

LiCS 2023

27th June 2023

EVIDENTIAL DECISION THEORY VIA PARTIAL MARKOV CATEGORIES

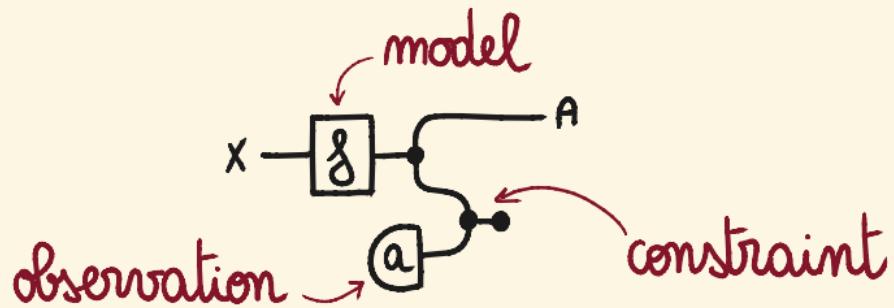
Elena Di Stefano and Mario Román

Tallinn University of Technology



PARTIALITY FOR OBSERVATIONS

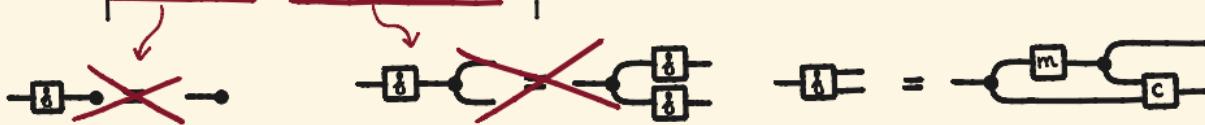
Updating a model on an observation means restricting the model to scenarios that are compatible with this observation.



constraints \rightarrow cannot be total computations
because $\neq \equiv$.

OVERVIEW

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



Add the discrete structure to express equality checking.



Morphisms $x - \boxed{\delta} - A$ are partial stochastic channels

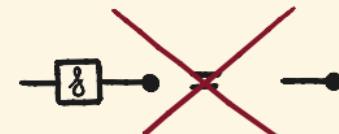
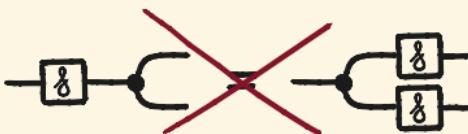
$\delta(a|x)$ = "probability of a given x "

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

PARTIAL MARKOV CATEGORIES

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

COPY - DISCARD STRUCTURE



CONDITIONALS

$$\boxed{\delta} = \text{---} \circlearrowleft \text{---} \circlearrowright \text{---}$$

$$\text{---} \circlearrowleft \text{---} \circlearrowright \text{---} = \text{---} \circlearrowleft \text{---} \circlearrowright \text{---}$$

quasi-totality

domain of definition

EXAMPLES : PARTIAL STOCHASTIC PROCESSES

Partial stochastic processes form a partial Markov category.
↓
maybe monad on a Mäkrov category

THEOREM

{ cf Markov category with conditionals and coproducts
{ some ugly technical conditions
⇒ $\text{Kl}(\cdot + 1)$ is a partial Markov category.

EXAMPLES

- $\text{Kl}(\mathcal{D}(\cdot + 1))$ → finitary subdistributions
- $\text{Kl}(\text{dgiry}_{\mathbb{B}}(\cdot + 1))$ → subdistributions on standard Borel spaces

BAYES INVERSION & NORMALISATION

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

$$\text{Diagram: } \sigma \xrightarrow{g} A \quad B = \text{Diagram: } \sigma \xrightarrow{g} \text{circle} \xrightarrow{g^+} A \quad B \quad \text{marginal}$$
$$g_\sigma^+(b|a) := \frac{g(a|b)\sigma(b)}{\sum_{b' \in B} g(a|b')\sigma(b')}$$

g^+ is a Bayes inversion of g w.r.t. σ

$$\text{Diagram: } x \xrightarrow{g} A = \text{Diagram: } x \xrightarrow{g} \text{circle} \xrightarrow{g^-} A \quad \text{marginal}$$
$$g^-(a|x) := \frac{g(x|a)}{1 - g(\perp|a)}$$

\bar{g} is a normalisation of g

DISCRETE PARTIAL MARKOV CATEGORIES

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

COPY - DISCARD STRUCTURE



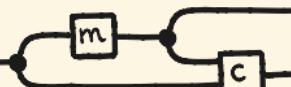
-



CONDITIONALS



=



PARTIAL FROBENIUS STRUCTURE



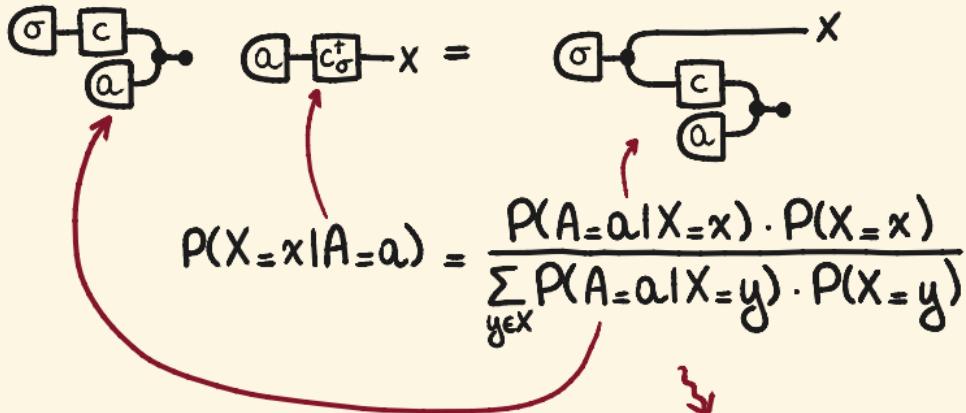
= —



COMPARATOR

SYNTHETIC BAYES THEOREM

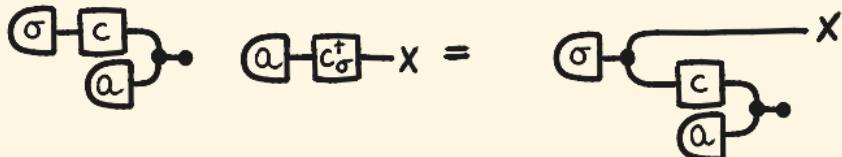
A deterministic observation $a: I \rightarrow A$ from a prior $\sigma: I \rightarrow X$ through a channel $c: X \rightarrow A$ determines an update proportional to the Bayes inversion c_σ^+ evaluated on a .



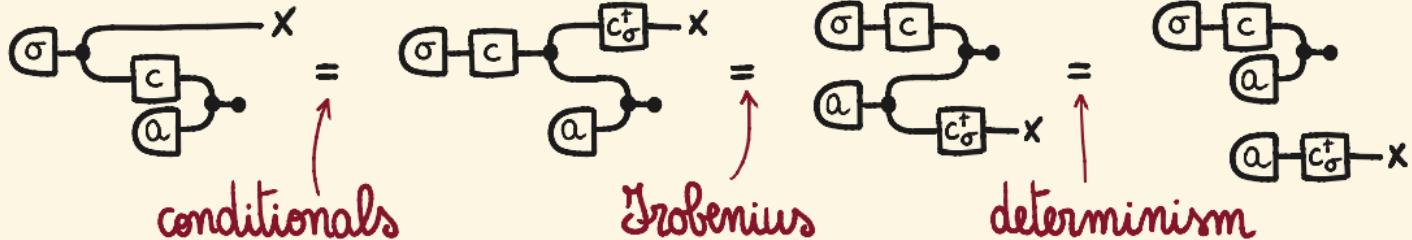
classical formula
for Bayes theorem

SYNTHETIC BAYES THEOREM

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



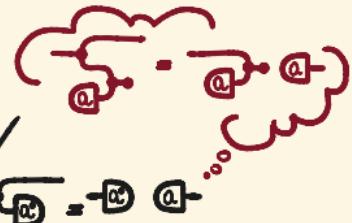
PROOF



□

PROCESSES WITH EXACT OBSERVATIONS

For a Markov category \mathcal{C} with conditionals, we construct a partial Markov category $\text{exOb}(\mathcal{C})$.



$$\text{exOb}(\mathcal{C}) = (\mathcal{C} + \{\mathbf{A} - \square | \square - \mathbf{A} \text{ deterministic}\}) / \text{partial Grobenius}$$

embeds faithfully into $(\mathcal{C} + \rightarrow) / \text{partial Grobenius}$

Conditionals and normalisations are computed in \mathcal{C}

normalisation of \mathbf{g}

$$-\square \mathbf{g} = - \bullet \begin{array}{c} \square \mathbf{g} \\ \square h \\ \square a \end{array}$$

conditional of $\bar{\mathbf{g}}$

$$-\square \bar{\mathbf{g}} = - \bullet \begin{array}{c} \square \mathbf{g} \\ \square c \end{array}$$

SUMMARY

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

$$\text{Diagram showing the composition of two morphisms: } \sigma \xrightarrow{c} a \text{ and } a \xrightarrow{c^+} x \text{ resulting in } \sigma \xrightarrow{c} x.$$

Synthetic Bayes theorem

They are copy-discard categories with conditionals and comparators.

$$\begin{array}{ccc} \text{copy} & \text{conditional} & \text{discard} \\ \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} \end{array}$$

Diagrams illustrating the properties of the copy, conditional, and discard operations:

- The copy operation (δ) is shown as a horizontal line with a square box labeled δ above it. It is equated to a morphism that first copies the input into a box labeled m , then splits it into two paths, one leading to a box labeled c and the other being discarded.
- The conditional operation (c) is shown as a horizontal line with a square box labeled c above it. It is equated to a morphism that splits the input into two paths, both entering a box labeled c and then merging back into a single output line.
- The discard operation (---) is shown as a horizontal line with a square box below it. It is equated to a morphism that discards the input.
- The identity operation (---) is shown as a horizontal line with a square box below it. It is equated to a morphism that does nothing to the input.
- The inverse operation (---) is shown as a horizontal line with a square box below it. It is equated to a morphism that inverts the input.

NEWCOMB'S PROBLEM

I PREDICT THAT
THE AGENT WILL ...

"ONE-BOX" $\Rightarrow X = 10\ 000$
"TWO-BOX" $\Rightarrow X = 0$



PREDICTOR

very accurate:
it is right 90%
of the times



OPAQUE
BOX WITH $X \in$



TRANSPARENT
BOX WITH 1€

SHOULD I
"ONE-BOX" OR
"TWO-BOX" ?



AGENT

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€ .

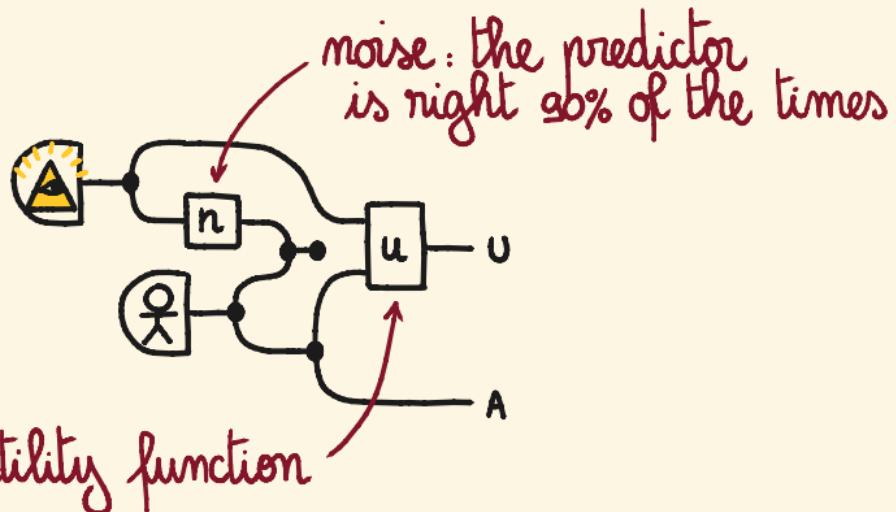
⇒ I will one-box

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

MOST LIKELY



NEWCOMB'S PROBLEM CATEGORICALLY



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?"

