Differential programming, *probably*

This talk references work that is part of a few different collaborations, including with Robin Cockett, Geoff Cruttwell, JS Lemay, Ben MacAdam, and Dorette Pronk.

A Puzzle from Tversky and Kahneman

"Linda is 31 years old, single, outspoken, and very bright. She did her PhD in behavioural neuroscience. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in climate change demonstrations."

- Which of the following is more likely?

1) Linda is the lead researcher of a neuroscience research firm.

2) Linda is the lead researcher of a neuroscience research firm who, in her free time, actively contributes to anti-discrimination programs.

A Puzzle from Tversky and Kahneman

"Linda is 31 years old, single, outspoken, and very bright. She did her PhD in behavioural neuroscience. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in climate change demonstrations."

- Which of the following is more likely?

1) Linda is the lead researcher of a neuroscience research firm.

2) Linda is the lead researcher of a neuroscience research firm who, in her free time, actively contributes to anti-discrimination programs.

A Puzzle from Tversky and Kahneman

“Linda is 31 years old, single, outspoken, … her PhD in behavioural neuroscience. As a student, … ed with issues of discrimination and social justice, … in climate change demonstrations.”

Linda can’t be ***just*** a researcher!!!

- Which of the following is more likely?

1) Linda is the lead researcher of a neuroscience research firm.

2) Linda is the lead researcher of a neuroscience research firm who, in her free time, actively contributes to anti-discrimination programs.

Causal bias – a case for probabilistic reasoning

- Tversky and Kahneman’s work on rational actors – brains (NNs) aren’t inherently rational
- Humans rank ***Causality*** as more important than ***Statistics***
  - E.g. Causal relationships and input-output type relationships
- Explicit probabilistic reasoning improves reliability of conclusions
  - Even instinctively or without calculations!

Causal bias – a case for probabilistic reasoning

- Tversky and Kahneman's work on rational actors – brains (NNs) aren't inherently rational
- Humans rank ***Causality*** as more important than ***Statistics***
  - E.g. Causal relationships and input-output type relationships
- Explicit probabilistic reasoning improves reliability of conclusions
  - Even instinctively!

Aiming to be probably correct is more reliable than aiming to be correct (PAC)

Simpson’s paradox – a case for causal reasoning

- Two treatments for kidney stones: Treatment A and Treatment B.
- Treatment A is 78% effective
- Treatment B is 83% effective
- The difference is statistically significant
- Which is the better treatment?

https://www.bmj.com/content/292/6524/879

Simpson’s paradox – a case for causal reasoning

- Treatment A is 78% effective
- Treatment B is 83% effective
- BUT

	Treatment A	Treatment B
Small stones	93% effective	87% effective
Large stones	73% effective	69% effective

https://www.bmj.com/content/292/6524/879

Simpson's paradox – a case for causal reasoning

- The cure to Simpson's paradox is causality!
- The problem can be formalized in terms of a Bayesian network
  - Simpson's paradox can be tested for algorithmically
- If identified one can form a 'story' of what is most important

- Explicit causal reasoning can improve reliability of conclusions
  - And makes avoiding Simpson's paradox solvable

https://ftp.cs.ucla.edu/pub/stat_ser/r414.pdf

Probabilities in (A)NNs / ML / differential programming

- Differential programming is winning:
  - Software 2.0: use the derivative everywhere to learn stuff – Tesla's AI director [1]
- Probabilistic programming is also winning!
- Basic algorithms in diff. prog. use probability
  - e.g. Stochastic gradient descent
  - e.g. Prob. neural networks are more accurate
- Basic algorithms in probabilistic prog. use the derivative
  - e.g. Variational method [2] – universally approximates prob. distributions

[1] https://karpathy.medium.com/software-2-0-a64152b37c35 [2]:https://pyro.ai/

Main goal

- Categorical models for differential programming with higher-order functions and recursion (also polymorphism)
  - A DiLL pickle

- Extension to probabilistic setting via effect algebras

- Pure locally presentable view of probabilistic + differential
  - Also gives a new way to obtain models of diff. prog. with function spaces and recursion

Another day

A simple differential programming language [1]

$$T := x \in \mathsf{Var} \,|\, r \in \mathbb{R} \,|\, \sum_{i=1}^{n} T \,|\, f(T, \ldots, T)$$

$$|\, (T, \ldots, T) \,|\, \mathsf{let}(x_1, \ldots, x_n) = T \mathsf{in} T$$

$$|\, \mathsf{if}\; T \;\mathsf{then}\; T \;\mathsf{else}\; T$$

$$|\, \frac{\partial^R T}{\partial x}(T) \cdot T$$

$$|\, \mathsf{letrec} f(x) = T \mathsf{in} T$$

[1] https://dl.acm.org/doi/10.1145/3371106

A partial sidebar

- We use restriction categories to encode partiality
- However, we only need partiality to determine the *objects* of the convex and effect module categories
  - The morphisms are necessarily total, and hence the categories will be ordinary categories
- The required knowledge of restriction categories is thus minimized
  - Much like the situation of relative category theory used in [1]

[1] Categorical treatment of pre- and post-conditions

Restriction categories overview

Definition. *A **restriction category** $\mathbb{X}$ is a category such that for each morphism $A \xrightarrow{f} B$ there is an idempotent $A \xrightarrow{\overline{f}} A$ satisfying certain axioms.*

Important: every restriction category $\mathbb{X}$ comes with a subcategory of total maps denoted $\mathsf{Total}(\mathbb{X})$. Also every restriction category is partial order enriched.

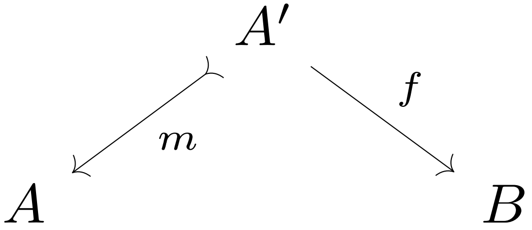
Definition. *A **cartesian restriction category** is a restriction category $\mathbb{X}$ with a monoidal structure $\mathbb{X} \times \mathbb{X} \xrightarrow{\text{-}\times\text{-}} \mathbb{X}$ and $1 \xrightarrow{1} \mathbb{X}$ that satisfies axioms making it behave like a partial product (natural contraction, non-natural weakening).*

Important: If $\mathbb{X}$ is a cartesian restriction category then $\mathsf{Total}(\mathbb{X})$ is a cartesian category.

[1] Restriction categories basics [2] Cartesian restriction categories and restriction limits

Restriction categories of partial maps

Lemma. *If* $\mathbb{X}$ *is a category with a display system of monics* $\mathcal{M}$ *that is closed to composition and isomorphism, then* $\mathsf{Par}(\mathbb{X}, \mathcal{M})$*, the category of equivalence classes of spans whose domain leg is in* $\mathcal{M}$*, is a restriction category.*



Gradients: The reverse derivative

Definition. *A **reverse differential restriction category** is a cartesian left additive restriction category with an operation on maps*

$$\frac{A \xrightarrow{f} B}{A \times B \xrightarrow[R[f]]{} A}$$

that satisfies certain axioms.

The are connected to differential restriction categories by the following

Proposition. *Reverse differential restriction categories are exactly differential restriction categories such that maps $A \times B \to C$ that linear and total in B have an involutive dagger operation sending them to maps $A \times C \to B$ and that plays nicely with reindexing.*

Semantics for Simple DPL

- We do not require all the axioms of a CRDC to model Simple DPL. In fact, we may drop the 2nd, 6th and 7th axioms. This is called a *basic CRDC*. We do however require that the category has joins of families compatible maps (this includes joins of directed sets).

Theorem. *Any reverse differential restriction category with joins models the operational semantics of Simple DPL.*

Intermediate DPL

- Add higher-order functions!

$$T := x \in \mathsf{Var} \mid r \in \mathbb{R} \mid \sum_{i=1}^{n} T \mid f(T, \ldots, T)$$

$$\mid (T, \ldots, T) \mid \mathsf{let}(x_1, \ldots, x_n) = T \mathsf{in} T$$

$$\mid \mathsf{if}\, T \,\mathsf{then}\, T \,\mathsf{else}\, T$$

$$\mid \frac{\partial T}{\partial x}(T) \cdot T$$

$$\mid \mathsf{letrec} f(x) = T \mathsf{in} T$$

$$\mid T(T) \mid \lambda x.T$$

Note we dropped the R in the derivative. Here we use the forward derivative.

Models of reverse differentiation with function spaces are subtle to obtain, and there are no known models with additionally recursion.

Semantics of Intermediate DPL

- The main source of models comes from synthetic differential geometry at the moment.
- For function spaces of partial maps: note that hom(X,1) is the object of ‘open’ subsets of X – and this is not a vector space type of thing.
- To handle these more general types of spaces we move to the setting of tangent categories. Note, also that we didn’t require some of the axioms for a differential restriction category.
  - This allows us to work in a slight generalization of a tangent category known as a cartesian proto-tangent category.

Proto-tangent categories

- Recall from the talk on Monday i.e. Garner [1] that tangent categories are exactly categories with an action by Weil algebras where the action preserves connected limits of Weil algebras.

Definition. *A **proto-tangent** category is a category with an action by* Weil_1. *It is **cartesian** when* $\mathbb{X}$ *has finite products and every action preserves products.*

Cartesian proto-tangent categories correspond to basic cartesian differential categories.

[1] Embedding theorem for tangent categories

Formally etale partiality

Definition. *A **formally etale subobject** $m : A \rightarrowtail B$ in a proto-tangent category is a subobject such that the following naturality square is a pullback*

$$\begin{array}{ccc} TA & \xrightarrow{Tm} & TB \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ A & \xrightarrow[m]{} & B \end{array}$$

Formally etale subobjects were introduced in Manifolds in formal differential geometry to build manifolds internally in SDG.

Formally etale partiality

Lemma. *A display system of formal etale monics* $\mathcal{M}$ *has the property that* $\mathsf{Par}(\mathbb{X}, \mathcal{M})$ *inherits the* Weil_1 *action. And*

1. *If* $\mathbb{X}$ *is in fact a tangent category then* $\mathsf{Par}(\mathbb{X}, \mathcal{M})$ *is a tangent restriction category.*

2. *If* $\mathbb{X}$ *is a cartesian tangent category then* $\mathsf{Par}(\mathbb{X}, \mathcal{M})$ *is a cartesian restriction category.*

Formal etaleness is necessary for the partial map category to inherit the proto-tangent structure.

Formally etale partiality

Definition. *A cartesian proto-tangent category is **coherently closed** when each cotensorial strength*

$$T_U([A,B]) \to^{\psi} [A, T_U(B)]$$

induced by strength of the proto-tangent functor, has ψ an isomorphism.

Definition. *A coherently closed proto-tangent category is called formally etale classified when there is a monad P such that*

$$\mathsf{Par}(\mathbb{X}, \mathcal{M})(A,B) \simeq \mathbb{X}(A, PB)$$

This implies that $P(1)$ is a formal etale subojbect classifier.

A model of intermediate DPL

Theorem. *If* $\mathbb{X}$ *is a coherently closed, formally etale classified, proto-tangent category, such that* $\mathsf{Par}(\mathbb{X}, \mathcal{M})$ *has joins, then* $\mathsf{Par}(\mathbb{X}, \mathcal{M})$ *is a model for the operational semantics of intermediate DPL.*

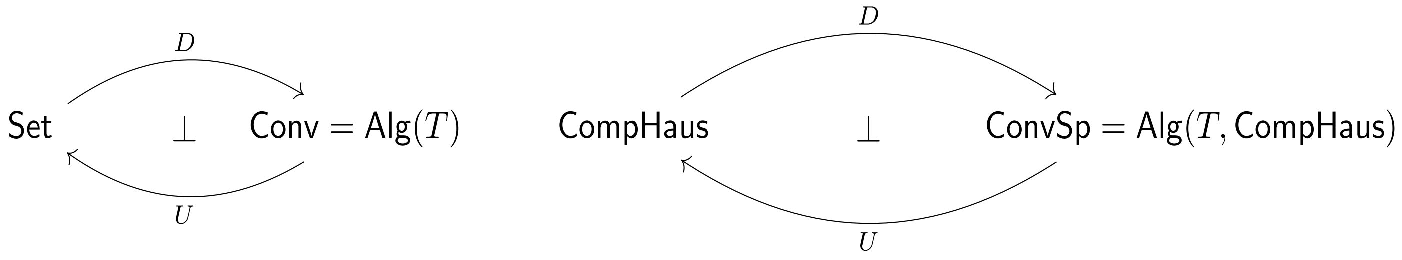
Models exist!

Proposition. *Suppose that $\mathcal{E}$ is a well-adapted, Grothendieck topos model of SDG, that has the amazing right adjoint property: $TA = D \Rightarrow A$ is both a right and left adjoint. Then $\mathcal{E}$ is a coherently closed, formally etale classified proto-tangent category.*

Extension to probabilistic setting via effects

Idea

- Meng [1] defines convex categories abstractly as algebras of a certain theory in Sets and Compact Hausdorff spaces



  - Then uses the induced *probability monad* to model stochastic programming as an extension of ordinary programming

[1] X. Meng. PhD thesis. Copy available from Nlab MengThesis

Idea

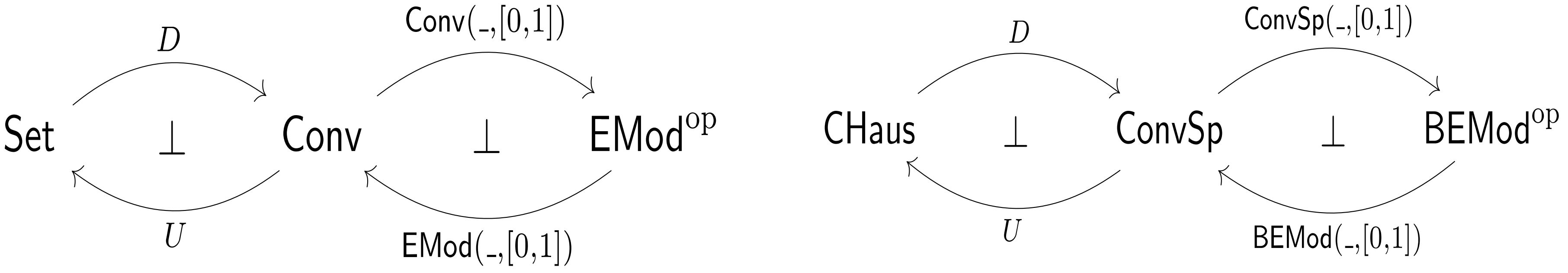
- Lucyshyn-Wright [2] refined the approach of Meng, and showed how to define convexity for any ordered commutative ring in a cartesian closed category

$$\mathsf{C} \underset{U}{\overset{D}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathsf{Aff}(\mathsf{C})$$

- Enables interpreting probabilistic extensions more generally

Idea

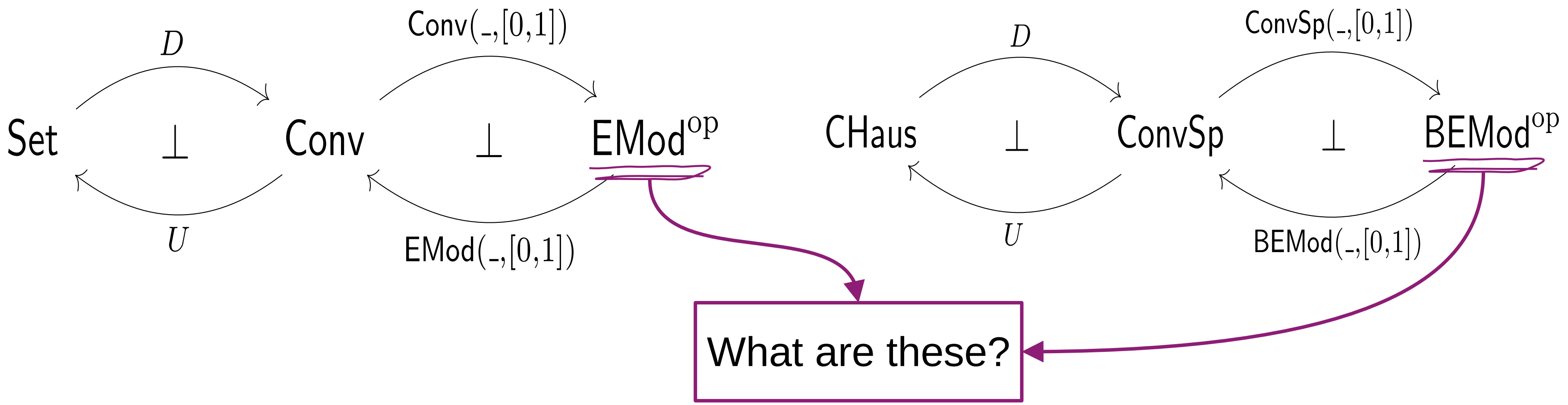
- However, often in practice one needs the expectation monad too.
- This has been developed in a series of papers by Jacobs (e.g. [1,2] for an overview)



[1] Probabilities, distribution monads, and convex categories [2] The expectation monad in quantum foundations

Idea

- However, often in practice one needs the expectation monad too [3].
- This has been developed in a series of papers by Jacobs as the monad from the composite adjunctions (e.g. [1,2] for an overview)



[1] Probabilities, distribution monads, and convex categories [2] The expectation monad in quantum foundations [3] The Hakaru language

Effect modules

- The categories EMod and BEMod are effect modules and Banach effect modules respectively.

- These don't fit the known algebraic story, as effect modules are underlied by partial monoids
  - Can't simply take Aff(C) of Lucyshyn-Wright and determine the appropriate effect modules.

- However, they do fit a partialized version of a theorem in Lucyshyn-Wright.

Goal refined

- Redevelop the notion of convexity purely in terms of effect algebras and modules in an arbitrary *restriction category*.
- Restriction categories are abstract categories of partial maps.
- Then we will tweak this development to work with differentiation.

Partial commutative monoids

Definition. *A **partial commutative monoid** (PCM) in a cartesian restriction category* $\mathbb{X}$ *is a monoid object* $(A, \oslash, \zeta)$ *in the monoidal category underlying* $\mathbb{X}$.

Example. *A partial commutative monoid is classically defined to be a monoid in the cartesian restriction category of sets and partial functions.*

Example. *The unit interval* $[0, 1]$ *under* $+$ *with unit* 0 *is a partial commutative monoid.*

Pre-Effect algebras I

An effect algebra is usually defined as a PCM with a map $A \xrightarrow{(\_)^{\perp}} A$ such that

1. $x^{\perp}$ is unique such that $x + x^{\perp} = 0^{\perp}$;
2. If $x + 0^{\perp}$ is defined then $x = 0$.

Pre-Effect algebras I

An effect algebra is usually defined as a PCM with a map $A \xrightarrow{(\text{-})^\perp} A$ such that

1. $x^\perp$ is unique such that $x + x^\perp = 0^\perp$;
2. If $x + 0^\perp$ is defined then $x = 0$.

Not algebraic!

Pre-Effect algebras I

An effect algebra is usually defined as a PCM with a map $A \xrightarrow{(\_)^{\perp}} A$ such that

1. $x^{\perp}$ is unique such that $x + x^{\perp} = 0^{\perp}$;
2. If $x + 0^{\perp}$ is defined then $x = 0$.

Lemma. *If A is an effect algebra, then there is a partial difference defined by*

$$a - b := (a^{\perp} \ovee b)^{\perp}$$

If $a + b$ is defined then

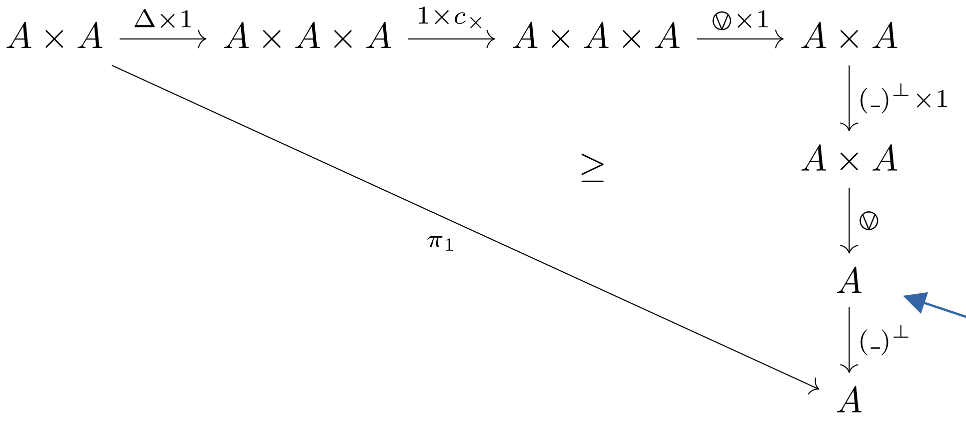
$$(a + b) - a = b$$

Moreover, the uniquness of $b^{\perp}$ is equivalent to this property of the difference together with the involutiveness of $(\_)^{\perp}$.

Pre-effect algebras II

Definition. *A **pre-effect algebra** in a cartesian restriction category is a PCM $(A, \oslash, \zeta)$ with a map $A \xrightarrow{(\_)^\perp} A$ such that*

1. *$(\_)^\perp$ is an involution;*
2. *$x + x^\perp = \zeta^\perp$;*
3. *We have*



4. *Name the map clockwise around the diagram in 3 h. We have $\overline{h} = \overline{\oslash}$.*

$(x + 0^\perp)\downarrow$ implies $x = 0$ forces partial order ... conflicts with application to SDG

Define $1 = 0^\perp$

Encodes the lemma that determines uniqueness

Pre-Effect Algebras

- $[0,1]$ $\quad\quad$ $\mathrm{x}^{\perp} = 1 - x$
- The lattice of closed subsets of a Hilbert space ("quantum effects") with complement and join
- R $\quad$ $\mathrm{x}^{\perp} = 1 - x$ $\quad\quad$ $(\mathrm{x}{+}\mathrm{y})\downarrow \leftrightarrow x, y, x + y \in [0, 1]$
- R as a total ea

$$\mathrm{x}^{\perp} = 1 - x \qquad \mathrm{x}\ `\sigma` y := \begin{cases} 0 & x + y \leq 0 \\ 1 & x + y \geq 1 \\ x + y & \text{else} \end{cases}$$

Quick analogy

- The free commutative monoid on a set M can be described as

$$\mathrm{B(A)} := \{ \mathrm{f} : \mathrm{A} \to \mathbb{N} \mid \mathsf{supp}(f) \text{ is finite} \}$$

- This can also be written as finite formal sums of A

Discrete probability distribution monad

- The discrete probability distributions on a set can be described as

$$D(A) := \{f : A \to [0,1] \mid \mathsf{supp}(f) \text{ is finite }, \sum_{x \in A} f(x) = 1\}$$

$$\simeq \left\{ \sum_i a_i r_i \mid \sum_i r_i = 1 \right\}$$

- There is a continuous version too – one takes Radon measures on compact Hausdorff spaces

The prob. dist. monad of an eff. alg.

- The discrete distribution valued in an eff. alg. E is described as

$$\mathrm{D}_E(A) := \{f : A \to E \,|\, \mathsf{supp}(f) \text{ is finite },\ \sum_{x \in \mathsf{supp}(f)} f(x) = 1\}$$

- This implies that the support is summable, or pairwise defined

Fuzzy predicates and duals of prob. monads

- For the discrete probability monad, we have that its algebras are convex sets, and hence by freeness [1,2]

$$\mathsf{Conv}(D(X), [0,1]) \simeq \mathsf{Alg}(D)(D(X), [0,1]) \simeq \mathsf{Set}(X, [0,1])$$

- Similarly, for the continuous probability monad, its algebras are convex spaces, and again by freeness [1,2]

$$\mathsf{ConvSp}(D(X), [0,1]) \simeq \mathsf{CompHaus}(X, [0,1])$$

where convex spaces are compact, convex subspaces of LCTVS with affine maps

[1] https://link.springer.com/article/10.1007/s11083-016-9404-1 [2] https://www.sciencedirect.com/science/article/pii/S0304397511002817

Fuzzy predicates and duals of prob. monads

- For the discrete probability monad, we have that its algebras are convex sets, and hence by freeness [1,2]

$$\mathsf{Conv}(D(X), [0,1]) \simeq \mathsf{Alg}(D)(D(X), [0,1]) \simeq \mathsf{Set}(X, [0,1])$$

- Similarly, for the continuous probability monad, its alge convex spaces, and again by freeness [1,2]

$$\mathsf{ConvSp}(D(X), [0,1]) \simeq \mathsf{CompHaus}(X, [0,1])$$

Fuzzy predicates

where convex spaces are compact, convex subspaces of LCTVS with affine maps

[1] https://link.springer.com/article/10.1007/s11083-016-9404-1 [2] https://www.sciencedirect.com/science/article/pii/S0304397511002817

Effect reflexivity determines distributions

Lemma. *Suppose that $D(X)$ is effect reflexive i.e.* $(D(X) \multimap I) \multimap I \simeq D(X)$
Then
$$D(X) \simeq [X, I] \multimap I$$

where

$A \multimap B$ is the internal hom in $\mathsf{Alg}(D)$

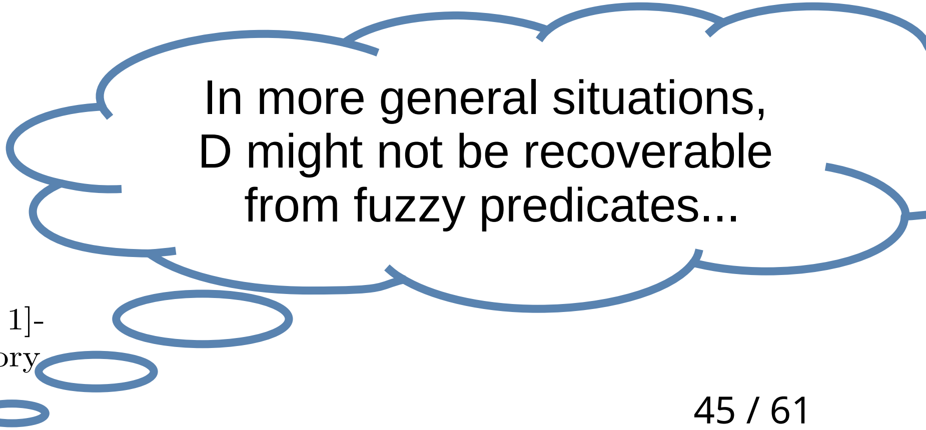
$[A, B]$ is the internal hom in the base category (set or comp. haus).

When D(X) is effect reflexive, the probability distributions on it can be recovered as the 'dual' of fuzzy predicates.

Creating the convex category abstractly

Consider

$$\mathsf{CompHaus} \underset{U}{\overset{D}{\rightleftarrows}} \mathsf{ConvSp} \quad (D \dashv U) \qquad\qquad \mathsf{Set} \underset{U}{\overset{D}{\rightleftarrows}} \mathsf{Conv} \quad (D \dashv U)$$



We want to generally get the convex category as a kind of enriched $[0, 1]$-affine space. I.e. want a canonical way to generate D and the convex category over any base.

The category of effect modules

- A morphism of effect modules is classically taken to be a total function of the underlying sets such that certain diagrams commute.
- This has a simple characterization from restriction categories.

RCat

The 2-category RCat has

0-cells: Restriction categories

1-cells: Restriction preserving functors

2-cells: Natural transformations that are total at each component

The category of restriction functors and componentwise total transformations between restriction categories has only a trivial restriction structure. I.e. it's a mere category.

The category of effect algebras

Lemma. *The category of pre-effect algebras in* $\mathbb{X}$ *is exactly the category with objects:*

$$T_{\mathsf{eff}} \to \mathbb{X} \qquad \textit{product preserving}$$

and morphisms: 2-cells in RCat.

where T_{eff} is the generic pre-effect algebra.

Commutative pre-effect monoids

Definition. *A **commutative pre-effect monoid** in a restriction category $\mathbb{X}$ is a commutative monoid in the category of effect pre-algebras.*

Note, this means that the multiplication is a total operation.

Example. *The unit interval $[0,1]$ is a commutative pre-effect monoid with multiplication as in $\mathbb{R}$.*

Effect modules over a commutative effect monoid

Definition. *A pre-effect module A (in a cartesian restriction category) over a commutative pre-effect monoid E is a pre-effect algebra A with a total action*

$$A \times E \stackrel{\cdot}{\to} A$$

such that the usual axioms of a module hold.

Lemma. *If $\mathbb{X}$ is a cartesian restriction category such that $\mathsf{Total}(\mathbb{X})$ is cartesian closed and hence self-enriched, then a commutative effect monoid E in $\mathbb{X}$ gives rise to an $\mathbb{X}$-theory T_E, and the category of effect modules over E is the category of producct preserving functors $T_E \to \mathbb{X}$.*

Proof. Follows the same lines as in [1] for R-modules. □

[1] Functional distribution monads in functional-analytic contexts

Pre-effect convex spaces

Lemma. *For any commutative effect monoid E, there is a sub theory $T^{\mathsf{aff}}_{\mathsf{eff}}$ whose arrows are matrices whose rows are defined and sum to 1.*

Lemma. *For any commutative pre-effect monoid in $\mathbb{X}$ with $\mathsf{Total}(\mathbb{X})$ cartesian closed, we have that $\mathsf{Mod}(\mathsf{T}_{\mathsf{eff}}, \mathbb{X})$ and $\mathsf{Mod}(T^{\mathsf{aff}}_{\mathsf{eff}}, \mathbb{X})$ are monoidal closed, and we have an adjunction*

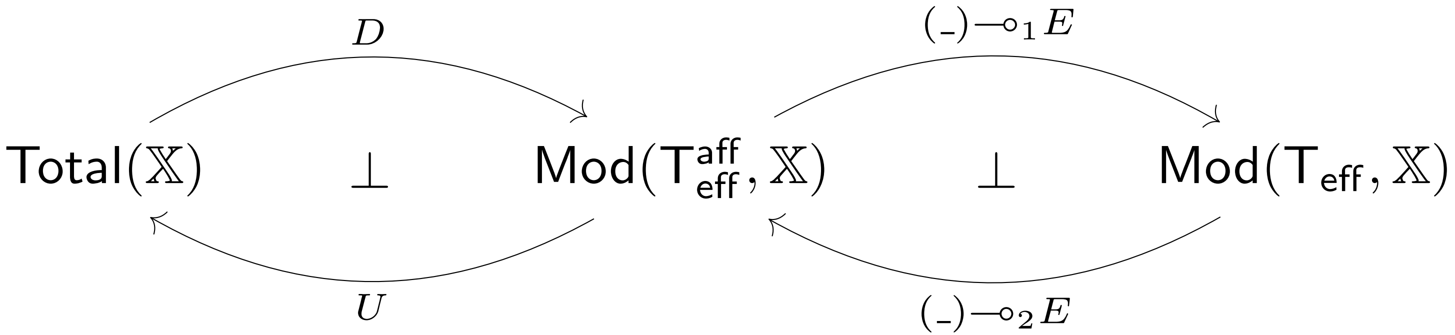
$$\mathsf{Mod}(\mathsf{T}^{\mathsf{aff}}_{\mathsf{eff}}, \mathbb{X}) \underset{(\_)\multimap_2 E}{\overset{(\_)\multimap_1 E}{\rightleftarrows}} \mathsf{Mod}(\mathsf{T}_{\mathsf{eff}}, \mathbb{X}) \qquad \bot$$

$\multimap_1$ is the internal hom in $\mathsf{Mod}(T^{\mathsf{aff}}_{\mathsf{eff}}, \mathbb{X})$

$\multimap_2$ is the internal hom in $\mathsf{Mod}(T_{\mathsf{eff}}, \mathbb{X})$

Free pre-effect convex space

Lemma. *If* $\mathbb{X}$ *is such that* $\mathsf{Total}(\mathbb{X})$ *is cocomplete and cartesian closed then one can form the free* E *pre-effect affine space. We write the left adjoint as* D*. Then we have the following adjunctions:*



$\multimap_1$ is the internal hom in $\mathsf{Mod}(T_{\mathsf{eff}}^{\mathsf{aff}}, \mathbb{X})$

$\multimap_2$ is the internal hom in $\mathsf{Mod}(T_{\mathsf{eff}}, \mathbb{X})$

Summary so far

- We have shown how to extend Jacobs' framework for probabilistic programming using effect algebras to obtain models in cartesian restriction categories.

- In the remainder of the talk, we will show how this technology can be applied to SDG, to extend a model of differential programming to the probabilistic setting too.

Recall the well-adapted model of diff. prog.

- In the well-adapted model of differential programming used earlier, we assumed a well-adapted model of synthetic differential geometry that satisfies the amazing right adjoint property.

- We will now make an additional assumption, that the model satisfies the order axiom.

Order axiom in SDG

Definition. *A model of SDG $\mathcal{E}$ satisfies the order axiom, when $\mathbb{R}$ is a preoder – that is has a binary relation that is reflexive and transitive, and where the order is compatible with the ring structure*

1. $x \leq y$ *implies* $x + z \leq y + z$
2. $x \leq y$ *and* $0 \leq t$ *implies* $xt \leq yt$
3. $0 \leq 1$

And finally that nilpotents are small; i.e. if $d^k = 0$ *for some* k *then both* $0 \leq d$ *and* $d \leq 0$*.*

With the order axiom define an interval $[a, b] := \{x \mid a \leq x \leq b\} \subseteq \mathbb{R}$.

Lemma. *The interval* $[0, 0]$ *contains the subobject* D_∞ *which is a formal etale subobject of* $\mathbb{R}$*.*

Microlinear intervals

- Kock [1] in [III.11] the following is shown.

Proposition. *There is a well-adapted model of SDG satisfying the amazing right adjoint that also satisfies the order axiom that additionally has a model of the upper half-plane H. As a subobject of $\mathbb{R}$, $H \subseteq \mathbb{R}$ is a formally-etale subobject.*

- If follows from [2] that

Corollary. *$H \subseteq R$ is microlinear.*

[1] Synthetic differential geometry [2] Manifolds in formal differential geometry

Microlinear Intervals

In this model of SDG, $-H := \{-a \,|\, a \in H\}$ and $-H+1 := \{-a+1 \,|\, a \in H\}$ are isomorphic to H and hence formal etale subobjects. The intersection of formally etale subobjects is formally etale. Hence

Lemma. *In this model of SDG we have*

$$[0,1] := H \cap (-H+1)$$

is a formally etale subobject of $\mathbb{R}$ *and hence microlinear.*

Microlinear pre-effect monoid

- The addition the real line can be pulled back to give

Lemma. *The span induced by the pullback gives a partial commutative monoid structure on* $[0,1]$.

$$
\begin{array}{ccc}
X & \rightarrowtail & [0,1] \times [0,1] \\
 & & \downarrow \\
\ovee \downarrow & & R \times R \\
 & & \downarrow + \\
[0,1] & \rightarrowtail & R
\end{array}
$$

This extends to a commutative pre-effect monoid with multiplication as in R *and* $x^{\perp} := 1 - x$.

Model of differential probabilistic programming

Let $\mathcal{E}$ denote the model of SDG mentioned above, and let $\widetilde{\mathcal{E}}$ the category of partial maps with respect to formal-etale monics. Then from the above we have

Lemma. *There is the following series of adjunctions*

$$\mathcal{E} \underset{U}{\overset{D}{\rightleftarrows}} \; \bot \; \mathsf{Mod}(\mathsf{T}^{\mathsf{aff}}_{\mathsf{eff}}, \widetilde{\mathcal{E}}) \underset{(\_)\multimap_2 E}{\overset{(\_)\multimap_1 E}{\rightleftarrows}} \; \bot \; \mathsf{Mod}(\mathsf{T}_{\mathsf{eff}},, \widetilde{\mathcal{E}})$$

Future work

- There are straightforward notions of KL pre-effect modules and convex spaces.

  For KL modules $V \to T_0(V)$ is an isomorphism. For KL pre-effect modules, it's natural to consider $V \to T_0(V)$ to be formally etale. This then induces a partial isomorphism $TV \to V \times V$ with a total inverse.

- Can the theory of microlinear enriched algebraic theories be extended so that we can see KL pre-effect modules and convex spaces as also being categories of models?
  - This would allow creating a free KL effect affine space, and having a microlinear probability monad!

Future work

- This story seems close to the story of storage and free differential objects in a tangent category.

- Note that the distribution and expectation monads have the relationship

$$\mathrm{D}(\mathrm{A}) \rightarrowtail (A \Rightarrow [0,1]) \multimap [0,1]$$

  This feels like the case of storage.

- KL effect modules have a notion of differentiation that is similar to KL modules. Can categories of KL effect modules be axiomatized directly, and tied to the distribution monad in a way similar to storage CDCs?