Order-Theoretic Semantics for ProbNetKAT*

1 Notation and Definitions

We use lower case letters $a, b, c \subseteq \mathsf{H}$ to denote history sets, uppercase letters $A, B, C \subseteq 2^{\mathsf{H}}$ to denote (measurable) sets (*i.e.*, sets of history sets), and calligraphic letters $\mathcal{B}, \mathcal{O}, \dots \subseteq 2^{2^{\mathsf{H}}}$ to denote sets of measurable sets. For a set X, we let $\wp_\omega(X) \triangleq \{Y \subseteq X \mid |Y| < \infty\}$ denote the finite subsets of X and $\mathbf{1}_X$ the characteristic function of X. For a statement φ, we let

$$[\varphi] \triangleq \begin{cases} 1 & \text{if } \varphi \text{ is true} \\ 0 & \text{otherwise} \end{cases}$$

For a function $f : X \to Y$, we let $f^{-1} : 2^Y \to 2^X$ denote the set function

$$f^{-1}(A) \triangleq \{x \in X \mid f(x) \in A\} \qquad\qquad f^{-1}(y) \triangleq f^{-1}(\{y\})$$

For $h \in \mathsf{H}$ and $b \in 2^{\mathsf{H}}$, define

$$B_h \triangleq \{c \mid h \in c\} \qquad\qquad B_b \triangleq \bigcap_{h \in b} B_h = \{c \mid b \subseteq c\}. \tag{1.1}$$

Then $B_h = B_{\{h\}}$. The sets B_h and $\sim B_h$ are the subbasic open sets of the Cantor space topology on 2^{H}. The family of Borel sets $\mathcal{B}$ is the smallest σ-algebra containing the Cantor-open sets. Let $\mathcal{B}_b$ be the Boolean subalgebra of $\mathcal{B}$ generated by $\{B_h \mid h \in b\}$.

Lemma 1.

(i) $b \subseteq c \Leftrightarrow B_c \subseteq B_b$

(ii) $B_b \cap B_c = B_{b \cup c}$

(iii) $B_\varnothing = 2^{\mathsf{H}}$

---

*Technical notes for the publication "Cantor Meets Scott: Foundations for Probabilistic Networks." Steffen Smolka, Praveen Kumar, Nate Foster, Dexter Kozen, and Alexandra Silva. In *POPL* 2017, Paris.

(iv) $\mathcal{B}_{2^{\mathsf{H}}} = \bigcup_{b \in \wp_\omega(H))} \mathcal{B}_b$.

If b is finite, so is $\mathcal{B}_b$. The atoms of $\mathcal{B}_b$ are in one-to-one correspondence with the subsets $a \subseteq b$, the subset a determining which B_h occur positively in the construction of the atom:

$$A_{ab} \triangleq \bigcap_{h \in a} B_h \cap \bigcap_{h \in b-a} {\sim}B_h = B_a - \bigcup_{a \subset c \subseteq b} B_c = \{c \in 2^{\mathsf{H}} \mid c \cap b = a\}, \tag{1.2}$$

where $\subset$ denotes proper subset. The sets A_{ab} are the basic open sets of the Cantor space. The notation A_{ab} is reserved for such sets.

Lemma 2. *For b finite and $a \subseteq b$, $B_a = \bigcup_{a \subseteq c \subseteq b} A_{cb}$.*

Proof. By (1.2),

$$\bigcup_{a \subseteq c \subseteq b} A_{cb} = \bigcup_{a \subseteq c \subseteq b} \{d \in 2^{\mathsf{H}} \mid d \cap b = c\} = \{d \in 2^{\mathsf{H}} \mid a \subseteq d\} = B_a.$$

□

Lemma 3.

(i) *$A_{ab} \subseteq A_{a'b'}$ iff $a' \subseteq a$ and $b' - a' \subseteq b - a$.*

(ii) *Either $A_{ab} \subseteq A_{a'b'}$, $A_{a'b'} \subseteq A_{ab}$, or $A_{ab} \cap A_{a'b'} = \varnothing$.*

Proof. (i) By (1.2), the assumption $A_{ab} \subseteq A_{a'b'}$ reduces to

$$\forall d \;\; d \cap b = a \;\Rightarrow\; d \cap b' = a' \tag{1.3}$$

Taking $d = a$ in (1.3), we have

$$a \cap b = a \Rightarrow a \cap b' = a' \Rightarrow a' \subseteq a.$$

Taking $d = {\sim}(b - a)$ in (1.3), we have

$$\begin{aligned} {\sim}(b-a) \cap b = a &\Rightarrow {\sim}(b-a) \cap b' = a' \\ &\Rightarrow (b-a) \cap b' = b' - a' \\ &\Rightarrow b' - a' \subseteq b - a. \end{aligned}$$

Conversely, if $a' \subseteq a$ and $b' - a' \subseteq b - a$, then

$$\begin{aligned} b' = (b' - a') \cup a' \subseteq (b-a) \cup a = b \\ b' \cap (a - a') = (b' - a') \cap a \subseteq (b-a) \cap a = \varnothing. \end{aligned}$$

Thus if $d \cap b = a$, then

$$d \cap b' = d \cap b \cap b' = a \cap b' = (b' \cap (a - a')) \cup (b' \cap a') = a',$$

so (1.3) holds.

(ii) Either $A_{ab} \subseteq A_{a'b'}$, $A_{a'b'} \subseteq A_{ab}$, or $A_{ab} \cap A_{a'b'} = \varnothing$.

□

Lemma 4 (see [3, Theorem III.13.A]). *Any probability measure is uniquely determined by its values on B_b for b finite.*

Proof. For b finite, the atoms of $\mathcal{B}_b$ are of the form (1.2). By the inclusion-exclusion principle,

$$\mu(A_{ab}) = \mu(B_a - \bigcup_{a \subset c \subseteq b} B_c) = \sum_{a \subseteq c \subseteq b} (-1)^{|c-a|} \mu(B_c). \tag{1.4}$$

Thus μ is uniquely determined on the atoms of $\mathcal{B}_b$, therefore on all of $\mathcal{B}_b$. As $\mathcal{B}_{2^{\mathsf{H}}}$ is the union of the $\mathcal{B}_b$ for b finite, μ is uniquely determined on $\mathcal{B}_{2^{\mathsf{H}}}$. By the monotone class theorem, the Borel sets $\mathcal{B}$ are the smallest monotone class containing $\mathcal{B}_{2^{\mathsf{H}}}$, and since $\mu(\bigcup_n A_n) = \sup_n \mu(A_n)$ and $\mu(\bigcap_n A_n) = \inf_n \mu(A_n)$, μ is determined on all Borel sets. □

Theorem 1. *A function $\mu : \{B_b \mid b \text{ finite}\} \to [0,1]$ extends to a measure $\mu : \mathcal{B} \to [0,1]$ iff for all finite b and all $a \subseteq b$*

$$\sum_{a \subseteq c \subseteq b} (-1)^{|c-a|} \mu(B_c) \geq 0$$

Moreover, the extension to $\mathcal{B}$ is unique.

Proof. The condition is clearly necessary by (1.4).

For sufficiency and uniqueness, we use the Carathéodory extension theorem. For each atom A_{ab} of $\mathcal{B}_b$, $\mu(A_{ab})$ is already determined uniquely by (1.4) and nonnegative by assumption. For each $B \in \mathcal{B}_b$, write B uniquely as a union of atoms and define $\mu(B)$ to be the sum of the $\mu(A_{ab})$ for all atoms A_{ab} of $\mathcal{B}_b$ contained in B.

We must show that $\mu(B)$ is well-defined. Note that the definition is given in terms of b, and we must show that the definition is independent of the choice of b. It suffices to show that the calculation using atoms of $b' = b \cup \{h\}$, $h \notin b$, gives the same result. Each atom of $\mathcal{B}_b$ is the disjoint union of two atoms of $\mathcal{B}_{b'}$:

$$A_{ab} = A_{a \cup \{h\}, b \cup \{h\}} \cup A_{a, b \cup \{h\}}$$

and it suffices to show that the sum of their measures is the measure of A_{ab}:

$$\begin{aligned} \mu(A_{a,b\cup\{h\}}) &= \sum_{a \subseteq c \subseteq b \cup \{h\}} (-1)^{|c-a|} \mu(B_c) \\ &= \sum_{a \subseteq c \subseteq b} (-1)^{|c-a|} \mu(B_c) + \sum_{a \cup \{h\} \subseteq c \subseteq b \cup \{h\}} (-1)^{|c-a|} \mu(B_c) \\ &= \mu(A_{ab}) - \mu(A_{a \cup \{h\}, b \cup \{h\}}). \end{aligned}$$

To apply the Carathéodory extension theorem, we must show that μ is countably additive, *i.e.* that $\mu(\bigcup_n A_n) = \sum_n \mu(A_n)$ for any countable sequence $A_n \in \mathcal{B}_{2^\mathsf{H}}$ of pairwise disjoint sets whose union is in $\mathcal{B}_{2^\mathsf{H}}$.

For a finite sequences $A_n \in \mathcal{B}_{2^\mathsf{H}}$, write each A_n uniquely as a disjoint union of atoms of $\mathcal{B}_b$ for some sufficiently large b such that all $A_n \in \mathcal{B}_b$. Then $\bigcup_n A_n \in \mathcal{B}_b$, the value of the atoms are given by (1.4), and the value of $\mu(\bigcup_n A_n)$ is well-defined and equal to $\sum_n \mu(A_n)$.

We cannot have an infinite set of pairwise disjoint nonempty $A_n \in \mathcal{B}_{2^\mathsf{H}}$ whose union is in $\mathcal{B}_{2^\mathsf{H}}$ by compactness. All elements of $\mathcal{B}_{2^\mathsf{H}}$ are clopen in the Cantor topology. If $\bigcup_n A_n = A \in \mathcal{B}_{2^\mathsf{H}}$, then $\{A_n \mid n \geq 0\}$ would be an open cover of A with no finite subcover. $\square$

2 DCPOs and the Scott Topology

We assume familiarity with basic domain theory; see [1] for an introduction.

A *partial order* $(D, \sqsubseteq)$ is a set D together with a reflexive, transitive, and antisymmetric relation $\sqsubseteq$ on D. For two elements $x, y \in D$ we let $x \sqcup y$ denote their $\sqsubseteq$-least upper bound (*i.e.*, their supremum), provided it exists. Analogously, the least upper bound of a subset $C \subseteq D$ is denoted $\bigsqcup C$, provided it exists. A subset $C \subseteq D$ is *directed* if for any two $x, y \in C$ there exists *some* upper bound $x, y \sqsubseteq z \in C$. A partial order (PO) is called a *directed complete partial order* (DCPO) if every directed subset $C \subseteq D$ has a supremum $\bigsqcup C$ in D. If a PO has a least element is is denoted by $\bot$; if it has a greatest element it is denoted by $\top$.

The set 2^H forms a DCPO under the subset order $\subseteq$ with suprema $\bigsqcup C = \bigcup C$, least element $\bot = \varnothing$, and greatest element $\top = \mathsf{H}$. The nonnegative real numbers with infinity $\mathbb{R}_+ = \{r \in \mathbb{R} \mid r \geq 0\} \cup \{\infty\}$ form a DCPO under the natural order $\leq$ with suprema $\bigsqcup C = \sup C$, least element $\bot = 0$, and greatest element $\top = \infty$. The unit interval is a DCPO under the same order, but with $\top = 1$.

Let D be a DCPO. A subset $A \subseteq D$ is called *up-closed* (or an *upper set*) if $b \in A$ whenever $a \in A$ and $a \sqsubseteq b$. The smallest up-closed set containing $A \subseteq D$ is called its *up-closure* and is denoted $A\!\uparrow$.

A subset $A \subseteq D$ is (*Scott-*)*open* if it is up-closed and intersects every directed subset $B \subseteq D$ such that $\bigsqcup B \in A$. The Scott-open sets form a topology on D called the *Scott topology*.

A function $f : D \to E$ between DCPOs is (*Scott-*)*continuous* if it is continuous with respect to the Scott topologies on D and E.

The Scott-open sets of $\mathbb{R}_+$ are the upper semi-infinite intervals $(r, \infty]$, $r \in \mathbb{R}_+$. The Scott-open sets of 2^H are denoted $\mathcal{O}$ and are characterized in Lemma 6 below.

An element a of a DCPO is called *finite* ([1] uses the term *compact*) if for any directed set A, if $a \sqsubseteq \bigsqcup A$, then there exists $b \in A$ such that $a \sqsubseteq b$. Equivalently,

a is finite if its up-closure $\{a\}\uparrow$ is Scott-open.

A DCPO is called *algebraic* if for every element b, the finite elements $\sqsubseteq$-below b form a directed set and b is the supremum of this set. The DCPO 2^{H} under the subset order is algebraic, the finite elements of 2^{H} being the finite subsets of H, and $\{a\}\uparrow = B_a$.

A set in a topological space is *compact-open* if it is compact (every open cover has a finite subcover) and open.

Here we recall some general facts. These are all well-known, but we state them as a lemma for future reference.

Lemma 5.

(i) *The cartesian product of any collection of DCPOs is a DCPO under the componentwise order.*

(ii) *A function $f : D \to E$ between DCPOs is Scott-continuous iff the following two properties hold:*

- *f is monotone, that is, if $a \sqsubseteq b$, then $f(a) \sqsubseteq f(b)$;*
- *f preserves suprema of directed sets in the sense that $f(\bigsqcup D) = \bigsqcup_{a\in D} f(a)$*

(see [1, Proposition 2.3.4]).

(iii) *If E is a DCPO and D is any set, then the family of functions $f : D \to E$ is a DCPO under the pointwise order $f \sqsubseteq g$ iff for all $a \in D$, $f(a) \sqsubseteq g(a)$. The supremum of a directed set $\mathcal{D}$ of functions $D \to E$ is the function $(\bigsqcup \mathcal{D})(a) = \bigsqcup_{f\in\mathcal{D}} f(a)$.*

(iv) *Let E be a DCPO and D_1, D_2 sets. There is a homeomorphism (bicontinuous bijection)*

$$\mathsf{curry} : (D_1 \times D_2 \to E) \to (D_1 \to D_2 \to E)$$

between the DCPOs $D_1 \times D_2 \to E$ and $D_1 \to D_2 \to E$, where all function spaces in question are ordered pointwise. The inverse of curry *is denoted* uncurry*.*

(v) *If D, E are DCPOs, then the supremum of a directed set of continuous functions $D \to E$ is continuous. Thus the family of continuous functions $D \to E$ form a DCPO. This space is denoted $[D \to E]$.*

(vi) *In an algebraic DCPO, the open sets $\{a\}\uparrow$ for finite a form a base for the Scott topology.*

(vii) *An element of an algebraic DCPO is compact-open iff it is a finite union of basic open sets $\{a\}\uparrow$.*

Let $\mathcal{O}$ denote the family of Scott-open sets of 2^{H}.

Lemma 6. *A subset $B \subseteq 2^{\mathsf{H}}$ is Scott-open iff there exists $F \subseteq \wp_\omega(H)$ such that $B = \bigcup_{a \in F} B_a$.*

Proof. This follows from the observation that $(2^{\mathsf{H}}, \subseteq)$ is an algebraic DCPO whose finite elements are the finite sets $\wp_\omega(H)$, and that their up-closures $\{a\}\uparrow$, $a \in \wp_\omega(H)$, form a base for the Scott topology (Lemma 5(vi)). ☐

Here are some facts about the Scott topology on 2^{H}.

- The Scott topology is weaker than the Cantor space topology. For example, $\sim B_h$ is Cantor-open but not Scott-open. However, the Borel sets of the two topologies are the same, as $\sim B_h$ is a Π^0_1 Borel set.[1]
- The open sets $\mathcal{O}$ ordered by the subset relation forms an ω-complete lattice with bottom element $B_H = \{H\}$ and top element $B_\varnothing = 2^{\mathsf{H}}$.
- The sets B_a for $a \in \wp_\omega(H)$ are a base for the Scott topology. The sets B_h for $h \in \mathsf{H}$ are a subbase.
- The finite sets $a \in \wp_\omega(H)$ are dense and countable, thus the space is separable.
- Every open set is a countable union of basic open sets B_a, $a \in \wp_\omega(H)$.
- Unlike the Cantor topology, the space is not Hausdorff, metrizable, or compact. It is not Hausdorff, as any nonempty open set contains H. However, it is satisfies the weaker T_0 separation property: For any pair of points a, b with $a \not\subseteq b$, $a \in B_a$ but $b \notin B_a$.
- There is an up-closed Π^0_2 Borel set with an uncountable set of minimal elements (Lemma 33 below).
- There are up-closed Borel sets with no minimal elements; for example, the family of cofinite subsets of H. This is a Σ^0_3 Borel set.
- The compact-open sets are those of the form $F\uparrow$, where F is a finite set of finite sets. There are plenty of open sets that are not compact-open, e.g. $B_\varnothing - \{\varnothing\} = \bigcup_{h \in \mathsf{H}} B_h$.

Lemma 7. *Let $b \in 2^{\mathsf{H}}$ be finite and let $F \subseteq 2^b$. Then $\bigcup_{a \in F} A_{ab} \in \mathcal{O}$ iff $F = F\uparrow \cap 2^b$.*

[1] References to the Borel hierarchy Σ^0_n and Π^0_n refer to the Scott topology. Although the Scott and Cantor topologies generate the same Borel sets, their Borel hierarchies are different.

Proof. Suppose $F = F{\uparrow} \cap 2^b$. By Lemma 2,

$$\bigcup_{a\in F} A_{ab} = \bigcup_{a\in F} \bigcup_{a\subseteq c\subseteq b} A_{cb} = \bigcup_{a\in F} B_a \in \mathcal{O}.$$

Conversely, suppose $\bigcup_{a\in F} A_{ab} \in \mathcal{O}$. Then for any $a \in F$ and $a \subseteq c \subseteq b$,

$$A_{cb} \subseteq B_c \subseteq B_a \subseteq \bigcup_{a\in F} A_{ab}.$$

The first inclusion is by Lemma 2; the second is by the fact that $a \subseteq c$; and the third is because $a \in A_{ab}$, $\bigcup_{a\in F} A_{ab}$ is in $\mathcal{O}$ and therefore up-closed in 2^{H}, and $B_a = \{a\}{\uparrow}$ is the smallest up-closed set containing a. Since the atoms of $\mathcal{B}_b$ are pairwise disjoint and A_{cb} is nonempty (it contains c), it follows that $c \in F$. ☐

3 Cantor Meets Scott

In this section we establish a strong correspondence between the Cantor and Scott topologies on 2^{H}.

Consider the infinite triangular matrix E and its inverse E^{-1} with rows and columns indexed by the finite subsets of H, where

$$E_{ac} = [a \subseteq c] \qquad\qquad E^{-1}_{ac} = (-1)^{|c-a|}[a \subseteq c].$$

These matrices are indeed inverses: For $a, d \in \wp_\omega(H)$,

$$\begin{aligned}(E \cdot E^{-1})_{ad} &= \sum_c E_{ac} \cdot E^{-1}_{cd} = \sum_c [a \subseteq c] \cdot [c \subseteq d] \cdot (-1)^{|d-c|} \\ &= \sum_{a\subseteq c\subseteq d} (-1)^{|d-c|} = [a = d],\end{aligned}$$

thus $E \cdot E^{-1} = I$, and similarly $E^{-1} \cdot E = I$.

Recall that the Cantor basic open sets are the elements A_{ab} for b finite and $a \subseteq b$. Those for fixed finite b are the atoms of the Boolean algebra $\mathcal{B}_b$. They form the basis of a $2^{|b|}$-dimensional linear space. The Scott basic open sets B_a for $a \subseteq b$ are another basis for the same space. The two bases are related by the matrix $E[b]$, the $2^b \times 2^b$ submatrix of E with rows and columns indexed by subsets of b. One can show that the finite matrix $E[b]$ is invertible with inverse $E[b]^{-1} = (E^{-1})[b]$.

Lemma 8. *Let μ be a measure on 2^{H} and $b \in \wp_\omega(H)$. Let X, Y be vectors indexed by subsets of b such that $X_a = \mu(B_a)$ and $Y_a = \mu(A_{ab})$ for $a \subseteq b$. Let $E[b]$ be the $2^b \times 2^b$ submatrix of E. Then $X = E[b] \cdot Y$.*

Proof. For any $a \subseteq b$,

$$\begin{aligned} X_a = \mu(B_a) &= \sum_{a \subseteq c \subseteq b} \mu(A_{cb}) \\ &= \sum_c [a \subseteq c] \cdot [c \subseteq b] \cdot \mu(A_{cb}) \\ &= \sum_c E[b]_{ac} \cdot Y_c = (E[b] \cdot Y)_a. \end{aligned}$$

□

The matrix-vector equation $X = E[b] \cdot Y$ captures the fact that for $a \subseteq b$, B_a is the disjoint union of the atoms A_{cb} of $\mathcal{B}_b$ for $a \subseteq c \subseteq b$, and consequently $\mu(B_a)$ is the sum of $\mu(A_{cb})$ for these atoms. The inverse equation $X = E[b]^{-1} \cdot Y$ captures the inclusion-exclusion principle for $\mathcal{B}_b$.

In fact, more can be said about the structure of E. For any $b \in 2^{\mathsf{H}}$, finite or infinite, let $E[b]$ be the submatrix of E with rows and columns indexed by the subsets of b. If $a \cap b = \varnothing$, then $E[a \cup b] = E[a] \otimes E[b]$, where $\otimes$ denotes Kronecker product. The formation of the Kronecker product requires a notion of pairing on indices, which in our case is given by disjoint set union. For example,

$$E[\{h_1\}] = \begin{array}{c} \\ \varnothing \\ \{h_1\} \end{array} \begin{array}{c} \begin{array}{cc} \varnothing & \{h_1\} \end{array} \\ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \end{array} \qquad E[\{h_2\}] = \begin{array}{c} \\ \varnothing \\ \{h_2\} \end{array} \begin{array}{c} \begin{array}{cc} \varnothing & \{h_2\} \end{array} \\ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \end{array}$$

$$E[\{h_1, h_2\}] = \begin{array}{c} \\ \varnothing \\ \{h_1\} \\ \{h_2\} \\ \{h_1,h_2\} \end{array} \begin{array}{c} \begin{array}{cccc} \varnothing & \{h_1\} & \{h_2\} & \{h_1,h_2\} \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} = E[\{h_1\}] \otimes E[\{h_2\}].$$

As $(E \otimes F)^{-1} = E^{-1} \otimes F^{-1}$ for Kronecker products of invertible matrices, we also have

$$E[\{h_1\}]^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \qquad E[\{h_2\}]^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$E[\{h_1, h_2\}]^{-1} = E[\{h_1\}]^{-1} \otimes E[\{h_2\}]^{-1} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The matrix E can be viewed as the infinite Kronecker product $\bigotimes_{h \in \mathsf{H}} E[\{h\}]$.

Theorem 2. *The probability measures on* $(2^{\mathsf{H}}, \mathcal{B})$ *are in one-to-one correspondence with pairs of matrices* $M, N \in \mathbb{R}^{\wp_\omega(H) \times \wp_\omega(H)}$ *such that*

(i) *M is diagonal with entries in* $[0, 1]$,

(ii) *N is nonnegative, and*

(iii) $N = E^{-1}ME$.

The correspondence associates the probability measure μ with the matrices

$$N_{ab} = \mu(A_{ab}) \qquad\qquad M_{ab} = [a = b] \cdot \mu(B_a). \tag{3.1}$$

Proof. Given a probability measure μ, certainly (i) and (ii) hold of the matrices M and N formed from μ by the rule (3.1). For (iii), we calculate:

$$\begin{aligned}(E^{-1}ME)_{ab} &= \sum_{c,d} E^{-1}_{ac} M_{cd} E_{db} \\ &= \sum_{c,d} [a \subseteq c] \cdot (-1)^{|c-a|} \cdot [c = d] \cdot \mu(M_{cd}) \cdot [d \subseteq b] \\ &= \sum_{a \subseteq c \subseteq b} (-1)^{|c-a|} \cdot \mu(B_c) = \mu(A_{ab}) = N_{ab}.\end{aligned}$$

That the correspondence is one-to-one is immediate from Theorem 1. □

4 A DCPO on Markov Kernels

For measures μ, ν on 2^{H}, define $\mu \sqsubseteq \nu$ if $\mu(B) \le \nu(B)$ for all $B \in \mathcal{O}$. Equivalently, $\mu \sqsubseteq \nu$ iff the signed measure $\nu - \mu$ takes only nonnegative values on $\mathcal{O}$ (it may take negative values on Borel sets not in $\mathcal{O}$). This order was first defined by Saheb-Djahromi [7].

Theorem 3. *The probability measures on* $(2^{\mathsf{H}}, \mathcal{B})$ *ordered by* $\sqsubseteq$ *form a DCPO with bottom element* $\delta_\varnothing$ *and top element* δ_H.

S-D proof is more general, but very long (5 journal pages). Ours is considerably shorter due to the extension theorem. –DCK

The relation $\sqsubseteq$ is a partial order. Reflexivity and transitivity are clear, and antisymmetry follows from Lemma 4.

Proof. To show that suprema of directed sets exist, let $\mathcal{D}$ be a directed set of measures, and define

$$(\bigsqcup \mathcal{D})(B) \triangleq \sup_{\mu \in \mathcal{D}} \mu(B), \;\; B \in \mathcal{O}.$$

This is clearly the supremum of $\mathcal{D}$, provided it defines a valid measure.[2] To show this, choose a countable chain $\mu_0 \sqsubseteq \mu_1 \sqsubseteq \cdots$ in $\mathcal{D}$ such that $\mu_m \sqsubseteq \mu_n$ for all $m < n$ and $(\bigsqcup \mathcal{D})(B_c) - \mu_n(B_c) \le 1/n$ for all c such that $|c| \le n$. Then for all finite $c \in 2^{\mathsf{H}}$, $(\bigsqcup \mathcal{D})(B_c) = \sup_n \mu_n(B_c)$.

Then $\bigsqcup \mathcal{D}$ is a measure by Theorem 1 because for all finite b and $a \subseteq b$,

$$\begin{aligned}\sum_{a\subseteq c\subseteq b} (-1)^{|c-a|}(\bigsqcup \mathcal{D})(B_c) &= \sum_{a\subseteq c\subseteq b} (-1)^{|c-a|} \sup_n \mu_n(B_c) \\ &= \lim_n \sum_{a\subseteq c\subseteq b} (-1)^{|c-a|}\mu_n(B_c) \\ &\ge 0.\end{aligned}$$

To show that $\delta_\varnothing$ is $\sqsubseteq$-minimum, observe that for all $B \in \mathcal{O}$,

$$\delta_\varnothing(B) = [\varnothing \in B] = [B = B_\varnothing = 2^{\mathsf{H}}]$$

as $B_\varnothing = 2^{\mathsf{H}}$ is the only up-closed set containing $\varnothing$. Thus for all measures μ, $\delta_\varnothing(2^{\mathsf{H}}) = 1 = \mu(2^{\mathsf{H}})$, and for all $B \in \mathcal{O}$, $B \ne 2^{\mathsf{H}}$, $\delta_\varnothing(B) = 0 \le \mu(B)$.

Finally, to show that δ_H is $\sqsubseteq$-maximum, observe that every nonempty $B \in \mathcal{O}$ contains H because it is up-closed. Therefore, δ_H is the constant function 1 on $\mathcal{O} - \{\varnothing\}$, making it $\sqsubseteq$-maximum. □

Lemma 9. *$\mu \sqsubseteq \mu \mathbin{\&} \nu$ and $\nu \sqsubseteq \mu \mathbin{\&} \nu$.*

Proof. For any up-closed measurable set B,

$$\begin{aligned}\mu(B) &= \mu(B)\cdot\nu(2^{\mathsf{H}}) = (\mu\times\nu)(B\times 2^{\mathsf{H}}) \\ &= (\mu\times\nu)(\{(b,c) \mid b\in B\}) \\ &\le (\mu\times\nu)(\{(b,c) \mid b\cup c\in B\}) = (\mu \mathbin{\&} \nu)(B).\end{aligned}$$

and similarly for ν. □

Surprisingly, despite Lemma 9, the probability measures do not form an upper semilattice under $\sqsubseteq$, although counterexamples are difficult to construct. A counterexample is given in Appendix C.

Now we lift the order $\sqsubseteq$ to Markov kernels $P : 2^{\mathsf{H}} \times \mathcal{B} \to [0,1]$. The order is defined pointwise on kernels regarded as functions $2^{\mathsf{H}} \times \mathcal{O} \to [0,1]$; that is,

$$P \sqsubseteq Q \overset{\triangle}{\Leftrightarrow} \forall a \in 2^{\mathsf{H}}\ \forall B \in \mathcal{O}\ P(a,B) \le Q(a,B).$$

There are several equivalent ways of viewing the lifted order $\sqsubseteq$:

[2] This is actually quite subtle. One might be tempted to define

$$(\bigsqcup \mathcal{D})(B) \triangleq \sup_{\mu\in\mathcal{D}} \mu(B),\ \ B\in\mathcal{B}$$

However, this definition would not give a valid probability measure in general. In particular, an increasing chain of measures does not generally converge to its supremum pointwise. However, it *does* converge pointwise on $\mathcal{O}$.

Lemma 10. *The following are equivalent:*

(i) $P \sqsubseteq Q$*, that is, for all* $a \in 2^{\mathsf{H}}$ *and* $B \in \mathcal{O}$*,* $P(a, B) \le Q(a, B)$*;*

(ii) *for all* $a \in 2^{\mathsf{H}}$*,* $P(a, -) \sqsubseteq Q(a, -)$ *in the DCPO of measures* $\mathcal{M}(2^{\mathsf{H}})$*;*

(iii) *for all* $B \in \mathcal{O}$*,* $P(-, B) \sqsubseteq Q(-, B)$ *in the DCPO of functions* $2^{\mathsf{H}} \to [0, 1]$*;*

(iv) $\mathsf{curry}\, P \sqsubseteq \mathsf{curry}\, Q$ *in the DCPO of functions* $2^{\mathsf{H}} \to \mathcal{M}(2^{\mathsf{H}})$*.*

Proof. To show that (i), (ii), and (iv) are equivalent,

$$\begin{aligned}
&\forall a \in 2^{\mathsf{H}}\ \forall B \in \mathcal{O}\ P(a, B) \le Q(a, B) \\
&\quad \Leftrightarrow \forall a \in 2^{\mathsf{H}}\ (\forall B \in \mathcal{O}\ P(a, B) \le Q(a, B)) \\
&\quad \Leftrightarrow \forall a \in 2^{\mathsf{H}}\ P(a, -) \sqsubseteq Q(a, -) \\
&\quad \Leftrightarrow \forall a \in 2^{\mathsf{H}}\ (\mathsf{curry}\, P)(a) \sqsubseteq (\mathsf{curry}\, Q)(a) \\
&\quad \Leftrightarrow \mathsf{curry}\, P \sqsubseteq \mathsf{curry}\, Q.
\end{aligned}$$

To show that (i) and (iii) are equivalent,

$$\begin{aligned}
&\forall a \in 2^{\mathsf{H}}\ \forall B \in \mathcal{O}\ P(a, B) \le Q(a, B) \\
&\quad \Leftrightarrow \forall B \in \mathcal{O}\ (\forall a \in 2^{\mathsf{H}}\ P(a, B) \le Q(a, B)) \\
&\quad \Leftrightarrow \forall B \in \mathcal{O}\ P(-, B) \sqsubseteq Q(-, B).
\end{aligned}$$

□

ProbNetKAT programs will be interpreted as *continuous* kernels. A Markov kernel $P : 2^{\mathsf{H}} \times \mathcal{B} \to [0, 1]$ *continuous* if it is Scott-continuous in its first argument; that is, for any fixed $A \in \mathcal{O}$, $P(a, A) \le P(b, A)$ whenever $a \subseteq b$, and for any directed set $D \subseteq 2^{\mathsf{H}}$, $P(\bigcup D, A) = \sup_{a \in D} P(a, A)$. This is the same as saying that its curried version $\mathsf{curry}\, P : 2^{\mathsf{H}} \to \mathcal{M}(2^{\mathsf{H}})$ is Scott-continuous as a function from the DCPO 2^{H} ordered by $\subseteq$ to the DCPO of probability measures ordered by $\sqsubseteq$.

We will show in §4.3 that all ProbNetKAT programs give rise to continuous kernels.

Theorem 4. *The continuous kernels* $P : 2^{\mathsf{H}} \times \mathcal{B} \to [0, 1]$ *ordered by* $\sqsubseteq$ *form a continuous DCPO with basis consisting of kernels of the form* $b\, ;\, P\, ;\, d$ *for* P *an arbitrary continuous kernel and* b, d *filters on finite sets* b *and* d*; that is, kernels that drop all input packets except for those in* b *and all output packets except those in* d*.*

Proof. We must show that the supremum of any directed set of continuous Markov kernels is a continuous Markov kernel. In general, the supremum of a directed set of continuous functions between DCPOs is continuous, as noted in Lemma 5(v). Given a directed set $\mathcal{D}$ of continuous kernels, we apply this to the directed set $\{\mathsf{curry}\,P : 2^{\mathsf{H}} \to \mathcal{M}(2^{\mathsf{H}}) \mid P \in \mathcal{D}\}$ to derive that $\bigsqcup_{P\in\mathcal{D}} \mathsf{curry}\,P$ is continuous, then use the fact that curry is continuous to infer that $\bigsqcup_{P\in\mathcal{D}} \mathsf{curry}\,P = \mathsf{curry} \bigsqcup\mathcal{D}$, therefore $\mathsf{curry} \bigsqcup\mathcal{D}$ is continuous. This says that the function $P : 2^{\mathsf{H}} \times \mathcal{B} \to [0,1]$ is continuous in its first argument.

We must still argue that the supremum $\bigsqcup\mathcal{D}$ is a Markov kernel, that is, a measurable function in its first argument and a probability measure in its second argument. The first statement follows from the fact that any continuous function is measurable with respect to the Borel sets generated by the topologies of the two spaces. For the second statement, we appeal to Theorem 3 and the continuity of curry:

$$(\mathsf{curry} \textstyle\bigsqcup\mathcal{D})(a) = (\bigsqcup_{P\in\mathcal{D}} \mathsf{curry}\,P)(a) = \bigsqcup_{P\in\mathcal{D}} (\mathsf{curry}\,P)(a),$$

which is a supremum of a directed set of probability measures, therefore by Theorem 3 is itself a probability measure.

To show that it is a continuous DCPO with basis of the indicated form, we note that for any $a \in 2^{\mathsf{H}}$ and $B \in \mathcal{O}$,

$$(b \,;\, P \,;\, d)(a, B) = P(a \cap b, \{c \mid c \cap d \in B\}). \tag{4.1}$$

Every element of the space is the supremum of a directed set of such elements. Given a continuous kernel P, consider the directed set $\mathcal{D}$ of all elements $b \,;\, P \,;\, d$ for b, d finite. Then for any $a \in 2^{\mathsf{H}}$ and $B \in \mathcal{O}$,

$$\begin{aligned}(\textstyle\bigsqcup\mathcal{D})(a, B) &= \sup_{b,d\in\wp_\omega(H)} P(a \cap b, \{c \mid c \cap d \in B\}) && (4.2)\\ &= \sup_{d\in\wp_\omega(H)} P(a, \{c \mid c \cap d \in B\}) && (4.3)\\ &= P(a, B), && (4.4)\end{aligned}$$

the inference (4.2) from (4.1), the inference (4.3) from the fact that P is continuous in its first argument, and the inference (4.3) from the fact that the sets $\{c \mid c \cap d \in B\}$ for $d \in \wp_\omega(H)$ form a directed set of Scott-open sets whose union is B and that P is a measure in its second argument. □

It is not true that the space of continuous kernels is algebraic with finite elements $b \,;\, P \,;\, d$. A counterexample is given in Appendix E.

4.1 Products and Integration

As pointed out by Jones [4, §3.6], the product σ-algebra of the Borel sets of two topological spaces X, Y is in general not the same as the Borel sets of the topological product $X \times Y$, although this property does hold for the Cantor space, as its basic open sets are clopen. More importantly, as also observed in [4, §3.6], the Scott topology on the product of DCPOs with the componentwise order is not necessarily the same as the product topology. However, in our case, the two topologies coincide.

Theorem 5. *Let D_α, $\alpha < \kappa$, be a collection of algebraic DCPOs with F_α the finite elements of D_α. Then the product $\prod_{\alpha<\kappa} D_\alpha$ with the componentwise order is an algebraic DCPO with finite elements*

$$F = \{c \in \textstyle\prod_\alpha F_\alpha \mid \pi_\alpha(c) = \bot \text{ for all but finitely many } \alpha\}.$$

Proof. The projections $\pi_\beta : \prod_\alpha D_\alpha \to D_\beta$ are easily shown to be continuous with respect to the componentwise order. For any $d \in \prod_{\alpha<\kappa} D_\alpha$, the set $\{d\}\downarrow \cap F$ is directed, and $d = \bigsqcup(\{d\}\downarrow \cap F)$: for any α, the set $\pi_\alpha(\{d\}\downarrow \cap F) = \{\pi_\alpha(d)\}\downarrow \cap F_\alpha$ is directed, thus

$$\pi_\alpha(d) = \bigsqcup(\{\pi_\alpha(d)\}\downarrow \cap F_\alpha) = \bigsqcup(\pi_\alpha(\{d\}\downarrow \cap F)) = \pi_\alpha(\bigsqcup(\{d\}\downarrow \cap F)),$$

and as α was arbitrary, $d = \bigsqcup(\{d\}\downarrow \cap F)$.

It remains to show that $\{c\}\uparrow = \prod_{\alpha<\kappa}\{\pi_\alpha(c)\}\uparrow$ is open for $c \in F$. Let A be a directed set with $\bigsqcup A \in \{c\}\uparrow$. For each α, $\{\pi_\alpha(a) \mid a \in A\}$ is directed, and

$$\bigsqcup_{a\in A} \pi_\alpha(a) = \pi_\alpha(\bigsqcup A) \in \pi_\alpha(\{c\}\uparrow) = \{\pi_\alpha(c)\}\uparrow,$$

so there exists $a_\alpha \in A$ such that $\pi_\alpha(a_\alpha) \in \{\pi_\alpha(c)\}\uparrow$. Since A is directed, there is a single $a \in A$ that majorizes the finitely many a_α such that $\pi_\alpha(c) \neq \bot$. Then $\pi_\alpha(a) \in \{\pi_\alpha(c)\}\uparrow$ for all α, thus $a \in \{c\}\uparrow$. □

Corollary 1. *The Scott topology on a product of algebraic DCPOs with respect to the componentwise order coincides with the product topology induced by the Scott topology on each component.*

Proof. Let $\prod_{\alpha<\kappa} D_\alpha$ be a product of algebraic DCPOs with $\mathcal{O}_0$ the product topology and $\mathcal{O}_1$ the Scott topology. As noted in the proof of Theorem 5, the projections $\pi_\beta : \prod_\alpha D_\alpha \to D_\beta$ are continuous with respect to $\mathcal{O}_1$. By definition, $\mathcal{O}_0$ is the weakest topology on the product such that the projections are continuous, so $\mathcal{O}_0 \subseteq \mathcal{O}_1$.

For the reverse inclusion, we use the observation that the sets $\{c\}\uparrow$ for finite elements $c \in F$ as defined in Theorem 5 form a base for the topology $\mathcal{O}_1$. These sets are also open in $\mathcal{O}_0$, since they are finite intersections of sets of the form $\pi_\alpha^{-1}(\{\pi_\alpha(c)\}\uparrow)$, and $\{\pi_\alpha(c)\}\uparrow$ is open in D_α since $\pi_\alpha(c) \in F_\alpha$. As $\mathcal{O}_1$ is the smallest topology containing its basic open sets, $\mathcal{O}_1 \subseteq \mathcal{O}_0$. □

A function $g : 2^{\mathsf{H}} \to \mathbb{R}_+$ is *$\mathcal{O}$-simple* if it is a finite linear combination of the form $\sum_{A\in F} r_A \mathbf{1}_A$, where F is a finite subset of $\mathcal{O}$. Let S denote the set of $\mathcal{O}$-simple functions.

Theorem 6. *Let f be a bounded Scott-continuous function $f : 2^{\mathsf{H}} \to \mathbb{R}_+$. Then*

$$\sup_{\substack{g\in S\\ g\le f}} \int g\,d\mu = \int f\,d\mu = \inf_{\substack{g\in S\\ f\le g}} \int g\,d\mu$$

under Lebesgue integration.

Proof. Let $\varepsilon > 0$ and $r_N = \sup_{a \in 2^{\mathsf{H}}} f(a)$. Let

$$0 = r_0 < r_1 < \cdots < r_N$$

such that $r_{i+1} - r_i < \varepsilon$, $0 \leq i \leq N-1$, and set

$$A_i = \{a \mid f(a) > r_i\} = f^{-1}((r_i, \infty)) \in \mathcal{O}, \quad 0 \leq i \leq N.$$

Then $A_{i+1} \subseteq A_i$ and

$$A_i - A_{i+1} = \{a \mid r_i < f(a) \leq r_{i+1}\} = f^{-1}((r_i, r_{i+1}]).$$

Let

$$f_\bullet = \sum_{i=0}^{N-1} r_i \mathbf{1}_{A_i - A_{i+1}} \qquad\qquad f^\bullet = \sum_{i=0}^{N-1} r_{i+1} \mathbf{1}_{A_i - A_{i+1}}.$$

For $a \in A_i - A_{i+1}$,

$$\begin{aligned} f_\bullet(a) &= \sum_{i=0}^{N-1} r_i \mathbf{1}_{A_i - A_{i+1}}(a) = r_i < f(a) \\ &\leq r_{i+1} = \sum_{i=0}^{N-1} r_{i+1} \mathbf{1}_{A_i - A_{i+1}}(a) = f^\bullet(a), \end{aligned}$$

and as a was arbitrary, $f_\bullet \leq f \leq f^\bullet$ pointwise. Thus

$$\int f_\bullet \, d\mu \leq \int f \, d\mu \leq \int f^\bullet \, d\mu.$$

Moreover,

$$\begin{aligned} \int f^\bullet \, d\mu - \int f_\bullet \, d\mu &= \sum_{i=0}^{N-1} r_{i+1} \mu(A_i - A_{i+1}) - \sum_{i=0}^{N-1} r_i \mu(A_i - A_{i+1}) \\ &= \sum_{i=0}^{N-1} (r_{i+1} - r_i) \mu(A_i - A_{i+1}) \\ &< \varepsilon \cdot \sum_{i=0}^{N-1} \mu(A_i - A_{i+1}) = \varepsilon \cdot \mu(2^{\mathsf{H}}) = \varepsilon, \end{aligned}$$

so the integral is approximated arbitrarily closely from above and below by the $f^\bullet$ and $f_\bullet$. Finally, we argue that $f_\bullet$ and $f^\bullet$ are $\mathcal{O}$-simple. Using the fact that

$r_0 = 0$ and $A_N = \varnothing$ to reindex,

$$\begin{aligned} f_\bullet &= \sum_{i=0}^{N-1} r_i \mathbf{1}_{A_i - A_{i+1}} = \sum_{i=0}^{N-1} r_i \mathbf{1}_{A_i} - \sum_{i=0}^{N-1} r_i \mathbf{1}_{A_{i+1}} \\ &= \sum_{i=0}^{N-1} r_{i+1} \mathbf{1}_{A_{i+1}} - \sum_{i=0}^{N-1} r_i \mathbf{1}_{A_{i+1}} = \sum_{i=0}^{N-1} (r_{i+1} - r_i) \mathbf{1}_{A_{i+1}}, \\ f^\bullet &= \sum_{i=0}^{N-1} r_{i+1} \mathbf{1}_{A_i - A_{i+1}} = \sum_{i=0}^{N-1} r_{i+1} \mathbf{1}_{A_i} - \sum_{i=0}^{N-1} r_{i+1} \mathbf{1}_{A_{i+1}} \\ &= \sum_{i=0}^{N-1} r_{i+1} \mathbf{1}_{A_i} - \sum_{i=0}^{N-1} r_i \mathbf{1}_{A_i} = \sum_{i=0}^{N-1} (r_{i+1} - r_i) \mathbf{1}_{A_i}, \end{aligned}$$

and both functions are $\mathcal{O}$-simple since all A_i are in $\mathcal{O}$. □

Lemma 11.

(i) *For any Scott-continuous function* $f : 2^{\mathsf{H}} \to \mathbb{R}_+$, *the map*

$$\mu \mapsto \int f \, d\mu \tag{4.5}$$

is Scott-continuous with respect to the order $\sqsubseteq$ *on* $\mathcal{M}(2^{\mathsf{H}})$.

(ii) *For any probability measure* μ, *the map*

$$f \mapsto \int f \, d\mu \tag{4.6}$$

is Scott-continuous with respect to the order $\sqsubseteq$ *on* $[2^{\mathsf{H}} \to [0, 1]]$.

Thus the Lebesgue integral is Scott-continuous in both arguments.

Proof. (i) We prove the result first for $\mathcal{O}$-simple functions. If $\mu \sqsubseteq \nu$, then for any $\mathcal{O}$-simple function $g = \sum_A r_A \mathbf{1}_A$,

$$\begin{aligned} \int g \, d\mu &= \int \sum_A r_A \mathbf{1}_A \, d\mu = \sum_A r_A \mu(A) \\ &\leq \sum_A r_A \nu(A) = \int \sum_A r_A \mathbf{1}_A \, d\nu = \int g \, d\nu. \end{aligned}$$

Thus the map (4.5) is monotone. If $\mathcal{D}$ is a directed set of measures with respect to $\sqsubseteq$, then

$$\begin{aligned} \int g \, d(\textstyle\bigsqcup \mathcal{D}) &= \int \sum_A r_A \mathbf{1}_A \, d(\textstyle\bigsqcup \mathcal{D}) = \sum_A r_A (\textstyle\bigsqcup \mathcal{D})(A) \\ &= \sup_{\mu \in \mathcal{D}} \sum_A r_A \mu(A) = \sup_{\mu \in \mathcal{D}} \int \sum_A r_A \mathbf{1}_A \, d\mu = \sup_{\mu \in \mathcal{D}} \int g \, d\mu. \end{aligned}$$

Now consider an arbitrary Scott-continuous function $f : 2^{\mathsf{H}} \to \mathbb{R}_+$. Let S is the family of $\mathcal{O}$-simple functions. By Theorem 6, if $\mu \sqsubseteq \nu$, we have

$$\int f\, d\mu = \sup_{\substack{g \in S \\ g \le f}} \int g\, d\mu \le \sup_{\substack{g \in S \\ g \le f}} \int g\, d\nu = \int f\, d\nu,$$

and if $\mathcal{D}$ is a directed set of measures with respect to $\sqsubseteq$, then

$$\begin{aligned} \int f\, d(\bigsqcup \mathcal{D}) &= \sup_{\substack{g \in S \\ g \le f}} \int g\, d(\bigsqcup \mathcal{D}) = \sup_{\substack{g \in S \\ g \le f}} \sup_{\mu \in \mathcal{D}} \int g\, d\mu \\ &= \sup_{\mu \in \mathcal{D}} \sup_{\substack{g \in S \\ g \le f}} \int g\, d\mu = \sup_{\mu \in \mathcal{D}} \int f\, d\mu. \end{aligned}$$

(ii) The result is a straightforward consequence of Theorem 6.

Actually, this is just the monotone convergence theorem for Lebesgue Integration. In fact, it holds more generally for $f : 2^{\mathsf{H}} \to [0, \infty]$, *i.e.* functions that may take on infinite value. –Steffen

□

4.2 Continuous Operations on Measures

Lemma 12. *For any probability measure μ on an algebraic DCPO and open set B, the value $\mu(B)$ is approximated arbitrarily closely from below by $\mu(C)$ for compact-open sets C.*

Proof. Since the sets $\{a\}\uparrow$ for finite a form a base for the topology, and every compact-open set is a finite union of such sets, the set $\mathcal{K}(B)$ of compact-open subsets of B is a directed set whose union is B. Then

$$\mu(B) = \mu(\bigcup \mathcal{K}(B)) = \sup\{\mu(C) \mid C \in \mathcal{K}(B)\}.$$

□

Lemma 13. *The product operator on measures in algebraic DCPOs is Scott-continuous in each argument.*

Proof. The difficult part of the argument is monotonicity. Once we have that, then for any $B, C \in \mathcal{O}$, we have $(\mu \times \nu)(B \times C) = \mu(B) \cdot \nu(C)$. Thus for any directed set D of measures,

$$\begin{aligned} (\bigsqcup D \times \nu)(B \times C) &= (\bigsqcup D)(B) \cdot \nu(C) = (\sup_{\mu \in D} \mu(B)) \cdot \nu(C) \\ &= \sup_{\mu \in D} (\mu(B) \cdot \nu(C)) = \sup_{\mu \in D} ((\mu \times \nu)(B \times C)) \\ &= (\bigsqcup_{\mu \in D} (\mu \times \nu))(B \times C). \end{aligned}$$

By Theorem 5, the sets $B \times C$ for $B, C \in \mathcal{O}$ form a basis for the Scott topology on the product space $2^{\mathsf{H}} \times 2^{\mathsf{H}}$, thus $\bigsqcup D \times \nu = \bigsqcup_{\mu \in D} (\mu \times \nu)$.

To show monotonicity, we use approximability by compact-open sets (Lemma 12). We wish to show that if $\mu_1 \sqsubseteq \mu_2$, then $\mu_1 \times \nu \sqsubseteq \mu_2 \times \nu$. By Lemma 12, it suffices to show that

$$(\mu_1 \times \nu)(\bigcup_n B_n \times C_n) \leq (\mu_2 \times \nu)(\bigcup_n B_n \times C_n),$$

where the index n ranges over a finite set, and B_n and C_n are open sets of the component spaces. Consider the collection of all atoms A of the Boolean algebra generated by the C_n. For each such atom A, let

$$N(A) = \{n \mid C_n \text{ occurs positively in } A\}.$$

Then

$$\bigcup_n B_n \times C_n = \bigcup_A (\bigcup_{n \in N(A)} B_n) \times A.$$

The right-hand side is a disjoint union, since the A are pairwise disjoint. Then

$$\begin{aligned}
(\mu_1 \times \nu)(\bigcup_n B_n \times C_n) &= (\mu_1 \times \nu)(\bigcup_A (\bigcup_{n \in N(A)} B_n) \times A) \\
&= \sum_A (\mu_1 \times \nu)((\bigcup_{n \in N(A)} B_n) \times A) \\
&= \sum_A \mu_1(\bigcup_{n \in N(A)} B_n) \cdot \nu(A) \\
&\leq \sum_A \mu_2(\bigcup_{n \in N(A)} B_n) \cdot \nu(A) \\
&= (\mu_2 \times \nu)(\bigcup_n B_n \times C_n).
\end{aligned}$$

□

Let S and T be measurable spaces and $f : S \to T$ a measurable function. For a measure μ on S, the *push-forward measure* $f_*(\mu)$ is the measure $\mu \circ f^{-1}$ on T.

Lemma 14. *If $f : (2^{\mathsf{H}})^\kappa \to 2^{\mathsf{H}}$ is Scott-continuous with respect to the subset order, then the push-forward operator $f_* : \mathcal{M}((2^{\mathsf{H}})^\kappa) \to \mathcal{M}(2^{\mathsf{H}})$ is Scott-continuous with respect to $\sqsubseteq$.*

Proof. Let $\mu, \nu \in \mathcal{M}((2^{\mathsf{H}})^\kappa)$, $\mu \sqsubseteq \nu$. If $B \in \mathcal{O}$, then $f^{-1}(B)$ is Scott-open in $(2^{\mathsf{H}})^\kappa$, so $f_*(\mu)(B) = \mu(f^{-1}(B)) \leq \nu(f^{-1}(B)) = f_*(\nu)(B)$. As $B \in \mathcal{O}$ was arbitrary, $f_*(\mu) \sqsubseteq f_*(\nu)$. Similarly, if D is any $\sqsubseteq$-directed set in $\mathcal{M}((2^{\mathsf{H}})^\kappa)$, then so is $\{f_*(\mu) \mid \mu \in D\}$, and

$$\begin{aligned} f_*(\bigsqcup D)(B) &= (\bigsqcup D)(f^{-1}(B)) = \sup_{\mu \in D} \mu(f^{-1}(B)) \\ &= \sup_{\mu \in D} f_*(\mu)(B) = (\bigsqcup\nolimits_{\mu \in D} f_*(\mu))(B) \end{aligned}$$

for any $B \in \mathcal{O}$, thus $f_*(\bigsqcup D) = \bigsqcup_{\mu \in D} f_*(\mu)$. □

Lemma 15. *Parallel composition of measures (&) is Scott-continuous in each argument.*

Proof. By definition, $\mu \,\&\, \nu = (\mu \times \nu) \,;\, \bigcup^{-1}$, where $\bigcup : 2^{\mathsf{H}} \times 2^{\mathsf{H}} \to 2^{\mathsf{H}}$ is the set union operator. The set union operator is easily shown to be continuous with respect to the Scott topologies on $2^{\mathsf{H}} \times 2^{\mathsf{H}}$ and the 2^{H}. By Lemma 14, the push-forward operator with respect to union is Scott-continuous with respect to $\sqsubseteq$. By Lemma 13, the product operator is Scott-continuous in each argument with respect to $\sqsubseteq$. The operator & is the composition of these two Scott continuous operators, therefore is itself Scott-continuous. □

4.3 Continuous Kernels

In this section we prove that all ProbNetKAT programs give rise to continuous kernels.

Lemma 16. *The deterministic kernel associated with any Scott-continuous function $f : D \to E$ is a continuous kernel.*

Proof. Recall from [2] that deterministic kernels are those whose output measures are Dirac measures (point masses). Any measurable function $f : D \to E$ uniquely determines a deterministic kernel P_f such that $P_f(a, -) = \delta_{f(a)}$ and vice versa (this was shown in [2] for $D = E = 2^{\mathsf{H}}$). We show that if in addition f is Scott-continuous, then the kernel P_f is continuous.

Let $f : D \to E$ be Scott-continuous. For any open B, if $a \sqsubseteq b$, then $f(a) \sqsubseteq f(b)$ since f is monotone. Since B is up-closed, if $f(a) \in B$, then $f(b) \in B$. Thus

$$P_f(a, B) = [f(a) \in B] \le [f(b) \in B] = P_f(b, B).$$

If $A \subseteq D$ is a directed set, then $f(\bigsqcup A) = \bigsqcup_{a \in A} f(a)$. Since B is open, $\bigsqcup_{a \in A} f(a) \in B$ iff there exists $a \in A$ such that $f(a) \in B$. Then

$$\begin{aligned} P_f(\bigsqcup A, B) &= [f(\bigsqcup A) \in B] = [\bigsqcup\nolimits_{a \in A} f(a) \in B] \\ &= \sup_{a \in A} [f(a) \in B] = \sup_{a \in A} P_f(a, B). \end{aligned}$$

□

All atomic ProbNetKAT programs are deterministic and Scott-continuous. The assignment $x := n$ takes a set a and returns (the Dirac measure on) the set $\{\pi[n/x]\!:\!h \mid \pi\!:\!h \in a\}$. The test $x = n$ takes a set a and returns the set $\{\pi\!:\!h \in a \mid \pi(x) = n\}$. The generalized test b takes a set a and returns the set $a \cap b$. The `dup` operator takes a set a and returns the set $\{\pi\!:\!\pi\!:\!h \mid \pi\!:\!h \in a\}$. All these functions are of the form $a \mapsto \{f(h) \mid h \in a\}$, where f is a partial function $H \rightharpoonup H$, and anything of this form is Scott-continuous:

- If $a \subseteq b$, then $\{f(h) \mid h \in a\} \subseteq \{f(h) \mid h \in b\}$; and
- If $D \subseteq 2^{\mathsf{H}}$ is a directed set, then $\{f(h) \mid h \in \bigcup D\} = \bigcup_{a \in D}\{f(h) \mid h \in a\}$.

Lemma 17. *Let P be a continuous Markov kernel and $f : 2^{\mathsf{H}} \to \mathbb{R}_+$ a Scott-continuous function. Then the map*

$$a \mapsto \int_{c \in 2^{\mathsf{H}}} f(c) \cdot P(a, dc) \tag{4.7}$$

is Scott-continuous.

Proof. The map (4.7) is the composition of the maps

$$a \mapsto P(a, -) \qquad\qquad P(a, -) \mapsto \int_{c \in 2^{\mathsf{H}}} P(a, dc) \cdot f(c),$$

which are Scott-continuous by Lemmas 25 and 11, respectively, and the composition of Scott-continuous maps is Scott-continuous. □

Lemma 18. *Product preserves continuity of Markov kernels: If P and Q are continuous, then so is $P \times Q$.*

Proof. We wish to show that if $a \subseteq b$, then $(P \times Q)(a, -) \sqsubseteq (P \times Q)(b, -)$, and if A is a directed subset of 2^{H}, then $(P \times Q)(\bigcup A) = \sup_{a \in A}(P \times Q)(a, -)$. For the first statement, using Lemma 13 twice,

$$\begin{aligned}(P \times Q)(a, -) &= P(a, -) \times Q(a, -) \sqsubseteq P(b, -) \times Q(a, -)\\ &\sqsubseteq P(b, -) \times Q(b, -) = (P \times Q)(b, -).\end{aligned}$$

For the second statement, for A a directed subset of 2^{H},

$$\begin{aligned}(P \times Q)(\textstyle\bigsqcup A, -) &= P(\textstyle\bigsqcup A, -) \times Q(\textstyle\bigsqcup A, -)\\ &= (\textstyle\bigsqcup_{a \in A} P(a, -)) \times (\textstyle\bigsqcup_{b \in A} Q(b, -))\\ &= \textstyle\bigsqcup_{a \in A} \textstyle\bigsqcup_{b \in A} P(a, -) \times Q(b, -)\\ &= \textstyle\bigsqcup_{a \in A} P(a, -) \times Q(a, -)\\ &= \textstyle\bigsqcup_{a \in A} (P \times Q)(a, -).\end{aligned}$$

□

Lemma 19. *Sequential composition preserves continuity of Markov kernels: If P and Q are continuous, then so is $P\,;Q$.*

Proof. We have

$$(P\,;Q)(a,A) = \int_{c\in 2^{\mathsf{H}}} P(a,dc)\cdot Q(c,A).$$

Since Q is a continuous kernel, it is Scott-continuous in its first argument, thus so is $P\,;Q$ by Lemma 17. □

Lemma 20. *Parallel composition preserves continuity of Markov kernels: If P and Q are continuous, then so is $P\,\&\,Q$.*

Proof. Suppose P and Q are continuous. By definition, $P\,\&\,Q = (P\times Q)\,;\bigcup$. By Lemma 18, $P\times Q$ is continuous, and $\bigcup : 2^{\mathsf{H}}\times 2^{\mathsf{H}} \to 2^{\mathsf{H}}$ is continuous. Thus their composition is continuous by Lemma 19.

Need to generalize lem:seqcomp-preserves –DCK

□

Lemma 21. *The probabilistic choice operator $(\oplus_r)$ preserves continuity of kernels.*

Proof. If P and Q are continuous, then $P\oplus_r Q = rP + (1-r)Q$. If $a\subseteq b$, then

$$\begin{aligned}(P\oplus_r Q)(a,-) &= rP(a,-) + (1-r)Q(a,-)\\ &\le rP(b,-) + (1-r)Q(b,-) = (P\oplus_r Q)(b,-).\end{aligned}$$

If $A\subseteq 2^{\mathsf{H}}$ is a directed set, then

$$\begin{aligned}(P\oplus_r Q)(\textstyle\bigcup A,-) &= rP(\textstyle\bigcup A,-) + (1-r)Q(\textstyle\bigcup A,-)\\ &= \textstyle\bigsqcup_{a\in A}(rP(a,-) + (1-r)Q(a,-))\\ &= \textstyle\bigsqcup_{a\in A}(P\oplus_r Q)(a,-).\end{aligned}$$

□

Lemma 22. *The iteration operator $(*)$ preserves continuity of kernels.*

Proof. Suppose P is continuous. It follows inductively using Lemmas 20 and 19 that $P^{(n)}$ is continuous. Since $P^* = \bigsqcup_n P^{(n)}$ and since the supremum of a directed set of continuous kernels is continuous by Theorem 4, P^* is continuous. □

Theorem 7. *All ProbNetKAT programs give continuous Markov kernels.*

Proof. By Lemmas 16, 20, 19, 21, and 22, all primitive programs are continuous kernels, and continuity is preserved by all the program operators. □

4.4 Continuous Operations on Kernels

In this section we show that all program operators are continuous functions on the DCPO of continuous Markov kernels.

Lemma 23. *The product operation on kernels* $(\times)$ *is Scott-continuous in each argument.*

Proof. We use Lemma 13. If $P_1 \sqsubseteq P_2$, then for all $a \in 2^{\mathsf{H}}$,

$$(P_1 \times Q)(a, -) = P_1(a, -) \times Q(a, -) \sqsubseteq P_2(a, -) \times Q(a, -) = (P_2 \times Q)(a, -).$$

Since a was arbitrary, $P_1 \times Q \sqsubseteq P_2 \times Q$. For a directed set $\mathcal{D}$ of kernels,

$$\begin{aligned}(\textstyle\bigsqcup \mathcal{D} \times Q)(a, -) &= (\textstyle\bigsqcup \mathcal{D})(a, -) \times Q(a, -) \\ &= \textstyle\bigsqcup_{P \in \mathcal{D}} P(a, -) \times Q(a, -) \\ &= \textstyle\bigsqcup_{P \in \mathcal{D}} (P(a, -) \times Q(a, -)) \\ &= \textstyle\bigsqcup_{P \in \mathcal{D}} (P \times Q)(a, -) \\ &= (\textstyle\bigsqcup_{P \in \mathcal{D}} (P \times Q))(a, -).\end{aligned}$$

Since a was arbitrary, $\bigsqcup \mathcal{D} \times Q = \bigsqcup_{P \in \mathcal{D}} (P \times Q)$. ☐

Lemma 24. *Parallel composition of kernels* $(\&)$ *is Scott-continuous in each argument.*

Proof. By definition, $P \,\&\, Q = (P \times Q) \,;\, \bigcup$. By Lemmas 23 and 26, the product operation and sequential composition are continuous in both arguments, thus their composition is. ☐

Lemma 25. *Let P be a continuous Markov kernel. The map* $\mathsf{curry}\, P$ *is Scott-continuous with respect to the subset order on* 2^{H} *and the order* $\sqsubseteq$ *on* $\mathcal{M}(2^{\mathsf{H}})$.

Proof. We have $(\mathsf{curry}\, P)(a) = P(a, -)$. Since P is monotone in its first argument, if $a \subseteq b$ and $B \in \mathcal{O}$, then $P(a, B) \leq P(b, B)$. As $B \in \mathcal{O}$ was arbitrary,

$$(\mathsf{curry}\, P)(a) = P(a, -) \sqsubseteq P(b, -) = (\mathsf{curry}\, P)(b).$$

This shows that $\mathsf{curry}\, P$ is monotone.

Let $D \subseteq 2^{\mathsf{H}}$ be a directed set. By the monotonicity of $\mathsf{curry}\, P$, so is the set $\{(\mathsf{curry}\, P)(a) \mid a \in D\}$. Then for any $B \in \mathcal{O}$,

$$\begin{aligned}(\mathsf{curry}\, P)(\textstyle\bigcup D)(B) = P(\textstyle\bigcup D, B) &= \sup_{a \in D} P(a, B) \\ &= \sup_{a \in D} (\mathsf{curry}\, P)(a)(B) = (\textstyle\bigsqcup_{a \in D} (\mathsf{curry}\, P)(a))(B),\end{aligned}$$

thus $(\mathsf{curry}\, P)(\bigcup D) = \bigsqcup_{a \in D} (\mathsf{curry}\, P)(a)$. ☐

Lemma 26. *Sequential composition of kernels (;) is Scott-continuous in each argument.*

Proof. To show that ; is continuous in its first argument, we wish to show that if P_1, P_2, Q are any continuous kernels with $P_1 \sqsubseteq P_2$, and if $\mathcal{D}$ is any directed set of continuous kernels, then

$$P_1 \,;\, Q \leq P_2 \,;\, Q \qquad\qquad (\textstyle\bigsqcup \mathcal{D}) \,;\, Q = \bigsqcup_{P \in \mathcal{D}} (P \,;\, Q).$$

We must show that for all $a \in 2^{\mathsf{H}}$ and $B\mathcal{O}$,

$$\int_c P_1(a, dc) \cdot Q(c, B) \leq \int_c P_2(a, dc) \cdot Q(c, B)$$
$$\int_c (\textstyle\bigsqcup \mathcal{D})(a, dc) \cdot Q(c, B) = \sup_{P \in \mathcal{D}} \int_c P(a, dc) \cdot Q(c, B).$$

By Lemma 10, for all $a \in 2^{\mathsf{H}}$, $P_1(a, -) \sqsubseteq P_2(a, -)$ and $(\bigsqcup \mathcal{D})(a, -) = \bigsqcup_{P \in \mathcal{D}} P(a, -)$, and $Q(-, B)$ is a Scott-continuous function by assumption. The result follows from Lemma 11(i).

The argument that ; is continuous in its second argument is similar, using Lemma 11(ii). We wish to show that if P, Q_1, Q_2 are any continuous kernels with $Q_1 \sqsubseteq Q_2$, and if $\mathcal{D}$ is any directed set of continuous kernels, then

$$P \,;\, Q_1 \leq P \,;\, Q_2 \qquad\qquad P \,;\, \textstyle\bigsqcup \mathcal{D} = \bigsqcup_{Q \in \mathcal{D}} (P \,;\, Q).$$

We must show that for all $a \in 2^{\mathsf{H}}$ and $B \in \mathcal{O}$,

$$\int_c P(a, dc) \cdot Q_1(c, B) \leq \int_c P(a, dc) \cdot Q_2(c, B)$$
$$\int_c P(a, dc) \cdot (\textstyle\bigsqcup \mathcal{D})(c, B) = \sup_{Q \in \mathcal{D}} \int_c P(a, dc) \cdot Q(c, B).$$

By Lemma 10, for all $B \in \mathcal{O}$, $Q_1(-, B) \sqsubseteq Q_2(-, B)$ and $(\bigsqcup \mathcal{D})(-, B) = \bigsqcup_{Q \in \mathcal{D}} Q(-, B)$. The result follows from Lemma 11(ii). $\square$

Lemma 27. *The probabilistic choice operator applied to kernels ($\oplus_r$) is continuous in each argument.*

Proof. If P and Q are continuous, then $P \oplus_r Q = rP + (1-r)Q$. If $P_1 \sqsubseteq P_2$, then for any $a \in 2^{\mathsf{H}}$ and $B \in \mathcal{O}$,

$$\begin{aligned}(P_1 \oplus_r Q)(a, B) &= rP_1(a, B) + (1-r)Q(a, B)\\ &\leq rP_2(a, B) + (1-r)Q(a, B)\\ &= (P_2 \oplus_r Q)(a, B),\end{aligned}$$

so $P_1 \oplus_r Q \sqsubseteq P_2 \oplus_r Q$. If $\mathcal{D}$ is a directed set of kernels and $B\mathcal{O}$, then

$$\begin{aligned}(\bigsqcup\mathcal{D} \oplus_r Q)(a,B) &= r(\bigsqcup\mathcal{D})(a,B) + (1-r)Q(a,B)\\ &= \sup_{P\in\mathcal{D}} (rP(a,B) + (1-r)Q(a,B))\\ &= \sup_{P\in\mathcal{D}} (P \oplus_r Q)(a,B).\end{aligned}$$

□

Lemma 28. *The iteration operator applied to kernels (*) is continuous.*

Proof. It is a straightforward consequence of Lemma 30 and Theorem 10 that if $P \sqsubseteq Q$, then $P^* \sqsubseteq Q^*$. Now let $\mathcal{D}$ be a directed set of kernels. It follows by induction using Lemmas 24 and 26 that the operator $P \mapsto P^{(n)}$ is continuous, thus

$$(\bigsqcup\mathcal{D})^* = \bigsqcup_n(\bigsqcup\mathcal{D})^{(n)} = \bigsqcup_n \bigsqcup_{P\in\mathcal{D}} P^{(n)} = \bigsqcup_{P\in\mathcal{D}} \bigsqcup_n P^{(n)} = \bigsqcup_{P\in\mathcal{D}} P^*.$$

□

5 Relation to the Semantics of [2]

In this section the notation P^* refers to the star operator in the semantics of [2]. We need a different notation. –DCK

The following results are needed to connect the semantics of iteration presented in [2] with the least fixpoint semantics presented here. Recall from [2] the approximants

$$P^{(0)} = \texttt{skip} \qquad\qquad P^{(m+1)} = \texttt{skip} \mathbin{\&} P \,;\, P^{(m)}.$$

It was shown in [2] that for any $c \in 2^{\mathsf{H}}$, the measures $P^{(m)}(c,-)$ converge weakly to $P^*(c,-)$; that is, for any bounded (Cantor-)continuous real-valued function f on 2^{H}, the expected values of f with respect to the measures $P^{(m)}(c,-)$ converge to the expected value of f with respect to $P^*(c,-)$:

$$\lim_{m\to\infty} \int_{a\in 2^{\mathsf{H}}} f(a)\cdot P^{(m)}(c,da) \;=\; \int_{a\in 2^{\mathsf{H}}} f(a)\cdot P^*(c,da).$$

Steffen's short proof –DCK

Theorem 8. *The kernel $Q = \bigsqcup_{n\in\mathbb{N}} P^{(n)}$ is the unique fixpoint of $(\lambda Q.\ \texttt{skip} \mathbin{\&} P \,;\, Q)$ such that $P^{(n)}(a)$ weakly converges to $Q(a)$ (with respect to the Cantor topology) for all $a \in 2^{\mathsf{H}}$.*

Proof. Let P^* denote any fixpoint of $(\lambda Q.\ \texttt{skip}\ \&\ P\ ;\ Q)$ such that the measure $\mu_n = P^{(n)}(a)$ weakly converges to the measure $\mu = P^*(a)$, i.e. such that for all (Cantor-)continuous bounded functions $f : 2^{\mathsf{H}} \to \mathbb{R}$

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu$$

for all $a \in 2^{\mathsf{H}}$. Let $\nu = Q(a)$. Fix an arbitrary Scott-open set V. Since 2^{H} is a Polish space under the Cantor topology, there exists an increasing chain of compact sets

$$C_1 \subseteq C_2 \subseteq \cdots \subseteq V \quad \text{such that} \quad \sup_{n\in\mathbb{N}} \mu(C_n) = \mu(V).$$

By Urysohn's lemma (see [5, 6]), there exist continuous functions $f_n : 2^{\mathsf{H}} \to [0,1]$ such that $f_n(x) = 1$ for $x \in C_n$ and $f(x) = 0$ for $x \in {\sim}V$. We thus have

$$\begin{aligned}
\mu(C_n) &= \int \mathbf{1}_{C_n} d\mu \\
&\leq \int f_n d\mu && \text{by monotonicity of } \int \\
&= \lim_{m\to\infty} \int f_n d\mu_m && \text{by weak convergence} \\
&\leq \lim_{m\to\infty} \int \mathbf{1}_V d\mu_m && \text{by monotonicity of } \int \\
&= \lim_{m\to\infty} \mu_m(V) \\
&= \nu(V) && \text{by pointwise convergence on } \mathcal{O}
\end{aligned}$$

Taking the supremum over n, we get that $\mu(V) \leq \nu(V)$. Since ν is the $\sqsubseteq$-*least* fixpoint, the measures must therefore agree on V, which implies that they are equal by Theorem 1. Thus, any fixpoint of $(\lambda Q.\ \texttt{skip}\ \&\ P\ ;\ Q)$ with the weak convergence property must be equal to Q. But the fixpoint P^* defined in previous work *does* enjoy the weak convergence property, and therefore so does $Q = P^*$. ☐

Dexter's longer proof –DCK

Lemma 29. *In a Polish space D, the values of*

$$\int_{a\in D} f(a) \cdot \mu(da)$$

for continuous $f : D \to [0,1]$ determine μ uniquely.

Proof. Let A be a Borel set. Since we are in a Polish space, $\mu(A)$ is approximated arbitrarily closely from below by $\mu(C)$ for compact sets $C \subseteq A$ and from above by $\mu(U)$ for open sets $U \supseteq A$. By Urysohn's lemma (see [5, 6]), there exists a continuous function $f : D \to [0,1]$ such that $f(a) = 1$ for all $a \in C$ and $f(a) = 0$ for all $a \notin U$. We thus have

$$\mu(C) = \int_{a \in C} f(a) \cdot \mu(da) \leq \int_{a \in D} f(a) \cdot \mu(da) = \int_{a \in U} f(a) \cdot \mu(da) \leq \mu(U)$$
$$\mu(C) \leq \mu(A) \leq \mu(U),$$

thus

$$\left| \mu(A) - \int_{a \in D} f(a) \cdot \mu(da) \right| \leq \mu(U) - \mu(C),$$

and the right-hand side can be made arbitrarily small. □

By Lemma 29, if P, Q are two Markov kernels and

$$\int_{a \in 2^{\mathsf{H}}} f(a) \cdot P(c, da) = \int_{a \in 2^{\mathsf{H}}} f(a) \cdot Q(c, da)$$

for all Cantor-continuous $f : 2^{\mathsf{H}} \to [0,1]$, then $P(c, -) = Q(c, -)$. If this holds for all $c \in 2^{\mathsf{H}}$, then $P = Q$.

Theorem 9. *Let A be a directed set of probability measures with respect to $\sqsubseteq$ and let $f : 2^{\mathsf{H}} \to [0,1]$ be a Cantor-continuous function. Then*

$$\lim_{\mu \in A} \int_{c \in 2^{\mathsf{H}}} f(c) \cdot d\mu = \int_{c \in 2^{\mathsf{H}}} f(c) \cdot d(\bigsqcup A).$$

Proof. Let $\varepsilon > 0$. Since all continuous functions on a compact space are uniformly continuous, for sufficiently large finite b and for all $a \subseteq b$, the value of f does not vary by more than ε on A_{ab}; that is, $\sup_{c \in A_{ab}} f(c) - \inf_{c \in A_{ab}} f(c) < \varepsilon$. Then for any μ,

$$\int_{c \in A_{ab}} f(c) \cdot \mu(dc) - \int_{c \in A_{ab}} \inf_{c \in A_{ab}} f(c) \cdot \mu(dc)$$
$$\leq \int_{c \in A_{ab}} (\sup_{c \in A_{ab}} f(c) - \inf_{c \in A_{ab}} f(c)) \cdot \mu(dc) < \varepsilon \cdot \mu(A_{ab}).$$

Moreover,

$$(\bigsqcup A)(A_{ab}) = \sum_{a \subseteq c \subseteq b} (-1)^{|c-a|} (\bigsqcup A)(B_c) = \sum_{a \subseteq c \subseteq b} (-1)^{|c-a|} \sup_{\mu \in A} \mu(B_c)$$
$$= \lim_{\mu \in A} \sum_{a \subseteq c \subseteq b} (-1)^{|c-a|} \mu(B_c) = \lim_{\mu \in A} \mu(A_{ab}),$$

so for sufficiently large $\mu \in A$, $\mu(A_{ab})$ does not differ from $(\bigsqcup A)(A_{ab})$ by more than $\varepsilon \cdot 2^{-|b|}$. Then for any constant $r \in [0,1]$,

$$\begin{aligned}&\left|\int_{c\in A_{ab}} r \cdot (\bigsqcup A)(dc) - \int_{c\in A_{ab}} r \cdot \mu(dc)\right| = r \cdot |(\bigsqcup A)(A_{ab}) - \mu(A_{ab})| \\ &\leq |(\bigsqcup A)(A_{ab}) - \mu(A_{ab})| < \varepsilon \cdot 2^{-|b|}.\end{aligned}$$

Combining these observations,

$$\begin{aligned}&\left|\int_{c\in 2^{\mathsf{H}}} f(c) \cdot (\bigsqcup A)(dc) - \int_{c\in 2^{\mathsf{H}}} f(c) \cdot \mu(dc)\right| \\ &= \left|\sum_{a\subseteq b} \int_{c\in A_{ab}} f(c) \cdot (\bigsqcup A)(dc) - \sum_{a\subseteq b} \int_{c\in A_{ab}} f(c) \cdot \mu(dc)\right| \\ &\leq \sum_{a\subseteq b} \Bigg(\left|\int_{c\in A_{ab}} f(c) \cdot (\bigsqcup A)(dc) - \int_{c\in A_{ab}} \inf_{c\in A_{ab}} f(c) \cdot (\bigsqcup A)(dc)\right| \\ &\quad + \left|\int_{c\in A_{ab}} \inf_{c\in A_{ab}} f(c) \cdot (\bigsqcup A)(dc) - \int_{c\in A_{ab}} \inf_{c\in A_{ab}} f(c) \cdot \mu(dc)\right| \\ &\quad + \left|\int_{c\in A_{ab}} \inf_{c\in A_{ab}} f(c) \cdot \mu(dc) - \int_{c\in A_{ab}} f(c) \cdot \mu(dc)\right|\Bigg) \\ &\leq \sum_{a\subseteq b} \Big(\varepsilon \cdot (\bigsqcup A)(A_{ab}) + \varepsilon \cdot 2^{-|b|} + \varepsilon \cdot \mu(A_{ab})\Big) \\ &= 3\varepsilon.\end{aligned}$$

As $\varepsilon > 0$ was arbitrary,

$$\lim_{\mu\in A} \int_{c\in 2^{\mathsf{H}}} f(c) \cdot \mu(dc) = \int_{c\in 2^{\mathsf{H}}} f(c) \cdot (\bigsqcup A)(dc).$$

□

Lemma 30. *If $P \sqsubseteq Q$ then $P^{(n)} \sqsubseteq Q^{(n)}$.*

Proof. By induction on $n \in \mathbb{N}$. The claim is trivial for $n = 0$. For $n > 0$, we assume that $P^{(n-1)} \sqsubseteq Q^{(n-1)}$ and deduce

$$P^{(n)} = \texttt{skip} \mathbin{\&} P \,;\, P^{(n-1)} \sqsubseteq \texttt{skip} \mathbin{\&} Q \,;\, Q^{(n-1)} = Q^{(n)}$$

by monotonicity of sequential and parallel composition (Lemmas 26 and 24, respectively). □

Lemma 31. *If $m \leq n$ then $P^{(m)} \sqsubseteq P^{(n)}$.*

Proof. We have $P^{(0)} \sqsubseteq P^{(1)}$ by Lemmas 9 and 10. Proceeding by induction using Lemma 30, we have $P^{(n)} \sqsubseteq P^{(n+1)}$ for all n. The result follows from transitivity. □

Let P^* be the iterate of P as defined in [2]. As proved in that paper, P^* satisfies the fixpoint equation

$$P^* = \texttt{skip} \mathbin{\&} P \mathbin{;} P^*.$$

Theorem 10. $P^* = \bigsqcup_n P^{(n)}$.

Proof. Consider the continuous transformation

$$T_P(Q) \triangleq \texttt{skip} \mathbin{\&} P \mathbin{;} Q$$

on the DCPO of continuous Markov kernels. The continuity of T_P follows from Lemmas 24 and 26. The bottom element $\bot$ is $\texttt{drop}$ in this space, and

$$T_P(\bot) = \texttt{skip} = P^{(0)} \qquad T_P(P^{(n)}) = \texttt{skip} \mathbin{\&} P \mathbin{;} P^{(n)} = P^{(n+1)},$$

thus $T_P^{n+1}(\bot) = P^{(n)}$, so $\bigsqcup T_P^n(\bot) = \bigsqcup_n P^{(n)}$, and this is the least fixpoint of T_P. As shown in [2], P^* is also a fixpoint of T_P, so it remains to show that $P^* = \bigsqcup_n P^{(n)}$.

Let $c \in 2^{\mathsf{H}}$. As shown in [2], the measures $P^{(n)}(c, -)$ converge weakly to $P^*(c, -)$; that is, for any Cantor-continuous function $f : 2^{\mathsf{H}} \to [0, 1]$, the expected values of f relative to $P^{(n)}$ converge to the expected value of f relative to P^*:

$$\lim_n \int f(a) \cdot P^{(n)}(c, da) = \int f(a) \cdot P^*(c, da).$$

But by Theorem 9, we also have

$$\lim_n \int f(a) \cdot P^{(n)}(c, da) = \int f(a) \cdot (\bigsqcup_n P^{(n)})(c, da),$$

thus

$$\int f(a) \cdot P^*(c, da) = \int f(a) \cdot (\textstyle\bigsqcup_n P^{(n)})(c, da).$$

As f was arbitrary, we have $P^*(c, -) = (\bigsqcup_n P^{(n)})(c, -)$ by Lemma 29, and as c was arbitrary, we have $P^* = \bigsqcup_n P^{(n)}$. ☐

6 Approximation

For a program p, define the n-th approximant $\langle p \rangle_n$ inductively as

$$\begin{aligned}
\langle p \rangle_n &\triangleq p \quad \text{(for } p \text{ primitive)} \\
\langle q \oplus_r r \rangle_n &\triangleq \langle q \rangle_n \oplus_r \langle r \rangle_n \\
\langle q \mathbin{\&} r \rangle_n &\triangleq \langle q \rangle_n \mathbin{\&} \langle r \rangle_n \\
\langle q \mathbin{;} r \rangle_n &\triangleq \langle q \rangle_n \mathbin{;} \langle r \rangle_n \\
\langle q^* \rangle_n &\triangleq q^{(n)}
\end{aligned}$$

Intuitively, $\langle p \rangle_n$ is just p where indefinite iteration $-^*$ is replaced by bounded iteration $-^{(n)}$.

Theorem 11. *The approximants of a program p form a $\sqsubseteq$-increasing chain with supremum p, that is $\langle p \rangle_1 \sqsubseteq \langle p \rangle_2 \sqsubseteq \dots$ and $\bigsqcup_{n \geq 0} [\![\langle p \rangle_n]\!] = [\![p]\!]$.*

Proof. By structural induction on p. □

This means that any program can be approximated using only finite distributions!

References

[1] Samson Abramsky and Achim Jung. Domain theory. In S. Abramsky, Dov M. Gabbay, and T.S.E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3, pages 1–168. Clarendon Press, 1994.

[2] Nate Foster, Dexter Kozen, Konstantinos Mamouras, Mark Reitblatt, and Alexandra Silva. Probabilistic NetKAT. In *ESOP*, pages 282–309, April 2016.

[3] P. R. Halmos. *Measure Theory.* Van Nostrand, 1950.

[4] Claire Jones. *Probabilistic Non-determinism.* PhD thesis, University of Edinburgh, August 1989.

[5] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis.* Prentice Hall, 1970.

[6] M. M. Rao. *Measure Theory and Integration.* Wiley-Interscience, 1987.

[7] N. Saheb-Djahromi. CPOs of measures for nondeterminism. *Theoretical Computer Science*, 12:19–37, 1980.

A Questions

Let $\widehat{\mathcal{O}}$ be the monotone class generated by $\mathcal{O}$; that is, the smallest class of sets containing $\mathcal{O}$ and closed under union of countable ascending chains and intersection of countable descending chains. Every element of $\widehat{\mathcal{O}}$ is Borel, but there are Borel sets not in $\widehat{\mathcal{O}}$, as every element of $\widehat{\mathcal{O}}$ is up-closed.

Lemma 32. *$\mu \sqsubseteq \nu$ iff $\mu(B) \leq \nu(B)$ for all $B \in \widehat{\mathcal{O}}$.*

Proof. If $B_0 \subseteq B_1 \subseteq B_2 \subseteq \cdots$ and $\mu(B_n) \leq \nu(B_n)$ for all n, then

$$\mu(\bigcup_n B_n) = \sup_n \mu(B_n) \leq \sup_n \nu(B_n) = \nu(\bigcup_n B_n).$$

Similarly, if $B_0 \supseteq B_1 \supseteq B_2 \supseteq \cdots$ and $\mu(B_n) \leq \nu(B_n)$ for all n, then

$$\mu(\bigcap_n B_n) = \inf_n \mu(B_n) \leq \inf_n \nu(B_n) = \nu(\bigcap_n B_n).$$

□

Conjecture 1. *$\mu \sqsubseteq \nu$ iff $\mu(B) \leq \nu(B)$ for all up-closed sets $B \in \mathcal{B}$.*

Proof. Surely ($\Leftarrow$) holds, since every element of $\mathcal{O}$ is up-closed. Converse? □

Conjecture 2. *The up-closed sets of $\mathcal{B}$ are just $\widehat{\mathcal{O}}$.*

Proof. Surely ($\supseteq$) holds. Converse? □

By Lemma 32, a positive answer to Conjecture 2 would imply a positive answer to Conjecture 1. All up-closed sets with a countable set of minimal elements are in $\widehat{\mathcal{O}}$, but this is not all of them:

Lemma 33. *There is an element of $\widehat{\mathcal{O}}$ with no countable set of minimal elements.*

Proof. Let A_n be the subset of 2^{H} described by the characteristic strings

$$(1(0+1)+(0+1)1)^n(0+1)^\omega.$$

Then A_n is up-closed and its minimal elements are the antichain $(10+01)^n 0^\omega$. This is a finite set of minimal elements for A_n, thus $A_n \in \mathcal{O}$. Note also that $A_{n+1} \subseteq A_n$, since

$$\begin{aligned}(10+01)^{n+1}0^\omega &= (10+01)^n(10+01)0^\omega \\ &\geq (10+01)^n(00+00)0^\omega = (10+01)^n 0^\omega.\end{aligned}$$

The intersection of the A_n is $(1(0+1)+(0+1)1)^\omega$ with uncountably many minimal elements $(10+01)^\omega$. □

Using Lemma 33, it follows from a cardinality argument that there exist up-closed non-Borel sets. However, it is more difficult to show that there exist up-closed non-Lebesgue measurable sets. The Lebesgue measurable sets are the completion of the Borel sets by nullsets. A construction of an up-closed non-Lebesgue measurable set is given in Appendix B.

B An Up-closed Non-Lebesgue Measurable Set

Define $\lambda(A_{ab}) = 2^{-|b|}$ for $b \in \wp_\omega(H)$. By Theorem 1, λ extends uniquely to a measure on $\mathcal{B}$, called the *uniform* or *Lebesgue measure* on 2^{H}. A set is a *nullset* if it is a subset of a Borel set of Lebesgue measure 0. The *Lebesgue measurable* *sets* are the smallest σ-algebra containing the Borel sets and the nullsets. A set is Lebesgue measurable iff it differs from a Borel set by a nullset; that is, iff there exists a Borel set B such that $A \oplus B$ is a nullset, where $\oplus$ denotes the symmetric difference operator on sets [3, Theorem 13.B]. The measure λ extends uniquely to all Lebesgue measurable sets by assigning $\lambda(N) = 0$ for all nullsets N.

For $c \in 2^{\mathsf{H}}$, define the map $T_c : 2^{2^{\mathsf{H}}} \to 2^{2^{\mathsf{H}}}$ by $T_c(B) = \{a \oplus c \mid a \in B\}$.

Lemma 34.

(i) *T_c is an involution.*

(ii) *For $b \in 2^{\mathsf{H}}$ finite and $a \subseteq b$, $T_c(A_{ab}) = A_{a\oplus(b\cap c),b}$.*

(iii) *T_c preserves the Borel sets and Lebesgue measurable sets.*

(iv) *Lebesgue measure is invariant under T_c.*

Proof. (iv) In light of the definition of λ on atoms A_{ab}, this is immediate from (ii) and Theorem 1. ☐

Lemma 35 ([3, Theorem 16.A])**.** *For any nonnull Lebesgue measurable set B and for all $\varepsilon > 0$, there exists a basic Cantor open set A_{ab} such that $\lambda(A_{ab} \cap B)/\lambda(A_{ab}) \geq 1 - \varepsilon$.*

Proof. The Cantor space 2^{H} is a Polish space, therefore outer regular. This means that for any probability measure μ, any Borel set is approximated in measure arbitrarily closely from above by Cantor open sets; that is, for any Borel set B and $\varepsilon > 0$ there exists a Cantor open set A such that $B \subseteq A$ and $\mu(A - B) \leq \varepsilon$.

Suppose $\lambda(B) > 0$. For $\varepsilon > 0$, let A be a Cantor open set such that $B \subseteq A$ and

$$\lambda(A - B) < \varepsilon \cdot \lambda(B) \leq \varepsilon \cdot \lambda(A). \tag{B.1}$$

The Cantor open set A can be represented as a countable disjoint union of basic Cantor open sets A_{ab}. For each point $c \in A$, let A_{ab} be a maximal Cantor basic open set such that $c \in A_{ab} \subseteq A$. These sets exist by Lemma 3. The collection of all such sets is the desired countable disjoint union $A = \bigcup_n A_n$. If all $\lambda(A_n - B) \geq \varepsilon \cdot \lambda(A_n)$, then

$$\begin{aligned} \lambda(A - B) &= \lambda(\bigcup_n A_n - B) = \sum_n \lambda(A_n - B) \\ &\geq \sum_n \varepsilon \cdot \lambda(A_n) = \varepsilon \cdot \lambda(\bigcup_n A_n) = \varepsilon \cdot \lambda(A), \end{aligned}$$

contradicting (B.1). Thus $\lambda(A_n - B) < \varepsilon \cdot \lambda(A_n)$ for some n, from which the result for Borel sets B follows.

For a Lebesgue measurable set C, let B be a Borel set such that $B \oplus C \subseteq D$, where D is Borel and $\lambda(D) = 0$. Approximate the Borel set $B \cup D$ from above with a Cantor basic open set A. As

$$B - D \subseteq C \subseteq B \cup D \qquad\qquad (B \cup D) - (B - D) \subseteq D,$$

we have

$$\begin{aligned}
\lambda(A - C) &\leq \lambda(A - (B - D)) \\
&= \lambda(A - (B \cup D)) + \lambda((B \cup D) - (B - D)) \\
&\leq \lambda(A - (B \cup D)) + \lambda(D) \\
&\leq \lambda(A - (B \cup D)) \\
&\leq \varepsilon \cdot \lambda(A).
\end{aligned}$$

□

A *filter* on 2^{H} is a nonempty subset of 2^{H} that is up-closed, closed under finite intersection, and does not contain $\varnothing$. A filter is an *ultrafilter* if, whenever $a \cup b \in F$, either $a \in F$ or $b \in F$. An ultrafilter is *principal* if it is of the form $\{a\}\!\uparrow$ for some $a \in 2^{\mathsf{H}}$, *nonprincipal* otherwise. An ultrafilter is nonprincipal iff it contains no finite sets. Nonprincipal ultrafilters can be constructed from the cofinite sets by Zorn's lemma.

Lemma 36. *If c is finite and U is a nonprincipal ultrafilter, then $T_c(U) = U$.*

Proof. If $a \in U$ and c is finite, then $\sim\! c$ is cofinite, therefore $\sim\! c \in U$. Since U is closed under intersection, $a - c \in U$. Since U is up-closed, $a \oplus c = (a - c) \cup (c - a) \in U$. Since $a \in U$ was arbitrary, we have shown that $T_c(U) \subseteq U$. Since T_c is an involution, by monotonicity we also have $U = T_c(T_c(U)) \subseteq T_c(U)$, therefore $T_c(U) = U$. □

Theorem 12. *There exists an up-closed non-Lebesgue measurable set.*

Proof. Let U be a nonprincipal ultrafilter on 2^{H}. Then U is up-closed. We claim that U is not Lebesgue measurable. Suppose it were. By Lemma 35, if $\lambda(U) > 0$, then there is a Cantor basic open set A_{ab} such that $\lambda(A_{ab} \cap U)/\lambda(A_{ab}) \geq 1 - \varepsilon$. Applying T_c for $c \subseteq b$, we have

$$\lambda(A_{ab} \cap U) = \lambda(T_c(A_{ab} \cap U)) = \lambda(T_c(A_{ab}) \cap T_c(U)) = \lambda(A_{a\oplus c,b} \cap U),$$

thus $\lambda(A_{cb} \cap U) = \lambda(A_{ab} \cap U)$ for all $c \subseteq b$. Summing over all atoms of $\mathcal{B}_b$,

$$\begin{aligned}
\lambda(U) = \lambda(\bigcup_{c \subseteq b} A_{cb} \cap U) &= \sum_{c \subseteq b} \lambda(A_{cb} \cap U) = 2^{|b|} \cdot \lambda(A_{ab} \cap U) \\
&\geq 2^{|b|} \cdot (1 - \varepsilon) \cdot \lambda(A_{ab}) = 2^{|b|} \cdot (1 - \varepsilon) \cdot 2^{-|b|} = 1 - \varepsilon.
\end{aligned}$$

As $\varepsilon > 0$ was arbitrary, $\lambda(U) = 1$.

We have argued that if $\lambda(U) > 0$, then $\lambda(U) = 1$. Thus either $\lambda(U) = 0$ or $\lambda(U) = 1$. But both are impossible, because

$$T_H(U) = \{H \oplus a \mid a \in U\} = \{\sim a \mid a \in U\} = 2^{\mathsf{H}} \oplus U = 2^{\mathsf{H}} - U,$$

since for all $a \in 2^{\mathsf{H}}$, $a \in U$ iff $\sim a \notin U$; so by Lemma 36,

$$\lambda(U) = \lambda(T_H(U)) = \lambda(2^{\mathsf{H}} - U) = 1 - \lambda(U),$$

thus $\lambda(U) = 1/2$. This is a contradiction. □

C $\mathcal{M}, \sqsubseteq$ is not an Upper Semilattice

Despite the fact that $(\mathcal{M}, \sqsubseteq)$ is a directed set (Lemma 9), it is not an upper semilattice. Here is a counterexample.

Let $b = \{\pi, \sigma, \tau\}$, where π, σ, τ are distinct packets. Let

$$\mu_1 = \tfrac{1}{2}\delta_{\{\pi\}} + \tfrac{1}{2}\delta_{\{\sigma\}} \qquad \mu_2 = \tfrac{1}{2}\delta_{\{\sigma\}} + \tfrac{1}{2}\delta_{\{\tau\}} \qquad \mu_3 = \tfrac{1}{2}\delta_{\{\tau\}} + \tfrac{1}{2}\delta_{\{\pi\}}.$$

The measures μ_1, μ_2, and μ_3 would be the output measures of the programs $\pi! \oplus \sigma!$, $\sigma! \oplus \tau!$, and $\tau! \oplus \pi!$, respectively.

We claim that $\mu_1 \sqcup \mu_2$ does not exist. To see this, define

$$\nu_1 = \tfrac{1}{2}\delta_{\{\tau\}} + \tfrac{1}{2}\delta_{\{\pi,\sigma\}} \qquad \nu_2 = \tfrac{1}{2}\delta_{\{\pi\}} + \tfrac{1}{2}\delta_{\{\sigma,\tau\}} \qquad \nu_3 = \tfrac{1}{2}\delta_{\{\sigma\}} + \tfrac{1}{2}\delta_{\{\tau,\pi\}}.$$

All ν_i are $\sqsubseteq$-upper bounds for all μ_j. (In fact, any convex combination $r\nu_1 + s\nu_2 + t\nu_3$ for $0 \le r, s, t$ and $r + s + t = 1$ is an upper bound for any convex combination $u\mu_1 + v\mu_2 + w\mu_3$ for $0 \le u, v, w$ and $u + v + w = 1$.) But we show by contradiction that there cannot exist a measure that is both $\sqsubseteq$-above μ_1 and μ_2 and $\sqsubseteq$-below ν_1 and ν_2. Suppose ρ was such a measure. Since $\rho \sqsubseteq \nu_1$ and $\rho \sqsubseteq \nu_2$, we have

$$\rho(B_{\sigma\tau}) \le \nu_1(B_{\sigma\tau}) = 0 \quad \rho(B_{\tau\pi}) \le \nu_1(B_{\tau\pi}) = 0 \quad \rho(B_{\pi\sigma}) \le \nu_2(B_{\pi\sigma}) = 0.$$

Since $\mu_1 \sqsubseteq \rho$ and $\mu_2 \sqsubseteq \rho$, we have

$$\rho(B_\pi) \ge \mu_1(B_\pi) = \tfrac{1}{2} \qquad \rho(B_\sigma) \ge \mu_1(B_\sigma) = \tfrac{1}{2} \qquad \rho(B_\tau) \ge \mu_2(B_\tau) = \tfrac{1}{2}.$$

But then

$$\begin{aligned}
\rho(A_{\pi b}) &= \rho(B_\pi) - \rho(B_{\pi\sigma} \cup B_{\tau\pi}) \ge \tfrac{1}{2} \\
\rho(A_{\sigma b}) &= \rho(B_\sigma) - \rho(B_{\sigma\tau} \cup B_{\pi\sigma}) \ge \tfrac{1}{2} \\
\rho(A_{\tau b}) &= \rho(B_\tau) - \rho(B_{\tau\pi} \cup B_{\sigma\tau}) \ge \tfrac{1}{2},
\end{aligned}$$

which is impossible, because ρ would have total weight at least $\frac{3}{2}$.

D Negative Attempts of Approximation

Given a sequence of sets $X_1, X_2, \ldots$, we write $X_i \uparrow X$ whenever $X_1 \subseteq X_2 \subseteq \ldots$ and $\bigcup_{i \in \mathbb{N}} X_i = X$; and we write $X_i \downarrow X$ whenever $X_1 \supseteq X_2 \supseteq \ldots$ and $\bigcap_{i \in \mathbb{N}} X_i = X$.

Fix an arbitrary sequence of finite history sets $b_i \in \wp_\omega(\mathsf{H})$ such that $b_i \uparrow \mathsf{H}$. We investigate the ability of the boolean subalgebras $\mathcal{B}_{b_i}$ of $\mathcal{B}$ to approximate measurable sets $A \in \mathcal{B}$.

Lemma 37. *Let $b \subseteq \mathsf{H}$ and $h^* \in \mathsf{H} - b$. Then all events $A \in \mathcal{B}_b$ are indifferent with respect to h^*, that is for all history sets $a \subseteq \mathsf{H}$, it holds that*

$$a \in A \iff a \cup \{h^*\} \in A \tag{D.1}$$

Proof. It suffices to show the claim for all events B_h with $h \in b$ since they generate the boolean algebra $\mathcal{B}_b$ and since (D.1) is preserved by union, intersection, and negation of events.

Recall that a set is in B_h iff it contains the history h. If a contains h then cleary $a \cup \{h^*\}$ also contains h. Conversely, if $a \cup \{h^*\}$ contains h then a must contain h, because by assumption $h \in b$ and $h^* \notin b$, that is $h \neq h^*$. □

Lemma 38. *Not all up-closed, measurable sets $A \in \mathcal{B}$ can be approximated from below by sets $A_i \in \mathcal{B}_{b_i}$ in the sense that $A_i \uparrow A$.*

Proof. Consider an infinite sets of histories $a \subseteq \mathsf{H}$ and take $A \triangleq B_a$ to be the set of all its supersets, an up-closed, measurable set. Now suppose we had a sequence of events $A_i \in \mathcal{B}_{b_i}$ with $A_i \subseteq A$. Recall that the b_i are finite by assumption – hence we can pick a history $h_i^* \in a - b_i$ for each i such that, by Lemma 37, all events in $\mathcal{B}_{b_i}$ are indifferent with respect to h_i^*. But this means that A_i must be empty, for if there existed a history set $a_i \in A_i$, we would also have $a_i - \{h_i^*\} \in A_i$ (D.1), contradicting the assumption that $A_i \subseteq A$. But if all A_i are empty, we do not have $A_i \uparrow A$. □

Lemma 39. *Not all up-closed, measurable sets $A \in \mathcal{B}$ can be approximated from above by sets $A_i \in \mathcal{B}_{b_i}$ in the sense that $A_i \downarrow A$.*

Proof. Let $c \triangleq \{\pi^n \in \mathsf{H} \mid n > 0\}$ denote the set of all histories build from a single packet π. For each finite history set b, let h_b denote a history that is not in $c \cup b$ – since b is finite, such a history always exists. Define a "fake" c for any b by $\bar{c}_b \triangleq (c \cap b) \uplus h_b$. Finally, define the measurable set $A \triangleq \bigcup_{b \text{ finite}} B_{\bar{c}_b}$, the smallest upward closed set containing all fakes. Observe that the "real" c is not in A.

Unfortunately, the boolean algebras $\mathcal{B}_{b_i}$ fail to discriminate between c and its fakes. In particular, for any b_i, the atom $A_{c \cap b_i, b_i}$ contains both the fake $\bar{c}_{b_i}$ and the real c. Therefore we must have $A_i \supseteq A_{c \cap b_i, b_i}$, and as a result $c \in A_i$. It follows that $\bigcap_i A_i \neq A$, since A does not contain c. □

Non-Theorem 1. *Let $f : 2^{\mathsf{H}} \to \mathbb{R}$ be a measurable function that is monotone with respect to the subset order on 2^{H} (that is, $a \subseteq b \Rightarrow f(a) \leq f(b)$), and let μ be a probability measure on $(2^{\mathsf{H}}, \mathcal{B})$. Then*

$$\int_{c \in 2^{\mathsf{H}}} f(c) \cdot \mu(dc) = \inf_{b \text{ finite}} \sum_{a \subseteq b} \sup_{c \in A_{ab}} f(c) \cdot \mu(A_{ab})$$

Counterexample. Let $c \triangleq \{\pi^n \in \mathsf{H} \mid n > 0\}$ denote the set of all histories build from a single packet π. For each finite history set b, let h_b denote a history that is not in $c \cup b$ – since b is finite, such a history always exists. Define a "fake" c for any b by $\overline{c}_b \triangleq (c \cap b) \uplus h_b$. Finally, define the measurable set $S \triangleq \bigcup_{b \text{ finite}} B_{\overline{c}_b}$, the smallest upward closed set containing all fakes. Observe that the "real" c is not in S.

We now consider the characteristic function $f \triangleq \mathbf{1}_S$ of S together with the dirac measure $\mu \triangleq \delta_c$ on c. Since S is upward-closed, f is monotone as required. Because $c \notin S$, we have

$$\int f d\mu = \mu(S) = 0$$

Unfortunately, the boolean algebras $\mathcal{B}_b$ fail to discriminate between c and its fakes. In particular, for any finite b, the atom $A_{c \cap b, b}$ contains both c and the fake $\overline{c}_b$; as a result,

$$\sum_{a \subseteq b} \sup_{d \in A_{ab}} f(d) \cdot \mu(A_{ab}) \geq \sup_{d \in A_{c \cap b, b}} f(d) \cdot \mu(A_{c \cap b, b}) = 1$$

and therefore

$$\inf_{b \text{ finite}} \sum_{a \subseteq b} \sup_{d \in A_{ab}} f(d) \cdot \mu(A_{ab}) \geq 1 > \int f d\mu = 0$$

□

E Non-Algebraicity

Here is a counterexample to the conjecture that the elements continuous DCPO of continuous kernels is algebraic with finite elements $b \ ; P \ ; d$. Let σ, τ be packets and let $\sigma!$ and $\tau!$ be the programs that set the current packet to σ or τ, respectively. For $r \in [\frac{1}{2}, 1]$, let $P_r = (\sigma! \oplus_r \tau!) \ \& \ (\tau! \oplus_r \sigma!)$. On any nonempty input, P_r produces $\{\sigma\}$ with probability $r(1-r)$, $\{\tau\}$ with probability $r(1-r)$, and $\{\sigma, \tau\}$ with probability $r^2 + (1-r)^2$. In particular, P_1 produces $\{\sigma, \tau\}$ with probability 1. The kernels P_r for $1/2 \leq r < 1$ form a directed set whose supremum is P_1, yet $\{\sigma\} \ ; P_1 \ ; \{\sigma, \tau\}$ is not $\sqsubseteq$-bounded by any P_r for $r < 1$, therefore the up-closure of $\{\sigma\} \ ; P_1 \ ; \{\sigma, \tau\}$ is not an open set.