Discrete Random Variables Over Domains, Revisited

Michael Mislove(✉)

Department of Computer Science, Tulane University, New Orleans, LA 70118, USA
mislove@tulane.edu

Abstract. We revisit the construction of discrete random variables over domains from [15] and show how Hoare's "normal termination" symbol ✓ can be used to achieve a more expressive model. The result is a natural model of flips of a coin that supports discrete and continuous (sub)probability measures. This defines a new random variables monad on BCD, the category of bounded complete domains, that can be used to augment semantic models of demonic nondeterminism with probabilistic choice. It is the second such monad, the first being Barker's monad for randomized choice [3]. Our construction differs from Barker's monad, because the latter requires the source of randomness to be shared across multiple users. The monad presented here allows each user to access a source of randomness that is independent of the sources of randomness available to other users. This requirement is useful, e.g., in models of crypto-protocols.

Keywords: Domain random variable · Sequential domain monoids · Continuous random variables

1 Introduction and Related Work

About ten years ago, the author presented a model for finite random variables over domains [15]. That model was based on work of Varacca [21,22], whose *indexed valuations* monads for probabilistic choice enjoy distributive laws over the standard power domains at the price of weakening one of the laws for probabilistic choice [11]. The model in [15] is arcane, but it nevertheless inspired an attempt to extend the ideas to a model that would support continuous probability measures over domains [9], an approach that was unfortunately flawed [16,17]. Here we present an improved construction for the model described in [15] that has the advantage of supporting all sub-probability measures – including both discrete and continuous – over sequences of flips of a random coin, yielding a new model for computational processes that involve probabilistic choice.

The last assertion can be understood by considering a natural model for sequences of coin tosses, the full binary tree, $\mathbb{CT} = \{0,1\}^* \cup \{0,1\}^\omega$. The root represents the starting point, and the n^{th} level $\mathcal{C}_n$ of n-bit words represents

M. Mislove—Work partially supported by AFOSR Grant FA9550-13-1-0135-1.

T. Gibson-Robinson et al. (Eds.): Roscoe Festschrift, LNCS 10160, pp. 185–202, 2017.
DOI: 10.1007/978-3-319-51046-0_10

the possible outcomes of n flips of the coin. A probability distribution over this family is then a probability measure on $\mathbb{CT}$. Endowed with the prefix order, $\mathbb{CT}$ is a domain whose set of maximal elements is a Cantor set, $\mathcal{C}$, and so this model is called the *Cantor tree*. If we endow $\mathcal{C}_n$ with a probability distribution μ_n representing the chances for a specific sequence of outcomes of n flips of the coin, and if $\pi_{m,n}\colon \mathcal{C}_n \longrightarrow \mathcal{C}_m$ satisfies $\pi_{m,n}\,\mu_n = \mu_m$ for $m \leq n$, then the sequence $\mu_n \longrightarrow_w \mu$ has a limit μ in the weak topology which is concentrated on $\mathcal{C}$. Likewise, any such measure μ gives rise to an associated sequence of measures, $\pi_n\,\mu$, where $\pi_n\colon \mathcal{C} \longrightarrow \mathcal{C}_n$ is the natural projection. Everything appears to be fine, until one tries to construct a monad based on these ideas, and then the construction falters when one tries to define a Kleisli lift (for details, see [16,17]).

In more detail, the flaw in the definition of the Kleisli lift in [9] was its use of concatenation of strings, which is not monotone in its first argument. Our remedy is to replace the Cantor tree $\mathbb{CT}$ with a domain monoid where composition is Scott continuous. Our construction yields a new monad on domains using a *domain monoid* $\mathbb{M}\{0,1\} = \{x\checkmark, x\bot \mid x \in \{0,1\}^*\} \cup \{0,1\}^\omega$. This domain monoid utilizes an idea first devised by Hoare that appears prominently in models of CSP [5]: a $\checkmark$-symbol that denotes normal termination. Algebraically, $\checkmark$ is an identity for multiplication, and making strings ending in $\checkmark$ maximal makes the multiplication in the monoid Scott continuous. Adding infinite strings requires a least element $\bot$ with strings ending in $\bot$ denoting terms that might diverge.

Probability is introduced by applying the sub-probability monad, $\mathbb{V}$; then monoid structure on $\mathbb{M}\{0,1\}$ then induces an affine domain monoid structure on $\mathbb{V}\,\mathbb{M}\{0,1\}$ where multiplication is convolution, induced by the monoid multiplication on $\mathbb{M}\{0,1\}$ with $\delta_\checkmark$ the identity. Moreover, since $\mathbb{M}\{0,1\}$ is a tree, it follows that $\mathbb{V}\,\mathbb{M}\,\{0,1\}$ is a bounded complete domain (cf. Corollary 1). The remainder of the construction follows along lines similar to those in [9]. However, restricting to random variables defined only on antichains as in [3,9] is not necessary for our construction, and this simplifies things somewhat.

1.1 The Plan of the Paper

The next section begins with a review of some background material from domain theory and other areas we need, including a result about the probability monad on the category of compact Hausdorff spaces and continuous maps and on the subcategory of compact monoids and continuous monoid homomorphisms. We also introduce our new "sequential domain monoid" construction, which is inspired by sequential composition from the process calculus world, and which forms a monad $\mathbb{M}$ on various categories of domains. Then we show that following $\mathbb{M}$ with the subprobability monad $\mathbb{V}$ yields a monad that supports convolution of subprobability measures as a Scott-continuous operation. While the facts that $\mathbb{M}$ and $\mathbb{VM}$ are monads are not necessary to show the main results of the paper, we include them to show that our constructions are canonical. In any case, Sect. 3 contains the main results of the paper, where we give the construction of our new monad, CRV, of random variables, and the paper concludes with a summary and comments about future work.

2 Background

2.1 Domains

Most of the results we need about domain theory can be found in [1] or [8]; we give specific references for those that appear elsewhere.

To start, a *poset* is a partially ordered set. A subset $S \subseteq P$ is *directed* if each finite subset of S has an upper bound in S, and P is *directed complete* if each of P's directed subsets has a least upper bound. A directed complete partial order is called a *dcpo*. The relevant maps between dcpos are the monotone maps that also preserve suprema of directed sets; these maps are usually called *Scott continuous*. The resulting category is denoted DCPO.

These notions can be presented in a purely topological fashion: a subset $U \subseteq P$ of a poset is *Scott open* if (i) $U = \uparrow U \equiv \{x \in P \mid (\exists u \in U)\ u \leq x\}$ is an upper set, and (ii) if $\sup S \in U$ implies $S \cap U \neq \emptyset$ for each directed subset $S \subseteq P$. It is routine to show that the family of Scott-open sets forms a topology on any poset; this topology satisfies $\downarrow x \equiv \{y \in P \mid y \leq x\} = \overline{\{x\}}$ is the closure of a point, so the Scott topology is always T_0, but it is T_1 iff P is a flat poset. In any case, a mapping between dcpos is Scott continuous in the order-theoretic sense iff it is a monotone map that is continuous with respect to the Scott topologies on its domain and range. The category DCPO is Cartesian closed.

If P is a poset, and $x, y \in P$, then x *approximates* y iff for every directed set $S \subseteq P$, if $\sup S$ exists and if $y \leq \sup S$, then there is some $s \in S$ with $x \leq s$. In this case, we write $x \ll y$ and we let $\twoheaddownarrow y = \{x \in P \mid x \ll y\}$. A *basis* for a poset P is a family $B \subseteq P$ satisfying $\twoheaddownarrow y \cap B$ is directed and $y = \sup(\twoheaddownarrow y \cap B)$ for each $y \in P$. A *continuous poset* is a poset that has a basis, and a dcpo P is a *domain* if P is a continuous dcpo. An element $k \in P$ is *compact* if $x \ll x$, and P is *algebraic* if $KP \equiv \{k \in P \mid k \ll k\}$ forms a basis. Domains are sober spaces in the Scott topology (cf. [14]).

We let DOM denote that category of domains and Scott continuous maps; this is a full subcategory of DCPO, but it is not Cartesian closed. Nevertheless, DOM has several Cartesian closed full subcategories. For example, there are the full subcategories SDOM of Scott domains, and BCD, its continuous analog: a *Scott domain* is an algebraic domain P for which KP is countable, and every non-empty subset of P has a greatest lower bound, or equivalently, every subset of P with an upper bound has a least upper bound. A domain is *bounded complete* if every non-empty subset has a greatest lower bound; BCD denotes the category of bounded complete domains and Scott-continuous maps.

Domains also have a Hausdorff refinement of the Scott topology which will play a role in our work. The *weak lower topology* on a poset P has the sets of the form $O = P \setminus \uparrow F$ as a basis, where $F \subset P$ is a finite subset. The *Lawson topology* on a domain P is the common refinement of the Scott- and weak lower topologies on P. This topology has the family

$$\{U \setminus \uparrow F \mid U \text{ Scott open } \& \ F \subseteq P \text{ finite}\}$$

as a basis. The Lawson topology on a domain is always Hausdorff. A domain is *coherent* if its Lawson topology is compact. We denote the closure of a subset

$X \subseteq P$ of a domain in the Lawson topology by $\overline{X}^{\Lambda}$, and Coh denotes the category of coherent domains and Scott-continuous maps. While the subcategory of Coh of coherent domains is Cartesian, and the subcategory of coherent domains having least elements is closed under arbitrary products, the category Coh is not Cartesian closed.

Example 1. This example is used extensively in [3,9]. Let $\mathcal{C}$ denote the middle third Cantor set from the unit interval. This is a Stone space, and so it can be realized as a projective limit of finite spaces $\mathcal{C} \simeq \varprojlim_{\alpha \in A} \mathcal{C}_\alpha$. Since $\mathcal{C}$ is second countable, we can define a countable family of finite spaces $\mathcal{C}_n$ for which $\mathcal{C} \simeq \varprojlim_n \mathcal{C}_n$. Indeed, we can take $\mathcal{C} = \{0,1\}^\omega$ and $\mathcal{C}_n = \{0,1\}^n$ for each n.

From a domain-theoretic perspective, $\mathbb{CT} = \bigcup_n \mathcal{C}_n \cup \mathcal{C} = \{0,1\}^* \cup \{0,1\}^\omega$, the finite and infinite words over $\{0,1\}$ in the prefix order. The finite words form the set of compact elements, $K\mathbb{CT}$, and so $\mathbb{CT}$ is an algebraic domain. It is called the *Cantor Tree*, and it can be viewed as the state space of the outcomes of flipping a coin: the root is the starting point, and with 0 denoting *Tails* and 1 *Heads*, the outcomes as we work our way up the tree give all possible results of flipping a coin some number of times. For example, the family $\mathbb{CT}_n = \bigcup_{m \leq n} \mathcal{C}_m$ gives the finite tree of possible outcomes of n flips of the coin.

As we commented in the introduction, $\mathbb{CT}$ is alluring as a model for the outcomes of tossing a coin, but it does not work well as a computational model. In particular, viewing $\mathbb{CT}_n$ as the possible outcomes of n tosses of a coin, the "obvious" mechanism to compose one sequence of tosses with another is concatenation, the operation used in [9]. But concatenation is not monotone in its first argument, and this undermines the approach. We define an alternative model of coin flips below as the family $\mathbb{M}\{0,1\}$. This is the heart of our model for probabilistic choice.

There is one technical result we will need, which comes from [8]:

Lemma 1. *If $f\colon B \longrightarrow E$ is a monotone map from a basis for a domain D into a dcpo E, then $\widehat{f}\colon D \longrightarrow E$ defined by $\widehat{f}(x) = \sup f(\mathop{\downarrow\!\!\downarrow} x \cap B)$ defines the largest Scott-continuous map below f. Moreover, if for each $x \in D$ there is a directed set $B_x \subseteq \mathop{\downarrow\!\!\downarrow} x \cap B$ with $x = \sup B_x$ and $\sup \widehat{f}(B_x) = f(x)$, then $\widehat{f}$ extends f.*

Proof. This is Lemma IV-9.23 of [8].

2.2 $\mathbb{M}\{0,1\}$ as a Domain Monoid

In this section we define a domain monoid $\mathbb{M}\{0,1\}$ based on the finite and infinite words over $\{0,1\}$.

Proposition 1. *We define $\mathbb{M}\{0,1\} \equiv (\{x\checkmark, x\bot \mid x \in \{0,1\}\}^* \cup \{0,1\}^\omega, \leq)$, where $\leq$ is defined by:*

- *If $x \in \{0,1\}^*\checkmark, y \in \mathbb{M}\{0,1\}$, then $x \leq y$ iff $x = y$;*
- *If $x \in \{0,1\}^*\bot, y \in \mathbb{M}\{0,1\}$, then $x \leq y$ iff $(\exists m \leq n < \omega)\, x \in \{0,1\}^m\bot, y \in \{0,1\}^n\bot \cup \{0,1\}^n\checkmark \cup \{0,1\}^\omega$ and $x_i \leq y_i$ for all $i \leq m$; and*
- *If $x \in \{0,1\}^\omega, y \in \mathbb{M}\{0,1\}$, then $x \leq y$ iff $x = y$.*

Then $\mathbb{M}D$ *is a bounded complete algebraic domain whose set of compact elements is* $K\mathbb{M}\{0,1\} = \{x\checkmark, x\bot \mid x \in \{0,1\}^*\}$.

Proof. It is routine to show that the partial order defined above endows $\mathbb{M}\{0,1\}$ with a tree structure whose root is $\bot$ and whose leaves (=maximal elements) are $\{x\checkmark \mid x \in \{0,1\}^*\} \cup \{0,1\}^\omega$. It's then obvious that the elements $x\checkmark$ and $x\bot$ are compact for x finite, and that each infinite word x satisfies $x = \sup_n x_1 \cdots x_n\bot$.

Theorem 1. *Endowed with the Lawson topology,* $(\mathbb{M}\{0,1\}, \leq, \cdot)$ *is a compact ordered monoid under the multiplication given by:*

$$x \cdot y = \begin{cases} x'y, & \textit{if } x = x'\checkmark \in \{0,1\}^*\checkmark, \\ x, & \textit{if } x \in \{0,1\}^*\bot \cup \{0,1\}^\omega. \end{cases}$$

Proof. Proposition 1 implies $\mathbb{M}\{0,1\}$ is a bounded complete algebraic domain, which implies it is coherent. If $x_1 < x_2 \in \mathbb{M}\{0,1\}$, then $x_1 \in \{0,1\}^*\bot$, so $x_1 \cdot y_1 = x_1 < x_2 \leq x_2 \cdot y_2$ for any $y_1 \leq y_2$. On the other hand, if x_1 is maximal, then $x_1 = x_2$. If $x_1 \in \{0,1\}^*\checkmark$, then $x_1 \cdot y_1 = x'y_1 \leq x'y_2 = x_1 \cdot y_2$, if $y_1 \leq y_2$. And if $x_1 \in \{0,1\}^\omega$, then $x_1 \cdot y_1 = x_1 = x_1 \cdot y_2$. It follows that the multiplication is monotone. By definition, $\checkmark$ is an identity for the multiplication. So it only remains to prove multiplication is jointly Lawson continuous.

It's straightforward to show multiplication is Scott continuous in each variable separately, which implies it is jointly Scott continuous. For Lawson continuity, it's sufficient to show that, given $z \in K\mathbb{M}\{0,1\}$, $A = \{(x,y) \mid x \cdot y \in \uparrow z\}$ is Scott compact. But $z \in K\mathbb{M}\{0,1\}$ implies $z = z'\checkmark$ or $z = z'\bot$, for a finite $z' \in \{0,1\}^*$. From this if follows that there are only finitely many ways to write z' is a concatenation of a prefix $p \in \{0,1\}^*$ and a suffix $s \in \{0,1\}^*$, and then

$$z = \begin{cases} p\checkmark \cdot s\bot & \text{if } z \in \{0,1\}^*\bot, \\ z = p\checkmark \cdot s\checkmark & \text{if } z \in \{0,1\}^\checkmark. \end{cases}$$

Then $z \leq x \cdot y$ implies there is some factorization $z = p\checkmark \cdot s\bot$ or $z = p\checkmark \cdot s\bot$ with $p\checkmark = x$ and either $s\checkmark \leq y$ or $s\bot \leq y$. Then A is a finite union of sets of the form $\uparrow(p\checkmark, s'\checkmark)$ or $\uparrow(p\checkmark, s\bot)$.

2.3 The Subprobability Monad

Probability on Comp and Dom. It is well known that the family of probability measures on a compact Hausdorff space is the object level of a functor which defines a monad on Comp, the category of compact Hausdorff spaces and continuous maps (Theorem 2.13 of [7]). As outlined in [10], this monad gives rise to two related monads:

1. On Comp, it associates to a compact Hausdorff space X the free *barycentric algebra* over X, the name deriving from the counit $\epsilon\colon \mathsf{Prob}(S) \longrightarrow S$ which assigns to each measure μ on a probabilistic algebra S its barycenter $\epsilon(\mu)$ (cf. Theorem 5.3 of [13], which references [20]).

2. A *compact affine monoid* is a compact monoid S for which there also is a continuous mapping $\cdot\colon [0,1] \times S \times S \longrightarrow S$ satisfying the property that translations by elements of S are affine maps (cf. Sect. 1.1ff. of [10]). On the category $\mathsf{CompMon}$ of compact monoids and continuous monoid homomorphisms, Prob gives rise to a monad that assigns to a compact monoid S the free compact affine monoid over S (cf. Corollary 7.4 of [10]).

Remarkably, these results have analogs in domain theory. Before we describe them, we first review some basic facts about (sub)probability measures on domains. Most of these results can be found [11].

Definition 1. *A* valuation *on a dcpo D is a mapping $\mu\colon \Sigma(D) \longrightarrow [0,1]$, where $\Sigma(D)$ denotes the Scott-open subsets of D, satisfying:*

Strictness: $\mu(\emptyset) = 0$.
Monotonicity: $U \subseteq V$ *Scott-open imliess* $\mu(U) \leq \mu(V)$.
Modularity: $\mu(U \cup V) + \mu(U \cap V) = \mu(U) + \mu(V),\ \forall U, V \in \Sigma(D)$,
Continuity: *If* $\{U_i\}_{i\in I} \subseteq \Sigma(D)$ *is* $\subseteq$*-directed, then* $\sup_i \mu(U_i) = \mu(\bigcup_i U_i)$.

If $\mu(D) = 1$, then μ is normalized. *We let $\mathbb{V}(D)$ denote the family of valuations on D under the pointwise order: $\mu \sqsubseteq \nu$ iff $\mu(U) \leq \nu(U)$ for all $U \in \Sigma(D)$; $\mathbb{V}_1(D)$ denotes the family of normalized valuations.*

It was first shown by Sahib-Djarhomi [19] that $\mathbb{V}(D)$ is a dcpo if D is one. The main result describing the domain structure of $\mathbb{V}(D)$ is the following:

Theorem 2 *(Splitting Lemma [11]). Let D be a domain with basis B. Then $\mathbb{V}(D)$ is a domain with a basis consisting of the simple measures with supports in B. Moreover, for simple measures $\mu = \sum_{x\in F} r_r \delta_x$ and $\nu = \sum_{y\in G} s_y \delta_y$, the following are equivalent:*

- *$\mu \leq \nu$ (respectively, $\mu \ll \nu$).*
- *There are non-negative* transport numbers *$\langle t_{x,y}\rangle_{(x,y)\in F\times G}$ satisfying:*
  1. *$r_x = \sum_{y\in G} t_{x,y}\ \forall x \in F$,*
  2. *$\sum_{x\in F} t_{x,y} \leq s_y\ \forall y \in G$,*
  3. *$t_{x,y} > 0$ implies $x \leq y$ (respectively, $x \ll y$) $\forall (x,y) \in F \times G$.*

  Moreover, if μ and ν are probability measures, then we can refine (ii) above to (ii') *$\sum_{x\in F} t_{x,y} = s_y\ \forall y \in G$.*

It is well-known that each Borel subprobability measure on a domain D gives rise to a unique valuation in the obvious way. Conversely, it was shown by Alvarez-Manilla, Edalat and Sahib-Djarhomi [2] that the converse holds, so we can identify the family of Borel subprobability measures on D with the family of valuations, including the order structure. Throughout this paper, we will refer to (sub)probability measures, rather than valuations, but the order structure is the one defined from valuations; for coherent domains, using the traditional functional-analytic approach to defining measures, the order can be realized as: For $\mu, \nu \in \mathbb{V}(D)$, $\mu \leq \nu$ iff $\int_D f d\mu \leq \int_D f d\nu$ for all $f\colon D \longrightarrow \mathbb{R}_+$ *monotone* and Lawson continuous.

Now for the analogs of (i) and (ii) at the start of this subsection:

Proposition 2. *Let D be a domain. Then*

1. *$\mathbb{V}$ defines a monad on DCPO.*
2. *$\mathbb{V}$ defines an endofunctor on Coh, the category of coherent domains and Scott-continuous maps.*
3. *If D is a domain with a Scott-continuous multiplication $\cdot\colon D \times D \longrightarrow D$ under which D is a topological semigroup, then there is a Scott-continuous* convolution operation $*\colon \mathbb{V}(D) \times \mathbb{V}(D) \longrightarrow \mathbb{V}(D)$ *defined by $(\mu * \nu)(U) = (\mu \times \nu)\{(x, y) \in D \times D \mid x \cdot y \in U\}$. Under this operation, $\mathbb{V}(D)$ is an affine topological semigroup.*

Proof. The result in (i) is contained in [11], and (ii) is from [12]. For (iii), it is well-known that the family of simple subprobability measures $\{\sum_{x\in F} r_x\delta_x \mid \sum_{x\in F} r_x \leq 1 \ \& \ F \subseteq S \text{ finite}\}$ is a semigroup under convolution if S is a semigroup. Since the operation $*$ is nothing more than $\mathbb{V}(\cdot)$, it is Scott-continuous on $\mathbb{V}(D \times D)$ if D is a domain semigroup. And since the simple measures contain a basis for $\mathbb{V}(D \times D)$, it follows that convolution is associative on all of $\mathbb{V}(D \times D)$. Thus $(\mathbb{V}(D), *)$ is a domain semigroup. The fact that $\mathbb{V}$ defines a monad on Dom means the only thing left to show is that each component of the unit $\eta\colon 1_{\mathsf{Coh}} \stackrel{\cdot}{\longrightarrow} |\ | \circ \mathbb{V}$ is a semigroup homomorphism. Since $\eta_D(d) = \delta_d$, this amounts to showing that $\delta_x * \delta_y = \delta_{x\cdot y}$ for each $x, y \in D$, for D a domain semigroup. But given $x, y \in D$, and $U \in \Sigma(D)$, we have

$$\begin{aligned}(\delta_x * \delta_y)(U) &= \delta_x \times \delta_y(\{(r, s) \in D \times D \mid r \cdot s \in U\}) \\ &= \begin{cases} 1 & \text{iff } x \cdot y \in U \\ 0 & \text{otherwise} \end{cases} = \delta_{x\cdot y}(U).\end{aligned}$$

The final claim that $\mathbb{V}(D)$ is an affine semigroup is clear.

Remark 1. There is a wealth of material on the semigroup of probability measures on a compact or locally compact (semi)group, but the assumption is invariably that the (semi)group is Hausdorff. The results above show that basic facts still hold if one generalizes to subprobability measures over domain semigroups endowed with the Scott topology. It turns out that, if the domain D is coherent and the multiplication $\cdot\colon D \times D \longrightarrow D$ is Lawson continuous, then one can retrieve the "classic" result, extended to subprobability measures.

$\mathbb{V}(\boldsymbol{D})$ is in BCD if $\boldsymbol{D}$ is a Tree Domain. Our next goal is to show that $\mathbb{V}\,\mathbb{M}\{0, 1\}$ is in BCD. Then, using a function space construction, if E is in BCD, we can define a monad of random variables over BCD. We begin with the following result; it is stated in [12], although no proof is provided; we include one here for completeness sake:

Lemma 2. *If T is a finite tree, then $\mathbb{V}(T)$ is closed under finite infima. Hence $\mathbb{V}(T) \in \mathsf{BCD}$.*

Proof. We prove that if $\mu, \nu \in \mathbb{V}(T)$, then $\mu \wedge \nu \in \mathbb{V}(T)$. We proceed by induction on $|T|$: the case that $|T| = 1$ is obvious, since $\mathbb{V}(\{*\}) \simeq [0,1]$. So suppose the result holds for $|T| \leq n$, and let T be a tree with $n+1$ elements. If $T' = T \backslash \{\bot_T\}$, then T' is a forest of k trees, $T' = \dot{\bigcup}_{i \leq k} T'_i$. The inductive hypothesis implies $\mathbb{V}(T'_i)$ is closed under finite infima for each i. So, if $\mu, \nu \in \mathbb{V}(T)$, then $\mu|_{T'_i} \wedge \mu|_{T'_i} \in \mathbb{V}(T'_i)$ for each $i \leq k$, and since the T'_is are pairwise disjoint, it follows that $\mu|_{T'} \wedge \nu|_{T'} \in \mathbb{V}(T')$. So, for any open sets $U, V \subseteq T'$, we have

$$\mu \wedge \nu\, (U \cup V) + \mu \wedge \nu\, (U \cap V) = \mu \wedge \nu\, (U) + \mu \wedge \nu\, (V).$$

The only remaining case to show $\mu \wedge \nu \in \mathbb{V}(T)$ is when $U = T$ or $\nu = T$; without loss of generality, we assume $U = T$. In that case,

$$\mu \wedge \nu\, (U \cup V) + \mu \wedge \nu\, (U \cap V) = \mu \wedge \nu\, (T) + \mu \wedge \nu\, (V) = \mu \wedge \nu\, (U) + \mu \wedge \nu\, (V).$$

Corollary 1. *If $T \simeq$ bilim$_n T_n$ is a the bilimit of finite trees, then $\mathbb{V}(T)$ is in BCD. In particular, $\mathbb{V}\,\mathbb{M}\{0,1\} \in$ BCD.*

Proof. If $T \simeq \mathsf{bilim}_n T_n$ with each T_n a finite tree, then the continuity of the functor $\mathbb{V}$ implies $\mathbb{V}(T) \simeq \lim_n \mathbb{V}(T_n)$, and since BCD is closed under limits, the result follows. The final claim follows from our remark in the proof of Proposition 1.

2.4 Domains of Partial Maps

In the last subsection we alluded to a “function space construction” that we’d need in our random variables model. We address that issue in this subsection, where we give some results about partial maps defined on the non-empty Scott-closed subsets of a domain. The results are needed for our analysis of sub-probabilities on domains.

To begin, recall that the *support* of a finite positive Borel measure μ on a topological space X is the smallest closed set $C \subseteq X$ satisfying $\mu(C) = \mu(X)$. For measures on a domain D, we let $\mathrm{supp}_\Sigma\, \mu$ denote the support of μ with respect to the Scott topology, and $\mathrm{supp}_\Lambda\, \mu$ denote the support of μ with respect to the Lawson topology. The appropriate domain for the random variables we plan to study is $\mathrm{supp}_\Sigma\, \mu$, the smallest Scott-closed subset X satisfying $\mu(D \backslash X) = 0$, where μ is the measure assigned to the domain of the random variable.

Recall that the lower power domain over a domain D is the family $\Gamma(D)$ of non-empty Scott-closed subsets of D ordered by inclusion; in fact, $\mathcal{P}_L(D) = (\Gamma(D), \subseteq)$ is the free sup-semilattice domain over D. $\mathcal{P}_L(D)$ defines a monad on every Cartesian closed category of domains; in fact, $\mathcal{P}_L(D)$ is bounded complete for any domain D. This leads us to an important construction that we need, one which we believe should be useful in applications of domains; although we have made only a cursory effort at best to locate the result, we find it surprising that we have been unable to find it in the literature.

Proposition 3. *Let D and E both be BCD. Let*

$$[D \rightharpoonup E] \overset{def}{=} \{f \colon C \longrightarrow E \mid C \in \mathcal{P}_L(D)\}$$

denote the family of Scott-continuous partial maps defined on non-empty Scott-closed subsets of D. We order $[D \rightharpoonup E]$ by

$$f \leq_L g \quad \textit{iff} \quad \textit{dom } f \subseteq \textit{dom } g \textit{ and } f \leq g|_{\textit{dom } f}.$$

Then $[D \rightharpoonup E]$ is a bounded complete domain.

Proof (Outine). The proof can be broken down into three claims:

1. *$[D \rightharpoonup E]$ is a dcpo:* Given $\mathcal{F} \subseteq [D \rightharpoonup E]$ directed, one first shows $\sup \mathcal{F} \in [C \longrightarrow E]$, where $C = \bigsqcup_{f \in \mathcal{F}} \mathrm{dom}\ f$, the Scott-closure of the union of the Scott-closed sets $\{\mathrm{dom}\ f \mid f \in \mathcal{F}\}$. This is done by noting that $X = \bigcup_{f \in \mathcal{F}} \mathrm{dom}\ f$ is a directed union of Scott-closed subsets of the domain D, so it is a lower set that has a basis, which implies X is a continuous poset. Then $F \colon X \longrightarrow E$ by $F(x) = \sup\{f(x) \mid x \in \mathrm{dom}\ f \in \mathcal{F}\}$ is well-defined because $\mathcal{F}$ is directed, so the same is true of those $f \in \mathcal{F}$ for which $x \in \mathrm{dom}\ f$. In fact, F is Scott-continuous, because, given $Y \subseteq X$ directed for which $x = \sup Y \in X$ exists, then $x \in \mathrm{dom}\ f$ for some $f \in \mathcal{F}$. Since $f \colon \mathrm{dom}\ f \longrightarrow E$ is Scott-continuous, we have $f|_{\downarrow x} \colon \downarrow x \longrightarrow E$ is Scott-continuous. Thus $F|_{\downarrow x} = \sup\{f|_{\downarrow x} \mid f \in \mathcal{F}\ \&\ x \in \mathrm{dom}\ f\}$ is the supremum of a directed family of Scott-continuous functions on $\downarrow x$, so it also is Scott-continuous on $\downarrow x$. Thus $F(\sup Y) = \sup F(Y)$ since $Y \subseteq \downarrow \sup Y = \downarrow x$. Then this continuous map extends continuously to the round ideal completion of X, and one argues this extension satisies $F = \sup \mathcal{F}$, so $[D \rightharpoonup E]$ is a dcpo.
2. *Next show $[D \rightharpoonup E]$ is a domain:* The set $[D \rightharpoonup E] = \bigcup_{C \in \Gamma(D)} [C \longrightarrow E]$ is the directed union of domains $[C \longrightarrow E]$ for $C \in \Gamma(D)$, and each of these domains has a basis, $\mathcal{B}_C \subseteq [C \longrightarrow E]$. We let $\mathcal{B} = \bigcup_{C \in \Gamma(D)} \mathcal{B}_C$ be the (directed) union of these function families. It follows that $\mathcal{B}$ is a basis for $[D \rightharpoonup E]$.
3. *Finally, validate the category claims.* If D, E are both in BCD, then given $f, g \in [D \rightharpoonup E]$, we can define $f \wedge g$ by $\mathrm{dom}\ f \wedge g = \mathrm{dom}\ f \cap \mathrm{dom}\ g$, and for $x \in \mathrm{dom}\ f \wedge g$, we let $(f \wedge g)(x) = f(x) \wedge g(x)$. That $f \wedge g$ is the inf of f and g follows from the fact that $h \leq f, g$ implies $\mathrm{dom}\ h \subseteq \mathrm{dom}\ f \cap \mathrm{dom}\ g$, and then the result is clear.

2.5 Domain Random Variables

With the results of the previous subsection in hand, we're now ready to begin our construction of domain random variables. We start with a lemma that will underpin our main result.

Lemma 3. *Let D be a domain and let $\mu, \nu \in \mathbb{V}(D)$. Then $\mu \sqsubseteq \nu$ implies $\mathit{supp}_\Sigma\, \mu \subseteq \mathit{supp}_\Sigma\, \nu$. Moreover, if $\{\mu_i\}_{i \in I} \subseteq \mathbb{V}(D)$ is directed with $\sup_i \mu_i = \mu$, then $\sup_i \mathit{supp}_\Sigma\, \mu_i = \mathit{supp}_\Sigma\, \mu$.*

Proof. For the first claim, $\mu \sqsubseteq \nu$ iff $\mu(U) \leq \nu(U)$ for each Scott-open set U. So, $\mu(U) = 0$ if $\nu(U) = 0$, and it follows that $\operatorname{supp}_\Sigma \mu \subseteq \operatorname{supp}_\Sigma \nu$. For the second claim, since $(\Gamma(D), \subseteq)$ (the family of Scott-closed subsets of D) is a domain, the first result implies $\sup_i \operatorname{supp}_\Sigma \mu_i \subseteq \operatorname{supp}_\Sigma \mu$. Conversely, $A = D \setminus \sup_i \operatorname{supp}_\Sigma \mu_i$ is Scott-open, and $\mu(A) > 0$ would violate $\sup_i \mu_i = \mu$. So $\mu(A) = 0$, which implies $\operatorname{supp}_\Sigma \mu \subseteq D \setminus A = \sup_i \operatorname{supp}_\Sigma \mu_i$.

We now define domain random variables based on a given domain D.

Definition 2. *Let D be a domain. A* domain random variable on D *is a mapping $X \colon supp_\Sigma \mu \longrightarrow E$ where μ is a subprobability measure on D and $X \colon supp_\Sigma \longrightarrow E$ is a Scott-continuous map. Given domains D and E, we define*

$$\begin{array}{rcl} RV(D,E) & \overset{def}{=} & \{(\mu, X) \in \mathbb{V}(D) \times [supp_\Sigma \mu \longrightarrow E]\} \quad where \\ (\mu, X) \leq (\nu, Y) & iff & \mu \sqsubseteq \nu \ \& \ X \leq Y|_{supp_\Sigma \mu} \end{array}$$

Proposition 4. *If D and E are domains, then*

- *$RV(D,E)$ is a dcpo.*
- *If D and E are in a CCC of domains, then $RV(D,E)$ is a domain.*
- *If $\mathbb{V}(D), D$ and E are all in a CCC C of domains, then so is $RV(D,E)$.*

Proof. The fact that the relation $\leq$ on $RV(D,E)$ is well defined follows from part (i) of Lemma 3. The proof of the first statement is straightforward, using part (ii) of Lemma 3. For the second part, first note that Proposition 3 implies $\mathbb{V}(D) \times [D \rightharpoonup E]$ is a domain since $\mathbb{V}(D)$ is one. The first part implies $RV(D,E) \subseteq \mathbb{V}(D) \times [D \rightharpoonup E]$ is closed under directed suprema. Moreover, for $(\mu, X) \in RV(D,E)$,

$$\Downarrow(\mu, X) \supseteq \{(\mu', X') \mid \mu' \ll \mu \ \& \ X' \ll X|_{\operatorname{supp}_\Sigma \mu'} \text{ in } [\operatorname{supp}_\Sigma \mu' \longrightarrow E]\},$$

and that the right-hand set is directed with supremum (μ, X). This implies $RV(D,E)$ is a domain. The third statement is then clear.

Theorem 3. *Fix a domain D. Then $RV(D, -)$ is the object level of a continuous endofunctor on DCPO that leaves each CCC of domains that contains $\mathbb{V}(D)$ invariant.*

Proof. Given a Scott-continuous map $f \colon E \longrightarrow F$ between domains E and F, define $RV(D,f) \colon RV(D,E) \longrightarrow RV(D,F)$ by $RV(D,f)(\mu, X) = (\mu, f \circ X)$. The third part of Proposition 4 then implies that $RV(D, -)$ is an endofunctor on any CCC of domains that contains D and $\mathbb{V}(D)$. This endofunctor is continuous because its components are.

3 A Monad of Continuous Random Variables

The development so far has been about domain theory with only a passing reference to a particular model of computation. We now focus more narrowly to obtain a monad of continuous random variables designed to model the prototypical source of randomness, the tosses of a fair (and also an unfair) coin. Such a model underlies the work in [3,9], for example, where the focus is on measures μ on the Cantor tree for which $\mathrm{supp}_\Lambda\, \mu$ is an antichain. We begin with a discussion around such a model.

3.1 Modeling Coin Flips

We have chosen the sequential domain $\mathbb{M}\{0,1\}$ because it provides a model for a series of coin flips that might occur during a computation. Our intuitive view is that a random choice of some element of from a semantic domain D would consist of a coin flip followed by a choice of the element from D based on the outcome. So, it is essentially a two-step process: the random process flips the coin resulting in a 0 (tails) or a 1 (heads), and then successfully terminates, adding a $\checkmark$ to the outcome, and then a random variable X is applied to this result to select the element of D. Note that a sequence of coin flips followed by a choice is a process that iterates the coin flips a prescribed number of times, represented by $x_1\checkmark \cdot x_2\checkmark \cdots x_n\checkmark = x_1 \cdots x_n\checkmark$, followed by the application of the random variable X.

3.2 The Inevitability of Nondeterminism

Our choice of $\mathbb{M}\{0,1\}$ to model coin flips naturally leads to the question of how to combine sequences of coin flips by two processes combined under sequential composition. The multiplication operation is used here, but it raises an additional issue.

Example 2. Suppose we have processes, P and Q, both of which employ probabilistic choice, and that we want to form the sequential composition $P; Q$. Let's assume P can flip a fair coin 0, 1 or 2 times, and on each toss, if there is a 0, then an action a is executed, while if there is a 1, then action b is executed, and control then is passed to Q. Likewise, suppose Q can toss a fair coin 0, 1 or 2 times and if the result is a 0, it executes an action c, while if a 1 appears, an action d is executed, and again, Q terminates normally after each execution. In our approach based on $\mathbb{M}\,\{0,1\}$, if we represent P as (μ, X), and Q as (ν, Y). then the composition $P; Q = (\mu * \nu, X \odot Y)$, where $*$ represents a convolution operator on measures, and $\odot$ an appropriate operation on the random variable components.

Consider now the value of $X \odot Y$ on the outcome of two 0s. This outcome could arise in any of three ways:

- P could terminate without making any coin tosses, and Q could toss its coin twice, then normally terminate. This would produce the value $X \odot Y(00) = cc\checkmark$;
- P could toss its coin once, pass control, and Q could toss its coin once, and terminate. This would produce $X \odot Y(00) = ac\checkmark$;
- P could toss its coin twice, pass control to Q, which terminates normally without tossing its coin. This would produce $X \odot Y(00) = aa\checkmark$.

Since we have no way of knowing which of the three possibilities occurred, we must allow $X \odot Y$ to account for all three possibilities. This means $X \odot Y(00)$ must contain all three outcomes. The traditional way of representing such multiple outcomes is through a model of nondeterministic choice, i.e., a power domain.

Notation 1. Throughout the remainder of the paper, we assume the semantic domains D where random variables take their values are bounded complete domains, and the inf-operation models probabilistic choice. Thus, such domains support a Scott-continuous nondeterministic choice operator– the inf-operation – which we denote by $\oplus_D$.

3.3 Constructing a Monad

We now focus more narrowly by restricting random variables to be defined on a particular probability space, namely, $\mathbb{M}\{0,1\}$. This amounts to restricting the functor to $RV(\mathbb{M}\{0,1\},-)$. However, this restriction is not enough to define a monad – we must restrict the measures on $\mathbb{M}\{0,1\}$ that are allowed. We do this by restricting the simple measures that are allowed, and then taking the smallest subdomain of $\mathbb{VM}\{0,1\}$ containing them.

Definition 3. *We say a simple measure $\sum_{x\in F_\mu} r_x\delta_x$ on $\mathbb{M}\{0,1\}$ is* normal *if $F_\mu \subseteq \{0,1\}^*\checkmark$. We denote the set of normal simple measures by $\mathbb{V}_N\mathbb{M}\{0,1\}$.*

Since each normal measure is concentrated on a subset of $\{0,1\}^*\checkmark \subseteq \operatorname{Max}\mathbb{M}\{0,1\}$, the suprema of a directed set of normal simple measures is another such. However, the following will be useful:

Proposition 5. *Let $\mu_n \in \mathbb{V}_N\mathbb{M}\{0,1\}$ be a sequence of normal measures. Then the following are equivalent:*

1. *$\mu_n \longrightarrow \mu$ in the weak topology on $\mathbb{VM}\{0,1\}$.*
2. *$\mu_n \longrightarrow \mu$ in the Lawson topology on $\mathbb{VM}\{0,1\}$.*
3. *The sequence $\{\inf_{m\geq n}\mu_m \mid n\geq 1\}$ satisfies $\mu = \sup_n(\inf_{m\geq n}\mu_m)$.*

Proof. From Corollary 1 we know $\mathbb{VM}\{0,1\} \in \mathsf{BCD}$, and hence it is a coherent domain. The equivalence of (i) and (ii) is then Corollary 15 of [4], while the equivalence of (ii) and (iii) is Proposition VII-3.10 of [8].

Theorem 4. *If $\mu \in \mathbb{VM}\{0,1\}$ is concentrated on $\{0,1\}^\omega$, then there are normal simple measures $\mu_n \in \mathbb{V}_N\mathbb{M}\{0,1\}$ with $\mu_n \longrightarrow \mu$ in the weak topology.*

Proof. Define $\phi_n \colon \{0,1\}^\omega \longrightarrow \{0,1\}^n\checkmark$ by $\phi_n(x_1\cdots) = x_1\cdots x_n\checkmark$. This is Lawson continuous between compact Hausdorff spaces (in the relative topologies), and then Proposition 2 of [6] implies $\mathsf{Prob}\,\{0,1\}^\omega \simeq \lim_n(\mathsf{Prob}\,\{0,1\}^n\checkmark, \phi_n)$. But the same argument *verbatim* shows $\mathbb{V}(\{0,1\}^\omega) \simeq \lim_n(\mathbb{V}(\{0,1\}^n\checkmark), \phi_n)$. Since $\mathbb{VM}\{0,1\}$ is a coherent domain and all the measures $\mu, \phi_n\,\mu$ are concentrated on $\operatorname{Max}\,\mathbb{VM}\{0,1\}$, the relative Lawson topology agrees with the weak topology on these measures.

Definition 4. *If D is a dcpo, we define the family of* random variables *on D to be*

$$CRV(D) \overset{def}{=} RV(\mathbb{M}\{0,1\}, D) = \{(\mu, X) \in \mathbb{V}\,\mathbb{M}\{0,1\} \times [supp_{\Sigma}\, \mu \longrightarrow D]\}.$$

Theorem 5. *If D is a dcpo, then so is $CRV(D)$. Moreover, if D is in BCD, then so is $CRV(D)$. Finally, CRV extends to a continuous endofunctor on BCD.*

Proof. Proposition 4 implies $CRV(D)$ is a dcpo if D is one. Together with Corollary 1, it also implies $CRV(D)$ is in BCD if D is, since $\mathbb{V}\mathbb{M}\{0,1\} \in \mathsf{BCD}$.

As for the final claim, If $f\colon D \longrightarrow E$, then define $CRV\, f\colon CRV(D) \longrightarrow CRV(E)$ by $CRV\, f(\mu, X) = (\mu, f \circ X)$. It's clear that this makes CRV a functor, and the comments above show it's an endofunctor on BCD. It's also continuous because its components are.

In the general theory we often couch the discussion in terms of sub-probability measures, with the implicit assumption that any mass unallocated is associated with nontermination. Since we have an explicit nontermination symbol in the current situation, this is a convenient place to describe the relationship between sub-probability measures and probability measures on the same domain.

Notation 2. If D is a domain, we let

$$PRV(D) = \{(\mu, X) \in \mathsf{Prob}\,\mathbb{M}\{0,1\} \times [\mathrm{supp}_{\Sigma}\, \mu \longrightarrow D]\},$$

We call $PRV(D)$ the family of *probabilistic random variables over D*.

Proposition 6. *If D is a domain in BCD, then the mapping*

$$(\mu, X) \mapsto (\mu \oplus (1 - ||\mu||)\delta_{\perp}, X)\colon CRV(D) \longrightarrow PRV(D)$$

is a closure operator on $CRV(D)$, and its image, $PRV(D)$, also is a domain in BCD. Moreover, a basis for $PRV(D)$ is the family $\{(\mu, X) \in PRV(D) \mid \mu$ is simple$\}$.

Proof. It's straightforward to show that the mapping $\mu \mapsto \mu + (1 - ||\mu||)\delta_{\perp}$ is a closure operator on $\mathbb{V}\mathbb{M}\{0,1\}$, and clearly its image is $\mathsf{Prob}\,\mathbb{M}\{0,1\}$, which is a dcpo. It follows from Corollary I-2.5 of [8] that $\mathsf{Prob}\,\mathbb{M}\{0,1\}$ is a domain, and that $\mu \ll \nu \in VD$ implies $\mu + (1 - ||\mu||)\delta_{\perp} \ll \nu + (1|||\nu||)\delta_{\perp}$. This last point implies $\mathsf{Prob}\,\mathbb{M}\{0,1\}$ has a basis of simple measures. It now follows that $(\mu, X) \mapsto (\mu + (1 - ||\mu||)\delta_{\perp}, X)$ is a closure operator on $CRV(D)$; note that $X(\perp)$ is well-defined since D is bounded complete. Thus, the image of $CRV(D)$ is $PRV(D)$, and the same result from [8] applies to finish the proof.

The Structure of $\mathbb{V}\mathbb{M}\{\mathbf{0},\mathbf{1}\}$. Since $\mathbb{M}\{0,1\} = \{x\checkmark, x\perp \mid x \in \{0,1\}^*\} \cup \{0,1\}^{\omega}$, we can exploit the structure of $\mathbb{M}\{0,1\}$, and the structure this induces on $\mathbb{V}\mathbb{M}\{0,1\}$, as follows:

Definition 5. *For each $n \geq 1$, we let $\mathbb{M}_n = \cup_{k \leq n}\{x\checkmark, x\bot \mid x \in \{0,1\}^k\}$. We also define $\pi_n \colon \mathbb{M}\{0,1\} \longrightarrow \mathbb{M}_n$ by $\pi_n(x) = \begin{cases} x & \text{if } x \in \mathbb{M}_n, \\ x_1 \cdots x_n\bot & \text{if } x \notin \mathbb{M}_n. \end{cases}$*

If $m \leq n$, let $\pi_{m,n} \colon \mathbb{M}_n \longrightarrow \mathbb{M}_m$ by $\pi_{m,n}(x) = \begin{cases} x & \text{if } x \in \mathbb{M}_m, \\ x_1 \cdots x_m\bot & \text{if } x \in \mathbb{M}_n \backslash \mathbb{M}_m. \end{cases}$

Note that $\pi_m = \pi_{m,n} \circ \pi_n$ for $m \leq n$.

Proposition 7. *$\mathbb{M}\{0,1\} \simeq \mathsf{bilim}\,(\mathbb{M}_n, \pi_{m,n}, \iota_{m,n})$, where $\iota_{m,n} \colon \mathbb{M}_m \hookrightarrow \mathbb{M}_n$ is the inclusion. Moreover, $\mathbb{V}\,\mathbb{M}\{0,1\} = \operatorname{proj}\lim_n(\mathbb{V}\mathbb{M}_n, \mathbb{V}\pi_{m,n})$.*

Proof. It's straightforward to verify that $\iota_{m,n} \colon \mathbb{M}_m \longrightarrow \mathbb{M}_n \colon \pi_{m,n}$ forms an embedding–projection pair for $m \leq n$, and then it follows that $\mathbb{M}\{0,1\} = \mathsf{bilim}\,(\mathbb{M}_n, \pi_{m,n}, \iota_{m,n})$. This implies $\mathbb{M}\{0,1\} \simeq \operatorname{proj}\lim_n(\mathbb{M}_n, \pi_{m,n})$ in the Scott topologies, and the same argument as in the proof of Theorem 1 shows this also holds for the Lawson topology. Then the same argument we used in the proof of Theorem 4 implies $\mathsf{Prob}\,\mathbb{M}\{0,1\} \simeq \lim_n(\mathsf{Prob}\,\mathbb{M}_n, \mathsf{Prob}\,\pi_{m,n})$ and $\mathbb{V}\mathbb{M}\{0,1\} \simeq \operatorname{proj}\lim_n(\mathbb{V}\mathbb{M}_n, \mathbb{V}\pi_{m,n})$.

Corollary 2. *If D is a domain, we define:*

- $CRV_n(D) = \{(\mathbb{V}\pi_n\,\mu, X|_{supp_\Sigma\,\mathbb{V}\pi_n\,\mu}) \mid (\mu, X) \in CRV(D)\}$, *and*
- $\Pi_n \colon CRV(D) \longrightarrow CRV_n(D)$ *by* $\Pi_n(\mu, X) = (\mathbb{V}\pi_n\,\mu, X|_{supp_\Sigma\,\mathbb{V}\pi_n\,\mu})$.

Then $CRV_n(D) \subseteq CRV(D)$ and $\mathbf{1}_{CRV(D)} = \sup_n \Pi_n$.

Proof. This follows from Propositions 3 and 7.

For CRV to define a monad, we have to show how to lift a mapping $h \colon D \longrightarrow CRV(E)$ to a mapping $h^\dagger \colon CRV(D) \longrightarrow CRV(E)$ satisfying the laws listed in Lemma 4 below. Corollary 2 reduces the problem to showing the following:

Given $h \colon D \longrightarrow CRV(E)$, let $h_n = \Pi_n \circ h \colon D \longrightarrow CRV_n(D)$. Then there is a mapping $h_n^\dagger \colon CRV_n(D) \longrightarrow CRV_n(E)$, satisfying the monad laws listed in Lemma 4 below for each n.

Since $CRV_n(E)$ has two components, we let $h_n^\dagger = (h_{n,1}, h_{n,2})$. Using this notation, we note the following:

If $(\sum_{x \in F} r_x \delta_x, X) \in CRV_n(D)$, then for each $x \in F$

$$h_{n,1}(X(x)) = \sum_{y \in G_x} s_y \delta_y \in \mathbb{V}\mathbb{M}_n,$$

where G_x denotes the set on which the simple measure $h_{n,1}X(x)$ is concentrated for each $x \in F$. Moreover,

$$\begin{aligned} h_{n,1}(\sum_{x\in F} r_x\delta_x, X) &= \sum_{x\in F} r_x(\delta_x * \sum_{y\in G_x} s_y\delta_y) \\ &= \sum_{x\in F}\sum_{y\in G_x} r_x s_y \delta_{x\cdot y}. \end{aligned}$$

This implies $\operatorname{supp}_\Sigma h_{n,1}(\sum_{x\in F} r_x\delta_x) = \bigcup_{x\in F} \downarrow x \cdot G_x = \bigcup_{x\in F\,\&\,y\in G_x} \downarrow(x\cdot y)$.

Definition 6. *We define* $h_n^\dagger = (h_{n,1}, h_{n,2})\colon CRV_n(D) \longrightarrow CRV_n(E)$, *where*
- $h_{n,1}(\sum_{x\in F} r_x\delta_x, X) = \sum_{x\in F} r_x(\delta_x * h_{n,1}(X(x)))$
  $= \sum_{x\in F} r_x(\delta_x * \sum_{y\in G_x} s_y\delta_y)$, *and*
- $h_{n,2}(\sum_{x\in F} r_x\delta_x, X)\colon\ supp_{\sum_{x\in F} r_x(\delta_x * \sum_{y\in G_x} s_y\delta_y)} \longrightarrow E$ *by*
  $h_{n,2}(\sum_{x\in F} r_x\delta_x, X)(z)$
  $= \wedge\{h_{n,2}(X(x'))(y') \mid z \le x'\cdot y', x' \le x \in F, y' \le y \in G_x\}$

Lemma 4. *Given* $h\colon D \longrightarrow CRV(E)$, *the mapping* $h_n^\dagger\colon CRV_n(D) \longrightarrow CRV_n(E)$ *satisfies the monad laws:*
(1) If $\eta_D\colon D \longrightarrow CRV(D)$ *is* $\eta_D(d) = (\delta_\checkmark, \chi_d)$, *then*
$\eta_D^\dagger\colon CRV_n(D) \longrightarrow CRV_n(D)$ *is the identity;*
(2) $h_n^\dagger \circ \eta_D = h_n$; *and*
(3) If $k\colon CRV(E) \longrightarrow CRV(P)$ *and* $k_n = \Pi_n \circ k$, *then* $(k_n^\dagger \circ h_n)^\dagger = k_n^\dagger \circ h_n^\dagger$.

Proof. (1) Note that $\eta_D(D) \subseteq CRV_n(D)$ for each $n \ge 1$, so $\Pi_n \circ \eta_D = \eta_D$. Then $(\eta_D^\dagger)_1(\sum_{x\in F} r_x\delta_x, X) = \sum_{x\in F} r_x(\delta_x * \delta_\checkmark) = \sum_{x\in F} r_x\delta_x$, and $(\eta_D^\dagger)_2(\sum_{x\in F} r_x\delta_x, X)(z) = \wedge\{(\eta_D)_2(X(x'))(y') \mid z \le x'\cdot y', x' \le x \in F, y' \le \checkmark\}$

$$= \begin{cases} \chi_{X(z)}(\checkmark) & \text{if } z \in \{0,1\}^*\checkmark \\ \chi_{X(z)}(\bot) & \text{if } z \in \{0,1\}^*\bot \end{cases} = X(z).$$

(2) If $h_{n,1}(d) = \sum_{x\in F} r_x\delta_x$, then
$h_{n,1}(\delta_\checkmark, \chi_d) = \sum_{x\in F} r_x(\delta_x * \delta_\checkmark) = \sum_{x\in F} r_x\delta_x$. Likewise,
$h_{n,2}(\delta_\checkmark, \chi_d)(z)$

$$\begin{aligned} &= \wedge\{h_{n,2}(\chi_d(x'))(y') \mid z \le x'\cdot y', x' \le \checkmark, y' \le y \in G_\checkmark\} \\ &= h_{n,2}(d)(y') = \begin{cases} h_{n,2}(d)(z) & \text{if } \bot < z \\ h_{n,2}(d)(\bot) & \text{if } z = \bot \end{cases} = h_{n,2}(\eta_D(d)). \end{aligned}$$

(3) $k_n^\dagger \circ h_n^\dagger(\sum_{x\in F} r_x\delta_x, X)$
$= k_n^\dagger(h_{n,1}(\sum_{x\in F} r_x\delta_x, X), h_{n,2}(\sum_{x\in F} r_x\delta_x, X))$
$= (k_{n,1}(h_{n,1}(\sum_{x\in F} r_x\delta_x, X), h_{n,2}(\sum_{x\in F} r_x\delta_x, X)),$
$k_{n,2}(h_{n,1}(\sum_{x\in F} r_x\delta_x, X), h_{n,2}(\sum_{x\in F} r_x\delta_x, X)))$.
Now, $k_{n,1}(h_{n,1}(\sum_{x\in F} r_x\delta_x, X), h_{n,2}(\sum_{x\in F} r_x\delta_x, X))$
$= k_{n,1}(\sum_{x\in F} r_x(\delta_x * (\sum_{y\in G_x} s_y\delta_y)), h_{n,2}(\sum_{x\in F} r_x\delta_x, X))$
$= k_{n,1}(\sum_{x\in F}\sum_{y\in G_x} r_x s_y \delta_{x\cdot y}, h_{n,2}(\sum_{x\in F} r_x\delta_x, X))$
$= \sum_{x\in F\,\&\,y\in G_x} r_x s_y(\delta_{x\cdot y} * k_{n,1}(h_{n,2}(\sum_{x\in F} r_x\delta_x, X)(x\cdot y)))$.
On the other hand,
$(k_n^\dagger \circ h_n)^\dagger(\sum_{x\in F} r_x\delta_x, X)$

$$\begin{aligned}
&= ((k_n^\dagger \circ h_n)_1(\textstyle\sum_{x\in F} r_x\delta_x, X), (k_n^\dagger \circ h_n)_2(\textstyle\sum_{x\in F} r_x\delta_x, X))\\
&= (\textstyle\sum_{x\in F} r_x(\delta_x * k_n^\dagger \circ h_{n,1}(X(x)),\\
&\qquad (k_n^\dagger \circ h_n)_2(\textstyle\sum_{x\in F} r_x\delta_x, X))\\
&= (\textstyle\sum_{x\in F} r_x(\delta_x * (k_n^\dagger)_1(\textstyle\sum_{y\in G_x} s_y\delta_y, h_{n,2}(\textstyle\sum_{x\in F} r_x\delta_x, X)),\\
&\qquad (k_n^\dagger \circ h_n)_2(\textstyle\sum_{x\in F} r_x\delta_x, X))\\
&= (\textstyle\sum_{x\in F} r_x(\delta_x * (\textstyle\sum_{y\in G_x} s_y(\delta_y * k_{n,1}(h_{n,2}(\textstyle\sum_{x\in F} r_x\delta_x, X)(x\cdot y)))),\\
&\qquad (k_n^\dagger \circ h_n)_2(\textstyle\sum_{x\in F} r_x\delta_x, X)))
\end{aligned}$$

We conclude that

$$\begin{aligned}
&(k_n^\dagger \circ h_n^\dagger)_1(\textstyle\sum_{x\in F} r_x\delta_x, X)\\
&\quad = k_{n,1}(h_{n,1}(\textstyle\sum_{x\in F} r_x\delta_x, X), h_{n,2}(\textstyle\sum_{x\in F} r_x\delta_x, X))\\
&\quad = \textstyle\sum_{x\in F\,\&\,y\in G_x} r_x s_y(\delta_{x\cdot y} * k_{n,1}(h_{n,2}(\textstyle\sum_{x\in F} r_x\delta_x, X)(x\cdot y))))\\
&\quad = \textstyle\sum_{x\in F} r_x(\delta_x * (\textstyle\sum_{y\in G_x} s_y(\delta_y * k_{n,1}(h_{n,2}(\textstyle\sum_{x\in F} r_x\delta_x, X)(x\cdot y))))\\
&\quad = (k_n^\dagger \circ h_n)_1(\textstyle\sum_{x\in F} r_x\delta_x, X),
\end{aligned}$$

which shows the first components of $k_n^\dagger \circ h_n^\dagger$ and $(k_n^\dagger \circ h_n)^\dagger$ agree. A similar (laborious) argument proves the second components agree as well.

Theorem 6. *The functor CRV defines a monad on BCD.*

Proof. This follows from Lemma 4 and Corollary 2.

Remark 2. As noted earlier, if M is a compact monoid, convolution satisfies $(\mu*\nu)(A) = (\mu\times\nu)\{(x,y)\in M\times M \mid xy\in A\}$, so it is a mapping $*\colon \mathsf{Prob}(M)\times \mathsf{Prob}(M) \longrightarrow \mathsf{Prob}(M)$. Our use of $*$ in Theorem 6 is of a different character, since we are integrating along a measure μ to obtain a measure $\widehat{f}(\mu, X) = \int_x d\, f\circ X(x) * \mu(x)$, where $f\colon CRV(D) \longrightarrow \mathbb{V}\mathbb{M}\{0,1\}$.

3.4 *CRV* and Continuous Probability Measures

An accepted model for probabilistic choice is a probabilistic Turing machine, a Turing machine equipped with an second infinite tape containing a random sequence of bits. As a computation unfolds, the machine can consult the random tape from time to time and use the next random bit as a mechanism for making a choice. The source of randomness is not usually defined, and in a sense, it's immaterial. But it's common to assume that the same source is used throughout the computation – i.e., there's a single probability measure that's governing the sequence of random bits written on the tape.

In the models described in [9] and in [3], the idea of the random tape is captured by the Cantor tree $\mathbb{CT} = \bigcup_n \mathcal{C}_n \cup \mathcal{C}$, where the "single source of randomness" arises naturally as a measure μ on the Cantor set (at the top). That measure μ can be realized as $\mu = \sup_n \phi_n\,\mu$, where $\phi_n\colon \mathcal{C} \longrightarrow \mathcal{C}_n$ is the natural projection. As a concrete example, one can take μ to be Haar measure on $\mathcal{C}$ regarded as an infinite product of two-point groups, and then μ_n is the normalized uniform measure on $\mathcal{C}_n \simeq 2^n$. Then the possible sequence of outcomes of coin tosses on a particular computation are represented by a single path through the tree $\mathbb{CT}$, and the results at the n^{th}-level are governed by the distribution

$\phi_n\,\mu$. The outcome at that level is used to define choices in the semantic domain D via a random variable $X\colon \mathcal{C}_n \longrightarrow D$ for each n.

The same ideas permeate our model CRV, but our structure is different. The mappings $\phi_n\colon \mathbb{CT} \longrightarrow \mathcal{C}_n$ are replaced in our model by the mappings

$$\pi_n\colon \mathbb{M}\{0,1\} \longrightarrow \mathbb{M}_n \text{ given by } \pi_n(x) = \begin{cases} x & \text{if } x \in \mathbb{M}_n \\ x_1 \cdots x_n\bot & \text{if } |x \not\in \mathbb{M}_n \end{cases}$$

described in Definition 5. Then $\mathbb{M}_n$ is a retract of $\mathbb{M}\{0,1\}$ under π_n.

To realize *any* measure μ concentrated on $\mathcal{C} \subseteq \mathbb{M}\{0,1\}$, and the measures μ_n, we define new projections $\rho_n\colon \mathcal{C} \longrightarrow \mathcal{C}_n\checkmark$ from the Cantor set of maximal elements of $\mathbb{M}\{0,1\}$ to the n-bit words ending with $\checkmark$ in the obvious fashion. These mappings are continuous, but their images $\mu_n = \rho_n\,\mu$ are incomparable (since the set $\mathcal{C}_n\checkmark \subseteq \operatorname{Max}\mathbb{M}\{0,1\}$ for each n). Nevertheless, Proposition 5 implies the sequence $\rho_n\,\mu \longrightarrow \mu$ in $\mathbb{M}\{0,1\}$ in the Lawson topology. From a computational perspective, we can consider the related measures $\pi_m\,\mu = \nu_m$, then $\nu_m \leq \mu_n$ for each m and each $n \geq m$. But $\mu = \sup_m \nu_m$ since $\mathbf{1}_{\mathbb{M}\{0,1\}} = \sup_m \pi_m$, and then $\nu_m \leq \mu_n$ for $n \geq m$ implies $\mu_n \longrightarrow \mu$ in the Scott topology, since $\nu_m \ll \mu$ for each m.

4 Summary and Future Work

We have constructed a new monad for probabilistic choice using domain theory. The model consists of pairs (μ, X), where $\mu \in \mathbb{VM}\,\{0,1\}$ and $X\colon\ \operatorname{supp}_\Sigma \mu \longrightarrow D$ is a Scott-continuous random variables that defines the choices in the semantic domain D. The fact that CRV forms a monad relies crucially on the convolution operation on $\mathbb{VM}\{0,1\}$ that arises from the monoid operation on $\mathbb{M}\{0,1\}$, and the new order on $\mathbb{M}\{0,1\}$, rather than the prefix order on the set of finite and infinite words over $\{0,1\}$.

Our construction is focused on bounded complete domains, in order to utilize the inf-operation to define the Kleisli lift – in particular, in the random variable component of a pair (μ, X). One fault that was identified in the monad $\mathbb{V}$ is its lack of a distributive law over and of the power domains, which model nondeterministic choice. But here we see that we must assume the domain of values for our random variables already must support nondeterminism, since it arises naturally when one composes random variables (cf. Subsect. 3.2).

With all the theory, one might rightfully ask for some examples. An obvious target would be the models of CSP, starting with the seminal paper [5]. Morgan, McIver and their colleagues [18] have developed an extensive theory of CSP with probabilistic choice modeled by applying the sub-probability monad $\mathbb{V}$ to CSP models. It would be interesting to compare the model developed here, as applied, e.g., to the model in [18].

Acknowledgement. The author wishes to thank Tyler Barker for some very helpful discussions on the topic of monads of random variables.

References

1. Abramsky, S., Jung, A.: Domain Theory. In: Handbook of Logic in Computer Science, pp. 1–168. Clarendon Press, Oxford (1994)
2. Alvarez-Manilla, M., Edalat, A., Saheb-Djahromi, N.: An extension result for continuous valuations. J. Lond. Math. Soc. **61**(2), 629–640 (2000)
3. Barker, T.: A Monad for Randomized Algorithms. Tulane University Ph.D. dissertation (2016)
4. van Breugel, F., Mislove, M., Ouaknine, J., Worrell, J.: Domain theory, testing and simulations for labelled Markov processes. Theor. Comput. Sci. **333**, 171–197 (2005)
5. Brookes, S.D., Hoare, C.A.R., Roscoe, A.W.: A theory of communicating sequential processes. J. ACM **31**, 560–599 (1984)
6. Fedorchuk, V.: Covariant functors in the category of compacta, absolute retracts, and Q-manifolds. Russ. Math. Surv. **36**, 211–233 (1981)
7. Fedorchuk, V.: Probability measures in topology. Russ. Math. Surv. **46**, 45–93 (1991)
8. Gierz, G., Hofmann, K.H., Lawson, J.D., Mislove, M., Scott, D.: Continuous Lattices and Domains. Cambridge University Press, Cambridge (2003)
9. Goubault-Larrecq, J., Varacca, D.: Continuous random variables. In: LICS 2011, pp. 97–106. IEEE Press (2011)
10. Hofmann, K.H., Mislove, M.: Compact affine monoids, harmonic analysis and information theory. In: Mathematical Foundations of Information Flow, AMS Symposia on Applied Mathematics, vol. 71, pp. 125–182 (2012)
11. Jones, C.: Probabilistic nondeterminism, Ph.D. thesis. University of Edinburgh (1988)
12. Jung, A., Tix, R.: The troublesome probabilistic powerdomain. ENTCS **13**, 70–91 (1998)
13. Keimel, K.: The monad of probability measures over compact ordered spaces and its Eilenberg-Moore algebras, preprint (2008). http://www.mathematik.tu-darmstadt.de/~keimel/Papers/probmonadfinal1.pdf
14. Mislove, M.: Topology, domain theory and theoretical computer science. Topol. Appl. **89**, 3–59 (1998)
15. Mislove, M.: Discrete random variables over domains. Theor. Comput. Sci. **380**, 181–198 (2007). Special Issue on Automata, Languages and Programming
16. Mislove, M.: Anatomy of a domain of continuous random variables I. Theor. Comput. Sci. **546**, 176–187 (2014)
17. Mislove, M.: Anatomy of a domain of continuous random variables ll. In: Coecke, B., Ong, L., Panangaden, P. (eds.) Computation, Logic, Games, and Quantum Foundations. The Many Facets of Samson Abramsky. LNCS, vol. 7860, pp. 225–245. Springer, Heidelberg (2013). doi:10.1007/978-3-642-38164-5_16
18. Morgan, C.A., McIver, K., Seidel, J.W.: Saunders: a probabilistic process algebra including demonic nondeterminism. Formal Aspects Comput. **8**, 617–647 (1994)
19. Saheb-Djahromi, N.: CPOs of measures for nondeterminism. Theor. Comput. Sci. **12**, 19–37 (1980)
20. Swirszcz, T.: Monadic functors and convexity. Bulletin de l'Académie Polonaise des Sciences, Série des sciences math. astr. et phys. **22**, 39–42 (1974)
21. Varacca, D.: Two denotational models for probabilistic computation, Ph.D. thesis. Aarhus University (2003)
22. Varacca, D., Winskel, G.: Distributing probabililty over nondeterminism. Math. Struct. Comput. Sci. **16**(1), 87–113 (2006)