Probabilistic Observations and Valuations (Extended Abstract) [1]

Matthias Schröder Alex Simpson

LFCS, School of Informatics,
University of Edinburgh, UK

Abstract

We give a universal property for an "abstract probabilistic powerdomain" based on an analysis of observable properties of probabilistic computation. The universal property determines an abstract notion of integration satisfying the expected equational properties. In the category of topological spaces, the abstract probabilsitic powerdomain is given explicitly as the space of continuous probability valuations with weak topology. It follows that our abstract notion of integration coincides with the usual integration with respect to probability valuations. We end by discussing how our approach might adapt to provide "abstract effect spaces" for other computational effects.

Key words: Domain theory, valuations, probability measures, integration, computational effects

1 Introduction

Topological spaces provide a mathematical model of the notion of datatype, with open sets correponding to "observable properties" of data [7]. In particular, Sierpinski space $\mathbb{S} = \{\bot, \top\}$ acts as a result space for "observations", where $\top$ represents a computation that halts, and $\bot$ represents one that loops. The Sierpinski topology arises naturally: $\{\top\}$ is open because termination is observable, whereas $\{\bot\}$ is not open because nontermination is not observable.

In the context of probabilistic computation, termination occurs with some probability $\lambda \in [0, 1]$. Thus it is natural to replace $\mathbb{S}$ with $[0, 1]$ as test space, where λ represents a computation that halts with probability λ and loops with probability $1 - \lambda$. The sensible observable properties in this case are the sets

$$\{[0, 1]\} \cup \{(\lambda, 1] \mid 0 \leq \lambda \leq 1\} \ .$$

[1] Research Supported by an EPSRC Research Grant "Topological Models of Computational Metalanguages" and an EPSRC Advanced Research Fellowship (Simpson).

These sets determine the topology of lower semicontinuity (equivalently Scott topology) on $[0, 1]$. We write $\mathbb{I}_<$ for this space, and consider it to be our basic space of probabilistic observations

In this abstract we explain how such probabilistic observations can be used to induce an abstract characterization of a probabilistic powerdomain $\mathcal{P}_{\mathrm{prob}}(X)$ over an arbitrary topological space X. The characterization gives a universal property for $\mathcal{P}_{\mathrm{prob}}(X)$, which can be used to define integration and establish its basic properties independently of any concrete construction of $\mathcal{P}_{\mathrm{prob}}(X)$. Nevertheless, $\mathcal{P}_{\mathrm{prob}}(X)$ can be described explicitly: it is the space of probabilistic continuous valuations over X, endowed with the weak topology. It follows that the abstract theory of integration for $\mathcal{P}_{\mathrm{prob}}(X)$ developed here coincides with the established theory for valuations.

2 Abstract Probabilistic Powerdomains

The probabilistic powerdomain $\mathcal{P}_{\mathrm{prob}}(X)$ should model a notion of probabilistic process outputting values in X. Our aim is to characterize $\mathcal{P}_{\mathrm{prob}}(X)$ in terms of its expected properties without specifying details of its construction.

A minimal requirement on a topological space Y for it to model a sensible collection "probabilistic processes" is that the collection of such processes should be closed under fair probabilistic choices. Thus for processes $\mu_1, \mu_2 \in Y$ there should be a process $\mu_1 \oplus \mu_2 \in Y$ representing the process that tosses an unbiased coin and then, depending on the outcome of the toss, continues either as process μ_1 or as process μ_2. Moreover, there is a uniform (computable) mechanism of going from the pair μ_1, μ_2 to the process $\mu_1 \oplus \mu_2$, so the operation $\oplus \colon Y \times Y \to Y$ should be (jointly) continuous.

Henceforth, a structure $(Y, \oplus)$, where Y is a topological space and the operation $\oplus \colon Y \times Y \to Y$ is continuous is called a *choice algebra*. A *homomorphism* $h \colon (Y, \oplus) \to (Y', \oplus')$ between two choice algebras is a continuous function $h \colon Y \to Y'$ satisfying $h(x \oplus y) = h(x) \oplus' h(y)$.

The observation space $\mathbb{I}_<$ carries a natural choice algebra structure:

$$\lambda_1 \oplus \lambda_2 \;=\; \frac{1}{2}(\lambda_1 + \lambda_2) \ .$$

Henceforth, whenever we write $(\mathbb{I}_<, \oplus)$, the algebra structure is always that defined above.

The notion of choice algebra only places weak requirements on a space of probabilistic processes. For example, no equational properties are required of the $\oplus$ operation. Also, the existence of $\oplus$ alone only guarantees that A models unbiased two-way probabilistic choices, whereas there are many other forms of probabilistic choice that can arise. Nevertheless, the notion of choice algebra is sufficient to next ask:

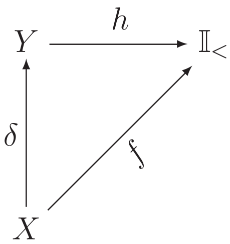
When does a choice algebra $(Y, \oplus)$ constitute a reasonable space of probabilistic processes outputting values in X?

We answer this question by placing two further requirements on A

Requirement 1 There is a distinguished (continuous) map $X \xrightarrow{\delta} Y$.

The intuition is that $\delta(x) \in Y$ is the deterministic process that outputs the value x with probability 1.

Requirement 2 For every map $X \xrightarrow{f} \mathbb{I}_<$ there exists a unique homomorphism $h\colon (Y, \oplus) \to (\mathbb{I}_<, \oplus)$ such that the diagram below commutes.



This requirement can be motivated as follows. First, given f, which performs an observation on X, we can use f to perform an observation on Y, by simply running any $\mu \in Y$ and applying f to any resulting value $x \in X$ output by μ. The so-induced observation h on Y is a homomorphism, because the probability of termination accumulates according to the probabilistic choices made during the execution of μ. Moreover, it is clear that the diagram commutes. It remains to explain the uniqueness requirement. This expresses that the only way of performing an observation h on any $\mu \in Y$ in such a way that probabilistic choices in μ are respected (i.e. so that h is a homomorphism) is by performing an observation f on the resulting values in X of μ.

We combine the requirements above into a definition.

Definition 2.1 An *abstract (probabilistic) choice structure* over X is given by a choice algebra $(Y, \oplus)$ together with a map $X \xrightarrow{\delta} Y$ such that Requirement 2 holds.

The notion of abstract choice structure suffers from the same weaknesses as the notion of choice algebra. However, we can use it to define a "completeness" property for choice algebras. This will guarantee completeness in two senses. First, the operation $\oplus$ will satisfy all the expected equational properties. Second, the space will be "complete" enough to interpret all possible forms of probabilistic choice.

Definition 2.2 A choice algebra $(A, \oplus)$ is said to be *complete* if, for every abstract choice structure $X \xrightarrow{\delta} (Y, \oplus)$ and map $X \xrightarrow{f} A$ there exists a unique homomorphism $(Y, \oplus) \xrightarrow{h} (A, \oplus)$ such that the diagram below commutes.

$$\begin{array}{ccc} Y & \xrightarrow{h} & A \\ {\scriptstyle\delta}\uparrow & \nearrow{\scriptstyle f} & \\ X & & \end{array}$$

Note that it is immediate from the definition of abstract choice structure that the choice algebra $(\mathbb{I}_<, \oplus)$ is complete.

The definition of completeness rather directly formalizes A being a space of probabilistic processes that is complete enough to interpret all forms of probabilistic choice. Explicitly, it says that any program f mapping values in X to probabilistic processes in A, extends uniquely to a choice-respecting program h translating probabilistic processes over X to probabilistic processes in A.

We next state a sequence of results about complete choice algebras. The first result is technical, but important. It states a fundamental property needed in the proofs of several of the subsequent results.

Proposition 2.3 (Parametrization) *If $(A, \oplus)$ is a complete choice algebra then, for every abstract choice structure $X \xrightarrow{\delta} (Y, \oplus)$ and map $Z \times X \xrightarrow{f} A$, there exists a unique continuous $Z \times Y \xrightarrow{h} A$, homomorphic in its right argument, such that:*

$$\begin{array}{ccc} Z \times Y & \xrightarrow{h} & A \\ {\scriptstyle \mathsf{id}_Z \times \delta}\uparrow & \nearrow{\scriptstyle f} & \\ Z \times X & & \end{array}$$

If the above property did not hold, then it would be possible instead to build it directly into the notions themeselves by parametrizing the definitions of abstract choice structure and complete choicec algebra.

Proposition 2.4 *If $(A, \oplus)$ is a complete choice algebra then the topological space A is sober.*

Proposition 2.5 *The forgetful functor* **CCA** $\to$ **Top** *(where* **CCA** *is category of complete choice algebras and homomorphisms) creates limits.*

Thus, for example, $\mathbb{I} = [0, 1]$ with the Euclidean topology is a complete choice algebra, because it arises as an equalizer

$$\mathbb{I} \xrightarrow{\lambda \mapsto (\lambda, 1-\lambda)} \mathbb{I}_< \times \mathbb{I}_< \underset{(\lambda_1,\lambda_2)\mapsto \lambda_1 \oplus \lambda_2}{\overset{(\lambda_1,\lambda_2)\mapsto \frac{1}{2}}{\rightrightarrows}} \mathbb{I}_<$$

of homomorphisms.

The next proposition shows that completeness does indeed have equational consequences. In fact, complete choice algebras inherit their equational theory from $(\mathbb{I}_<, \oplus)$.

Proposition 2.6 *If $(A, \oplus)$ is a complete choice algebra then the following equations hold:*

$$x \oplus x = x$$
$$x \oplus y = y \oplus x$$
$$(x \oplus y) \oplus (z \oplus w) = (x \oplus z) \oplus (y \oplus w) \ .$$

The above proposition states that $(A, \oplus)$ is a *midpoint algebra* in the sense of [2].

Proposition 2.7 *If $(A, \oplus)$ is a complete choice algebra then the space A carries a unique continuous map $+ \colon \mathbb{I} \times A \times A \longrightarrow A$ (where $\mathbb{I}$ has the Euclidean toplogy) satisfying:*

$$x +_0 y = x$$
$$x +_\lambda x = x$$
$$x +_\lambda y = y +_{(1-\lambda)} x$$
$$x +_\lambda (y +_{\lambda'} z) = (x +_{\frac{\lambda(1-\lambda')}{1-\lambda\lambda'}} y) +_{\lambda\lambda'} z$$

such that $x +_{\frac{1}{2}} y = x \oplus y$. Thus A is a "convex space" with $x +_\lambda y$ expressing the convex combination $(1-\lambda)x + \lambda y$. Further, every homomorphism of complete choice algebras is affine (i.e. preserves convex combinations).

It is possible to also show that A has uniquely determined (continuous) countable convex combinations. Rather than pursuing this direction here, we proceed instead to giving our abstract definition of probabilistic powerdomain.

Definition 2.8 The *abstract probabilistic powerdomain* over X, if it exists, is given by an abstract choice structure $X \xrightarrow{\delta} (\mathcal{P}_{\text{prob}}(X), \oplus)$ such that $(\mathcal{P}_{\text{prob}}(X), \oplus)$ is a complete choice algebra.

The abstract probabilistic powerdomain is characterized up to isomorphism in two complementary ways:

(i) $X \xrightarrow{\delta} (\mathcal{P}_{\text{prob}}(X), \oplus)$ is final amongst abstract choice structures over X.

(ii) $X \xrightarrow{\delta} (\mathcal{P}_{\text{prob}}(X), \oplus)$ exhibits $(\mathcal{P}_{\text{prob}}(X), \oplus)$ as the free (i.e. initial) complete choice algebra over X.

Theorem 2.9 *The abstract probabilistic powerdomain over X exists, for every topological space X.*

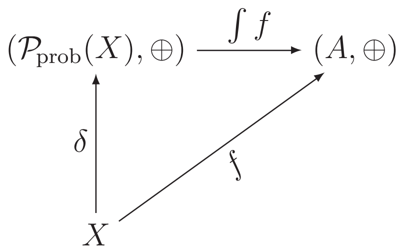
An equivalent statement is that the forgetful functor **CCA** $\to$ **Top** has a left adjoint.

We shall briefly discuss the proof of Theorem 2.9 in Section 4.

3 Abstract Integration

We have motivated $\mathcal{P}_{\text{prob}}(X)$ as an abstract space of "probabilistic processes" over X. Alternatively, one can think of it as an abstract space of "probability measures" over X, where "measure" here is, for the moment, to be understood in an intuitive rather than technical sense. In this section, we develop this viewpoint, by developing a theory of integration relative to the "probability measures" in $\mathcal{P}_{\text{prob}}(X)$.

For any complete choice algebra $(A, \oplus)$ and continuous $f\colon X \to A$, we write $\int f$ for the unique homomorphism such that the diagram below commutes, and we write $\int f\, d\mu$ for $(\int f)(\mu)$ etc.



This definition gives us a notion of integration with respect to abstract probability measures in $\mathcal{P}_{\text{prob}}(X)$, for functions taking values in any complete choice algebra A. For example, we obtain Euclidean-valued integration by taking $\mathbb{I}$ for A, and lower semicontinuous integration by taking $\mathbb{I}_<$ for A.

Many of the expected properties of integration fall out straightforwardly from the universal property of $\mathcal{P}_{\text{prob}}(X)$. It is not necessary to know any concrete description of $\mathcal{P}_{\text{prob}}(X)$.

Proposition 3.1 *Using the convex space structure of A (Proposition 2.7),*

$$\int (x \mapsto a)\, d\mu = a$$

$$\int ((1-\lambda)f + \lambda g)\, d\mu = (1-\lambda)(\int f\, d\mu) + \lambda(\int g\, d\mu) \ .$$

Proposition 3.2 (Monotonicity) *If $f \sqsubseteq g$ pointwise in the specialization order on A then $\int f\, d\mu \sqsubseteq \int g\, d\mu$.*

Proposition 3.3 (Monotone convergence) *If $\{f_d\}_{d \in D}$ is a directed set of continuous functions from X to A then*

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

using the sobriety of A (Proposition 2.4) to find the suprema.

Lemma 3.4 *For topological spaces X, Y, there is a unique continuous map $\otimes\colon \mathcal{P}_{\text{prob}}(X) \times \mathcal{P}_{\text{prob}}(Y) \to \mathcal{P}_{\text{prob}}(X \times Y)$ that is a bihomomorphism (i.e. a homomorphism in each argument separately) and satisfies $\delta(x) \otimes \delta(y) = \delta(x, y)$.*

Proposition 3.5 (Fubini) *For any continuous* $X \times Y \xrightarrow{f} A$*,*

$$\int_{x\in X}\int_{y\in Y} f(x,y)\,d\nu\,d\mu \;=\; \int_{y\in Y}\int_{x\in X} f(x,y)\,d\mu\,d\nu$$
$$= \int_{(x,y)\in X\times Y} f(x,y)\,d(\mu\otimes\nu)\ ,$$

using the operation $\otimes$ *from Lemma 3.4 in the third integral.*

4 Probabilistic Valuations

In this section we give a concrete presentation of the space $\mathcal{P}_{\mathrm{prob}}(X)$.

Definition 4.1 A *(continuous) probability valuation* on a space X is a continuous function $\nu\colon \mathcal{O}(X) \to \mathbb{I}_<$ (where $\mathcal{O}(X)$ is the lattice of open sets of X endowed with the Scott topology) satisfying:

(i) $\nu(\emptyset) = 0$

(ii) $\nu(U) + \nu(V) = \nu(U \cup V) + \nu(U \cap V)$ *(modularity)*

(iii) $\nu(X) = 1$.

We write $\mathcal{V}_1(X)$ for the set of probability valuations on X, and we give it the *weak topology* (cf. [4]), which has subbasic opens

$$\{\nu \mid \nu(U) > \lambda\}$$

generated by open $U \subseteq X$ and $\lambda \in [0,1)$. Define $X \xrightarrow{\delta} \mathcal{V}_1(X)$ by:

$$\delta(x)(U) = \begin{cases} 1 \text{ if } x \in U \\ 0 \text{ if } x \notin U \end{cases}$$

and $\mathcal{V}_1(X) \times \mathcal{V}_1(X) \xrightarrow{\oplus} \mathcal{V}_1(X)$ by:

$$(\nu_1 \oplus \nu_2)(U) = \nu_1(U) \oplus \nu_2(U)$$

Theorem 4.2 *For any topological space* X*, the structure* $X \xrightarrow{\delta} (\mathcal{V}_1(X), \oplus)$ *is an abstract probabilistic powerdomain over* X*.*

The proof is quite involved. We mention only that the full Axiom of Choice is used to prove the uniqueness part of Requirement 2 in order to show that $X \xrightarrow{\delta} (\mathcal{V}_1(X), \oplus)$ is an abstract choice structure. A similar argument also provides a positive solution to Problem 1 of [4]. The details will appear in a full version of this extended abstract.

Of course, Theorem 2.9 follows from Theorem 4.2. Using results from the literature, it also follows that, for good classes of spaces, the abstract probabilistic powerdomain coincides with standard constructions of probability spaces.

In domain theory [3], one works with dcpos with the Scott topology. For any dcpo D the set of probability valuations again $\mathcal{V}_1(D)$ again forms a dcpo, cf. [5]. We refer to $\mathcal{V}_1(D)$ with the Scott topoplogy as the *domain-theoretic probabilistic powerdomain*. (In contrast to [5], we are considering probability valuations rather than subprobability valuations.)

Corollary 4.3 *The abstract probabilistic powerdomain over a continuous pointed dcpo D carries the Scott topology, and hence coincides with the domain-theoretic probabilistic powerdomain over D.*

This result follows from [4]

In analysis, for any compact Hausdorff space X, one considers a space $\mathcal{M}_1(X)$ of *regular Borel probability measures* (also known as *Radon measures*) endowed with the *weak topology* (also known as the *vague topology*), which is again compact. This construction generalizes further to *stably compact spaces*, which are the T_0 analogues of compact Hausdorff spaces citeajk.

Corollary 4.4 *The abstract probabilistic powerdomain over a stably compact space X is homeomorphic to the space $\mathcal{M}_1(X)$ of regular Borel probability measures with the weak topology.*

This follows easily from results in [1].

It would be interesting to generalize the second corollary above to include all locally compact sober spaces, since this would then subsume both locally compact Hausdorff spaces from analysis and continuous dcpos from domain theory.

5 Other Computational Effects

The general approach we have taken to characterizing $\mathcal{P}_{\mathrm{prob}}(X)$ has nothing to do with probability! It potentially adapts to other “computational effects”, so long as these are created by a collection of algebraic operations in the sense of Plotkin and Power [6].

We assume a signature Σ of basic operations and a (topological) Σ-algebra $\mathcal{O}$ for the signature, acting as algebra of observations. Given this one can define successively

abstract effect structure — analogously to abstract choice structure,

complete Σ-algebra — analogously to complete choice algebra, and

abstract effect space — analogously to abstract probabilistic powerdomain.

In the case dealt with above, Σ contains just one binary operation, $\oplus$, and $\mathcal{O}$ is the algebra $(\mathbb{I}_<, \oplus)$.

Other forms of nondeterministic choice are potentially addressed by retaining a single binary operation, but varying the observation algebra. Using $(\mathbb{S}, \vee)$ One should obtain a "lower powerdomain" using $(\mathbb{S}, \vee)$ for observations (recall that $\mathbb{S}$ is Sierpinski space), an "upper powerdomain" using $(\mathbb{S}, \wedge)$, and a "convex powerdomain" using $(\{\{\bot\}, \{\bot, \top\}, \{\top\}\}, \cup)$.

Other examples require different signatures. For example, for nontermination, only a single constant $\bot$ is needed, and $(\mathbb{S}, \bot)$ is the natural observation algebra. To combine nontermination and probabilistic choice, take the signature containing one binary operation for probabilistic choice and one constant for nontermination and use $(\mathbb{I}_<, \oplus, 0)$ for the observation algebra. For this example, we have calculated the associated abstract effect space in **Top**, and proved that it is $\mathcal{V}_{\leq 1}(X)$ of *subprobability valuations* on X (again with the weak topology).

Finally, we mention that none of the basic ideas above are at all dependent on working in **Top** as the ambient category. The notion of abstract effect space makes sense in any category with finite products, all that is needed is a chosen object representing an algebra of observations. It would be interesting to see if there are other interesting mathematical constructions that can be captured as abstract effect spaces in appropriate categories.

Postscript

While producing this extended abstract for the MFPS proceedings we learnt with sadness of the untimely death of Claire Jones in October 2005. In her PhD thesis of 1990 [5], Claire established the definition of probabilistic powerdomain for arbitrary dcpos, and proved many of the fundamental results in the area. Claire was always modest about the achievements of her PhD. The second author recalls telling her several times how much he liked her thesis, to which she would always respond: "Ah, but have you read it? There's not much in it!" Time has told a different story. Fifteen years on, Claire's thesis rightly remains the primary reference in the field.

References

[1] Alvarez-Manilla, M., A. Jung and K. Keimel, *The probabilistic powerdomain for stably compact spaces*, Theoretical Computer Science **328** (2004), p. 221.

[2] Escardó, M. and A. Simpson, *A universal characterisation of the closed Euclidean interval*, in: *Proceedings of 16th Annual IEEE Symposium on Logic in Computer Science*, 2001, pp. 115–125.

[3] Gierz, G., K.H.Hofmann, K.Keimel, J.D.Lawson, M. Mislove and D. Scott, "Continuous Lattices and Domains," CUP, 2003.

[4] Heckmann, R., *Spaces of valuations*, Annals of the New York Academy of Sciences **806** (1996), pp. 174–200.

[5] Jones, C., "Probabilistic Non-determinism," PhD thesis, University of Edinburgh, 1990.

[6] Plotkin, G. and A. Power, *Computational effects and operations: An overview*, Electr. Notes Theor. Comput. Sci. **73**, pp. 149–163.

[7] Smyth, M., *Topology*, , **I**, OUP, 1992 pp. 641–761.