ItaCa Fest 27 April 2023

EVIDENTIAL DECISION THEORY VIA PARTIAL MARKOV CATEGORIES

Elena Di Lavore and Mario Román

Tallinn University of Technology

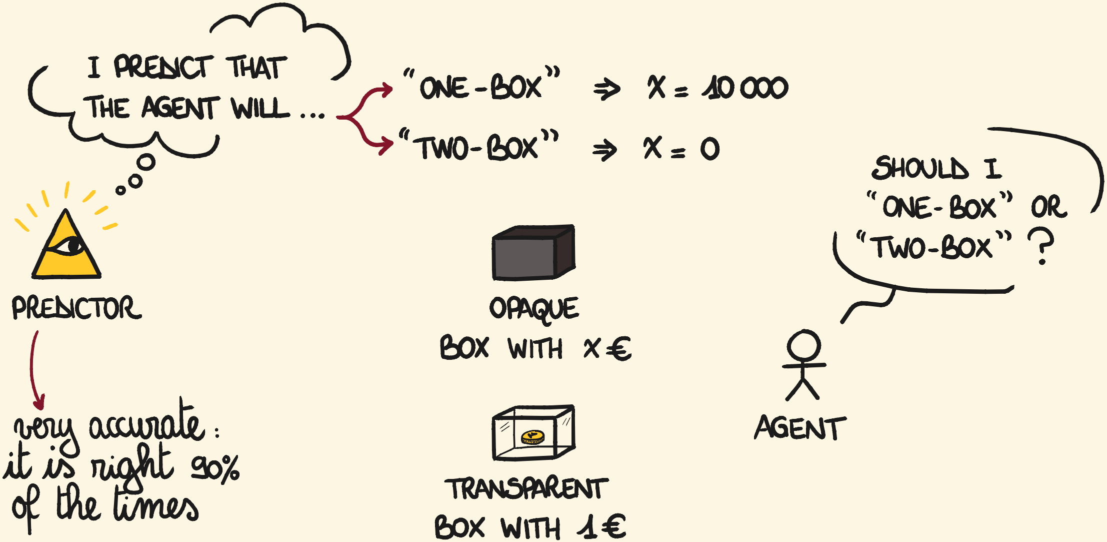
OUTLINE

[• motivation: Evidential Decision Theory]

• (discrete) partial Markov categories

• Bayes, observations & updates

NEWCOMB'S PROBLEM

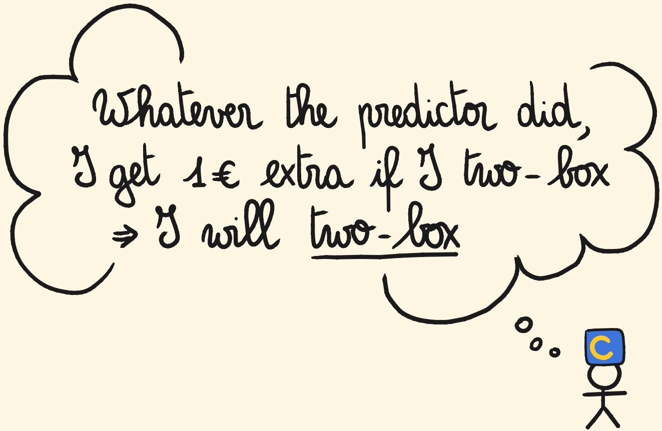


(drawing inspired by Mario's slides for NWPT '22)

CAUSAL DECISION THEORY

Causal decision theory answers :

"Which action would cause the best-case scenario ? "



BEST !

AGENT / PREDICTOR	ONE-BOX	TWO-BOX
ONE-BOX	10 000 €	10 001€
TWO-BOX	0 €	1 €

EVIDENTIAL DECISION THEORY

Evidential decision theory answers :

"Which action would be evidence for the best-case scenario ?"

My action is evidence for the prediction :
if I one-box I expect 10 000 €,
if I two-box I expect 1 € .
⇒ I will one-box

PREDICTOR \ AGENT	ONE-BOX	TWO-BOX
ONE-BOX	10 000 €	10 001 €
TWO-BOX	0 €	1 €

MOST LIKELY

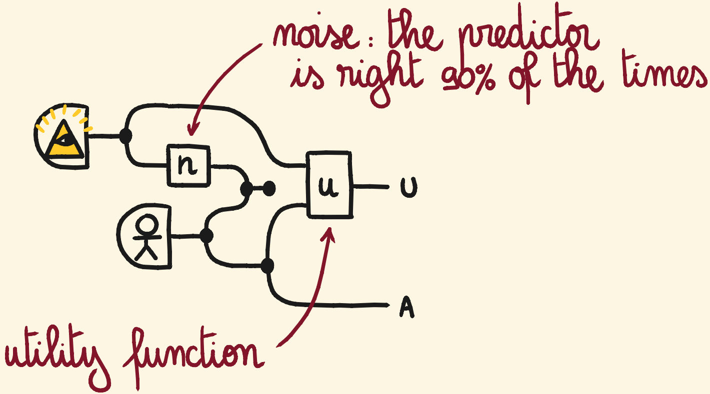
EVIDENTIAL VS CAUSAL DECISION THEORY

CAUSAL DECISION THEORIST

EVIDENTIAL DECISION THEORIST

	CAUSAL DECISION THEORIST	EVIDENTIAL DECISION THEORIST
EXPECTED UTILITY	0.9×1 € $+\ 0.1 \times 10\,001$ € $= 1\,001$ €	$0.9 \times 10\,000$ € $+\ 0.1 \times 0$ € $= 9\,000$ €

A DIAGRAM FOR NEWCOMB



AGENT / PREDICTOR	ONE-BOX	TWO-BOX
ONE-BOX	10 000 €	10 001 €
TWO-BOX	0 €	1 €

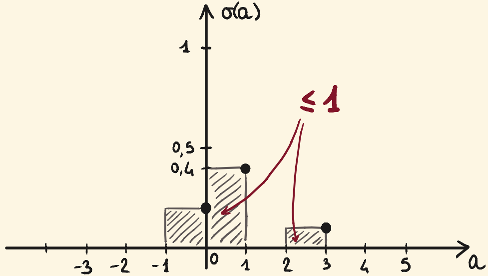
SUBDISTRIBUTIONS

A <u>subdistribution</u> σ on A is a distribution on $A+1$:

$\sigma \in \mathcal{D}_{\leq 1}(A)$ is a function $\sigma : A \to [0,1]$ such that

- its support, $\mathrm{supp}(\sigma) := \{a \in A \mid \sigma(a) > 0\}$, is finite, and
- its total probability mass is at most 1, $\sum_{a \in A} \sigma(a) \leq 1$.

<u>ex</u> $A = \mathbb{N}$, $\sigma =$



Subdistributions give a monad $\mathcal{D}_{\leq 1} : \mathbf{Set} \to \mathbf{Set}$

(there's a distributive law $\mathcal{D}(-)+1 \longrightarrow \mathcal{D}(-+1)$)

SUBDISTRIBUTIONS

A morphism $X - \boxed{f} - A$ in $Kl\mathcal{D}_{\leq 1}$ is a function $X \to \mathcal{D}_{\leq 1}(A)$

$$f(a|x) = \text{"probability of } a \text{ given } x\text{"}$$

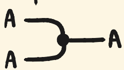
$$f(\perp|x) = \text{"probability of failure"}$$

Composition is:

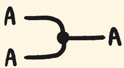
$$X - \boxed{f} - \boxed{g} - B\ (b|x) := \sum_{a \in A} f(a|x) \cdot g(b|a)$$

$$X - \boxed{f} - \boxed{g} - B\ (\perp|x) := f(\perp|x) + \sum_{a \in A} f(a|x) \cdot g(\perp|a)$$

Equality checks are lifted from partial functions:

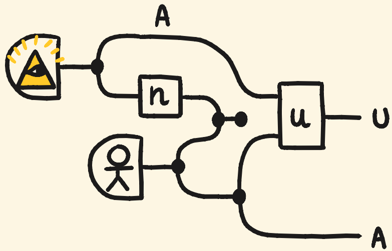


$$(a|a_1, a_2) := \begin{cases} 1 & a_1 = a_2 = a \\ 0 & \text{otherwise} \end{cases}$$



$$(\perp|a_1, a_2) := \begin{cases} 0 & a_1 = a_2 \\ 1 & a_1 \neq a_2 \end{cases}$$

READING NEWCOMB'S DIAGRAM

$N =$ 

$$N(v,a\mid *) = \sum_{p\in A} \triangle(p\mid *)\cdot n(a\mid p)\cdot \curlywedge(a\mid *)\cdot u(v\mid p,a)$$

$$N(\perp\mid *) = \sum_{a\neq a'\in A}\ \sum_{p\in A} \triangle(p\mid *)\cdot n(a'\mid p)\cdot \curlywedge(a\mid *)$$

WANTED: a calculus to reason with diagrams of this kind:

1. stochastic processes
2. partiality & constraints

OUTLINE

- motivation: Evidential Decision Theory

[- (discrete) partial Markov categories]

- Bayes, observations & updates

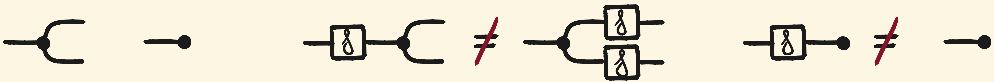
(DISCRETE) PARTIAL MARKOV CATEGORIES

- add partiality to Markov categories

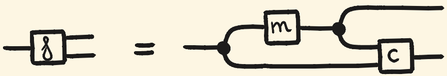
OR

- add probability to (discrete) cartesian restriction categories

COPY-DISCARD STRUCTURE



CONDITIONALS



PARTIAL FROBENIUS STRUCTURE

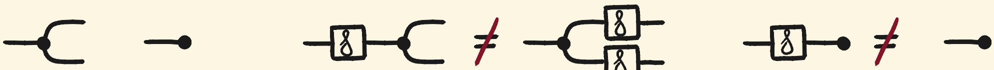
(DISCRETE) PARTIAL MARKOV CATEGORIES

- REMOVE TOTALITY FROM ~~add partiality to~~ Markov categories

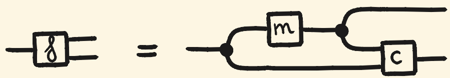
OR

- REMOVE DETERMINISM FROM ~~add probability to~~ (discrete) cartesian restriction categories

COPY-DISCARD STRUCTURE



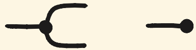
CONDITIONALS



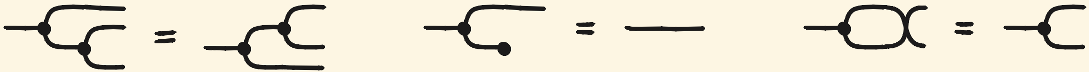
PARTIAL FROBENIUS STRUCTURE

COPY-DISCARD CATEGORIES

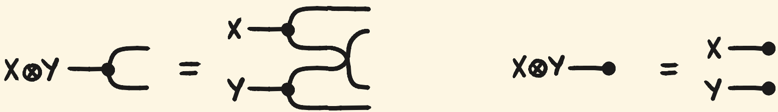
A copy-discard category is a symmetric monoidal category where every object is a uniform cocommutative comonoid.



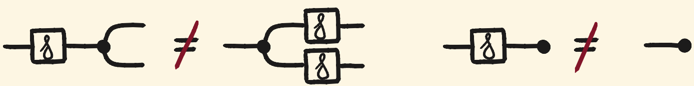
COCOMMUTATIVE COMONOID



UNIFORMITY



NO NATURALITY REQUIRED

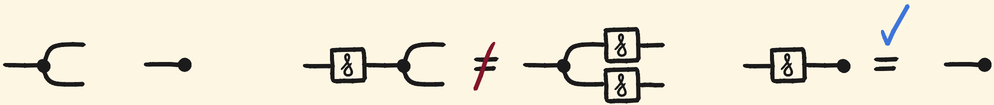


[Corradini & Gadducci 1999]

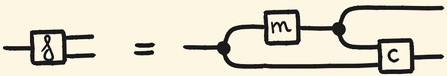
MARKOV CATEGORIES & CONDITIONALS

A Markov category with conditionals is a copy-discard category with conditionals where all morphisms are total.

COPY - DISCARD STRUCTURE



CONDITIONALS

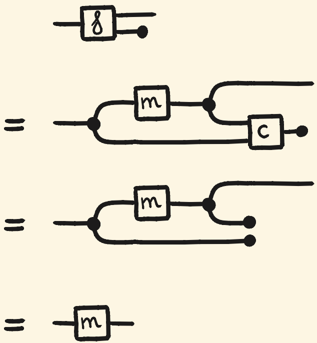


[Fritz 2020]

MARGINALS IN MARKOV CATEGORIES

Marginals in Markov categories are as expected :

PROOF



conditionals :

⇝ totality

□

DROPPING TOTALITY

We want to keep the nice marginals of Markov categories.

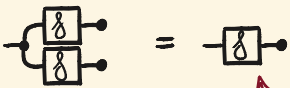
Should we ask conditionals to be total? ✗ NO

⇝ too strong: total conditionals fail to exist in $\mathrm{Kl}(\mathcal{D}_{\leq 1})$.

Can we ask conditionals to be quasi-total? ✓ YES

⇝ sweet spot: quasi-total conditionals usually exist and give nice marginals.

QUASI-TOTAL MORPHISM (in a copy-discard category)



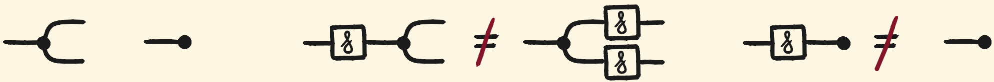
⇝ failure is deterministic

domain of definition

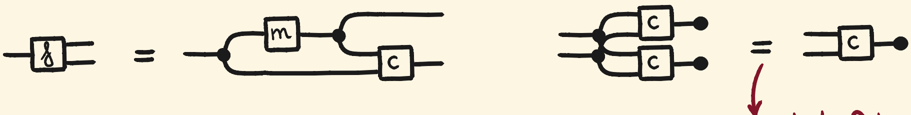
PARTIAL MARKOV CATEGORIES

A partial Markov category is a copy-discard category with quasi-total conditionals.

COPY-DISCARD STRUCTURE



CONDITIONALS

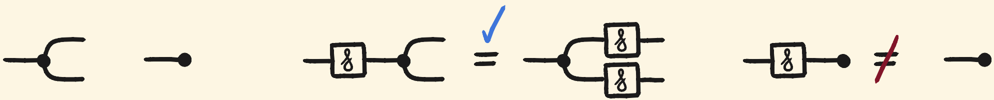


quasi-totality

CONSTRAINTS VIA PARTIAL FROBENIUS

A discrete cartesian restriction category is a copy-discard category with comparators where all morphisms are deterministic.

COPY-DISCARD STRUCTURE



PARTIAL FROBENIUS STRUCTURE

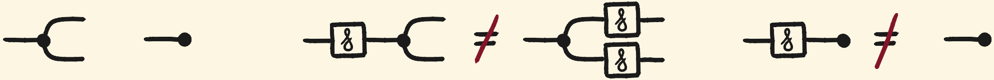
COMPARATOR

[Cockett & Lack 2003, Cockett, Guo & Hofstra 2012]

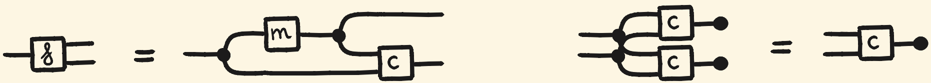
DISCRETE PARTIAL MARKOV CATEGORIES

A discrete partial Markov category is a copy-discard category with quasi-total conditionals and comparators.

COPY - DISCARD STRUCTURE



CONDITIONALS



PARTIAL FROBENIUS STRUCTURE

EXAMPLES : PARTIAL STOCHASTIC PROCESSES

A <u>partial stochastic process</u> is a stochastic process that may fail.

↳ Maybe monad

your favourite Markov category with conditionals

PROPOSITION

Partial stochastic processes form a partial Markov category.

$\begin{cases} \mathcal{C} \text{ Markov category with conditionals and coproducts} \\ \text{Some ugly technical conditions} \end{cases}$

$\Rightarrow \mathrm{Kl}(\cdot + 1)$ is a partial Markov category.

EXAMPLES

- $\mathrm{Kl}(\mathcal{D}(\cdot + 1))$ ⇝ finitary subdistributions
- $\mathrm{Kl}(\mathcal{G}iry_B(\cdot + 1))$ ⇝ subdistributions on standard Borel spaces

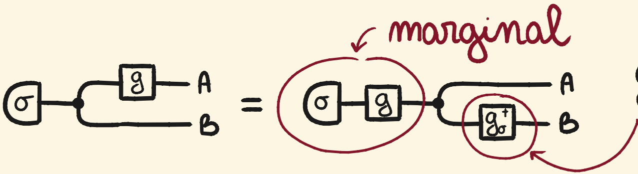
OUTLINE

- motivation: Evidential Decision Theory
- (discrete) partial Markov categories

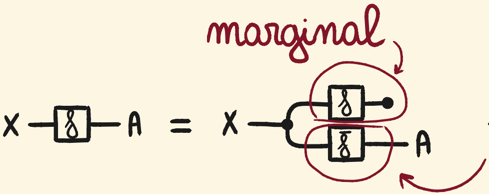
[• Bayes, observations & updates]

BAYES INVERSION & NORMALISATION

Bayes inversions and normalisations are particular cases of quasi-total conditionals:



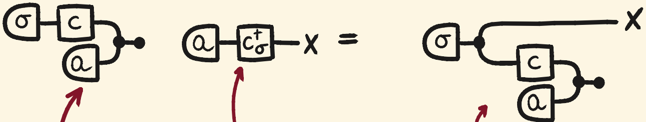
$g^\dagger_\sigma$ is a <u>Bayes inversion</u> of g w.r.t. σ



$\bar{f}$ is a <u>normalisation</u> of f

SYNTHETIC BAYES THEOREM

A deterministic observation $a : I \to A$ from a prior $\sigma : I \to X$ through a channel $c : X \to A$ determines an update proportional to the Bayes inversion $c^\dagger_\sigma$ evaluated on a.

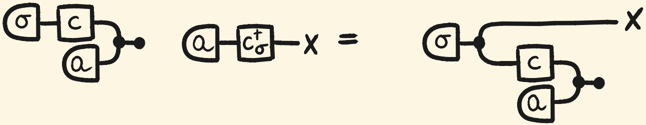


$$P(X = x \mid A = a) = \frac{P(A = a \mid X = x) \cdot P(X = x)}{\sum_{y \in X} P(A = a \mid X = y) \cdot P(X = y)}$$

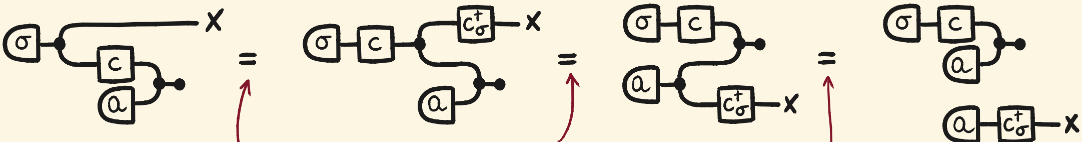
classical formula
for Bayes theorem

SYNTHETIC BAYES THEOREM

A deterministic observation $a : I \to A$ from a prior $\sigma : I \to X$ through a channel $c : X \to A$ determines an update proportional to the Bayes inversion $c^\dagger_\sigma$ evaluated on a.



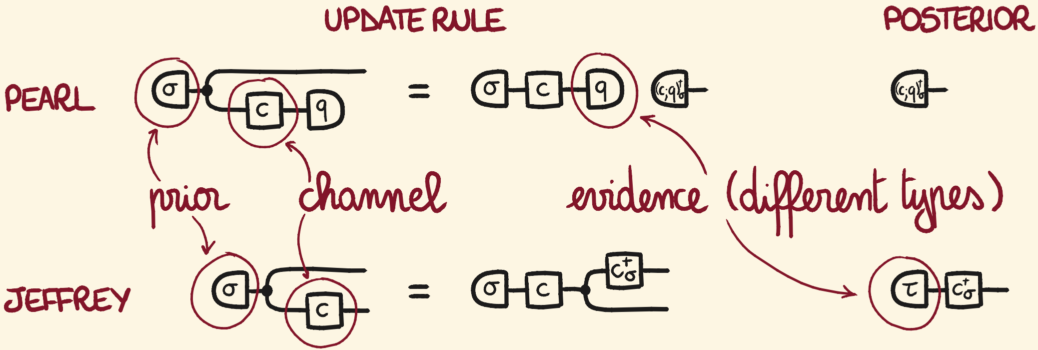
PROOF

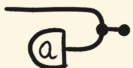
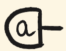
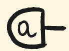


conditionals Frobenius determinism

□

PEARL'S VS JEFFREY'S UPDATES

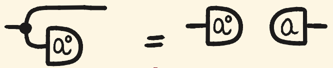


Pearl's update on  coincides with Jeffrey's update on , whenever  is deterministic.

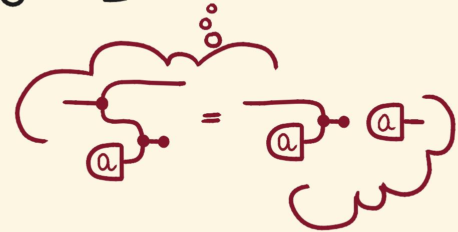
[Jacobs 2019]

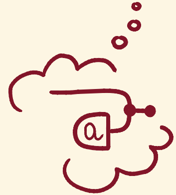
PROCESSES WITH EXACT OBSERVATIONS

We construct a partial Markov category $\mathrm{exOb}(\mathcal{C})$ on top of a Markov category $\mathcal{C}$ with conditionals by freely adding, for every deterministic state a — A in $\mathcal{C}$, a costate A — a° and quotienting by



.





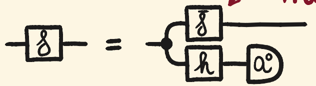
$\mathrm{exOb}(\mathcal{C}) = (\mathcal{C} + \{ A \text{—} a^\circ \mid a \text{—} A \text{ deterministic} \}) /$

embeds faithfully into ($\mathcal{C}$ +

) / partial Frobenius

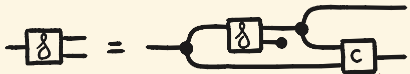
COMPUTING PROCESSES WITH EXACT OBSERVATIONS

Morphisms in exOb($\mathcal{C}$) have a normal form



normalisation of f

that can be computed by conditioning in $\mathcal{C}$,
and they have conditionals

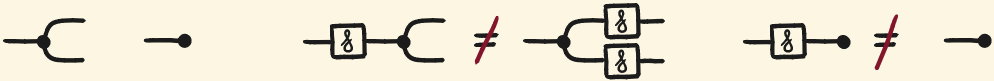


conditional of $\bar{f}$

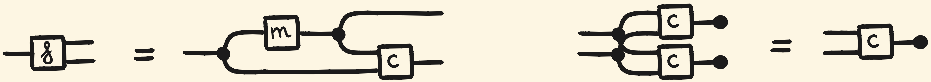
that can be computed by conditioning in $\mathcal{C}$.

SUMMARY : DISCRETE PARTIAL MARKOV CATS

COPY - DISCARD STRUCTURE



CONDITIONALS



PARTIAL FROBENIUS STRUCTURE

THANKS FOR LISTENING !