A Probabilistic Powerdomain of Evaluations

C. Jones G. D. Plotkin

Department of Computer Science
University of Edinburgh

Abstract

We give a probabilistic powerdomain construction on the category of inductively complete partial orders; it is the partial order of continuous $[0,1]$-valued evaluations on the Scott topology. By means of a theory of integration with respect to such evaluations, the powerdomain is shown to be a monad, and indeed one gets a model for Moggi's computational lambda-calculus. One can also solve recursive domain equations involving the powerdomain, and all this gives a meta-language for programming languages with probabilistic features. This is used to give the semantics of a language with a probabilistic parallel construct. We show the construction generalises previous work on partial orders of measures, and indeed restricts to the category of continuous partial orders, where it can be characterised as a free continuous finitary algebra.

1 Introduction

In order to study the semantics of probabilistic programming languages one is naturally led to investigate probabilistic powerdomains (perhaps by analogy with the introduction of powerdomains for non-determinism [15]). For this one minimally needs a category of "datatypes" closed under the powerdomain and any other needed functors and also allowing recursive definitions of elements and datatypes. Probabilistic program logics are also very important and have a wide literature, but fall outside the scope of this paper.

The seminal work here is by Saheb-Djahromi [14] who gave a probabilistic powerdomain $\mathcal{P}_D$ of ω-algebraic cpos, consisting of all probability distributions on the Borel sets of the Scott topology. However that was not algebraic in general. Then Plotkin [13] and Graham [3] showed the category of finitely continuous posets (retracts of **SFP** objects) was closed under the powerdomain and it can indeed be used for semantical purposes as above. There is closely related work by Kozen [6] and Yamada [17]. These authors typically consider all measurable functions; any topological or partial order structure is derived. It has not been shown how to fulfill the above minimal requirements with this approach.

This paper presents a natural and straightforward theory of a probabilistic powerdomain $\mathcal{E}(P)$ of any directed complete poset based on open sets (following [15]) rather than general Borel sets, and hence evaluations rather than measures. Section 2 gives basic definitions, section 3 presents a theory of integration of upper continuous functions relative to evaluations and section 4 organises the structure associated with $\mathcal{E}$ showing it to be a monad (see [2],[9],[8] for the importance of monads for probability theory). Section 5 shows how to solve recursive domain equations involving

$\mathcal{E}$. Section 6, following recent ideas of Moggi [10] shows the monad is strong and so one can probabilistically interpret his computational λ-calculus, providing (the basis of) a very suitable metalanguage for the denotational semantics of probabilistic programming languages; section 7 gives a semantics for such a language with probabilistic concurrency. Section 8 connects up with the other approaches showing evaluations extend to measures whenever P is continuous [4], and, finally, section 9 extends Graham's finitary algebraic characterisation of the powerdomain [3] to the continuous case.

2 Evaluations

We work with the category **IPO** with objects the *inductively complete partial orders* (ipos) those having lubs of directed subsets (there need be no least element); the morphisms are the continuous functions, those preserving directed lubs. The category is cartesian closed, complete and co-complete. The *Scott topology* on an ipo consists of those upper sets inaccessible by directed joins [4].

To specify a probabilistic computation over an ipo P it is natural to consider $[0,1]$-valued functions on sets of "events". For computer science purposes we consider as such events, tests on data and identify these with open sets rather than the more general Borel sets. There are natural and well-known axioms for such functions.

Definition Let X be a topological space and $\Omega(X)$ be its (complete) lattice of open subsets. A function $\mu : \Omega(X) \to [0,1]$ is an *evaluation* iff

Monotonicity If $U \subseteq V$ then $\mu(U) \leq \mu(V)$

Strictness $\mu(\emptyset) = 0$

Modularity $\mu(U)+\mu(V) = \mu(U \cup V)+\mu(U \cap V)$

Also, such a function μ is *continuous* iff for any directed set U_λ ($\lambda \in \Lambda$) of open sets, $\mu(\bigcup_\lambda U_\lambda) = \sqcup_\lambda \mu(U_\lambda)$. This natural condition was introduced by Lawson [7].

Examples 1. For any x in P the point-mass evaluation $\eta_P(x)$ is defined by:

$$\eta_P(x)(U) = \begin{cases} 1 & \text{if } x \text{ is in } U \\ 0 & \text{otherwise} \end{cases}$$

2. If $\sum_{i=1}^{n} r_i \leq 1$ where $r_i \geq 0$ then $\sum r_i \mu_i$ is an evaluation if the μ_i are; its value at U is $\sum r_i \mu_i(U)$.

Definition Let P be an ipo. Then its *probabilistic powerdomain* is $\mathcal{E}(P)$ the set of continuous evaluations on P partially ordered by: $\mu \leq \nu$ iff $\forall U$ in $\Omega(P)$, $\mu(U) \leq \nu(U)$.

Theorem 2.1 *$\mathcal{E}(P)$ is an ipo with directed lubs defined pointwise, and with a least element, the constantly zero evaluation.*

Remark We could instead have worked with the ω-continuous evaluations, those preserving lubs of countably directed sets. This would make no difference to the course of the general theory in sections 3, 4, 5, 6 and 8 except that the Monotone Convergence theorem below would be restricted to *countable* directed sets of continuous functions. In the definition of $\mathcal{D}(P)$ we would take all $[0,1]$-valued measures. However the results given for continuous ipos (Fubini's theorem and theorems 8.2 and 9.1) would be replaced by the corresponding results for the smaller class of ω-continuous ipos (defined as in [16], but without the bottom element). These results would be a special case of the results given below as when P is ω-continuous, all ω-continuous evaluations are actually continuous.

3 Integration & Evaluation

Previous work on probabilistic powerdomains has involved the use of measures and integration. We consider integration (with respect to an evaluation) of *continuous functions* to $[0,1]$ regarded as an ipo under its usual ordering.

We start with *simple functions*, those taking finitely many values. By [12] or [5, p23], any evaluation μ extends to a finitely additive set function $\bar{\mu}$ on the ring generated by the open sets. So if s is a simple function, taking values $\alpha_1, \ldots, \alpha_n$, one can define:

$$\int s\, d\mu = \sum_{i=1}^{n} \alpha_i \bar{\mu}(s^{-1}\{\alpha_i\})$$

Any simple function can be written as $\sum r_i \chi_{U_i}$ ($r_i \geq 0$) and it can be seen that $\int s\, d\mu = \sum r_i \mu(U_i)$. Now the integral of a general continuous function $f : P \to [0,1]$ has the standard form of definition:

$$\int f\, d\mu = \sup\{\int s\, d\mu \mid s \text{ is simple and } s \leq f\}$$

Theorem 3.1 (Monotone Convergence) *For a bounded, directed set of continuous functions $f_\lambda : P \to [0,1]$ ($\lambda \in \Lambda$)*

$$\int \bigvee f_\lambda\, d\mu = \bigvee \int f_\lambda\, d\mu$$

Remark In proving this one makes use of the fact that every continuous $f : P \to [0,1]$ is the sup and uniform limit of an increasing sequence of simple functions.

Theorem 3.2 (Linearity) *For continuous $f_i : P \to [0,1]$ and non-negative real numbers r_i with $\sum r_i \leq 1$*

$$\int \sum_{i=1}^{n} r_i f_i\, d\mu = \sum_{i=1}^{n} r_i \int f_i\, d\mu$$

As an application of these results one sees that $\int f\, d\eta_P(x) = f(x)$. For this is trivial when $f = \chi_U$ and then it follows for general f by Monotone Convergence, Linearity and the above remark. As another example of this one can show

Theorem 3.3 *If D is a directed set of evaluations, then*

$$\int f\, d(\bigsqcup_{\mu \in D} \mu) = \sup_{\mu \in D} \int f\, d\mu$$

4 $\mathcal{E}$ is a monad

There are a number of natural functions which arise in using $\mathcal{E}$ to give the semantics of probabilistic programming languages. They can be described categorically by showing $\mathcal{E}$ is a monad. The monadic structure can be nicely presented from the "Kleisli point of view".

Definition Let $\mathbf{K}$ be a category. A *Kleisli triple* is a structure $(F, \eta, (\cdot)^*)$ where $F : Obj(\mathbf{K}) \to Obj(\mathbf{K})$, and $\eta_x : x \to Fx$ (for any object x) is the *unit map* and $f^* : Fx \to Fy$ (for any $f : x \to Fy$) is the *extension* of f, such that

1. $\eta_x^* = \mathrm{id}_{Fx}$ (for x in $Obj(\mathbf{K})$)
2. $f^* \circ \eta = f$ (for $f : x \to Fy$)
3. $g^* \circ f^* = (g^* \circ f)^*$ (for $f : x \to Fy$ and $g : y \to Fz$)

As is well-known [9] to every such triple $(F, \eta, (\cdot)^*)$ corresponds the monad (F, η, μ) where $F(f) = (\eta_Y \circ f)^*$ (for $f : x \to y$) and $\mu_x = (\mathrm{id}_{Fx})^*$. (And, conversely, given such a monad we can define a triple with $f^* = \mu_y \circ Ff$, and moreover these correspondences are mutual inverses.)

In our case $\eta_P(x)$ is as above and the extension of $f : P \to \mathcal{E}(Q)$ is defined by:

$$f^*(\mu)(V) = \int_{x \in P} f(x)(V)\, d\mu$$

(meaning $\int k\,d\mu$ where $k(x) = f(x)(V)$ for x in P). One shows that f^* is well-defined by using the results of section 3. We just verify the third property of a Kleisli triple.

We have to prove for $f : P \to \mathcal{E}(Q)$ and $g : Q \to \mathcal{E}(R)$ that

$$\int_{y\in Q} g(y)V\, d(W \mapsto \int_{x\in P} f(x)W\, d\mu) =$$

$$\int_{x\in P}\int_{y\in Q} g(y)V\, df(x)\, d\mu$$

which would follow from the more general

$$\int_{y\in Q} h(y)\, d(W \mapsto \int_{x\in P} f(x)W\, d\mu) =$$

$$\int_{x\in P}\int_{y\in Q} h(y)\, df(x)\, d\mu$$

where $h : Q \to [0,1]$ is continuous. It suffices to prove it for $h = \chi_W$ for open W which is immediate since both sides reduce to the expression

$$\int_{x\in P} f(x)(W)\, d\mu$$

Finally we note that the action of $\mathcal{E}$ on morphisms is given by: $\mathcal{E}(f)(\mu)(V) = \mu(f^{-1}(V))$.

5 Recursive Domain Equations

We apply the theory of [16] to the category **pIPO** of ipos and continuous partial functions, those $f : P \rightharpoonup Q$ such that $f^{-1}(V)$ is open whenever V is. This is a partial ccc in the sense of Moggi [11]. (More accurately we can define the domain structure $\mathcal{M}_{\text{open}}$ of open subobjects of **IPO** so that $[e]$ is in $\mathcal{M}_{\text{open}}(Q)$ for $e : P \to Q$ iff (i) $ex \sqsubseteq ey$ iff $x \sqsubseteq y$ for x, y in P and (ii) $e(P)$ is open, and then **pIPO** $\cong P(\mathbf{IPO}, \mathcal{M}_{\text{open}})$.)

The product and sum functors extend from **IPO** to **pIPO** by: $f \times g(\langle x,y\rangle) \simeq \langle fx, gy\rangle$ (which exists iff fx and gx do) and $f + g(\langle i,x\rangle) \simeq \langle 0, fz\rangle$ (if $i = 0$) and $\simeq \langle 1, gz\rangle$ (if $i = 1$). There is a partial function-space functor defined by $f \rightharpoonup g(h) = g \circ h \circ f$. However a certain "$\mathcal{E}$ function space" functor $\to_{\mathcal{E}}$ is more important to us. First note the natural inclusion functor α : **pIPO** $\to$ $\mathbf{IPO}_{\mathcal{E}}$ (the latter being the Kleisli category) where $\alpha(f)(x)(V) = 1$ (if $fx\downarrow$ and $fx \in V$) $= 0$ (otherwise). Then $\to_{\mathcal{E}}$ is defined on **pIPO** by $P \to_{\mathcal{E}} Q =_{\text{def}} P \to \mathcal{E}(Q)$ and for $f : P \leftharpoondown P'$ and $g : Q \rightharpoonup Q'$, $f \to_{\mathcal{E}} g(h) = \alpha(g) \circ h \circ \alpha(f)$.

Now **pIPO** is an **O**-category in the sense of [16] if we partially order the hom-sets by: $f \sqsubseteq g$ iff $\forall x \in P. f(x)\downarrow \supset g(x)\downarrow \wedge f(x) \sqsubseteq g(x)$, where $f, g : P \rightharpoonup Q$. For if $f_n : P \rightharpoonup Q$ is an increasing sequence it has a lub given by $(\sqcup_n f_n)(x) \simeq \sqcup_n f_n(x)$ (where this is taken to exist iff $f_k(x)\downarrow$ for some k and then it is taken to be $\sqcup_{n\geq k} f_n(x)$). Composition is continuous in that these ω-lubs are preserved. Note that the totally undefined function, $\emptyset$, is always least and composition is strict (preserves $\emptyset$) in each argument; also the empty ipo is a null object, both initial and final. Next **pIPO** has ω-limits. Suppose $\Delta = \langle P_n, f_n\rangle$ is an ω^{op}-chain. Then the limit is the set $\{x : \omega \rightharpoonup \bigcup_n P_n \mid (\forall n, x_n\downarrow \supset x_n \in P_n \wedge x_n \simeq f_n(x_{n+1}))$ and $\exists n, x_n\downarrow\}$ partially ordered by $x \sqsubseteq y$ iff $\forall n, x_n\downarrow \supset y_n\downarrow \wedge x_n \sqsubseteq y_n$. The universal cone $\rho : P \to \Delta$ is $\rho_n(x) \simeq x_n$.

With all this established the theory of [16] directly applies to **pIPO** and we can solve systems of recursive domain equations involving locally continuous functors: sums, products and $\to_{\mathcal{E}}$ are all easily be seen to be locally continuous.

6 Computational λ-calculus

We want to give the semantics of probabilistic programming languages using a λ-calculus for probabilistically nondeterministic functions. An extraordinarily suitable general tool for this purpose has been provided by Moggi's λ_c-calculus ([10]). To interpret the calculus we have to provide a *tensorial strength*, which is a natural transformation:

$$P \times \mathcal{E}Q \stackrel{t_{P,Q}}{\to} \mathcal{E}(P \times Q)$$

satisfying certain diagrams. It turns out that as **IPO** has enough points $t_{P,Q}$ is determined uniquely by

$$t_{P,Q}(x,\nu)(W) = \int_{y \in Q} \chi_W(x,y)\, d\nu$$

and that this indeed defines a tensorial strength. This gives a λ_c-model in Moggi's sense as η_P is easily shown a mono.

The tensorial strength gives rise to a natural transformation $\psi_{P,Q} : (\mathcal{E}(P) \times \mathcal{E}(Q)) \to \mathcal{E}(P \times Q)$ which can be used to interpret the pairing of two computations. An explicit formula for $\psi_{P,Q}$ is given by

$$\psi_{P,Q}(\mu,\nu) = W \mapsto \int_{x \in P} \int_{y \in Q} \chi_W(x,y)\, d\nu\, d\mu$$

(Integrating in the other order gives the dual notion of pairing ($\tilde{\psi}$ in [10]). The two are equal iff Fubini's theorem holds in this setting — which we conjecture to be false in general, but which does hold when P and Q are both continuous.)

We present (an extension of) the λ_c-calculus and its semantics in **IPO** somewhat informally. There will be expressions and type expressions. The latter comprise constants for all the ipos (and we identify the two) and product and function space type expressions (and we shall add sum type expressions too) They denote what one would expect except perhaps for function spaces where if σ,τ denote P,Q then $\sigma \to \tau$ denotes $P \to_{\mathcal{E}} Q$. The former include variables, local declarations, abstractions, applications, tuples, projections, injections and case statements, we also add recursion and constants as required. Expressions are typed given types for their free variables, and if e has type P it is to denote an element of $\mathcal{E}(P)$ and e will denote a continuous function of its free variables. A *local declaration* **let** $x \in \sigma$ **be** e **in** e' denotes $f^*(\mu)$ where e denotes $\mu \in \mathcal{E}(P)$ (and σ denotes P) and e' denotes $f : P \to_{\mathcal{E}} Q$ as a function of P. A *λ-abstraction* $\lambda x \in \sigma.e$ where e has type τ denotes $\eta_{P \to_{\mathcal{E}} Q}(f)$ where σ, τ denote P, Q and $f(x) = e$. An *application* $e'(e)$ (where $e' : \sigma \to \tau$ denotes $f \in \mathcal{E}(P \to_{\mathcal{E}} Q)$ and $e : \sigma$ denotes $x \in \mathcal{E}(P)$) denotes $\alpha(f,x)$ where α is the composition

$$\mathcal{E}(P \to \mathcal{E}(Q)) \times \mathcal{E}(P) \xrightarrow{\psi}$$

$$\mathcal{E}((P \to \mathcal{E}(Q)) \times P) \xrightarrow{\mathrm{eval}^*} \mathcal{E}(Q)$$

A *pair* $\langle e, e' \rangle$ denotes $\psi(\mu,\nu)$, where e,e' denote μ,ν and projections **fst**(e), **snd**(e) have evident denotations. For sums if e denotes $\mu \in \mathcal{E}(P+Q)$ and e_1, e_2 denote $f : P \to_{\mathcal{E}} R$, $g : Q \to_{\mathcal{E}} R$ respectively, as functions of x then the *case analysis* **cases** e **fst** $x \in \sigma.e_1$ **snd** $y \in \tau.e_2$ denotes $[f,g]^*(\mu)$. The injections **inl**(e), **inr**(e) have evident denotations. Finally, functions can be defined recursively by the construct $\mu f : \sigma \to \tau.\lambda x : \sigma.e$ where e must have type τ. Suppose σ, τ denote P, Q and e denotes $F : (P \to_{\mathcal{E}} Q) \times P \to_{\mathcal{E}} Q$ as a function of f and x. Then the construct denotes $\eta_{P \to_{\mathcal{E}} Q}(g)$ where g is the least fixed point of Curry(F). We will define functions recursively by writing $f \Leftarrow \lambda x : \sigma.e$ meaning $f = \mu f : \sigma \to \tau.\lambda x : \sigma.e$, for appropriate σ, τ.

7 A Language with Probabilistic Concurrency

Consider the language *Com* of commands c given by:

$$c \ ::= \ a \mid c;c' \mid \mathtt{skip} \mid \mathtt{if}\ b\ \mathtt{then}\ c\ \mathtt{else}\ c' \mid \mathtt{while}\ b\ \mathtt{do}\ c \mid c + c' \mid c \parallel c'$$

where a,b range over sets $ACom$,$BExp$ of atomic commands and Boolean expressions. The first five clauses present a simple iterative language, the sixth a probabilistic choice between c and c', the last a probabilistic scheduled concurrency. For the semantics we assume an ipo, S, of states and given denotational functions $\mathcal{A} : ACom \to (S \to_{\mathcal{E}} S)$ and $\mathcal{B} : BExp \to (S \to_{\mathcal{E}} \mathsf{T})$ where T is $\mathbb{1} + \mathbb{1}$ (and $\mathbb{1}$ is the one-point ipo).

We define the denotational semantics $\mathcal{C} : Com \to R$ where the ipo R of *resumptions* is the solution as considered above to the recursive domain equation:

$$R \cong S \to_{\mathcal{E}} (S + (S \times R))$$

In what follows we treat the isomorphism as an actual equality.

Atomic Commands
$\mathcal{C}[\![a]\!] = \lambda\sigma \in S.\mathbf{inl}(\mathcal{A}[\![a]\!](\sigma))$

Sequence
$\mathcal{C}[\![c; c']\!] = \mathcal{C}[\![c]\!] \star \mathcal{C}[\![c']\!]$ where $\star : R \to_{\mathcal{E}} (R \to_{\mathcal{E}} R)$ is defined recursively by

$$\begin{aligned} \star \Leftarrow & \lambda r \in R\, \lambda r' \in R \quad \lambda\sigma \in S \ \mathbf{cases}\ r(\sigma) \\ & \mathbf{fst}\ \sigma' \in S.\mathbf{inr}\,(\langle \sigma', r' \rangle) \\ & \mathbf{snd}\ x \in S \times R.\mathbf{inr}\,(\langle \mathbf{fst}(x), \mathbf{snd}(x) \star r' \rangle) \end{aligned}$$

Conditionals
$\mathcal{C}[\![\mathtt{if}\ b\ \mathtt{then}\ c\ \mathtt{else}\ c']\!] = \lambda\sigma \in S.\mathbf{if}\ \mathcal{B}[\![b]\!](\sigma)\ \mathbf{then}\ \mathcal{C}[\![c]\!](\sigma)\ \mathbf{else}\ \mathcal{C}[\![c']\!](\sigma)$
(where $\mathbf{if}\ e\ \mathbf{then}\ e'\ \mathbf{else}\ e''$ abbreviates $\mathbf{cases}\ e\ \mathbf{fst}\ x \in \mathbb{1}.e'\ \mathbf{snd}\ x \in \mathbb{1}.e''$ with x not free in e' or e'')

Loops

$$\begin{aligned} \mathcal{C}[\![\mathtt{while}\ b\ \mathtt{do}\ c]\!] = \mu\theta : R\, \lambda\sigma \in S\ & \mathbf{if}\ \ \mathcal{B}[\![b]\!](\sigma) \\ & \mathbf{then}\ \ (\theta \star \mathcal{C}[\![c]\!])(\sigma) \\ & \mathbf{else}\ \ \mathbf{inl}(\sigma) \end{aligned}$$

Probabilistic Choice
$\mathcal{C}[\![c + c']\!] = \lambda\sigma.\frac{1}{2}\mathcal{C}[\![c]\!](\sigma) + \frac{1}{2}\mathcal{C}[\![c']\!](\sigma)$

Probabilistic Concurrency We parameterise on a probabilistic scheduler, κ, which decides, given a state, which process runs next; the scheduler also has its own internal state. We take the ipo Sch of schedulers to be the solution to the recursive domain equation:

$$Sch \cong S \to_{\mathcal{E}} (\mathsf{T} \times Sch)$$

(and treat the isomorphism as an equality) and define recursively $PAR : R \to_{\mathcal{E}} (R \to_{\mathcal{E}} (Sch \to_{\mathcal{E}} R))$ by:

$$\begin{aligned} PAR \Leftarrow & \lambda r \in R\, \lambda r' \in R\, \lambda\kappa \in Sch\, \lambda\sigma \in S \\ & \quad \mathbf{let}\ x \in \mathsf{T} \times Sch\ \mathbf{be}\ \kappa(\sigma)\ \mathbf{in} \\ & \quad \mathbf{if}\ \mathbf{fst}(x) \\ & \quad \mathbf{then\ cases}\ r(\sigma) \\ & \qquad \mathbf{fst}\ \sigma' \in S.\mathbf{inr}\,(\langle \sigma', r' \rangle) \\ & \qquad \mathbf{snd}\ y \in S \times R.\mathbf{inr}\,(\langle \mathbf{fst}(y), \\ & \qquad\qquad PAR(\mathbf{snd}(y))(r')(\mathbf{snd}(x)) \rangle) \\ & \quad \mathbf{else\ cases}\ \ r'(\sigma) \\ & \qquad \mathbf{fst}\ \sigma' \in S.\mathbf{inr}\,(\langle \sigma', r \rangle) \\ & \qquad \mathbf{snd}\ y \in S \times R.\mathbf{inr}\,(\langle \mathbf{fst}(y), \\ & \qquad\qquad PAR(r)(\mathbf{snd}(y))(\mathbf{snd}(x)) \rangle) \end{aligned}$$

and then,

$$\mathcal{C}[\![c \parallel c']\!] = PAR(\mathcal{C}[\![c]\!])(\mathcal{C}[\![c']\!])(\kappa)$$

It would be interesting to investigate this semantics in detail especially comparing it with a random walk operational semantics. It would also be interesting to give a probabilistic semantics to a language like Milner's CCS.

8 Measure Theory

From the point of view of probability theory it is more natural to use measures than evaluations. In fact we can obtain a similar construction with measures to the one with evaluations. We use the Borel sets of the Scott topology on an ipo, which form the least σ-field generated by the open sets, well-known in measure theory.

Every $[0,1]$-valued measure restricts to an evaluation $\tau_P(\mu)$ and this is a $1-1$ map as measures are determined by their values on open sets.

Definition For an ipo $P, \mathcal{D}(P)$ is the set of $[0,1]$-valued measures on the Borel sets of P that restrict to a continuous evaluation. It is partially ordered by: $\mu \sqsubseteq \eta$ iff $\tau_P(\mu) \sqsubseteq \tau_P(\eta)$.

Theorem 8.1 *$\mathcal{D}(P)$ is an ipo with least element.*

The theorem is proved by showing that for a directed set of measures $\{\mu_i\}$ that the evaluation $\sqcup_i \tau_P(\mu_i)$ extends to a measure (and so restriction is continuous).

We can show $\mathcal{D}$ is a monadic functor, the definitions are as before, substituting Borel sets for open ones:

$$\eta_P(x)(T) = \begin{cases} 1 & \text{if } x \text{ is in } T \\ 0 & \text{otherwise} \end{cases}$$

and

$$f^*(\mu)(T) = \int_{x \in P} f(x)(T)\, d\mu$$

Note if f is continuous, $\lambda x.f(x)(T)$ is measurable and this equation defines a measure by linearity and continuity of integration.

Restriction is a morphism $\tau : \mathcal{D} \to \mathcal{E}$ of monads, and we consider when the two monads coincide in that τ is an isomorphism. This is equivalent to asking when an evaluation can be extended to a measure. By a result of Lawson [7] this holds whenever $\Omega(P)$ is ω-continuous and so when P is. The next theorem generalises to the continuous case and yields further information on evaluations.

Theorem 8.2 *If P is continuous, every continuous evaluation extends to a measure and $\mathcal{D}(P)$ is continuous (with basis the finite rational linear combinations of point measures of any given basis of P).*

Proof Every linear combination of point evaluations clearly extends to a measure so every directed lub of such measures extends to the lub of the corresponding extensions (as τ is continuous). One then shows that every evaluation is such a lub (this was shown for the ω-algebraic case [16] by Saheb-Djahromi [14]). As every continuous ipo is a retract of an algebraic one [4], we can restrict our attention to the case where P is algebraic. The key lemma is the following:

Lemma 8.3 *Let ν be a continuous evaluation, let $O_1, \ldots, O_n$ be open sets and choose r with $0 < r < 1$. Then there is a linear combination μ of point measures of finite elements of P such that:*

1. *$\mu(O_i) \geq r\nu(O_i)$ (for $i = 1, \ldots, n$).*
2. *$\mu(V) \ll \nu(V)$ (for every open set V).*

Proof First we show how to approximate ν on a crescent $C = U \setminus V$ (U and V are open). As ν is continuous and P is algebraic,

$$\bar{\nu}(C) = \sup_{a_1, \ldots, a_n \in C^\circ} \bar{\nu}(C \cap \bigcup_i a_i\!\uparrow)$$

where C° is the set of finite elements of C and $a_i\!\uparrow$ is $\{x \in P \mid x \sqsupseteq a_i\}$. Now pick $a_1, \ldots, a_n$ in C° so that $\bar{\nu}(C \cap \bigcup_i a_i\!\uparrow) \geq \sqrt{r}\bar{\nu}(C)$. Put

$$\mu_C = \sum_{i=1}^{n} \sqrt{r}\bar{\nu}(C \cap (a_i\!\uparrow \setminus \bigcup_{j<i} a_j\!\uparrow))\eta_P(a_i)$$

we get $\bar{\mu}_C(C) \geq r\bar{\nu}(C)$ and $\mu_C(V) \ll \bar{\nu}(V \cap C)$ for every open set V.

Now let $\mathcal{P}$ be the partition of P into the crescents $(\bigcup_{i \in S} O_i) \setminus (\bigcup_{i \notin S} O_i)$ (where $S \subset \{1, \ldots, n\}$). Then the required μ is the sum of the μ_C over the C in $\mathcal{P}$.∎

We need only show now that the measures given by the lemma form a directed set. To this end we use a lemma to the effect that if

μ is a linear combination of point measures of finite elements of P then there are open sets $U_1, \ldots, U_m$ ($m \geq 0$) such that $\mu(U_i) \neq 0$ ($i = 1, \ldots, m$) and for any measure μ', $\mu \sqsubseteq \mu'$ iff $\mu(U_i) \leq \mu'(U_i)$ for $i = 1, \ldots, m$. Now let μ_1, μ_2 be two measures as in the lemma and let $U_1^1, \ldots, U_{m_1}^1, U_1^2, \ldots, U_{m_2}^2$ be the corresponding sequences of open sets. Set $r_j = \max_{i=1,m_j}(\mu_j(U_i^j)/\nu(U_i^j))$ (which is well-defined), take $r = \max(r_1, r_2)$ and apply the lemma to obtain the desired μ.

Finally, as every measure is a directed lub of linear combinations of point evaluations, it is enough to show that, in turn, these are directed lubs of elements of the proposed basis way below them.∎

9 A Finitary Characterisation

By general nonsense, $\mathcal{E}(P)$ is the free $\mathcal{E}$-algebra over P. Generalising [3] we characterise $\mathcal{E}(P)$ in finitary terms for continuous P by proving in theorem 9.1 that the categories of $\mathcal{E}$-algebras and finitary ones coincide.

Definition An *abstract probabilistic domain* is an ipo P with least element $\perp$ and a continuous function $+ : [0,1] \times P^2 \to P$ satisfying

1. (Commutativity) $a +_r b = b +_{1-r} a$
2. (Associativity) $(a +_r b) +_s c = a +_{rs} (b +_{\frac{s(1-r)}{1-sr}} c)$ $(sr < 1)$
3. (Absorption) $a +_r a = a$
4. (One) $a +_1 b = a$

Here $[0,1]$ has the usual Hausdorff topology and P, P^2 the Scott topologies.

We form the category **APD** with abstract probabilistic domains as the objects and continuous functions which are linear, (i.e. $f(\perp) = \perp$ and $f(a +_r b) = f(a) +_r f(b)$) as morphisms. Then there is a forgetful functor $U : \mathbf{IPO}^{\mathcal{V}} \to \mathbf{APD}$, where $U(P, \alpha) = (P, \perp, +)$ where $\perp = \alpha(W \mapsto 0)$ and $a +_r b = \alpha(r\eta_P(a) + (1-r)\eta_P(b))$ and U is the identity on morphisms.

Theorem 9.1 *U cuts down to an isomorphism on the full subcategory of objects whose domains are continuous.*

Proof One difficulty in proving this is to define the object part of the inverse functor (it will be the identity on morphisms). To that end, given $(P, \perp, +)$ with P continuous we want to define $\alpha : \mathcal{E}(P) \to P$ by

$$\alpha(\mu) = \bigsqcup\{\sum_{b \in B} r_b b \mid \sum_{b \in B} r_b \eta_P(b) \ll \mu\}$$

where the finite weighted sums are defined using $+$ and $\perp$. The following key lemma ensures the set on the right is directed.

Lemma 9.2 *Let P be an ipo. Then for finite sets B and C $\sum_{b \in B} s_b \eta_P(b) \sqsubseteq \sum_{c \in C} t_c \eta_P(c)$ iff each sum can be divided into an equal number of parts in correspondence so that the corresponding parts are of the form $r\eta_P(x)$ and $s\eta_P(y)$ with $r \leq s$ and $x \sqsubseteq y$.*

Proof One half is by the monotonicity of the operation of taking finite linear combinations. For the other we require a set of values $r_{b,c}$ with the three properties: $r_{b,c} = 0$ unless $b \sqsubseteq c$, $\sum_{c \in C} r_{b,c} = s_b$ and $\sum_{b \in B} r_{b,c} \leq t_c$.

We prove the theorem by applying the Max-cut, Min-flow theorem for directed graphs [1]. We have a node for each $c \in C$ and $b \in B$, a source and a sink. We link the source to each b with an edge of capacity s_b and each c to the sink with an edge of capacity t_c, then we link b and c by an edge with capacity 1 if $b \sqsubseteq c$. A flow of value $\sum_{b \in B} s_b$ gives a value for $r_{b,c}$ as the flow from b to c. By the theorem, such a flow exists iff every cut has value at least $\sum_{b\in B} s_b$. (A cut is just a set of nodes, containing the source but not the sink, and its value is the sum of the capacities of all edges from a node in the cut to one not in it.) We now call upon a further lemma, that if $\sum_{b\in B} s_b\eta(b) \sqsubseteq \sum_{c\in C} t_c\eta(c)$, then for any subset $K \subset B$ upward closed in B,

$$\sum_{b\in K} s_b \leq \sum_{c\in C \text{ s.t. } \exists b\in K, b\sqsubseteq c} t_c$$

which is proved by considering the values of the measures on an open set containing only those b's and c's greater than (or equal to) something in K.

Then one can show that any cut has value at least $\sum_{b\in B} s_b$ by first disregarding any cut whose value contains a edge of capacity 1 since its total value will then be greater than $\sum_{b\in B} s_b$. So, given a cut L which, if it contains some b and $b \sqsubseteq c$ must also contain c, we take as K, the upper closure in B of the b's in the cut and the value of the cut is $\sum_{b\notin L} s_b+\sum_{c\in L} t_c$ but the second term is at least $\sum_{b\in K} s_b$ by the lemma above, hence the value of the cut is at least $\sum_{b\in B} s_b$ as required.∎

The main difficulty in the rest of the proof is to show that α is linear in the sense that $\alpha(\sum_i r_i\mu_i) = \sum_i r_i\alpha(\mu_i)$. To this end one uses a variant of of lemma 9.2 that holds for continuous P with $\ll$ replacing both $\sqsubseteq$ and $\leq$.∎

References

[1] Bela Bollobas. *Graph Theory, an Introductory Course.* Springer-Verlag, 1979.

[2] M. Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Categorical Aspects of Topology and Analysis.* Springer-Verlag, Lecture Notes in Mathematics 915, 1981.

[3] Steven K. Graham. Closure properties of a probabilistic domain construction. Computer Science Program, University of Missouri, 1987.

[4] P. T. Johnstone. *Stone Spaces*, volume 3 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, 1982.

[5] J. L. Kelley and T. P. Srinivasan. *Measure and Integral, Volume 1*, volume 116 of *Graduate Texts in Mathematics.* Springer-Verlag, 1988.

[6] D. Kozen. Semantics of probabilistic programs. *Journal of Computer and System Sciences*, 22, 1981.

[7] J. D. Lawson. Valuations on continuous lattices. In Rudolf-Eberhard Hoffman, editor, *Continuous Lattices and Related Topics.* Universitat Bremen, 1982. Mathematik Arbeitspapiere 27.

[8] F. N. Lawvere. The category of probabilistic mappings. Preprint, 1962.

[9] Ernest G. Manes. *Algebraic Theories.* Graduate Texts in Mathematics 26. Springer-Verlag, 1976.

[10] E. Moggi. Computational lambda-calculus and monads. To appear in LICS 1989.

[11] E. Moggi. Partial morphisms in categories of effective objects. *Information and Computation*, 76(2/3), 1988.

[12] B. J. Pettis. On the extension of measures. *Ann. of Math.*, 54, 1951.

[13] G. D. Plotkin. Probabilistic powerdomains. In *Proceedings CAAP*, 1982.

[14] N. Saheb-Djahromi. CPO's of measures for non-determinism. *Theoretical Computer Science*, 12(1), 1980.

[15] M. B. Smyth. Powerdomains and predicate transformers: a topological veiw. In J. Diaz, editor, *Proceedings ICALP 1983, LNCS 154*. Springer-Verlag, 1983.

[16] M. B. Smyth and G. D. Plotkin. The category-theoretic solution of recursive domain equation solutions. *SIAM J. Computing*, 11(4), November 1982.

[17] Shinichi Yamada. A mathematical theory of semantic domains of randomized algorithms. Waseda University, Tokyo, 1987.