Categories of partial probabilistic computations

Márk Széles

DutchCATS, Nijmegen – February 7, 2025

Radboud Universiteit

Discrete probability distributions

Definition

A **discrete probability distribution** on a set X is a function $\varphi : X \to [0, 1]$ such that

1. $\operatorname{supp}(\varphi) = \{x \in X : \varphi(x) \neq 0\}$ is a finite set.
2. $\sum_{x \in X} \varphi(x) = 1$.

Discrete probability distributions

Definition

A **discrete probability distribution** on a set X is a function $\varphi : X \to [0,1]$ such that

1. $\mathrm{supp}(\varphi) = \{x \in X : \varphi(x) \neq 0\}$ is a finite set.
2. $\sum_{x \in X} \varphi(x) = 1$.

We often write such distributions as formal sums:

$$\mathit{flip} : \mathcal{D}(\{H, T\})$$
$$\mathit{flip} = \frac{1}{2}\Big|H\Big\rangle + \frac{1}{2}\Big|T\Big\rangle$$

Discrete probability distributions

Definition

A **discrete probability distribution** on a set X is a function $\varphi : X \to [0,1]$ such that

1. $\mathrm{supp}(\varphi) = \{x \in X : \varphi(x) \neq 0\}$ is a finite set.
2. $\sum_{x \in X} \varphi(x) = 1$.

We often write such distributions as formal sums:

$$\mathit{flip} : \mathcal{D}(\{H, T\})$$
$$\mathit{flip} = \frac{1}{2}\Big|H\Big\rangle + \frac{1}{2}\Big|T\Big\rangle$$

There is a monad $\mathcal{D} : \mathbf{Sets} \to \mathbf{Sets}$.

Disintegration

Disintegration

$$\omega : \mathcal{D}(\mathbb{A} \times \mathbb{N})$$

$$\omega = \frac{1}{3}\Big|a, 1\Big\rangle + \frac{1}{3}\Big|b, 1\Big\rangle + \frac{1}{3}\Big|b, 3\Big\rangle$$

Disintegration

$$\omega : \mathcal{D}(\mathbb{A} \times \mathbb{N})$$
$$\omega = \frac{1}{3}\Big|a, 1\Big\rangle + \frac{1}{3}\Big|b, 1\Big\rangle + \frac{1}{3}\Big|b, 3\Big\rangle$$

Compute the conditional probability $\omega|_{\mathbb{A}} : \mathbb{A} \to \mathcal{D}(\mathbb{N})$

Disintegration

$$\omega : \mathcal{D}(\mathbb{A} \times \mathbb{N})$$
$$\omega = \frac{1}{3}\big|a, 1\big\rangle + \frac{1}{3}\big|b, 1\big\rangle + \frac{1}{3}\big|b, 3\big\rangle$$

Compute the conditional probability $\omega|_{\mathbb{A}} : \mathbb{A} \to \mathcal{D}(\mathbb{N})$

$$\frac{1}{3}\big|a, 1\big\rangle + \frac{1}{3}\big|b, 1\big\rangle + \frac{1}{3}\big|b, 3\big\rangle \quad \mapsto \quad \text{(observe that the letter is } b\text{)}$$

$$\frac{1}{3}\big|a, 1\big\rangle + \frac{1}{3}\big|b, 1\big\rangle + \frac{1}{3}\big|b, 3\big\rangle \quad \mapsto \quad \text{renormalise}$$

$$\frac{1}{2}\big|b, 1\big\rangle + \frac{1}{2}\big|b, 3\big\rangle \quad = \omega|_{\mathbb{A}}(b)$$

Disintegration is partial

$$\omega : \mathcal{D}(\mathbb{A} \times \mathbb{N})$$
$$\omega = \frac{1}{3}\Big|a, 1\Big\rangle + \frac{1}{3}\Big|b, 1\Big\rangle + \frac{1}{3}\Big|b, 3\Big\rangle$$

Disintegration is partial

$$\omega : \mathcal{D}(\mathbb{A} \times \mathbb{N})$$
$$\omega = \frac{1}{3}\big|a, 1\big\rangle + \frac{1}{3}\big|b, 1\big\rangle + \frac{1}{3}\big|b, 3\big\rangle$$

Compute the conditional probability $\omega|_{\mathbb{A}} : \mathbb{A} \to \mathcal{D}(\mathbb{N})$

Disintegration is partial

$$\omega : \mathcal{D}(\mathbb{A} \times \mathbb{N})$$
$$\omega = \frac{1}{3}\big|a, 1\big\rangle + \frac{1}{3}\big|b, 1\big\rangle + \frac{1}{3}\big|b, 3\big\rangle$$

Compute the conditional probability $\omega|_{\mathbb{A}} : \mathbb{A} \to \mathcal{D}(\mathbb{N})$

$$\frac{1}{3}\big|a, 1\big\rangle + \frac{1}{3}\big|b, 1\big\rangle + \frac{1}{3}\big|b, 3\big\rangle \quad \mapsto \quad \text{(observe that the letter is } c\text{)}$$

$$\frac{1}{3}\big|a, 1\big\rangle + \frac{1}{3}\big|b, 1\big\rangle + \frac{1}{3}\big|b, 3\big\rangle \quad \mapsto \quad \text{renormalise}$$

???

Discrete probability subdistributions

Definition

A **discrete probability subdistribution** on a set X is a function $\varphi : X \to [0,1]$ such that

1. $\mathrm{supp}(\varphi) = \{x \in X : \varphi(x) \neq 0\}$ is a finite set.
2. $\sum_{x \in X} \varphi(x) \leq 1$.

There is a monad $\mathcal{D}_{\leq} : \mathbf{Sets} \to \mathbf{Sets}$. Its Kleisli-category is symmetric monoidal.
Write $f : X \circ\!\!\to Y$ for a Kleisli-map $f : X \to \mathcal{D}_{\leq}(Y)$.

The structure of $\mathcal{K}l(\mathcal{D}_{\leq})$

- Copier maps:

$$\Delta_X : X \multimapdotinv X \times X$$
$$\Delta_X(x) = 1|x, x\rangle$$

The structure of $\mathcal{K}l(\mathcal{D}_{\leq})$

- Copier maps:

$$\Delta_X : X \multimap X \times X$$
$$\Delta_X(x) = 1|x, x\rangle$$

- Discard maps:

$$d_X : X \multimap 1$$
$$d_X(x) = 1|0\rangle$$

The structure of $\mathcal{K}l(\mathcal{D}_{\leq})$

- Copier maps:

$$\Delta_X : X \multimap X \times X$$
$$\Delta_X(x) = 1|x, x\rangle$$

- Discard maps:

$$d_X : X \multimap 1$$
$$d_X(x) = 1|0\rangle$$

- Comparator maps:

$$\nabla_X : X \times X \multimap X$$
$$\nabla_X(x, y) = \begin{cases} 1|x\rangle & \text{if } x = y \\ \mathbf{0} & \text{otherwise} \end{cases}$$

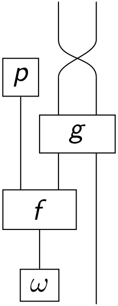
Copy-discard-compare (CDC) categories

Definition

A **copy-discard-compare category** is a symmetric monoidal category $(C, \otimes, I)$ with **copier** $\Delta_X : X \to X \otimes X$, **discard** $d_X : X \to I$, and **comparator** $\nabla_X : X \otimes X \to X$ maps such that...

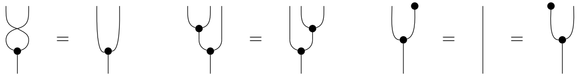
The anatomy of a string diagram

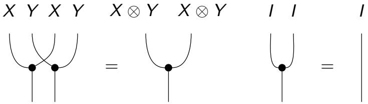
Let $(\mathbf{C}, \otimes, I, \alpha, \lambda, \rho, \sigma)$ be symmetric monoidal.



The axioms of copy-discard-compare (CDC) categories

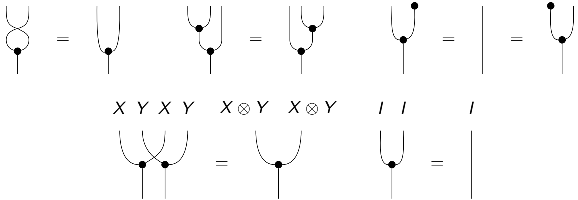
- Copy and discard:



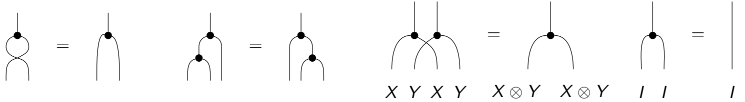


The axioms of copy-discard-compare (CDC) categories

- Copy and discard:

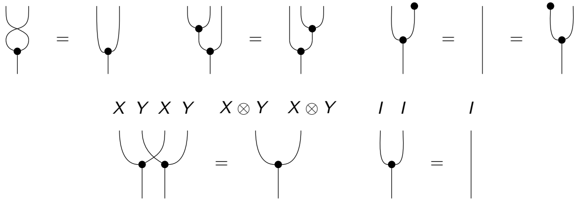


- Compare:

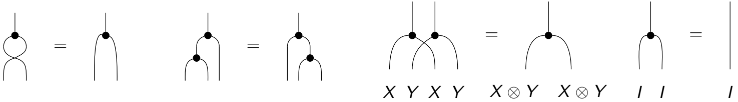


The axioms of copy-discard-compare (CDC) categories

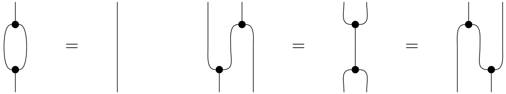
- Copy and discard:



- Compare:



- Copy-compare interaction:



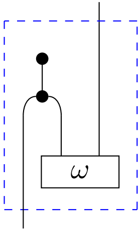
Examples of CDC-categories

- Sets and subprobability channels: $\mathcal{K}\ell(\mathcal{D}_{\leq})$
- Finite dimensional vector spaces and linear maps: **FinVect**
- Sets and relations: **Rel**
- Standard Borel spaces and subprobability kernels: $\textbf{BorelStoch}_{\leq}$
- ...

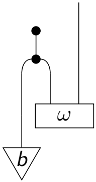
Our disintegration problem revisited

From $\omega : 1 \multimap \mathbb{A} \times \mathbb{N}$
extract $\omega|_{\mathbb{A}} : \mathbb{A} \multimap \mathbb{N}$:

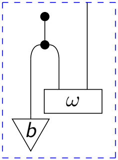
$\omega|_{\mathbb{A}} \quad :=$



Normalisation

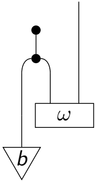


$\mapsto$

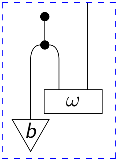


$$\frac{1}{3}\Big|a,1\Big\rangle + \frac{1}{3}\Big|b,1\Big\rangle + \frac{1}{3}\Big|b,3\Big\rangle \quad\mapsto\quad \frac{1}{2}\Big|b,1\Big\rangle + \frac{1}{2}\Big|b,3\Big\rangle$$

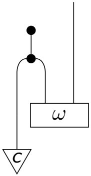
Normalisation



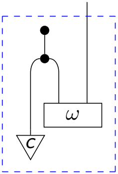
$\mapsto$



$$\tfrac{1}{3}\big|a,1\big\rangle + \tfrac{1}{3}\big|b,1\big\rangle + \tfrac{1}{3}\big|b,3\big\rangle \quad \mapsto \quad \tfrac{1}{2}\big|b,1\big\rangle + \tfrac{1}{2}\big|b,3\big\rangle$$



$\mapsto$



$$\tfrac{1}{3}\big|a,1\big\rangle + \tfrac{1}{3}\big|b,1\big\rangle + \tfrac{1}{3}\big|b,3\big\rangle \quad \mapsto \quad \mathbf{0}$$

Normalisation

- One can axiomatise the normalisation boxes purely in terms of CD-categories

Normalisation

- One can axiomatise the normalisation boxes purely in terms of CD-categories
- These axioms uniquely characterise normalisation boxes when they exist.

Normalisation

- One can axiomatise the normalisation boxes purely in terms of CD-categories
- These axioms uniquely characterise normalisation boxes when they exist.
- The normalisation boxes select a least normalisation with respect to a partial order on normalisations.

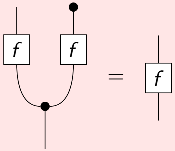
Normalisation

- One can axiomatise the normalisation boxes purely in terms of CD-categories
- These axioms uniquely characterise normalisation boxes when they exist.
- The normalisation boxes select a least normalisation with respect to a partial order on normalisations.
- The dashed boxes enjoy many compositional properties.

Self-normalising maps

Definition

We call a map $f : X \to Y$ **self-normalising** if



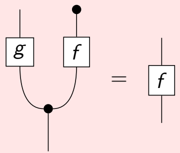
For a subchannel $f : X \multimap Y$ in $\mathcal{K}l(\mathcal{D}_{\leq})$ this translates to

$$\forall x \in X. \sum_{y \in Y} f(x)(y) \in \{0, 1\}$$

The 'normalised by' relation

Definition

A map $g : X \to Y$ **normalises** $f : X \to Y$ if



In this case, we write $f \preceq g$.

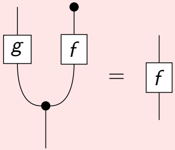
For subchannels $f, g : X \multimap Y$ in $\mathcal{K}\ell(\mathcal{D}_{\leq})$ this translates to

$$f(x) \neq \mathbf{0} \implies \forall y \in Y . g(x)(y) = \frac{f(x)(y)}{\sum_{y'} f(x)(y')}$$

The 'normalised by' relation

Definition

A map $g : X \to Y$ **normalises** $f : X \to Y$ if



In this case, we write $f \preceq g$.

For subchannels $f, g : X \multimap Y$ in $\mathcal{K}\ell(\mathcal{D}_{\leq})$ this translates to

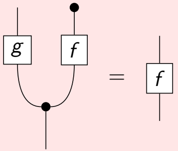
$$f(x) \neq \mathbf{0} \implies \forall y \in Y. g(x)(y) = \frac{f(x)(y)}{\sum_{y'} f(x)(y')}$$

- The relation $\preceq$ is a partial order on self-normalising maps.

The ‘normalised by’ relation

Definition

A map $g : X \to Y$ **normalises** $f : X \to Y$ if



In this case, we write $f \preceq g$.

For subchannels $f, g : X \rightsquigarrow Y$ in $\mathcal{K}\ell(\mathcal{D}_{\leq})$ this translates to

$$f(x) \neq \mathbf{0} \Longrightarrow \forall y \in Y . g(x)(y) = \frac{f(x)(y)}{\sum_{y'} f(x)(y')}$$

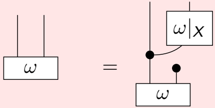
- The relation $\preceq$ is a partial order on self-normalising maps.
- The dashed box should select the least normalisation.

Disintegration

'How to factorise $P(X, Y) = P(X) \cdot P(Y|X)$?'.

Definition

If $\omega : I \to X \otimes Y$, then a *disintegration* of ω is a map $\omega|_X : X \to Y$ that satisfies



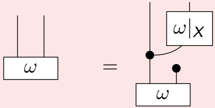
This can be generalised to maps with arbitrary domain.

Disintegration

'How to factorise $P(X, Y) = P(X) \cdot P(Y|X)$?'.

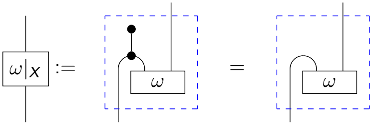
Definition

If $\omega : I \to X \otimes Y$, then a *disintegration* of ω is a map $\omega|_X : X \to Y$ that satisfies



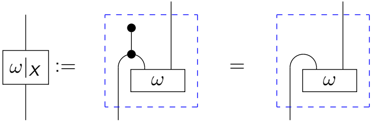
This can be generalised to maps with arbitrary domain.

Can we use comparators and normalisation to compute a disintegration?



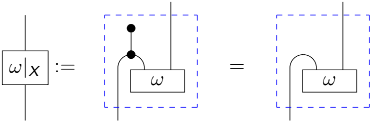
Deriving disintegration

Can we use comparators and normalisation to compute a disintegration?

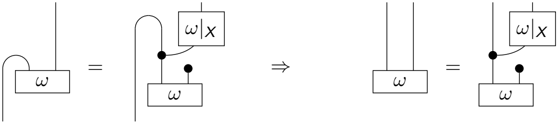


Deriving disintegration

Can we use comparators and normalisation to compute a disintegration?

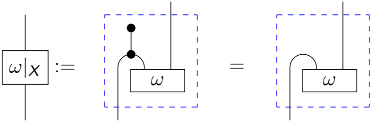


Yes, when the following implication holds:

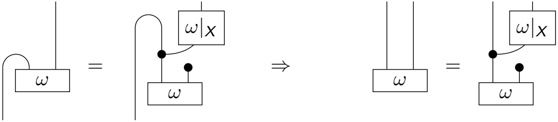


Deriving disintegration

Can we use comparators and normalisation to compute a disintegration?



Yes, when the following implication holds:



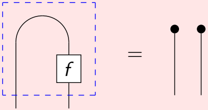
Remark

This does not recover disintegration in $\mathbf{BorelStoch}_{\leq}$

Outlook: notions of support

Definition

A map $f : X \to Y$ has **everywhere full support** if

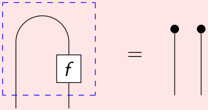


- For $f : X \rightsquigarrow Y$ this means $f(x)(y) > 0$ for all $x \in X$, $y \in Y$.

Outlook: notions of support

Definition

A map $f : X \to Y$ has **everywhere full support** if

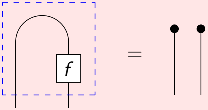


- For $f : X \multimap Y$ this means $f(x)(y) > 0$ for all $x \in X$, $y \in Y$.
- If f has everywhere full support, updating f never fails.

Outlook: notions of support

Definition

A map $f : X \to Y$ has **everywhere full support** if

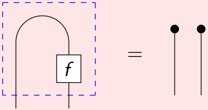


- For $f : X \multimap Y$ this means $f(x)(y) > 0$ for all $x \in X$, $y \in Y$.
- If f has everywhere full support, updating f never fails.
- Everywhere full support maps are closed under many operations: composition, tensor, marginalisation, normalisation, disintegration.

Outlook: notions of support

Definition

A map $f : X \to Y$ has **everywhere full support** if



- For $f : X \multimap Y$ this means $f(x)(y) > 0$ for all $x \in X$, $y \in Y$.
- If f has everywhere full support, updating f never fails.
- Everywhere full support maps are closed under many operations: composition, tensor, marginalisation, normalisation, disintegration.
- We use this notion for a compositional graphical calculus of disintegrations in a future article with Bart Jacobs, and Dario Stein.

Outlook: connections to effectus theory

- Effectus theory is a category theoretic framework for quantum and probability theory, focusing on the structure of well-behaved coproducts.

Outlook: connections to effectus theory

- Effectus theory is a category theoretic framework for quantum and probability theory, focusing on the structure of well-behaved coproducts.
- Effectuses need not have a monoidal structure. Adding that creates a very rich setting.

Outlook: connections to effectus theory

- Effectus theory is a category theoretic framework for quantum and probability theory, focusing on the structure of well-behaved coproducts.
- Effectuses need not have a monoidal structure. Adding that creates a very rich setting.
- Effectuses come in a total and partial flavour:

$$\mathbf{TotEf} \underset{\mathcal{T}ot}{\overset{\mathcal{K}\ell(-+1)}{\underset{\cong}{\rightleftarrows}}} \mathbf{ParEf}$$

References

- K. Cho and B. Jacobs, "Disintegration and Bayesian inversion via string diagrams," Mathematical Structures in Computer Science, vol. 29, no. 7, pp. 938–971, 2019. doi:10.1017/S0960129518000488
- E. Di Lavore and M. Román, "Evidential Decision Theory via Partial Markov Categories," 2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), pp. 1-14, 2023. doi: 10.1109/LICS56636.2023.10175776
- R. Lorenz and S. Tull, "Causal models in string diagrams," arXiv preprint, 2023. doi:10.48550/arXiv.2304.07638
- K. Cho, B. Jacobs, B. Westerbaan, and A. Westerbaan . "An introduction to effectus theory," arXiv preprint arXiv:1512.05813, 2015.