Full Abstraction for Non-Deterministic and Probabilistic Extensions of PCF I: the Angelic Cases

Jean Goubault-Larrecq[☆]

ENS Cachan, 61 avenue du président Wilson, 94230 Cachan, France

Abstract

We examine several extensions and variants of Plotkin's language PCF, including non-deterministic and probabilistic choice constructs. For each, we give an operational and a denotational semantics, and compare them. In each case, we show soundness and computational adequacy: the two semantics compute the same values at ground types. Beyond this, we establish full abstraction (the observational preorder coincides with the denotational preorder) in a number of cases. In the probabilistic cases, this requires the addition of so-called statistical termination testers to the language.

Keywords: domain theory; PCF; full abstraction

1. Introduction

There are at least two common ways of giving semantics to a programming language. *Denotational* semantics defines domains of values, typically directed-complete partial orders (dcpos), and maps each program to its value, directly. *Operational* semantics define the semantics of a program by a fixed collection of computation rules. The latter can be thought of as abstractions of elementary computation steps that a machine would implement to run a given program.

Now, given two semantics for a programming language, one denotational, one operational, how do these two relate to each other? To our knowledge, this question was first asked, and answered, in G. Plotkin's seminal paper [38]. The language studied here, which came later to be known as PCF, is a prototypical higher-order, typed, side-effect free programming language.

[☆]LSV, ENS Cachan, CNRS, INRIA.

Plotkin showed that the relation between the two semantics covered several distinct questions. The simplest question is *soundness*, which holds whenever any program M that evaluates, operationally, to some normal form V, has the same denotational value as V. A harder question is *computational adequacy*, which is a form of converse: if a program M has the same value as some normal form V, then the operational semantics will run M until it terminates, with V as normal form. In PCF, this holds for all programs at ground types, such as the type of natural numbers, but the question would be hopeless at higher types. Instead, *full abstraction* holds if and only if equality of denotations of two programs M, N is equivalent to *observational equivalence*, i.e., to the fact, intuitively, that M and N behave operationally in the same ways under any evaluation context E. Plotkin then showed that PCF was *not* fully abstract, but that PCF plus the so-called parallel or construct was [38].

Our purpose is to extend the study of these questions to higher-order, typed programming languages that implement *choice*: either non-deterministic (angelic, demonic, erratic) or probabilistic, or both. Our extensions of PCF with choice will be side-effect free, just like PCF. The questions of soundness, computational adequacy, and full abstraction, in the presence of side-effects (assignments and global stores, input/output for example) fall outside the scope of this paper, and already involve dealing with subtle issues such as snapback (see [35, 36] or [39, 33] for example).

On the other hand, we shall concentrate on non-deterministic and probabilistic choice. As A. Jung noticed, our present work can be seen as a test case for our notions of *previsions*, which we proposed as a denotational model for both kinds of choice [11].

Outline. We give an overview of related work in Section 2, and necessary domain-theoretic preliminaries in Section 3. We recap our notions of previsions in Section 4, then proceed to define our languages PCF_S in Section 5, which are variants of PCF with various forms of non-deterministic choice and probabilistic choice. While we shall eventually only deal with full abstraction in the angelic cases $S = \mathtt{A}$ and $S = \mathtt{AP}$, it takes no more effort to give semantics, prove soundness (Section 6) and computational adequacy (Section 7) for every possible value of S. In Section 8, we show that no language PCF_S that has probabilistic features (in particular $\mathrm{PCF}_{\mathtt{AP}}$, but not $\mathrm{PCF}_{\mathtt{A}}$) can be fully abstract. This prompts us to introduce so-called statistical termination testers in Section 9 . With the help of a few additional topological facts, re-

lating the Scott and weak topologies on spaces of previsions (Section 10), this allows us to prove our final full abstraction results in Section 11: $\mathrm{PCF}_{\mathtt{AP}}$ plus statistical termination testers is fully abstract (Theorem 11.7), and $\mathrm{PCF}_{\mathtt{A}}$, even without termination testers, is fully abstract as well (Theorem 11.9).

2. Related Work

Although the question of full abstraction for PCF with choice is natural, it does not seem there has been any study of the question since Plotkin's seminal paper [38], at least when the denotational side of the matter consists of *domain semantics*, i.e., semantics based on dcpos, in the style of D. S. Scott.

One of Plotkin's result is that PCF alone is not fully abstract with respect with its domain semantics. The key to the problem lies in the undefinability (through PCF terms) of the so-called parallel or map from $\mathbb{N}_\perp \times \mathbb{N}_\perp$ to $\mathbb{N}_\perp$, defined as mapping $(0,0)$ to 0, the pairs $(0,1)$, $(1,0)$ and $(1,1)$ to 1, and all other pairs to the "undefined" element $\perp$. But PCF with an extra `por` primitive that implements this function *is* fully abstract, as shown, again, by Plotkin [38].

A lot of research has been devoted to finding fully abstract semantics models for PCF, without `por`. Milner showed that there was only one such model enriched over the category of dcpos, up to isomorphism [27]. Since PCF is sequential, this involved attempts at characterizing sequentiality in denotational models, of which one of the first is Berry's *stable* domain semantics [4]. To our knowledge, the only attempts that were successful in proving full abstraction for domain-theoretic models are due to O'Hearn and Riecke [34], for a call-by-name language, and to Riecke and Sandholm [40], for a call-by-value language. Both are based on refinements of domain-theoretic models that rest on Kripke logical relations of varying arity [22]. Chapters 14 and 15 of Streicher's book [42] give a complete and readable exposition of the matter.

A completely different line of research led to *game semantics* [19, 1]. These are sometimes claimed to provide fully abstract models for PCF, but one should be aware that full abstraction is obtained only after the game model is quotiented by the intrinsic, a.k.a., observational preorder, as in the Milner model.

As far as extensions of PCF with non-deterministic or probabilistic choice are concerned, one must cite Harmer and McCusker [16], who proposed a fully

abstract game semantics for Idealized Algol with non-deterministic choice. (Idealized Algol is essentially PCF with mutable references.) One must also cite Danos and Harmer, who proposed a fully abstract game semantics for Idealized Algol with probabilistic choice [6]. Again, full abstraction is obtained after a quotienting step.

The initial difficulties with PCF and parallel or, and the subsequent successes of game semantics may explain why similar questions were not considered with standard domain semantics instead, i.e., using dcpo-based semantics. Another plausible reason is that the treatment of probabilities in domain semantics (using Jones and Plotkin's probabilistic powerdomain [21]) has its share of problems, too [23]: there is simply no known Cartesian-closed category of continuous domain that would be closed under Jones and Plotkin's probabilistic powerdomain. Although we shall prove soundness and computational adequacy for probabilistic extensions of PCF, with domain-theoretic denotational semantics, the latter difficulties will prevent us from attacking the question of full abstraction with probabilistic choice *alone*.

Curiously, these difficulties simply vanish if we decide to work with extensions of PCF that include non-deterministic choice, either alone or in combination with probabilistic choice. E.g., the category of bc-domains is Cartesian-closed, and closed under both the lower and upper prevision powerdomains of [11], which we shall rely on; see [12].

It is the purpose of this paper to examine questions of soundness, computational adequacy, and full abstraction in extensions of PCF with mixed non-deterministic and probabilistic choice. It has come as a surprise to some that no such work seems to have been done earlier. One possible explanation is that adequate domain-theoretic models of mixed choice have only been invented rather recently. One family of such models is due to Mislove *et al.* [29, 30] and independently by Tix *et al.* [44, 45]. Another one is due to the author, under the name of previsions [11]. In categories of coherent pointed continuous domains, the two kinds of models are isomorphic [12, 15]. Precursors of these models include work by Morgan, McIver and coauthors in computer science [32, 26], or by Walley in mathematical finance and statistics for example [46].

Note added to the final version.. While this paper was being reviewed, Ehrhard, Tasson and Pagani independently produced a denotational model of probabilistic PCF that is fully abstract [7], based on probabilistic coherence spaces. This is a very clever model, of an algebraic rather than a domain-theoretic

nature, where morphisms are power series rather than continuous maps. There are many other differences with our result: we incorporate angelic non-determinism, they do not; in the probabilistic cases, we need statistical termination testers, they do not; our notion of full abstraction is the strong, inequational form (the observational preorder coincides with the denotational preorder), while only the weaker equational form holds in their case. Nonetheless, our two models have in common that, unlike with game semantics, they are fully abstract by themselves, without requiring any further quotient.

3. Preliminaries

We refer the reader to [10, 2, 28] for background on domain theory and topology.

Domain Theory. A set with a partial ordering $\leq$ is a *poset*. We write $\uparrow E$ for $\{y \in X \mid \exists x \in E \,.\, x \leq y\}$, $\downarrow E = \{y \in X \mid \exists x \in E \,.\, y \leq x\}$. A *dcpo* is a poset in which every directed family $(x_i)_{i \in I}$ has a least upper bound (a.k.a., supremum or *sup*) $\sup_{i \in I} x_i$. Symmetrically, we call *inf* (or infimum) any greatest lower bound. A family $(x_i)_{i \in I}$ is *directed* if and only if it is non-empty, and any two elements have an upper bound in the family. Any poset can be equipped with the *Scott topology*, whose opens are the upward closed sets U such that whenever $(x_i)_{i \in I}$ is a directed family that has a least upper bound in U, then some x_i is in U already. A dcpo X is *pointed* if and only if it has a least element, which we shall always write $\bot$.

We shall always consider $\mathbb{R}^+$ (the set of non-negative reals), or $\overline{\mathbb{R}^+} = \mathbb{R}^+ \cup \{+\infty\}$, or $[0, a]$ with $a \in \mathbb{R}^+$—in particular $\mathbb{I} = [0, 1]$— as posets, and implicitly endow them with the Scott topology of their ordering $\leq$. The opens of $\mathbb{R}^+$ in its Scott topology are the intervals $(r, +\infty)$, $r \in \mathbb{R}^+$, together with $\mathbb{R}^+$ and $\emptyset$. Those of $[0, 1]$ are $\emptyset$, $[0, 1]$, and $(r, 1]$, $r \in [0, 1)$. Those of $\overline{\mathbb{R}^+}$ are $\emptyset$, $\overline{\mathbb{R}^+}$, and $(r, +\infty]$, $r \in \mathbb{R}^+$.

Given two dcpos X and Y, a map $f : X \to Y$ is *Scott-continuous* if and only if it is monotonic and $f(\sup_{i \in I} x_i) = \sup_{i \in I} f(x_i)$ for every directed family $(x_i)_{i \in I}$ in X. The space $[X \to Y]$ of all Scott-continuous maps from X to Y is again a dcpo with the pointwise ordering.

The *way-below* relation $\ll$ on a poset X is defined by $x \ll y$ if and only if, for every directed family $(z_i)_{i \in I}$ that has a least upper bound z such that $y \leq z$, then $x \leq z_i$ for some $i \in I$ already. We also say that x

approximates y. Note that $x \ll y$ implies $x \leq y$, and that $x' \leq x \ll y \leq y'$ implies $x' \ll y'$. However, $\ll$ is not reflexive or irreflexive in general. Write $\uparrow\!\!\uparrow E = \{y \in X \mid \exists x \in E \,.\, x \ll y\}$, $\downarrow\!\!\downarrow E = \{y \in X \mid \exists x \in E \,.\, y \ll x\}$. X is *continuous* if and only if, for every $x \in X$, $\downarrow\!\!\downarrow x$ is a directed family, and has x as least upper bound. One may be more precise: A *basis* is a subset B of X such that any element $x \in X$ is the least upper bound of a directed family of elements way-below x *in* B. Then X is continuous if and only if it has a basis, and in this case X itself is the largest basis.

In a continuous poset with basis B, *interpolation holds*: if $x_1, \ldots, x_n$ are finitely many elements way-below x, then there is a $b \in B$ such that $x_1, \ldots, x_n$ are way-below b, and $b \ll x$. (See for example [28, Section 4.2].) Moreover, the Scott-opens are exactly the unions of sets of the form $\uparrow\!\!\uparrow b$, $b \in B$.

The category of continuous dcpos and Scott-continuous maps is not Cartesian-closed. However, several full subcategories are known that are. We shall be especially interested in the category of *bc-domains*. A bc-domain (for bounded complete domain) is a pointed, continuous dcpo in which every pair of elements that has an upper bound has a least upper bound. We shall also be interested in the full subcategory of *continuous lattices*, i.e., of complete lattices that are also continuous. Alternatively, a continuous lattice is exactly a bc-domain with a top element. The unit interval $\mathbb{I}$, and every interval $[0, a]$, $a \in \mathbb{R}^+$, is a continuous lattice, for example.

We shall sometimes need to deal with continuous dcpos (not necessarily pointed) in which every pair of elements that has an upper bound has a least upper bound. Call these *nbc-domains*, for nearly bounded complete domains. The typically example is $\mathbb{N}$, with equality as ordering. Given a dcpo X, $X_\perp$ denotes its *lifting*, i.e., X plus a fresh element $\perp$ below all elements of X. It is clear that X is an nbc-domain if and only if $X_\perp$ is a bc-domain.

Topology. A topology $\mathcal{O}$ on a set X is a collection of subsets of X, called the *opens*, such that any union and any finite intersection of opens is open. The *interior* of a subset A of X is the largest open included in A. A *closed* subset is the complement of an open subset. The *closure* of A is the smallest closed subset containing A.

A topology $\mathcal{O}_1$ is *finer* than $\mathcal{O}_2$ if and only if it contains all opens of $\mathcal{O}_2$. We also say that $\mathcal{O}_2$ is *coarser* than $\mathcal{O}_1$.

A *base* $\mathcal{B}$ (not a basis) of $\mathcal{O}$ is a collection of opens such that every open is a union of elements of the base. Equivalently, a family $\mathcal{B}$ of opens is a base if and only if for every $x \in X$, for every open U containing x, there is a $V \in \mathcal{B}$

such that $x \in V \subseteq U$. A *subbase* of $\mathcal{O}$ is a collection of opens such that the finite intersections of elements of the subbase form a base; equivalently, the coarsest topology containing the elements of the subbase is $\mathcal{O}$, and then we say that $\mathcal{O}$ is *generated* by the subbase.

The specialization preorder of a space X is defined by $x \leq y$ if and only if for every open subset U of X that contains x, U also contains y. For every subbase $\mathcal{B}$ of the topology of X, it is equivalent to say that $x \leq y$ if and only if every $U \in \mathcal{B}$ that contains x also contains y. The specialization preorder of a dcpo X, with ordering $\leq$, in its Scott topology, is $\leq$. A topological space is T_0 if and only if $\leq$ is a partial ordering, not just a preorder. A subset A of X is *saturated* if and only if it is upward-closed in the specialization preorder $\leq$.

A subset K of a topological space X is *compact* if and only if every open cover of K has a finite subcover. A map f from X to Y is *continuous* if and only if $f^{-1}(V)$ is open in X for every open subset V of Y. The image $f[K]$ of any compact subset of X by any continuous map $f \colon X \to Y$ is compact in Y. When X and Y are posets in their Scott topology, $f \colon X \to Y$ is continuous if and only if it is Scott-continuous.

A topological space X is *locally compact* if and only if for every $x \in X$, for every open subset U of X containing x, there is a compact subset $K \subseteq U$ whose interior contains x. Every continuous dcpo is locally compact in its Scott topology.

4. Powerdomains, Previsions

We shall design languages PCF_S, where S is a non-empty set of letters among `A` (for angelic non-determinism), `D` (for demonic non-determinism), and `P` (probabilistic choice). Such sets S are written as the mere concatenation of its elements, so that, e.g., $\mathrm{PCF}_{\mathtt{A}}$ is PCF plus angelic non-deterministic choice, $\mathrm{PCF}_{\mathtt{AP}}$ is PCF plus demonic non-deterministic and probabilistic choice.

Our denotational model for PCF_S will rest on relevant monads from the literature, or variants thereof. For purposes of uniformity, the space $\mathbb{P}_S(X)$ of "choices" of elements of X will always be a dcpo of continuous maps from $[X \to \mathbb{I}]$ to $\mathbb{I}$, where we remind the reader that $\mathbb{I} = [0, 1]$.

Given a space X, let $\mathbf{1}$ denote the constant map from X to $\mathbb{I}$ with value 1. For every $a \in \mathbb{I}$, $a\,\mathbf{1}$ is the constant map with value a.

Definition 4.1 (Prevision). *Let X be a topological space. A* prevision *on X is any continuous map F from $[X \to \mathbb{I}]$ to $\mathbb{I}$ such that:*

- $F(ah) = aF(h)$ *for all* $a \in \mathbb{I}$, $h \in [X \to \mathbb{I}]$;
- $F(a\,\mathbf{1} + (1-a)h) \leq a + (1-a)F(h)$ *for all* $a \in \mathbb{I}$, $h \in [X \to \mathbb{I}]$.

Previsions, and more generally all the functionals we shall consider, are ordered pointwise: $F \leq F'$ if and only if $F(h) \leq F'(h)$ for every $h \in [X \to \mathbb{I}]$.

Remark 4.2. *Up to isomorphism, these are the same objects as those which we called continuous, subnormalized previsions in [11].*

The latter were defined as continuous maps F from the poset $\langle X \to \mathbb{R}^+\rangle$ of all bounded Scott-continuous maps from X to $\mathbb{R}^+$ to $\mathbb{R}^+$ (with the pointwise ordering), such that $F(ah) = aF(h)$ and $F(a+h) \leq a+F(h)$ for all $a \in \mathbb{R}^+$, $h \in \langle X \to \mathbb{R}^+\rangle$. Such a map yields a prevision in the sense of Definition 4.1 by restricting it to $[X \to \mathbb{I}]$, noticing that $F(h)$ must then be in $\mathbb{I}$ for every $h \in [X \to \mathbb{I}]$. (Indeed, $F(0\,\mathbf{1}) = 0$ since $F(0h) = 0.F(h)$ for every h; so $F(\mathbf{1}) = F(\mathbf{1} + 0\,\mathbf{1}) \leq 1 + F(0\,\mathbf{1}) = 1$, whence we infer $F(h) \leq F(\mathbf{1}) = 1$.) Conversely, any prevision F in the sense of Definition 4.1 yields a continuous, subnormalized prevision in the sense of [11] by extending it to every $h \in \langle X \to \mathbb{R}^+\rangle$ as yielding $aF(h/a)$, for any upper bound $a \in \mathbb{R}^+$ on the values of h.

To deal with the non-probabilistic cases, we introduce the following. A map between pointed dcpos is called *strict* if and only if it maps bottom to bottom. In particular, a strict map $f\colon \mathbb{I} \to \mathbb{I}$ is one that maps 0 to 0.

Definition 4.3 (Discrete Prevision). *Let X be a topological space. A functional $F\colon [X \to \mathbb{I}] \to \mathbb{I}$ is* discrete *if and only if, for every $h \in [X \to \mathbb{I}]$ and every strict Scott-continuous map $f\colon \mathbb{I} \to \mathbb{I}$, $F(f \circ h) = f(F(h))$. A* discrete prevision *is a prevision that is discrete in this sense.*

Discreteness implies $F(ah) = aF(h)$ for every $a \in \mathbb{I}$, by taking $f(t) = at$. Therefore, in verifying that a functional is a discrete prevision, we only need to verify that it is continuous, discrete, and satisfies the second property of Definition 4.1, not the first one.

Discreteness also implies that F sends every continuous map $h\colon X \to \{0,1\}$ to a value in $\{0,1\}$: letting f map 0 to 0 and every $t \in (0,1]$ to 1, discreteness together with $f \circ h = h$ implies $F(h) = f(F(h)) \in \{0,1\}$. More

precisely, the restriction $F_{|[X\to\{0,1\}]}$ is an element of the poset $[[X \to \{0,1\}] \to \{0,1\}]_0$ of strict continuous functionals from $[X \to \{0,1\}]$ to $\{0,1\}$. Here, the set $\{0,1\}$ is equipped with the ordering induced from $\mathbb{I}$; this is sometimes called Sierpiński space.

Lemma 4.4. *The restriction map that sends every discrete prevision F on X to $F_{|[X\to\{0,1\}]} \in [[X \to \{0,1\}] \to \{0,1\}]_0$ is an order isomorphism. Its inverse maps $\Phi \in [[X \to \{0,1\}] \to \{0,1\}]_0$ to F defined by $F(h) = \sup_{t\in\mathbb{I}} t\Phi(\chi_{h^{-1}(\Uparrow t)})$.*

In order to clarify the notation $\Uparrow t$ in the last statement, note that the way-below relation $\ll$ on $\mathbb{I}$ is defined by $t \ll u$ if and only if $t = 0$ or $t < u$. Accordingly, $\Uparrow t$ is equal to $(t, 1]$ if $t \neq 0$, and to $[0, 1]$ if $t = 0$. Also, we write χ_U for the characteristic map of U.

Proof. We first draw a few consequences of the fact that F is discrete. χ_U is a continuous map if and only if U is open. Also, all the continuous maps $h\colon X \to \{0,1\}$ are of this form, with $U = h^{-1}(0,1]$.

For $h \in [X \to \mathbb{I}]$, using the definition of discrete previsions with $f = \chi_{\Uparrow t}$, so that $f \circ h = \chi_{h^{-1}(\Uparrow t)}$, we obtain $F(\chi_{h^{-1}(\Uparrow t)}) = \chi_{\Uparrow t}(F(h))$. Multiplying by t and taking sups, while noticing that the pointwise supremum $\sup_{t\in\mathbb{I}} t\chi_{\Uparrow t}$ is the identity map on $\mathbb{I}$, $F(h) = \sup_{t\in\mathbb{I}} tF(\chi_{h^{-1}(\Uparrow t)})$.

We can therefore define the inverse map as mapping $\Phi \in [[X \to \{0,1\}] \to \{0,1\}]_0$ to the functional F defined by $F(h) = \sup_{t\in\mathbb{I}} t\Phi(\chi_{h^{-1}(\Uparrow t)})$. Note that, conversely, the restriction of the latter to $[X \to \{0,1\}]$ is Φ: for every $h \in [X \to \{0,1\}]$, write $h = \chi_U$ for some open subset U of X, observe that $h^{-1}(\Uparrow t)$ is equal to X if $t = 0$, to U if $0 < t < 1$ and to $\emptyset$ if $t = 1$, so that $F(h) = \sup_{t\in(0,1)} t\Phi(\chi_U) = \Phi(h)$.

The only remaining non-obvious claim is that F really is a discrete prevision. Scott-continuity is easy and $F(ah) = aF(h)$ will follow from discreteness, as explained in the remark following Definition 4.3.

We deal with the case where $\Phi(\chi_X) = 0$ separately. In that case, Φ is identically zero, so $F = 0$ as well, and F is therefore a discrete prevision. Otherwise, assume $\Phi(\chi_X) \neq 0$. Then $F(h)$ is the supremum of all $t \in \mathbb{I}$ such that $\Phi(\chi_{h^{-1}(\Uparrow t)}) = 1$. This is a non-empty family, and it is easily seen to be directed.

Let us show that F is discrete. Consider a strict Scott-continuous function $f\colon \mathbb{I} \to \mathbb{I}$. We first show that $f(F(h)) \leq F(f \circ h)$. To this end, pick any $t \ll f(F(h))$, and by interpolation find v such that $t \ll v \ll f(F(h))$. Since

$F(h)$ is the (directed) sup of all $u \in \mathbb{I}$ such that $\Phi(\chi_{h^{-1}(\uparrow u)}) = 1$, and f is continuous, $f(F(h))$ is the sup of all $f(u)$ for $u \in \mathbb{I}$ such that $\Phi(\chi_{h^{-1}(\uparrow u)}) = 1$. Since v is way-below this sup, there is an $u \in \mathbb{I}$ such that $v \leq f(u)$ and $\Phi(\chi_{h^{-1}(\uparrow u)}) = 1$. Since $t \ll v \leq f(u)$, $u \in f^{-1}(\Uparrow t)$, so $\Uparrow u \subseteq f^{-1}(\Uparrow t)$, from which we obtain $h^{-1}(\Uparrow u) \subseteq h^{-1}(f^{-1}(\Uparrow t)) = (f \circ h)^{-1}(\Uparrow t)$. It follows that $\Phi(\chi_{(f\circ h)^{-1}(\uparrow t)}) = 1$. By definition $F(f \circ h)$ must then be at least t. Taking sups over all $t \ll f(F(h))$, we obtain $f(F(h)) \leq F(f \circ h)$.

For the converse inequality, pick any $t \ll F(f \circ h)$, so there is an element $u \in \mathbb{I}$ such that $t \leq u$ and $\Phi(\chi_{(f\circ h)^{-1}(\uparrow u)}) = 1$. The open subset $f^{-1}(\Uparrow u)$ is non-empty: if it were empty, then $\chi_{(f\circ h)^{-1}(\uparrow u)} = \chi_{h^{-1}(f^{-1}(\uparrow u))}$ would be identically 0, so $\Phi(\chi_{(f\circ h)^{-1}(\uparrow u)})$ would be equal to 0 by strictness, not 1. As an open subset of $\mathbb{I}$, $f^{-1}(\Uparrow u)$ is the union of all subsets of the form $\Uparrow v$ where $v \in f^{-1}(\Uparrow u)$, and this union is non-empty (because $f^{-1}(\Uparrow u)$ is non-empty) and directed (because $\mathbb{I}$ is a chain). It follows that $\chi_{(f\circ h)^{-1}(\Uparrow u)} = \chi_{h^{-1}(f^{-1}(\uparrow u))}$ is the sup of the directed family of all maps $\chi_{h^{-1}(\uparrow v)}$, $v \in f^{-1}(\Uparrow u)$. Since Φ maps it to 1, and is Scott-continuous, there is a $v \in f^{-1}(\Uparrow u)$ such that $\Phi(\chi_{h^{-1}(\uparrow v)}) = 1$. By definition, $F(h) \geq v$, and $v \in f^{-1}(\Uparrow u)$ implies that $f(v) \geq u \geq t$, so $f(F(h)) \geq t$. Taking sups over all $t \ll F(f \circ h)$, $f(F(h)) \geq F(f \circ h)$.

We must finally show that $F(a\,\mathbf{1} + (1-a)h) \leq a + (1-a)F(h)$. When $a = 1$, this means $F(\mathbf{1}) \leq 1$, which is clear. Otherwise, it will be easier to consider an equivalent definition of F, viz., $F(h) = \sup_{t\in\mathbb{I}} t\Phi(\chi_{h^{-1}(t,1]})$. Indeed $(t, 1]$ only differs from $\Uparrow t$ when $t = 0$, a case that contributes nothing to the sup. For every $t \in \mathbb{I}$, $(a\,\mathbf{1} + (1-a)h)^{-1}(t, 1] = h^{-1}(\frac{t-a}{1-a}, 1]$ if $t \geq a$ and $(a\,\mathbf{1} + (1-a)h)^{-1}(t, 1] = X$ if $t < a$, so $F(a\,\mathbf{1} + (1-a)h) = \sup(\sup_{t\in[0,a)} t\Phi(\mathbf{1}), \sup_{t\in[a,1]} t\Phi(\chi_{h^{-1}(\frac{t-a}{1-a},1]}))$. Now $\Phi(\mathbf{1}) \leq 1$, and renaming $\frac{t-a}{1-a}$ as u, $\sup_{t\in[a,1]} t\Phi(\chi_{h^{-1}(\frac{t-a}{1-a},1]}) = \sup_{u\in[0,1]}(a + (1-a)u)\Phi(\chi_{h^{-1}(u,1]}) \leq a + (1-a)\sup_{u\in\mathbb{I}} u\Phi(\chi_{h^{-1}(u,1]}) = a + (1-a)F(h)$. □

In particular, we could have simply defined the discrete previsions as the strict Scott-continuous maps from $[X \to \{0, 1\}]$ to $\{0, 1\}$. One advantage of not doing so is the uniformity of the following definition.

Definition 4.5 ($\mathbb{P}_S$). *Let X be a topological space.*

($\mathbb{P}_\mathtt{P}$) $\mathbb{P}_\mathtt{P}(X)$ *is the set of all previsions F on X that are* linear*, i.e., such that $F(ah + (1-a)h') = aF(h) + (1-a)F(h')$ for all $a \in \mathbb{I}$, $h, h' \in [X \to \mathbb{I}]$.*

($\mathbb{P}_{\mathtt{AP}}$) $\mathbb{P}_{\mathtt{AP}}(X)$ *is the set of all previsions* F *on* X *that are* sublinear*, i.e., such that* $F(ah + (1-a)h') \leq aF(h) + (1-a)F(h')$ *for all* $a \in \mathbb{I}$, $h, h' \in [X \to \mathbb{I}]$.

($\mathbb{P}_{\mathtt{A}}$) $\mathbb{P}_{\mathtt{A}}(X)$ *is the set of all sublinear discrete previsions on* X.

($\mathbb{P}_{\mathtt{DP}}$) $\mathbb{P}_{\mathtt{DP}}(X)$ *is the set of all previsions* F *on* X *that are* superlinear*, i.e., such that* $F(ah + (1-a)h') \geq aF(h) + (1-a)F(h')$ *for all* $a \in \mathbb{I}$, $h, h' \in [X \to \mathbb{I}]$.

($\mathbb{P}_{\mathtt{D}}$) $\mathbb{P}_{\mathtt{D}}(X)$ *is the set of all superlinear discrete previsions on* X.

($\mathbb{P}_{\mathtt{ADP}}$) $\mathbb{P}_{\mathtt{ADP}}(X)$ *is the set of all* forks *on* X*, i.e., of all pairs* $(F^-, F^+) \in \mathbb{P}_{\mathtt{DP}}(X) \times \mathbb{P}_{\mathtt{AP}}(X)$ *that satisfy* Walley's condition*:*

$$F^-(ah + (1-a)h') \leq aF^-(h) + (1-a)F^+(h') \leq F^+(ah + (1-a)h')$$

for all $a \in \mathbb{I}$, $h, h' \in [X \to \mathbb{I}]$.

($\mathbb{P}_{\mathtt{AD}}$) $\mathbb{P}_{\mathtt{AD}}(X)$ *is the set of discrete forks, i.e., of forks* (F^-, F^+) *where both* F^- *and* F^+ *are discrete previsions.*

Walley's condition implies, but is not equivalent to, $F^- \leq F^+$. We shall see in Lemma 4.7 3 that Walley's condition simplifies somewhat in the discrete ($\mathbb{P}_{\mathtt{ADP}}$) case. In the general case, Walley's condition has the following geometrical interpretation. Imagine you draw F^- and F^+ as curves, with h on the x-axis (see Figure 1). Walley's condition means that the line segment from the point on the F^- curve with x value h to the point on the F^+ curve with x value h' remains entirely between the two curves (as shown left), and never crosses the curves (as shown right).

Recall that previsions are ordered pointwise. Forks are ordered componentwise: $(F_1^-, F_1^+) \leq (F_2^-, F_2^+)$ if and only if $F_1^- \leq F_2^-$ and $F_1^+ \leq F_2^+$. Then $\mathbb{P}_S(X)$ is a pointed dcpo, whatever S is.

Up to the isomorphism mentioned in Remark 4.2, the elements of $\mathbb{P}_S(X)$ coincide with the linear (resp., upper, lower) continuous, subnormalized previsions on X of [11] when $S = \mathtt{P}$ (resp., $\mathtt{D}$, $\mathtt{A}$), which were proposed as denotational models of probabilistic (resp., probabilistic plus demonic non-deterministic, probabilistic plus angelic non-deterministic) choice in op. cit. Notably, $\mathbb{P}_{\mathtt{P}}(X)$, is isomorphic to Jones and Plotkin's (sub)probabilistic powerdomain $\mathbf{V}_{\leq 1}(X)$ [21, 20]. The latter is the dcpo of all subnormalized continuous valuations, that is, all Scott-continuous maps ν from the dcpo $\mathcal{O}(X)$

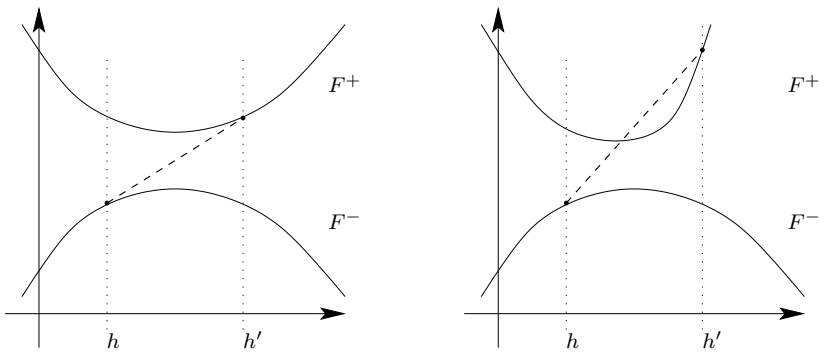

Figure 1: Walley's condition: OK (left), fails (right)

of open subsets of X to $\mathbb{I}$ such that $\nu(\emptyset) = 0$ and $\nu(U \cup V) + \nu(U \cap V) = \nu(U) + \nu(V)$ for all $U, V \in \mathcal{O}(X)$. The isomorphism maps $F \in \mathbb{P}_{\mathtt{P}}(X)$ to $\nu \in \mathbf{V}_{\leq 1}(X)$ defined by $\nu(U) = F(\chi_U)$, where χ_U is the characteristic map of U. Conversely, one retrieves F from ν by $F(h) = \int_{x \in X} h(x) d\nu$.

Remark 4.6. *A prevision F is* normalized *if and only if $F(a\,\mathbf{1} + (1-a)h) = a + (1-a)F(h)$ for all $a \in \mathbb{I}$, $h \in [X \to \mathbb{I}]$. We write $\mathbb{P}^1_S(X)$ for the subset of those normalized elements of $\mathbb{P}_S(X)$.*

We have chosen to use sub*normalized previsions (or valuations), not normalized ones (or probability valuations). This allows us to account for nontermination, represented as the constant 0 prevision. We might also have decided on using normalized previsions on $X_\perp$ instead of X, where $X_\perp$ is X plus a fresh least element $\perp$. The theory would essentially be equivalent, since $\mathbb{P}_S(X)$ is isomorphic to $\mathbb{P}^1_S(X_\perp)$. We leave this as an exercise: map $F \in \mathbb{P}_S(X)$ to $(h' \in [X_\perp \to \mathbb{I}] \mapsto F(h'_{|X} - h'(\perp)\,\mathbf{1}) + h'(\perp))$, where $h'_{|X}$ is the restriction of h' to X, and conversely, map $F' \in \mathbb{P}^1_S(X_\perp)$ to $(h \in [X \to \mathbb{I}] \mapsto F'(\widehat{h}))$, where $\widehat{h}(x) = h(x)$ for every $x \in X$ and $\widehat{h}(\perp) = 0$.*

When no probabilistic choice is involved, the domains of choices are usually given by the Hoare, Smyth, and Plotkin powerdomains, rather than the dcpos $\mathbb{P}_{\mathtt{A}}(X)$, $\mathbb{P}_{\mathtt{D}}(X)$, $\mathbb{P}_{\mathtt{AD}}(X)$ described above. We claim that these are the same dcpos, up to isomorphism, in all practical situations. Historically, one may trace the idea of characterizing these powerdomains as domains of second-order functionals to Heckmann [17].

We start by observing that the conditions of linearity, sublinearity, superlinearity, and Walley's condition simplify somehow in the discrete case.

Lemma 4.7. *The isomorphism of Lemma 4.4 specializes to one between:*

1. *sublinear discrete previsions and those strict continuous functionals* $\Phi \in [[X \to \{0,1\}] \to \{0,1\}]_0$ *that are* sup-preserving*, namely such that for all open subsets* U, V *of* X*,* $\Phi(\chi_{U\cup V}) = \max(\Phi(\chi_U), \Phi(\chi_V))$*;*
2. *superlinear discrete previsions and strict continuous functionals that are* inf-preserving*, namely such that* $\Phi(\chi_{U\cap V}) = \min(\Phi(\chi_U), \Phi(\chi_V))$ *for all opens* U, V*;*
3. *discrete forks and* Heckmann pairs*, namely pairs* (Φ^-, Φ^+) *of an inf-preserving strict continuous functional* Φ^- *and a sup-preserving strict continuous functional* Φ^+*, such that for all opens* U, V *of* X*, if* $\Phi^+(\chi_V) = 0$ *then* $\Phi^-(\chi_{U\cup V}) = \Phi^-(\chi_U)$*, and if* $\Phi^-(\chi_U) = 1$ *then* $\Phi^+(\chi_{U\cap V}) = \Phi^+(\chi_V)$*.*

In the latter case, note that Heckmann pairs satisfy $\Phi^- \leq \Phi^+$. Indeed, for every open subset V, if $\Phi^+(\chi_V) = 0$, then $\Phi^-(\chi_V) = \Phi^-(\chi_{\emptyset\cup V}) = \Phi^-(\chi_\emptyset) = 0$. Heckmann pairs are a direct reformulation of Heckmann's A-valuations [18], without the normalisation condition.

Proof. Write F for an arbitrary discrete prevision, and Φ for its restriction to $[X \to \{0,1\}]$.

1. If F is sublinear, then taking $a = 1/2$, $h = \chi_U$ and $h' = \chi_V$, $F(1/2\chi_U + 1/2\chi_V) \leq 1/2\Phi(\chi_U) + 1/2\Phi(\chi_V)$. Since $\chi_{U\cup V} \leq \chi_U + \chi_V$, $1/2\Phi(\chi_{U\cup V}) \leq 1/2\Phi(\chi_U) + 1/2\Phi(\chi_V)$, which implies that if $\Phi(\chi_U) = \Phi(\chi_V) = 0$, then $\Phi(\chi_{U\cup V}) = 0$. If $\Phi(\chi_U)$ or $\Phi(\chi_V)$ is equal to 1, then $\Phi(\chi_{U\cup V}) = 1$ by monotonicity. In any case, $\Phi(\chi_{U\cup V}) = \max(\Phi(\chi_U), \Phi(\chi_V))$.

Conversely, assume Φ is sup-preserving. Given any non-empty family $(U_i)_{i\in I}$ of open subsets of X, we can write their union as the directed union of all finite unions $\bigcup_{i\in J} U_i$, when J ranges over the non-empty subsets of I. Since Φ is Scott-continuous and $\cup$-preserving, $\Phi(\chi_{\bigcup_{i\in I} U_i}) = \sup_{i\in I} \Phi(\chi_{U_i})$: we say that Φ preserves non-empty sups.

Recall that F is obtained from Φ by the formula $F(h) = \sup_{t\in\mathbb{I}} t\Phi(\chi_{h^{-1}(\uparrow t)})$.

For every element x in the open subset $(ah + (1-a)h')^{-1}(\uparrow t)$, $t \ll ah(x) + (1-a)h'(x)$, which implies that there are elements $u, u' \in \mathbb{I}$ such that $u \ll h(x)$, $u' \ll h'(x)$, and $t \leq au + (1-a)u'$. It follows that $(ah + (1-a)h')^{-1}(\uparrow t)$ is included in the union $\bigcup_{u,u'\in\mathbb{I}, t\leq au+(1-a)u'} (h^{-1}(\uparrow u) \cap h'^{-1}(\uparrow u'))$. Since Φ preserves non-empty sups, $\Phi(\chi_{(ah+(1-a)h')^{-1}(\uparrow t)})$ is equal to $\sup_{u,u'\in\mathbb{I}, t\leq au+(1-a)u'}$

$\Phi(\chi_{h^{-1}(\uparrow u)\cap h'^{-1}(\uparrow u')})$, and therefore:

$$\begin{aligned} F(ah+(1-a)h') &= \sup_{\substack{t,u,u'\in\mathbb{I}\\ t\leq au+(1-a)u'}} t\Phi(\chi_{h^{-1}(\uparrow u)\cap h'^{-1}(\uparrow u')}) \\ &\leq \sup_{\substack{t,u,u'\in\mathbb{I}\\ t\leq au+(1-a)u'}} (au+(1-a)u')\Phi(\chi_{h^{-1}(\uparrow u)\cap h'^{-1}(\uparrow u')}) \\ &\leq \sup_{\substack{t,u,u'\in\mathbb{I}\\ t\leq au+(1-a)u'}} (au\Phi(\chi_{h^{-1}(\uparrow u)})+(1-a)u'\Phi(\chi_{h'^{-1}(\uparrow u')})) \\ &= aF(h)+(1-a)F(h'). \end{aligned}$$

Thus 1 is proved. Before we proceed to 2, let us notice the following Claim $(*)$: for every discrete prevision F, with restriction $\Phi = F_{|[X\to\{0,1\}]}$, and for all open subsets U and V of X, $F(1/2\chi_U + 1/2\chi_V) = 1/2\Phi(\chi_{U\cup V}) + 1/2\Phi(\chi_{U\cap V})$. This is proved as follows. For $h = 1/2\chi_U + 1/2\chi_V$, we note that $h^{-1}(\Uparrow t)$ is equal to $\emptyset$ if $t = 1$, to $U\cap V$ if $1/2 \leq t < 1$, to $U\cup V$ if $0 < t < 1/2$, and to X if $t = 0$. Using the formula $F(h) = \sup_{t\in\mathbb{I}} t\Phi(\chi_{h^{-1}(\Uparrow t)})$, we obtain $F(h) = \sup(\sup_{1/2\leq t<1} t\Phi(\chi_{U\cap V}), \sup_{0<t<1/2} t\Phi(\chi_{U\cup V})) = \sup(\Phi(\chi_{U\cap V}), 1/2\Phi(\chi_{U\cup V}))$. The latter can only have three possible values: 1 if $\Phi(\chi_{U\cap V}) = 1$ (hence also $\Phi(\chi_{U\cup V}) = 1$), 1/2 if $\Phi(\chi_{U\cap V} = 0)$ and $\Phi(\chi_{U\cup V}) = 1$, 0 if $\Phi(\chi_{U\cap V}) = \Phi(\chi_{U\cup V}) = 0$. These are the same values as for $1/2\Phi(\chi_{U\cup V}) + 1/2\Phi(\chi_{U\cap V})$.

2. If F is superlinear, then $F(1/2\chi_U + 1/2\chi_V) \geq 1/2\Phi(\chi_U) + 1/2\Phi(\chi_V)$. Using $(*)$, $1/2\Phi(\chi_{U\cup V})+1/2\Phi(\chi_{U\cap V}) \geq 1/2\Phi(\chi_U)+1/2\Phi(\chi_V)$. If both $\Phi(\chi_U)$ and $\Phi(\chi_V)$ equal 1, then this inequality implies that $\Phi(\chi_{U\cap V}) = \Phi(\chi_{U\cup V}) = 1$. In all other cases, $\Phi(\chi_{U\cap V}) = 0$ by monotonicity. In any case, $\Phi(\chi_{U\cap V}) = \min(\Phi(\chi_U), \Phi(\chi_V))$.

Conversely, assume that Φ is inf-preserving. We compute $aF(h) + (1-a)F(h') = \sup_{u,u'\in\mathbb{I}}(au\Phi(\chi_{h^{-1}(\uparrow u)})+(1-a)u'\Phi(\chi_{h'^{-1}(\uparrow u')}))$. Using the fact that Φ takes its values in $\{0,1\}$, one checks that this is the sup over the values $u, u' \in \mathbb{I}$ such that $\Phi(\chi_{h^{-1}(\uparrow u)}) = \Phi(\chi_{h'^{-1}(\uparrow u')}) = 1$ of the quantity $au+(1-a)u'$ (the other terms of the supremum do not contribute any larger value). For every such pair u, u', since Φ is inf-preserving, $\Phi(\chi_{h^{-1}(\uparrow u)\cap h'^{-1}(\uparrow u')}) = 1$. Since $h^{-1}(\Uparrow u) \cap h'^{-1}(\Uparrow u') \subseteq (ah+(1-a)h')^{-1}(\Uparrow(au+(1-a)u'))$, we obtain that $aF(h) + (1-a)F(h')$ is less than or equal to the sup of the quantity $au + (1-a)u'$ over all pairs $u, u' \in \mathbb{I}$ such that $\Phi(\chi_{(ah+(1-a)h')^{-1}(\uparrow(au+(1-a)u'))}) = 1$; this is at most the sup of all $t \in \mathbb{I}$ such that $\Phi(\chi_{(ah+(1-a)h')^{-1}(\uparrow t)}) = 1$, namely $F(ah+(1-a)h')$.

3. Given a pair (F^-, F^+) of discrete previsions, write Φ^- (Φ^+) for the restriction of F^- (F^+) to $[X \to \{0,1\}]$.

If (F^-, F^+) is a fork, the first half of Walley's condition with $a = 1/2$, $h = \chi_U$, $h' = \chi_V$ yields $F^-(1/2\chi_U + 1/2\chi_V) \leq 1/2\Phi^-(\chi_U) + 1/2\Phi^+(\chi_V)$. In particular, if $\Phi^+(\chi_V) = 0$, then $\Phi^-(\chi_{U\cup V}) \leq 2F^-(1/2\chi_U + 1/2\chi_V) \leq \Phi^-(\chi_U)$. Since Φ^- is monotonic, we obtain $\Phi^-(\chi_{U\cup V}) = \Phi^-(\chi_U)$. The second half of Walley's condition yields $1/2\Phi^-(\chi_U) + 1/2\Phi^+(\chi_V) \leq F^+(1/2\chi_U + 1/2\chi_V)$. The right-hand side, $F^+(1/2\chi_U + 1/2\chi_V)$, is equal to $1/2\Phi^+(\chi_{U\cup V}) + 1/2\Phi^+(\chi_{U\cap V})$ by $(*)$, so, if $\Phi^-(\chi_U) = 1$ (hence also $\Phi^-(\chi_{U\cup V}) = 1$), then $\Phi^+(\chi_V) \leq \Phi^+(\chi_{U\cap V})$, whence $\Phi^+(\chi_V) = \Phi^+(\chi_{U\cap V})$.

Conversely, if (Φ^-, Φ^+) is a Heckmann pair, then we claim that (F^-, F^+) is a fork.

We first prove $F^-(ah + (1-a)h') \leq aF^-(h) + (1-a)F^+(h')$, which is the more difficult of the two inequalities that we must establish. Consider an arbitrary $t \in \mathbb{I}$ such that $\Phi^-(\chi_{(ah+(1-a)h')^{-1}(\uparrow t)}) = 1$. As in 1, $(ah + (1-a)h')^{-1}(\uparrow t)$ is included in the union $\bigcup_{u,u' \in \mathbb{I}, t \leq au+(1-a)u'} (h^{-1}(\uparrow u) \cap h'^{-1}(\uparrow u'))$. We split this union into the terms such that $\Phi^+(\chi_{h'^{-1}(\uparrow u')}) = 1$ and those such that $\Phi^+(\chi_{h'^{-1}(\uparrow u')}) = 0$. We take an even larger subset by replacing the terms $h^{-1}(\uparrow u) \cap h'^{-1}(\uparrow u')$ of the first kind by $h^{-1}(\uparrow u)$, and those of the second kind by $h'^{-1}(\uparrow u')$. To sum up, $(ah + (1-a)h')^{-1}(\uparrow t)$ is included in the union $U \cup V$, where:

- U is the union of the subsets $h^{-1}(\uparrow u)$ where u ranges over those elements of $\mathbb{I}$ such that, for some $u' \in \mathbb{I}$, $t \leq au + (1-a)u'$ and $\Phi^+(\chi_{h'^{-1}(\uparrow u')}) = 1$;
- V is the union of the subsets $h'^{-1}(\uparrow u')$ when u' ranges over those elements of $\mathbb{I}$ such that $\Phi^+(\chi_{h'^{-1}(\uparrow u')}) = 0$.

The union defining V is directed, since the sets $h'^{-1}(\uparrow u')$ form a chain ($\mathbb{I}$ is a chain), and a non-empty one, because $u' = 1$ satisfies $\Phi^+(\chi_{h'^{-1}(\uparrow u')}) = 0$, since Φ^+ is strict. Since Φ^+ is Scott-continuous, $\Phi^+(\chi_V) = 0$.

Recalling that $\Phi^-(\chi_{(ah+(1-a)h')^{-1}(\uparrow t)}) = 1$, and because Φ^- is monotonic, $\Phi^-(\chi_{U\cup V}) = 1$. Using the definition of a Heckmann pair, $\Phi^-(\chi_U) = 1$. In particular, U is non-empty, since Φ^- is strict. This implies that the union defining U is also non-empty, and as it is a chain, it is directed. Since Φ^- is Scott-continuous and $\Phi^-(\chi_U) = 1$, one of the terms in the union defining U must have a Φ^- value of 1, that is: there is a $u \in \mathbb{I}$ such that $\Phi^-(\chi_{h^{-1}(\uparrow u)}) = 1$ and, for some $u' \in \mathbb{I}$, $t \leq au + (1-a)u'$ and $\Phi^+(\chi_{h'^{-1}(\uparrow u')}) = 1$. It follows

that $t \leq aF^-(h) + (1-a)F^+(h')$, and by taking sups over those t such that $\Phi^-(\chi_{(ah+(1-a)h')^{-1}(\Uparrow t)}) = 1$, $F^-(ah + (1-a)h') \leq aF^-(h) + (1-a)F^+(h')$.

It remains to establish the inequality $aF^-(h) + (1-a)F^+(h') \leq F^+(ah + (1-a)h')$. We should in principle distinguish four cases, depending on whether there is a $u \in \mathbb{I}$ such that $\Phi^-(\chi_{h^{-1}(\Uparrow u)}) = 1$, and whether there is a $u' \in \mathbb{I}$ such that $\Phi^+(\chi_{h'^{-1}(\Uparrow u')}) = 1$. If there is no such u, then $F^-(h) = 0$, and the claim is obvious. If there is no such u', then $F^+(h') = 0$, and the claim is obvious as well. So assume there is a $u \in \mathbb{I}$ such that $\Phi^-(\chi_{h^{-1}(\Uparrow u)}) = 1$ and a $u' \in \mathbb{I}$ such that $\Phi^+(\chi_{h'^{-1}(\Uparrow u')}) = 1$, and consider arbitrary such values. Using the definition of a Heckmann pair, $\Phi^+(\chi_{h^{-1}(\Uparrow u) \cap h'^{-1}(\Uparrow u')}) = \Phi^+(\chi_{h'^{-1}(\Uparrow u')}) = 1$. Every element of $h^{-1}(\Uparrow u) \cap h'^{-1}(\Uparrow u')$ is in $(ah + (1-a)h')^{-1}(\Uparrow t)$, where $t = au + (1-a)u'$. So $\Phi^+(\chi_{(ah+(1-a)h')^{-1}(\Uparrow t)}) = 1$. The latter implies $F^+(ah + (1-a)h') \geq t = au + (1-a)u'$. Taking sups over u, u', we obtain $F^+(ah + (1-a)h') \geq aF^-(h) + (1-a)F^+(h')$. □

The standard model of angelic non-deterministic choice is the *Hoare powerdomain* $\mathcal{H}(X)$. This is the dcpo of all closed non-empty subsets of X, ordered by inclusion. In order to account for non-termination again, we shall consider the *lifted* Hoare powerdomain $\mathcal{H}_\perp(X)$. This is the dcpo of all closed subsets of X, including the empty one $\emptyset$, which acts as bottom element. (And again, $\mathcal{H}_\perp(X)$ is isomorphic to $\mathcal{H}(X_\perp)$.) The essence of the following result is due to Heckmann [17, Proposition 18.3.1].

Proposition 4.8. *Let X be a topological space. (i) $\mathbb{P}_{\mathtt{A}}(X)$ and $\mathcal{H}_\perp(X)$ are isomorphic. (ii) $\mathbb{P}^1_{\mathtt{A}}(X)$ and $\mathcal{H}(X)$ are isomorphic.*

Proof. In view of Lemma 4.7 (1), we can equate $\mathbb{P}_{\mathtt{A}}(X)$ with the strict, sup-preserving continuous functionals. Heckmann showed that the latter formed a poset that is isomorphic to $\mathcal{H}_\perp(X)$. Here is the argument. Given any poset Z, we can equate the Scott-continuous maps from Z to $\{0,1\}$ with (the characteristic maps of) open subsets of Z: $[Z \to \{0,1\}] \cong \mathcal{O}(Z)$. Modulo this identification, $[[X \to \{0,1\}] \to \{0,1\}]_0$ is the collection of Scott-open subsets of $\mathcal{O}(X)$ not containing $\emptyset$. The sup-preserving functionals are equated with what Heckmann calls the *open grills*, i.e., the non-trivial Scott-open subsets G of $\mathcal{O}(X)$ such that, for all $U, V \in \mathcal{O}(X)$ with $U \cup V \in G$, U or V is in G. Given any open grill G, the union of any family of open sets not in G must still be outside G, so there is a largest open subset not in G. The isomorphism of (i) maps G (or the isomorphic sublinear discrete prevision

F) to the complement of this largest open subset. The isomorphism of (ii) is obtained by restriction to the relevant subsets. □

Remark 4.9. *Explicitly, the isomorphism maps F to the complement of the largest open subset U of X such that $F(\chi_U) = 0$. Its inverse maps every closed subset C of X to the open grill of all the open subsets U that intersect C; equivalently, to the sublinear discrete prevision that maps $h \in [X \to \{0,1\}]$ to $\sup_{x \in C} h(x)$. (The actual formula from Lemma 4.4 is $\sup\{t \in \mathbb{I} \mid h^{-1}(\uparrow t)$ intersects $C\}$. But $h^{-1}(\uparrow t)$ intersects C if and only if for some $x \in C$, $t \ll h(x)$, and this is equivalent to $t \ll \sup_{x \in C} h(x)$. The sup of these numbers t is exactly $\sup_{x \in C} h(x)$.)*

The standard model of demonic non-deterministic choice is the *Smyth powerdomain* $\mathcal{Q}(X)$ of all non-empty compact saturated subsets of X, ordered by reverse inclusion. We consider its lifted version: $\mathcal{Q}_\perp(X)$ is $\mathcal{Q}(X)$ plus an extra least element. (Again, this is isomorphic to $\mathcal{Q}(X_\perp)$, where the least element of the latter is $X_\perp$ itself.) The following is a slight variant of a result by Heckmann [17, Proposition 19.2.1].

Proposition 4.10. *Let X be a sober topological space (e.g., a continuous dcpo). (i) $\mathbb{P}_\mathtt{D}(X)$ and $\mathcal{Q}_\perp(X)$ are isomorphic. (ii) $\mathbb{P}^1_\mathtt{D}(X)$ and $\mathcal{Q}(X)$ are isomorphic.*

We will not define what sober means, see [10, Definition O-5.6] or [2, Section 7.2.1]. In this paper, we only use the fact that every continuous dcpo is sober in its Scott topology, see [10, Corollary II-1.12] or [2, Proposition 7.2.27].

Proof. Using Lemma 4.7 (2), and equating $[Z \to \{0,1\}]$ with $\mathcal{O}(Z)$ as above, $\mathbb{P}_\mathtt{D}(X)$ is isomorphic to the poset of all Scott-open subsets $\mathcal{F}$ of $\mathcal{O}(X)$ that do not contain $\emptyset$ and are closed under binary intersections: if $U_1, U_2 \in \mathcal{F}$ then $U_1 \cap U_2 \in \mathcal{F}$. Apart from the empty set, these subsets $\mathcal{F}$ are exactly the Scott-open *filters* of open subets of X, and the Hofmann-Mislove Theorem [10, Theorem II-1.20] states that $\mathcal{F}$ must then be equal to the set of open neighborhoods of Q, where Q is some uniquely determined compact saturated subset of X; precisely, Q is the intersection $\bigcap \mathcal{F}$ of all the elements of $\mathcal{F}$. Moreover, since the empty set does not belong to $\mathcal{F}$, Q is non-empty. Mapping the Scott-open filters $\mathcal{F}$ to $\bigcap \mathcal{F}$, and the empty family to $\perp$ therefore yields an isomorphism with image $\mathcal{Q}_\perp(X)$. By composition with the

isomorphism of Lemma 4.7 (2), we obtain the isomorphism mentioned in (i) (and, by restriction, (ii) as well). □

Remark 4.11. *The isomorphism maps every superlinear discrete prevision F to $\bot$ if F is identically 0, and to $\bigcap\{U \in \mathcal{O}(X) \mid F(\chi_U) = 1\} \in \mathcal{Q}(X)$ otherwise. Its inverse maps $\bot$ to the constant 0 prevision, and every $Q \in \mathcal{Q}(X)$ to the prevision mapping $h \in [X \to \mathbb{I}]$ to $\min_{x \in Q} h(x)$. The latter requires some proof. We first realize that $\min_{x \in Q} h(x)$ is attained: indeed the image $h[Q]$ of Q by h is compact, so $\uparrow h[Q]$ is compact saturated in $\mathbb{I}$; but the compact saturated subsets of $\mathbb{I}$ are the closed intervals $[a, 1]$, so there is an $x \in Q$ such that $h(x) = a = \min_{x \in Q} h(x)$. Then, the formula of Lemma 4.4 yields that h should be mapped to $\sup\{t \in \mathbb{I} \mid Q \subseteq h^{-1}(\mathord{\uparrow} t)\}$. For every t, $Q \subseteq h^{-1}(\mathord{\uparrow} t)$ if and only if $t \ll \min_{x \in Q} h(x)$, and taking sups over $t \in \mathbb{I}$, we obtain that $\sup\{t \in \mathbb{I} \mid Q \subseteq h^{-1}(\mathord{\uparrow} t)\} = \min_{x \in Q} h(x)$.*

A similar treatment of the standard model of erratic non-deterministic choice—the *Plotkin powerdomain* $\mathcal{P}\ell(X)$—is possible as well. The easiest route is through Heckmann's *A-valuations* [18]. Let $\mathbf{A}$ be the poset $\{0, \mathsf{M}, 1\}$, ordered by $0 < \mathsf{M} < 1$. A *sub-A-valuation* on a space X is a strict Scott-continuous map $\alpha \colon \mathcal{O}(X) \to \mathbf{A}$ such that, for all $U, V \in \mathcal{O}(X)$, if $\alpha(V) = 0$ then $\alpha(U \cup V) = \alpha(U)$, and if $\alpha(U) = 1$ then $\alpha(U \cap V) = \alpha(V)$. An *A-valuation* is a sub-A-valuation such that $\alpha(X) = 1$.

Sub-A-valuations look a lot like Heckmann pairs (Lemma 4.7 3), and indeed every Heckmann pair (Φ^-, Φ^+) yields a sub-A-valuation α, defined by $\alpha(U) = 0$ if $\Phi^+(\chi_U) = 0$, $\alpha(U) = \mathsf{M}$ if $\Phi^-(\chi_U) = 0$ and $\Phi^+(\chi_U) = 1$, and $\alpha(U) = 1$ if $\Phi^-(\chi_U) = 1$. Conversely, every sub-A-valuation α yields a Heckmann pair (Φ^-, Φ^+), where $\Phi^-(\chi_U) = 0$ if $\alpha(U) = 0$, $\Phi^-(\chi_U) = 1$ otherwise, and $\Phi^+(\chi_U) = 1$ if $\alpha(U) = 1$, $\Phi^+(\chi_U) = 0$ otherwise. This defines an isomorphism of posets, once sub-A-valuations are ordered pointwise, which restricts to an isomorphism between normalized Heckmann pairs (i.e., those such that $\Phi^-(\mathbf{1}) = 1$) and A-valuations.

Proposition 4.12. *Let X be a topological space. (i) $\mathbb{P}_{\mathtt{AD}}(X)$ and the poset of sub-A-valuations on X are isomorphic. (ii) $\mathbb{P}^1_{\mathtt{AD}}(X)$ and the poset of A-valuations on X are isomorphic.*

One sees that the poset of sub-A-valuations on X is also isomorphic to the poset of A-valuations on $X_\bot$, by similar arguments as in the Hoare and Smyth cases. The relevance to the Plotkin powerdomain $\mathcal{P}\ell(X)$, which is

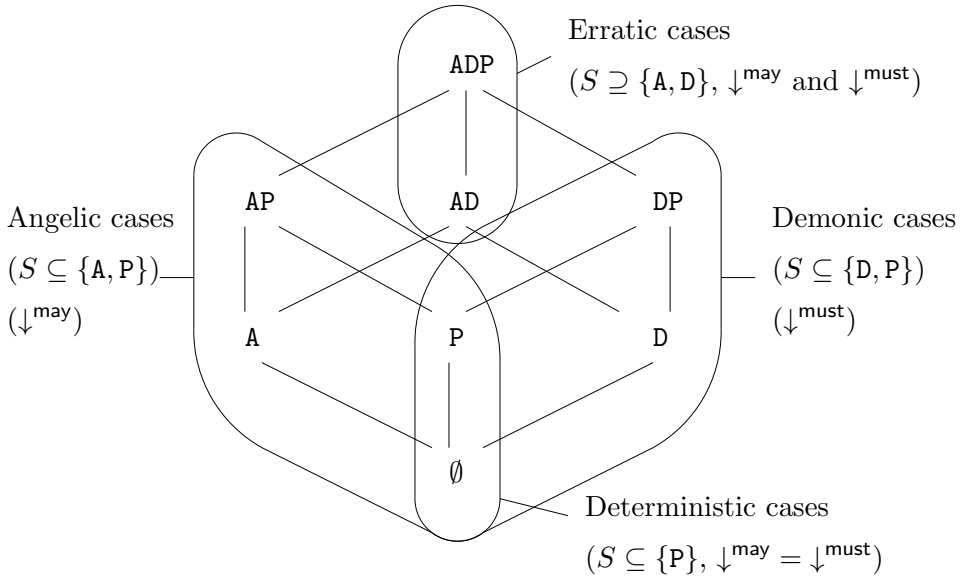

Figure 2: The Eight Languages PCF_S, $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$

defined rather differently [2, Section 6.2.1], is that the poset of A-valuations on X is isomorphic to it in common cases: when X is a continuous dcpo [18, Corollary 6.2], a Hausdorff space [18, Theorem 5.1], or a stably compact space [13, Proposition 5.3].

5. PCF with Choice

Let us proceed to the definition of our languages PCF_S, where S is one among the seven non-empty subsets of $\{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$. Whatever S, PCF_S will be an extension of PCF [38], a simply-typed λ-calculus with support for primitive arithmetic and recursion.

5.1. Syntax

The *ground types* are:

$$\gamma ::= \mathtt{Nat} \mid \mathtt{Unit}.$$

The standard presentation of PCF only uses `Nat`. We find it practical to use the *unit type* `Unit` as well. Intuitively, while the values of `Nat` are the natural numbers, `Unit` has just one value $*$, representing termination. In

usual presentations of PCF, `Unit` is usually emulated by `Nat`, representing $*$ as 0, for instance.

The *general types* are:

$$\sigma, \tau ::= \gamma \mid \sigma \to \tau$$

There is no dedicated type for choice sets, as one would require if we used the syntax of Moggi's monadic metalanguage [31]. Instead, all types will be thought as choice sets of actual elements of the type. For example, `Nat` will be interpreted denotationally as $\mathbb{P}_S(\mathbb{N})$, where $\mathbb{N}$ is the dcpo of all natural numbers, ordered by equality.

The terms of PCF_S are those of PCF, plus binary choice constructs $\mathbin{⩒}$ and $\oplus$. We also include a `let` construct that allows one to force a sequential order of evaluation. In particular, while PCF_S is a call-by-name language, just like PCF, it also includes some features of call-by-value: notably, one can encode the call-by-value application of M to N as $\mathtt{let}\, x \Leftarrow N\, \mathtt{in}\, Mx$.

Just like PCF, PCF_S is a typed language. We write $M : \tau$ to abbreviate that M is a term of type τ. Explicitly, the terms of PCF_S are inductively defined by the rules:

- There are countably many variables x_τ of type τ, for each type τ. We shall usually abbreviate x_τ as x when the type is clear or irrelevant, as in, e.g., $\lambda x_\tau \,.\, x$.
- For every natural number n, $\underline{n}$ is a term of type `Nat`.
- $\underline{*}$ is a term of type `Unit`.
- For every term $M : \sigma \to \tau$ and every term $N : \sigma$, the application MN is a term of type τ.
- For every term $M : \tau$, for every variable x_σ, $\lambda x_\sigma \,.\, M$ is a term of type $\sigma \to \tau$.
- For every type τ, for every term $M : \tau \to \tau$, $\mathtt{Y}M$ is a term of type τ.
- For every $M :$ `Nat`, $\mathtt{succ}\, M$ and $\mathtt{pred}\, M$ are terms of type `Nat`.
- For every type τ, for all terms $M :$ `Nat`, $N : \tau$, $P : \tau$, $\mathtt{ifz}\ M\ N\ P$ is a term of type τ.

- For all terms $M : \sigma$ and $N : \tau$, $\mathtt{let}\, x_\sigma \Leftarrow M\ \mathtt{in}\ N$ is a term of type τ.
- In PCF_S where S contains at least one of `A` or `D` (non-deterministic choice), for every type τ, for all terms $M : \tau$ and $N : \tau$, $M \otimes N$ is a term of type τ.
- In PCF_S where S contains `P`, for every type τ, for all terms $M : \tau$ and $N : \tau$, $M \oplus N$ is a term of type τ.

The variable x is bound in $\lambda x\,.\,M$ and in $\mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ N$; in the latter case, its scope is the term N. We take the usual conventions on α-renaming of bound variables.

Among these, the *values* V are $\underline{*}$, $\underline{n}$, and expressions of the form $\lambda x_\sigma\,.\,M$.

5.2. Operational Semantics

We define the *operational semantics* of PCF_S using a context machine, similar for example to the Krivine abstract machine [5]. A distinctive aspect of such machines is that they run by rewriting configurations to other configurations by elementary steps, where a *configuration* is a pair $C \cdot M$ of an *evaluation context* C and of a term M. The evaluation contexts C are certain terms with a distinguished unique occurrence of a placeholder $\_$, called the *hole*. The term $C[M]$ is obtained by replacing the hole by the term M, and a configuration $C \cdot M$ encodes the fact that we are attempting to find the value of $C[M]$, together with the fact that the current focus is on the subterm M.

Evaluation contexts are typed, too, and we shall reserve the notation $C : \sigma \vdash \tau$ to state that C has type τ provided we assume the hole $\_$ has type σ.

For reasons of proof, it is more profitable to define evaluation contexts as finite sequences of *elementary* evaluation contexts, which only apply one PCF_S construct to the hole (and possibly other terms).

Definition 5.1 (Evaluation Context). *An* elementary evaluation context *is any formal expression of one of the following forms:*

- $[\_N]$*, of type* $(\sigma \to \tau) \vdash \tau$*, for any term* $N : \sigma$*, and all types* σ*,* τ*;*
- $[\mathtt{succ}\ \_]$ *and* $[\mathtt{pred}\ \_]$*, of type* $\mathtt{Nat} \vdash \mathtt{Nat}$*;*
- $[\mathtt{ifz}\ \_\ N\ P]$*, of type* $\mathtt{Nat} \vdash \tau$*, for all terms* $N, P : \tau$*, and every type* τ*;*
- $[\mathtt{let}\, x_\sigma \Leftarrow \_\ \mathtt{in}\ N]$*, of type* $\sigma \vdash \tau$*, for every term* $N : \mathtt{Nat}$*;*

$$C \cdot MN \to C[\_N] \cdot M \qquad C \cdot \mathtt{pred}\, N \to C[\mathtt{pred}\, \_] \cdot N$$
$$C \cdot \mathtt{succ}\, N \to C[\mathtt{succ}\, \_] \cdot N \quad C \cdot \mathtt{ifz}\ M\ N\ P \to C[\mathtt{ifz}\ \_\ N\ P] \cdot M$$
$$C \cdot \mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ N \to C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ N] \cdot M$$

Figure 3: Redex Discovery Rules in PCF_S

$$C[\_N] \cdot \lambda x\,.\,P \to C \cdot P[x := N] \qquad C \cdot \mathtt{Y}N \to C \cdot N(\mathtt{Y}N)$$
$$C[\mathtt{pred}\, \_] \cdot \underline{n+1} \to C \cdot \underline{n} \qquad C[\mathtt{succ}\, \_] \cdot \underline{n} \to C \cdot \underline{n+1}$$
$$C[\mathtt{ifz}\ \_\ N\ P] \cdot \underline{0} \to C \cdot N \qquad C[\mathtt{ifz}\ \_\ N\ P] \cdot \underline{n+1} \to C \cdot P$$
$$C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ N] \cdot V \to C \cdot N[x := V] \qquad (V \text{ a value})$$

Figure 4: Computation Rules in PCF_S

An evaluation context C *of type* $\sigma \vdash \tau$ *is any finite sequence* $E_n E_{n-1} \ldots E_2 E_1$ *of elementary evaluation contexts* $E_1 : \sigma_0 \vdash \sigma_1$, $E_2 : \sigma_1 \vdash \sigma_2$, $\ldots$, $E_{n-1} : \sigma_{n-2} \vdash \sigma_{n-1}$, $E_n : \sigma_{n-1} \vdash \sigma_n$, *where* $\sigma_0 = \sigma$ *and* $\sigma_n = \tau$. *(When* $n = 0$, *we write* $\_$ *for the* empty *evaluation context, which has all types of the form* $\sigma \vdash \sigma$.*)*

A configuration *is a pair* $C \cdot M$ *of an evaluation context* C *of type* $\tau \vdash \mathtt{Unit}$ *and of a term* $M : \tau$.

Given any elementary evaluation context E_1 of type $\sigma \to \tau$, and any term $M : \sigma$, the result $E_1[M]$ of replacing the hole $\_$ in E_1 by M yields a term of type τ. We extend this to arbitrary evaluation contexts $C = E_n E_{n-1} \ldots E_2 E_1$, by defining $C[M]$ as $E_n[E_{n-1}[\ldots E_2[E_1[M]]\ldots]]$.

One should note that our machine will only work on configurations $C \cdot M$, such that $C[M]$ has type $\mathtt{Unit}$. One might as well replace the latter by any type in the definition of the machine itself, but it will be important that it be a ground type, when we attack questions of computational adequacy and of full abstraction.

The rules of the PCF_S machine will comprise four kinds of rules. We first find the *redex discovery rules* of Figure 3, whose rule is to shift the focus until one finds a place where computation can happen. Deterministic computation occurs through the *computation rules* of Figure 4, which contract a redex under focus.

As far as choice ($\Cup$, $\oplus$) is concerned, rules such that $M \Cup N \to M$,

$$\frac{C' \cdot M' \downarrow^m a}{C \cdot M \downarrow^m a} \text{ if } C \cdot M \to C' \cdot M' \qquad \frac{}{\_ \cdot \underline{*} \downarrow^m a} \ (a \in [0,1)) \qquad \frac{}{C \cdot M \downarrow^m 0}$$

$$\frac{C \cdot M \downarrow^{\mathsf{may}} a}{C \cdot M \mathbin{\textcircled{\scriptsize \&}} N \downarrow^{\mathsf{may}} a} \qquad \frac{C \cdot N \downarrow^{\mathsf{may}} a}{C \cdot M \mathbin{\textcircled{\scriptsize \&}} N \downarrow^{\mathsf{may}} a} \ \text{(both only if } \mathtt{A} \in S)$$

$$\frac{C \cdot M \downarrow^{\mathsf{must}} a \quad C \cdot N \downarrow^{\mathsf{must}} a}{C \cdot M \mathbin{\textcircled{\scriptsize \&}} N \downarrow^{\mathsf{must}} a} \text{ (only if } \mathtt{D} \in S) \qquad \frac{C \cdot M \downarrow^m a \quad C \cdot N \downarrow^m b}{C \cdot M \oplus N \downarrow^m \frac{1}{2}(a+b)} \text{ (only if } \mathtt{P} \in S)$$

Figure 5: Termination Semantics for PCF_S

$M \mathbin{\textcircled{\scriptsize \&}} N \to N$ would be in order, but using termination judgments in the style of Pitts and Stark [37] will prove to be more convenient. In such a semantics, we give rules to derive judgments of the form $C \cdot M{\downarrow}$ stating that M terminates when evaluated in evaluation context C. In the presence of non-determinism, this judgment must be split in two: $C \cdot M{\downarrow}^{\mathsf{may}}$ states that M *may* terminate when evaluated in evaluation context C, i.e., that some execution of $C \cdot M$ eventually terminates, and $C \cdot M{\downarrow}^{\mathsf{must}}$ states that M *must* terminate when evaluated in evaluation context C, i.e., that all possible executions of $C \cdot M$ do terminate.

In the presence of probabilistic choice, it is reasonable to ask whether $C \cdot M$ terminates with probability greater than some fixed number a. By "greater" we mean way-greater, the opposite of the way-below relation on $\mathbb{I}$, which will be mathematically convenient. We write this judgment $C \cdot M \downarrow a$, and recall that, due to the way $\ll$ behaves on $\mathbb{I}$, this means: the probability that $C \cdot M$ terminates is strictly greater than a, when $a > 0$; the probability that $C \cdot M$ terminates is greater than or equal to 0, when $a = 0$. The latter statement looks like a triviality, yet we need a specific rule to ensure that it holds for all configurations, including non-terminating configurations: see the top right rule in Figure 5. For reasons of implementability, a should be taken rational. It should also be taken in $[0, 1)$, excluding 1, because a probability cannot be way-greater than 1.

In the presence of both forms of choice, we use judgments of the form $C \cdot M \downarrow^{\mathsf{may}} a$ ("there is a strategy for resolving all non-deterministic choices

to come, such that $C \cdot M$ terminates with probability way-greater than a, when evaluated under this strategy") and $C \cdot M \downarrow^{\mathsf{must}} a$ ("for every strategy, $C \cdot M$ will necessarily terminate with probability way-greater than a, when evaluated under this strategy"). In general, our judgments will be of the form $C \cdot M \downarrow^m a$ where the *mode* m is either may or must, and a is a rational number in $[0, 1)$: see Figure 5. Note that, depending on the subscript S in the name of the language (PCF_S) that we consider, some of the rules may be absent. E.g., the two middle rules are only present in $\mathrm{PCF}_{\mathtt{A}}$, $\mathrm{PCF}_{\mathtt{AD}}$, $\mathrm{PCF}_{\mathtt{AP}}$ and $\mathrm{PCF}_{\mathtt{ADP}}$.

The last four rules implement choice. The top left rule incorporates all our deterministic rules from Figure 3 and Figure 4. For example, it implies the following derivation:

$$\frac{C \cdot P[x := N] \downarrow^m a}{C[\_N] \cdot \lambda x \,.\, P \downarrow^m a}$$

which states that to ensure that $\lambda x \,.\, P$ terminates (may, must; with probability larger than a) when applied to N in evaluation context C, it suffices to show that $P[x := N]$ terminates in evaluation context C.

The middle top rule of Figure 5 is the *final state* rule: it states that any configuration of the form $\_ \cdot \underline{*}$ is final, i.e., terminates (may, must) with probability exactly 1. Since the type of any evaluation context C appearing in a configuration is of the form $\tau \vdash \mathtt{Unit}$, $\_ \cdot \underline{*}$ is the only configuration that we can sensibly call final.

Given a configuration $C \cdot M$, the set of numbers $a \in [0, 1)$ such that $C \cdot M \downarrow^m a$ is derivable:

- is non-empty: 0 is in it, by the top right rule of Figure 5;
- is downward closed: for $a, b \in [0, 1)$, if $C \cdot M \downarrow^m a$ is derivable, and $b \leq a$, then $C \cdot M \downarrow^m b$ is derivable, too;
- and is such that if $a > 0$ is in it, then there is an even larger $b > a$ in it: for $a \in (0, 1)$ such that $C \cdot M \downarrow^m a$ is derivable, there is a $b \in (a, 1)$ such that $C \cdot M \downarrow^m b$ is derivable, too.

The latter two properties are proved by induction on the derivation. These properties imply that the set of values $a \in [0, 1)$ such that $C \cdot M \downarrow^m a$ is derivable is an interval, and one that is entirely described by their supremum. Temporarily calling this supremum A, this interval is $\{0\}$ if $A = 0$, and $[0, A)$ otherwise. We reserve the following notation for this supremum.

Definition 5.2. *Given any PCF$_S$ term $M : \tau$, and any evaluation context C of type $\tau \vdash \mathtt{Unit}$, let $Pr(C \cdot M{\downarrow}^m)$ be* $\sup\{a \in \mathbb{Q} \in [0,1) \mid C \cdot M \downarrow^m a$ *is derivable*$\}$*, where sups are taken in* $[0,1]$.

Proposition 5.3. *The following equalities hold:*

$$\begin{array}{rcl} Pr(C \cdot M \oplus N{\downarrow}^m) & = & \frac{1}{2}(Pr(C \cdot M{\downarrow}^m) + Pr(C \cdot N{\downarrow}^m)) \\ Pr(C \cdot M \mathbin{\textcircled{\scriptsize \&}} N{\downarrow}^{\mathsf{may}}) & = & \max(Pr(C \cdot M{\downarrow}^{\mathsf{may}}), Pr(C \cdot N{\downarrow}^{\mathsf{may}})) \\ Pr(C \cdot M \mathbin{\textcircled{\scriptsize \&}} N{\downarrow}^{\mathsf{must}}) & = & \min(Pr(C \cdot M{\downarrow}^{\mathsf{must}}), Pr(C \cdot N{\downarrow}^{\mathsf{must}})). \end{array}$$

Proof. Define $\Downarrow A$ as $[0, A)$ when $A \in (0,1]$, $\{0\}$ when $A = 0$. We have seen that $\Downarrow \Pr(C \cdot M{\downarrow}^m)$ is the set of values $a \in [0,1)$ such that $C \cdot M \downarrow^m a$ is derivable.

Let $A = \Pr(C \cdot M{\downarrow}^m)$, $B = \Pr(C \cdot N{\downarrow}^m)$. The set of numbers of the form $\frac{1}{2}(a+b)$, $a \in \Downarrow A$, $b \in \Downarrow B$, is equal to $\Downarrow \frac{1}{2}(A+B)$. Using the bottom right rule of Figure 5, it follows that $\Downarrow \Pr(C \cdot M \oplus N{\downarrow}^m) = \Downarrow \frac{1}{2}(A+B)$, whence $\Pr(C \cdot M \oplus N{\downarrow}^m) = \frac{1}{2}(A+B)$. The other two equalities are proved by noting that $\Downarrow A \cup \Downarrow B = \Downarrow \max(A,B)$ and $\Downarrow A \cap \Downarrow B = \Downarrow \min(A,B)$. □

5.3. Denotational Semantics

We define the semantics of type τ in PCF$_S$ as dcpos $[\![\tau]\!]_S$, as follows. First, for every type τ, $[\![\tau]\!]_S$ is defined as $\mathbb{P}_S([\![\tau]\!]_S^\circ)$, so that every type is interpreted as a space of previsions (or forks). Then we let:

$$\begin{array}{rcl} [\![\mathtt{Nat}]\!]_S^\circ & = & \mathbb{N} \\ [\![\mathtt{Unit}]\!]_S^\circ & = & \{*\} \\ [\![\sigma \to \tau]\!]_S^\circ & = & [[\![\sigma]\!]_S \to [\![\tau]\!]_S] \end{array}$$

where $\mathbb{N}$ is considered as a dcpo with ordering taken as equality. Notice that $[\![\mathtt{Nat}]\!]_S$ is $\mathbb{P}_S(\mathbb{N})$: we incorporate the possibility of making non-deterministic, resp. probabilistic choices on natural numbers, through the recourse to $\mathbb{P}_S$.

To distinguish the function that maps values v in a domain D to values $e(v)$ from the syntax $\lambda v\,.\,e(v)$, write such functions as $(v \in D \mapsto e(v))$. For each non-empty $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$, $\mathbb{P}_S$ defines the functor part of a monad (more on this in Remark 5.4). Using notations inspired from Moggi [31], define:

$$\begin{array}{rcll} val_S\, a & = & (h \in [X \to \mathbb{I}] \mapsto h(a)) & (a \in X) \\ let_S\, v \Leftarrow F \text{ in } G(v) & = & (h \in [Y \to \mathbb{I}] \mapsto F(v \in X \mapsto G(v)(h))) & \end{array}$$

whenever $S \not\supseteq \{\mathtt{A}, \mathtt{D}\}$. One easily checks that $val_S\, a$ is in $\mathbb{P}_S(X)$ for every $a \in X$, and that $val_S \colon X \to \mathbb{P}_S(X)$ is Scott-continuous. Also, for every $F \in \mathbb{P}_S(X)$, and for every continuous map G from X to $\mathbb{P}_S(Y)$, $let_S\, v \Leftarrow F$ *in* $G(v)$ is an element of $\mathbb{P}_S(Y)$. This was shown in the non-discrete cases (namely, $\mathtt{P} \in S$) in [11]. Clearly, $let_S\, v \Leftarrow F$ *in* $G(v)$ is a discrete prevision when F is discrete and when $G(v)$ is discrete for every $v \in X$, which covers the remaining cases. Finally, $let_S\, v \Leftarrow F$ *in* $G(v)$ is Scott-continuous in F and G, which will be required to make sure that the denotation of every term, evaluation context, and value, is continuous, and therefore that the denotation of $\mathtt{Y}N$ makes sense.

When $S \supseteq \{\mathtt{A}, \mathtt{D}\}$, where elements of $\mathbb{P}_S(X)$ are forks, not previsions, the analogous quantities are computed componentwise. I.e., letting $S^- = S \cap \{\mathtt{D}, \mathtt{P}\}$ and $S^+ = S \cap \{\mathtt{A}, \mathtt{P}\}$, we let:

$$\begin{array}{rcl} val_S\, a & = & (val_{S^-}\, a, val_{S^+}\, a) \\ let_S\, v \Leftarrow F \text{ in } G(v) & = & (let_{S^-}\, v \Leftarrow F^- \text{ in } G^-(v), let_{S^+}\, v \Leftarrow F^+ \text{ in } G^+(v)) \\ & & \text{where } F = (F^-, F^+), G(v) = (G^-(v), G^+(v)) \end{array}$$

Remark 5.4. *Declaring that $\mathbb{P}_S$ is the functor part of a monad only makes sense once we have defined the ambient category. In this paper, this will always be the category of dcpos and Scott-continuous maps, and we shall always think of $\mathbb{P}_S(X)$ as a* dcpo *of previsions.*

Remark 5.5. *The monads $\mathbb{P}_S$ all have (order-preserving and) order-reflecting units. In other words, $val_S\, a \leq val_S\, b$ if and only if $a \leq b$. Indeed, $val_S\, a \leq val_S\, b$ implies that $\chi_U(a) \leq \chi_U(b)$ for every open subset U, i.e., that a is below b in the specialization preorder. In particular, on T_0 spaces (dcpos, for example), val_S is injective, and therefore an order embedding.*

Remark 5.6. *The monads $\mathbb{P}_S$ are all strong. In the non-erratic cases $S \not\supseteq \{\mathtt{A}, \mathtt{D}\}$, the so-called tensorial strength $t_{X,Y} \colon \mathbb{P}_S(X) \times Y \to \mathbb{P}_S(X \times Y)$ maps (F, w) to $let_S\, v \Leftarrow F$ in $val_S(v, w)$, i.e., to $(h \in [X \times Y \to \mathbb{I}] \mapsto F(v \in X \mapsto h(v, w)))$. (In the erratic cases $S \supseteq \{\mathtt{A}, \mathtt{D}\}$, it maps $((F^-, F^+), w)$ to $(let_S\, v \Leftarrow F^-$ in $val_S(v, w), let_S\, v \Leftarrow F^+$ in $val_S(v, w))$.) The fact that we are working in a category of dcpos, not general topological spaces,is important here: the formula $(h \in [X \times Y \to \mathbb{I}] \mapsto F(v \in X \mapsto h(v, w)))$ is separately continuous in F and w, and joint continuity follows because joint and separate continuity coincide on dcpos.*

Remark 5.7. *The monads $\mathbb{P}_{\mathtt{A}}$, $\mathbb{P}_{\mathtt{D}}$, $\mathbb{P}_{\mathtt{AD}}$ and $\mathbb{P}_{\mathtt{P}}$ are not just strong but even commutative, meaning that for $S = \mathtt{A}$, $S = \mathtt{D}$, $S = \mathtt{AD}$ or $S = \mathtt{P}$, the equation:*

$$\begin{aligned} & let_S\, v \Leftarrow F \text{ in } (let_S\, w \Leftarrow G \text{ in } val_S(v,w)) \qquad (1)\\ = \;& let_S\, w \Leftarrow G \text{ in } (let_S\, v \Leftarrow F \text{ in } val_S(v,w)) \end{aligned}$$

holds. (Equivalently, $let_S\, v \Leftarrow F$ in $(let_S\, w \Leftarrow G$ in $f(v,w)) = let_S\, w \Leftarrow G$ in $(let_S\, v \Leftarrow F$ in $f(v,w))$ for every continuous map f.) Expanding the definitions in the non-erratic cases $S \not\supseteq \{\mathtt{A}, \mathtt{D}\}$, $\mathbb{P}_S$ is commutative if and only if, for all dcpos X, Y, for all $F \in \mathbb{P}_S(X)$, $G \in \mathbb{P}_S(Y)$, $h \in [X \times Y \to \mathbb{I}]$,

$$\begin{aligned} & F(v \in X \mapsto G(w \in Y \mapsto h(v,w))) \qquad (2)\\ = \;& G(w \in Y \mapsto F(v \in X \mapsto h(v,w))). \end{aligned}$$

For $\mathbb{P}_{\mathtt{P}}$, this is a form of Fubini's Theorem, which one can obtain by carrying Jones's version for continuous valuations [20, Theorem 3.17] along the isomorphism between $\mathbb{P}_{\mathtt{P}}(X)$ and $\mathbf{V}_{\leq 1}(X)$.

For $\mathbb{P}_{\mathtt{A}}$, we use Proposition 4.8 and the fact that the Hoare powerdomain monad is commutative; using Remark 4.9, the essence of the proof is the fact that, given two closed sets C, C', $\sup_{x \in C} \sup_{x' \in C'} h(x,x') = \sup_{x' \in C'} \sup_{x \in C} h(x,x')$.

For $\mathbb{P}_{\mathtt{D}}$, a similar argument using Proposition 4.10 works provided that X and Y are sober topological spaces, and the proof boils down to the fact that given two compact saturated subsets Q, Q', $\min_{x \in Q} \min_{x' \in Q'} h(x,x') = \min_{x' \in Q'} \min_{x \in Q} h(x,x')$. We deal with the general case, establishing (2) when X and Y are general topological spaces, as follows. Embed X into its sobrification *X^s (see [10, Exercise V-4.9] or [14, Section 8.2]). Every prevision $F \in \mathbb{P}_S(X)$ extends to a prevision $\hat{F} \in \mathbb{P}_S(X^s)$ by $\hat{F}(h) = F(h_{|X})$, where $h_{|X}$ is the restriction of $h \in [X^s \to \mathbb{I}]$ to X. Note in particular that $\hat{F}$ is in $\mathbb{P}_S(X^s)$: if F is discrete, resp. sublinear, resp. superlinear, then so is $\hat{F}$. Since X^s and Y^s are sober, we have seen that $\hat{F}(v \in X^s \mapsto \hat{G}(w \in Y^s \mapsto h'(v,w))) = \hat{G}(w \in Y^s \mapsto \hat{F}(v \in X^s \mapsto h'(v,w)))$ for every $h' \in [X^s \times Y^s \to \mathbb{I}]$. For every $h \in [X \times Y \to \mathbb{I}]$, h extends to a unique continuous map h' from $(X \times Y)^s \to \mathbb{I}$ [14, Theorem 8.2.44], and noting that $(X \times Y)^s = X^s \times Y^s$ up to natural isomorphism [14, Theorem 8.4.8], the latter equality specializes to (2).*

Showing that $\mathbb{P}_{\mathtt{AD}}$ is commutative reduces to the cases of $\mathbb{P}_{\mathtt{A}}$ and $\mathbb{P}_{\mathtt{D}}$.

The monads $\mathbb{P}_{\mathtt{AP}}$, $\mathbb{P}_{\mathtt{DP}}$ and $\mathbb{P}_{\mathtt{ADP}}$ are not *commutative, and this reflects the fact that non-deterministic and probabilistic choices do not commute. For a*

$$[\![\underline{*}]\!]^\circ_S\,\rho = * \qquad [\![\underline{n}]\!]^\circ_S\,\rho = n \qquad [\![\lambda x_\sigma\,.\,M]\!]^\circ_S\,\rho = (G \in [\![\sigma]\!]_S \mapsto [\![M]\!]_S\,(\rho[x := G]))$$

Figure 6: Denotational Semantics: Values

$$\begin{array}{rcl}
[\![x]\!]_S\,\rho = \rho(x) \qquad [\![V]\!]_S\,\rho = val_S\,[\![V]\!]^\circ_S\,\rho \quad (V \text{ a value}) \\
[\![MN]\!]_S\,\rho & = & let_S\,f \Leftarrow [\![M]\!]_S\,\rho\ in\ f([\![N]\!]_S\,\rho) \\
[\![\mathtt{Y}N]\!]_S\,\rho & = & \sup_{n\in\mathbb{N}} f^n(0) \\
 & & \text{where } f(F) = let_S\,g \Leftarrow [\![N]\!]_S\,\rho\ in\ g(F) \\
[\![\mathtt{pred}\,M]\!]_S\,\rho & = & let_S\,n \Leftarrow [\![M]\!]_S\,\rho\ in \begin{cases} val_S\,(n-1) & \text{if } n \neq 0 \\ \bot & \text{if } n = 0 \end{cases} \\
[\![\mathtt{succ}\,M]\!]_S\,\rho & = & let_S\,n \Leftarrow [\![M]\!]_S\,\rho\ in\ val_S\,(n+1) \\
[\![\mathtt{ifz}\ M\ N\ P]\!]_S\,\rho & = & let_S\,m \Leftarrow [\![M]\!]_S\,\rho\ in \begin{cases} [\![N]\!]_S\,\rho & \text{if } m = 0 \\ [\![P]\!]_S\,\rho & \text{if } m \neq 0 \end{cases} \\
[\![\mathtt{let}\,x \Leftarrow M\ \mathtt{in}\ N]\!]_S\,\rho & = & let_S\,v \Leftarrow [\![M]\!]_S\,\rho\ in\ [\![N]\!]_S\,(\rho[x := val_S\,v]) \\
[\![M \mathbin{\textcircled{\scriptsize\&}} N]\!]_S\,\rho & = & (h \mapsto \max([\![M]\!]_S\,\rho\,(h), [\![N]\!]_S\,\rho\,(h)))(\text{if } \mathtt{A} \in S, \mathtt{D} \notin S) \\
[\![M \mathbin{\textcircled{\scriptsize\&}} N]\!]_S\,\rho & = & (h \mapsto \min([\![M]\!]_S\,\rho\,(h), [\![N]\!]_S\,\rho\,(h)))\ (\text{if } \mathtt{D} \in S, \mathtt{A} \notin S) \\
[\![M \mathbin{\textcircled{\scriptsize\&}} N]\!]_S\,\rho & = & (h \mapsto \min(F^-(h), G^-(h)), \qquad (\text{if } \mathtt{A} \in S, \mathtt{D} \in S) \\
 & & \ h \mapsto \max(F^+(h), G^+(h))) \\
 & & \text{where } [\![M]\!]_S\,\rho = (F^-, F^+),\ [\![N]\!]_S\,\rho = (G^-, G^+) \\
[\![M \oplus N]\!]_S\,\rho & = & (h \mapsto \tfrac{1}{2}([\![M]\!]_S\,\rho\,(h) + [\![N]\!]_S\,\rho\,(h))) \qquad (\text{if } \mathtt{P} \in S)
\end{array}$$

Figure 7: Denotational Semantics of PCF_S

counter-example, let $X = Y = \mathbb{B}_\bot$, the dcpo containing three elements tt, ff *and $\bot$, with* tt *and* ff *incomparable and above $\bot$. For $h \in [\mathbb{B}_\bot \to \mathbb{I}]$, let $F(h) = \max(h(\mathtt{tt}), h(\mathtt{ff}))$ (in the angelic case; $\min(h(\mathtt{tt}), h(\mathtt{ff}))$ in the demonic case), $G(h) = \frac{1}{2}(h(\mathtt{tt}) + h(\mathtt{ff}))$. Consider now $h \in [X \times Y \to \mathbb{I}]$ defined as the "equals to" function: $h(v, \bot) = h(\bot, w) = \bot$, while for $v, w \neq \bot$, $h(v, w)$ equals 1 if $v = w$ and 0 otherwise. Then $F(v \in X \mapsto G(w \in Y \mapsto h(v, w))) = \max(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}$ (in the angelic case; $\min(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}$ in the demonic case), while $G(w \in Y \mapsto F(v \in X \mapsto h(v, w))) = \frac{1}{2}(1 + 1) = 1$ (in the angelic case, $\frac{1}{2}(0 + 0) = 0$ in the demonic case). As usual, the erratic case reduces to considering both the angelic and demonic cases together.*

The denotational semantics $[\![M]\!]_S\,\rho$ of a PCF_S term $M : \tau$ in the *environment* ρ, which maps variables x_σ of type σ to values in $[\![\sigma]\!]_S$, is defined

in Figure 7. In the non-erratic cases $S \not\supseteq \{\mathtt{A}, \mathtt{D}\}$, $[\![M]\!]_S\,\rho$ is a prevision, so it makes sense to apply it to a function h, yielding a non-negative real number that we write $[\![M]\!]_S\,\rho\ (h)$. The notation $([\![M]\!]_S\,\rho)(h)$ would have had the advantage of making clearer what is the function and what is its argument, but would have been cumbersome.

We use an auxiliary definition of $[\![V]\!]^\circ_S\,\rho$ for values $V : \tau$ as well, see Figure 6. The environment $\rho[x := V]$ is the map that takes x to V and all other variables $y \neq x$ to $\rho(y)$. The $\bot$ symbol in the definition of $[\![\mathtt{pred}\,M]\!]_S$ denotes the bottom element of $\mathbb{P}_S(\mathbb{N})$: recall that $\mathbb{P}_S(X)$ is always pointed; pointedness is the reason we have chosen to work with subnormalized, not normalized previsions. On the third line of Figure 7, 0 denotes the constant 0 prevision, or the fork $(0, 0)$, depending on S. We also agree formulae such as $\frac{1}{2}(A + B)$ are interpreted componentwise when $A = (A^-, A^+)$ and $B = (B^-, B^+)$ are pairs; this serves to interpret the last line in the case $S = \mathtt{ADP}$.

It is easy to check that $[\![M]\!]_S\,\rho$ is in $[\![\tau]\!]_S$ for every $M : \tau$, and that $[\![V]\!]^\circ_S\,\rho$ is in $[\![\tau]\!]^\circ$ for every value $V : \tau$, whatever $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$.

6. Soundness

In PCF, soundness states that if M evaluates to a normal form V (in the operational semantics), then M and V have the same semantics. This translates to the following in our setting, where we write $\blacklozenge$ for the constant 1 map from $\{*\}$ to $\mathbb{I}$. For any term $P : \mathtt{Unit}$, the application of the functional $[\![P]\!]_S\,\rho$ to $\blacklozenge$, which we shall write $[\![P]\!]_S\,\rho\,\blacklozenge$, is to be understood as the probability that P terminates (if $S = \mathtt{P}$), the least probability that P terminates under all possible strategies (if $S = \mathtt{DP}$ or $S = \mathtt{D}$), or the supremum of all probabilities that P terminates (if $S = \mathtt{AP}$ or $S = \mathtt{A}$).

In our case, we claim that soundness is the following fact, which states that the denotational probability of terminating is at least the operational probability of terminating. By "$[\![C[M]]\!]_S\,\rho$ makes sense", we mean that it must be defined according to Figure 7; for example, this excludes cases where the symbol $\oplus$ would occur in $C[M]$ but $\mathtt{P} \notin S$.

Lemma 6.1 (Soundness). *Let C be an evaluation context of type $\sigma \vdash \mathtt{Unit}$, M be a term of type σ, ρ be an arbitrary environment, and S be a non-empty subset of $\{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$ be such that $[\![C[M]]\!]_S\,\rho$ makes sense.*

$(S \subseteq \{\mathtt{A}, \mathtt{P}\})$ $[\![C[M]]\!]_S\,\rho\,\blacklozenge \geq Pr(C \cdot M{\downarrow^{\mathsf{may}}})$.

$(S \subseteq \{\mathtt{D}, \mathtt{P}\})$ $[\![C[M]]\!]_S\,\rho\,\blacklozenge \geq Pr(C \cdot M{\downarrow}^{\mathsf{must}})$.

$(\{\mathtt{A}, \mathtt{D}\} \subseteq S)$ *Letting* $(F^-, F^+) = [\![C[M]]\!]_S\,\rho$, *the two inequalities* $F^+(\blacklozenge) \geq Pr(C \cdot M{\downarrow}^{\mathsf{may}})$ *and* $F^-(\blacklozenge) \geq Pr(C \cdot M{\downarrow}^{\mathsf{must}})$ *hold.*

Proof. Let us first extend the notation $[\![C]\!]_S$ to the case where C is an evaluation context, not a term: let $[\![C]\!]_S\,\rho$ be the map $[\![\lambda x\,.\,C[x]]\!]^\circ_S = (G \mapsto [\![C[x]]\!]_S\,\rho[x := G])$, where x is some fresh variable. It is easy to see that: (a) $[\![C[M]]\!]_S\,\rho = [\![C]\!]_S\,\rho\,([\![M]\!]_S\,\rho)$ for any term M such that $[\![C[M]]\!]_S\,\rho$ makes sense. Also: (b) if C is the concatenation $C'E$ of an evaluation context C' with an elementary evaluation context E, then $[\![C]\!]_S\,\rho = [\![C']\!]_S\,\rho \circ [\![E]\!]_S\,\rho$.

Let us prove the Lemma. More precisely, we shall show that, for $a \in [0, 1)$:

$(S \subseteq \{\mathtt{A}, \mathtt{P}\})$ If $C \cdot M \downarrow^{\mathsf{may}} a$ is derivable, then $a \ll [\![C[M]]\!]_S\,\rho\,\blacklozenge$.

$(S \subseteq \{\mathtt{D}, \mathtt{P}\})$ If $C \cdot M \downarrow^{\mathsf{must}} a$ is derivable, then $a \ll [\![C[M]]\!]_S\,\rho\,\blacklozenge$.

$(\{\mathtt{A}, \mathtt{D}\} \subseteq S)$ If $C \cdot M \downarrow^{\mathsf{must}} a^-$ and $C \cdot M \downarrow^{\mathsf{may}} a^+$ are derivable, then $a^- \ll F^-(\blacklozenge)$ and $a^+ \ll F^+(\blacklozenge)$, where $(F^-, F^+) = [\![C[M]]\!]_S\,\rho$.

We only prove the first case. The second one is similar, and the third one is proved similarly to the conjunction of the first two. If $a = 0$, this is obvious, so we assume $a \neq 0$. It remains to show that if $C \cdot M \downarrow^{\mathsf{may}} a$ is derivable and $0 < a < 1$, then $[\![C[M]]\!]_S\,\rho\,\blacklozenge > a$. This is by structural induction on derivations.

$C \cdot M \downarrow^{\mathsf{may}} a$ cannot have been derived by the top right rule of Figure 5, since $a \neq 0$.

If $C \cdot M \downarrow^{\mathsf{may}} a$ was derived by the top left rule of Figure 5, then there is a redex discovery rule or a computation rule $C \cdot M \to C' \cdot M'$, and we have derived $C' \cdot M' \downarrow^{\mathsf{may}} a$. By induction hypothesis, $[\![C'[M']]\!]_S\,\rho\,\blacklozenge > a$. If $C \cdot M \to C' \cdot M'$ was a redex discovery rule (Figure 3), then we conclude right away that $[\![C[M]]\!]_S\,\rho\,\blacklozenge > a$, since $C[M] = C'[M']$. If this is a computation rule (Figure 4), then we must analyze each of the seven cases in turn. Let us deal with the case of $\mathtt{pred}$, where $C \cdot M$ is of the form $C_1[\mathtt{pred}\ \_] \cdot \underline{n+1}$, and $C' \cdot M'$ equals $C_1 \cdot \underline{n}$. We check that $[\![\mathtt{pred}(\underline{n+1})]\!]_S\,\rho = [\![\underline{n}]\!]_S\,\rho$. The

left-hand side is:

$$\begin{array}{rl} & [\![\mathtt{pred}(\underline{n+1})]\!]_S\,\rho \\ = & let_S\, n' \Leftarrow val_S\,(n+1)\ in \left\{ \begin{array}{ll} val_S\,(n'-1) & \text{if } n \neq 0 \\ \bot & \text{if } n = 0 \end{array} \right. \\ = & (h \in [\mathbb{N} \to \mathbb{I}] \mapsto val_S\,(n+1)(n' \mapsto \left\{ \begin{array}{ll} val_S\,(n'-1)\ (h) & \text{if } n' \neq 0 \\ 0 & \text{if } n' = 0 \end{array} \right.)) \\ = & (h \in [\mathbb{N} \to \mathbb{I}] \mapsto val_S((n+1)-1)\ (h)) \\ = & (h \in [\mathbb{N} \to \mathbb{I}] \mapsto val_S\, n\ (h)) = val_S\, n = [\![\underline{n}]\!]_S\,\rho. \end{array}$$

The induction hypothesis states that $[\![C_1[\underline{n}]]\!]_S\,\rho\ \blacklozenge > a$. By Claim (a) above, $[\![C_1[\underline{n}]]\!]_S\,\rho = [\![C_1]\!]_S\,\rho\ ([\![\underline{n}]\!]_S\,\rho)$, and this is equal to $[\![C_1]\!]_S\,\rho\ ([\![\mathtt{pred}(\underline{n+1})]\!]_S\,\rho) = [\![C_1[\mathtt{pred}\ \_][\underline{n+1}]]\!]_S\,\rho$, by (a) again.

The other seven computation rules are similar, and reduce to proving that $[\![(\lambda x\,.\,P)N]\!]_S\,\rho = [\![P[x := N]]\!]_S\,\rho$ (which we do by first noticing that $[\![P[x := N]]\!]_S\,\rho = [\![P]\!]_S\,(\rho[x := [\![N]\!]_S\,\rho])$), that $[\![\mathtt{Y}N]\!]_S\,\rho = [\![N(\mathtt{Y}N)]\!]_S\,\rho$, that $[\![\mathtt{succ}\,\underline{n}]\!]_S\,\rho = val_S\,(n+1) = [\![\underline{n+1}]\!]_S\,\rho$, that $[\![\mathtt{ifz}\,\underline{0}\ N\ P]\!]_S\,\rho = [\![N]\!]_S\,\rho$ and $[\![\mathtt{ifz}\,\underline{n+1}\ N\ P]\!]_S\,\rho = [\![P]\!]_S\,\rho$, and finally that for every value $V : \sigma$, $[\![\mathtt{let}\,x_\sigma \Leftarrow V\ \mathtt{in}\ N]\!]_S\,\rho = [\![N[x := V]]\!]_S\,\rho$.

The middle top rule of Figure 5 is trivial. If $C \cdot M \downarrow^{\mathsf{may}} a$ using this rule, then C is empty and $M = \underline{*}$. Then $[\![C[M]]\!]_S\,\rho\ \blacklozenge = val_S\,*(\blacklozenge) = \blacklozenge(*) = 1 > a$.

We finally deal with the four rules that implement choice ($\curlyvee$, $\oplus$). This relies on the following auxiliary claim, which states that the semantics $[\![C]\!]_S\,\rho\ (F)\ (h)$ of an arbitrary evaluation context C, applied to a prevision F and a function h, is given by applying F to some function, depending only on C, ρ and h. Precisely, the claim is:

> $(*)$ For every evaluation context C of type $\tau \vdash \mathtt{Unit}$, for every $h \in [[\![\mathtt{Unit}]\!]_S \to \mathbb{I}]$, and every environment ρ, there is a function $C._\rho h \in [[\![\tau]\!]_S \to \mathbb{I}]$ such that, for every $F \in [\![\tau]\!]_S$ (necessarily a prevision—or a fork when $\{\mathtt{A}, \mathtt{D}\} \subseteq S$), $[\![C]\!]_S\,\rho\ (F)\ (h) = F(C._\rho h)$ (resp., $[\![C]\!]_S\,\rho\ (F^-, F^+)\ (h) = (F^-(C._\rho h), F^+(C._\rho h))$ in the case of forks).

We shall prove $(*)$ below. Using $(*)$, we show that evaluation of evaluation contexts commutes with max, min and averaging operations. E.g., still

assuming $S \subseteq \{\mathtt{A}, \mathtt{P}\}$,

$$\begin{array}{rcll}
[\![C[M \mathbin{\textcircled{\scriptsize ?}} N]]\!]_S\,\rho\,(h) & = & [\![C]\!]_S\,\rho\,([\![M \mathbin{\textcircled{\scriptsize ?}} N]\!]_S\,\rho])\,(h) & \\
 & & \text{(by Claim } (a)\text{, in the prologue of this proof)} & \\
 & = & [\![M \mathbin{\textcircled{\scriptsize ?}} N]\!]_S\,\rho\,(C._\rho h) & \text{(by } (*)\text{)} \\
 & = & \max([\![M]\!]_S\,\rho\,(C._\rho h), [\![N]\!]_S\,\rho\,(C._\rho h)) & \\
 & = & \max([\![C]\!]_S\,\rho\,([\![M]\!]_S\,\rho)(h), [\![C]\!]_S\ ([\![N]\!]_S\,\rho)(h)) & \text{(by } (*)\text{)} \\
 & = & \max([\![C[M]]\!]_S\,\rho\,(h), [\![C[N]]\!]_S\,\rho\,(h)) &
\end{array}$$

Therefore, if $C \cdot M \mathbin{\textcircled{\scriptsize ?}} N \downarrow^{\mathsf{may}} a$ is deduced from $C \cdot M \downarrow^{\mathsf{may}} a$, then by induction hypothesis, and letting $h = \blacklozenge$, $[\![C[M]]\!]_S\,\rho\,\blacklozenge > a$, so $[\![C[M \mathbin{\textcircled{\scriptsize ?}} N]]\!]_S\,\rho\,\blacklozenge > a$. The case $S \subseteq \{\mathtt{D}, \mathtt{P}\}$ (with $\downarrow^{\mathsf{must}}$ instead of $\downarrow^{\mathsf{may}}$) works similarly, trading max for min. In the erratic cases $\{\mathtt{A}, \mathtt{D}\} \subseteq S$, we consider both max and min.

Finally, in the probabilistic cases, if $C \cdot M \oplus N \downarrow^m \frac{1}{2}(a+b)$ $(0 < \frac{1}{2}(a+b) < 1)$ is deduced from $C \cdot M \downarrow^m a$ and $C \cdot N \downarrow^m b$, then: if $a \neq 0$ then $[\![C[M]]\!]_S\,\rho\,(h) > a$, and if $b \neq 0$ then $[\![C[N]]\!]_S\,\rho\,(h) > b$. We do not know whether $a \neq 0$, or $b \neq 0$, here, but one of a, b must be non-zero since $\frac{1}{2}(a+b) \neq 0$, and in any case $[\![C[M \oplus N]]\!]_S\,\rho\,(h) = \frac{1}{2}([\![C[M]]\!]_S\,\rho\,(h) + [\![C[N]]\!]_S\,\rho\,(h)) > \frac{1}{2}(a+b)$.

It only remains to prove $(*)$. By Claim (b) in the prologue of this proof, it is enough to show this when C is an elementary evaluation context E: in the general case where $C = E_n E_{n-1} \dots E_2 E_1$, we shall take $C._\rho h = E_1._\rho E_2._\rho \dots ._\rho E_{n-1}._\rho E_n._\rho h$.

All the cases are similar. E.g., let E be $[\_N]$. Then:

$$\begin{array}{rcl}
[\![E]\!]_S\,\rho\,(F)\,(h) & = & [\![xN]\!]_S\,(\rho[x := F])\,(h) \\
 & = & [\mathit{let}_S\, f \Leftarrow F\ \mathit{in}\ f([\![N]\!]_S\,\rho)]\,(h) = F(f \mapsto f([\![N]\!]_S\,\rho)(h))
\end{array}$$

so one can take $E._\rho h$ to be the map $f \mapsto f([\![N]\!]_S\,\rho)(h)$ in this case. In general, whatever elementary evaluation context E we consider, $[\![E]\!]_S\,\rho\,(F)$ is of the form $\mathit{let}_S\, f \Leftarrow F\ \mathit{in}\ g(f)$ for some expression $g(f)$, and we can then define $E._\rho h$ as being $f \mapsto g(f)(h)$. □

7. Computational Adequacy

Computational adequacy is the converse of soundness, and says that if M and a normal form N have the same denotational value, then M will evaluate, operationally, to N. In PCF, this only holds at ground types. The situation is similar in PCF_S, where we shall prove this at type `Unit`, under the empty evaluation context. Together with soundness, this will mean that

$[\![M]\!]_S\,\rho\,\blacklozenge$ is equal to $\Pr(\_ \cdot M{\downarrow}^{\mathsf{may}})$, or to $\Pr(\_ \cdot M{\downarrow}^{\mathsf{must}})$, or to the pair of these two quantities, depending on S.

This is also traditionally harder to prove than soundness. Our proof strategy is adapted from Streicher [42]. A term M is *closed* if and only if it has no free variable. We define *closed evaluation contexts* similarly (the hole $\_$ is *not* a variable.) A *substitution* θ is a map from variables to terms of the same type, with finite domain. We write $M\theta$ for the result of *applying* the (capture-avoiding) substitution θ to M. The substitution θ is *closed* if and only if $x\theta$ is closed for every variable x.

Definition 7.1 ($\underset{\sim}{\sqsubset}^m_\sigma$). *For any two closed terms M, N of type σ, for any mode $m \in \{\mathsf{may}, \mathsf{must}\}$, let $M \underset{\sim}{\sqsubset}^m_\sigma N$ if and only if for every closed evaluation context C of type $\sigma \vdash \mathtt{Unit}$, $Pr(C \cdot M{\downarrow}^m) \leq Pr(C \cdot N{\downarrow}^m)$.*

Extend this to (not necessarily closed) terms by $M \underset{\sim}{\sqsubset}^m_\sigma N$ if and only if $M\theta \underset{\sim}{\sqsubset}^m_\sigma N\theta$ for every closed substitution θ.

On closed terms, it is equivalent to define $M \underset{\sim}{\sqsubset}^m_a N$ if and only if, for every rational number $a \in [0,1)$, if $C \cdot M \downarrow^m a$ is derivable, then $C \cdot N \downarrow^m a$ is derivable.

We shall use the following argument again and again. Assume $C \cdot M \to C' \cdot M'$. Then, for every $a \in \mathbb{Q} \cap [0,1)$, every derivation of $C' \cdot M'{\downarrow}^m a$ can be extended to one of $C \cdot M{\downarrow}^m a$, by using the top left rule of Figure 5. It follows that $\Pr(C \cdot M{\downarrow}^m) \geq \Pr(C' \cdot M'{\downarrow}^m)$. For example, since $C \cdot \mathtt{Y}N \to C \cdot N(\mathtt{Y}N)$ (for any term $N : \tau \to \tau$), $\Pr(C \cdot \mathtt{Y}N{\downarrow}^m) \geq \Pr(C \cdot N(\mathtt{Y}N){\downarrow}^m)$. In particular, $N(\mathtt{Y}N) \underset{\sim}{\sqsubset}^m_\tau \mathtt{Y}N$.

When M is closed, $[\![M]\!]_S\,\rho$ is independent of ρ. Let us write it $[\![M]\!]_S$, for short, then. Similarly, when M is a term with at most x as a free variable, we shall write $[\![M]\!]_S\,[x := V]$ for $[\![M]\!]_S\,\rho$ where ρ is any environment such that $\rho(x) = V$.

Proposition 7.2. *Let M be a closed term of type $\mathtt{Unit}$, and $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$ be such that $[\![M]\!]_S$ makes sense. Then:*

$(S \subseteq \{\mathtt{A}, \mathtt{P}\})$ $Pr(\_ \cdot M{\downarrow}^{\mathsf{may}}) = [\![M]\!]_S\ \blacklozenge$.

$(S \subseteq \{\mathtt{D}, \mathtt{P}\})$ $Pr(\_ \cdot M{\downarrow}^{\mathsf{must}}) = [\![M]\!]_S\ \blacklozenge$.

$(\{\mathtt{A}, \mathtt{D}\} \subseteq S)$ $Pr(\_ \cdot M{\downarrow}^{\mathsf{must}}) = F^-(\blacklozenge)$ and $Pr(\_ \cdot M{\downarrow}^{\mathsf{may}}) = F^+(\blacklozenge)$, where $(F^-, F^+) = [\![M]\!]_S$.

PROOF. We concentrate on the case $S \subseteq \{\mathtt{A}, \mathtt{P}\}$, where the mode m is may. The other two cases are similar, and we shall evoke them briefly near the end of the proof.

Define a binary logical relation R_τ between closed terms M of type τ (up to α-renaming) and previsions (or forks) in $[\![\tau]\!]_S$, by induction on types, as follows. Remembering that $[\![\tau]\!]_S$ is equal to $\mathbb{P}_S([\![\tau]\!]^\circ_S)$, we shall also define auxiliary relations R°_τ between closed terms M of type τ and values in $[\![\tau]\!]^\circ_S$, and $R^\perp_\tau$ between closed evaluation contexts C of type $\tau \vdash \mathtt{Unit}$ and functions $h \in [[\![\tau]\!]^\circ_S \to \mathbb{I}]$. For short, we say "for all C $R^\perp_\tau$ h" instead of "for every evaluation context C of type $\tau \vdash \mathtt{Unit}$, every function $h \in [[\![\tau]\!]^\circ_S \to \mathbb{I}]$, if C $R^\perp_\tau$ h then".

- M R_τ F if and only if for all C $R^\perp_\tau$ h, $\Pr(C \cdot M{\downarrow^m}) \geq F(h)$.
- C $R^\perp_\tau$ h if and only if for all M R°_τ u, $\Pr(C \cdot M{\downarrow^m}) \geq h(u)$.
- M $R^\circ_{\mathtt{Unit}}$ u if and only if ($u = *$ and) $* \sqsubseteq^m_{\sim_{\mathtt{Unit}}} M$.
- M $R^\circ_{\mathtt{Nat}}$ n if and only if $\underline{n} \sqsubseteq^m_{\sim_{\mathtt{Nat}}} M$.
- M $R^\circ_{\sigma\to\tau}$ f if and only if there is a term $M_1 : \tau$ such that $\lambda x\,.\,M_1 \sqsubseteq^m_{\sim_{\sigma\to\tau}} M$ and, for all N R_σ G, $M_1[x := N]$ R_τ $f(G)$.

A crucial item in the definition is the two-tiered definition of M R_τ F, by quantification over all C $R^\perp_\tau$ h, while C $R^\perp_\tau$ h is defined by quantification over all M R°_τ u. This pattern is similar to the technique of $\top\top$-lifting, and particularly to Katsumata's $\top\top$-logical predicates [24].

It is easy to show that, for every closed term M of type τ, M $R_\tau{=} \{F \in [\![\tau]\!]_S \mid M\ R_\tau\ F\}$ is Scott-closed, i.e., downward closed and stable under sups of directed families, and contains 0.

Also, we claim that, if $M \sqsubseteq^m_{\sim_\tau} N$ and M R_τ F, then N R_τ F, and similarly for R°_τ. I.e., $R_\tau\ F = \{M : \tau \mid M\ R_\tau\ F\}$ and $R^\circ_\tau\ u = \{M : \tau \mid M\ R^\circ_\tau\ u\}$ are upward closed in $\sqsubseteq^m_{\sim_\tau}$. This is proved as follows: If $M \sqsubseteq^m_{\sim_\tau} N$ and M R_τ F, then for all C $R^\perp_\tau$ h, $\Pr(C \cdot M{\downarrow^m}) \geq F(h)$. Since $M \sqsubseteq^m_{\sim_\tau} N$, $\Pr(C \cdot N{\downarrow^m}) \geq \Pr(C \cdot M{\downarrow^m}) \geq F(h)$. So N R_τ F. That $R^\circ_\tau\ u$ is upward closed follows from the transitivity of $\sqsubseteq^m_{\sim_\tau}$.

Finally, we observe that: (‡) whenever M R°_τ v, then M R_τ $val_S\, v$. Indeed, for all C $R^\perp_\tau$ h, the assumption implies $\Pr(C \cdot M{\downarrow^m}) \geq h(v)$. But

$val_S\, v\ (h) = h(v)$, so we have proved that $\Pr(C \cdot M{\downarrow}^m) \geq val_S\, v\ (h)$ for all $C\ R_\tau^\perp\ h$, i.e., $M\ R_\tau\ val_S\, v$.

We can now prove the *basic lemma*: if θ is a substitution and ρ is an environment such that for every variable x in the domain of θ (say, of type σ), $x\theta\ R_\sigma\ \rho(x)$—in which case we write $\theta\ R_*\ \rho$—, and if M is a term of type τ, then $M\theta\ R_\tau\ [\![M]\!]_S\, \rho$. This is by induction on the typing derivation.

When M is a value V, $[\![M]\!]_S\, \rho$ equals $val_S\, ([\![V]\!]_S^\circ\, \rho)$. We first show that for each value $V : \tau$, $V\theta\ R_\tau^\circ\ [\![V]\!]_S^\circ\, \rho$:

- When $V = \underline{n}$, this amounts to showing $\underline{n}\ R_{\mathtt{Nat}}^\circ\ n$, or equivalently $\underline{n} \sqsubseteq^m_{\sim_{\mathtt{Nat}}} \underline{n}$, and this is clear.
- When $V = \underline{*}$, similarly, this reduces to $\underline{*} \sqsubseteq^m_{\sim_{\mathtt{Unit}}} \underline{*}$.
- When $V = \lambda x_\sigma\ .\ M_1$, then $[\![V]\!]_S^\circ\, \rho = (G \in [\![\sigma]\!]_S\, \rho \mapsto [\![M_1]\!]_S\, (\rho[x := G]))$, and we need to show that for all $N\ R_\sigma\ G$, $(M_1\theta)[x := N]\ R_\tau\ [\![M_1]\!]_S\, (\rho[x := G])$. We assume that V has been α-renamed, so that x is not in the domain of θ, and not free in any term of the form $\theta(y)$. In particular, $(M_1\theta)[x := N]$ is equal to $M_1(\theta[x := N])$, where $\theta[x := N]$ is the substitution mapping x to N and each y in the domain of θ to $y\theta$. Clearly $\theta[x := N]\ R_*\ \rho[x := G]$, so $M_1(\theta[x := N])\ R_\tau\ [\![M_1]\!]_S\, (\rho[x := G])$, and we are done.

Since $V\theta\ R_\tau^\circ\ [\![V]\!]_S^\circ\, \rho$, we conclude immediately that $V\theta\ R_\tau\ val_S\, [\![V]\!]_S^\circ\, \rho = [\![V]\!]_S\, \rho$ by (‡). This finishes the case where M is a value V.

The cases of `succ`, `pred`, `ifz`, `let`, applications, and `Y` remain. We deal with `pred`, applications, `Y` and `let` only, since they are perhaps a bit subtler than the other cases.

- For `pred`, it suffices to show that if $M\ R_{\mathtt{Nat}}\ F$, then $\mathtt{pred}\, M\ R_{\mathtt{Nat}}\ let_S\, n \Leftarrow F\ in \begin{cases} val_S\, (n-1) & \text{if } n \neq 0 \\ \bot & \text{if } n = 0 \end{cases}$—the difficulty comes from some juggling we shall need to do with $R_{\mathtt{Nat}}$, $R_{\mathtt{Nat}}^\perp$, and $R_{\mathtt{Nat}}^\circ$.

  Assume $C\ R_{\mathtt{Nat}}^\perp\ h$. We must show that $\Pr(C \cdot \mathtt{pred}\, M{\downarrow}^m) \geq \left(let_S\, n \Leftarrow F\ in \begin{cases} val_S\, (n-1) & \text{if } n \neq 0 \\ 0 & \text{if } n = 0 \end{cases} \right) (h)$, in other words, that $\Pr(C \cdot \mathtt{pred}\, M{\downarrow}^m) \geq F(k)$, where k is the function $n \in \mathbb{N} \mapsto \begin{cases} val_S\, (n-1)(h) & \text{if } n \neq 0 \\ 0 & \text{if } n = 0 \end{cases}$.

Alternatively, k is the function that maps 0 to 0 and every $n \neq 0$ to $h(n-1)$.

Since every derivation of $C[\mathtt{pred}\ \_] \cdot M \downarrow^m a$ can be completed to the following:

$$\frac{C[\mathtt{pred}\ \_] \cdot M \downarrow^m a}{C \cdot \mathtt{pred}\, M \downarrow^m a}$$

using the top left rule of Figure 5 and the redex discovery rule $C \cdot \mathtt{pred}\, M \to C[\mathtt{pred}\ \_] \cdot M$, we obtain $\Pr(C \cdot \mathtt{pred}\, M\downarrow^m) \geq \Pr(C[\mathtt{pred}\ \_] \cdot M\downarrow^m)$. Therefore, it suffices to show that $\Pr(C[\mathtt{pred}\ \_] \cdot M\downarrow^m) \geq F(k)$.

We claim that $C[\mathtt{pred}\ \_]\ R^{\perp}_{\mathtt{Nat}}\ k$. Before we prove the claim, notice that the definition of $R_{\mathtt{Nat}}$, and the fact that $M\ R_{\mathtt{Nat}}\ F$, will imply the desired inequality.

To prove the claim, we must show that for all $N\ R^{\circ}_{\mathtt{Nat}}\ n$, $\Pr(C[\mathtt{pred}\ \_] \cdot N\downarrow^m)$ is larger or equal to $h(n-1)$ if $n \geq 1$, or to 0 if $n = 0$. The case $n = 0$ is clear, otherwise write $n = n' + 1$. The definition of $N\ R^{\circ}_{\mathtt{Nat}}\ n$ means that $\underline{n'+1} \sqsubseteq^m_{\sim_{\mathtt{Nat}}} N$, in particular $\Pr(C[\mathtt{pred}\ \_] \cdot N\downarrow^m) \geq \Pr(C[\mathtt{pred}\ \_] \cdot \underline{n'+1}\downarrow^m)$. Since every derivation of $C \cdot \underline{n}' \downarrow^m a$ can be completed to:

$$\frac{C \cdot \underline{n}' \downarrow^m a}{C[\mathtt{pred}\ \_] \cdot \underline{n'+1} \downarrow^m a}$$

by using the computation rule $C[\mathtt{pred}\ \_] \cdot \underline{n'+1} \to C \cdot \underline{n}'$, $\Pr(C[\mathtt{pred}\ \_] \cdot \underline{n'+1}\downarrow^m) \geq \Pr(C \cdot \underline{n}'\downarrow^m)$. The latter is greater than or equal to $h(n') = h(n-1)$ since $C\ R^{\perp}_{\mathtt{Nat}}\ h$. So $\Pr(C[\mathtt{pred}\ \_] \cdot N\downarrow^m) \geq h(n-1)$, as desired.

- As far as applications are concerned, we show more generally that: $(*)$ if $M\ R_{\sigma\to\tau}\ F$, and $N\ R_{\sigma}\ G$, then $MN\ R_{\tau}\ let_S\, f \Leftarrow F\ in\ f(G)$. Assume $C\ R^{\perp}_{\tau}\ h$: we must show that $\Pr(C \cdot MN\downarrow^m) \geq [let_S\, f \Leftarrow F\ in\ f(G)]\ (h) = F(f \mapsto f(G)(h))$. Since every derivation of $C[\_N] \cdot M\downarrow^m a$ can be completed to one of $C \cdot MN \downarrow^m a$ by the redex discovery rule $C \cdot MN \to C[\_N] \cdot M$, $\Pr(C \cdot MN\downarrow^m) \geq \Pr(C[\_N] \cdot M\downarrow^m)$, so it suffices to show $\Pr(C[\_N] \cdot M\downarrow^m) \geq F(f \mapsto f(G)(h))$. To this end, since $M\ R_{\sigma\to\tau}\ F$, it remains to show $C[\_N]\ R^{\perp}_{\sigma\to\tau}\ (f \mapsto f(G)(h))$. That is, let $P\ R^{\circ}_{\sigma\to\tau}\ f$, and let us show that $\Pr(C[\_N] \cdot P\downarrow^m) \geq f(G)(h)$. By definition of $R^{\circ}_{\sigma\to\tau}$, there is a term P_1 such that $\lambda x\,.\,P_1 \sqsubseteq^m_{\sim_{\sigma\to\tau}} P$ and (since $N\ R_{\sigma}\ G$), $P_1[x := N]\ R_{\tau}\ f(G)$. Since $C\ R^{\perp}_{\tau}\ h$, $\Pr(C \cdot P_1[x :=$

$N]\downarrow^m) \geq f(G)(h)$. Now $\Pr(C[\_N] \cdot P\downarrow^m) \geq \Pr(C[\_N] \cdot \lambda x\,.\,P_1\downarrow^m)$ since $\lambda x\,.\,P_1 \sqsubseteq^m_{\sim_{\sigma\to\tau}} P$, and $\Pr(C[\_N] \cdot \lambda x\,.\,P_1\downarrow^m) \geq \Pr(C \cdot P_1[x := N]\downarrow^m)$ since every derivation ending in $C \cdot P_1[x := N] \downarrow^m a$ can be completed to:

$$\frac{C \cdot P_1[x := N] \downarrow^m a}{C[\_N] \cdot \lambda x\,.\,P_1 \downarrow^m a}$$

by using the computation rule $C[\_N] \cdot \lambda x\,.\,P_1 \to C \cdot P_1[x := N]$. So $\Pr(C[\_N] \cdot P\downarrow^m) \geq f(G)(h)$, as claimed.

- For $\mathtt{Y}$, we show that if $N\ R_{\tau\to\tau}\ G$, then $\mathtt{Y}N\ R_\tau\ \sup_{n\in\mathbb{N}} f^n(0)$, where $f(F) = \mathit{let}_S\, g \Leftarrow G\ \mathit{in}\ g(F)$. To this end, since $\mathtt{Y}N\ R_{\tau\to\tau}$ is Scott-closed, it suffices to show that $\mathtt{Y}N\ R_\tau\ f^n(0)$ for every $n \in \mathbb{N}$. This is by induction on n.

  If $n = 0$, then $\mathtt{Y}N\ R_\tau\ 0$ is clear. Otherwise, assume $\mathtt{Y}N\ R_\tau\ f^n(0)$. By $(*)$ (in the previous item), and since $N\ R_{\tau\to\tau}\ G$, $N(\mathtt{Y}N)\ R_\tau\ \mathit{let}_S\, g \Leftarrow G\ \mathit{in}\ g(f^n(0)) = f^{n+1}(0)$. Remember that $N(\mathtt{Y}N) \sqsubseteq^m_{\sim_\tau} \mathtt{Y}N$; this was due to the computation rule $C\cdot\mathtt{Y}N \to C\cdot N(\mathtt{Y}N)$. Since $R_\tau\ f^{n+1}(0)$ is upward-closed, $\mathtt{Y}N\ R_\tau\ f^{n+1}(0)$.

- For `let` expressions, we must show that if $M\ R_\sigma\ F$, then $\mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ (N\theta)\ R_\tau\ \mathit{let}_S\, v \Leftarrow F\ \mathit{in}\ [\![N]\!]_S\,(\rho[x := \mathit{val}_S\, v])$, under the induction hypothesis that for all $\theta'\ R_*\ \rho'$, $N\theta'\ R_\tau\ [\![N]\!]_S\,\rho'$. By α-renaming, choose x fresh, in particular neither in the domain of θ or free in any term of the form $\theta(y)$.

  Let $C\ R^\perp_\tau\ h$: we must show that $\Pr(C \cdot \mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ (N\theta)\downarrow^m) \geq F(v \mapsto [\![N]\!]_S\,(\rho[x := \mathit{val}_S\, v])\ (h))$. By the redex discovery rule $C \cdot \mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ (N\theta) \to C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)]\cdot M$, $\Pr(C\cdot\mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ (N\theta)\downarrow^m) \geq \Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot M\downarrow^m)$, so it is enough to show $\Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot M\downarrow^m) \geq F(v \mapsto [\![N]\!]_S\,(\rho[x := \mathit{val}_S\, v])\ (h))$. Since $M\ R_\sigma\ F$, this reduces to showing that $C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)]\ R^\perp_\sigma\ (v \mapsto [\![N]\!]_S\,(\rho[x := \mathit{val}_S\, v])\ (h))$. To this end, fix $Q\ R^\circ_\sigma\ v$, and show that $\Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot Q\downarrow^m) \geq [\![N]\!]_S\,(\rho[x := \mathit{val}_S\, v])\ (h)$. We make a case analysis on the type σ:

  - If $\sigma = \mathtt{Unit}$, then $v = *$ and $\underline{*} \sqsubseteq^m_{\mathtt{Unit}} Q$, so $\Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot Q\downarrow^m) \geq \Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot \underline{*}\downarrow^m)$. By the computation rule $C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot \underline{*} \to C \cdot N\theta[x := \underline{*}]$, we obtain

$\Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot \underline{*}\downarrow^m) \geq \Pr(C \cdot N\theta[x := \underline{*}]\downarrow^m)$. Since $\underline{*}\ R^\circ_{\mathtt{Unit}}\ *$ by definition, $\underline{*}\ R_{\mathtt{Unit}}\ val_S * = val_S\, v$ using (‡). It follows that $\theta' = \theta[x := \underline{*}]$ is such that $\theta'\ R_*\ \rho[x := val_S\, v]$. By induction hypothesis, $N\theta[x := \underline{*}]\ R_\tau\ [\![N]\!]_S\,(\rho[x := val_S\, v])$. Since $C\ R^\perp_\tau\ h$, $\Pr(C \cdot N\theta[x := \underline{*}]\downarrow^m) \geq [\![N]\!]_S\,(\rho[x := val_S\, v])\ (h)$. We conclude that $\Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot Q\downarrow^m) \geq [\![N]\!]_S\,(\rho[x := val_S\, v])\ (h)$, as desired.

- If $\sigma = \mathtt{Nat}$, the argument is similar.
- If σ is of the form $\sigma_1 \to \sigma_2$, then $Q\ R^\circ_{\sigma_1\to\sigma_2}\ v$ means that there is a term Q_1 such that $\lambda y \cdot Q_1 \sqsubseteq^m_{\sim_{\sigma_1\to\sigma_2}} Q$ and for all $P\ R_{\sigma_1}\ G$, $Q_1[y := P]\ R_{\sigma_2}\ v(G)$. The latter entails immediately that $\lambda y\ .\ Q_1\ R^\circ_{\sigma_1\to\sigma_2}\ v$, so $\lambda y\ .\ Q_1\ R_{\sigma_1\to\sigma_2}\ val_S\, v$ by (‡). It follows that $\theta[x := \lambda y\ .\ Q_1]\ R_*\ \rho[x := val_S\, v]$, so by induction hypothesis $N\theta[x := \lambda y\ .\ Q_1]\ R_\tau\ [\![N]\!]_S\,(\rho[x := val_S\, v])$. Since $C\ R^\perp_\tau\ h$, $\Pr(C \cdot N\theta[x := \lambda y\ .\ Q_1]\downarrow^m) \geq [\![N]\!]_S\,(\rho[x := val_S\, v])\ (h)$. We complete the argument by noting that $\Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot Q\downarrow^m) \geq \Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot \lambda y\ .\ Q_1\downarrow^m) \geq \Pr(C \cdot (N\theta)[x := \lambda y\ .\ Q_1]\downarrow^m)$, using the fact that $\lambda y.Q_1 \sqsubseteq^m_{\sim_{\sigma_1\to\sigma_2}} Q$, then the computation rule $C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot \lambda y\ .\ Q_1 \to C \cdot (N\theta)[x := \lambda y\ .\ Q_1]$. So $\Pr(C[\mathtt{let}\, x \Leftarrow \_\ \mathtt{in}\ (N\theta)] \cdot Q\downarrow^m) \geq [\![N]\!]_S\,(\rho[x := val_S\, v])$, as desired.

Finally, we deal with $\otimes$ and $\oplus$. Assume $M\ R_{\mathtt{T}\tau}\ F$ and $N\ R_{\mathtt{T}\tau}\ G$, and let $C\ R^\perp_\tau\ h$. We must show that $\Pr(C \cdot M \otimes N\downarrow^{\mathsf{may}}) \geq \max(F(h), G(h))$ and $\Pr(C \cdot M \oplus N\downarrow^{\mathsf{may}}) \geq \frac{1}{2}(F(h) + G(h))$. By definition, $\Pr(C \cdot M\downarrow^{\mathsf{may}}) \geq F(h)$ and $\Pr(C \cdot N\downarrow^{\mathsf{may}}) \geq G(h)$. We have already seen that $\Pr(C \cdot M \otimes N\downarrow^{\mathsf{may}}) = \max(\Pr(C \cdot M\downarrow^{\mathsf{may}}), \Pr(C \cdot N\downarrow^{\mathsf{may}}))$, so $\Pr(C \cdot M \otimes N\downarrow^{\mathsf{may}}) \geq \max(F(h), G(h))$. Similarly, $\Pr(C \cdot M \oplus N\downarrow^{\mathsf{may}}) = \frac{1}{2}(\Pr(C \cdot M\downarrow^{\mathsf{may}}) + \Pr(C \cdot N\downarrow^{\mathsf{may}})) \geq \frac{1}{2}(F(h) + G(h))$.

This terminates the proof of the basic lemma in the angelic cases.

To conclude the proof of the proposition in the case $S \subseteq \{\mathtt{A}, \mathtt{P}\}$, we now note that: (†) $\_\ R^\perp_{\mathtt{Unit}}\ \blacklozenge$. To show this, we need to show that, for all $Q\ R^\circ_{\mathtt{Unit}}\ v$, $\Pr(\_ \cdot Q\downarrow^m) \geq \blacklozenge(v)$. The assumption $Q\ R^\circ_{\mathtt{Unit}}\ v$ means that $v = *$ and $\underline{*} \sqsubseteq^m_{\sim_{\mathtt{Unit}}} Q$, so $\Pr(\_ \cdot Q\downarrow^m) \geq \Pr(\_ \cdot \underline{*}\downarrow^m) = 1$, while $\blacklozenge(*) = 1$, whence the claim.

Apply the basic lemma to closed terms M of type σ, with an arbitrary environment ρ and an arbitrary substitution θ: for all $C\ R^\perp_\sigma\ h$, $\Pr(C \cdot M\downarrow^m) \geq$

$[\![M]\!]_S(h)$. In particular, by (†), when $\sigma = \mathtt{Unit}$, $\Pr(\_ \cdot M{\downarrow^m}) \geq [\![M]\!]_S$ ♦. The converse inequality is by soundness (Lemma 6.1).

The above arguments apply to the case $S \subseteq \{\mathtt{A}, \mathtt{P}\}$. When $S \subseteq \{\mathtt{D}, \mathtt{P}\}$, the same arguments apply, replacing $\downarrow^{\mathsf{may}}$ by $\downarrow^{\mathsf{must}}$. The only change is in the proof of the basic lemma, case of the $\otimes$ operator, where we use min instead of max: $\Pr(C \cdot M \otimes N{\downarrow^{\mathsf{must}}}) = \min(\Pr(C \cdot M{\downarrow^{\mathsf{must}}}), \Pr(C \cdot N{\downarrow^{\mathsf{must}}})) \geq \min(F(h), G(h))$. When $\{\mathtt{A}, \mathtt{D}\} \subseteq S$, the reasoning proceeds by proving the conjunction of the previous two cases simultaneously. □

Proposition 7.2 is especially useful in view of the following lemma.

Lemma 7.3. *For every type τ, for every term $M : \tau$, for every evaluation context C of type $\tau \vdash \mathtt{Unit}$, for every mode $m \in \{\mathsf{may}, \mathsf{must}\}$, $Pr(\_ \cdot C[M]{\downarrow^m}) = Pr(C \cdot M{\downarrow^m})$.*

Proof. We claim that $\_ \cdot C[M] \downarrow^m a$ is derivable if and only if $C \cdot M \downarrow^m a$ is derivable. The key step is showing that if CE is an evaluation context, obtained as the concatenation of C and of the elementary evaluation context E, then $C \cdot E[M] \downarrow^m a$ if and only if $CE \cdot M \downarrow^m a$, for every $M : \sigma$, every elementary evaluation context E of type $\sigma \vdash \tau$, and every evaluation context C of type $\tau \vdash \mathtt{Unit}$.

We enumerate all five possible cases for E. They are all straightforward. E.g., when $E = [\_N]$, then if $C[\_N] \cdot M \downarrow^m a$ is derivable, we may produce the following derivation using the redex discovery rule $C \cdot MN \to C[\_N] \cdot M$:

$$\frac{\begin{array}{c}\vdots\\ C[\_N] \cdot M \downarrow^m a\end{array}}{C \cdot MN \downarrow^m a}$$

Conversely, if $C \cdot MN \downarrow^m a$ is derivable, then, by inspection of the rules, the proof must be exactly as above: so $C[\_N] \cdot M \downarrow^m a$ is derivable as well. □

Proposition 7.2 then reads:

Theorem 7.4 (Computational Adequacy). *Let $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$. Let C be an evaluation context of type $\tau \vdash \mathtt{Unit}$, and M be a PCF_S term τ such that $C[M]$ is closed and such that $[\![C[M]]\!]_S$ makes sense. Then:*

$(S \subseteq \{\mathtt{A}, \mathtt{P}\})$ $Pr(C \cdot M{\downarrow^{\mathsf{may}}}) = Pr(\_ \cdot C[M]{\downarrow^{\mathsf{may}}}) = [\![C[M]]\!]_S$ ♦.

$(S \subseteq \{\mathtt{D}, \mathtt{P}\})$ $Pr(C \cdot M{\downarrow}^{\mathsf{must}}) = Pr(\_ \cdot C[M]{\downarrow}^{\mathsf{must}}) = [\![C[M]]\!]_S$ $\blacklozenge$.

$(\{\mathtt{A}, \mathtt{D}\} \subseteq S)$ $Pr(C{\cdot}M{\downarrow}^{\mathsf{must}}) = Pr(\_{\cdot}C[M]{\downarrow}^{\mathsf{must}}) = F^-(\blacklozenge)$ *and* $Pr(C{\cdot}M{\downarrow}^{\mathsf{may}}) = Pr(\_ \cdot C[M]{\downarrow}^{\mathsf{must}}) = F^+(\blacklozenge)$, *where* $(F^-, F^+) = [\![C[M]]\!]_S$.

8. The Failure of Full Abstraction: Statistical Termination Testers

Full abstraction is a result that relates the denotational ordering with the so-called contextual preorder. Slightly more precisely, such a result would state that for any two closed terms M, N of type σ, $[\![M]\!]_S \leq [\![N]\!]_S$ if and only if $M \precsim^m_\sigma N$, where $\precsim^m_\sigma$ is defined as follows.

Definition 8.1 (Contextual Preorder $\precsim^m_\sigma$). *For any two closed terms M, N of type σ, for any mode $m \in \{\mathsf{may}, \mathsf{must}\}$, let $M \precsim^m_\sigma N$ if and only if for every closed term $P : \sigma \to \mathtt{Unit}$, $Pr(\_ \cdot PM{\downarrow}^m) \leq Pr(\_ \cdot PN{\downarrow}^m)$.*

A variant on Jung's proof of Milner's context lemma [42, Theorem 8.1] shows that this is a familiar relation:

Proposition 8.2. *For all closed terms M, N of type σ, for every mode $m \in \{\mathsf{may}, \mathsf{must}\}$, $M \precsim^m_\sigma N$ if and only if $M \sqsubseteq^m_\sigma N$.*

In any case, we leave the language PCF_S implicit. For $m = \mathsf{may}$, we may take $S = \mathtt{A}$, $S = \mathtt{AP}$, or $S = \mathtt{ADP}$. For $m = \mathsf{must}$, $S = \mathtt{D}$, $S = \mathtt{DP}$, or $S = \mathtt{ADP}$.

Proof. We first note that for every term $Q : \sigma$, for every subset S of $\{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$ such that the following terms make sense, $[\![(\lambda x \,.\, C[x])Q]\!]_S = [\![C[Q]]\!]_S$. Indeed, using the definition yields $[\![(\lambda x \,.\, C[x])M]\!]_S = [\![C[x]]\!]_S\,[x := [\![M]\!]_S]$, and this is equal to $[\![C[Q]]\!]_S$ by a simple induction on C.

If $M \precsim^m_\sigma N$, then in particular, for every closed evaluation context C of type $\sigma \vdash \mathtt{Unit}$, letting $P = \lambda x \,.\, C[x]$, $\Pr(\_ \cdot (\lambda x \,.\, C[x])M{\downarrow}^m) \leq \Pr(\_ \cdot (\lambda x \,.\, C[x])N{\downarrow}^m)$. Using computational adequacy (Theorem 7.4) and the fact that $[\![(\lambda x \,.\, C[x])M]\!]_S = [\![C[M]]\!]_S$ and $[\![(\lambda x \,.\, C[x])N]\!]_S = [\![C[N]]\!]_S$, we obtain $\Pr(\_ \cdot C[M]{\downarrow}^m) \leq \Pr(\_ \cdot C[N]{\downarrow}^m)$, hence $\Pr(C \cdot M{\downarrow}^m) \leq \Pr(C \cdot N{\downarrow}^m)$. So $M \sqsubseteq^m_\sigma N$.

Conversely, assume $M \sqsubseteq^m_\sigma N$. We reuse the logical relation R_τ that we defined in the proof of Proposition 7.2, and recall that we proved that $M \; R_\sigma \; [\![M]\!]_S$, and also that if $M \; R_\sigma \; F$ and $M \sqsubseteq^m_\sigma N$, then $N \; R_\sigma \; F$. So $N \; R_\sigma \; [\![M]\!]_S$. We also proved (†): $\_ \; R^\perp_{\mathtt{Unit}} \; \blacklozenge$; (‡), which implies: if $Q \; R^\circ_{\sigma \to \mathtt{Unit}} \; f$

then $Q\ R_{\sigma\to\mathtt{Unit}}\ val_S\, f$; and $(*)$, which together with the latter, implies: if $Q\ R^{\circ}_{\sigma\to\mathtt{Unit}}\ f$ and $N\ R_\sigma\ G$ for some G, then $QN\ R_{\mathtt{Unit}}\ let_S\, f \Leftarrow val_S\, f\ in\ f(G)$, that is, $QN\ R_{\mathtt{Unit}}\ f(G)$.

Let P be a closed term of type $\sigma\to\mathtt{Unit}$. In particular, $P\ R_{\sigma\to\mathtt{Unit}}\ [\![P]\!]_S$. For all $Q\ R^{\circ}_{\sigma\to\mathtt{Unit}}\ f$, the last observation of the last paragraph entails that $QN\ R_{\mathtt{Unit}}\ f([\![M]\!]_S)$. By $(\dagger)$, $\Pr(\_\cdot QN{\downarrow^m}) \geq f([\![M]\!]_S)(\blacklozenge)$, and by Lemma 7.3, $\Pr([\_N]\cdot Q{\downarrow^m}) \geq f([\![M]\!]_S)(\blacklozenge)$. As this holds for all $Q\ R^{\circ}_{\sigma\to\mathtt{Unit}}\ f$, we obtain $[\_N]\ R^{\perp}_{\sigma\to\mathtt{Unit}}\ (f\mapsto f([\![M]\!]_S)(\blacklozenge))$, by definition of $R^{\perp}_{\sigma\to\mathtt{Unit}}$. Since $P\ R_{\sigma\to\mathtt{Unit}}\ [\![P]\!]_S$, we deduce that $\Pr([\_N]\cdot P{\downarrow^m}) \geq [\![P]\!]_S\,(f\mapsto f([\![M]\!]_S)(\blacklozenge))$. Using Lemma 7.3, $\Pr([\_N]\cdot P{\downarrow^m}) = \Pr(\_\cdot PN{\downarrow^m})$, while $[\![P]\!]_S\,(f\mapsto f([\![M]\!]_S)(\blacklozenge)) = [\![PM]\!]_S\ \blacklozenge$. So $\Pr(\_\cdot PN{\downarrow^m}) \geq [\![PM]\!]_S\ \blacklozenge$. By Proposition 7.2, $[\![PM]\!]_S\ \blacklozenge = \Pr(\_\cdot PM{\downarrow^m})$, so $\Pr(\_\cdot PM{\downarrow^m}) \leq \Pr(\_\cdot PN{\downarrow^m})$, and as P is arbitrary, $M \precsim^m N$. □

The same technique yields the following useful lemma.

Lemma 8.3. *Let $m\in\{\mathsf{may},\mathsf{must}\}$. Let $\lambda x_\sigma\,.\,M$ and $\lambda x_\sigma\,.\,N$ be two closed values of type $\sigma\to\tau$, and assume that for every closed term $P:\sigma$, $M[x:=P] \precsim^m_\tau N[x:=P]$. Then $\lambda x_\sigma\,.\,M \precsim^m_{\sigma\to\tau} \lambda x_\sigma\,.\,N$.*

Proof. In view of Proposition 8.2, we shall show that $\lambda x_\sigma\,.\,M \sqsubset^m_{\sim_{\sigma\to\tau}} \lambda x_\sigma\,.\,N$ under the assumption that for every closed term $P:\sigma$, $M[x:=P] \sqsubset^m_{\sim_\tau} N[x:=P]$.

We claim that for all $P\ R_\sigma\ F$, $N[x:=P]\ R_\tau\ [\![M]\!]_S\,[x:=F]$. This follows from $M[x:=P] \sqsubset^m_{\sim_\tau} N[x:=P]$ and $M[x:=P]\ R_\tau\ [\![M]\!]_S\,[x:=F]$, the latter being a consequence of $\lambda x\,.\,M\ R^{\circ}_{\sigma\to\tau}\ [\![\lambda x\,.\,M]\!]^{\circ}_S$.

From this, we infer that $\lambda x\,.\,N\ R^{\circ}_{\sigma\to\tau}\ [\![\lambda x\,.\,M]\!]^{\circ}_S$. Indeed, for all $P\ R_\sigma\ F$, $N[x:=P]\ R_\tau\ [\![M]\!]_S\,[x:=F] = [\![\lambda x\,.\,M]\!]^{\circ}_S(F)$. So $\lambda x\,.\,N\ R_{\sigma\to\tau}\ val_S\,([\![\lambda x\,.\,M]\!]^{\circ}_S) = [\![\lambda x\,.\,M]\!]_S$.

For every closed evaluation context C of type $(\sigma\to\tau)\vdash\mathtt{Unit}$, $[\![C[\lambda x\,.\,M]]\!]_S = [\![C[y]]\!]_S\,[y:=[\![\lambda x\,.\,M]\!]_S]$. By the basic lemma, $C[y][y:=\lambda x\,.\,N]\ R_{\mathtt{Unit}}\ [\![C[y]]\!]_S\,[y:=[\![\lambda x\,.\,M]\!]_S]$, i.e., $C[\lambda x\,.\,N]\ R_{\mathtt{Unit}}\ [\![C[\lambda x\,.\,M]]\!]_S$. Since $\_\ R^{\perp}_{\mathtt{Unit}}\ \blacklozenge$, $\Pr(\_\cdot C[\lambda x\,.\,N]{\downarrow^m}) \geq [\![C[\lambda x\,.\,M]]\!]_S\ \blacklozenge$. But $\Pr(\_\cdot C[\lambda x\,.\,N]{\downarrow^m}) = \Pr(C\cdot\lambda x\,.\,N{\downarrow^m})$ by Lemma 7.3, and $[\![C[\lambda x\,.\,M]]\!]_S\ \blacklozenge = \Pr(C\cdot\lambda x\,.\,M{\downarrow^m})$ by Theorem 7.4. So $\Pr(C\cdot\lambda x\,.\,N{\downarrow^m}) \geq \Pr(C\cdot\lambda x\,.\,M{\downarrow^m})$. Since C is arbitrary, $M \sqsubset^m_{\sim_{\sigma\to\tau}} N$. □

By analogy with the case of PCF, one would expect PCF_S to not be fully abstract. However, the reasons of the failure of full abstraction are different:

we are not lacking a parallel or construct, but so-called statistical termination testers, as we shall see.

We should indeed ponder the fact that, given two terms $M, N : \mathtt{Nat}$, the $\mathrm{PCF}_{\mathtt{A}}$-definable term $(M \,\mathtt{or}\, N) \oslash (N \,\mathtt{or}\, M)$ implements a variant of parallel or, where $M \,\mathtt{or}\, N = \mathtt{ifz}\ M\ N\ \underline{1}$ is the sequential short-circuit or. Contrarily to Plotkin's, this form of parallel or terminates even when M and N have values other than 0 or 1. Plotkin's actual parallel or of M and N can be defined as $(M \,\mathtt{or}'\, N) \oslash (N \,\mathtt{or}'\, M)$, where $\mathtt{or}'$ is more complicated:

$$\begin{array}{rcl} M \,\mathtt{or}'\, N & = & \mathtt{ifz}\ M\ (\mathtt{ifz}\ N\ \underline{0}\ (\mathtt{ifz}\ (\mathtt{pred}\, N)\ \underline{1}\ \Omega_{\mathtt{Nat}})) \\ & & \qquad (\mathtt{ifz}\ (\mathtt{pred}\, M) \\ & & \qquad\quad (\mathtt{ifz}\ N\ \underline{1}\ (\mathtt{ifz}\ (\mathtt{pred}\, N)\ \underline{1}\ \Omega_{\mathtt{Nat}})) \\ & & \qquad\quad \Omega_{\mathtt{Nat}} \end{array}$$

and $\Omega_{\mathtt{Nat}}$ is a non-terminating term of type $\mathtt{Nat}$ (see below).

Another difference with PCF is that full abstraction will in fact fail only in the probabilistic cases, where $\mathtt{P} \in S$.

We shall keep in mind the following lemma, which will be useful in understanding what happens at type $\mathtt{Unit}$.

Lemma 8.4. *Let $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$, with $\{\mathtt{A}, \mathtt{D}\} \not\subseteq S$.*

- *If $\mathtt{P} \in S$, then $\mathbb{P}_S\{*\}$ is isomorphic to $\mathbb{I}$: the elements of $\mathbb{P}_S\{*\}$ are all of the form $\alpha\, val_S\, *$, $\alpha \in \mathbb{I}$ and the isomorphism maps them to α.*
- *If $\mathtt{P} \notin S$, then $\mathbb{P}_S\{*\}$ is isomorphic to $\{0, 1\}$: the elements of $\mathbb{P}_S\{*\}$ are the zero map and $val_S\, *$.*

Proof. Recall that $val_S\, *$ is the map $(h \mapsto h(*))$. The continuous maps $h \in [\{*\} \to \mathbb{I}]$ are all constant maps, with value $h(*)$. For $F \in \mathbb{P}_S\{*\}$, let $\alpha = F(\mathbf{1})$: then $F(h) = \alpha h(*) = \alpha\, val_S\, *\, (h)$. In the second case, since F is discrete, α can only be equal to 0 or 1. □

Proposition 8.5 (Full Abstraction Fails in $\mathbf{PCF}_S$). *Let $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$, and assume that $\mathtt{P} \in S$. Let $m = \mathsf{may}$ if $S = \mathtt{AP}$, must if $S = \mathtt{DP}$, or any one of may, must if $S = \mathtt{ADP}$. Consider the following PCF_S terms:*

$$\begin{array}{rcl} \Omega_\tau & = & \mathtt{Y}(\lambda x_\tau\, .\, x) \\ M_{8.5} & = & \lambda g_{\mathtt{Unit}\to\mathtt{Unit}}\, .\, g(\Omega_{\mathtt{Unit}} \oplus \underline{*}) \\ N_{8.5} & = & \lambda g_{\mathtt{Unit}\to\mathtt{Unit}}\, .\, g(\Omega_{\mathtt{Unit}}) \oplus \underline{*} \end{array}$$

Then $M_{8.5} \precsim^m_{(\mathtt{Unit}\to\mathtt{Unit})\to\mathtt{Unit}} N_{8.5}$, but $[\![M_{8.5}]\!]^\circ_S \not\leq [\![N_{8.5}]\!]^\circ_S$.

As a consequence, $[\![M_{8.5}]\!]_S = val_S\,[\![M_{8.5}]\!]^\circ_S \not\leq val_S\,[\![N_{8.5}]\!]^\circ_S = [\![N_{8.5}]\!]_S$, since our monads have order-reflecting units (Remark 5.5).

PROOF. We only deal with the non-erratic cases $S \not\supseteq \{\mathtt{A}, \mathtt{D}\}$. As usual, the erratic cases consist in playing the same arguments as in the corresponding angelic and demonic cases simultaneously.

First step: defining a logical relation. Let $\triangleright$ be any binary relation on $\mathbb{I}$ such that $0 \triangleright 0$, which is inductive in the sense that if $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are any two non-decreasing chains of elements of $\mathbb{I}$ such that $a_n \triangleright b_n$ for every $n \in \mathbb{N}$, then $\sup_{n\in\mathbb{N}} a_n \triangleright \sup_{n\in\mathbb{N}} b_n$; and which is closed under max, min, and means: if $a_1 \triangleright a_2$ and $b_1 \triangleright b_2$, then $\max(a_1, b_1) \triangleright \max(a_2, b_2)$, $\min(a_1, b_1) \triangleright \min(a_2, b_2)$, and $\frac{1}{2}(a_1 + b_1) \triangleright \frac{1}{2}(a_2 + b_2)$.

Define binary logical relations $(\triangleright)_\tau$, $(\triangleright)^\circ_\tau$, $(\triangleright)^\perp_\tau$ between previsions in $[\![\tau]\!]_S$, between continuous maps in $[[\![\tau]\!]^\circ_S \to \mathbb{I}]$, and between values in $[\![\tau]\!]^\circ_S$ respectively, as follows:

- $F_1\ (\triangleright)_\tau\ F_2$ if and only if for all $h_1\ (\triangleright)^\perp_\tau\ h_2$, $F_1(h_1) \triangleright F_2(h_2)$;
- $h_1\ (\triangleright)^\perp_\tau\ h_2$ if and only if for all $v_1\ (\triangleright)^\circ_\tau\ v_2$, $h_1(v_1) \triangleright h_2(v_2)$;
- $*\ (\triangleright)^\circ_{\mathtt{Unit}}\ *$;
- $n_1\ (\triangleright)^\circ_{\mathtt{Nat}}\ n_2$ if and only if $n_1 = n_2$;
- $f_1\ (\triangleright)^\circ_{\sigma\to\tau}\ f_2$ if and only if for all $F_1\ (\triangleright)_\sigma\ F_2$, $f_1(F_1)\ (\triangleright)_\tau\ f_2(F_2)$.

Given two environments ρ_1, ρ_2, write $\rho_1\ (\triangleright)^*\ \rho_2$ if and only if for every variable x_σ, $\rho_1(x_\sigma)\ (\triangleright)_\sigma\ \rho_2(x_\sigma)$. We prove the basic lemma: for every term $M : \tau$, if $\rho_1\ (\triangleright)^*\ \rho_2$, then $[\![M]\!]_S\,\rho_1\ (\triangleright)_\tau\ [\![M]\!]_S\,\rho_2$. This is by induction on M, and we leave the verification of this claim to the reader. (As hints, first show the following facts: (a) if $v_1\ (\triangleright)^\circ_\sigma\ v_2$, then $val_S\,v_1\ (\triangleright)_\sigma\ val_S\,v_2$; (b) $0\ (\triangleright)_\tau\ 0$; (c) $(\triangleright)_\tau$ is inductive for every type τ, i.e., for all non-decreasing chains of previsions $(F_{1n})_{n\in\mathbb{N}}$ and $(F_{2n})_{n\in\mathbb{N}}$ such that $F_{1n}\ (\triangleright)_\tau\ F_{2n}$ for every $n \in \mathbb{N}$, then $\sup_{n\in\mathbb{N}} F_{1n}\ (\triangleright)_\tau\ \sup_{n\in\mathbb{N}} F_{2n}$. The former is used when M is a constant, the latter are used in the case where M is of the form $\mathtt{Y}N$. The fact that $\triangleright$ is closed under max, min, and means is used when M is of the form $N \mathbin{\mathbb{V}} P$, or $N \oplus P$.)

In particular, when P is a closed term of type τ, $[\![P]\!]_S\ (\triangleright)_\tau\ [\![P]\!]_S$.

A specific logical relation. Specialize the above construction to the case where $\triangleright$ is defined by: $a_1 \triangleright a_2$ if and only if $a_2 \leq (1-a)a_1 + a$, where a is some fixed constant in $[0,1]$. (Later, we shall be interested in $a = \frac{1}{2}$, but there is no point now in taking a specific value for a.) Let us make $(\triangleright)_{\mathtt{Unit}}$ explicit.

First, $h_1 \ (\triangleright)^{\perp}_{\mathtt{Unit}} \ h_2$ if and only if $h_2(*) \leq (1-a)h_1(*) + a$. If we equate $[\{*\} \to \mathbb{I}]$ with $\mathbb{I}$ (i.e., if we equate $\blacklozenge$ with 1), hence h_i with the real number $h_i(*) \in \mathbb{I}$, we have $a_1 \ (\triangleright)^{\perp}_{\mathtt{Unit}} \ a_2$ if and only if $a_2 \leq (1-a)a_1 + a$, if and only if $a_1 \triangleright a_2$.

Let us turn to $(\triangleright)_{\mathtt{Unit}}$. At type $\mathtt{Unit}$, $\mathbb{P}_S\{*\}$ is isomorphic to $\mathbb{I}$, by Lemma 8.4: any $F \in \mathbb{P}_S\{*\}$ can be written in a unique way as $\alpha \, val_S *$ for some $\alpha \in \mathbb{I}$.

We claim that $\alpha_1 \, val_S * \ (\triangleright)_{\mathtt{Unit}} \ \alpha_2 \, val_S *$ if and only if $\alpha_1 \triangleright \alpha_2$. By definition, $\alpha_1 \, val_S * \ (\triangleright)_{\mathtt{Unit}} \ \alpha_2 \, val_S *$ if and only if for all $a_1 \triangleright a_2$, $\alpha_1 \, val_S *(a_1 \mathbf{1}) \triangleright \alpha_2 \, val_S *(a_2 \mathbf{1})$, that is, $\alpha_1 a_1 \triangleright \alpha_2 a_2$. If this is the case, namely if $\alpha_1 a_1 \triangleright \alpha_2 a_2$ for all $a_1 \triangleright a_2$, then taking $a_1 = a_2 = 1$ we obtain $\alpha_1 \triangleright \alpha_2$. Conversely, if $\alpha_1 \triangleright \alpha_2$, then for all $a_1 \triangleright a_2$:

$$\begin{array}{rcll}
\alpha_2 a_2 & \leq & ((1-a)\alpha_1 + a)((1-a)a_1 + a) & \text{(since } \alpha_1 \triangleright \alpha_2, \ a_1 \triangleright a_2\text{)} \\
& = & (\alpha_1 + a(1-\alpha_1))((1-a)a_1 + a) & \\
& = & \alpha_1((1-a)a_1 + a) + a(1-\alpha_1)((1-a)a_1 + a) & \\
& \leq & \alpha_1((1-a)a_1 + a) + a(1-\alpha_1)((1-a) + a) & \\
& = & \alpha_1((1-a)a_1 + a) + a(1-\alpha_1) = (1-a)\alpha_1 a_1 + a. &
\end{array}$$

Relating two continuations. With the above definition of $\triangleright$, and taking $a = \frac{1}{2}$ (so that $a_1 \triangleright a_2$ if and only if $a_2 \leq \frac{1}{2}a_1 + \frac{1}{2}$), we claim that $(f \in [\mathbb{P}_S\{*\} \to \mathbb{P}_S\{*\}] \mapsto f(0)(h)) \ (\triangleright)^{\perp}_{\mathtt{Unit}\to\mathtt{Unit}} \ (f \in [\mathbb{P}_S\{*\} \to \mathbb{P}_S\{*\}] \mapsto f(\frac{1}{2} \, val_S *)(h))$, for every $h \in [\{*\} \to \mathbb{I}]$.

To this end, let $f_1 \ (\triangleright)_{\mathtt{Unit}\to\mathtt{Unit}} \ f_2$, and let us show that $f_1(0)(h) \triangleright f_2(\frac{1}{2} \, val_S *)(h)$. Since $f_1 \ (\triangleright)_{\mathtt{Unit}\to\mathtt{Unit}} \ f_2$, for all $\alpha_1 \triangleright \alpha_2$ (i.e., whenever $\alpha_1 \, val_S * \ (\triangleright)_{\mathtt{Unit}} \ \alpha_2 \, val_S *$), it must hold $f_1(\alpha_1 \, val_S *) \ (\triangleright)_{\mathtt{Unit}} \ f_2(\alpha_2 \, val_S *)$. Let $\alpha_1 = 0$, $\alpha_2 = \frac{1}{2}$: clearly, and since $a = \frac{1}{2}$, $\alpha_1 \triangleright \alpha_2$. So $f_1(0) \ (\triangleright)_{\mathtt{Unit}} \ f_2(\frac{1}{2} \, val_S *)$. Now, $h \ (\triangleright)^{\perp}_{\mathtt{Unit}} \ h$: this indeed means that $h(*) \leq \frac{1}{2}h(*) + \frac{1}{2}$, which is clear since $h(*) \leq 1$. By definition of $(\triangleright)_{\mathtt{Unit}}$, we conclude that $f_1(0)(h) \triangleright f_2(\frac{1}{2} \, val_S *)(h)$.

Comparing $M_{8.5}$ and $N_{8.5}$. For every closed term $P : \mathtt{Unit} \to \mathtt{Unit}$, the above claim entails that $[\![P]\!]_S \, (f \mapsto f(0)(h)) \triangleright [\![P]\!]_S \, (f \mapsto f(\frac{1}{2} \, val_S *)(h))$,

since $[\![P]\!]_S \ (\triangleright)_{\mathtt{Unit}\to\mathtt{Unit}} \ [\![P]\!]_S$; i.e., $[\![P]\!]_S\,(f \mapsto f(\frac{1}{2}\,val_S\,*)(h)) \leq \frac{1}{2}\,[\![P]\!]_S\,(f \mapsto f(0)(h)) + \frac{1}{2}$.

Using the fact that $[\![\Omega_\tau]\!]_S = 0$ at every type τ (which we leave as an exercise to the reader), $[\![\Omega_{\mathtt{Unit}} \oplus \underline{*}]\!]_S = \frac{1}{2}\,val_S\,*$. So:

$$\begin{array}{rcl} [\![P(\Omega_{\mathtt{Unit}} \oplus \underline{*})]\!]_S\,(h) & = & [\![P]\!]_S\,(f \mapsto f([\![\Omega_{\mathtt{Unit}} \oplus \underline{*}]\!]_S)(h)) \\ & = & [\![P]\!]_S\,(f \mapsto f(\frac{1}{2}\,val_S\,*)(h)) \end{array}$$

and similarly:

$$[\![P(\Omega_{\mathtt{Unit}}) \oplus \underline{*}]\!]_S\,(h) \quad = \quad \tfrac{1}{2}\,[\![P]\!]_S\,(f \mapsto f(0)(h)) + \tfrac{1}{2}$$

So $[\![P(\Omega_{\mathtt{Unit}} \oplus \underline{*})]\!]_S \leq [\![P(\Omega_{\mathtt{Unit}}) \oplus \underline{*}]\!]_S$. For every closed evaluation context C of type $\mathtt{Unit} \vdash \mathtt{Unit}$, it follows that $\Pr(C{\cdot}P(\Omega_{\mathtt{Unit}} \oplus \underline{*})\downarrow^m) \leq \Pr(C{\cdot}P(\Omega_{\mathtt{Unit}}) \oplus \underline{*}\downarrow^m)$, using computational adequacy (Theorem 7.4). So $P(\Omega_{\mathtt{Unit}} \oplus \underline{*}) \sqsubseteq^m_{\sim_{\mathtt{Unit}}} P(\Omega_{\mathtt{Unit}}) \oplus \underline{*}$, whence $P(\Omega_{\mathtt{Unit}} \oplus \underline{*}) \precsim^m_{\mathtt{Unit}} P(\Omega_{\mathtt{Unit}}) \oplus \underline{*}$ by Proposition 8.2. Since P is arbitrary, Lemma 8.3 tells us that $\lambda g\,.\,g(\Omega_{\mathtt{Unit}} \oplus \underline{*}) \precsim^m_{(\mathtt{Unit}\to\mathtt{Unit})\to\mathtt{Unit}} \lambda g\,.\,g(\Omega_{\mathtt{Unit}}) \oplus \underline{*}$, in other words $M_{8.5} \precsim^m_{(\mathtt{Unit}\to\mathtt{Unit})\to\mathtt{Unit}} N_{8.5}$.

Comparing the denotations of $M_{8.5}$ and $N_{8.5}$. For every $G \in [\![\mathtt{Unit} \to \mathtt{Unit}]\!]_S$, by computations similar to the ones we have done above:

$$\begin{array}{rcl} [\![M_{8.5}]\!]^\circ_S\,(G)\ \blacklozenge & = & [\![g(\Omega_{\mathtt{Unit}} \oplus \underline{*})]\!]_S\,[g := G]\ \blacklozenge \\ & = & G(f \mapsto f(\frac{1}{2}\,val_S\,*)\ \blacklozenge) \\ [\![N_{8.5}]\!]^\circ_S\,(G)\ \blacklozenge & = & [\![g(\Omega_{\mathtt{Unit}}) \oplus \underline{*}]\!]_S\,[g := G]\ \blacklozenge \\ & = & \frac{1}{2}G(f \mapsto f(0)\ \blacklozenge) + \frac{1}{2}. \end{array}$$

For any $b \in [0,1]$, let $[> b]$ be the map from $\mathbb{P}_S\{*\}$ to $\mathbb{P}_S\{*\}$ that sends $\alpha\,val_S\,*$ to $val_S\,*$ if $\alpha > b$, and to 0 otherwise. (If we equate $\mathbb{P}_S\{*\}$ with $\mathbb{I}$, this is just the characteristic function $\chi_{(b,1]}$.) This is a continuous map. By taking $G = val_S[> \frac{1}{4}]$, so that $G(f \mapsto f(\alpha\,val_S\,*)\ \blacklozenge) = [> \frac{1}{4}](\alpha)$, we obtain $[\![M_{8.5}]\!]^\circ_S\,(G)\ \blacklozenge = 1$ and $[\![N_{8.5}]\!]^\circ_S\,(G)\ \blacklozenge = \frac{1}{2}$, so $[\![M_{8.5}]\!]^\circ_S \not\leq [\![N_{8.5}]\!]^\circ_S$. □

Call a prevision F in $[\![\tau]\!]_S$ *definable* in PCF_S if and only if $F = [\![M]\!]_S$ for some closed term $M : \tau$. Call a value $v \in [\![\tau]\!]^\circ_S$ definable if and only if $val_S\,v$ is definable. If $[> \frac{1}{4}]$ were definable as a term $P : \mathtt{Unit} \to \mathtt{Unit}$, then $[\![M_{8.5}P]\!]_S\ \blacklozenge = [\![M_{8.5}]\!]_S\,(f \mapsto f([\![P]\!]_S)\ \blacklozenge) = [\![M_{8.5}]\!]^\circ_S\,([\![P]\!]_S)\ \blacklozenge = [\![M_{8.5}]\!]^\circ_S\,(val_S[> \frac{1}{4}])\ \blacklozenge$, and similarly for $N_{8.5}$. The last part of the proof shows that $[\![M_{8.5}]\!]^\circ_S\,(val_S[> \frac{1}{4}])\ \blacklozenge \not\leq [\![N_{8.5}]\!]^\circ_S\,(val_S[> \frac{1}{4}])\ \blacklozenge$, but we had shown earlier that $[\![M_{8.5}P]\!]_S = [\![P(\Omega_{\mathtt{Unit}} \oplus \underline{*})]\!]_S \leq [\![P(\Omega_{\mathtt{Unit}}) \oplus \underline{*}]\!]_S = [\![N_{8.5}P]\!]_S$, contradiction.

This fact generalizes to $[> b]$ for every $b \neq 1$, not just $b = \frac{1}{4}$:

Proposition 8.6 (Failure of Definability). *Let $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$, and assume $\mathtt{P} \in S$. No value $[> b] \in [\![\mathtt{Unit} \to \mathtt{Unit}]\!]^\circ_S$ is definable, for any $b \in [0, 1)$.*

The assumption $\mathtt{P} \in S$ is important: if $\mathtt{P} \notin S$, then $[\![\mathtt{Unit}]\!]_S$ is isomorphic to $\{*\}$ (Lemma 8.4), there are only three values in $[\![\mathtt{Unit} \to \mathtt{Unit}]\!]^\circ_S$, and they are definable as $\lambda x_{\mathtt{Unit}} \,.\, \Omega_{\mathtt{Unit}}$, $\lambda x_{\mathtt{Unit}} \,.\, x_{\mathtt{Unit}}$, and $\lambda x_{\mathtt{Unit}} \,.\, \underline{*}$ respectively.

Proof. As we saw in the course of the proof of Proposition 8.5, if $[> b]$ were definable, then we would have $[> b]\ (\triangleright)^\circ_{\mathtt{Unit}\to\mathtt{Unit}}\ [> b]$, where $\triangleright$ is the relation defined by $a_1 \triangleright a_2$ if and only if $a_2 \leq (1-a)a_1 + a$, and where $a \in [0, 1]$ is a fixed, but arbitrary constant. Since $[> b]\ (\triangleright)^\circ_{\mathtt{Unit}\to\mathtt{Unit}}\ [> b]$, for all $\alpha_1 \triangleright \alpha_2$ (hence $\alpha_1\, val_S *\ (\triangleright)_{\mathtt{Unit}}\ \alpha_2\, val_S *$), $[> b](\alpha_1\, val_S *)\ (\triangleright)_{\mathtt{Unit}}\ [> b](\alpha_2\, val_S *)$. Now pick any a in $(b, 1)$, and let $\alpha_1 = 0$, $\alpha_2 = a$. Then $\alpha_1 \triangleright \alpha_2$, but $[> b](\alpha_1\, val_S *) = 0$, $[> b](\alpha_2\, val_S *) = val_S *$, and it is wrong that $0\ (\triangleright)_{\mathtt{Unit}}\ val_S *$, since $1 \not\leq (1-a).0 + a$. So $[> b]$ is not definable in PCF_S if $\mathtt{P} \in S$. □

9. PCF_S Plus Statistical Termination Testers

To repair this, we shall therefore add operations $\bigcirc_{>b}$ to PCF_S whose semantics will be the missing functions $[> b]$, $0 < b < 1$. Given any closed term $M : \mathtt{Unit}$, if $\mathtt{P} \in S$ then $[\![M]\!]_S$ is a prevision of the form $\alpha\, val_S *$, and α is the probability that M terminates. $[> b](\alpha\, val_S *)$ tests whether the probability that M terminates is larger than b, and this is also the semantics we shall associate with $\bigcirc_{>b} M$. We call the $\bigcirc_{>b}$ operators *statistical termination testers*.

A theoretician's implementation of $\bigcirc_{>b} M$ would consist in guessing a derivation of $\_ \cdot M \downarrow^m b$, and checking that it is well-formed. If one is found, then the computation may proceed, else we loop forever. This amounts to *simulating* the recursively enumerable set of all possible runs of M, and can be implemented in a more practical way by counting the proportion of those runs of M that terminate, say within N execution steps, and checking whether this proportion exceeds b in the limit $N \to +\infty$. We omit the details, which are out of scope of the current paper.

The operator $\bigcirc_{>b}$ is related to, and inspired from, Escardó's operator $\bigcirc$ [9, Section 4.2]. In that paper, Escardó has another, clever strategy for semi-deciding testing, and compiles his non-deterministic and probabilistic extensions of PCF to (a fragment of) Real PCF, an extension of PCF with

real numbers. Using a similar compilation strategy, we might implement $\bigcirc_{>b} M$ as $b < \bigcirc M$, where $<$ is the Real PCF comparison operator on $\mathbb{I}$.

Call $\mathrm{PCF}_S + \bigcirc$ the natural extension of PCF_S with $\bigcirc_{>b}$ operators, $b \in \mathbb{Q} \cap (0,1)$. The types of $\mathrm{PCF}_S + \bigcirc$ are as for PCF_S. The definition of the terms of $\mathrm{PCF}_S + \bigcirc$ is as for PCF_S (see Section 5.1), except for the replacement of "PCF_S by "$\mathrm{PCF}_S + \bigcirc$" throughout, and for the additional clause:

- for every rational number b in $(0,1)$, for every $\mathrm{PCF}_S + \bigcirc$ term $M : \mathtt{Unit}$, $\bigcirc_{>b} M$ is a $\mathrm{PCF}_S + \bigcirc$ term, of type $\mathtt{Unit}$.

The operational semantics is as for PCF_S (see Figure 5), with the following additional rule:

$$\frac{\_ \cdot M \downarrow^m b \qquad C \cdot \underline{*} \downarrow^m a}{C \cdot \bigcirc_{>b} M \downarrow^m a} (\bigcirc)$$

($a \in \mathbb{Q} \cap [0,1)$, $b \in \mathbb{Q} \cap (0,1)$). We adapt Definition 5.2, and (re)define $\Pr(C \cdot M\downarrow^m)$ as $\sup\{a \in \mathbb{Q} \in [0,1) \mid C \cdot M \downarrow^m a \text{ derivable}\}$, where this time "derivable" means derivable in the extended system of rules.

We also extend the denotational semantics accordingly, by letting:

$$[\![\bigcirc_{>b} M]\!]_S \, \rho = \begin{cases} val_S * & \text{if } [\![M]\!]_S \, \rho \blacklozenge > b \\ 0 & \text{otherwise} \end{cases}$$

Soundness is established as for PCF_S (Lemma 6.1); note in particular that the evaluation contexts of $\mathrm{PCF}_S + \bigcirc$ are exactly those of PCF_S (we are *not* adding any elementary evaluation context of the form $[\bigcirc_{>b} \_]$), and the proof runs exactly as in PCF_S.

The situation is similar with computational adequacy. Using the notations used in the proof of Proposition 7.2, the only new case to be considered consists in showing that, if $M \; R_{\mathtt{Unit}} \; F$, then $\bigcirc_{>b} M \; R_{\mathtt{Unit}} \; G$, where $G = val_S * $ if $F(\blacklozenge) > b$ and $G = 0$ otherwise. Assume $C \; R^{\perp}_{\mathtt{Unit}} \; h$, and let us show that $\Pr(C \cdot \bigcirc_{>b} M\downarrow^m) \geq G(h)$. This is clear if $G = 0$. Otherwise, $G(h) = val_S * \, (h) = h(*)$, and $F(\blacklozenge) > b$. Since $\_ \; R^{\perp}_{\mathtt{Unit}} \; \blacklozenge$ (fact (†) in the proof of Proposition 7.2), $M \; R_{\mathtt{Unit}} \; F$ implies that $\Pr(\_ \cdot M\downarrow^m) \geq F(\blacklozenge) > b$. So $\_ \cdot M\downarrow^m b$ is derivable. Using rule $(\bigcirc)$ above, for every $a \in \mathbb{Q} \cap [0,1)$, if $C \cdot \underline{*}\downarrow^m a$ is derivable then so is $C \cdot \bigcirc_{>b} M \downarrow^m a$. So $\Pr(C \cdot \bigcirc_{>b} M\downarrow^m) \geq \Pr(C \cdot \underline{*}\downarrow^m)$. Since $C \; R^{\perp}_{\mathtt{Unit}} \; h$, for all $N \; R^{\circ}_{\mathtt{Unit}} \; v$, $\Pr(C \cdot N\downarrow^m) \geq h(v)$. This certainly holds for $N = \underline{*}$ and $v = *$, so $\Pr(C \cdot \bigcirc_{>b} M\downarrow^m) \geq \Pr(C \cdot \underline{*}\downarrow^m) \geq h(*) = G(h)$, and we conclude. To sum up:

Proposition 9.1 (Computational Adequacy). *Let $S \subseteq \{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$. Let M be a $PCF_S + \bigcirc$ term of type* `Unit`*. Then:*

$(S \subseteq \{\mathtt{A}, \mathtt{P}\})$ $Pr(\_ \cdot M{\downarrow}^{\mathsf{may}}) = [\![M]\!]_S$ $\blacklozenge$.

$(S \subseteq \{\mathtt{D}, \mathtt{P}\})$ $Pr(\_ \cdot M{\downarrow}^{\mathsf{must}}) = [\![M]\!]_S$ $\blacklozenge$.

$(\{\mathtt{A}, \mathtt{D}\} \subseteq S)$ $Pr(\_ \cdot M{\downarrow}^{\mathsf{must}}) = F^-(\blacklozenge)$ *and* $Pr(\_ \cdot M{\downarrow}^{\mathsf{must}}) = F^+(\blacklozenge)$*, where* $(F^-, F^+) = [\![M]\!]_S$.

10. Extra Topological Facts

Before we start proving any full abstraction result, we shall need to characterize the Scott topologies on the spaces we are interested in. On spaces of previsions, it will be important to note that they coincide with the weak topologies, defined as follows. We deal only with the cases $S \subseteq \{\mathtt{A}, \mathtt{P}\}$, since these are the only ones we shall be interested in in this paper, as far as full abstraction is concerned.

Definition 10.1 (Weak Topology). *For any space Y of previsions on X, define the* weak topology *on Y as the topology generated by the subsets $[h > b] = \{F \in Y \mid F(h) > b\}$, where $h \in [X \to \mathbb{I}]$ and $b \in \mathbb{R}^+$. When $Y = \mathbb{P}_S(X)$, we write $[h > b]_S$ for $[h > b]$, and when $Y = \mathbb{P}^1_S(X)$, we write $[h > b]^1_S$ for $[h > b]$ in case there is any need to make the ambient space clear.*

We write $\mathbb{P}_{S\,wk}(X)$, resp., $\mathbb{P}^1_{S\,wk}(X)$ for the space $\mathbb{P}_S(X)$, resp., $\mathbb{P}^1_S(X)$ with its weak topology.

We first explore the purely probabilistic case. The weak topology on $\mathbb{P}_{\mathtt{P}}(X)$ transports through the isomorphism between $\mathbb{P}_{\mathtt{P}}(X)$ and $\mathbf{V}_{\leq 1}(X)$ to a topology on $\mathbf{V}_{\leq 1}(X)$ that we call the weak topology on $\mathbf{V}_{\leq 1}(X)$, following Alvarez-Manilla *et al.* [3]. This is defined from the subbasic opens $[h > b] = \{\nu \in \mathbf{V}_{\leq 1}(X) \mid \int_{x \in X} h(x) d\nu > b\}$, $h \in [X \to \mathbb{I}]$, $b \in \mathbb{R}^+$. The following lemma is due to the above authors, who note that the weak topology on $\mathbf{V}_{\leq 1}(X)$ coincides with the so-called product topology, whose subbasic open subsets are $\{\nu \in \mathbf{V}_{\leq 1}(X) \mid \nu(U) > b\}$. Since it is short, we give a direct argument.

Lemma 10.2 (Alvarez-Manilla, Jung and Keimel). *Let X be any topological space. A subbase of the weak topology on $\mathbb{P}_{\mathtt{P}}(X)$ is given by the subsets of the form $[\chi_U > b]$, U open in X, $b \in \mathbb{R}^+$.*

PROOF. Every continuous map h is the sup of a directed family of step functions $h_m = \sum_{i=1}^{m} a_{im}\chi_{U_{im}}$ (e.g., taking $U_{im} = h^{-1}(i/m, 1]$, and $a_{im} = 1/m$), so that $[h > b] = \bigcup_{m=1}^{+\infty}[h_m > b]$, while $[\sum_{i=1}^{m} a_{im}\chi_{U_{im}} > b] = \bigcup_{\substack{b_0,b_1,\ldots,b_m \in \mathbb{I} \\ \sum_{i=1}^{m} a_{im}b_i \geq b}} \bigcap_{i=1}^{m}[\chi_{U_{im}} > b_i]$, using the fact that $\mathbb{P}_{\mathtt{P}}(X)$ is a space of linear maps and that no a_i is equal to zero. □

The *Kirch-Tix Theorem* states that the weak topology on $\mathbf{V}_{\leq 1}(X)$ coincides with the Scott topology when X is a continuous dcpo. See [3], who attribute it to Tix [43, Satz 4.10], who in turn attributes it to Kirch [25, Satz 8.6]; the latter two prove it for unbounded continuous valuations, not subprobability valuations, but the argument is unchanged in our case. Again up to the isomorphism between $\mathbb{P}_{\mathtt{P}}(X)$ and $\mathbf{V}_{\leq 1}(X)$, we obtain our first result on coincidence of topologies.

Proposition 10.3 (Kirch, Tix). *Let X be a continuous dcpo. The Scott and the weak topologies coincide on $\mathbb{P}_{\mathtt{P}}(X)$.*

The purely angelic case is even simpler.

Lemma 10.4. *Let X be a topological space. A subbase of the weak topology on $\mathbb{P}_{\mathtt{A}}(X)$, resp. $\mathbb{P}^1_{\mathtt{A}}(X)$, is given by the subsets $[\chi_U > 0]$, U open in X.*

PROOF. Let $h \in [X \to \mathbb{I}]$, and write h as the sup of a directed family of step functions $h_m = \sup_{i=1}^{m} a_{im}\chi_{U_{im}}$, where $U_{im} = h^{-1}(i/m, 1]$, and $a_{im} = i/m$. Then $[h > b]$ equals $\bigcup_{m=1}^{+\infty}[h_m > b]$. Now, given any F in $\mathbb{P}_{\mathtt{A}}(X)$, resp. $\mathbb{P}^1_{\mathtt{A}}(X)$, there is a closed subset C of X (resp., a non-empty closed subset C of X) such that $F = (h \mapsto \sup_{x \in C} h(x))$, see Remark 4.9. F belongs to $[h_m > b]$ if and only if there is an $x \in C$ and an i, $1 \leq i \leq m$, such that $a_{im} > b$ and $x \in U_{im}$. Moreover, F belongs to $[\chi_{U_{im}} > 0]$ if and only if there is an $x \in C$ such that $x \in U_{im}$. So $[h_m > b] = \bigcup_{\substack{1 \leq i \leq m \\ a_{im} > b}}[\chi_{U_{im}} > 0]$. □

This leads us to our second result on coincidence of topologies. Up to a few inessential details, this is due to Schalk [41, Section 6.3.3].

Proposition 10.5 (Schalk). *Let X be a continuous dcpo. The Scott and the weak topologies coincide on $\mathbb{P}_{\mathtt{A}}(X)$, and also on $\mathbb{P}^1_{\mathtt{A}}(X)$.*

PROOF. Up to the isomorphism between $\mathcal{H}(X)$ (resp., $\mathcal{H}_\perp(X)$) and $\mathbb{P}^1_{\mathtt{A}}(X)$ (resp., $\mathbb{P}_{\mathtt{A}}(X)$) that maps C to $(h \mapsto \sup_{x\in C} h(x))$ (see Remark 4.9), the weak topology on the latter transports to a topology on the former, whose subbasic open sets are of the form $\Diamond U = \{C \mid \sup_{x\in C} \chi_U(x) > 0\} = \{C \mid C \cap U \neq \emptyset\}$, U open in X. This topology is known as the *lower Vietoris topology* on $\mathcal{H}(X)$ (resp., $\mathcal{H}_\perp(X)$). When X is a continuous dcpo, the lower Vietoris topology coincides with the Scott topology. Indeed, clearly $\Diamond U$ is Scott-open, and conversely, a subbase of the Scott topology on $\mathcal{H}(X)$ (resp., $\mathcal{H}_\perp(X)$) is given by the subsets of the form $\Uparrow(\downarrow E)$, E finite and non-empty (resp., E finite). For every $C \in \mathcal{H}(X)$, $\downarrow E \ll C$ if and only if $\downarrow E \subseteq \Downarrow C$, as a consequence of [10, Corollary IV-8.7], and it is easy to see that a similar result holds in $\mathcal{H}_\perp(X)$. So $\Uparrow(\downarrow E) = \bigcap_{x\in E} \Diamond(\Uparrow x)$. □

We shall also need the following later.

Lemma 10.6. *Let X be a topological space. In $\mathbb{P}_{\mathtt{A}}(X)$ or in $\mathbb{P}^1_{\mathtt{A}}(X)$, $[\chi_- > 0]_{\mathtt{A}}$ commutes with unions: if U is a union of a family of opens U_i of X, then $[\chi_U > 0]_{\mathtt{A}} = \bigcup_{i\in I} [\chi_{U_i} > 0]_{\mathtt{A}}$;*

PROOF. Up to the aforementioned isomorphisms, this amounts to checking that $\Diamond(\bigcup_{i\in I} U_i) = \bigcup_{i\in I} \Diamond U_i$, i.e., that a closed subset C intersects $\bigcup_{i\in I} U_i$ if and only if it intersects some U_i. □

We turn to the case $S = \mathtt{AP}$. Our third, and final theorem on coincidence of topologies, is the following. A topological space is coherent if and only if the intersection of any two compact saturated subsets is again compact. For example, every bc-domain (and more generally, every FS-domain, see [2, Theorem 4.2.18]) is not only continuous, but also coherent.

Proposition 10.7 ([15], Proposition 3.42). *Let X be a continuous, coherent dcpo. The Scott topology coincides with the weak topology on $\mathbb{P}_{\mathtt{AP}}(X)$; also on $\mathbb{P}^1_{\mathtt{AP}}(X)$ if X is additionally assumed to be pointed.*

For the sake of completeness, here is an idea of the proof, for $\mathbb{P}_{\mathtt{AP}}(X)$. Write $\mathcal{H}_{\mathcal{V}}(Y)$ for $\mathcal{H}(Y)$ with its lower Vietoris topology. We show that $\mathbb{P}_{\mathtt{AP}\,wk}(X)$ is a retract of $\mathbb{P}^1_{\mathtt{A}\,wk}(\mathbb{P}_{\mathtt{P}\,wk}(X)) \cong \mathcal{H}_{\mathcal{V}}(\mathbb{P}_{\mathtt{P}\,wk}(X))$, through $r_{\mathtt{AP}} \colon F \mapsto (h \in [X \to \mathbb{I}] \mapsto \sup_{G\in F} G(h))$ [15, Proposition 3.11]. The same map defines a retract of $\mathcal{H}(\mathbb{P}_{\mathtt{P}}(X))$ onto $\mathbb{P}_{\mathtt{AP}}(X)$, that is, with respect to the Scott topologies instead of the weak topologies. Using the Kirch-Tix Theorem (Proposition 10.3)

and Schalk's Theorem (Proposition 10.5), $\mathcal{H}(\mathbb{P}_{\mathtt{P}}(X)) = \mathcal{H}_{\mathcal{V}}(\mathbb{P}_{\mathtt{P}\,wk}(X))$ (meaning that not only the underlying sets, but also the topologies, coincide), so $\mathbb{P}_{\mathtt{AP}}(X) = \mathbb{P}_{\mathtt{AP}\,wk}(X)$.

We also cite the following, which will come in handy later.

Proposition 10.8 ([15], Proposition 3.41). *Let X be a continuous, coherent dcpo. Then $\mathbb{P}_{\mathtt{AP}}(X)$ is a continuous dcpo, with basis given by the finite non-empty sups of simple linear previsions:*

$$h \mapsto \max_{i=1}^{m} \sum_{j=1}^{n_i} a_{ij} h(x_{ij})$$

where $m \geq 1$ and $\sum_{j=1}^{n_i} a_{ij} \leq 1$ for every i.

If X is a pointed continuous, coherent dcpo, then $\mathbb{P}^1_{\mathtt{AP}}(X)$ is a pointed continuous dcpo, with basis given by the finite sups of simple normalized previsions (i.e., $\sum_{j=1}^{n_i} a_{ij} = 1$ for every i). The least element is $h \mapsto h(\bot)$, where $\bot$ is the least element of X.

11. Full Abstraction

Redefine $\precsim^m_\sigma$ (Definition 8.1) on $\mathrm{PCF}_S + \bigcirc$ terms in the obvious way, by $M \precsim^m_\sigma N$ if and only if for every closed $\mathrm{PCF}_S + \bigcirc$ term $P : \sigma \to \mathtt{Unit}$, $\Pr(\_ \cdot PM{\downarrow^m}) \leq \Pr(\_ \cdot PN{\downarrow^m})$.

Given any two closed $\mathrm{PCF}_S + \bigcirc$ terms M, N of type σ, if $[\![M]\!]_S \leq [\![N]\!]_S$, necessarily $[\![PM]\!]_S \leq [\![PN]\!]_S$ for every closed term $P : \sigma \to \mathtt{Unit}$, as can be checked from the Definition (Figure 7), and the fact that all functions involved are monotonic. By computational adequacy (Proposition 9.1), $\Pr(\_ \cdot PM{\downarrow^m}) \leq \Pr(\_ \cdot PN{\downarrow^m})$ for any of the relevant modes m. Since P is arbitrary, $M \precsim^m_\sigma N$. So $[\![M]\!]_S \leq [\![N]\!]_S$ implies $M \precsim^m_\sigma N$. We embark on showing the converse implication.

The key idea is that, since $\leq$ is the specialization preorder of $[\![\sigma]\!]_S$, if $[\![M]\!]_S \not\leq [\![N]\!]_S$, then there is an open subset U that contains $[\![M]\!]_S$ but that does not contain $[\![N]\!]_S$. We can even take U from a well-chosen subbase of the Scott topology on $[\![\sigma]\!]_S$. We shall choose this subbase so that each of its elements can be defined in $\mathrm{PCF}_S + \bigcirc$, in a suitable sense.

In doing so, we shall also need to define a basis of the continuous dcpo $[\![\sigma]\!]_S$. The following easy and well-known lemma will help us.

Lemma 11.1. *Let X be a continuous poset, with a basis B_0. Let B be a subset of X such that every element of B_0 is the sup of a directed family of elements of B. Then B is also a basis of X.*

PROOF. For every $x \in X$, $x = \sup_{i \in I} x_i$ where $(x_i)_{i \in I}$ is the directed family of all elements of B_0 way-below x, and $x_i \ll x$ for every $i \in I$. Let us write x_i as the sup of some directed family $(x_{ij})_{j \in J_i}$ of elements of B. So x is the sup of the family $(x_{ij})_{i \in I, j \in J_i}$. Also, $x_{ij} \leq x_i \ll x$ for all $i \in I$, $j \in J_i$. It remains to show that this family is directed. Pick two elements $x_{i_1 j_1}$, $x_{i_2 j_2}$ from it. Since $x_{i_1}, x_{i_2} \ll x$, by interpolation, there is an element in B_0 way-below x (hence of the form x_i for some $i \in I$) such that $x_{i_1}, x_{i_2} \ll x_i$. Since $x_i = \sup_{j \in J_i} x_{ij}$, and the family $(x_{ij})_{j \in J_i}$ is directed, there is a $j \in J_i$ such that $x_{i_1}, x_{i_2} \leq x_{ij}$. In particular $x_{i_1 j_1}, x_{i_2 j_2} \leq x_{ij}$. □

To start with, here is the subbase $\mathcal{S}_{[X \to Y]}$ we choose for function spaces—the careful reader will notice that this implies that the Scott topology on $[X \to Y]$ coincides with the topology of pointwise convergence on $[X \to Y]$.

Proposition 11.2. *Let X and Y be bc-domains. Let B_X be a basis of X, $\mathcal{S}_X$ be a subbase of the Scott topology on X. Let B_Y be a basis of Y, and $\mathcal{S}_Y$ be a subbase of the Scott topology on Y. Then:*

- *The following set $B_{[X \to Y]}$ is a basis of $[X \to Y]$: $B_{[X \to Y]}$ is the set of all functions that can be written as the pointwise sup $\sup_{i=1}^{m} U_i \searrow y_i$, where each U_i is an intersection of finitely many elements from $\mathcal{S}_X$, $y_i \in B_Y$, and $U \searrow y$ denotes the map that sends each $x \in U$ to y and each $x \notin U$ to the bottom element $\bot$ of Y.*
- *The set $\mathcal{S}_{[X \to Y]}$ of all opens $[x \in V]$, $x \in B_X$, $V \in \mathcal{S}_Y$, is a subbase of the Scott topology on $[X \to Y]$. We write $[x \in V]$ for the open subset $\{f \in [X \to Y] \mid f(x) \in V\}$.*

PROOF. $[X \to Y]$ has a basis B_0 consisting of all step functions, which are by definition the finite pointwise sups (that exist) of step functions $U \searrow y$, U open in X, $y \in Y$ [10, Proposition II-4.20]. Let $\mathcal{B}_X$ be the base consisting of all finite intersections of opens from $\mathcal{S}_X$. To establish the first claim, using Lemma 11.1, it is enough to show that $U \searrow y$ is the sup of a directed family of elements from $\mathcal{B}_{[X \to Y]}$. By definition, U is the union of a family of elements U_i, $i \in I$, of $\mathcal{B}_X$, so U is also the union of the directed family

of opens $\bigcup_{i\in J} U_i$, where J ranges over the finite subsets of I. Also, y is the union of a directed family of elements y_k, $k \in K$, of B_Y. So $U \searrow y$ is the sup of the directed family of maps $\left(\bigcup_{i\in J} U_i\right) \searrow y_k$, J finite subset of I, $k \in K$. Now $\left(\bigcup_{i\in J} U_i\right) \searrow y_k = \sup_{i\in J} U_i \searrow y_k$ is in $B_{[X\to Y]}$.

For the second part, we use Lemma 5.16 of [13], which states that the subsets $[x \in V]$, $x \in X$, V open in Y, form a subbase of the topology of $[X \to Y]$, as soon as X is a continuous poset and Y is a bc-domain. Write x as the sup of the directed family $(x_k)_{k\in K}$ of elements of B_X. Then $[x \in V] = \bigcup_{k\in K}[x_k \in V]$: since f is Scott-continuous and V is open, $f(x) \in V$ is equivalent to the existence of $k \in K$ such that $f(x_k) \in V$. Write V as $\bigcup_{i\in I} \bigcap_{j\in J_i} V_{ij}$, where V_{ij} is in $\mathcal{B}_Y$ and each J_i is finite. Then $[x \in V] = \bigcup_{k\in K} \bigcup_{i\in I} \bigcap_{j\in J_i} [x_k \in V_{ij}]$, showing that $\mathcal{S}_{[X\to Y]}$ is indeed a subbase. □

In the cases of the base types `Unit` and `Nat`, we define the following. We take the opportunity to define a basis and a base of the topology of $\mathbb{I}$. This will serve in Proposition 11.4.

Definition 11.3. *A basis of $\{*\}$ is $B_{\{*\}} = \{*\}$, a subbase of its topology is $\mathcal{S}_{\{*\}} = \{\{*\}\}$.*

A basis of $\mathbb{N}$ is $B_{\mathbb{N}} = \mathbb{N}$, a subbase of its topology is $\mathcal{S}_{\mathbb{N}} = \{\{n\} \mid n \in \mathbb{N}\}$.

A basis of $\mathbb{I}$ is $B_{\mathbb{I}}$, the set of dyadic numbers in $[0, 1)$. A dyadic number *is one of the form $k/2^N$, $k, N \in \mathbb{N}$. A subbase of the Scott topology on $\mathbb{I}$ is $\mathcal{S}_{\mathbb{I}} = \{(b, 1] \mid b \in \mathbb{Q} \cap (0, 1)\}$.*

In the latter cases, we might take the elements of $B_{\mathbb{I}}$ rational, or take a subbase of opens of the form $(b, \mathbb{I}]$ with b dyadic instead. Choosing the elements of $B_{\mathbb{I}}$ dyadic will be needed for purposes of definability.

Proposition 11.4. *Let X be a continuous dcpo, B_X be a basis of X, and $\mathcal{S}_X$ be a subbase of the Scott topology on X. Then, for each of the following subsets S of $\{\mathtt{A}, \mathtt{D}, \mathtt{P}\}$, $B_{\mathbb{P}_S(X)}$ is a basis of $\mathbb{P}_S(X)$:*

- *($S = \mathtt{A}$) $B_{\mathbb{P}_\mathtt{A}(X)}$ is the set of discrete previsions of the form $(h \in [X \to \mathbb{I}] \mapsto \max(h(x_1), \ldots, h(x_n)))$, $n \geq 0$, $x_1, \ldots, x_n \in B_X$ (if $n = 0$, this is the constant zero map);*
- *($S = \mathtt{P}$) $B_{\mathbb{P}_\mathtt{P}(X)}$ is the set of linear previsions of the form $(h \in [X \to \mathbb{I}] \mapsto \sum_{i=1}^m a_i h(x_i))$, where $x_i \in B_X$, a_i is dyadic, and $\sum_{i=1}^m a_i \leq 1$;*
- *($S = \mathtt{AP}$) $B_{\mathbb{P}_\mathtt{AP}(X)}$ is the set of sublinear previsions of the form $(h \in [X \to \mathbb{I}] \mapsto \max_{i=1}^m G_i(h))$, where $m \geq 1$ and each G_i is in $B_{\mathbb{P}_\mathtt{P}(X)}$.*

In the case $S = \mathtt{A}$, $\mathcal{S}_{\mathbb{P}_\mathtt{A}(X)}$, defined as the family of all opens of the form $[\chi_U > 0]$, where U ranges over the finite intersections of elements of $\mathcal{S}_X$, is a subbase of the Scott topology on $\mathbb{P}_S(X)$.

In the last two (probabilistic) cases, $\mathcal{S}_{\mathbb{P}_S(X)}$, defined as the family of all opens of the form $[h > b]_S = \{F \in \mathbb{P}_S(X) \mid F(h) > b\}$, $h \in B_{[X \to \mathbb{I}]}$, $b \in \mathbb{Q} \cap (0,1)$, is a subbase of the Scott topology on $\mathbb{P}_S(X)$.

Moreover, $\mathbb{P}_\mathtt{A}(X)$ and $\mathbb{P}_\mathtt{AP}(X)$ are continuous lattices.

PROOF. ($S = \mathtt{A}$) We first observe that if Y is a continuous dcpo, with basis B_Y, then $\mathcal{H}(Y)$ has a basis given by the elements of the form $\downarrow E$, $E \subseteq Y$ finite and non-empty [10, Corollary IV-8.7]. One can even take $E \subseteq B_Y$ [17, Theorem 18.6.1], a fact that we may also derive using Lemma 11.1.

Now $Y = X_\perp$ is a continuous dcpo, so, using the isomorphism between $\mathcal{H}_\perp(Y)$ and $\mathcal{H}(Y_\perp)$ (which maps $C \in \mathcal{H}(Y_\perp)$ to $C \cap Y \in \mathcal{H}_\perp(Y)$), we deduce that $\mathcal{H}_\perp(X)$ has a basis of elements of the form $\downarrow E$, $E \subseteq X$ finite and possibly empty.

Recall from Proposition 4.8 that $\mathbb{P}_\mathtt{A}(X)$ is isomorphic to the lifted Hoare powerdomain $\mathcal{H}_\perp(X)$. The isomorphism maps each $C \in \mathcal{H}_\perp(X)$ to the discrete sublinear prevision $(h \in [X \to \mathbb{I}] \mapsto \sup_{x \in C} h(x))$. Through this isomorphism, the image of $\downarrow\{x_1, \ldots, x_n\}$ ($n \geq 0$) is $(h \in [X \to \mathbb{I}] \mapsto \max(h(x_1), \ldots, h(x_n)))$, so $B_{\mathbb{P}_\mathtt{A}(X)}$ is a basis.

Let us look at the topology. By Proposition 10.5 and Lemma 10.4, the subsets $[\chi_U > 0]$, U open in X, form a subbase of the topology of $\mathbb{P}_\mathtt{A}(X)$. Write U as a union $\bigcup_{i \in I} U_i$ where each U_i is a finite intersection of elements of $\mathcal{S}_X$. Then $[\chi_U > 0] = \bigcup_{i \in I} [\chi_{U_i} > 0]$ by Lemma 10.6. So $\mathcal{S}_{\mathbb{P}_\mathtt{A}(X)}$ is a subbase.

Finally, if X is a continuous dcpo, then $\mathbb{P}_\mathtt{A}(X) \cong \mathcal{H}_\perp(X)$ is not just a continuous dcpo, but a continuous lattice, because every finite union of closed sets is again a closed set.

($S = \mathtt{P}$) Recall that $\mathbb{P}_\mathtt{P}(X)$ is isomorphic to $\mathbf{V}_{\leq 1}(X)$, with the isomorphism mapping $F \in \mathbb{P}_\mathtt{P}(X)$ to $\nu \in \mathbf{V}_{\leq 1}(X)$ such that $\nu(U) = F(\chi_U)$, and conversely, maps ν to $F = (h \mapsto \int_{x \in X} h(x) d\nu)$. By the argument in the proof of Corollary 5.5 of [20], a basis of $\mathbf{V}_{\leq 1}(X)$ is given by the valuations of the form $\sum_{i=1}^m a_i \delta_{x_i}$, where each a_i is in $[0,1]$, $x_i \in B_X$, and $\sum_{i=1}^m a_i \leq 1$. (δ_x is the Dirac mass at x: $\delta_x(U)$ equals 1 if $x \in U$, 0 otherwise; the isomorphism maps it to the linear prevision $h \mapsto h(x)$.) So the elements of the form $(h \mapsto \sum_{i=1}^m a_i h(x_i))$, $x_i \in B_X$, $a_i \in [0,1]$, $\sum_{i=1}^m a_i \leq 1$, also form a basis of $\mathbb{P}_\mathtt{P}(X)$. (For future reference, call such elements *simple linear previsions*.)

Now $(h \mapsto \sum_{i=1}^m 2^n \lfloor \frac{a_i}{2^n} \rfloor h(x_i))_{n\in\mathbb{N}}$ forms a chain of elements of $B_{\mathbb{P}_{\mathtt{P}}(X)}$ whose sup is $h \mapsto \sum_{i=1}^m a_i h(x_i)$. So $B_{\mathbb{P}_{\mathtt{P}}(X)}$ is also a basis, using Lemma 11.1.

By Proposition 10.3 and Lemma 10.2, a subbase of $\mathbb{P}_{\mathtt{P}}(X)$ is given by the subsets $[h > b]$, $h \in [X \to \mathbb{I}]$, $b \in \mathbb{R}^+$. We can restrict to $b \in (0,1)$, since $[h > b]$ is empty for $b \geq 1$, to the whole of $\mathbb{P}_{\mathtt{P}}(X)$ for $b < 0$, and to $\bigcup_{n\geq 1}[h > 1/n]$ for $b = 0$. Since h is the sup of a directed family $(h_i)_{i\in I}$ in $B_{[X\to\mathbb{I}]}$, and b is the infimum of a family of rational numbers $(b_j)_{j\in J}$ in $(0,1)$, $[h > b] = \bigcup_{i\in I, j\in J}[h_i > b_j]$, so $\mathcal{S}_{\mathbb{P}_{\mathtt{P}}(X)}$ is a subbase of the topology of $\mathbb{P}_{\mathtt{P}}(X)$.

($S = \mathtt{AP}$) By Proposition 10.8, a basis of $\mathbb{P}_{\mathtt{AP}}(X)$ is given by the previsions of the form $\max_{i=1}^m \sum_{j=1}^{n_i} a_{ij} h(x_{ij})$, where $m \geq 1$ and $\sum_{j=1}^{n_i} a_{ij} \leq 1$ for every i. As in the $S = \mathtt{P}$ case, using Lemma 11.1, we also obtain a basis by requiring that each a_{ij} is dyadic and $x_{ij} \in B_X$. The resulting basic elements are exactly those of $B_{\mathbb{P}_{\mathtt{AP}}(X)}$.

By Proposition 10.7, the topology of $\mathbb{P}_{\mathtt{AP}}(X)$ is the weak topology. As in the $S = \mathtt{P}$ case, we can restrict b to $(0,1)$ in taking subbasic open subsets $[h > b]$, $h \in [X \to \mathbb{I}]$. Since h is the sup of a directed family $(h_i)_{i\in I}$ in $B_{[X\to\mathbb{I}]}$, and b is the infimum of a family of rational numbers $(b_j)_{j\in J}$ in $[0,1)$, $[h > b] = \bigcup_{i\in I, j\in J}[h_i > b_j]$, so $\mathcal{S}_{\mathbb{P}_{\mathtt{AP}}(X)}$ is a subbase of the topology of $\mathbb{P}_{\mathtt{AP}}(X)$.

Finally, $\mathbb{P}_{\mathtt{AP}}(X)$ is also a continuous lattice, not just a continuous dcpo, because the pointwise supremum of any family of sublinear previsions is again a sublinear prevision, as one checks easily. □

Recall that an nbc-domain is a not necessarily pointed continuous dcpo in which every pair of elements that has an upper bound has a least upper bound.

Corollary 11.5. *Let $S = \mathtt{A}$ or $S = \mathtt{AP}$. For every type τ: (i) $[\![\tau]\!]_S^\circ$ is an nbc-domain; (ii) $[\![\tau]\!]_S$ is a continuous lattice.*

Proof. By structural induction on τ, using the fact that $\{*\}$ and $\mathbb{N}$ are nbc-domains, that $\mathbb{P}_S(X)$ is a continuous lattice whenever X is an nbc-domain (Proposition 11.4, last part), and that $[X \to Y]$ is an nbc-domain, whenever X and Y are continuous lattices. The latter follows from [8, Proposition 2], which shows in particular that $[X \to L]$ is a bc-domain for every locally compact space X and every bc-domain L. □

The *values* are defined in $\mathrm{PCF}_S + \bigcirc$ as in PCF_S: $\underline{*}$ is the only value of type $\mathtt{Unit}$, the terms $\underline{n}$, $n \in \mathbb{N}$, are the values of type $\mathtt{Nat}$, and the values of type $\sigma \to \tau$ are the $\mathrm{PCF}_S + \bigcirc$ terms of the form $\lambda x_\sigma . M$, where $M : \tau$.

Lemma 11.6 (Definability, Case $S = \mathtt{AP}$). *For every type τ,*

1. *Every element v of $B_{[\![\tau]\!]^\circ_{\mathtt{AP}}}$ is definable by a value: there is a closed value $V : \tau$ such that $[\![V]\!]^\circ_{\mathtt{AP}} = v$.*
2. *Every open subset U in $\mathcal{S}_{[\![\tau]\!]^\circ_{\mathtt{AP}}}$ is definable by a value: there is a closed value $V : \tau \to \mathtt{Unit}$ such that, for every $v \in [\![\tau]\!]^\circ_{\mathtt{AP}}$, $[\![V]\!]^\circ_{\mathtt{AP}}(val_{\mathtt{AP}}\, v)$ is equal to $val_{\mathtt{AP}} *$ if $v \in U$, and to 0 otherwise.*
3. *Every element F of $B_{[\![\tau]\!]_{\mathtt{AP}}}$ is definable: there is a closed $PCF_{\mathtt{AP}} + \bigcirc$ term $M : \tau$ such that $[\![M]\!]_{\mathtt{AP}} = F$.*
4. *Every open subset U in $\mathcal{S}_{[\![\tau]\!]_{\mathtt{AP}}}$ is definable by a value: there is a closed value $V : \tau \to \mathtt{Unit}$ such that, for every $F \in [\![\tau]\!]_{\mathtt{AP}}$, $[\![V]\!]^\circ_{\mathtt{AP}}(F)$ is equal to $val_{\mathtt{AP}} *$ if $F \in U$, and to 0 otherwise.*

Proof. Using Corollary 11.5, $[\![\tau]\!]^\circ_{\mathtt{AP}}$ and $[\![\tau]\!]_{\mathtt{AP}}$ are continuous dcpos for every type τ, so that Proposition 11.4 applies to these spaces, a fact we shall use several times.

We note that 3 and 4 follow from 1 and 2.

$1 \Rightarrow 3$. Let $F \in B_{[\![\tau]\!]_{\mathtt{AP}}}$, and recall that $[\![\tau]\!]_{\mathtt{AP}} = \mathbb{P}_{\mathtt{AP}}([\![\tau]\!]^\circ_{\mathtt{AP}})$. By Proposition 11.4 (case $S = \mathtt{AP}$), F is of the form $(h \in [X \to \mathbb{I}] \mapsto \max_{i=1}^m G_i(h))$, where $m \geq 1$ and each G_i is of the form $(h \in [X \to \mathbb{I}] \mapsto \sum_{j=1}^{m_i} a_{ij} h(v_{ij}))$, $m_j \geq 0$, $v_{ij} \in B_{[\![\tau]\!]^\circ_{\mathtt{AP}}}$, $\sum_{j=1}^{m_i} a_{ij} \leq 1$, and each a_{ij} is dyadic. Write a_{ij} as $k_{ij}/2^n$, with the same n for all i, j, and where $k_{ij} \in \mathbb{N}$. By 1, there is a value V_{ij} such that $[\![V_{ij}]\!]^\circ_{\mathtt{AP}} = v_{ij}$.

Call *n-sum* any term of the form $M_1 \oplus M_2 \oplus \ldots \oplus M_{2^n}$. Let N_i be the n-sum of k_{i1} times the term V_{i1}, k_{i2} times the term V_{i2}, …, k_{im_i} times the term V_{im_i}, and $2^n - \sum_{j=1}^{m_i} k_{ij}$ times the term $\Omega_\tau = \mathtt{Y}(\lambda x_\tau . x)$. Since $[\![\Omega_\tau]\!]_{\mathtt{AP}} = 0$, and $[\![V_{ij}]\!]_{\mathtt{AP}} = val_{\mathtt{AP}}\, v_{ij} = (h \mapsto h(v_{ij}))$, $[\![N_i]\!]_{\mathtt{AP}} = (h \mapsto \frac{1}{2^n} \sum_{j=1}^{m_i} k_{ij} h(v_{ij})) = G_i$.

Let now M be $N_1 \mathbin{\bigcirc\!\!\!\!\vee} N_2 \mathbin{\bigcirc\!\!\!\!\vee} \ldots \mathbin{\bigcirc\!\!\!\!\vee} N_m$. Then $[\![M]\!]_{\mathtt{AP}} = F$, as desired.

$2 \Rightarrow 4$. Let $U \in \mathcal{S}_{[\![\tau]\!]_{\mathtt{AP}}}$. By Proposition 11.4 (case $S = \mathtt{AP}$), U is of the form $[h > b]$, where $h \in B_{[[\![\tau]\!]^\circ_{\mathtt{AP}} \to \mathbb{I}]}$ and b is in $\mathbb{Q} \cap (0,1)$. By Proposition 11.2, h can be written as $\sup_{i=1}^m U_i \searrow b_i$, where U_i is a finite intersection $\bigcap_{j=1}^{m_i} U_{ij}$ of elements U_{ij} of $\mathcal{S}_{[\![\tau]\!]^\circ_{\mathtt{AP}}}$, and b_i is a dyadic number in $[0,1)$. Write each b_i as $k_i/2^n$, with the same n, and with $k_1, \ldots, k_m \in \mathbb{N}$.

By 2, there is a closed value $V_{ij} = \lambda x\, .\, M_{ij}$ of type $\tau \to \mathtt{Unit}$ such that for every $v \in [\![\tau]\!]^\circ_{\mathtt{AP}}$, $[\![V_{ij}]\!]^\circ_{\mathtt{AP}}(val_{\mathtt{AP}}\, v)$ is equal to $val_{\mathtt{AP}} *$ if $v \in U_{ij}$, and to 0 otherwise.

We implement intersection by sequential composition. I.e., for terms M, N of type $\mathtt{Unit}$, define $M \wedge N$ as $\mathtt{let}\, y \Leftarrow M\ \mathtt{in}\ N$, where y is a dummy

variable, not free in M or N. Notice that if $[\![M]\!]_{\mathtt{AP}}\,\rho = a\,val_{\mathtt{AP}}\,*$ and $[\![N]\!]_{\mathtt{AP}}\,\rho = b\,val_{\mathtt{AP}}\,*$, then $[\![M \wedge N]\!]_{\mathtt{AP}}\,\rho = ab\,val_{\mathtt{AP}}\,*$. In particular, if $a, b \in \{0,1\}$, then ab is the logical and of a and b.

Let $M_i = M_{i1} \wedge M_{i2} \wedge \ldots \wedge M_{im_i} \wedge N_i$, where N_i is the n-sum of k_i times $\underline{*}$ and $2^n - k_i$ times $\Omega_{\mathtt{Unit}}$. For every $v \in [\![\tau]\!]^\circ_{\mathtt{AP}}$, $[\![M_i]\!]_{\mathtt{AP}}\,[x := val_{\mathtt{AP}}\,v]\ \blacklozenge$ is equal to $k_i/2^n = b_i$ if $v \in U_i$, to 0 otherwise. In other words, M_i implements $U_i \searrow b_i$.

Let $M = M_1 \mathbin{\mathbb{O}} \ldots \mathbin{\mathbb{O}} M_m$. Since the semantics of $\mathbb{O}$ is a sup, M implements h in the sense that for every $v \in [\![\tau]\!]^\circ_{\mathtt{AP}}$, $[\![M]\!]_{\mathtt{AP}}\,[x := val_{\mathtt{AP}}\,v]\ \blacklozenge = h(v)$.

We now define the value $V = \lambda z\,.\bigcirc_{>b} \mathtt{let}\,x \Leftarrow z\ \mathtt{in}\ M$, where z is a fresh variable of type τ, different from x and not free in M. (This is the unique place where we need the $\bigcirc_{>b}$ operator.) For every $F \in [\![\tau]\!]_{\mathtt{AP}}$, $[\![V]\!]^\circ_{\mathtt{AP}}(F) = [\![\bigcirc_{>b} \mathtt{let}\,x \Leftarrow z\ \mathtt{in}\ M]\!]_{\mathtt{AP}}\,[z := F]$ is equal to $val_{\mathtt{AP}}\,*$ if $[\![\mathtt{let}\,x \Leftarrow z\ \mathtt{in}\ M]\!]_{\mathtt{AP}}\,[z := F]\ \blacklozenge > b$, and to 0 otherwise. However:

$$\begin{aligned}[\![\mathtt{let}\,x \Leftarrow z\ \mathtt{in}\ M]\!]_{\mathtt{AP}}\,[z := F]\ \blacklozenge &= (let_{\mathtt{AP}}\,v \Leftarrow F\ in\ [\![M]\!]_{\mathtt{AP}}\,[x := val_{\mathtt{AP}}\,v])\,(\blacklozenge)\\ &= F(v \mapsto [\![M]\!]_{\mathtt{AP}}\,[x := val_{\mathtt{AP}}\,v]\ \blacklozenge) = F(h).\end{aligned}$$

So $[\![V]\!]^\circ_{\mathtt{AP}}(F)$ is equal to $val_{\mathtt{AP}}\,*$ if $F \in [h > b]$, and to 0 otherwise.

We now prove 1 and 2 by induction on τ. When $\tau = \mathtt{Unit}$: 1. $v = *$, so we take $V = \underline{*}$; 2. $U = \{*\}$, so we take V equal to $\lambda x_{\mathtt{Unit}}\,.\,\underline{*}$. When $\tau = \mathtt{Nat}$: 1. $v = n$ for some $n \in \mathbb{N}$, and we take $V = \underline{n}$; 2. by Definition 11.3, $U = \{n\}$ for some $n \in \mathbb{N}$, and we take $V = \lambda x_{\mathtt{Nat}}\,.\,\mathtt{ifz}\ \mathtt{pred}^n\,x\ \underline{*}\ \Omega_{\mathtt{Unit}}$, where $\mathtt{pred}^n$ is defined by $\mathtt{pred}^0\,M = M$, $\mathtt{pred}^{n+1}\,M = \mathtt{pred}\,(\mathtt{pred}^n\,M)$.

On function types $\sigma \to \tau$:

1. Every element of $B_{[\![\sigma\to\tau]\!]^\circ_{\mathtt{AP}}}$ is of the form $\sup_{i=1}^m U_i \searrow y_i$, where each U_i is an intersection of finitely many elements $U_{i1}, \ldots, U_{im_i}$ from $\mathcal{S}_{[\![\sigma]\!]_S}$, $y_i \in B_{[\![\tau]\!]_S}$ (Proposition 11.2). We proceed as in the construction of a term denoting h in the proof of 4 above. By induction hypothesis, each U_{ij} is defined by a closed value $\lambda x\,.\,M_{ij} : \sigma \to \mathtt{Unit}$, and y_i is defined by a closed value $V_i : \tau$. Let $M_i = \mathtt{let}\,y \Leftarrow M_{i1} \wedge \ldots \wedge M_{im_i}\ \mathtt{in}\ V_i$, where y is a dummy variable ($M_i = V_i$ if $m_i = 0$). For every $v \in [\![\sigma]\!]^\circ_{\mathtt{AP}}$, $[\![M_i]\!]\,[x := val_{\mathtt{AP}}\,v]$ is equal to y_i if $v \in U_i$, and to the zero prevision otherwise. So $M = M_1 \mathbin{\mathbb{O}} \ldots \mathbin{\mathbb{O}} M_m$ implements $\sup_{i=1}^m U_i \searrow y_i$, in the sense that for every $v \in [\![\sigma]\!]^\circ_{\mathtt{AP}}$, $[\![M]\!]\,[x := val_{\mathtt{AP}}\,v]$ is equal to $(\sup_{i=1}^m U_i \searrow y_i)(v)$. Finally, we take $V = \lambda x\,.\,M$.

2. Let U be an open subset, in $\mathcal{S}_{[\![\sigma\to\tau]\!]^\circ_{\mathtt{AP}}}$. By Proposition 11.2, U is of the form $[G \in W]$, where $G \in B_{[\![\sigma]\!]_{\mathtt{AP}}}$ and $W \in \mathcal{S}_{[\![\tau]\!]_{\mathtt{AP}}}$. By induction hypothesis, there is a closed term $M : \sigma$ such that $[\![M]\!]_{\mathtt{AP}} = G$, and there is a closed value

$V : \tau \to \mathtt{Unit}$ such that for every $F \in [\![\tau]\!]_{\mathtt{AP}}$, $[\![V]\!]^{\circ}_{\mathtt{AP}}(F)$ is equal to $val_{\mathtt{AP}} *$ if $F \in W$, to 0 otherwise. Consider the value $\lambda x_{\sigma\to\tau}.V(xM)$, where x is a fresh variable. We claim this is the desired value. Indeed, for every $f \in [\![\sigma \to \tau]\!]^{\circ}_{\mathtt{AP}}$,

$$\begin{array}{rcl}
[\![\lambda x_{\sigma\to\tau}\,.\,V(xM)]\!]^{\circ}_{\mathtt{AP}}(val_{\mathtt{AP}}\, f) & = & [\![V(xM)]\!]_{\mathtt{AP}}[x := val_{\mathtt{AP}}\, f] \\
& = & let_{\mathtt{AP}}\, h \Leftarrow [\![V]\!]_{\mathtt{AP}}\ in\ h([\![xM]\!]_{\mathtt{AP}}[x := val_{\mathtt{AP}}\, f]) \\
& = & [\![V]\!]^{\circ}_{\mathtt{AP}}([\![xM]\!]_{\mathtt{AP}}[x := val_{\mathtt{AP}}\, f]) \\
& = & [\![V]\!]^{\circ}_{\mathtt{AP}}(let_{\mathtt{AP}}\, h \Leftarrow val_{\mathtt{AP}}\, f\ in\ h([\![M]\!]_{\mathtt{AP}})) \\
& = & [\![V]\!]^{\circ}_{\mathtt{AP}}(f([\![M]\!]_{\mathtt{AP}})) = [\![V]\!]^{\circ}_{\mathtt{AP}}(f(G))
\end{array}$$

and this is equal to $val_{\mathtt{AP}} *$ if $f(G)$ is in W, i.e., if $f \in [G \in W] = U$, and to 0 otherwise. □

Theorem 11.7 (Full Abstraction, Case $S = \mathtt{AP}$). *For all closed $PCF_{\mathtt{AP}}+\bigcirc$ terms M and N of type σ, $M \precsim^{\mathsf{may}}_{\sigma} N$ if and only if $[\![M]\!]_{\mathtt{AP}} \leq [\![N]\!]_{\mathtt{AP}}$.*

Proof. We have already seen that $[\![M]\!]_{\mathtt{AP}} \leq [\![N]\!]_{\mathtt{AP}}$ implies $M \precsim^{\mathsf{may}}_{\sigma} N$. Conversely, if $[\![M]\!]_{\mathtt{AP}} \not\leq [\![N]\!]_{\mathtt{AP}}$, there there is an open subset U, taken from $\mathcal{S}_{[\![\sigma]\!]}$, such that M is in U but N is not. By definability (Lemma 11.6, Item 4), there is a closed value $V : \sigma \to \mathtt{Unit}$ such that for every $F \in [\![\sigma]\!]_{\mathtt{AP}}$, $[\![V]\!]_{\mathtt{AP}}(F)$ is equal to $val_{\mathtt{AP}} *$ if $F \in U$, to 0 otherwise. Then $[\![VM]\!]_{\mathtt{AP}} = let_{\mathtt{AP}}\, f \Leftarrow [\![V]\!]_{\mathtt{AP}}\ in\ f([\![M]\!]_{\mathtt{AP}}) = [\![V]\!]^{\circ}_{\mathtt{AP}}([\![M]\!]_{\mathtt{AP}}) = val_{\mathtt{AP}} *$, while $[\![VN]\!]_{\mathtt{AP}} = [\![V]\!]^{\circ}_{\mathtt{AP}}([\![N]\!]_{\mathtt{AP}}) = 0$. By computational adequacy (Proposition 9.1), $\Pr(\_ \cdot VM{\downarrow}^{\mathsf{may}}) = 1$, $\Pr(\_ \cdot VN{\downarrow}^{\mathsf{may}}) = 0$, so it is certainly not the case that $M \precsim^{\mathsf{may}}_{\sigma} N$. □

The non-probabilistic case $S = \mathtt{A}$ is simpler, and yields a stronger result. A hint is that termination testers $\bigcirc_{>b}$ are definable in $PCF_{\mathtt{A}}$, by $\bigcirc_{>b} M = M$ (recall that we are dealing with discrete previsions here, which can only take the values 0 or 1 when evaluated on a $\{0,1\}$-valued function, such as ⧫).

Lemma 11.8 (Definability, Case $S = \mathtt{A}$). *For every type τ,*

1. *Every element v of $B^{\circ}_{[\![\tau]\!]_{\mathtt{A}}}$ is definable by a value: there is a closed value $V : \tau$ such that $[\![V]\!]^{\circ}_{\mathtt{A}} = v$.*
2. *Every open subset U in $\mathcal{S}^{\circ}_{[\![\tau]\!]_{\mathtt{A}}}$ is definable by a value: there is a closed value $V : \tau \to \mathtt{Unit}$ such that, for every $v \in [\![\tau]\!]^{\circ}_{\mathtt{A}}$, $[\![V]\!]^{\circ}_{\mathtt{A}}(val_{\mathtt{A}}\, v)$ is equal to $val_{\mathtt{A}} *$ if $v \in U$, and to 0 otherwise.*
3. *Every element F of $B_{[\![\tau]\!]_{\mathtt{A}}}$ is definable: there is a closed $PCF_{\mathtt{A}}$ term $M : \tau$ such that $[\![M]\!]_{\mathtt{A}} = F$.*

4. *Every open subset U in $\mathcal{S}_{[\![\tau]\!]_{\mathtt{A}}}$ is definable by a value: there is a closed value $V : \tau \to \mathtt{Unit}$ such that, for every $F \in [\![\tau]\!]_{\mathtt{A}}$, $[\![V]\!]^{\circ}_{\mathtt{A}}(F)$ is equal to $val_{\mathtt{A}} *$ if $F \in U$, and to 0 otherwise.*

PROOF. The proof is as for Lemma 11.6. Only the proof of 4, knowing 2, changes. This was the case where we required the operator $\bigcirc_{>b}$, and we will not need it any longer. Let $U \in \mathcal{S}_{[\![\tau]\!]_{\mathtt{A}}}$. By Proposition 11.4 (case $S = \mathtt{A}$), U is of the form $[\chi_W > 0]$, where W is a finite intersection $W_1 \cap \ldots \cap W_m$, with $W_i \in \mathcal{S}_{[\![\tau]\!]^{\circ}_{\mathtt{A}}}$. By 2, there is a value $\lambda x_\tau \,.\, M_i$ such that for every $v \in [\![\tau]\!]^{\circ}_{\mathtt{A}}$, $[\![\lambda x \,.\, M_i]\!]^{\circ}_{\mathtt{A}}(val_{\mathtt{A}}\, v)$ is equal to $val_{\mathtt{A}} *$ if $v \in W_i$, and to 0 otherwise. W is then definable as $\lambda x_\tau \,.\, M$, where $M = M_1 \wedge \cdots \wedge M_m$ (or $\underline{*}$ when $m = 0$), where $\wedge$ is defined as sequential composition at type $\mathtt{Unit}$, as in the proof of Lemma 11.6. That is, $[\![M]\!]_{\mathtt{A}}[x := val_{\mathtt{A}}\, v]$ is equal to $val_{\mathtt{A}} *$ if $v \in W$, and to 0 otherwise, or in other words, $[\![M]\!]_{\mathtt{A}}[x := val_{\mathtt{A}}\, v] = \chi_W(v).\, val_{\mathtt{A}} *$. The desired value, defining U, is then $V = \lambda z \,.\, let_{\mathtt{A}}\, x \Leftarrow z \ in \ M$. Indeed, $[\![V]\!]^{\circ}_{\mathtt{A}}(F) = [\![let_{\mathtt{A}}\, x \Leftarrow z \ in \ M]\!]_{\mathtt{A}}[z := F] = let_{\mathtt{A}}\, v \Leftarrow F \ in \ [\![M]\!]_{\mathtt{A}}[x := val_{\mathtt{A}}\, v] = let_{\mathtt{A}}\, v \Leftarrow F \ in \ \chi_W(v).\, val_{\mathtt{A}} * = (h \in [\{*\} \to \mathbb{I}] \mapsto F(v \mapsto \chi_W(v).\, val_{\mathtt{A}} *\, (h)) = (h \in [\{*\} \to \mathbb{I}] \mapsto F(h(*).\chi_W)) = F(\chi_W).\, val_{\mathtt{A}} *$. If $F \in U$, then $F(\chi_W) > 0$, which entails $F(\chi_W) = 1$ (see comments after Definition 4.3), hence $[\![V]\!]^{\circ}_{\mathtt{A}}(F) = val_{\mathtt{A}} *$; otherwise, $F(\chi_W) = 0$, so $[\![V]\!]^{\circ}_{\mathtt{A}}(F) = 0$. □

It follows that $\mathrm{PCF}_{\mathtt{A}}$ is fully abstract (even without $\bigcirc_{>b}$):

Theorem 11.9 (Full Abstraction, Case $S = \mathtt{A}$). *For all closed $PCF_{\mathtt{A}}$ terms M and N of type σ, $M \precsim^{\mathsf{may}}_{\sigma} N$ if and only if $[\![M]\!]_{\mathtt{A}} \leq [\![N]\!]_{\mathtt{A}}$.*

PROOF. The proof is as for Theorem 11.7, using Lemma 11.8 (and Theorem 7.4 for computational adequacy in $\mathrm{PCF}_{\mathtt{A}}$). □

We should also note that, through the isomorphism of Proposition 4.8 (i), $\mathrm{PCF}_{\mathtt{A}}$ has an equivalent semantics where $[\![\tau]\!]_{\mathtt{A}}$ is now defined as $\mathcal{H}_{\perp}([\![\tau]\!]^{\circ}_{\mathtt{A}})$, the dcpo of closed subsets of $[\![\tau]\!]^{\circ}_{\mathtt{A}}$; this semantics is given in Figure 8, and is the usual semantics one would expect for a variant of PCF with (angelic) non-deterministic choice. (The supremum of a family of closed subsets, which occurs several times there, is the closure of the union, not the union.) It follows that $\mathrm{PCF}_{\mathtt{A}}$ is also sound, adequate, and fully abstract for this semantics.

$$
\begin{array}{rcl}
[\![*]\!]^\circ_{\mathtt{A}}\,\rho = * \qquad [\![\underline{n}]\!]^\circ_{\mathtt{A}}\,\rho = n \qquad [\![\lambda x_\sigma\,.\,M]\!]^\circ_{\mathtt{A}}\,\rho = (C \in [\![\sigma]\!]_{\mathtt{A}} \mapsto [\![M]\!]_S\,(\rho[x := C])) \\
[\![x]\!]_{\mathtt{A}}\,\rho = \rho(x) \qquad [\![V]\!]_S\,\rho = \downarrow [\![V]\!]^\circ_S\,\rho \quad (V \text{ a value}) \\
[\![MN]\!]_{\mathtt{A}}\,\rho &=& \sup_{\substack{f\in[\![M]\!]_{\mathtt{A}}\rho \\ v\in[\![N]\!]_{\mathtt{A}}\rho}} f(v) \\
[\![\mathtt{Y}N]\!]_{\mathtt{A}}\,\rho &=& \sup_{n\in\mathbb{N}} f^n(\emptyset) \\
& & \text{where } f(C) = \sup_{g\in[\![N]\!]_{\mathtt{A}}\rho} g(C) \\
[\![\mathtt{pred}\,M]\!]_{\mathtt{A}}\,\rho &=& \{n-1 \mid n \in [\![M]\!]_{\mathtt{A}}\,\rho, n \neq 0\} \\
[\![\mathtt{succ}\,M]\!]_{\mathtt{A}}\,\rho &=& \{n+1 \mid n \in [\![M]\!]_{\mathtt{A}}\,\rho\} \\
[\![\mathtt{ifz}\ M\ N\ P]\!]_{\mathtt{A}}\,\rho &=& \begin{cases} \emptyset & \text{if } [\![M]\!]_{\mathtt{A}}\,\rho = \emptyset \\ [\![N]\!]_{\mathtt{A}}\,\rho & \text{if } [\![M]\!]_{\mathtt{A}}\,\rho = \{0\} \\ [\![P]\!]_{\mathtt{A}}\,\rho & \text{if } [\![M]\!]_{\mathtt{A}}\,\rho \neq \emptyset \text{ and } 0 \notin [\![M]\!]_{\mathtt{A}}\,\rho \\ [\![N]\!]_{\mathtt{A}}\,\rho \cup [\![P]\!]_{\mathtt{A}}\,\rho & \text{otherwise} \end{cases} \\
[\![\mathtt{let}\,x \Leftarrow M\ \mathtt{in}\ N]\!]_{\mathtt{A}}\,\rho &=& \sup_{v\in[\![M]\!]_{\mathtt{A}}\rho} [\![N]\!]_{\mathtt{A}}\,(\rho[x := \downarrow v]) \\
[\![M \mathbin{\textcircled{\scriptsize\vee}} N]\!]_{\mathtt{A}}\,\rho &=& [\![M]\!]_{\mathtt{A}}\,\rho \cup [\![N]\!]_{\mathtt{A}}\,\rho
\end{array}
$$

Figure 8: Standard semantics for $\mathrm{PCF}_{\mathtt{A}}$

12. Concluding Remarks

In the purely non-deterministic, angelic case, Theorem 11.9 is rather remarkable: the domain-theoretic semantics of $\mathrm{PCF}_{\mathtt{A}}$ *is* fully abstract. No extra primitive (parallel or, statistical termination testers) is needed for that.

With both angelic non-determinism and probabilistic choice, a situation similar to that of PCF occurs: $\mathrm{PCF}_{\mathtt{AP}}$ is not fully abstract, but $\mathrm{PCF}_{\mathtt{AP}}$ plus a simple, natural operation—statistical termination testers $\bigcirc_{>b}$—is fully abstract (Theorem 11.7).

The other PCF_S languages are on our agenda. Their all enjoy soundness and adequacy, as we have seen, and some of the results seem to be at hand. Notably, full abstraction for $\mathrm{PCF}_{\mathtt{DP}} + \bigcirc$ (resp., $\mathrm{PCF}_{\mathtt{D}}$) seems to be entirely analogous to $\mathrm{PCF}_{\mathtt{AP}} + \bigcirc$ (resp., $\mathrm{PCF}_{\mathtt{A}}$), since the Scott and weak topologies also coincide on $\mathbb{P}_{\mathtt{DP}}(X)$ (resp., $\mathbb{P}_{\mathtt{D}}(X)$) for every continuous dcpo X [15, Proposition 3.44]; but the definability of weak opens is harder to realize. Once this is done, the erratic cases $S = \mathtt{ADP}$ and $S = \mathtt{AD}$ are low-hanging fruit. The case $S = \mathtt{P}$ is hardest: we do not even know a suitable Cartesian-closed category of *continuous* dcpos to interpret $\mathrm{PCF}_{\mathtt{P}}$ in [23], and continuity on X is needed to equate the Scott and weak topologies on $\mathbb{P}_{\mathtt{P}}(X)$.

Acknowledgments

Earlier versions of the results of this paper were presented at the Domains X workshop in Swansea. I would like to thank the participants of the workshop for their questions, and in particular Paul B. Levy. The last PCF_S languages I presented at Swansea were call-by-value, and this intrigued Paul. I thought back about it, and I switched to the call-by-name language presented here, which additionally has no explicit monadic construct. Call-by-value remains required, in the form of a `let` construct that is used again and again in the proof of definability, and which is essentially used to force sequential execution. This is not needed in Plotkin's original PCF, which does not have choice, and where sequential execution is hidden in calls to `ifz`.

One might also wonder whether one really needs higher-order lets, namely let expressions of the form $\mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ N$ where M is of non-ground type, for definability and full abstraction. Indeed, the only higher-order let we ever need in our proofs of the latter results is the term $V = \lambda z\,.\bigcirc_{>b} \mathtt{let}\, x \Leftarrow z\ \mathtt{in}\ M$, with z of arbitrary type τ (see Lemma 11.6). This is also the only place were we need the $\bigcirc_{>b}$ operator. This suggests an alternative language, where:

- $\mathtt{let}\, x \Leftarrow M\ \mathtt{in}\ N$ would only be available for M of type `Unit`; this would be synonymous with Escardó's unary conditional $\mathtt{if}\ M\ \mathtt{then}\ N$ [9, Section 3.1], a very undemanding construct, available in usual programming languages as sequential composition $M; N$,

- and $\bigcirc_{>b}$ would be absent. Instead, we would have an extra construction $Pr[x \Leftarrow N\ \mathtt{in}\ M{\downarrow}]_{>b}$, of type `Unit` (for x, N of type τ, M of type `Unit`), meant to terminate if and only if the probability that M terminates when x is sampled according to N is greater than b. This is the semantics of $\bigcirc_{>b} \mathtt{let}\, x \Leftarrow N\ \mathtt{in}\ M$ in our language $\mathrm{PCF}_S + \bigcirc$.

Formally, the denotational semantics would need to be modified by letting:

$$\begin{aligned} & [\![Pr[x \Leftarrow N\ \mathtt{in}\ M{\downarrow}]_{>b}]\!]_S\, \rho \\ = & \begin{cases} val_S * & \text{if } (let_S\, v \Leftarrow [\![N]\!]_S\, \rho\ in\ [\![M]\!]_S\, (\rho[x := val_S\, v])) \blacklozenge > b \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

and the operational rules for `let` and $\bigcirc_{>b}$ should be replaced by the following

two:

$$C \cdot Pr[x \Leftarrow N \,\mathtt{in}\, M\downarrow]_{>b} \to C[Pr[x \Leftarrow \_ \,\mathtt{in}\, M\downarrow]_{>b}] \cdot N$$

$$\frac{C \cdot M[x := V] \downarrow^m b \qquad C \cdot \underline{*} \downarrow^m a}{C[Pr[x \Leftarrow \_ \,\mathtt{in}\, M\downarrow]_{>b}] \cdot V \downarrow^m a} \,(V \text{ a value})$$

For $S = \mathtt{AP}$ or $S = \mathtt{A}$, proofs similar to those we have given in this paper should establish that the resulting language is again sound, adequate and fully abstract.

Finally, I would like to thank the anonymous referees, who displayed extraordinary acuteness. I also thank them for their patience.

This work was partly supported by the ANR Blanc project ANR-09-BLAN-0345-02 (CPP).

References

[1] S. Abramsky, R. Jagadeesan, P. Malacaria, Full abstraction for PCF, Information and Computation 163 (2000) 409–470.

[2] S. Abramsky, A. Jung, Domain theory, in: S. Abramsky, D.M. Gabbay, T.S.E. Maibaum (Eds.), Handbook of Logic in Computer Science, volume 3, Oxford University Press, 1994, pp. 1–168.

[3] M. Alvarez-Manilla, A. Jung, K. Keimel, The probabilistic powerdomain for stably compact spaces, Theoretical Computer Science 328 (2004) 221–244.

[4] G. Berry, Modèles Complètement Adéquats et Stables des Lambda-Calculs Typé, Thèse d'état, Université Paris 7, 1979.

[5] P.L. Curien, An abstract framework for environment machines, Theoretical Computer Science 82 (1991) 389–402.

[6] V. Danos, R. Harmer, Probabilistic game semantics, ACM Transactions on Computational Logic 3 (2002) 359–382.

[7] T. Ehrhard, C. Tasson, M. Pagani, Probabilistic coherence spaces are fully abstract for probabilistic PCF, in: S. Jagannathan, P. Sewell (Eds.), Proc. 41st Ann. ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL '14), pp. 309–320.

[8] T. Erker, M.H. Escardó, K. Keimel, The way-below relation of function spaces over semantic domains, Topology and Its Applications 89 (1998) 61–74.

[9] M. Escardó, Semi-decidability of may, must and probabilistic testing in a higher-type setting, ENTCS 249 (2009) 219–242.

[10] G. Gierz, K.H. Hofmann, K. Keimel, J.D. Lawson, M. Mislove, D.S. Scott, Continuous Lattices and Domains, volume 93 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, 2003.

[11] J. Goubault-Larrecq, Continuous previsions, in: J. Duparc, Th.A. Henzinger (Eds.), Proceedings of the 16th Annual EACSL Conference on Computer Science Logic (CSL'07), volume 4646, Springer Verlag Lecture Notes in Computer Science, Lausanne, Switzerland, 2007, pp. 542–557.

[12] J. Goubault-Larrecq, Prevision domains and convex powercones, in: R. Amadio (Ed.), Proceedings of the 11th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS'08), Springer-Verlag LNCS 4962, Budapest, Hungary, 2008, pp. 318–333.

[13] J. Goubault-Larrecq, De Groot duality and models of choice: Angels, demons, and nature, Mathematical Structures in Computer Science 20 (2010) 169–237.

[14] J. Goubault-Larrecq, Non-Hausdorff Topology and Domain Theory—Selected Topics in Point-Set Topology, volume 22 of *New Mathematical Monographs*, Cambridge University Press, 2013.

[15] J. Goubault-Larrecq, Isomorphism theorems between models of mixed choice, Mathematical Structures in Computer Science (2014). To appear.

[16] R. Harmer, G. McCusker, A fully abstract game semantics for finite nondeterminism, in: Proc. 14th IEEE Symposium on Logics in Computer Science (LICS'99), IEEE Computer Society Press, 1999, pp. 422–430.

[17] R. Heckmann, Power Domain Constructions (Potenzbereich-Konstruktionen), Ph.D. thesis, Universität des Saarlandes, 1990.

[18] R. Heckmann, Abstract valuations: A novel representation of Plotkin power domain and Vietoris hyperspace, Electronic Notes in Theoretical Computer Science 6 (1997). In Proc. 13th Intl. Symp. on Mathematical Foundations of Programming Semantics (MFPS'97).

[19] J.M.E. Hyland, L. Ong, On full abstraction for PCF: I, II and III, Information and Computation 163 (2000) 285–408.

[20] C. Jones, Probabilistic Non-Determinism, Ph.D. thesis, University of Edinburgh, 1990. Technical Report ECS-LFCS-90-105.

[21] C. Jones, G.D. Plotkin, A probabilistic powerdomain of evaluations, in: Proc. 4th IEEE Symposium on Logics in Computer Science (LICS'89), IEEE Computer Society Press, 1989, pp. 186–195.

[22] A. Jung, J. Tiuryn, A new characterization of lambda definability, in: Proc. Conf. Typed Lambda Calculi and Applications, Springer Verlag LNCS 664, 1993, pp. 230–244.

[23] A. Jung, R. Tix, The troublesome probabilistic powerdomain, in: A. Edalat, A. Jung, K. Keimel, M. Kwiatkowska (Eds.), Proc. 3rd Workshop on Computation and Approximation, volume 13 of *Electronic Lecture Notes in Computer Science*, Elsevier, 1998, pp. 70–91. 23pp.

[24] S.Y. Katsumata, A semantic formulation of $\top\top$-lifting and logical predicates for computational metalanguage, in: L. Ong (Ed.), Proc. 19th Intl. Workshop CSL 2005, 14th Ann. Conf. of the EACSL, Springer Verlag LNCS 3634, 2005, pp. 87–102.

[25] O. Kirch, Bereiche und Bewertungen, Master's thesis, Technische Hochschule Darmstadt, 1993.

[26] A. McIver, C. Morgan, Demonic, angelic and unbounded probabilistic choices in sequential programs, Acta Informatica 37 (2001) 329–354.

[27] R. Milner, Fully abstract models of typed λ-calculi, Theoretical Computer Science 4 (1977) 1–22.

[28] M. Mislove, Topology, domain theory and theoretical computer science, Topology and Its Applications 89 (1998) 3–59.

[29] M. Mislove, Nondeterminism and probabilistic choice: Obeying the law, in: Proc. 11th Conf. Concurrency Theory (CONCUR'00), Springer Verlag LNCS 1877, 2000, pp. 350–364.

[30] M. Mislove, J. Ouaknine, J. Worrell, Axioms for probability and nondeterminism, Electronic Notes in Theoretical Computer Science 91 (2003) 7–28. Proc. 10th Int. Workshop on Expressiveness in Concurrency (EXPRESS'03).

[31] E. Moggi, Notions of computation and monads, Information and Computation 93 (1991) 55–92.

[32] C. Morgan, A. McIver, K. Seidel, Probabilistic program transformers, ACM Transactions on Programming Languages and Systems 18 (1996) 325–353.

[33] P.W. O'Hearn, U.S. Reddy, Objects, interference and the Yoneda embedding, Theoretical Computer Science 228 (1999) 253–282. Journal version of *Objects, Interference and the Yoneda Embedding*, Proceedings of the 11th Intl. Conf. on Mathematical Foundations of Programming Semantics, Electronic Notes in Theoretical Computer Science, Main, M. and Brookes, S., eds, 1995.

[34] P.W. O'Hearn, J.G. Riecke, Kripke logical relations and PCF, Information and Computation 120 (1995) 107–116.

[35] P.W. O'Hearn, R.D. Tennent, Semantics of local variables, in: M.P. Fourman, P.T. Johnstone, A.M. Pitts (Eds.), Applications of Categories in Computer Science: Proceedings of the LMS Symposium, volume 177 of *LMS Lecture Note Series*, Cambridge University Press, Durham, UK, 1992, pp. 217–238.

[36] P.W. O'Hearn, R.D. Tennent, Parametricity and local variables, Journal of the ACM 42 (1995) 658–709.

[37] A.M. Pitts, I.D.B. Stark, Operational reasoning for functions with local state, in: A.D. Gordon, A.M. Pitts (Eds.), Higher-Order Operational Techniques in Semantics, Publications of the Newton Institute, Cambridge University Press, 1998, pp. 227–273.

[38] G.D. Plotkin, LCF considered as a programming language, Theoretical Computer Science 5 (1977) 223–255.

[39] U.S. Reddy, Global state considered unnecessary: Object-based semantics for interference-free imperative programs, Lisp and Symbolic Computation 9 (1996).

[40] J.G. Riecke, A. Sandholm, A relational account of call-by-value sequentiality, Information and Computation 179 (2002) 296–331.

[41] A. Schalk, Algebras for Generalized Power Constructions, Ph.D. thesis, Technische Universität Darmstadt, 1993.

[42] T. Streicher, Domain-Theoretic Foundations of Functional Programming, World Scientific, 2006.

[43] R. Tix, Stetige Bewertungen auf topologischen Räumen, Diplomarbeit, TH Darmstadt, 1995.

[44] R. Tix, Continuous D-Cones: Convexity and Powerdomain Constructions, Ph.D. thesis, Technische Universität Darmstadt, 1999.

[45] R. Tix, K. Keimel, G. Plotkin, Semantic domains for combining probability and non-determinism, Electronic Notes in Theoretical Computer Science 129 (2005) 1–104.

[46] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.