Commutative Monads for Probabilistic Programming Languages

Xiaodong Jia, Bert Lindenhovius, Michael Mislove and Vladimir Zamdzhiev

Applied Category Theory
16 July 2021
Motivation

- Probability and recursion are important computational effects.
- *Domain Theory* – staple of denotational study of recursion.
- Adding probability to domain-theoretic approach has been difficult.
- Canonical approach: Kleisli category of the *valuations monad* $\mathcal{V}$ [1].
- Two major open problems unsolved since 1989.
- Many experts have considered other denotational approaches to combining probability and recursion: probabilistic coherence spaces, quasi-Borel spaces, measurable cones and others.
- We show domain theory can combine probability and recursion in an *elegant* way.

Background: Domain Theory (Dcpo’s)

- Domain theory provides an order-theoretic view of computation and recursion.
- Two main classes of objects in domain theory: dcpo’s and domains.
- A nonempty subset $A$ of a poset $D$ is directed if each pair of elements in $A$ has an upper bound in $A$.
- A directed-complete partial order (dcpo) is a poset in which every directed subset $A$ has a supremum $\text{sup } A$.
  - **Example:** the unit interval $[0, 1]$ is a dcpo in the usual ordering.
  - **Example:** the open sets of a topological space in the inclusion order.
- A function $f : D \rightarrow E$ between two dcpo’s is **Scott-continuous** if it is monotone and preserves suprema of directed subsets.
- The category $\text{DCPO}$ of dcpo’s and Scott-continuous functions is **cartesian closed**, complete and cocomplete.
- The category $\text{DCPO}$ is very important for denotational semantics.
Background: Domain Theory (Domains)

- We say $x$ is **way-below** $y$ ($x \ll y$) iff for every directed set $A$ with $y \leq \sup A$, there is some $a \in A$, s.t. $x \leq a$.
- We write $\downarrow y = \{x \in D \mid x \ll y\}$.
- A **basis** for a dcpo $D$ is a subset $B$ satisfying $\downarrow x \cap B$ is directed and $x = \sup \downarrow x \cap B$, for each $x \in D$.
- A dcpo $D$ is **continuous** if it has a basis.
- Continuous dcpo’s are also called **domains**. The category of domains and Scott-continuous maps is denoted by $\text{DOM}$.
- Domains may be thought of as very well-behaved dcpo’s.
- **Problem:** The category $\text{DOM}$ is not cartesian closed.
The order on a dcpo $X$ induces a canonical topology $\sigma_X$, called the *Scott-topology*.

The *Scott topology* $\sigma_D$ on a dcpo $D$ consists of the upper subsets $U = \uparrow U = \{x \in D \mid \exists u \in U. \ u \leq x\}$ that are *inaccessible by directed suprema*: i.e., if $A \subseteq D$ is directed and $\sup A \in U$, then $A \cap U \neq \emptyset$.

The topological space $(D, \sigma_D)$ is also written as $\Sigma D$.

$f : X \rightarrow Y$ is Scott-continuous iff $f$ is continuous w.r.t. $\Sigma X$ and $\Sigma Y$. 

---

**Background: Domain Theory (Scott Topology)**

- The order on a dcpo $X$ induces a canonical topology $\sigma_X$, called the *Scott-topology*.

- The *Scott topology* $\sigma_D$ on a dcpo $D$ consists of the upper subsets $U = \uparrow U = \{x \in D \mid \exists u \in U. \ u \leq x\}$ that are inaccessible by directed suprema: i.e., if $A \subseteq D$ is directed and $\sup A \in U$, then $A \cap U \neq \emptyset$.

- The topological space $(D, \sigma_D)$ is also written as $\Sigma D$.

- $f : X \rightarrow Y$ is Scott-continuous iff $f$ is continuous w.r.t. $\Sigma X$ and $\Sigma Y$. 

---

**Background: Domain Theory (Scott Topology)**

- The order on a dcpo $X$ induces a canonical topology $\sigma_X$, called the *Scott-topology*.

- The *Scott topology* $\sigma_D$ on a dcpo $D$ consists of the upper subsets $U = \uparrow U = \{x \in D \mid \exists u \in U. \ u \leq x\}$ that are inaccessible by directed suprema: i.e., if $A \subseteq D$ is directed and $\sup A \in U$, then $A \cap U \neq \emptyset$.

- The topological space $(D, \sigma_D)$ is also written as $\Sigma D$.

- $f : X \rightarrow Y$ is Scott-continuous iff $f$ is continuous w.r.t. $\Sigma X$ and $\Sigma Y$. 

---

**Background: Domain Theory (Scott Topology)**
Background: Probability and Recursion

- How to talk about recursion \textit{and} probability?
- Why not just take $\text{Meas}(X)$, the set of subprobability measures on the Borel $\sigma$-algebra induced by the Scott-topology of a dcpo $X$?
- Because it is unclear how to extend the assignment $\text{Meas}(\_)$ to a monad over DCPO.
- A monadic semantics over DCPO seems very unlikely with this approach.
Background: Valuations

• The domain-theoretic approach to probability is based on valuations [1].
• A subprobability valuation on a dcpo $X$ is a Scott-continuous map $\nu : \sigma X \to [0, 1]$, which is strict ($\nu(\emptyset) = 0$) and modular ($\nu(U) + \nu(V) = \nu(U \cup V) + \nu(U \cap V)$).
  • Example: The always-zero valuation $0$.
  • Example: For $x \in X$, $\delta_x$ is defined as $\delta_x(U) = 1$ if $x \in U$ and $\delta_x(U) = 0$ otherwise.
• The set of subprobability valuations on a dcpo $X$, denoted $\mathcal{V}X$, is a pointed dcpo in the stochastic order: $\nu_1 \leq \nu_2$ iff $\forall U \in \sigma X. \nu_1(U) \leq \nu_2(U)$.
• Remark: Valuations are similar to Borel measures and in some cases coincide.

Background: Valuations Monad

- The assignment $\mathcal{V}(-)$ can be equipped with the structure of a *strong monad*.
- Given $h : D \to E$, define $\mathcal{V}(h) : \mathcal{V}D \to \mathcal{V}E :: \nu \mapsto \lambda U.\nu(h^{-1}(U))$.
- The unit of $\mathcal{V}$ is given by $\eta_D : D \to \mathcal{V}D :: x \mapsto \delta_x$.
- A notion of integration can be defined. Given $\nu \in \mathcal{V}X$ and $f : X \to [0, 1]$ Scott-continuous, we can define the *integral of f against* $\nu$ by:

\[
\int_{x \in X} f(x) d\nu \overset{\text{def}}{=} \int_{0}^{1} \nu(f^{-1}((t, 1))) dt.
\]

- The multiplication is given by $\mu_D : \mathcal{V}\mathcal{V}D \to \mathcal{V}D :: \varpi \mapsto \lambda U.\int_{\nu \in \mathcal{V}D} \nu(U) d\varpi$.
- The strength is $\tau_{DE} : D \times \mathcal{V}E \to \mathcal{V}(D \times E) :: (x, \nu) \mapsto \lambda U.\int_{y \in E} \chi_U(x, y) d\nu$. 
Background: Problems of the Valuations Monad

- The monad $\mathcal{V}$ is *strong* on DCPO and *commutative* on DOM [2].
- Two major open problems since 1989:
  - **Problem**: Is $\mathcal{V}$ a commutative monad on DCPO?
  - **Problem (Jung-Tix)**: Find a cartesian closed category of domains on which $\mathcal{V}$ is a commutative monad.
- Having a domain-theoretic model with a *commutative valuations monad* over a *cartesian closed category* is important for the semantics. Do they exist?

---

Our approach

• How to construct a domain-theoretic model for probability and recursion:
  • such that we have a \textit{commutative monad} of valuations; and
  • such that this monad is taken over a \textit{cartesian closed category}?

• Our approach and our results:
  • we describe a commutative monad of valuations $\mathcal{M}$ on DCPO (cartesian closed);
  • $\mathcal{M}X \subseteq \mathcal{V}X$ for every dcpo $X$; in fact, $\mathcal{M}$ is a \textit{submonad} of $\mathcal{V}$;
  • $\mathcal{M}$ coincides with $\mathcal{V}$ on domains;
  • $\mathcal{M}$ contains enough valuations for semantics: we show how to define a sound and (strongly) adequate interpretation of PFPC using $\mathcal{M}$;
  • we characterise the Eilenberg-Moore algebras of $\mathcal{M}$ over DOM by showing $\text{DOM}^\mathcal{M} = \text{DOM}^\mathcal{V}$ is isomorphic to the category of continuous Kegelspitzen [3];
  • our constructions use \textit{topological methods} and we construct \textit{two additional} such monads with all of the above properties.

Fubini $\iff$ Commutativity of $\mathcal{V}$

- Commutativity of the monad $\mathcal{V}$ is equivalent to showing the Fubini-style equation

$$\int_{x \in D} \int_{y \in E} \chi_U(x, y) d\xi d\nu = \int_{y \in E} \int_{x \in D} \chi_U(x, y) d\nu d\xi$$

for dcpo's $D$ and $E$, for $U \in \sigma(D \times E)$ and for $\nu \in \mathcal{V}D, \xi \in \mathcal{V}E$.

- This equation is known to hold for *simple* valuations, directed suprema of simple valuations, directed suprema of directed suprema of simple valuations, etc.
Simple Valuations

- $\mathcal{V}X$ has a convex structure: if $\nu_i \in \mathcal{V}X$ and $r_i \geq 0$, with $\sum_{i=1}^{n} r_i \leq 1$, then the convex sum $\sum_{i=1}^{n} r_i \nu_i \overset{\text{def}}{=} \lambda U$. $\sum_{i=1}^{n} r_i \nu_i(U)$ also is in $\mathcal{V}X$.

- The *simple valuations* on a dcpo $X$ are those of the form $\sum_{i=1}^{n} r_i \delta_{x_i}$, where $r_i \geq 0$ and $\sum_{i=1}^{n} r_i \leq 1$.

- The set of simple valuations on $X$ is denoted by $S\mathcal{X}$.

- $S\mathcal{X} \subseteq \mathcal{V}X$, but $S\mathcal{X}$ is not a dcpo in general.
A Commutative Monad of Valuations

- To interpret *discrete* probabilistic choice in programming, it suffices:
  1. to take a class of valuations that contains the simple valuations;
  2. this class of valuations should be closed under directed suprema (for recursion).

- **Definition:** For each dcpo $D$, we define $\mathcal{M}D$ to be the intersection of all sub-dcpo’s of $\mathcal{V}D$ that contain $\mathcal{S}D$.

- In other words, $\mathcal{M}D$ is the smallest sub-dcpo of $\mathcal{V}D$ that contains $\mathcal{S}D$.

- **Theorem:** $\mathcal{M}$ is a commutative monad on $\text{DCPO}$. Its monad operations are (co)restrictions of those of $\mathcal{V}$.

- **Remark:** $\mathcal{M}D$ is *not* the dcpo-completion of $\mathcal{S}D$, in general. It is a *topological completion* of $\mathcal{S}D$ within $\mathcal{V}D$. 
The Monad $\mathcal{M}$ as a Topological Completion

- Given a dcpo $D$, the $d$-topology on $D$ is the topology whose closed subsets consist of sub-dcpo’s of $D$.
- Given a subset $C \subseteq D$, the $d$-closure of $C$ in $D$ is the topological closure of $C$ w.r.t the $d$-topology on $D$.
- $\mathcal{M}D$ is precisely the $d$-closure of $SD$ in $VD$.
- This view is a lot more useful for establishing the required proofs.
- We obtain two additional commutative monads by taking suitable completions of $SD$ in $VD$. 
K-categories, Completions and Commutative Monads

- A $K$-category is a full subcategory of the category $T_0$ of $T_0$-spaces satisfying properties that imply it determines a completion of each of its objects.
- **Example:** The category $D$ of $d$-spaces and continuous maps.
- **Example:** The category $SOB \subseteq D$ of sober spaces and continuous maps.
- **Example:** The category $WF \subseteq D$ of well-filtered spaces and continuous maps.
- **Theorem:** Any $K$-category $K$ with $K \subseteq D$ determines a commutative valuations monad $\mathcal{V}_K$ on DCPO.
- The monad $M$ is recovered as $M = \mathcal{V}_D$.
- Two additional commutative monads: $P = \mathcal{V}_{SOB}$ and $W = \mathcal{V}_{WF}$.
- $SD \subseteq MD \subseteq WD \subseteq PD \subseteq VD$ for each dcpo $D$.
- All subsequent results hold for all three monads $M$, $W$ and $P$. 
Definition of Kegelspitzten

The EM-algebras of $\mathcal{M}$ and $\mathcal{V}$ over domains may be characterised using Kegelspitzten.

Definition

A barycentric algebra is a set $A$ equipped with a binary operation $a +_r b$ for $r \in [0, 1]$ such that for all $a, b, c \in A$ and $r, p \in [0, 1]$, the following equations hold:

\[
\begin{align*}
  a +_1 b &= a; \\
  a +_r b &= b +_{1-r} a; \\
  a +_r a &= a; \\
  (a +_p b) +_r c &= a +_{pr} (b +_{\frac{r-pr}{1-pr}} c) \quad \text{provided } r, p < 1.
\end{align*}
\]

Definition

A pointed barycentric algebra is a barycentric algebra $A$ with a distinguished element $\bot$. For $a \in A$ and $r \in [0, 1]$, we define $r \cdot a \overset{\text{def}}{=} a +_r \bot$. A map $f : A \to B$ between pointed barycentric algebras is called linear if $f(\bot_A) = \bot_B$ and $f(a +_r b) = f(a) +_r f(b)$ for all $a, b \in A, r \in [0, 1]$. 

Definition of Kegelspitzen (Contd.)

Definition
A Kegelspitze is a pointed barycentric algebra $K$ equipped with a directed-complete partial order such that, for every $r$ in the unit interval, the functions determined by convex combination $(a, b) \mapsto a + r \cdot b: K \times K \to K$ and scalar multiplication $(r, a) \mapsto r \cdot a: [0, 1] \times K \to K$ are Scott-continuous in both arguments. A continuous Kegelspitze is a Kegelspitze that is a domain in the equipped order.

- **Kegelspitzen** [3] are dcpo's equipped with a convex structure.
- **Example:** The real unit interval $[0, 1]$ is a continuous Kegelspitze.
- **Example:** For every dcpo $X$, both $\mathcal{M}X$ and $\forall X$ are Kegelspitzen. If $X$ is a domain then $\mathcal{M}X = \forall X$ is a continuous Kegelspitze.

---

Kegelspitzen and EM-algebras

- **Theorem:** The Eilenberg-Moore category $\text{DOM}^\mathcal{M}$ of $\mathcal{M}$ over $\text{DOM}$ is isomorphic to the category of continuous Kegelspitzen and Scott-continuous linear maps.

- **Remark:** $\text{DOM}^\mathcal{M} = \text{DOM}^\mathcal{V}$ and this corrects an error in the thesis of Jones.

- In every Kegelspitze $K$, one can define the *subconvex sum*: for $a_i \in K$, $r_i \in [0, 1]$, with $\sum_{i=1}^{n} r_i \leq 1$, then $\sum_{i=1}^{n} r_i a_i$ is also in $K$ and this expression is Scott continuous in each $r_i$ and $a_i$.

- A countable convex sum may also be defined: given $a_i \in K$ and $r_i \in [0, 1]$, for $i \in I$, with $\sum_{i \in I} r_i \leq 1$, let $\sum_{i \in I} r_i a_i \overset{\text{def}}{=} \sup\{\sum_{j \in J} r_j a_j \mid J \subseteq I \text{ and } J \text{ is finite}\}$. 
The Kleisli Category of $\mathcal{M}$

- The Kleisli category $\text{DCPO}_\mathcal{M}$ of $\mathcal{M}$ over $\text{DCPO}$:
  - Inherits coproducts from $\text{DCPO}$.
  - Has a symmetric monoidal structure induced by the commutative monad $\mathcal{M}$.
  - Contains the structure of a Kleisli exponential, because $\text{DCPO}$ is a CCC.
  - Is enriched over Kegelspitzen; the Kleisli adjunction is $\text{DCPO}$-enriched.
  - Has sufficient structure to solve recursive domain equations.

- This means $\text{DCPO}_\mathcal{M}$ has sufficient structure for the semantics of probabilistic programming languages with discrete probabilistic choice.
Denotational Semantics for PFPC

• PFPC is a type system with: function types, pair types, sum types, recursive types and (induced) term recursion, discrete probabilistic choice.
  • No restrictions on admissible logical polarities when forming recursive types.
• Judgements $\Gamma \vdash M : A$ are interpreted as Scott-continuous $\llbracket M \rrbracket : \llbracket \Gamma \rrbracket \to M[\llbracket A \rrbracket]$.
• **Theorem:** The system PFPC may be interpreted in the Kleisli category $\text{DCPO}_M$. This interpretation is sound and strongly adequate:

$$\llbracket M \rrbracket = \sum_{M \xrightarrow{p} M'} p \llbracket M' \rrbracket \quad \quad \llbracket M \rrbracket = \sum_{V \in \text{Val}(M)} P(M \to^* V) \llbracket V \rrbracket.$$

• **Remark:** The same results hold *verbatim* when $M$ is replaced by $\mathcal{P}$ or $\mathcal{W}$.
• **Remark:** The interpretation of every closed term is a *discrete* valuation.
Conclusion and Future Work

• Three commutative submonads of $\mathcal{V} : \text{DCPO} \to \text{DCPO}$.
• Characterised the EM-algebras of our monads (and $\mathcal{V}$) on domains as exactly the continuous Kegelspitzen.
• Sound and strongly adequate denotational semantics for PFPC.
• **Future Work:** *Continuous* probabilistic choice?
  • We recently discovered a fourth commutative submonad $\mathcal{Z} : \text{DCPO} \to \text{DCPO}$.
  • It is constructed using *algebraic* ideas, not topological ones.
  • $\mathcal{S}D \subseteq \mathcal{M}D \subseteq \mathcal{W}D \subseteq \mathcal{P}D \subseteq \mathcal{Z}D \subseteq \mathcal{V}D$ for each dcpo $D$.
  • $\mathcal{Z} = \mathcal{V}$ iff $\mathcal{V}$ is commutative (open problem for 32 years).
  • We believe $\mathcal{Z}$ could be suitable for continuous probabilistic choice (work-in-progress).
<table>
<thead>
<tr>
<th>Introduction</th>
<th>Commutative Monads of Valuations</th>
<th>EM-algebras and Kegelspitzen</th>
<th>Semantics</th>
<th>Conclusion</th>
</tr>
</thead>
</table>

Thank you for your attention!