Compositional Imprecise Probability

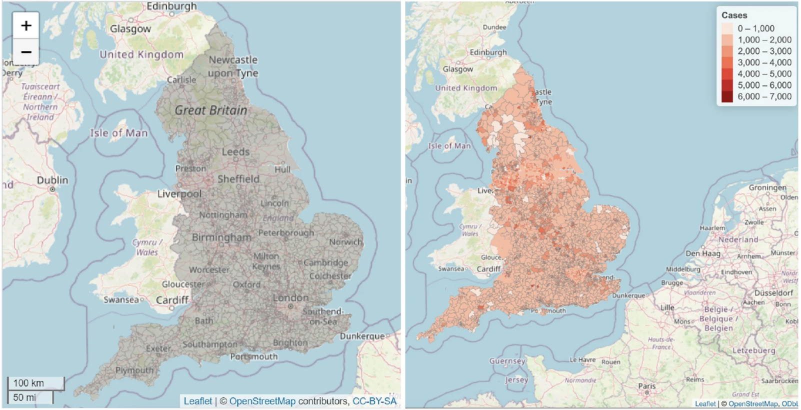
Jack Liell-Cock [1]   Sam Staton [1]

[1] University of Oxford

POPL 2025
22nd January, 2025

Context

- Probabilistic programming for modelling
- Programs model phenomena allowing automatic inference
- Successes in areas like UK covid predictions



Overview

This work:

- Design a fully compositional probabilistic programming language with Bernoulli and Knightian uncertainty
- Using graded monads and named Knightian choices
- **Theorem 1:** This gives a refined bound on uncertainty
- **Theorem 2:** It is maximal among compositional accounts

Overview

This work:

- Design a fully compositional probabilistic programming language with Bernoulli and Knightian uncertainty
- Using graded monads and named Knightian choices
- **Theorem 1:** This gives a refined bound on uncertainty
- **Theorem 2:** It is maximal among compositional accounts

Not this work:

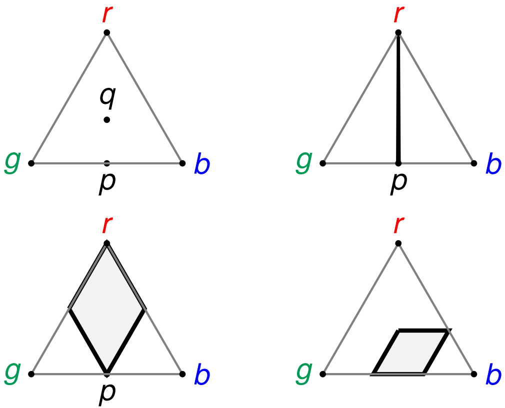
- Wider interest in probability with non-determinism [Dash and Staton 2020; Dahlqvist et al. 2018; Keimel et al. 2017; Dash 2024; Kozen et al. 2023; Varacca et al. 2006; Jacobs 2021]
- Our focus is in the setting of imprecise probability

Contents

Outline

Imprecise Probability

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



Applications of Imprecise Probability

Wider applicataions:

- Statistical robustness [Walley 1991; Huber 1981],
- Economics [Amarante et al. 2007; Battigalli, Cerreia-Vioglio, et al. 2015; Battigalli, Corrao, et al. 2019; Cerreia-Vioglio et al. 2023; Vicig 2008; Pavlidis et al. 2011]
- Engineering [Ferson et al. 2003; Aslett et al. 2022; Vasile 2021; Krpelík et al. 2021]

In machine learning:

- Bayesian learning [Caprio et al. 2024]
- Reinforcement learning [Oren et al. 2024; Zanger et al. 2023]
- Conformal prediction [Stutz et al. 2023; Javanmardi et al. 2024]
- Foundations of safe AI [Dalrymple 2024; Dalrymple et al. 2024]

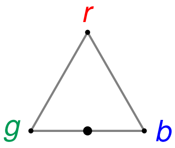
A First Language for Imprecise Probability

Our prototype language for imprecise probability will have:

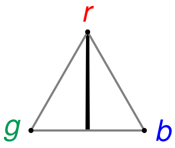
- If-then-else statements;
  - if b then t else u
- Let bindings;
  - $x \leftarrow t \; ; \; u$
- Two commands returning booleans:
  - ***bernoulli***: a fair Bernoulli choice;
  - ***knight***: a Knightian choice.

Examples with the Language

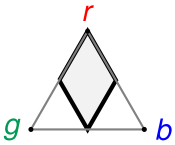
$z \leftarrow$ ***bernoulli*** ;
if z then g else b



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



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



Outline

Desiderata of the Language

Desideratum (1)

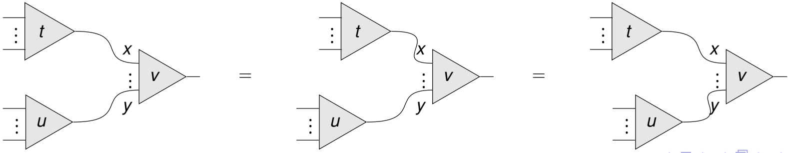
The language should be commutative:

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

for $x \notin \mathrm{free}(u)$ *and* $y \notin \mathrm{free}(t)$*; and affine:*

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

for $x \notin \mathrm{free}(u)$.



Desideratum (2)

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

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

$$=$$

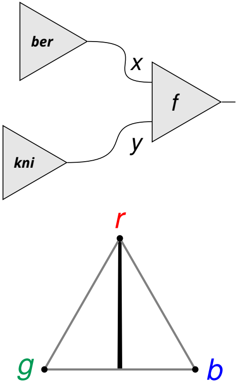
$$x \leftarrow t \; ; \; \text{if } \; b \; \text{then } u \; \text{else } v$$

for $x \notin \mathrm{free}(b)$.

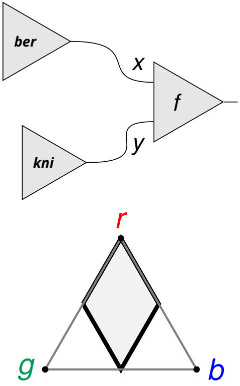
Outline

The Problem: Convex Powerset doesn't work

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



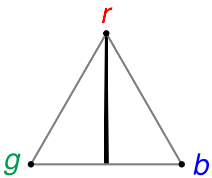
knight interpreted as $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

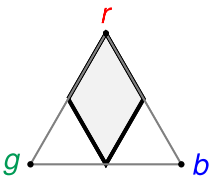


$$f(x,y) = \text{if } x \text{ then } (\text{if } y \text{ then } r \text{ else } g) \text{ else } (\text{if } y \text{ then } r \text{ else } b)$$

The Problem: Nothing Works

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



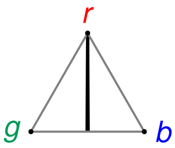


$z \leftarrow$ ***bernoulli*** ;
if z then (if $x \leftarrow$ ***knight*** ; x then *r* else *g*)
else (if $x \leftarrow$ ***knight*** ; x then *r* else *b*)

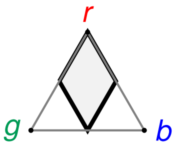
Outline

The Solution: Named Knightian choices

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



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



The Solution: Named Knightian choices

- Reader monad transformer of finite distributions monad

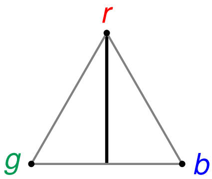
$$T_{2^A}(X) = [2^A \Rightarrow D(X)]$$

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

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

- Convex powerset recovered by taking convex hull of the image

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



A Graded Monad for Imprecise Probability

Definition

T is the $\mathbf{FinStoch}_{\mathrm{Surj}}$-graded version of:

$$T_A(X) = [A \Rightarrow D(X)]$$

T supports finite probability and finite non-determinism:

- $\boldsymbol{bernoulli} \in T_{2^{\emptyset}}(2)$ given by $\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$
- $\boldsymbol{knight}(a) \in T_{2^{\{a\}}}(2)$ given by $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

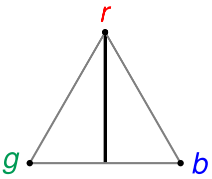
Function and monoidal composition use independent non-deterministic branches by the graded multiplication:

$$\mu_{A,B} : T_A \circ T_B \to T_{A \times B}$$

Outline

Theorem 1: Improved Uncertainty Bounds

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



- There is a mapping $T_A(X) \to \mathrm{CP}(X)$ by taking the convex hull of the image
- This induces an "op-lax functor":

$$R : [X \Rightarrow T_A(Y)] \to [X \Rightarrow \mathrm{CP}(Y)]$$

- That is, composition in our framework gives tighter bounds on the *Knightian uncertainty* than composition with CP

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

Theorem 2: Maximality of the Theory

- The mapping into convex powersets is not injective
- There is ”more stuff” in our theory than in convex powersets
- So can we quotient our theory any further?
- Not without:
  - Breaking connection with convex subsets (and imprecise probability); or
  - Compromising the compositional structure

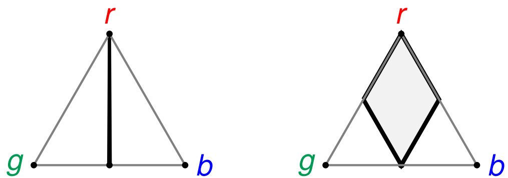
Conclusion

This work:

- Using graded monads and named 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

- Amarante, Massimiliano and Fabio Maccheroni (2007). "When an Event Makes a Difference". In: *Uncertainty and Risk: Mental, Formal, Experimental Representations*. Ed. by Mohammed Abdellaoui et al. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 27–32. ISBN: 978-3-540-48935-1. DOI: `10.1007/978-3-540-48935-1_3`.
- Aslett, Louis, Frank Coolen, and Jasper Bock (Jan. 2022). *Uncertainty in Engineering Introduction to Methods and Applications: Introduction to Methods and Applications*. ISBN: 978-3-030-83639-9. DOI: `10.1007/978-3-030-83640-5`.
- Battigalli, Pierpaolo, Simone Cerreia-Vioglio, et al. (Feb. 2015). "Self-Confirming Equilibrium and Model Uncertainty". In: *American Economic Review* 105.2, pp. 646–77. DOI: `10.1257/aer.20130930`.
- Battigalli, Pierpaolo, Roberto Corrao, and Martin Dufwenberg (2019). "Incorporating belief-dependent motivation in games". In: *Journal of Economic Behavior & Organization* 167, pp. 185–218. ISSN: 0167-2681. DOI: `10.1016/j.jebo.2019.04.009`.

References II

- Caprio, Michele et al. (2024). *Credal Bayesian Deep Learning*. arXiv: 2302.09656 [cs.LG].
- Cerreia-Vioglio, Simone, Roberto Corrao, and Giacomo Lanzani (July 2023). "Dynamic Opinion Aggregation: Long-run Stability and Disagreement". In: *The Review of Economic Studies* 91.3, pp. 1406–1447. ISSN: 0034-6527. DOI: 10.1093/restud/rdad072.
- Dahlqvist, Fredrik, Louis Parlant, and Alexandra Silva (2018). "Layer by layer: composing monads". In: *Proc. ICTAC 2018*.
- Dalrymple, David 'davidad' (Jan. 2024). *Safeguarded AI: constructing safety by design*. ARIA Programme Thesis. https://www.aria.org.uk/wp-content/uploads/2024/01/ARIA-Safeguarded-AI-Programme-Thesis-V1.pdf.
- Dalrymple, David 'davidad' et al. (2024). *Towards Guaranteed Safe AI: A Framework for Ensuring Robust and Reliable AI Systems*. arXiv: 2405.06624 [cs.AI].

References III

- 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*.
- Ferson, Scott et al. (2003). *Constructing Probability Boxes and Dempster-Shafer Structures*. Tech. rep. Technical Report SAND2002-4015. Sandia National Laboratiories.
- Huber, P J (1981). *Robust statistics*. Wiley.
- Jacobs, Bart (2021). “From multisets over distributions to distributions over multisets”. In: *Proc. LICS 2021*.
- Javanmardi, Alireza, David Stutz, and Eyke Hüllermeier (2024). *Conformalized Credal Set Predictors*. arXiv: `2402.10723 [stat.ML]`.
- Keimel, Klaus and Gordon D. Plotkin (2017). “Mixed powerdomains for probability and nondeterminism”. In: *Log. Methods Comput. Sci.* 13.

References IV

- Kozen, Dexter and Alexandra Silva (2023). *Multisets and Distributions*. arxiv:2301.10812.
- Krpelík, Danda and Tathagata Basu (Jan. 2021). “Introduction to Imprecise Probabilities”. In: chap. 13, pp. 35–79. ISBN: 978-3-030-60165-2. DOI: 10.1007/978-3-030-60166-9_2.
- Oren, Yaniv et al. (2024). *Epistemic Monte Carlo Tree Search*. arXiv: 2210.13455 [cs.LG].
- Pavlidis, E.G. et al. (Aug. 2011). “Nonlinear dynamics in economics and finance and unit root testing”. In: *The European Journal of Finance* 2011. DOI: 10.1080/1351847X.2011.607006.
- Stutz, David et al. (2023). *Conformal prediction under ambiguous ground truth*. arXiv: 2307.09302 [cs.LG].
- Varacca, D and G Winskel (2006). “Distributing probability over non-determinism”. In: *Math. Structures Comput. Sci.*, pp. 87–113.

References V

- Vasile, Massimiliano (July 2021). *Optimization Under Uncertainty with Applications to Aerospace Engineering*. ISBN: 978-3-030-60165-2.
- Vicig, Paolo (Sept. 2008). "Financial Risk Measurement with Imprecise Probabilities". In: *Int. J. Approx. Reasoning* 49.1, pp. 159–174. ISSN: 0888-613X. DOI: 10.1016/j.ijar.2007.06.009.
- Walley, Peter (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall.
- Zanger, Moritz A., Wendelin Böhmer, and Matthijs T. J. Spaan (2023). *Diverse Projection Ensembles for Distributional Reinforcement Learning*. arXiv: 2306.07124 [cs.LG].