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.

[1] Jones and Plotkin. *"A probabilistic powerdomain of evaluations."* **LICS 1989**.

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 $\sup A$.
    - **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 $f : D \to E$ between two dcpo's is *Scott-continuous* if it is monotone and preserves suprema of directed subsets.
- The category **DCPO** of dcpo's and Scott-continuous functions is *cartesian closed*, complete and cocomplete.
- The category **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 $\twoheaddownarrow y = \{x \in D \mid x \ll y\}$.
- A *basis* for a dcpo D is a subset B satisfying $\twoheaddownarrow x \cap B$ is directed and $x = \sup \twoheaddownarrow x \cap B$, 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 **DOM**.
- Domains may be thought of as very well-behaved dcpo's.
- **Problem:** The category **DOM** is *not* cartesian closed.

Background: Domain Theory (Scott Topology)

- The order on a dcpo X induces a canonical topology σX, called the *Scott-topology*.
- The *Scott topology* σD 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 ΣD.
- $f : X \to Y$ is Scott-continuous iff f is continuous w.r.t. ΣX and ΣY.

Background: Probability and Recursion

- How to talk about recursion *and* probability?
- Why not just take $\mathrm{Meas}(X)$, the set of subprobability measures on the Borel σ-algebra induced by the Scott-topology of a dcpo X?
- Because it is unclear how to extend the assignment $\mathrm{Meas}(-)$ 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 $\mathbf{0}$.
  - **Example:** For $x \in X$, δ_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.

[1] Jones and Plotkin. *"A probabilistic powerdomain of evaluations."* **LICS 1989**.

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* ν by:

$$\int_{x\in X} f(x) d\nu \stackrel{\text{def}}{=} \int_0^1 \nu(f^{-1}((t,1]))dt.$$

- 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?

Jones. *Probabilistic non-determinism*. **PhD Thesis, University of Edinburgh, 1990.**

Our approach

- How to construct a domain-theoretic model for probability and recursion:
  - such that we have a *commutative monad* of valuations; and
  - such that this monad is taken over a *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 *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 $\mathbf{DOM}^{\mathcal{M}} = \mathbf{DOM}^{\mathcal{V}}$ is isomorphic to the category of continuous Kegelspitzen [3];
  - our constructions use *topological methods* and we construct *two additional* such monads with all of the above properties.

[3] Keimel and Plotkin. *Mixed powerdomains for probability and nondeterminism*. **LMCS**, 2017.

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 $\sum_{i=1}^{n} r_i \nu_i \stackrel{\text{def}}{=} \lambda U. \sum_{i=1}^{n} r_i \nu_i(U)$ 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 $\mathcal{S}X$.
- $\mathcal{S}X \subseteq \mathcal{V}X$, but $\mathcal{S}X$ 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 **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 $\mathcal{S}D$ in $\mathcal{V}D$.
- This view is a lot more useful for establishing the required proofs.
- We obtain *two additional* commutative monads by taking suitable completions of $\mathcal{S}D$ in $\mathcal{V}D$.

$\mathbb{K}$-categories, Completions and Commutative Monads

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

Definition of Kegelspitzen

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

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:

$$a +_1 b = a; \qquad a +_r b = b +_{1-r} a; \qquad a +_r a = a;$$
$$(a +_p b) +_r c = a +_{pr} (b +_{\frac{r-pr}{1-pr}} c) \ \text{ provided } r, p < 1.$$

Definition

A *pointed barycentric algebra* is a barycentric algebra A with a distinguished element $\perp$. For $a \in A$ and $r \in [0,1]$, we define $r \cdot a \stackrel{\text{def}}{=} a +_r \perp$. A map $f \colon A \to B$ between pointed barycentric algebras is called *linear* if $f(\perp_A) = \perp_B$ 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 $(a, b) \mapsto a +_r b \colon K \times K \to K$ and scalar multiplication $(r, a) \mapsto r \cdot a \colon [0, 1] \times K \to K$ 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 $\mathcal{V}X$ are Kegelspitzen. If X is a domain then $\mathcal{M}X = \mathcal{V}X$ is a continuous Kegelspitze.

[3] Keimel and Plotkin. *Mixed powerdomains for probability and nondeterminism*. **LMCS**, 2017.

Kegelspitzen and EM-algebras

- **Theorem:** The Eilenberg-Moore category $\mathbf{DOM}^{\mathcal{M}}$ of $\mathcal{M}$ over $\mathbf{DOM}$ is isomorphic to the category of continuous Kegelspitzen and Scott-continuous linear maps.
- **Remark:** $\mathbf{DOM}^{\mathcal{M}} = \mathbf{DOM}^{\mathcal{V}}$ 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 $\sum_{i \in I} r_i a_i \stackrel{\text{def}}{=} \sup\{\sum_{j \in J} r_j a_j \mid J \subseteq I \text{ and } J \text{ is finite}\}$.

The Kleisli Category of $\mathcal{M}$

- The Kleisli category $\mathbf{DCPO}_{\mathcal{M}}$ of $\mathcal{M}$ over **DCPO**:
  - Inherits coproducts from **DCPO**.
  - Has a symmetric monoidal structure induced by the commutative monad $\mathcal{M}$.
  - Contains the structure of a Kleisli exponential, because **DCPO** is a CCC.
  - Is enriched over Kegelspitzen; the Kleisli adjunction is **DCPO**-enriched.
  - Has sufficient structure to solve recursive domain equations.
- This means $\mathbf{DCPO}_{\mathcal{M}}$ 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 $[\![M]\!] : [\![\Gamma]\!] \to \mathcal{M}[\![A]\!]$.
- **Theorem:** The system PFPC may be interpreted in the Kleisli category $\mathbf{DCPO}_{\mathcal{M}}$. This interpretation is sound and strongly adequate:

$$[\![M]\!] = \sum_{M \xrightarrow{p} M'} p[\![M']\!] \qquad\qquad [\![M]\!] = \sum_{V \in \mathrm{Val}(M)} P(M \to_* V)[\![V]\!].$$

- **Remark:** The same results hold *verbatim* when $\mathcal{M}$ 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 $\mathcal{V} : \mathbf{DCPO} \to \mathbf{DCPO}$.
- Characterised the EM-algebras of our monads (and $\mathcal{V}$) on domains as exactly the continuous Kegelspitzen.
- Sound and strongly adequate denotational semantics for $\mathtt{PFPC}$.
- **Future Work:** *Continuous* probabilistic choice?
  - We recently discovered a fourth commutative submonad $\mathcal{Z} : \mathbf{DCPO} \to \mathbf{DCPO}$.
  - 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).

Thank you for your attention!