Compositional Imprecise Probability

Early Announcement

Jack Liell-Cock [1]   Sam Staton [1]

[1]University of Oxford

MFPS2024
21st June, 2024

Preprint

arXiv > cs > arXiv:2405.09391

Search... All fields Search
Help | Advanced Search

Computer Science > Programming Languages

[Submitted on 15 May 2024]

Compositional imprecise probability

Jack Liell-Cock, Sam Staton

Imprecise probability is concerned with uncertainty about which probability distributions to use. It has applications in robust statistics and elsewhere. Imprecise probability can be modelled in various ways, including by convex sets of probability distributions.
We look at programming language models for imprecise probability. Our desiderata are that we would like our model to support all kinds of composition, categorical and monoidal, in other words, guided by dataflow diagrams. Another equivalent perspective is that we would like a model of synthetic probability in the sense of Markov categories.
There is already a fairly popular monad-based approach to imprecise probability, but it is not fully compositional because the monad involved is not commutative, which means that we do not have a proper monoidal structure. In this work, we provide a new fully compositional account. The key idea is to name the non-deterministic choices. To manage the renamings and disjointness of names, we use graded monads. We show that the resulting compositional model is maximal. We relate with the earlier monad approach, showing that we obtain tighter bounds on the uncertainty.

Comments: Draft. Feedback welcome
Subjects: **Programming Languages (cs.PL)**; Logic in Computer Science (cs.LO); Category Theory (math.CT); Probability (math.PR)
Cite as: arXiv:2405.09391 **[cs.PL]**
(or arXiv:2405.09391v1 **[cs.PL]** for this version)
https://doi.org/10.48550/arXiv.2405.09391

Submission history

From: Sam Staton [view email]
[v1] Wed, 15 May 2024 14:49:13 UTC (31 KB)

Access Paper:

- View PDF
- HTML (experimental)
- TeX Source
- Other Formats

view license

Current browse context:
cs.PL
< prev | next >
new | recent | 2024-05
Change to browse by:
cs
cs.LO
math
math.CT
math.PR

References & Citations

- NASA ADS
- Google Scholar
- Semantic Scholar

Export BibTeX Citation

Bookmark

Overview

This work:

- Using graded perspective and naming Knightian choices we give a fully compositional account of Bernoulli and Knightian uncertainty together
- **Theorem 1:** This gives a refined bound on uncertainty
- **Theorem 2:** It is maximal among compositional accounts

Overview

This work:

- Using graded perspective and naming Knightian choices we give a fully compositional account of Bernoulli and Knightian uncertainty together
- **Theorem 1:** This gives a refined bound on uncertainty
- **Theorem 2:** It is maximal among compositional accounts

Not this work:

- There is a broader interest in combining non-determinism and probability [Dash and Staton 2021; Dash and Staton 2020; Dahlqvist et al. 2018; Keimel et al. 2017; Dash 2024; Kozen et al. 2023; Varacca et al. 2006; Jacobs 2021]
- Our focus is in the setting of imprecise probability

Contents

1 Imprecise Probability

2 Desiderata

3 The Problem

4 The Solution: Named Knightian Choices

5 A Graded Category

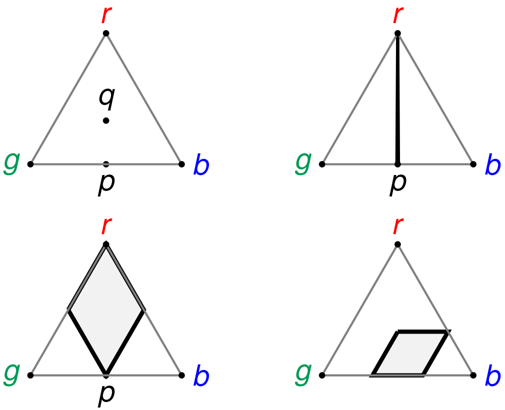
6 Results

Outline

1. Imprecise Probability
2. Desiderata
3. The Problem
4. The Solution: Named Knightian Choices
5. A Graded Category
6. Results

Imprecise Probability

- **Probability** = point in simplex
- **Imprecise probability** = convex set of points



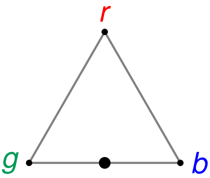
A First Language

Our prototype language for imprecise probability is a *first-order functional* language without recursion. We have:

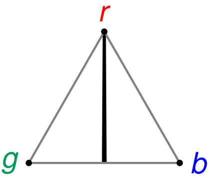
- If/then/else statements;
- Sequencing with immutable variable assignment;
- Two commands returning booleans:
  - ***bernoulli***: a fair Bernoulli choice;
  - ***knight***: a Knightian choice.

Examples

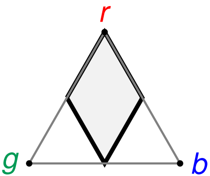
$z \leftarrow$ ***bernoulli*** ;
if z then return g else return b



$x \leftarrow$ ***knight*** ; $z \leftarrow$ ***bernoulli*** ;
if z then (if x then return r else return g)
else (if x then return r else return b)



$x \leftarrow$ ***knight*** ; $y \leftarrow$ ***knight*** ; $z \leftarrow$ ***bernoulli*** ;
if z then (if x then return r else return g)
else (if y then return r else return b)



Outline

1. Imprecise Probability
2. Desiderata
3. The Problem
4. The Solution: Named Knightian Choices
5. A Graded Category
6. Results

Desiderata

Desideratum (1)

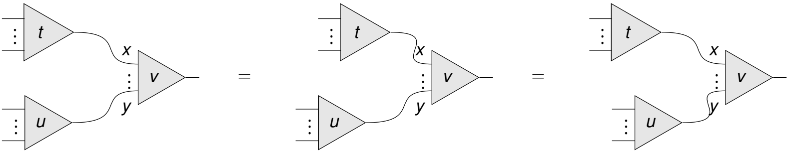
The language should be commutative:

$$x \leftarrow t \; ; \; y \leftarrow u \; ; \; v \quad = \quad y \leftarrow u \; ; \; x \leftarrow t \; ; \; v$$

for $x \notin fv(u)$ *and* $y \notin fv(t)$*; and affine:*

$$x \leftarrow t \; ; \; u \quad = \quad u$$

for $x \notin fv(u)$*.*



Desiderata

Desideratum (2)

Standard equational reasoning about if/then/else should apply:

$$\text{if } b \text{ then } (x \leftarrow t \ ; \ u) \text{ else } (x \leftarrow t \ ; \ v)$$
$$=$$
$$x \leftarrow t \ ; \ \text{if } b \text{ then } u \text{ else } v$$

for $x \notin fv(b)$.

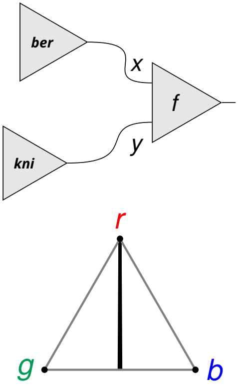
Outline

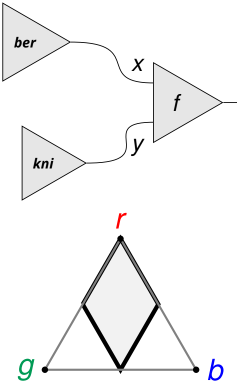
1. Imprecise Probability
2. Desiderata
3. The Problem
4. The Solution: Named Knightian Choices
5. A Graded Category
6. Results

The Problem: CP doesn't work [Mio et al. 2020]

bernoulli interpreted as $\left\{ \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \right\}$

knight interpreted as $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

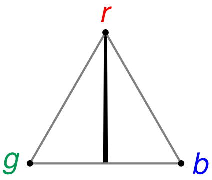


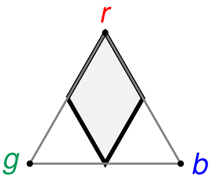


$f(x,y)$ = if x then (if y then return *r* else return *g*)
else (if y then return *r* else return *b*)

The Problem

Theorem: Any semantic model that satisfies our desiderata cannot distinguish the following convex sets of distributions.





$z \leftarrow$ ***bernoulli*** ;
if z then (if $x \leftarrow$ ***knight*** ; x then return *r* else return *g*)
else (if $x \leftarrow$ ***knight*** ; x then return *r* else return *b*)

Outline

1. Imprecise Probability
2. Desiderata
3. The Problem
4. The Solution: Named Knightian Choices
5. A Graded Category
6. Results

The Solution: Named Knightian choices

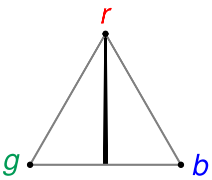
$x \leftarrow$ ***knight***(a_1) ; $z \leftarrow$ ***bernoulli*** ;
if z then (if x then return r else return g)
else (if x then return r else return b)

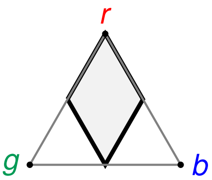
=

$x \leftarrow$ ***knight***(a_1) ; $y \leftarrow$ ***knight***(a_1) ; $z \leftarrow$ ***bernoulli*** ;
if z then (if x then return r else return g)
else (if y then return r else return b)

$\neq$

$x \leftarrow$ ***knight***(a_1) ; $y \leftarrow$ ***knight***(a_2) ; $z \leftarrow$ ***bernoulli*** ;
if z then (if x then return r else return g)
else (if y then return r else return b)





The Solution: Named Knightian choices

- Reader monad transformer of finite distributions monad

$$T_{2^A}(X) = [2^A \Rightarrow D(X)]$$

The Solution: Named Knightian choices

- Reader monad transformer of finite distributions monad

$$T_{2^A}(X) = [2^A \Rightarrow D(X)]$$

- **Knightian** choices given by *reading*

The Solution: Named Knightian choices

- Reader monad transformer of finite distributions monad

$$T_{2^A}(X) = [2^A \Rightarrow D(X)]$$

- **Knightian** choices given by *reading*
- **Bernoulli** choices given by *distributions*

The Solution: Named Knightian choices

- Reader monad transformer of finite distributions monad

$$T_{2^A}(X) = [2^A \Rightarrow D(X)]$$

- **Knightian** choices given by *reading*
- **Bernoulli** choices given by *distributions*
- We generalise the Knightian choices 2^A to arbitrary sets B

$$T_B(X) = [B \Rightarrow D(X)]$$

The Solution: Named Knightian choices

- Reader monad transformer of finite distributions monad

$$T_{2^A}(X) = [2^A \Rightarrow D(X)]$$

- **Knightian** choices given by *reading*
- **Bernoulli** choices given by *distributions*
- We generalise the Knightian choices 2^A to arbitrary sets B

$$T_B(X) = [B \Rightarrow D(X)]$$

- Convex powerset recovered by pushing forward maximal convex distribution on B

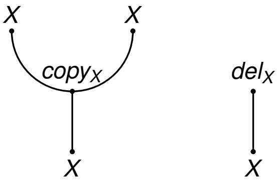
$$[\![t]\!]_B = \{p \gg\!=_D t \mid p \in D(B)\} \in \mathrm{CP}(X).$$

Outline

1. Imprecise Probability
2. Desiderata
3. The Problem
4. The Solution: Named Knightian Choices
5. A Graded Category
6. Results

Markov Categories

A **Markov category** [Fritz 2020] is a symmetric monoidal category such that every object is equipped with a commutative comonoid structure.



A **distributive Markov category** [Ackerman et al. 2024] is a Markov category with coproducts such that $X \otimes Z + Y \otimes Z \cong (X \otimes Y) + Z$ and injections commute with *copy* maps.

The graded Markov category **ImP**

Definition

FinStoch is the distributive Markov category of natural numbers and stochastic matrices.

The graded Markov category **ImP**

Definition

FinStoch is the distributive Markov category of natural numbers and stochastic matrices.

Definition

$\mathbf{FinStoch}_{\mathrm{Surj}}$ is the subcategory of **FinStoch** with the same objects but only surjective stochastic matrices.

The graded Markov category **ImP**

Definition

FinStoch is the distributive Markov category of natural numbers and stochastic matrices.

Definition

$\mathbf{FinStoch}_{\mathrm{Surj}}$ is the subcategory of **FinStoch** with the same objects but only surjective stochastic matrices.

Definition

ImP is the $\mathbf{FinStoch}_{\mathrm{Surj}}$-graded version of **FinStoch**.
That is, for $a \in \mathbf{FinStoch}_{\mathrm{Surj}}$ and $x, y \in \mathbf{FinStoch}$:

$$\mathbf{ImP}_a(X, Y) = \mathbf{FinStoch}(a \otimes X, Y)$$

The graded Markov category **ImP**

Definition

ImP is the $\mathbf{FinStoch}_{\mathrm{Surj}}$-graded version of **FinStoch**.
That is, for $a \in \mathbf{FinStoch}_{\mathrm{Surj}}$ and $x, y \in \mathbf{FinStoch}$:

$$\mathbf{ImP}_a(X, Y) = \mathbf{FinStoch}(a \otimes X, Y)$$

ImP supports finite probability and finite non-determinism:

- ***bernoulli*** is a morphism in $\mathbf{ImP}_1(1, 2)$ given by $\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$
- ***knight*** is a morphism in $\mathbf{ImP}_2(1, 2)$ given by unit diagonal

Composition uses independent non-deterministic branches:

$$\mathbf{ImP}_a(X, Y) \times \mathbf{ImP}_b(Y, Z) \to \mathbf{ImP}_{a \otimes b}(X, Z)$$

Monoidal structure too:

$$\mathbf{ImP}_b(X, Y) \times \mathbf{ImP}_a(X', Y') \to \mathbf{ImP}_{a \otimes b}(X \otimes X', Y \otimes Y')$$

Outline

1 Imprecise Probability

2 Desiderata

3 The Problem

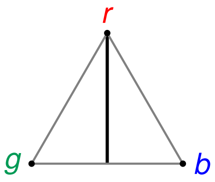
4 The Solution: Named Knightian Choices

5 A Graded Category

6 Results

Threorem 1: Improved Bounds

$$\begin{matrix} r: \\ g: \\ b: \end{matrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \\ 0 & 0.5 \end{pmatrix} \longrightarrow$$



- There is a mapping from $f \in \mathbf{ImP}_a(X, Y)$ to $R(f) : m \to \mathrm{CP}(n)$
- This is an 'op-lax' functor,

$$R : \mathbf{ImP} \to \mathrm{Kl}(\mathrm{CP})$$

- So composition in **ImP** gives tighter bounds on the *Knightian uncertainty* than composition in $\mathrm{Kl}(\mathrm{CP})$

$$R(g \circ f) \subseteq R(g) \circ R(f)$$

Threorem 2: Maximality

- The language gives rise to a compositional theory of equality
- This equational theory is maximal
- We can add no further equations without
  - Compromising imprecise probability connection (equating different convex subsets); or
  - Compromising the compositional structure

Conclusion

This work:

- Using graded perspective and naming Knightian choices we give a fully compositional account of Bernoulli and Knightian uncertainty together
- **Theorem 1:** This gives a refined bound on uncertainty
- **Theorem 2:** It is maximal among compositional accounts

Future work:

- Iteration and infinite dimensional structures
- Function spaces via quasi-Borel spaces
- Implementation and approximation of bounds

References I

- Ackerman, N. et al. (2024). "Probabilistic programming interfaces for random graphs: Markov categories, graphons, and nominal sets.". In: *Proc. POPL 2024.*
- Dahlqvist, Fredrik, Louis Parlant, and Alexandra Silva (2018). "Layer by layer: composing monads". In: *Proc. ICTAC 2018.*
- Dash, Swaraj (2024). "A Monadic Theory of Point Processes". PhD thesis. University of Oxford.
- Dash, Swaraj and Sam Staton (2020). "A monad for probabilistic point processes". In: *Proc. ACT 2020.*
- — (2021). "Monads for Measurable Queries in Probabilistic Databases". In: *Proc. MFPS 2021.*
- Fritz, T. (2020). "A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics". In: *Adv. Math.* 370.

References II

- Jacobs, Bart (2021). “From multisets over distributions to distributions over multisets”. In: *Proc. LICS 2021*.
- Keimel, Klaus and Gordon D. Plotkin (2017). “Mixed powerdomains for probability and nondeterminism”. In: *Log. Methods Comput. Sci.* 13.
- Kozen, Dexter and Alexandra Silva (2023). *Multisets and Distributions*. arxiv:2301.10812.
- Mio, Matteo and Valeria Vignudelli (2020). “Monads and Quantitative Equational Theories for Nondeterminism and Probability”. In: *Proc. CONCUR 2020*.
- Varacca, D and G Winskel (2006). “Distributing probability over non-determinism”. In: *Math. Structures Comput. Sci.*, pp. 87–113.