Lecture Notes
Part III of MPRI 2 – 02
2021 - 2021

Michele Pagani
pagani@irif.fr

February 4, 2022

Contents

List of exercises

These notes are a continuation of the lecture notes by Thomas Ehrhard, `https://www.irif.fr/~ehrhard/pub/mpri-2020-2021.pdf`.

1 The Probabilistic Extension pPCF of PCF

1.1 The Syntax of pPCF

Figure 1 sketches the probabilistic extension of PCF, written pPCF. Let Γ be a typing context and A be a type, we denote by Λ_Γ^A the set of all terms M such that $\Gamma \vdash M : A$. In the case where Γ is empty, and so the elements of Λ_Γ^A are closed, we use Λ_0^A to denote that set. A *program* will be a closed term of pPCF of ground type ι, i.e. an element of Λ_0^ι

By a simple inspection of the typing rules, the reader can check the following.

Remark: Let M be a term and Γ be a typing context. There is at most one type A such that $\Gamma \vdash M : A$.

Exercise 1. Give an example of expression M generated by the grammar of Figure 1b, such that M cannot be typed by the rules of Figure 1c. Can you find an M using only abstractions, applications and variables? and another M using only only variables, numerals, coin, branchings and $\mathsf{succ}(M)$?

Answer of Exercise 1. *The expressions* $\mathsf{succ}(\lambda x.x)$ *or* $\lambda x.\,(x)\,x$ *cannot be simply typed. By structural induction, one can prove that an expression generated with only variables, numerals,* coin*, branchings and* $\mathsf{succ}(M)$ *is always typable with the ground type* ι*.*

The *reduction relation* for evaluating pPCF terms is given in Figure 1d. In the β-rule (topmost leftmost rule of Figure 1d), the term $M\,[N/x]$ stands for M where the variable x is substituted with the term N, avoiding the capture of the free variables in N. If $M \xrightarrow{p} M'$ is the conclusion of one axiom rule (i.e. one of the rules in the first three lines of Figure 1d), then we call M the *redex* of the reduction, M' its *contractum* and p the *probability* to happen. This reduction is called *weak-head reduction* (or simply weak reduction) since it always reduces the leftmost outermost redex and never reduces redexes under abstractions. We say that M is *weak-normal*, or a *value*, if there is no reduction $M \xrightarrow{p} M'$.

Lemma 1 (Substitution) *Assume* $\Gamma, x : A \vdash M : B$ *and* $\Gamma \vdash N : A$*, then* $\Gamma \vdash M\,[N/x] : B$*.*

Exercise 2. Prove Lemma 1.

Answer of Exercise 2. *By induction on the derivation of* $\Gamma, x : A \vdash M : B$*.*

Proposition 2 (Subject reduction) *Assume* $M \xrightarrow{p} M'$*. If* $\Gamma \vdash M : A$*, then* $\Gamma \vdash M' : A$*.*

Exercise 3. Give a proof of Proposition 2.

Answer of Exercise 3. *By structural induction on a derivation of* $M \xrightarrow{p} M'$*. All cases are easy, but for the* β *and if-reduction, where the substitution lemma should be used.*

$$A, B, \ldots := \iota \mid A \Rightarrow B$$

(a) The grammar of types, ι is the *ground type* of natural numbers.

$$M, N, \ldots := \underline{n} \mid x \mid \mathsf{succ}(M) \mid \mathsf{if}(M, P, z \cdot R) \mid \lambda x^A\, M \mid (M)\, N \mid \mathsf{fix}(M) \mid \mathsf{coin}$$

(b) The grammar of terms, with $n \in \mathbb{N}$, $p \in [0,1]$, and x, $y \ldots$ variables.

$$\frac{}{\Gamma \vdash \underline{n} : \iota} \qquad \frac{}{\Gamma, x : A \vdash x : A} \qquad \frac{}{\Gamma \vdash \mathsf{coin} : \iota} \qquad \frac{\Gamma \vdash M : \iota}{\Gamma \vdash \mathsf{succ}(M) : \iota}$$

$$\frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x^A\, M : A \Rightarrow B} \qquad \frac{\Gamma \vdash M : A \Rightarrow B \qquad \Gamma \vdash N : A}{\Gamma \vdash (M)\, N : B}$$

$$\frac{\Gamma \vdash M : A \Rightarrow A}{\Gamma \vdash \mathsf{fix}(M) : A} \qquad \frac{\Gamma \vdash M : \iota \qquad \Gamma \vdash P : A \qquad \Gamma, z : \iota \vdash R : A}{\Gamma \vdash \mathsf{if}(M, P, z \cdot R) : A}$$

(c) The typing rules, with $\Gamma = y_1 : A_1, \ldots, y_k : A_k$ a typing context, $k \in \mathbb{N}$ and $y_i \neq y_j$ whenever $i \neq j$.

$$\frac{}{\left(\lambda x^A\, M\right) N \xrightarrow{1} M\,[N/x]} \qquad \frac{}{\mathsf{fix}(M) \xrightarrow{1} (M)\, \mathsf{fix}(M)}$$

$$\frac{}{\mathsf{succ}(\underline{n}) \xrightarrow{1} \underline{n+1}} \qquad \frac{}{\mathsf{if}(\underline{0}, P, z \cdot R) \xrightarrow{1} P} \qquad \frac{}{\mathsf{if}(\underline{n+1}, P, z \cdot R) \xrightarrow{1} R\,[\underline{n}/z]}$$

$$\frac{}{\mathsf{coin} \xrightarrow{1/2} \underline{0}} \qquad \frac{}{\mathsf{coin} \xrightarrow{1/2} \underline{1}}$$

$$\frac{M \xrightarrow{p} M'}{(M)\, N \xrightarrow{p} (M')\, N} \qquad \frac{M \xrightarrow{p} M'}{\mathsf{succ}(M) \xrightarrow{p} \mathsf{succ}(M')}$$

$$\frac{M \xrightarrow{p} M'}{\mathsf{if}(M, P, z \cdot R) \xrightarrow{p} \mathsf{if}(M', P, z \cdot R)}$$

(d) The reduction relation $M \xrightarrow{p} M'$, with $p \in [0,1]$, M, M' pPCF terms.

Figure 1: Résumé of pPCF.

Exercise 4. Give a counterexample to the inverse of subjection reduction, called subject expansion: give an example of reduction $M \xrightarrow{p} M'$ and of type A, environment Γ, such that $\Gamma \vdash M' : A$ but it is false that $\Gamma \vdash M : A$.

Answer of Exercise 4. $M = (\lambda x^{\iota}\, \underline{0})\, y \xrightarrow{d} \underline{0} = M'$. *We have* $\vdash M' : \iota$, *while* M *cannot be typed under the empty context.*

Exercise 5. Characterise the set of closed values of pPCF.

Answer of Exercise 5. *The closed values are either numerals or abstractions. In fact, these are normal forms for* $\xrightarrow{p}$. *Viceversa, if* M *is a closed normal form for* $\xrightarrow{p}$, *we prove that it is a numeral or an abstraction, by structural induction on* M.

Notice that M *cannot be a variable since it is closed, neither a fixpoint nor* coin, *otherwise it would reduce. If* $M = \mathsf{succ}(N)$ *for some closed term* N, *then* N *also must be a normal form (see rules Figure 1d) so that by induction hypothesis* N *is a numeral and hence* $M = \mathsf{succ}(N)$ *is not normal. The case* $M = \mathsf{if}(N, P, z \cdot R)$ *is similar. In case* $M = (P)\, Q$, *we have that* P *also is a closed normal form. By typing,* P *cannot be a numeral, so it is an abstraction and hence* M *is a* β*-redex.*

A *reduction sequence* from a term M to a term M' is a finite sequence $\varphi = (M_i)_{i=0}^{k}$ such that $M_0 = M$, $M_k = M'$ and for every $i < k$, $M_i \xrightarrow{p_i} M_{i+1}$ for some probability $p_i \in [0, 1]$. By inspection of the rules in Figure 1d, the reader can check that the probability p_i in $M_i \xrightarrow{p_i} M_{i+1}$ is unique, given M_i and M_{i+1}. The *length of* φ is k and the *probability* $\mathsf{p}(\varphi)$ *of* φ is the product $\prod_{i=0}^{k-1} p_i$.

We say that a term M *deterministically reduces* to a value V, written $M \to_{\mathsf{d}}{}^{*} V$, if there is a reduction sequence φ from M to V of probability 1. Notice that such a reduction is unique, i.e. any other reduction sequence starting from M is a prefix of φ. The following exercise exploits the deterministic fragment of pPCF.

Exercise 6. Define terms representing the following functions:

1. the predecessor function, i.e. a term pred such that:

$$(\mathsf{pred})\, \underline{n} \to_{\mathsf{d}}{}^{*} \begin{cases} \underline{0} & \text{if } n = 0 \\ \underline{n-1} & \text{if } n > 0 \end{cases}$$

2. the addition function, i.e. a term add such that:

$$(\mathsf{add})\, \underline{n}\, \underline{m} \to_{\mathsf{d}}{}^{*} \underline{n+m}$$

3. the exponential function, i.e. a term exp_2 such that:

$$(\mathsf{exp}_2)\, \underline{n} \to_{\mathsf{d}}{}^{*} \underline{2^n}$$

4. the comparison function, i.e. a term cmp such that:

$$(\mathsf{cmp})\, \underline{n}\, \underline{m} \to_{\mathsf{d}}{}^{*} \begin{cases} \underline{0} & \text{if } n \leq m \\ \underline{1} & \text{if } n > m \end{cases}$$

Answer of Exercise 6.

$$\mathsf{pred} = \lambda x^{\iota}\, \mathsf{if}(x, \underline{0}, z \cdot z) \qquad \mathsf{add} = \lambda x^{\iota}\, \mathsf{fix}(\lambda a^{\iota \Rightarrow \iota}\, \lambda y^{\iota}\, \mathsf{if}(y, x, z \cdot \mathsf{succ}((a)\, z)))$$

$$\mathsf{exp}_2 = \mathsf{fix}(\lambda e^{\iota \Rightarrow \iota}\, \lambda x^{\iota}\, \mathsf{if}(x, 1, z \cdot (\mathsf{add})\, (e)\, z\ (e)\, z))$$

$$\mathsf{cmp} = \mathsf{fix}(\lambda c^{\iota \Rightarrow \iota \Rightarrow \iota}\, \lambda x^{\iota}\, \lambda y^{\iota}\, \mathsf{if}(x, \underline{0}, z \cdot \mathsf{if}(y, \underline{1}, z' \cdot (c)\, z\, z')))$$

The constructor coin is the stochastic primitive of pPCF, leading to different outcomes. Given a term M and a value V, we define the set of different reduction sequences from M to V as:

$$\mathsf{Path}^{\leq n}(M, V) = \{\varphi \mid \varphi \text{ reduction sequence of length } \textit{at most } n \text{ from } M \text{ to } V\} \tag{1}$$

$$\mathsf{Path}(M, V) = \bigcup_{n \in \mathbb{N}} \mathsf{Path}^{\leq n}(M, V) \tag{2}$$

The quantity $\sum_{\varphi \in \mathsf{Path}(M,V)} \mathsf{p}(\varphi)$ defines the probability that M reduces to V. We will formalise this idea in Section 1.3 by representing the reduction relation as a discrete time Markov chain whose states are terms, weak-normal terms being stationary. Before that, let us recall some notions we need in the sequel.

1.2 Compendium of Markov Chains

Let S be a countable set and let $R \in [0,1]^{S \times S}$ be a matrix with S-indexed rows and columns. One says that R is *sub-stochastic* if $\forall i \in S, \sum_{j \in S} R_{i,j} \leq 1$, we call R *stochastic* whenever the previous sum is equal to 1 for all i. Given two such matrices R and T, their *product* RT is given by

$$\forall (i,j) \in S^2, (RT)_{i,j} = \sum_{k \in I} R_{i,k} T_{k,j}$$

which is also a (sub-)stochastic matrix. Given $n \in \mathbb{N}$, we denote by R^n the n-fold product of R, which is the diagonal matrix if $n = 0$.

A stochastic matrix represents a one-step evolution of a discrete-time Markov process. A typical example is a random-walk, as the following one.

Example. Let $S = \mathbb{N}$ and consider the following matrix over $[0,1]^{S \times S}$:

$$W_{i,j} = \begin{cases} 1 & \text{if } i = j = 0, \\ \frac{1}{2} & \text{if } i > 0 \text{ and } (j = i - 1 \text{ or } j = i + 1), \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Notice that W is stochastic. In fact, W defines a Markov process describing a particle travelling over $\mathbb{N}$: once the particle reaches 0, it will stay there, otherwise it will move $+1$ or -1 with equal probability $\frac{1}{2}$. The matrix W^n will then describe the state of the particle after n iterations.

Given a stochastic matrix R over S, the set of *stationary states* of R is defined by:

$$S_1^R = \{i \in S \mid R_{i,i} = 1\} \tag{4}$$

so that if $i \in S_1^R$ and $R_{i,j} \neq 0$ then $i = j$.

Let $(i,j) \in S \times S_1^R$. Then the n-indexed sequence $(R^n)_{i,j} \in [0,1]$ is monotone. Indeed, for all n we have

$$(R^{n+1})_{i,j} = \sum_{k \in S} (R^n)_{i,k} R_{k,j} \geq (R^n)_{i,j} R_{j,j} = (R^n)_{i,j}$$

So we can define a matrix $R^\infty \in [0,1]^{S\times S}$ as follows

$$(R^\infty)_{i,j} = \begin{cases} \sup_{n\in\mathbb{N}} (R^n)_{i,j} & \text{if } (i,j) \in S \times S_1^R \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The matrix S^∞ is a sub-stochastic matrix because, given $i \in I$

$$\begin{aligned} \sum_{j\in S} (R^\infty)_{i,j} &= \sum_{j\in S_1^R} \sup_{n\in\mathbb{N}} (R^n)_{i,j} \\ &= \sup_{n\in\mathbb{N}} \sum_{j\in S_1^R} (R^n)_{i,j} \quad \text{by the monotone convergence theorem} \\ &\leq \sup_{n\in\mathbb{N}} \sum_{j\in S} (R^n)_{i,j} = 1 \end{aligned}$$

1.3 The Markov Chain of pPCF

Given a context Γ and a type A, we consider Λ_Γ^A as a set of states, and we define the reduction relation as a stochastic matrix Red given by

$$\mathsf{Red}(\Gamma, A)_{M,M'} = \begin{cases} p & \text{if } M \xrightarrow{p} M' \\ 1 & \text{if } M \text{ is a value and } M' = M \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

We also use the notation $\mathsf{Red}(A)$ for the matrix $\mathsf{Red}(\Gamma, A)$ when the typing context is empty. Also, we will simply write Red if the typing annotation is irrelevant or clear from the context. The number $\mathsf{Red}(\Gamma, A)_{M,M'}$ is the probability of M to reduce to M' in one step. Notice that all weak-normal terms are stationary states of $\mathsf{Red}(\Gamma, A)$, but not all stationary states are weak-normal terms. Therefore, if V is a weak-normal form, then the n-fold product $\mathsf{Red}(\Gamma, A)^n_{M,V}$ gives the probability that M reduces to V in at most n steps. This is precised by the following proposition (recall notation (1)).

Proposition 3 *Let M be term and V be a value in Λ_Γ^A. One has*

$$\mathsf{Red}(\Gamma, A)^n_{M,V} = \sum_{\varphi \in \mathsf{Path}^{\leq n}(M,V)} \mathsf{p}(\varphi)\,.$$

Hence, $\mathsf{Red}(\Gamma, A)^\infty_{M,V} = \sum_{\varphi\in\mathsf{Path}(M,V)} \mathsf{p}(\varphi)$.

Exercise 7. Prove Proposition 3.

Answer of Exercise 7. *By induction on n. For $n = 0$, if $M = V$, we have $\mathsf{Red}(\Gamma, A)^0_{M,V} = 1$ by definition of diagonal matrix, and $\sum_{\varphi\in\mathsf{Path}^{\leq 0}(M,V)} \mathsf{p}(\varphi) = 1$ as $\varphi \in \mathsf{Path}^{\leq 0}(M,V)$ contains the empty path. If $M \neq V$, then $\mathsf{Red}(\Gamma, A)^0_{M,V} = 0$ as well as $\mathsf{Path}^{\leq 0}(M,V)$ is empty.*

For $n > 0$, we have:

$$\begin{aligned}
\mathsf{Red}(\Gamma, A)^n_{M,V} &= \sum_{M' \in \Lambda^A_\Gamma} \mathsf{Red}(\Gamma, A)_{M,M'} \mathsf{Red}(\Gamma, A)^{n-1}_{M',V} && \textit{by def.} \\
&= \sum_{M' \in \Lambda^A_\Gamma} \mathsf{Red}(\Gamma, A)_{M,M'} \Big(\sum_{\varphi \in \mathsf{Path}^{\leq n-1}(M',V)} \mathsf{p}(\varphi) \Big) && \textit{by IH} \\
&= \sum_{M' \in \Lambda^A_\Gamma} \sum_{\varphi \in \mathsf{Path}^{\leq n-1}(M',V)} \mathsf{Red}(\Gamma, A)_{M,M'} \mathsf{p}(\varphi) \\
&= \sum_{\varphi \in \mathsf{Path}^{\leq n}(M,V)} \mathsf{p}(\varphi) && \textit{by def.}
\end{aligned}$$

The last statement is immediate: $\mathsf{Red}(\Gamma, A)^\infty_{M,V} = \sup_n \mathsf{Red}(\Gamma, A)^n_{M,V} = \sup_n \sum_{\varphi \in \mathsf{Path}^{\leq n}(M,V)} \mathsf{p}(\varphi) = \sum_{\varphi \in \mathsf{Path}(M,V)} \mathsf{p}(\varphi)$.

Exercise 8. Does Red have stationary states that are not weak-head normal terms? and what about Red^2?

Answer of Exercise 8. *The only possible stationary states of* Red *are the weak-head normal terms: the proof is by inspection of the rules in Figure 1d, checking that whenever $M \xrightarrow{1} M'$, we have $M' \neq M$. Indeed, the case of β-reduction is not trivial (notice that in untyped λ-calculus we have that $(\lambda x\,(x)\,x)\,(\lambda x\,(x)\,x) \xrightarrow{1} (\lambda x\,(x)\,x)\,(\lambda x\,(x)\,x)$). In case of* pPCF*, if $M \xrightarrow{1} M$ by β-reduction we should have $M = \big(\lambda x^A\, M_1\big)\, M_2 = M_1[M_2/x]$. This means $M_1 = (P)\,Q$ with $P[M_2/x] = \lambda x^A\, M_1$ and $Q[M_2/x] = M_2$. Moreover, suppose that $\Gamma \vdash M : B$, so that $\Gamma, x : A \vdash M_1 : B$ and $\Gamma \vdash M_2 : A$, with $x : A$ not in Γ. We consider two cases:*

- *if $P = x$, then from $P[M_2/x] = \lambda x^A\, M_1$, we have $M_2 = \lambda x^A\, M_1$, so $A \Rightarrow B = A$, which is impossible;*
- *if $P \neq x$, then from $P[M_2/x] = \lambda x^A\, M_1$, $P = \lambda y^A\, P'$ with $P'[M_2/x] = M_1$. Since x is a free variable in M_1, this means that M_2 should have x free also. But this contradicts the fact that $\Gamma \vdash M_2 : A$, with x not in Γ.*

On the contrast, the term $\mathsf{fix}(\lambda x.x)$ is an example of not weak-head normal term but stationary for Red^2, in fact $\mathsf{fix}(\lambda x.x) \xrightarrow{1} (\lambda x.x)\,\mathsf{fix}(\lambda x.x) \xrightarrow{1} \mathsf{fix}(\lambda x.x)$.

Exercise 9. A stochastic program can have different notions of termination. Given a program M, we say that :

- *M strongly terminates* (ST), whenever the set $\bigcup_n \mathsf{Path}(M, \underline{n})$ is finite;
- *M positively almost surely terminates* (PAST), whenever the expected runtime

$$\sum_{n \in \mathbb{N}} \sum_{\varphi \in \mathsf{Path}(M, \underline{n})} \mathsf{p}(\varphi) \mathrm{length}(\varphi)$$

  is finite;
- *M almost surely terminates* (AST), whenever $\sum_n \mathsf{Red}^\infty_{M,\underline{n}} = 1$.

Prove that ST $\rightarrow$ PAST $\rightarrow$ AST and that no implication can be inverted. (*This exercise is not trivial. You can have a look at [1] to have some inspiration...*).

Answer of Exercise 9. ST $\to$ PAST *is immediate (notice that* $\bigcup_n \mathsf{Path}(M,\underline{n})$ *is a disjoint sum).*
As for PAST $\to$ AST*. By Proposition 3 we have that:*

$$\sum_{n\in\mathbb{N}}\sum_{\varphi\in\mathsf{Path}(M,\underline{n})}\mathsf{p}(\varphi)\mathrm{length}(\varphi)=\sum_{k=1}^{\infty}(1-\sum_{n\in\mathbb{N}}\mathsf{Red}^k_{M,\underline{n}})$$

Then, $\sum_{k=1}^{\infty}(1-\sum_{n\in\mathbb{N}}\mathsf{Red}^k_{M,\underline{n}})<\infty$ *implies* $\lim_{k\to\infty}(1-\sum_{n\in\mathbb{N}}\mathsf{Red}^k_{M,\underline{n}})=0$*, so* $\lim_{k\to\infty}\sum_{n\in\mathbb{N}}\mathsf{Red}^k_{M,\underline{n}}=1$ *and we conclude as* $\lim_{k\to\infty}\sum_{n\in\mathbb{N}}\mathsf{Red}^k_{M,\underline{n}}=\sum_{n\in\mathbb{N}}\mathsf{Red}^\infty_{M,\underline{n}}$.
For the counterexamples of the inversions, consider the terms:

$$M_1=\mathsf{fix}(\lambda x^{\iota}\,\mathsf{if}(\mathsf{coin},x,z\cdot\underline{0}))$$
$$M_2=\mathsf{fix}(\lambda f^{\iota\Rightarrow\iota}\,\lambda x^{\iota}\,\mathsf{if}(x,\mathsf{if}(\mathsf{coin},\underline{0},z\cdot(f)\,(\mathsf{exp}_2)\,x),z\cdot(f)\,z))\underline{1}$$

Clearly M_1 *is not ST. However, one can check that* M_1 *reduces to itself in* 4 *steps with probability* $\frac{1}{2}$ *and to* $\underline{0}$ *always in* 4 *steps with probability* $\frac{1}{2}$*. So that:*

$$\sum_{n\in\mathbb{N}}\sum_{\varphi\in\mathsf{Path}(M_1,\underline{n})}\mathsf{p}(\varphi)\mathrm{length}(\varphi)=\sum_{i=1}^{\infty}\frac{4i}{2^i}=2\sum_{i=1}^{\infty}\frac{i}{2^{i-1}}=8$$

so that M_1 *is PAST. Concerning* M_2*, one have that the expected runtime diverges as the reduction sequences are of length exponentials in the number of probabilistic choices.* M_2 *however is easily proven to be AST.*

1.4 Basic Examples

We illustrate the expressive power of pPCF by encoding in this language simple probabilistic algorithms. We explain intuitively the behaviour of these programs, but a formal proof of their soundness would require more sophisticated tools, like a denotational semantics. In fact, the next section will provide one of such semantics, based on probabilistic coherence spaces.

"Let" construction. This version of pPCF, which is globally call-by-name, offers however the possibility of handling integers in a call-by-value way. For instance, we can define the typical call-by-value "let" construction as follows

$$\mathsf{let}\ x\ \mathsf{be}\ M\ \mathsf{in}\ N=\mathsf{if}(M,N\left[\underline{0}/x\right],z\cdot N\left[\mathsf{succ}(z)/x\right]) \tag{7}$$

and this construction is restricted to the type of natural numbers; it can be typed as:

$$\frac{\Gamma\vdash M:\iota\qquad\Gamma,x:\iota\vdash N:A}{\Gamma\vdash\mathsf{let}\ x\ \mathsf{be}\ M\ \mathsf{in}\ N:A}$$

The effect of this construction is that, before replacing x with M in N, M must be evaluated to a value $\underline{n}$. This is particularly important in the case where M is a probabilistic integer since this construction allows to "roll the dice" only once and then provide N with as many copies of the result as needed.

In accordance with this intuition, one can also check that the following reduction inference is derivable from the rules of Figure 1d

$$\frac{M\xrightarrow{p}M'}{\mathsf{let}\ x\ \mathsf{be}\ M\ \mathsf{in}\ N\xrightarrow{p}\mathsf{let}\ x\ \mathsf{be}\ M'\ \mathsf{in}\ N} \tag{8}$$

whereas *it is not true* that

$$\frac{M\xrightarrow{p}M'}{N\left[M/x\right]\xrightarrow{p}N\left[M'/x\right]} \tag{9}$$

Exercise 10. Prove (8) and give a counterexample to (9).

Answer of Exercise 10. *One can notice that* (8) *is an instance of the contextual if-rule in Figure 1d. A counterexample of* (9) *is for* $N = (\mathsf{add})\, xx$ *and* $M = \mathsf{coin}$. *We have* $M \xrightarrow{\frac{1}{2}} \underline{0}$, *but* $N[M/x]$ *does not reduce to* $N[\underline{0}/x] = (\mathsf{add})\,\underline{0}\,\underline{0}$. *The only one-step contractums of* $N[M/x]$ *are* $(\mathsf{add})\,\underline{0}\,\mathsf{coin}$ *and* $(\mathsf{add})\,\underline{1}\,\mathsf{coin}$, *with probability* $\frac{1}{2}$. *From there we get, in several steps, the values* $\underline{0}$ *and* $\underline{2}$, *each with probability* $\frac{1}{4}$, *and* $\underline{1}$ *with probability* $\frac{1}{2}$. *On the contrast,* $N[\underline{0}/x]$ *deterministically evaluates to* $\underline{0}$, *and* $N[\underline{1}/x]$ *deterministically evaluates to* $\underline{2}$.

We have of course

$$\frac{}{\mathsf{let}\ x\ \mathsf{be}\ \underline{n}\ \mathsf{in}\ N \xrightarrow{1} N\,[\theta(n)/x]}$$

where $\theta(0) = \underline{0}$ and $\theta(n+1) = \mathsf{succ}(\underline{n})$ (which reduces to $\underline{n+1}$ in one deterministic step) by definition of this construction.

Random Generators. Using the functions defined in Exercice 6, we can define a closed term unif_2 of type $\iota \Rightarrow \iota$ which, given an integer n, yields a uniform probability distribution on the integers $0, \ldots, 2^n - 1$:

$$\mathsf{unif}_2 = \mathsf{fix}(\lambda f^{\iota \Rightarrow \iota}\, \lambda x^{\iota}\, \mathsf{if}(x, \underline{0}, z \cdot \mathsf{if}(\mathsf{coin}, (f)\, z, z' \cdot (\mathsf{add})\, (\mathsf{exp}_2)\, z\, (f)\, z))) \tag{10}$$

Observe that, when evaluating $(\mathsf{unif}_2)\, M$ (where $\vdash M : \iota$), the term M is evaluated only once thanks to the CBV feature of the conditional construct. Indeed, we do not want the upper bound of the interval on which we produce a probability distribution to change during the computation (the result would be unpredictable!).

Exercise 11. Using the unif_2 and let constructions, define a term unif which, given an integer n, yields a *uniform probability distribution* on the integers $0, \ldots, n$.

Answer of Exercise 11. *Given* $n \in \mathbb{N}$, *the idea is to apply iteratively* unif_2 *until the result is* $\leq n$:

$$\mathsf{unif} = \lambda x^{\iota}\, \mathsf{let}\ y\ \mathsf{be}\ x\ \mathsf{in}\ \ \mathsf{fix}(\lambda f^{\iota}\, \mathsf{let}\ z\ \mathsf{be}\ \ (\mathsf{unif}_2)\, y\ \mathsf{in}\ \mathsf{if}((\mathsf{cmp})\, z\, y, z, w \cdot f))$$

One checks easily that $\vdash \mathsf{unif} : \iota \Rightarrow \iota$. *It is not hard to check that the resulting distribution is uniform (with probability* $\frac{1}{n+1}$ *for each possible result). Notice that this algorithm is almost sure terminating, but not strongly terminating, as the recursive call does not decrease any parameter (see Exercice 9). What about its expected runtime?*

Exercise 12. Define a closed term binom of type $\iota \Rightarrow \iota$ which, given an integer n, yields a (fair) *binomial distribution* out of n trials, i.e. $(\mathsf{binom})\, \underline{n}$ evaluates to $\underline{k}$ with the probability of getting k-times $\underline{1}$ in a sequence of n independent evaluations of coin.

Answer of Exercise 12.

$$\mathsf{binom} = \mathsf{fix}(\lambda f^{\iota \Rightarrow \iota}\, \lambda x^{\iota}\, \mathsf{if}(x, \underline{0}, z \cdot \mathsf{if}(\mathsf{coin}, (f)\, z, w \cdot \mathsf{succ}((f)\, z))))$$

Notice in fact that $(\mathsf{binom})\, \underline{n}$ *will perform exactly* n *recursive calls, each recursive call being preceded by exactly one evaluation of a* coin *redex. So that we can represent the evaluation tree of* $(\mathsf{binom})\, \underline{n}$ *as a complete binary tree of height* n *where each branching is labelled by either* $\underline{0}$ *(if the corresponding evaluation of* coin *returns* $\underline{0}$*) or* $\underline{1}$. *Notice that* $(\mathsf{binom})\, \underline{n}$ *evaluates to* $\underline{k}$ *exactly on the branches where we have had* k *evaluations of* coin *to* $\underline{1}$, *independently from the order of the evaluations. Now, the number of different branches of this tree having exactly* k *evaluations of* coin *to* $\underline{1}$ *(independently from their order) is given by the binomial coeffiecient* $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. *Also, any branch happens with equal probability given by* $\frac{1}{2^n}$, *so that* $(\mathsf{binom})\, \underline{n}$ *evaluates to* $\underline{k}$ *with probability* $\frac{1}{2^n}\binom{n}{k}$, *this describing the binomial law.*

Las Vegas algorithms. A Las Vegas algorithm is a randomized algorithm that always gives the correct result but its running time depends on the draws from the random variables in the algorithm.

Exercise 13. One of the simplest example of a Las Vegas algorithm can be used to find zeros in a finite array: given a function $f : \mathbb{N} \to \mathbb{N}$ and $n \in \mathbb{N}$, find a $k \in \{0, \dots, n\}$ such that $f(k) = 0$. This can be done by iterating random choices of k until we get a value such that $f(k) = 0$. Define a closed term M of type $(\iota \Rightarrow \iota) \Rightarrow \iota \Rightarrow \iota$ that implements this algorithm.

Answer of Exercise 13.

$$M = \lambda f^{\iota \Rightarrow \iota}\, \lambda x^{\iota}\, \mathsf{fix}(\lambda r^{\iota}\, \mathsf{let}\ y\ \mathsf{be}\ \ (\mathsf{unif})\, x\ \mathsf{in}\ \mathsf{if}((f)\, y, y, z \cdot r))$$

with $\vdash M : (\iota \Rightarrow \iota) \Rightarrow \iota \Rightarrow \iota$.

One can notice that our CBV version of the conditional is fundamental in solving Exercise 13. In fact, we strongly believe that this algorithm cannot be written with the usual version of the conditional (as in standard PCF) but we didn't really try to prove this. Do you have some hints in proving (or disproving) this conjecture?

Random-walks. We can define a random-walk over $\mathbb{N}$ as a closed term W of type $\iota \Rightarrow \iota$, meaning that a particle at position $i \in \mathbb{N}$ will evolve in one step to position $j \in \mathbb{N}$ with the probability of $(W)\, \underline{i}$ to evaluate to $\underline{j}$.

Exercise 14. Define a closed term W of type $\iota \Rightarrow \iota$ representing the random-walk of Equation (3).

Answer of Exercise 14. $W = \lambda x^{\iota}\, \mathsf{if}(x, \underline{0}, z \cdot \mathsf{if}(\mathsf{coin}, \mathsf{succ}(\mathsf{succ}(z)), z' \cdot z))$

The following exercise give you an exemple of how natural is the use of higher-order combinators for probabilistic programming. One can in fact defines an iterator of random processes independently from the specific process to iterate.

Exercise 15. Define a closed term iter of type $(\iota \Rightarrow \iota) \Rightarrow \iota \Rightarrow \iota \ \Rightarrow \ \iota$ that takes a term W representing a random-walk, a numeral $\underline{n}$ and returns a term of type $\iota \Rightarrow \iota$ simulating n-iterations of W.

Answer of Exercise 15.

$$\mathsf{iter} = \lambda w^{\iota \Rightarrow \iota}\, \mathsf{fix}(\lambda f^{\iota \Rightarrow \iota \Rightarrow \iota}\, \lambda n^{\iota}\, \lambda x^{\iota}\, \mathsf{if}(n, x, z \cdot (w)\, (f)\, zx))$$

In the above exercices, we just argue intuitively that the solutions actually satisfy the required specification. In fact, proving the soundness formally can be quite burdensome: for example, try to prove that the term $(\mathsf{iter})\, W\, \underline{n}$, with W and iter defined in resp. Exercice 14 and 15, expresses in pPCF the matrix W^n, for W given in (3). The major difficulty is that the operational semantics of pPCF, i.e. the definition of the matrix Red^{∞} is not defined compositionally but with respect to a Makov chain (section 1.3). The next section will present the probabilistic coherence spaces as a denotational model of pPCF. One major feature of a denotational semantics is to be defined compositionally on the structure of a term. The adequacy theorem will then prove the equivalence between the denotational model and the definition of Red^{∞} on ground types, so allowing for compositional proofs of soundness.

2 The standard model of pPCF in $\mathbf{Pcoh}_!$

In order to interpret pPCF in a denotational model, we need:

1. a cartesian closed category, for modelling the simply typed λ-calculus (namely: variables, abstraction and application) and its β-reduction,
2. completely partially ordered hom-sets, for modelling the fix-point operator,
3. convex hom-sets, for sampling from random data,
4. and an object of numerals, in order to express numerals, successor and our zero-test conditional.

We consider the category $\mathbf{Pcoh}_!$, which is the Kleisli category associated with the !-comonad of **Pcoh**. We recall briefly the categorical structure of $\mathbf{Pcoh}_!$ from the linear logic structure of **Pcoh**. The benefit of starting from a linear logic category is to be able to express at a denotational level the linearity of some programming primitives of pPCF, which is a remarkable feature for a denotational semantics of a probabilistic programming language.

2.1 The structure of $\mathbf{Pcoh}_!$ out of that of Pcoh

The category $\mathbf{Pcoh}_!$. An *object* of $\mathbf{Pcoh}_!$ is a PCS $X = (|X|, \mathsf{P}X)$, and the set $\mathbf{Pcoh}_!(X,Y)$ of *morphisms from X to Y* is the set of matrices $f \in \mathbb{R}^{+\mathcal{M}_{\mathrm{fin}}(|X|)\times|Y|}$ such that

$$\forall x \in \mathsf{P}X, \quad \widehat{f}(x) = f \cdot x^{(!)} = \left(\sum_{m \in \mathcal{M}_{\mathrm{fin}}(|X|)} f_{m,b} x^m \right)_{b \in |Y|} \in \mathsf{P}Y \tag{11}$$

where $x^{(!)}$ is the vector in $\mathsf{P}!X$ defined by $x^{(!)}_m = x^m = \prod_{a \in \mathsf{supp}(m)} x_a^{m(a)}$, for $m \in \mathcal{M}_{\mathrm{fin}}(|X|)$.

Notice that the sum in (11) might diverge for arbitrary matrices $f \in \mathbb{R}^{+|X|\times|Y|}$ and vectors $x \in \mathbb{R}^{+|X|}$.

Exercise 16. Recall the PCSs $\mathsf{1} = (\{*\}, [0,1])$ and $\mathsf{Bool} = \mathsf{1} \oplus \mathsf{1} = (\{\mathtt{t}, \mathtt{f}\}, \{(\lambda_\mathtt{t}, \lambda_\mathtt{f}) \in [0,1]^2 \ ; \ \lambda_\mathtt{t} + \lambda_\mathtt{f} \leq 1\})$. Give the following examples of matrices in $\mathbb{R}^{+\mathcal{M}_{\mathrm{fin}}(|\mathsf{Bool}|)\times|\mathsf{1}|}$:

1. a matrix f such that $\widehat{f}$ is a total function from $\mathbb{R}^{+|\mathsf{Bool}|}$ to $\mathbb{R}^{+|\mathsf{1}|}$, but it does not map PBool into $\mathsf{P1}$, so $f \notin \mathbf{Pcoh}_!(\mathsf{Bool}, \mathsf{1})$;
2. a matrix g such that $\widehat{g}$ is a total function from PBool to $\mathsf{P1}$, so $g \in \mathbf{Pcoh}_!(\mathsf{Bool}, \mathsf{1})$, but $\widehat{g}$ diverges on some vectors of $\mathbb{R}^{+|\mathsf{Bool}|}$ outside PBool. (*Hint: recall the example of analytic function on the booleans given in Ehrhard's notes*).

Answer of Exercise 16.

1. *Take for example the function* $f_{m,*} = \begin{cases} 2 & \text{if } m = [], \\ 0 & \text{otherwise.} \end{cases}$. *We have* $\widehat{f}(x) = 2$, *so* $\widehat{f}$ *is well-defined on the whole* $\mathbb{R}^{+2}$, *however the codomain of* $\widehat{f}$ *is ouside* $\mathsf{P1} = [0,1]$.
2. *Take for example the function*

$$g_{[\mathtt{t}^n, \mathtt{f}^k], *} = \begin{cases} 2^n & \text{if } n = k \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

We have that $\widehat{g}(x) = \sum_{n=1}^{\infty} 2^n x_\mathtt{t}^n x_\mathtt{f}^n$. *If* $x \in \mathsf{PBool}$, *so* $x_\mathtt{t} + y_\mathtt{f} \leq 1$, *the maximal value of this function is reached when* $x_\mathtt{f} = 1 - x_\mathtt{t}$, *so that we can consider the function* $\lambda \mapsto \sum_{n=1}^{\infty} 2^n \lambda^n (1 -$

$\lambda)^n$, *with* $\lambda \in [0,1]$. *The quantity* $\lambda^n(1-\lambda)^n$ *is maximal for* $\lambda = \frac{1}{2}$, *so that* $\sum_{n=1}^{\infty} 2^n\lambda^n(1-\lambda)^n \leq \sum_{n=1}^{\infty} \frac{1}{2^n} \leq 1$ *and we have* $g \in \mathbf{Pcoh}_!(\mathsf{Bool}, 1)$. *On the contrast, if we take* $x = (1,1)$, *then of course* $\widehat{g}(x)$ *diverges.*

Exercise 17. Prove that $\mathbf{Pcoh}_!(X,Y) = \mathbf{Pcoh}(!X, Y)$. What is the difference between (11) and the condition necessary for inferring $f \in \mathbf{Pcoh}(!X, Y)$?

Answer of Exercise 17. $f \in \mathbf{Pcoh}(!X, Y)$ *means:*

$$\forall z \in \mathsf{P}(!X), f \cdot z \in \mathsf{P}Y$$

The above equation trivially implies (11), *as* $x^{(!)} \in \mathsf{P}(!X)$. *Let us prove the converse.*

Take $u \in \mathsf{P}(!X)$, $y \in \mathsf{P}Y$, *we have to prove that:* $\langle f \cdot u, y\rangle \leq 1$. *Notice that we have:*

$$\langle f \cdot u, y\rangle = \langle f, u \otimes y\rangle = \langle f^{\perp} \cdot y, u\rangle$$

By hypothesis we have moreover that $f^{\perp} \cdot y \in \{x^{(!)}\,;\, x \in \mathsf{P}X\}^{\perp} = (\mathsf{P}!X)^{\perp}$, *we conclude that* $\langle f^{\perp} \cdot y, u\rangle \leq 1$, *as* $u \in \mathsf{P}(!X)$.

The *identity on* X is given by the dereliction matrix $\mathsf{der}_X \in \mathbf{Pcoh}(!X, X)$:

$$\mathsf{Id}^{\mathsf{KI}}{}_{X\,m,a} = \mathsf{der}_{X\,m,a} = \begin{cases} 1 & \text{if } m = [a], \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

In fact, we have $\widehat{\mathsf{Id}^{\mathsf{KI}}{}_X}(x) = \mathsf{der}_X \cdot x^{(!)} = x$, for every $x \in \mathsf{P}X$.

The *composition* of a morphism $f \in \mathbf{Pcoh}_!(X,Y)$ and a morphism $g \in \mathbf{Pcoh}_!(Y,Z)$ is obtained via the matrix composition, the digging and the functorial promotion of **Pcoh**:

$$g \circ f = g(!f)\,\mathsf{dig}_X \tag{13}$$

where we recall that $\mathsf{dig}_X \in \mathbf{Pcoh}(!X, !!X)$ and $!f \in \mathbf{Pcoh}(!!X, !Y)$ are:

$$\mathsf{dig}_{X\,m,M} = \begin{cases} 1 & \text{if } m = \sum M, \\ 0 & \text{otherwise.} \end{cases} \qquad !f_{M,p} = \sum_{r \in L(M,p)} \frac{p!}{r!} f^r \tag{14}$$

with $f^r = \prod_{(m,b)\in\mathsf{supp}(r)} f_{m,b}^{r(m,b)}$ and $p! = \prod_{a\in|X|} p(a)!$ is the multiset factorial.

Exercise 18. Given $f \in \mathbf{Pcoh}_!(X,Y)$, $g \in \mathbf{Pcoh}_!(Y,Z)$ and $x \in \mathsf{P}X$, prove that $\widehat{g \circ f}(x) = \widehat{g}(\widehat{f}(x))$. (*Hint: use the categorical properties of* dig *and* !). Conclude that $g \circ f \in \mathbf{Pcoh}_!(X,Z)$.

Answer of Exercise 18.

$$\begin{aligned}
\widehat{g \circ f}(x) &= (g(!f)\,\mathsf{dig}_X) \cdot x^{(!)} && \textit{by definition} \\
&= (g(!f)) \cdot (\mathsf{dig}_X \cdot x^{(!)}) = (g(!f)) \cdot x^{(!)(!)} && \textit{by def. of } \mathsf{dig} \\
&= g \cdot ((!f) \cdot x^{(!)(!)}) = g \cdot (f \cdot x^{(!)})^{(!)} && \textit{by funct. of } ! \\
&= \widehat{g}(\widehat{f}(x)) && \textit{by def. of } \widehat{\ }
\end{aligned}$$

We can conclude that $g \circ f \in \mathbf{Pcoh}_!(X,Z)$, *since by hypothesis* $\widehat{f}(x) \in \mathsf{P}Y$ *and hence* $\widehat{g}(\widehat{f}(x)) \in \mathsf{P}Z$, *so condition* (11) *holds.*

Recall from Ehrhard's notes that a crucial feature of $\mathbf{Pcoh}_!$ is to be well-pointed, meaning that a matrix $f \in \mathbf{Pcoh}_!(X,Y)$ is univocally characterised by its behaviour as the map $\widehat{f}$:

Proposition 4 (Functional characterization) *Given two matrices $f, f' \in \mathbf{Pcoh}_!(X,Y)$, one has $f = f'$ (as matrices) iff $\widehat{f} = \widehat{f'}$ (as maps $\mathsf{P}X \to \mathsf{P}Y$).*

This property is extremely convenient, as one can define a morphism of $\mathbf{Pcoh}_!$ extensionally, without the need of giving the coefficients of the matrix associated with the morphisms. In fact, we will use this property in Figure **??**, when giving a functional definition of the denotation of the pPCF terms.

Cartesian closeness. The cartesian product of $\mathbf{Pcoh}_!$ is the same as that of $\mathbf{Pcoh}$, with the projections composed with der, that is, given a countable collection of PCSs $(X_i)_{i\in I}$, we have:

$$|\&_{i\in I} X_i| = \bigcup_{i\in I} \{i\} \times |X_i|$$

$$\mathsf{P}\&_{i\in I} X_i = \{x \in \mathbb{R}_{\geq 0}^{|\&_{i\in I} X_i|} \; ; \; \forall i \in I, (x_{(i,a)})_{a\in|X_i|} \in \mathsf{P}X_i\}$$

$$\pi_j^{\mathsf{KI}} = \pi_j \, \mathsf{der}_{\&_{i\in I} X_i} \in \mathbf{Pcoh}_!\big(\&_{i\in I} X_i, X_j\big) \quad \text{i.e. } (\pi_j^{\mathsf{KI}})_{m,a} = \begin{cases} 1 & \text{if } m = [(i,a)] \text{ and } j = i \\ 0 & \text{otherwise} \end{cases}$$

Exercise 19. Prove the universal property of the cartesian product in $\mathbf{Pcoh}_!$, i.e. given a collection $f_i \in \mathbf{Pcoh}_!(Y, X_i)$ for $i \in I$, the morphism $\langle f_i \rangle_{i\in I} \in \mathbf{Pcoh}_!(Y, \&_{i\in I} X_i)$ is the only one satisfying $(\pi_j^{\mathsf{KI}} \circ \langle f_i \rangle_{i\in I}) = f_j$ for every $j \in I$.

Answer of Exercise 19.

$$\begin{aligned} (\pi_j^{\mathsf{KI}} \circ \langle f_i \rangle_{i\in I}) &= \pi_j \, \mathsf{der}(!\langle f_i \rangle) \, \mathsf{dig} \\ &= \pi_j \langle f_i \rangle_{i\in I} \\ &= f_j \end{aligned}$$

The unicity follows from the unicity of $\langle f_i \rangle$ for π_j and the universal property of der *and* dig.

In the following we will use the infix notation $X \,\&\, Y$ and $\langle f, g \rangle$ for binary cartesian products. Also, we will denote by T the zero-ary product, which is the PCS of empty web.

A crucial ingredient necessary to lift the closeness structure of $\mathbf{Pcoh}$ to $\mathbf{Pcoh}_!$ is the strong monoidal isomorphisms $\mathsf{mat}(\mathsf{m}^0) \in \mathbf{Pcoh}(1, !\mathsf{T})$ and $\mathsf{mat}(\mathsf{m}^2_{|\mathsf{X}_1|,|\mathsf{X}_2|}) \in \mathbf{Pcoh}(!X_1 \otimes !X_2, !(X_1 \,\&\, X_2))$, transforming the tensor product of promoted spaces into the promotion of a cartesian product:

$$\mathsf{mat}(\mathsf{m}^0)_{*,[]} = 1 \qquad \mathsf{mat}(\mathsf{m}^2)_{(m_1,m_2),q} = \begin{cases} 1 & \text{if } q(i,a) = m_i(a) \\ & \text{for } i \in \{1,2\},\ a \in |X_i|, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

The *object of morphisms* is defined by Girard's decomposition:

$$X \Rightarrow Y = !X \multimap Y = \mathbf{Pcoh}(!X, Y) = \mathbf{Pcoh}_!(X, Y) \tag{16}$$

The *evaluation* morphism $\mathsf{ev}^{\mathsf{KI}} \in \mathbf{Pcoh}_!((X \Rightarrow Y) \,\&\, X, Y)$ and the *curryfication* $\mathsf{Cur}^{\mathsf{KI}}(f) \in \mathbf{Pcoh}_!(Z, X \Rightarrow Y)$, for every $f \in \mathbf{Pcoh}_!(Z \,\&\, X, Y)$ are then obtained by their corresponding constructions in $\mathbf{Pcoh}$ as follows:

$$\mathsf{ev}^{\mathsf{KI}} = \mathsf{ev}(\mathsf{der}_{X\Rightarrow Y} \otimes \mathsf{Id}_{!X})\mathsf{mat}(\mathsf{m}^2_{|\mathsf{X}\Rightarrow\mathsf{Y}|,|\mathsf{X}|})^{-1} \quad \text{i.e. } \mathsf{ev}^{\mathsf{KI}}{}_{(m,p),b} = \begin{cases} 1 & \text{if } m = [(p,b)], \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

$$\mathsf{Cur}^{\mathsf{KI}}(f) = \mathsf{Cur}(f\mathsf{mat}(\mathsf{m}^2_{|\mathsf{Z}|,|\mathsf{X}|})^{-1}) \qquad \text{i.e. } \mathsf{Cur}^{\mathsf{KI}}(f)_{m,(p,b)} = f_{(m,p),b} \tag{18}$$

Notice that in the above two equations we deliberately use the relational strong monoidal isomorphisms in order to represent with a pair (m, p) of two multisets a multiset over the disjoint union of the supports of m and p.

Exercise 20. By using the properties of the morphisms of **Pcoh**, prove that:

1. $\widehat{\mathsf{ev}^{\mathsf{Kl}}}(\langle f, x\rangle) = \widehat{f}(x)$
2. $(\widehat{\widehat{\mathsf{Cur}^{\mathsf{Kl}}}(f)}(x))(z) = \widehat{f}(\langle x, z\rangle)$

Answer of Exercise 20.

$$\begin{aligned}
\widehat{\mathsf{ev}^{\mathsf{Kl}}}(\langle f, x\rangle) &= (\mathsf{ev}(\mathsf{der} \otimes !X)\mathsf{mat}(\mathsf{m}^2)^{-1}) \cdot (\langle f, x\rangle)^{(!)} \\
&= (\mathsf{ev}(\mathsf{der} \otimes !X)) \cdot (\mathsf{mat}(\mathsf{m}^2)^{-1} \cdot (\langle f, x\rangle)^{(!)}) \\
&= (\mathsf{ev}(\mathsf{der} \otimes !X)) \cdot (f^{(!)} \otimes x^{(!)}) \\
&= \mathsf{ev} \cdot ((\mathsf{der} \otimes !X) \cdot (f^{(!)} \otimes x^{(!)})) \\
&= \mathsf{ev} \cdot (f \otimes x^{(!)}) \\
&= f \cdot x^{(!)} \\
&= \widehat{f}(x)
\end{aligned}$$

$$\begin{aligned}
(\widehat{\widehat{\mathsf{Cur}^{\mathsf{Kl}}}(f)}(x))(z) &= ((\mathsf{Cur}(f\mathsf{mat}(\mathsf{m}^2_{|\mathsf{Z}|,|\mathsf{X}|})^{-1})) \cdot x^{(!)}) \cdot z^{(!)} \\
&= (f\mathsf{mat}(\mathsf{m}^2_{|\mathsf{Z}|,|\mathsf{X}|})^{-1}) \cdot (x^{(!)} \otimes z^{(!)}) \\
&= f \cdot \langle x, z\rangle^{(!)} \\
&= \widehat{f}(\langle x, z\rangle)
\end{aligned}$$

Cpo-enriched hom-sets. A categorical model of a typed programming language associates the types with objects of the category and the programs with morphisms from the input type interpretation to the output type interpretation. Some programming primitives may need some structure on the hom-sets, for example the fix-point operator (giving recursion) needs the hom-set to be cpo-enriched.

There is actually an equivalence between the sets $\mathsf{P}X$ associated with PCSs X and the sets of the morphisms of **Pcoh** and $\mathbf{Pcoh}_!$. Namely, given a PCS X, $\mathsf{P}X$ is equivalent to the set $\mathbf{Pcoh}(1, X)$ as well as $\mathbf{Pcoh}_!(\mathsf{T}, X)$. Viceversa, the sets $\mathbf{Pcoh}(X, Y)$ and $\mathbf{Pcoh}_!(X, Y)$ are equivalent respectively to the sets $\mathsf{P}(X \multimap Y)$ and $\mathsf{P}(X \Rightarrow Y)$. Henceforth, studying the structure of $\mathsf{P}X$ for generic X corresponds to study the structure of the hom-sets of the categories **Pcoh** and $\mathbf{Pcoh}_!$, which is what we will do in this subsection.

Given a PCS X, recall that $\mathsf{P}X$ is endowed with the partial order defined component-wise:

$$x \leq x' \text{ iff } \forall a \in |X|, x_a \leq x'_a \tag{19}$$

Recall that the vectors in $\mathsf{P}X$ are bounded in a fixed direction, i.e. $\forall a \in |X|, \exists \lambda \in \mathbb{R}_{\geq 0}, \forall x \in \mathsf{P}X, x_a \leq \lambda$. Therefore, giving an increasing ω-*chain*, i.e. a countable increasing family of vectors in $\mathsf{P}X$, its limit can be defined as the component-wise supremum:

$$\text{given } (x_i)_{i\in\mathbb{N}} \in \mathsf{P}X \text{ s.t. } x_i \leq x_{i+1}, \text{ we define } \sup_i(x_i) = \big(\sup_i(x_{ia})\big)_{a\in|X|} \tag{20}$$

The following proposition states that $\mathbf{Pcoh}_!$ behaves well with such a notion of limit.

Proposition 5 (Scott continuity) *Let X, Y be PCSs, $(x_i)_{i\in\mathbb{N}} \in \mathsf{P}X$ be an increasing ω-chain,*

1. $\sup_i(x_i) \in \mathsf{P}X$,
2. *for every $f \in \mathbf{Pcoh}_!(X,Y)$, $(\widehat{f}(x_i))_{i\in\mathbb{N}}$ is increasing and $\widehat{f}(\sup_i(x_i)) = \sup_i\big(\widehat{f}(x_i)\big)$.*

Exercise 21. Prove Proposition 5.

Answer of Exercise 21. *For 1, given $y \in \mathsf{P}X^\perp$, we have: $\langle \sup_i(x_i), y\rangle = \sup_i\langle x_i, y\rangle \leq 1$.*

For 2, notice that $x_i \leq x_{i+1}$ implies $x_i^{(!)} \leq x_{i+1}^{(!)}$ and hence $f \cdot x_i^{(!)} \leq f \cdot x_{i+1}^{(!)}$ as addition and multiplication (with positive reals) are monotone increasing. We conclude that $\widehat{f}$ also is monotone increasing and so $(\widehat{f}(x_i))_{i\in\mathbb{N}}$ is an increasing ω-chain. Similarly, $\widehat{f}(\sup_i(x_i)) = \sup_i\big(\widehat{f}(x_i)\big)$ is an immediate consequence of Equation 11 and the fact that addition and multiplication (with positive reals) commutes with suprema.

An immediate consequence of the component-wise definition in (20) is that, given two ω-chains $(x_i)_{i\in\mathbb{N}} \in \mathsf{P}X$ and $(y_j)_{j\in\mathbb{N}} \in \mathsf{P}Y$, we have:

$$\langle \sup_{i\in\mathbb{N}} x_i, \sup_{j\in\mathbb{N}} y_j\rangle = \sup_{i\in\mathbb{N}}\sup_{j\in\mathbb{N}}\langle x_i, y_j\rangle = \sup_{j\in\mathbb{N}}\sup_{i\in\mathbb{N}}\langle x_i, y_j\rangle = \sup_{i\in\mathbb{N}}\langle x_i, y_i\rangle \in \mathsf{P}(X \,\&\, Y) \tag{21}$$

Exercise 22. Given increasing $(f_i)_{i\in\mathbb{N}} \in \mathsf{P}(X \Rightarrow Y)$ and $(x_i)_{i\in\mathbb{N}} \in \mathsf{P}X$, prove that:

$$\widehat{(\sup_i f_i)}(\sup_i(x_i)) = \sup_i(\widehat{f_i}(x_i)).$$

Answer of Exercice 22. *By Exercice 20, Equation* (21) *and Proposition 5:*

$$\widehat{(\sup_i f_i)}(\sup_i(x_i)) = \widehat{\mathsf{ev}^{\mathsf{Kl}}}(\sup_i\langle f_i, x_i\rangle) = \sup_i \widehat{\mathsf{ev}^{\mathsf{Kl}}}(\langle f_i, x_i\rangle) = \sup_i(\widehat{f_i}(x_i))$$

The two properties of Proposition 5 justifies the standard definition of the least fix-point operator for $\mathbf{Pcoh}_!$. Given a PCS X, we set $\mathsf{Y}_n \in \mathbf{Pcoh}_!(X \Rightarrow X, X)$ for any $n \in \mathbb{N}$ and its limit $\mathsf{Y} \in \mathbf{Pcoh}(X \Rightarrow X, X)$ as:

$$\mathsf{Y}_0 = 0, \qquad \mathsf{Y}_{n+1} = \mathsf{ev}^{\mathsf{Kl}} \circ \langle \mathsf{Id}, \mathsf{Y}_n\rangle, \qquad \mathsf{Y} = \sup_n \mathsf{Y}_n.$$

Exercise 23.

1. Prove that $(\mathsf{Y}_n)_n$ is a increasing chain in $\mathsf{P}((X \Rightarrow X) \Rightarrow X)$, so that $\mathsf{Y} = \sup_n \mathsf{Y}_n$ is well-defined.
2. Prove that, for any $n \in \mathbb{N}$, any $f \in \mathsf{P}(X \Rightarrow X)$, $\widehat{\mathsf{Y}}_{n+1}(f) = \widehat{f}(\widehat{\mathsf{Y}}_n(f))$. Conclude the fix-point equation: $\widehat{\mathsf{Y}}(f) = \widehat{f}(\widehat{\mathsf{Y}}(f))$.

Answer of Exercise 23.

1. *Remark that $\circ$ and pairing are monotone increasing. Therefore, by induction on n, we have $\mathsf{Y}_n \leq \mathsf{Y}_{n+1}$. The base of induction is trivial, since 0 is the minimum.*
2. *By definition*

$$\begin{aligned}
\widehat{\mathsf{Y}}_{n+1}(f) &= \widehat{(\mathsf{ev}^{\mathsf{Kl}} \circ \langle \mathsf{Id}, \mathsf{Y}_n\rangle)}(f) && \text{by definition}\\
&= \big(\widehat{\mathsf{ev}^{\mathsf{Kl}}}(\widehat{\langle \mathsf{Id}, \mathsf{Y}_n\rangle}(f))\big) && \text{by Ex. 18}\\
&= \big(\widehat{\mathsf{ev}^{\mathsf{Kl}}}(\langle f, \widehat{\mathsf{Y}_n}(f)\rangle)\big) && \text{by def. pairing}\\
&= \widehat{f}(\widehat{\mathsf{Y}}_n(f)) && \text{by Ex 20}
\end{aligned}$$

The fix-point equation is a trivial consequence of the above equality and Proposition 5.

This means that the standard least fix-point operator Y can be described as a power series, which is not completely obvious at first sight.

Convex hom-sets. Random data will be denoted by barycentric sums: for example, if $x, x' \in \mathsf{P}X$ will be the denotation of two values of some type X, and $\lambda \in [0,1]$, then $\lambda x + (1-\lambda)x'$ will represent a random program evaluating with probability λ to x, and with probability $(1-\lambda)$ to x'. The following proposition states then the PCSs are closed under barycentric sums:

Proposition 6 (Convexity) *Let X be a PCS, $\forall (x_i)_{i\in I} \in \mathsf{P}X$, $\forall (\lambda_i)_{i\in I} \in [0,1]$ s.t. $\sum_{i\in I} \lambda_i = 1$, we have: $\sum_{i\in I} \lambda_i x_i \in \mathsf{P}X$.*

Exercise 24. Prove Proposition 6.

Answer of Exercise 24. *Given $y \in \mathsf{P}X^{\perp}$, we have: $\langle \sum_i \lambda_i x_i, y\rangle = \sum_i \lambda_i \langle x_i, y\rangle \leq 1$.*

The object of numerals. The object of numerals is an object N associated with the ground type ι of natural numbers and having enough structure to express the basic operations of pPCF over ι: constants, successor and conditionals based on a zero testing.

In $\mathbf{Pcoh}_!$, one can define this object from standard constructions in the linear logic category $\mathbf{Pcoh}$. Namely, we let N to be the countable coproduct of the tensor unit:

$$\mathsf{N} = \bigoplus_{i\in\mathbb{N}} \mathsf{1}, \qquad \text{i.e. } \mathsf{N} = \left(\mathbb{N}, \left\{v \in [0,1]^{\mathbb{N}} \,;\, \sum_{i\in\mathbb{N}} v_i \leq 1\right\}\right) \tag{22}$$

First of all, notice that a numeral can be associated with a constant function $\overline{n}_X \in \mathbf{Pcoh}_!(X, \mathsf{N})$ by the weakening $\mathsf{w}_X \in \mathbf{Pcoh}(!X, 1)$ and the injections $\overline{\pi_n} \in \mathbf{Pcoh}(1, \mathsf{N})$:

$$\overline{n}_X = \overline{\pi_n}\,\mathsf{w}_X \qquad \text{i.e. } \overline{n_X}_{m,k} = \begin{cases} 1 & \text{if } m = [] \text{ and } k = n, \\ 0 & \text{otherwise,} \end{cases} \tag{23}$$

Another major benefit of this definition is to lift the structure of !-coalgebra of the tensor unit $\mathsf{1}$ to N, by the morphism $h_{\mathsf{N}} : \mathbf{Pcoh}(\mathsf{N}, !\mathsf{N})$:

$$(h_{\mathsf{N}})_{n,m} = \begin{cases} 1 & \text{if } m = k[n] \text{ for some } k \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

The following exercise shows that h_{N} allows to duplicate and erase "true" natural numbers e_n but not general elements of PN which can be considered as "computations" and not as "values".

Exercise 25. Prove that for any $n \in \mathbb{N}$, $h_{\mathsf{N}} \cdot e_n = e_n^{(!)}$. Moreover, observe that it is not true however that $\forall u \in \mathsf{PN}\ h_{\mathsf{N}} \cdot u = u^{(!)}$, in fact what we have is: $h_{\mathsf{N}} \cdot u = \sum_{n\in\mathbb{N}} u_n {e_n}^{(!)}$

Answer of Exercise 25. *In fact, if $m = k[n]$ for some k, then $(h_{\mathsf{N}} \cdot e_n)_m = 1 = (e_n)^m = (e_n^{(!)})_m$. Otherwise, if $m = [n'] + m'$ for some $n' \neq n$, then $(h_{\mathsf{N}} \cdot e_n)_m = 0 = (e_n)_{n'}(e_n)^{m'} = e_n^{(!)}{}_{[n']+m'}$.*

As for the second statement, consider $u = \frac{1}{2}e_0 + \frac{1}{2}e_1$. We have that $(h_{\mathsf{N}} \cdot u)_{[0,1]} = 0$ while $u^{(!)}_{[0,1]} = \frac{1}{4}$. In general, we have that $(h_{\mathsf{N}} \cdot u) = h_{\mathsf{N}} \cdot (\sum_{n\in\mathbb{N}} u_n e_n) = \sum_{n\in\mathbb{N}} u_n (h_{\mathsf{N}} \cdot e_n) = \sum_{n\in\mathbb{N}} u_n {e_n}^{(!)}$.

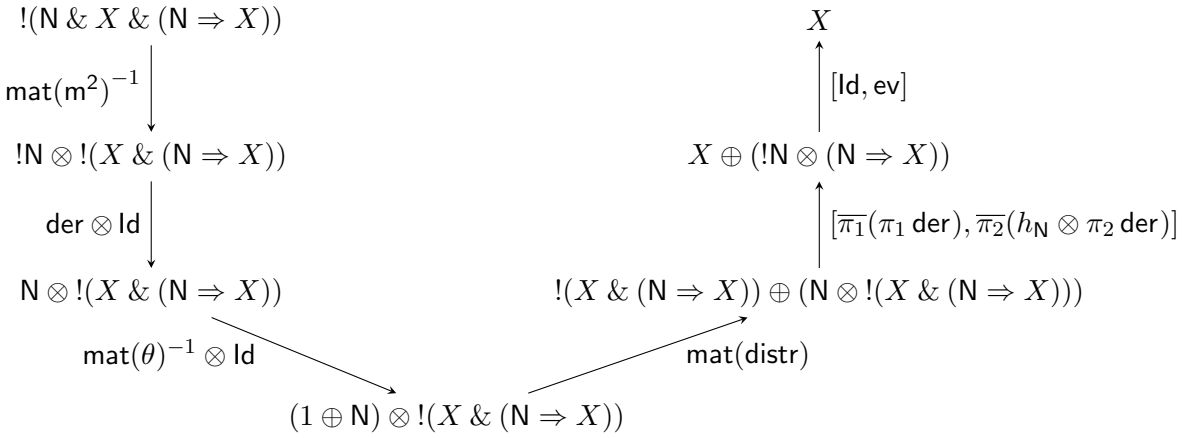
Finally, N enjoys the strong isos $\mathsf{mat}(\theta) \in \mathbf{Pcoh}(1 \oplus \mathsf{N}, \mathsf{N})$ given by the relation θ:

$$\begin{array}{rccc} \theta: & |1 \oplus \mathsf{N}| & \to & |\mathsf{N}| \\ & (1, *) & \mapsto & 0 \\ & (2, n) & \mapsto & n+1 \end{array}$$

The successor morphism $\overline{\mathsf{suc}} \in \mathbf{Pcoh}_!(\mathsf{N}, \mathsf{N})$ is then the composition of dereliction, the right injection and the above isomorphism:

$$\overline{\mathsf{suc}} = \mathsf{mat}(\theta)\overline{\pi_2}\,\mathsf{der} \quad \text{i.e. } \overline{\mathsf{suc}}_{m,n} = \begin{cases} 1 & \text{if } n > 0 \text{ and } m = [n-1], \text{ or } n = 0 \text{ and } m = [0], \\ 0 & \text{otherwise.} \end{cases}$$

Our conditional, which gathers a zero-test and a predecessor operation, is based on the inverse of $\mathsf{mat}(\theta)$ and the !-coalgebra morphism h_{N}. We define $\overline{\mathsf{If}} \in \mathbf{Pcoh}_!(\mathsf{N} \,\&\, X \,\&\, (\mathsf{N} \Rightarrow X))$ by:



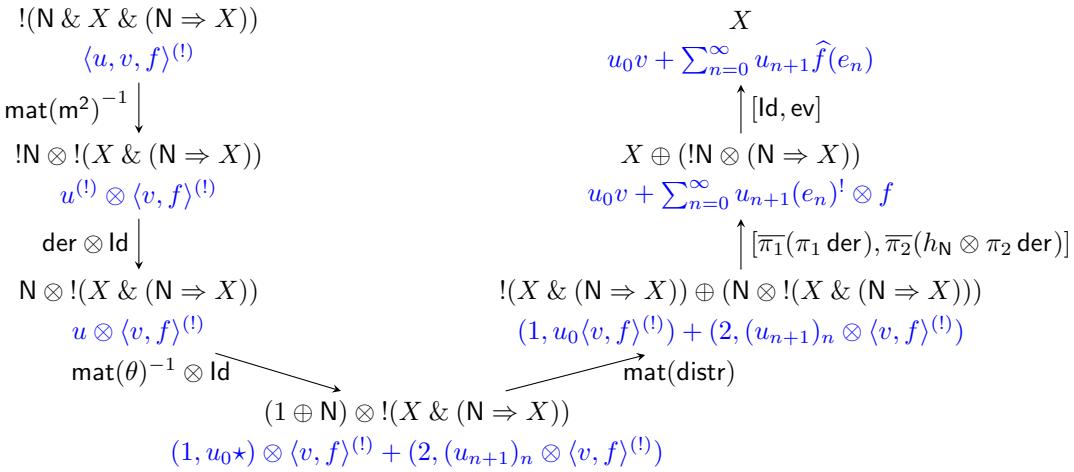
where we omit to explicit the associativity and neutrality isos of $\otimes$, $\mathsf{mat}(\mathsf{distr})_{X_1,X_2,Z} \in \mathbf{Pcoh}((X_1 \oplus X_2) \otimes Z, (X_1 \otimes Z) \oplus (X_2 \otimes Z))$ is the strong isos of the distributive lax of $\otimes$ over $\oplus$ given by the following relation:

$$\begin{array}{rccc} \mathsf{distr}: & |(X_1 \oplus X_2) \otimes Z| & \to & |(X_1 \otimes Z) \oplus (X_2 \otimes Z)| \\ & ((i, a), b) & \mapsto & (i, (a, b)) \end{array}$$

with also $\overline{\pi_i} \in \mathbf{Pcoh}(X_i, X_1 \oplus X_2)$ being the injection of the coproduct $X_1 \oplus X_2$, for $i \in \{1, 2\}$, and $[f_1, f_2] \in \mathbf{Pcoh}(X_1 \oplus X_2, Z)$ being the copairing of $f_i \in \mathbf{Pcoh}(X_i, Z)$.

Exercise 26. Given $u \in \mathsf{P}\mathsf{N}$, $v \in \mathsf{P}X$ and $f \in \mathsf{P}(\mathsf{N} \Rightarrow X)$, prove that $\widehat{\overline{\mathsf{If}}}(u, v, f) = u_0 v + \sum_{n=0}^{\infty} u_{n+1} \widehat{f}(e_n)$.

Answer of Exercise 26. *We sketch the proof by travelling through the diagram defining $\overline{\mathsf{If}}$, every single step being an easy consequence of the definitions.*


$!(\mathsf{N} \,\&\, X \,\&\, (\mathsf{N} \Rightarrow X))$
$\langle u, v, f\rangle^{(!)}$
$\mathsf{mat}(\mathsf{m}^2)^{-1}$
$!\mathsf{N} \otimes !(X \,\&\, (\mathsf{N} \Rightarrow X))$
$u^{(!)} \otimes \langle v, f\rangle^{(!)}$
$\mathsf{der} \otimes \mathsf{Id}$
$\mathsf{N} \otimes !(X \,\&\, (\mathsf{N} \Rightarrow X))$
$u \otimes \langle v, f\rangle^{(!)}$
$\mathsf{mat}(\theta)^{-1} \otimes \mathsf{Id}$
$(1 \oplus \mathsf{N}) \otimes !(X \,\&\, (\mathsf{N} \Rightarrow X))$
$(1, u_0\star) \otimes \langle v, f\rangle^{(!)} + (2, (u_{n+1})_n \otimes \langle v, f\rangle^{(!)})$
$\mathsf{mat}(\mathsf{distr})$
$!(X \,\&\, (\mathsf{N} \Rightarrow X)) \oplus (\mathsf{N} \otimes !(X \,\&\, (\mathsf{N} \Rightarrow X)))$
$(1, u_0\langle v, f\rangle^{(!)}) + (2, (u_{n+1})_n \otimes \langle v, f\rangle^{(!)})$
$[\mathsf{Id}, \mathsf{ev}]$
$X \oplus (!\mathsf{N} \otimes (\mathsf{N} \Rightarrow X))$
$u_0 v + \sum_{n=0}^{\infty} u_{n+1}(e_n)^! \otimes f$
$[\overline{\pi_1}(\pi_1\, \mathsf{der}), \overline{\pi_2}(h_{\mathsf{N}} \otimes \pi_2\, \mathsf{der})]$
X
$u_0 v + \sum_{n=0}^{\infty} u_{n+1} \widehat{f}(e_n)$

References

[1] Martin Avanzini, Ugo Dal Lago, and Akihisa Yamada. On probabilistic term rewriting. *Science of Computer Programming*, 185:102338, 2020.