LiCS 2023 27th June 2023

EVIDENTIAL DECISION THEORY VIA PARTIAL MARKOV CATEGORIES

Elena Di Lavore and Mario Román

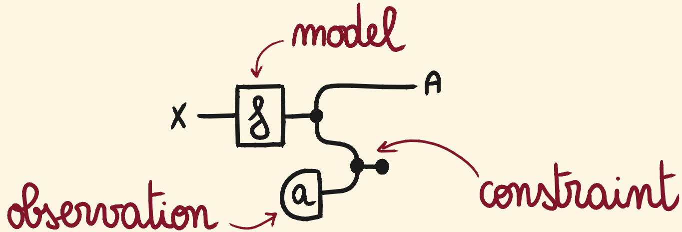
Tallinn University of Technology

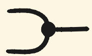
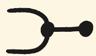
European Union
European Regional
Development Fund

Investing
in your future

PARTIALITY FOR OBSERVATIONS

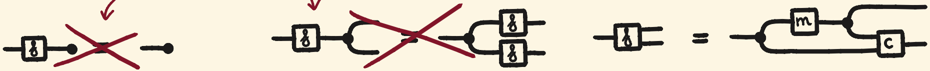
Updating a model on an observation means restricting the model to scenarios that are compatible with this observation.



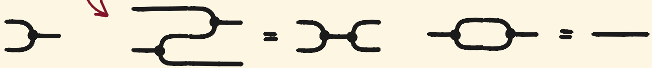
constraints  cannot be total computations because  $\neq$  .

OVERVIEW

Combine Markov and cartesian restriction categories to express partial stochastic processes.



Add the discrete structure to express equality checking.



Morphisms $X \xrightarrow{f} A$ are partial stochastic channels

$f(a|x)$ = "probability of a given x"

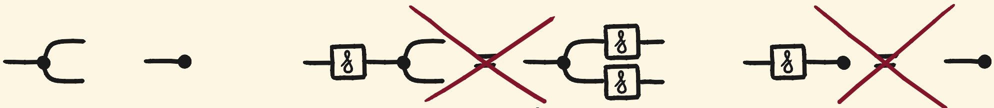
$f(\perp|x)$ = "probability of failure"

[Fritz 2020, Cockett & Lack 2007, Cockett, Guo & Hofstra 2012, Cho & Jacobs 2019]

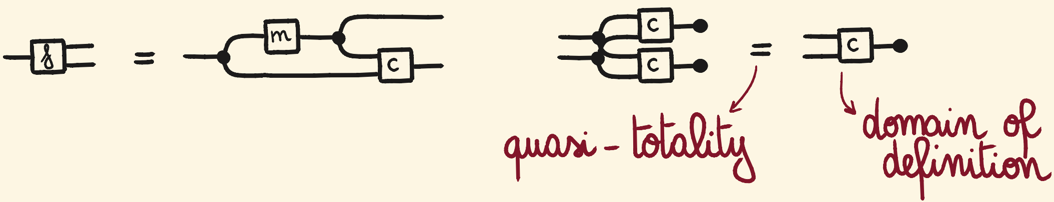
PARTIAL MARKOV CATEGORIES

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

COPY - DISCARD STRUCTURE



CONDITIONALS



EXAMPLES : PARTIAL STOCHASTIC PROCESSES

Partial stochastic processes form a partial Markov category.

maybe monad on a Markov category

THEOREM

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

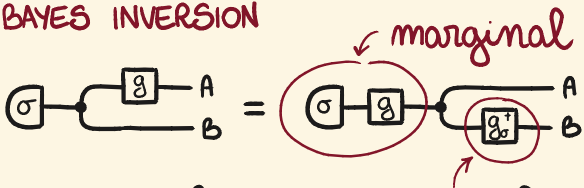
$\Rightarrow \mathrm{Kl}(\cdot + 1)$ is a partial Markov category.

EXAMPLES

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

BAYES INVERSION & NORMALISATION

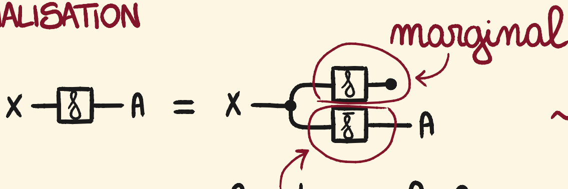
BAYES INVERSION



$\rightsquigarrow$ $g^\dagger_\sigma(b|a) := \dfrac{g(a|b)\,\sigma(b)}{\sum_{b'\in B} g(a|b')\,\sigma(b')}$

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

NORMALISATION



$\rightsquigarrow$ $\bar{f}(a|x) := \dfrac{f(x|a)}{1 - f(\perp|a)}$

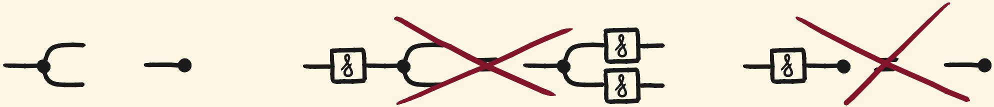
$\bar{f}$ is a normalisation of f

→ Both are particular cases of quasi-total conditionals.

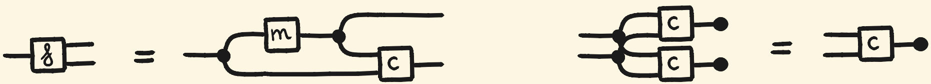
DISCRETE PARTIAL MARKOV CATEGORIES

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

COPY-DISCARD STRUCTURE



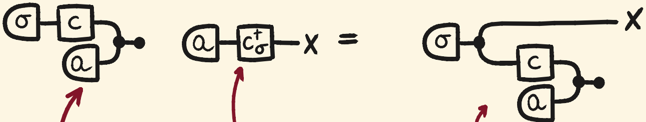
CONDITIONALS



PARTIAL FROBENIUS STRUCTURE

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.

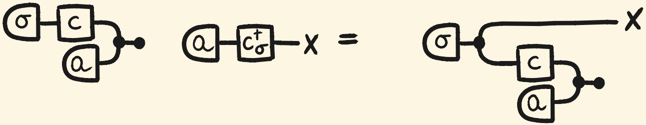


$$P(X{=}x \mid A{=}a) = \frac{P(A{=}a \mid X{=}x) \cdot P(X{=}x)}{\sum_{y \in X} P(A{=}a \mid X{=}y) \cdot P(X{=}y)}$$

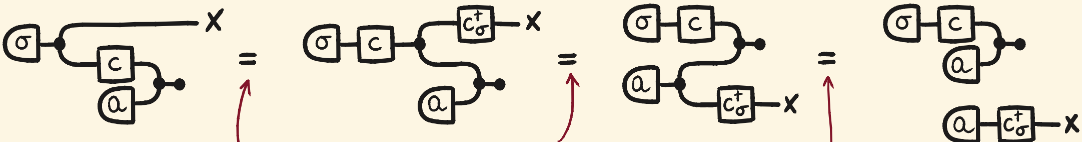
classical formula for Bayes theorem

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

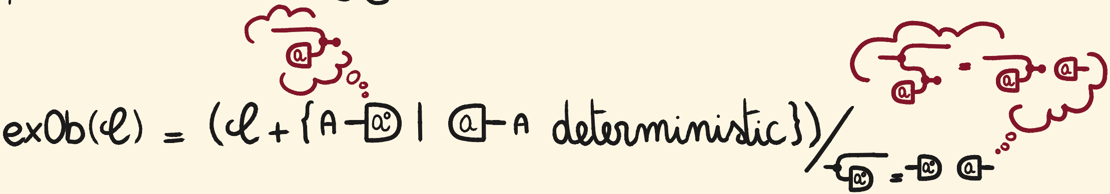
Frobenius

determinism

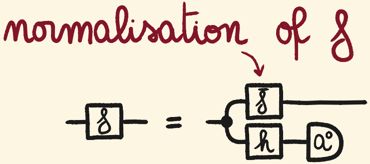
□

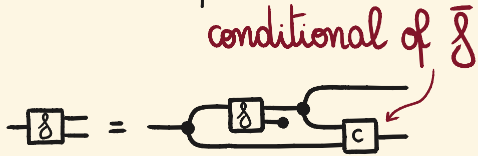
PROCESSES WITH EXACT OBSERVATIONS

For a Markov category $\mathscr{C}$ with conditionals, we construct a partial Markov category $\mathrm{exOb}(\mathscr{C})$:



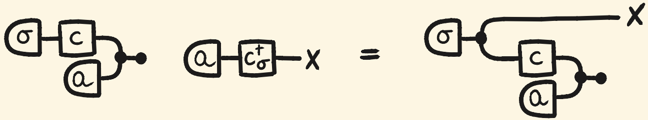
Conditionals and normalisations are computed in $\mathscr{C}$





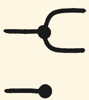
SUMMARY

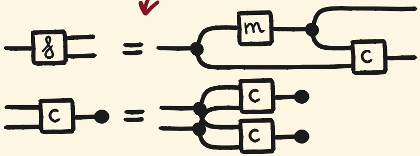
Discrete partial Markov categories express stochastic processes with observations and updates.

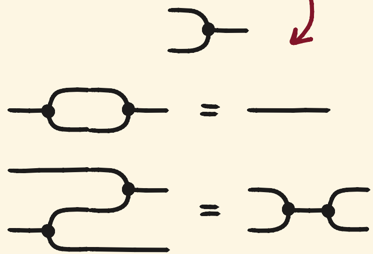


Synthetic Bayes theorem

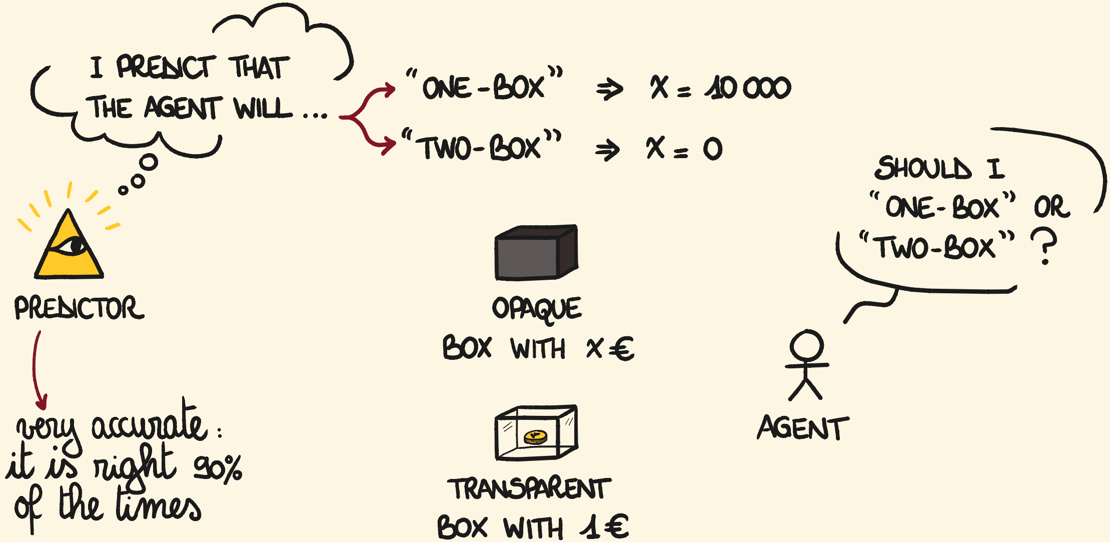
They are copy-discard categories with conditionals and comparators.







NEWCOMB'S PROBLEM



EVIDENTIAL DECISION THEORY

Evidential decision theory answers :

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

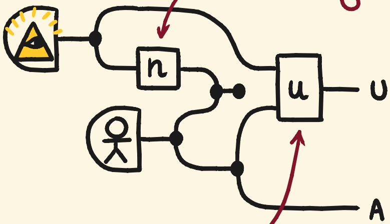
My action is evidence for the prediction :
if I one-box I expect 10 000 €,
if I two-box I expect 1 € .
⇒ I will one-box

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

MOST LIKELY

NEWCOMB'S PROBLEM CATEGORICALLY

noise: the predictor is right 90% of the times



utility function

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

SOLVING NEWCOMB'S PROBLEM

Evidential decision theory asks:
"Which action would be evidence for the best-case scenario?"
i.e. "Which action maximises the average of the state below?"