omega}$:
the Cantor tree.

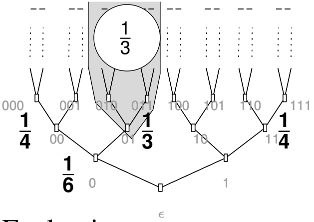


Evaluating
$\frac{1}{4}\delta_{00} + \frac{1}{6}\delta_0 + \frac{1}{3}\delta_{01} + \frac{1}{4}\delta_{11}$
on $\uparrow 0$

- Basic open sets: $\uparrow x$ for finite sequence x

The Sample Space Ω

We take $\Omega = \{0, 1\}^{\leq\omega}$: the Cantor tree.

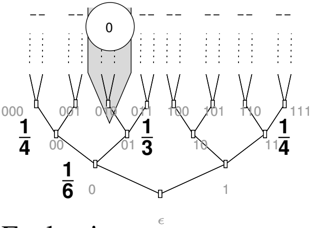


Evaluating $\frac{1}{4}\delta_{00} + \frac{1}{6}\delta_0 + \frac{1}{3}\delta_{01} + \frac{1}{4}\delta_{11}$ on $\uparrow 01$

- Basic open sets: $\uparrow x$ for finite sequence x

The Sample Space Ω

We take
$\Omega = \{0, 1\}^{\leq\omega}$:
the Cantor tree.

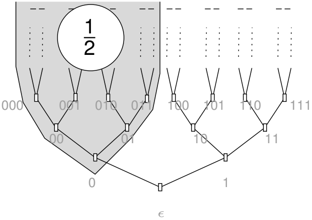


Evaluating
$\frac{1}{4}\delta_{00} + \frac{1}{6}\delta_{0} + \frac{1}{3}\delta_{01} + \frac{1}{4}\delta_{11}$
on $\uparrow 010$

- Basic open sets: $\uparrow x$ for finite sequence x

The Sample Space Ω

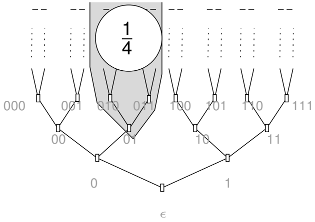
We take
$\Omega = \{0, 1\}^{\leq\omega}$:
the Cantor tree.



- Basic open sets: $\uparrow x$ for finite sequence x
- A fair coin induces the uniform valuation Λ (on the top elts)

The Sample Space Ω

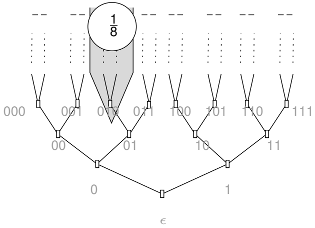
We take $\Omega = \{0, 1\}^{\leq \omega}$: the Cantor tree.



- Basic open sets: $\uparrow x$ for finite sequence x
- A fair coin induces the uniform valuation Λ (on the top elts)

The Sample Space Ω

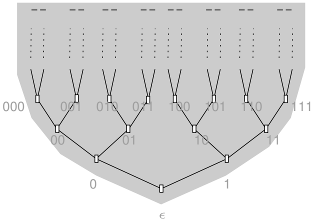
We take
$\Omega = \{0,1\}^{\leq\omega}$:
the Cantor tree.



- Basic open sets: $\uparrow x$ for finite sequence x
- A fair coin induces the uniform valuation Λ (on the top elts)

The Sample Space Ω

We take
$\Omega = \{0, 1\}^{\leq\omega}$:
the Cantor tree.

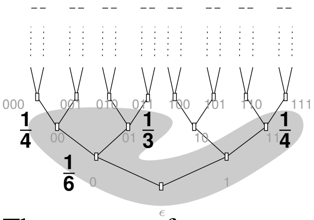


The support of Λ is
the whole Cantor tree

- Basic open sets: $\uparrow x$ for finite sequence x
- A fair coin induces the uniform valuation Λ (on the top elts)
- The support $\text{supp}\,\nu$, is the complement of the largest U such that $\nu(U) = 0$

The Sample Space Ω

We take $\Omega = \{0,1\}^{\leq\omega}$: the Cantor tree.



The support of $\frac{1}{4}\delta_{00} + \frac{1}{6}\delta_{0} + \frac{1}{3}\delta_{01} + \frac{1}{4}\delta_{11}$

- Basic open sets: $\uparrow x$ for finite sequence x
- A fair coin induces the uniform valuation Λ (on the top elts)
- The support $\mathrm{supp}\,\nu$, is the complement of the largest U such that $\nu(U) = 0$

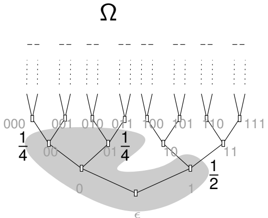
Quasi-Uniform Valuations

For every non-empty closed subset F of Ω,
there is a projection map $p_F : \omega \mapsto$ largest prefix of ω in F.

Definition

A valuation ν on Ω is quasi-uniform iff ν is image measure $p_F[\Lambda]$ of Λ onto some closed subset F of Ω such that $\forall \omega \notin F \cdot p_F(\omega) \in \mathrm{Max}\,\Omega$ (thinness condition).

"Flip all bits with probability $\frac{1}{2}$, independently; possibly stop"



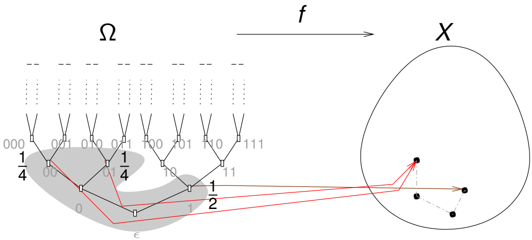
Continuous Random Variables

We fix the sample space Ω=Cantor tree.

Definition

A (uniform) continuous random variable (CRV) on X is a pair (ν, f) of a quasi-uniform valuation ν on Ω, and a continuous map

$$f : \mathrm{Max}\,\mathrm{supp}\,\nu \to X$$



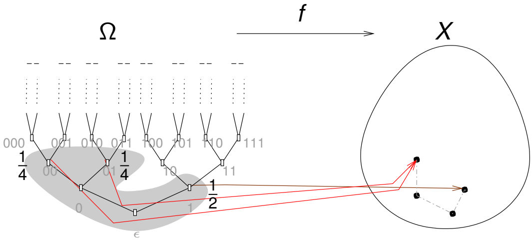
The Ordering on Uniform CRVs

Definition ($\leq$)

$(\nu, f) \leq (\nu', f')$ iff:

"increase supp" $\operatorname{supp} \nu \subseteq \operatorname{supp} \nu'$

"increase values" $f \circ p_{\operatorname{supp} \nu} \leq f'$



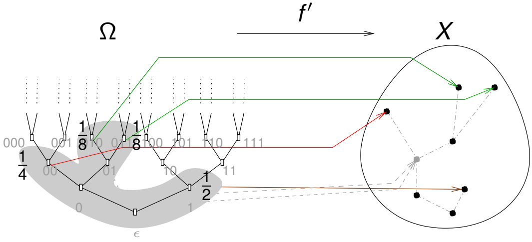
The Ordering on Uniform CRVs

Definition ($\leq$)

$(\nu, f) \leq (\nu', f')$ iff:

"increase supp" $\operatorname{supp}\nu \subseteq \operatorname{supp}\nu'$

"increase values" $f \circ p_{\operatorname{supp}\nu} \leq f'$



Domain-Theoretic Properties

Let $\upsilon\mathbf{R}(X)$ be the dcpo of (uniform) CRVs on X.

Theorem

The following categories are closed under $\upsilon\mathbf{R}$*:*

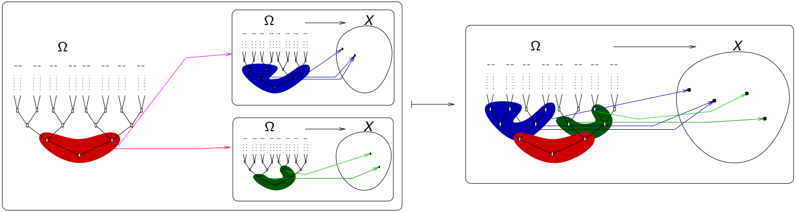
- *bc-domains (continuous + sups of finite bounded sets)*
- *Scott domains (algebraic + sups of finite bounded sets)*
- *continuous lattices*
- *algebraic lattices*

All these categories are Cartesian-closed

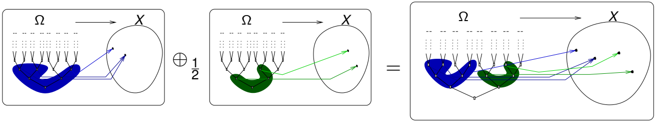
(very different from the case of valuations $\mathbf{V}_1$)

The Monad of CRVs

- Unit $\eta(x) = (\delta_\epsilon, \{\epsilon \mapsto x\})$ ("flip no coin, return x right away")
- Multiplication $\mu : \upsilon\mathbf{R}^2(X) \to \upsilon\mathbf{R}(X)$



- Random choice



Relationship to Valuations

A natural transformation $\upsilon\mathbf{R} \to \mathbf{V}_1$

Maps $(\nu, f) \in \upsilon\mathbf{R}(X)$ to $f[\nu]$

E.g., maps $(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1, \{0 \mapsto a, 1 \mapsto b\})$ to $\frac{1}{2}\delta a + \frac{1}{2}\delta b$

Note: if $a = b$, the CRV
keeps the information that we have flipped a coin

$$(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1, \{0, 1 \mapsto a\})$$

The valuation $f[\nu]$ has lost it:

$$\delta_a$$

Relationship to Indexed Valuations

A natural transformation $\upsilon\mathbf{R} \to \mathbf{V}_1$

Maps $(\nu, f) \in \upsilon\mathbf{R}(X)$ to $f[\nu]$
factors as $\upsilon\mathbf{R} \to \mathcal{IV} \to \mathbf{V}_1$

E.g., maps $(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1, \{0 \mapsto a, 1 \mapsto b\})$ to $\frac{1}{2}\delta a + \frac{1}{2}\delta b$
through the equivalence class $\{\frac{1}{2}\delta_a, \frac{1}{2}\delta_b\}$

Note: if $a = b$, the IV
keeps the information that we have flipped a coin

$$\{\tfrac{1}{2}\delta_a, \tfrac{1}{2}\delta_a\}$$

The valuation $f[\nu]$ has lost it:

$$\delta_a$$

(In)equational Characterization

Theory of $\mathbf{V}_1$

1. $x \oplus_p y = y \oplus_{1-p} x$
2. $x \oplus_p (y \oplus_q z) = (x \oplus_{\frac{p}{p+q-pq}} y) \oplus_{p+q-pq} z$
3. $x \oplus_1 y = x$, $x \oplus_0 y = y$
4. $x = x \oplus_p x$

with $x \oplus_p y$ continuous in $x, y, p \in [0, 1]$

(In)equational Characterization

Theory of $\mathcal{IV}$

1. $x \oplus_p y = y \oplus_{1-p} x$
2. $x \oplus_p (y \oplus_q z) = (x \oplus_{\frac{p}{p+q-pq}} y) \oplus_{p+q-pq} z$
3. $x \oplus_1 y = x$, $x \oplus_0 y = y$
4. $x \leq x \oplus_p x$

with $x \oplus_p y$ continuous in x, y

(In)equational Characterization

Theory of υ**R**

1. ~~$x \oplus_p y = y \oplus_{1-p} x$~~
2. ~~$x \oplus_p (y \oplus_q z) = (x \oplus_{\frac{p}{p+q-pq}} y) \oplus_{p+q-pq} z$~~
3. ~~$x \oplus_1 y = x, x \oplus_0 y = y$~~
4. $x \leq x \oplus_{\frac{1}{2}} x$

with $x \oplus_{\frac{1}{2}} y$ continuous in x, y

Combining Non-Determinism and CRVs

Theorem ([Beck69])

If S, T are monads, and $\ell : TS \to ST$ is a distributivity law, then ST is a monad.

- (Details omitted)
- Take $S = \mathcal{H}$, $T = \upsilon\mathbf{R}$.

A Distributivity Law

Distributing $\upsilon\mathbf{R}$ Over $\mathcal{H}$

$$\begin{array}{rcl} \ell_X : \upsilon\mathbf{R}(\mathcal{H}(X)) & \to & \mathcal{H}(\upsilon\mathbf{R}(X)) \\ (\nu, F) & \mapsto & cl\{(\nu, f) \mid f \text{ selection of } F\} \end{array}$$

where $f : \Omega \to X$ is a selection of $F : \Omega \to \mathcal{H}(X)$ iff for every $\omega \in \Omega$, $f(\omega) \in F(\omega)$.

$$\begin{array}{rcl} \ell_X : \upsilon\mathbf{R}(\mathcal{H}(X)) & \to & \mathcal{H}(\upsilon\mathbf{R}(X)) \\ (\nu, \{\omega_i \mapsto \downarrow E_i\}) & \mapsto & \downarrow\{(\nu, \{\omega_i \mapsto x_i\}) \mid \forall i \cdot x_i \in E_i)\} \end{array}$$

with E_i finite non-empty subsets, $\nu = \sum_i a_i \delta_{\omega_i}$

A Monad of Non-Det.+Prob. Choice

- Define $TX = \mathcal{H}(\upsilon\mathbf{R}(X))$
  do all non-deterministic choices first, then random choices
- Unit $\eta(x) = \downarrow\{(\epsilon, \{\epsilon \mapsto \delta_x\})\}$
- Multiplication: horrible
  No better formula than Beck's: $\mu_T = \mathcal{H}(\mu_{\upsilon\mathbf{R}}) \circ \mu_{\mathcal{H}} \circ \mathcal{H}(\ell)$
- Non-det. choice $F \uplus F' = F \cup F'$
- Probabilistic choice
  $F \oplus_{\frac{1}{2}} F' = cl\{(\nu, f) \oplus_{\frac{1}{2}} (\nu', f') \mid (\nu, f) \in F, (\nu', f') \in F'\}$
  —distributing probability over non-determinism

Outline

Mislove's Observation

Remember that

Distributing Non-Det. Over Prob. Implies Convexity

$$
\begin{aligned}
x \uplus y &= (x \uplus y) \oplus_{\frac{1}{2}} (x \uplus y) && \text{since } z \oplus_{\frac{1}{2}} z = z \\
&= (x \oplus_{\frac{1}{2}} x) \uplus (x \oplus_{\frac{1}{2}} y) \\
&\quad \uplus (y \oplus_{\frac{1}{2}} x) \uplus (y \oplus_{\frac{1}{2}} y) \\
&= x \uplus (x \oplus_{\frac{1}{2}} y) \uplus (y \oplus_{\frac{1}{2}} x) \uplus y && \text{since } z \oplus_{\frac{1}{2}} z = z
\end{aligned}
$$

"If you make a choice between x and y,
you cannot prevent possibly choosing $x \oplus_{\frac{1}{2}} y$ as well"

Mislove's Observation: Getting Around It

~~Remember that~~ In $\mathcal{JV}$ and CRVs,

Distributing Non-Det. Over Prob. ~~Implies Convexity~~

$$
\begin{array}{rcll}
x \uplus y & \neq & (x \uplus y) \oplus_{\frac{1}{2}} (x \uplus y) & \text{since } z \oplus_{\frac{1}{2}} z = z \\
& = & (x \oplus_{\frac{1}{2}} x) \uplus (x \oplus_{\frac{1}{2}} y) & \\
& & \uplus (y \oplus_{\frac{1}{2}} x) \uplus (y \oplus_{\frac{1}{2}} y) & \\
& \neq & x \uplus (x \oplus_{\frac{1}{2}} y) \uplus (y \oplus_{\frac{1}{2}} x) \uplus y & \text{since } z \oplus_{\frac{1}{2}} z = z
\end{array}
$$

"If you make a choice between x and y,
you ~~cannot~~ can prevent possibly choosing $x \oplus_{\frac{1}{2}} y$ as well"

Mislove's Observation

Remember that, with valuations:

Distributing Non-Det. Over Prob. Implies Convexity

$$\begin{aligned}
x \uplus y &= (x \uplus y) \oplus_{\frac{1}{2}} (x \uplus y) && \text{since } z \oplus_{\frac{1}{2}} z = z \\
&= (x \oplus_{\frac{1}{2}} x) \uplus (x \oplus_{\frac{1}{2}} y) \\
&\quad \uplus (y \oplus_{\frac{1}{2}} x) \uplus (y \oplus_{\frac{1}{2}} y) \\
&= x \uplus (x \oplus_{\frac{1}{2}} y) \uplus (y \oplus_{\frac{1}{2}} x) \uplus y && \text{since } z \oplus_{\frac{1}{2}} z = z
\end{aligned}$$

"If you make a choice between x and y,
you cannot prevent possibly choosing $x \oplus_{\frac{1}{2}} y$ as well,
nor $x \oplus_{\frac{1}{4}} y$, $x \oplus_{\frac{3}{4}} y$ ..."

The MOW-TKP Model

[Mislove, CONCUR00] [Mislove, Ouaknine, Worrell, ENTCS03]
[McIver, Morgan, TCS01]
[Tix, PhD99] [Tix, Keimel, Plotkin, ENTCS05]
(Presented differently, to make comparison with T easier)

- Let $\mathcal{H}^{\mathrm{cvx}}(C) = \{F \in \mathcal{H}(C) \mid F \text{ convex}\}$
  F convex iff $ax + (1-a)y \in F$ whenever $x, y \in F$, $a \in [0,1]$
  and C is a cone for this to make sense
- Now define $T^{\mathrm{cvx}}X = \mathcal{H}^{\mathrm{cvx}}(\mathbf{V}_1(X))$ instead of $\mathcal{H}(\upsilon\mathbf{R}(X))$
- Unit $\eta(x) = \downarrow\{\delta_x\}$
- Multiplication $\mu_T = \mathcal{H}^{\mathrm{cvx}}(\mu_{\mathbf{V}}) \circ \mu_{\mathcal{H}^{\mathrm{cvx}}} \circ \mathcal{H}^{\mathrm{cvx}}(\ell^{\mathrm{cvx}})$ (Beck)
  where $\ell^{\mathrm{cvx}}(\sum_i a_i \delta_{\downarrow E_i})) = \overline{conv}(\{\sum_i a_i \delta_{x_i} \mid \forall i \cdot x_i \in E_i\})$
- $F \uplus F' = \overline{\mathrm{conv}}(F \cup F'))$ (closed convex hull)
- $F \oplus_p F' = \overline{\mathrm{conv}}(\{p\nu + (1-p)\nu' \mid \nu \in F, \nu' \in F'\})$
  distributing probability over non-determinism

A More Abstract View: (Linear) Previsions

Given $\nu \in \mathbf{V}_1(X)$, define

$$\begin{aligned} rep(\nu) : [X \to \overline{\mathbb{R}^+}] &\to \overline{\mathbb{R}^+} \\ h &\mapsto \int_{x \in X} h(x) d\nu \end{aligned}$$

Lemma

- $rep(F)$ is Scott-continuous
- $rep(F)(a.h) = a.rep(F)(h)$ whenever $a \geq 0$ (positively homogeneous)
- $rep(F)(h + h') = rep(F)(h) + rep(F)(h')$ (linear)
- $rep(F)(a + h) = a + rep(F)(h)$, a cst. (normalisation)

A More Abstract View: (Convex) Previsions

Given $F \in \mathcal{H}^{\mathrm{cvx}}(\mathbf{V}_1(X))$, define

$$\begin{array}{rcl} rep(F) \colon [X \to \overline{\mathbb{R}^+}] & \to & \overline{\mathbb{R}^+} \\ h & \mapsto & \sup_{\nu \in F} \int_{x \in X} h(x) d\nu \end{array}$$

Lemma

- $rep(F)$ is Scott-continuous
- $rep(F)(a.h) = a.rep(F)(h)$ whenever $a \geq 0$ (positively homogeneous)
- $rep(F)(h + h') \leq rep(F)(h) + rep(F)(h')$ (sublinear, convex)
- $rep(F)(a + h) = a + rep(F)(h)$, a cst. (normalisation)

A More Abstract View: (Convex Discrete) Previsions

Given $F \in \mathcal{H}(X)$, define

$$rep(F) \colon [X \to \overline{\mathbb{R}^+}] \to \overline{\mathbb{R}^+}$$
$$h \mapsto \sup_{x \in F} h(x)$$

Lemma

- $rep(F)$ is Scott-continuous
- $rep(F)(a.h) = a.rep(F)(h)$ whenever $a \geq 0$ (positively homogeneous)
- $rep(F)(h + h') \leq rep(F)(h) + rep(F)(h')$ (sublinear, convex)
- $rep(F)(a + h) = a + rep(F)(h)$, a cst. (normalisation)
- $rep(F)$ maps $[X \to \{0 < 1\}]$ into $\{0 < 1\}$ (discrete)

Representation Theorems

Definition ([JGL, CSL07])

A prevision is a continuous, positively homogeneous map $P : [X \to \overline{\mathbb{R}^+}] \to \overline{\mathbb{R}^+}$.

$$\nu \in \mathbf{V}_1(X) \xrightarrow{rep} \text{linear prevision}$$

$$F \in \mathcal{H}(X) \xrightarrow{rep} \text{convex discrete prevision}$$

$$F \in \boldsymbol{T}^{\text{cvx}}(X) = \mathcal{H}^{\text{cvx}}(\mathbf{V}_1(X)) \xrightarrow{rep} \text{convex prevision}$$

Representation Theorems

Definition

A prevision is a continuous, positively homogeneous map $P : [X \to \overline{\mathbb{R}^+}] \to \overline{\mathbb{R}^+}$.

Theorem ([JGL, FOSSACS08])

The following are isomorphisms:

$$\nu \in \mathbf{V}_1(X) \xrightarrow{rep} \textit{linear prevision}$$

$$F \in \mathcal{H}(X) \xrightarrow{rep} \textit{convex discrete prevision}$$

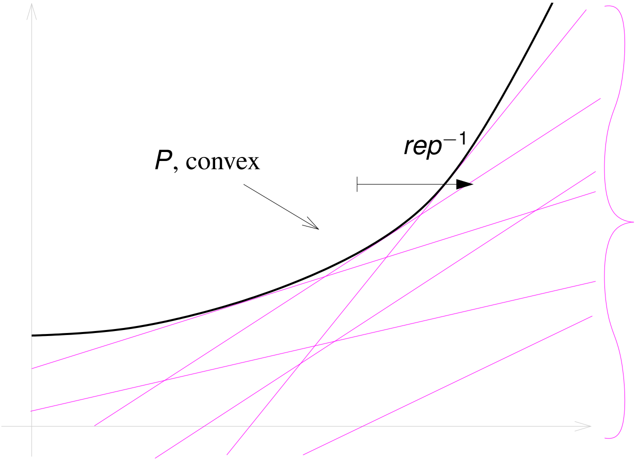
$$F \in \mathbf{T}^{cvx}(X) = \mathcal{H}^{cvx}(\mathbf{V}_1(X)) \xrightarrow{rep} \textit{convex prevision}$$

Idea of the Proof

Build rep^{-1}:

- (Non-det. case) P convex discrete prev. $\mapsto F \in \mathcal{H}(X)$:
  $F =$ complement of largest open U with $P(\chi_U) = 0$
- (Prob. case) P linear prev. $\mapsto \nu \in \mathbf{V}_1(X)$:
  $\nu = \lambda U \text{ open} \cdot P(\chi_U)$ [Tix, Diplom, 95]
- (Mixed case) P convex prev. $\mapsto F \in \boldsymbol{T}^{\text{cvx}}(X)$:
  see next slide

Idea of the Proof: “Subtangents”



set of
linear previsions
(=valuations)
below P:
closed,
convex,
non-empty
(Hahn-Banach
type argument)

Monads of Previsions

The point with previsions is that the monad structure is simple:

- Let $\mathrm{Prev}_\bullet(X) = [[X \to \overline{\mathbb{R}^+}] \to_\bullet \overline{\mathbb{R}^+}]$,
  where $\bullet$ can be "linear", "convex", "convex discrete"
- Unit $\eta(x) = \lambda h_{[X \to \overline{\mathbb{R}^+}]} \cdot h(x)$
- Multiplication: for $P : \mathrm{Prev}^2_\bullet(X)$,

$$\mu(P) = \lambda h_{[X \to \overline{\mathbb{R}^+}]} \cdot P(\lambda Q_{\mathrm{Prev}_\bullet(X)} \cdot Q(h))$$

- Non-det. choice $P \uplus P' = \lambda h_{[X \to \overline{\mathbb{R}^+}]} \cdot \max(P(h), P'(h))$
  (if $\bullet$ is "convex" or "convex discrete")
- Prob. choice $P \oplus_p P' = \lambda h_{[X \to \overline{\mathbb{R}^+}]} \cdot p.P(h) + (1-p).P'(h)$
  (if $\bullet$ is "convex" or "linear")

Monads of Previsions

The point with previsions is that the monad structure is simple:
...to make it simpler (to a logician), write $\neg X$ for $[X \to \overline{\mathbb{R}^+}]$

- Let $\text{Prev}_\bullet(X) = \neg_\bullet \neg X$,
  where $\bullet$ can be "linear", "convex", "convex discrete"
- Unit $\eta(x) = \lambda h_{\neg X} \cdot h(x)$
- Multiplication: for $P : \neg_\bullet \neg\ \neg_\bullet \neg X$,

$$\mu(P) = \lambda h_{\neg X} \cdot P(\lambda Q_{\neg_\bullet \neg X} \cdot Q(h))$$

- Non-det. choice $P \uplus P' = \lambda h_{\neg X} \cdot \max(P(h), P'(h))$
  (if $\bullet$ is "convex" or "convex discrete")
- Prob. choice $P \oplus_p P' = \lambda h_{\neg X} \cdot p.P(h) + (1-p).P'(h)$
  (if $\bullet$ is "convex" or "linear")

Monads of Previsions

The point with previsions is that the monad structure is simple:

...to make it simpler (to a logician), write $\neg X$ for $[X \to \overline{\mathbb{R}^+}]$

- Let $\mathrm{Prev}_\bullet(X) = \neg_\bullet \neg X$,

  where $\bullet$ can be "linear", "convex", "convex discrete"
- Unit $\eta(x) = \lambda h_{\neg X} \cdot h(x)$
- Multiplication: for $P : \neg_\bullet \neg \ \neg_\bullet \neg X$,

$$\mu(P) = \lambda h_{\neg X} \cdot P(\lambda Q_{\neg_\bullet \neg X} \cdot Q(h))$$

- Non-det. choice $P \uplus P' = \lambda h_{\neg X} \cdot \max(P(h), P'(h))$

  (if $\bullet$ is "convex" or "convex discrete")
- Prob. choice $P \oplus_p P' = \lambda h_{\neg X} \cdot p.P(h) + (1 - p).P'(h)$

  (if $\bullet$ is "convex" or "linear")

Yes, this is a continuation semantics [JGL, CSL07]

Outline

Full Abstraction as a Test Problem

- We shall define two variants of PCF [Plotkin77]:
  - $\text{PCF}_{\mathtt{A}}$: PCF with angelic non-deterministic choice
  - $\text{PCF}_{\mathtt{AP}}$: same + probabilistic (valuation) choice
- We shall imitate Plotkin's classic study:
  - soundness
  - computational adequacy
  - full abstraction (or failure thereof)
- Will work extremely well using previsions (for the denotational side)

Syntax

$$\begin{array}{rcll} \gamma & ::= & \mathtt{Nat} \mid \mathtt{S} & \text{Ground types} \\ \sigma, \tau & ::= & \gamma \mid \sigma \to \tau \mid \mathtt{T}\tau & \text{Types} \end{array}$$

Notes:

- $\mathtt{S}$ has only one (non-bottom) value
  = $\mathtt{unit}$ type, termination type

  Not required in principle, but practical

- $\mathtt{T}\tau$ type of processes computing value of type τ

  à la [Moggi91]

PCF_S Terms

Define $\mathrm{PCF}_{\mathtt{A}}$, $\mathrm{PCF}_{\mathtt{AP}}\ldots$ in general PCF_S where $S \subseteq \{\mathtt{A}, \mathtt{P}\}$

- PCF terms
  (λ-calculus + basic arithmetic + $\mathtt{ifz}$ + fixpoint Y)
- At $\mathtt{S}$ type:
  - $\underline{\top} : \mathtt{S}$
  - for every $M : \mathtt{Nat}$, $\mathtt{ignore}\, M : \mathtt{S}$
  - for all $M : \mathtt{S}$, $N : \sigma$, sequencing $M; N : \sigma$
- At $\mathtt{T}\tau$ types:
  - for each $M : \tau$, $\mathtt{val}\, M : \mathtt{T}\tau$
  - for all $M : \sigma$, $N : \mathtt{T}\tau$, $\mathtt{let}\, x \Leftarrow M\, \mathtt{in}\, N : \mathtt{T}\tau$
  - Non-det. choice $\circledvee : \mathtt{T}\tau \to \mathtt{T}\tau \to \mathtt{T}\tau$ (if $\mathtt{A} \in S$)
  - Prob. choice $\oplus : \mathtt{T}\tau \to \mathtt{T}\tau \to \mathtt{T}\tau$ (if $\mathtt{P} \in S$)

Operational Semantics

As a machine (a transition system)
working on configurations $E \cdot M$

- $M : \sigma$ is PCF(S) term
- Contexts $E : (\sigma \vdash \mathtt{TS})$ = stacks of pending operations:

$$
\begin{array}{rcl}
E & := & \_ \\
 & | & E[\_ N] \\
 & | & E[\mathtt{succ}\ \_] \mid E[\mathtt{pred}\ \_] \\
 & | & E[\mathtt{ifz}\ \_\ N\ P] \mid E[\_; N] \\
 & | & E[\mathtt{ignore}\ \_] \\
 & | & \mathtt{val}\ \_ \\
 & | & E[\mathtt{let}\ x \Leftarrow \_\ \mathtt{in}\ N]
\end{array}
$$

The PCF$_S$ Machine: 1. Redex Discovery Rules

Purpose: move top of M into context E, until redex appears

$$
\begin{array}{rcl}
E \cdot MN & \to & E[\_N] \cdot M \\
E[\_N] \cdot \mathtt{pred} & \to & E[\mathtt{pred}\,\_] \cdot N \\
E[\_N] \cdot \mathtt{succ} & \to & E[\mathtt{succ}\,\_] \cdot N \\
E[\_MNP] \cdot \mathtt{ifz} & \to & E[\mathtt{ifz}\,\_\,N\;P] \cdot M \\
E \cdot M; N & \to & E[\_; N] \cdot M \\
E[\_N] \cdot \mathtt{ignore} & \to & E[\mathtt{ignore}\,\_] \cdot N \\
\_ \cdot \mathtt{val}\,M & \to & \mathtt{val}\,\_ \cdot M \\
E \cdot \mathtt{let}\,x \Leftarrow M\,\mathtt{in}\,N & \to & E[\mathtt{let}\,x \Leftarrow \_\,\mathtt{in}\,N] \cdot M
\end{array}
$$

The PCF$_S$ Machine: 2. Computation

Redex = interaction between top of M and bottom of E

$$
\begin{array}{rcl}
E[\_N] \cdot \lambda x \cdot P & \to & E \cdot P[x := N] \\
E[\mathtt{pred}\,\_] \cdot \underline{n+1} & \to & E \cdot \underline{n} \\
E[\mathtt{succ}\,\_] \cdot \underline{n} & \to & E \cdot \underline{n+1} \\
E[\mathtt{ifz}\,\_\,N\,P] \cdot \underline{0} & \to & E \cdot N \\
E[\mathtt{ifz}\,\_\,N\,P] \cdot \underline{n+1} & \to & E \cdot P \\
E[\_;N] \cdot \underline{\top} & \to & E \cdot N \\
E[\mathtt{ignore}\,\_] \cdot \underline{n} & \to & E \cdot \underline{\top} \\
E[\_N] \cdot Y & \to & E \cdot N(YN) \\
E[\mathtt{let}\,x \Leftarrow \_\,\mathtt{in}\,P] \cdot \mathtt{val}\,N & \to & E \cdot P[x := N]
\end{array}
$$

The PCF$_S$ Machines: 3. Choice and Testing

The above rules (groups 1., 2.) cannot handle choice $\varovee$, $\oplus$
$\Rightarrow$ Use judgments $E \cdot M \downarrow a$, $a \in \mathbb{Q} \cap [0,1]$.

"The probability that $E \cdot M$ may terminate is $> a$"

groups 1., 2. (where $C \to C'$)	choice
$$\frac{C' \downarrow a}{C \downarrow a}$$	$$\frac{E \cdot M \downarrow a}{E[\_MN] \cdot \varovee \downarrow a} \qquad \frac{E \cdot N \downarrow a}{E[\_MN] \cdot \varovee \downarrow a}$$
termination	
$$\frac{}{\mathtt{val}\ \_ \cdot \underline{\top} \downarrow a}$$	$$\frac{E \cdot M \downarrow a \quad E \cdot N \downarrow b}{E[\_MN] \cdot \oplus \downarrow \frac{1}{2}(a+b)}$$

Probabilistic May-Testing

Definition

$\Pr(E \cdot M{\downarrow}) = \sup\{a \in \mathbb{Q} \in [0,1] \mid E \cdot M \downarrow a \text{ derivable}\}$

- $\Pr(\mathtt{val}\ \_ \cdot \top{\downarrow}) = 1$
- $\Pr(E[\_MN] \cdot \oplus{\downarrow}) = \frac{1}{2}(\Pr(E \cdot M{\downarrow}) + \Pr(E \cdot N{\downarrow}))$
- $\Pr(E[\_MN] \cdot \vee\!\!\!\!\bigcirc{\downarrow}) = \max(\Pr(E \cdot M{\downarrow}), \Pr(E \cdot N{\downarrow}))$

Denotational Semantics

$$[\![x]\!]_S = x \quad [\![\top]\!]_S = \top \quad [\![\underline{n}]\!]_S = n \in \mathbb{N}$$

$$[\![\lambda x \cdot M]\!]_S = (x \mapsto [\![M]\!]_S) \qquad [\![MN]\!]_S(\rho) = [\![M]\!]_S([\![N]\!]_S)$$

$$\begin{array}{rcll}
[\![Y]\!]_S & = & (f \mapsto \bigcup_{n\in\mathbb{N}} f^n(\bot)) & \\
[\![\mathtt{pred}]\!]_S & = & (v \in \mathbb{N} \smallsetminus \{0\} \mapsto v-1 \mid 0, \bot \mapsto \bot) & \\
[\![\mathtt{succ}]\!]_S & = & (v \in \mathbb{N} \mapsto v+1 \mid \bot \mapsto \bot) & \\
[\![\mathtt{ifz}]\!]_S & = & (0, t, e \mapsto t \mid n \in \mathbb{N} \smallsetminus \{0\}, t, e \mapsto e \mid \bot \mapsto \bot) & \\
[\![M; N]\!]_S & = & [\![N]\!]_S \text{ if } [\![M]\!]_S \neq \bot, \text{ else } \bot & \\
[\![\mathtt{ignore}]\!]_S & = & (n \in \mathbb{N} \mapsto \top \mid \bot \to \bot) & \\
[\![\mathtt{val}\, M : \mathtt{T}\sigma]\!]_S & = & (h \mapsto h([\![M]\!]_S)) & \text{(unit)} \\
[\![\mathtt{let}\, x \Leftarrow M \,\mathtt{in}\, N]\!]_S & = & (h \mapsto [\![M]\!]_S(x \mapsto [\![N]\!]_S(h)) & \text{(mult.)} \\
[\![\textcircled{\vee}]\!]_S & = & (F_1, F_2, h \mapsto \max(F_1(h), F_2(h)) & (\text{if } \mathtt{A} \in S) \\
[\![\oplus]\!]_S & = & (F_1, F_2, h \mapsto \frac{1}{2}(F_1(h) + F_2(h)) & (\text{if } \mathtt{P} \in S)
\end{array}$$

The Purely Non-Deterministic Case

When $S = \mathbb{A}$, retrieve the usual semantics, up to the *rep* iso:

$$[\![x]\!]_S = x \quad [\![\top]\!]_S = \top \quad [\![\underline{n}]\!]_S = n \in \mathbb{N}$$

$$[\![\lambda x \cdot M]\!]_S = (x \mapsto [\![M]\!]_S) \qquad [\![MN]\!]_S(\rho) = [\![M]\!]_S([\![N]\!]_S)$$

$$\begin{aligned}
[\![Y]\!]_S &= (f \mapsto \bigcup_{n\in\mathbb{N}} f^n(\bot)) \\
[\![\mathtt{pred}]\!]_S &= (v \in \mathbb{N} \setminus \{0\} \mapsto v-1 \mid 0, \bot \mapsto \bot) \\
[\![\mathtt{succ}]\!]_S &= (v \in \mathbb{N} \mapsto v+1 \mid \bot \mapsto \bot) \\
[\![\mathtt{ifz}]\!]_S &= (0, t, e \mapsto t \mid n \in \mathbb{N} \setminus \{0\}, t, e \mapsto e \mid \bot \mapsto \bot) \\
[\![M; N]\!]_S &= [\![N]\!]_S \text{ if } [\![M]\!]_S \neq \bot, \text{ else } \bot \\
[\![\mathtt{ignore}]\!]_S &= (n \in \mathbb{N} \mapsto \top \mid \bot \to \bot) \\
[\![\mathtt{val}\, M : \mathtt{T}\sigma]\!]_S &= \downarrow\{[\![M]\!]_S\} && \text{(unit)} \\
[\![\mathtt{let}\, x \Leftarrow M \,\mathtt{in}\, N]\!]_S &= \bigcup_{x \in [\![M]\!]_S} [\![N]\!]_S && \text{(mult.)} \\
[\![\ovee]\!]_S &= (F_1, F_2 \mapsto F_1 \cup F_2)
\end{aligned}$$

The Purely Probabilistic Case

When $S = \mathtt{P}$, retrieve the usual semantics, up to the *rep* iso:

$$[\![x]\!]_S = x \quad [\![\top]\!]_S = \top \quad [\![\underline{n}]\!]_S = n \in \mathbb{N}$$

$$[\![\lambda x \cdot M]\!]_S = (x \mapsto [\![M]\!]_S) \qquad [\![MN]\!]_S(\rho) = [\![M]\!]_S([\![N]\!]_S)$$

$$\begin{aligned}
[\![Y]\!]_S &= (f \mapsto \bigcup_{n \in \mathbb{N}} f^n(\bot)) \\
[\![\mathtt{pred}]\!]_S &= (v \in \mathbb{N} \smallsetminus \{0\} \mapsto v - 1 \mid 0, \bot \mapsto \bot) \\
[\![\mathtt{succ}]\!]_S &= (v \in \mathbb{N} \mapsto v + 1 \mid \bot \mapsto \bot) \\
[\![\mathtt{ifz}]\!]_S &= (0, t, e \mapsto t \mid n \in \mathbb{N} \smallsetminus \{0\}, t, e \mapsto e \mid \bot \mapsto \bot) \\
[\![M; N]\!]_S &= [\![N]\!]_S \text{ if } [\![M]\!]_S \neq \bot, \text{ else } \bot \\
[\![\mathtt{ignore}]\!]_S &= (n \in \mathbb{N} \mapsto \top \mid \bot \to \bot) \\
[\![\mathtt{val}\, M : \mathtt{T}\sigma]\!]_S &= \delta_{[\![M]\!]_S} && \text{(unit)} \\
[\![\mathtt{let}\, x \Leftarrow M \,\mathtt{in}\, N]\!]_S &= (U \text{ open} \mapsto \int_x [\![N]\!]_S(U) d\, [\![M]\!]_S) && \text{(mult.)} \\
[\![\oplus]\!]_S &= (\nu_1, \nu_2 \mapsto \tfrac{1}{2}\nu_1 + \tfrac{1}{2}\nu_2)
\end{aligned}$$

Soundness

In usual PCF, soundness states that if $M \to^* V$ then $[\![M]\!] = [\![V]\!]$.

Theorem (Soundness)

Let $\blacklozenge = \chi_{\{\top\}} : [\![\mathsf{S}]\!] \to I$ *map* $\bot$ *to* 0, $\top$ *to* 1 *(termination test)*

- $[\![E[M]]\!]_S(\blacklozenge) \geq Pr(E \cdot M\downarrow)$

Proof: easy; one has to notice that E has a special shape. □

Computational Adequacy

In usual PCF, $M \to^* V$ iff $[\![M]\!] = [\![V]\!]$, at ground types.
Here, use $E = \_$ (empty context, of type $\mathtt{TS} \vdash \mathtt{TS}$)

Theorem (Computational Adequacy)

- $[\![M]\!]_S(\blacklozenge) = Pr(\_ \cdot M\downarrow)$

Proof: Details omitted. The key point is defining a logical relation by double orthogonality on monadic types:

$M \, R_{\sigma\to\tau} \, f$ iff for all $N \, R_\sigma \, v$, $MN \, R_\tau \, f(v)$
$M \, R_{\mathrm{T}\sigma} \, F$ iff for all $E \, R_\sigma^\perp \, h$, $\Pr(E \cdot M\downarrow^m) \geq F(h)$
$E \, R_\sigma^\perp \, h$ iff for all $Q \, R_\sigma \, v$, $\Pr(E \cdot \mathtt{val}\ Q\downarrow^m) \geq h(v)$

Note: Pr is operational, $F(h)$, $h(v)$ is the prevision monad at work

Full Abstraction

Definition (Observational Preorder)

Let $M \precsim N$ iff $\Pr(E \cdot M\downarrow) \leq \Pr(E \cdot N\downarrow)$ for every E

Definition (Full Abstraction)

$M \precsim N$ iff $[\![M]\!]_S \leq [\![N]\!]_S$, at all types.

- Fails for PCF without choice [Plotkin77]
- True for PCF without choice+parallel or [Plotkin77]
- True for $\mathrm{PCF}_{\mathtt{A}}$ (with a restriction on public types), as we shall see
- Fails for $\mathrm{PCF}_{\mathtt{AP}}$, even with par. or [JGL, Domains X, 2011] ... we shall need statistical termination testers (and no parallel or)

Public Types

Definition (Public Types, τ_{pub})

$$
\begin{aligned}
\tau_{\text{pub}} &::= \mathtt{S} \mid \cancel{\mathtt{Nat}} \mid \tau_{\text{priv}} \to \tau_{\text{pub}} \mid \mathtt{T}\tau_{\circ} \\
\tau_{\circ} &::= \mathtt{S} \mid \mathtt{Nat} \mid \tau_{\text{priv}} \to \tau_{\circ} \mid \mathtt{T}\tau_{\circ} \\
\tau_{\text{priv}} &::= \mathtt{S} \mid \mathtt{Nat} \mid \tau_{\text{pub}} \to \tau_{\text{priv}} \mid \mathtt{T}\tau_{\text{priv}}
\end{aligned}
$$

- $M : \tau_{\text{pub}}$ promises it won't return a `Nat`;
  but can return a `T Nat`
- ...so cannot be denied the possibility of making choices
- Private terms guarantee they won't deny this possibility to their arguments either

Public Types

Definition (Public Types, τ_{pub})

$$\begin{array}{rcl} \tau_{\text{pub}} & ::= & \mathtt{S} \mid \cancel{\mathtt{Nat}} \mid \tau_{\text{priv}} \to \tau_{\text{pub}} \mid \mathtt{T}\tau_{\circ} \\ \tau_{\circ} & ::= & \mathtt{S} \mid \mathtt{Nat} \mid \tau_{\text{priv}} \to \tau_{\circ} \mid \mathtt{T}\tau_{\circ} \\ \tau_{\text{priv}} & ::= & \mathtt{S} \mid \mathtt{Nat} \mid \tau_{\text{pub}} \to \tau_{\text{priv}} \mid \mathtt{T}\tau_{\text{priv}} \end{array}$$

Mathematically, what's important is:

- $[\![\tau]\!]_S$ is a bc-domain for every type τ
- $[\![\tau_\circ]\!]_S$ has a basis of definable elements
- $[\![\tau_{\text{pub}}]\!]_S$ has a basis of definable elements and definable finite sups (through $\circledvee$, on monadic types $\mathtt{T}\tau$)
- The topology of $[\![\tau_{\text{priv}}]\!]_S$ has a basis of definable opens$^{(*)}$
  ($U \subseteq [\![\tau]\!]_S$ open is definable iff $\chi_U : [\![\tau \to \mathtt{S}]\!]_S$ definable)

The Case of $\text{PCF}_{\mathbb{A}}$

When $S = \mathbb{A}$, $(*)$ is literally true:

- The topology of $[\![\tau_{\text{priv}}]\!]_S$ has a basis of definable opens$^{(*)}$

Theorem (Full Abstraction, Non-Deterministic Case)

Let M, N be two closed $\text{PCF}_{\mathbb{A}}$ terms of public type τ_{pub}. Then $M \precsim N$ iff $[\![M]\!]_{\mathbb{A}} \leq [\![N]\!]_{\mathbb{A}}$.

(No need for parallel or. . . which is definable anyway)

The Case of PCF$_{\mathtt{AP}}$

When $S = \mathtt{AP}$, $(*)$ is wrong:

- Every definable map in $[\![\mathtt{T}\ \mathtt{S} \to \mathtt{S}]\!]_S$ is constant
- But $[\![\mathtt{T}\ \mathtt{S} \to \mathtt{S}]\!]_S$ contains plenty of non-constant maps
- I.e., $[\![\mathtt{T}\ \mathtt{S}]\!]_S$ contains plenty of non-trivial opens
  at least all $[h > r] = \{P \mid P(h) > r\}$, for $h\colon [\![\mathtt{S}]\!]_S \to \overline{\mathbb{R}^+}$

and full abstraction fails:

- Let $M = \lambda f_{\mathtt{T}\ \mathtt{S}\to\mathtt{S}} \cdot f(\mathtt{val}\ \underline{\top})$, $N = \lambda f_{\mathtt{T}\ \mathtt{S}\to\mathtt{S}} \cdot f(\mathtt{val}\ \Omega)$
- $M \precsim N$, in fact $\Pr(E \cdot M\!\downarrow) = \Pr(E \cdot N\!\downarrow)$
  because the only f we can feed them are constant
- But certainly $[\![M]\!]_S \not\leq [\![N]\!]_S$, since
  $[\![M]\!]_S([\blacklozenge > \frac{1}{2}]) = 1 > [\![N]\!]_S([\blacklozenge > \frac{1}{2}]) = 0$

The Case of $\text{PCF}_{\texttt{AP}}$

So add syntax for $[h > r]$:

Definition (Statistical Termination Testers)

For each $M : \texttt{TS}$, add $\texttt{Pr}(M > r)$, with following semantics

$$\frac{\_ \cdot M \downarrow r \quad E \cdot \underline{\top} \downarrow a}{E \cdot \texttt{Pr}(M > r) \downarrow a} \qquad [\![\texttt{Pr}(M > r)]\!]_{\texttt{AP}} = \begin{cases} \top & \text{if } [\![M]\!]_{\texttt{AP}}(\blacklozenge) > r \\ \bot & \text{otherwise} \end{cases}$$

Soundness, computational adequacy still OK, and:

Theorem (Full Abstraction, Non-Deterministic Case)

Let M, N be two closed $\text{PCF}_{\texttt{AP}} + \texttt{Pr}$ terms of public type τ_{pub}. Then $M \precsim N$ iff $[\![M]\!]_{\texttt{AP}} \leq [\![N]\!]_{\texttt{AP}}$.

Outline

What About Other Forms of Non-Determinism?

- Angelic (may-testing, this talk): everything works well . . . . ☺
- Demonic (must-testing):
  - MOW-TKP and prevision models . . . . ☺
  - Incompatible with $\mathcal{JV}$, CRV (no distribution law) ☹
  - Soundness, computational adequacy . . . . ☺
  - Full abstraction only with quirky `irq` primitive ☹
- Erratic (may-must-testing):
  - MOW-TKP and prevision models (with coherence) . . . . ☺
  - Incompatible with $\mathcal{JV}$, CRV (no distribution law) ☹
  - Soundness, computational adequacy . . . . ☺
  - Full abstraction as easy as angelic+demonic ☺
- No non-determinism at all, probabilities only:
  - a mess (Jung-Tix again + more problems) ☹

. . . and similarly for call-by-value forms (where all types are public) [JGL, Domains X, 2011]

What about $\mathcal{JV}$ and CRVs?

- What semantics would make CRV-based (not $\mathbf{V}_1$-based) probabilistic PCF (+non-det. choice+something?) fully abstract?

  ... we probably need a random bit generator rigger,
  replacing the first few (random) bits
  by deterministically generated ones

- Then statistical termination testers can be replaced by explicit Monte-Carlo tests.

Closing Word

- We now understand choice better
- Interesting pairs:
  - full non-determistic choice / random variables (or $\mathcal{JV}$)
  - convex non-deterministic choice / valuations
- Full abstraction works more smoothly that in the pure PCF case