

LiCS 2023

27<sup>th</sup> June 2023

# EVIDENTIAL DECISION THEORY VIA PARTIAL MARKOV CATEGORIES

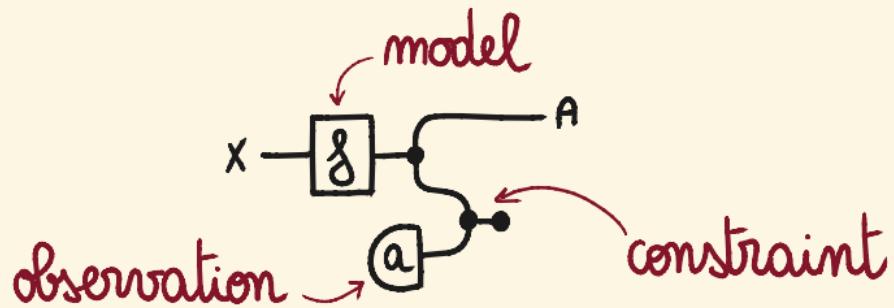
Elena Di Stefano and Mario Román

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.



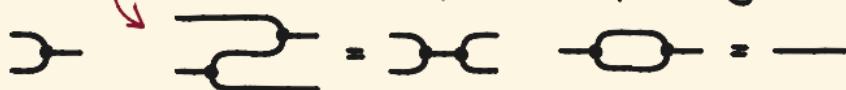
constraints  $\rightarrow$  cannot be total computations  
because  $\neq \equiv$ .

## OVERVIEW

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



Add the discrete structure to express equality checking.



Morphisms  $x - \delta - A$  are partial stochastic channels

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

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

[Trifunovic 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

---

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

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

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 INVERSION

$$\text{original DAG: } \sigma \rightarrow g \rightarrow A = \text{transformed DAG: } \sigma \rightarrow g \rightarrow \text{marginal oval} \rightarrow g^+ \rightarrow A$$
$$g_\sigma^+(b|a) := \frac{g(a|b)\sigma(b)}{\sum_{b' \in B} g(a|b')\sigma(b')}$$

$g_\sigma^+$  is a Bayes inversion of  $g$  w.r.t.  $\sigma$

## NORMALISATION

$$\text{original DAG: } x \rightarrow f \rightarrow A = \text{transformed DAG: } x \rightarrow f \rightarrow \text{marginal oval} \rightarrow f^- \rightarrow A$$
$$\bar{f}(a|x) := \frac{f(x|a)}{1 - f(\perp|a)}$$

$\bar{f}$  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

---

$$\text{---} \otimes \text{---} = \text{---} \xrightarrow{m} \text{---} \xrightarrow{c} \text{---}$$

$$\text{---} \xrightarrow{c} \text{---} = \text{---} \xrightarrow{c} \text{---}$$

---

## PARTIAL FROBENIUS STRUCTURE

---

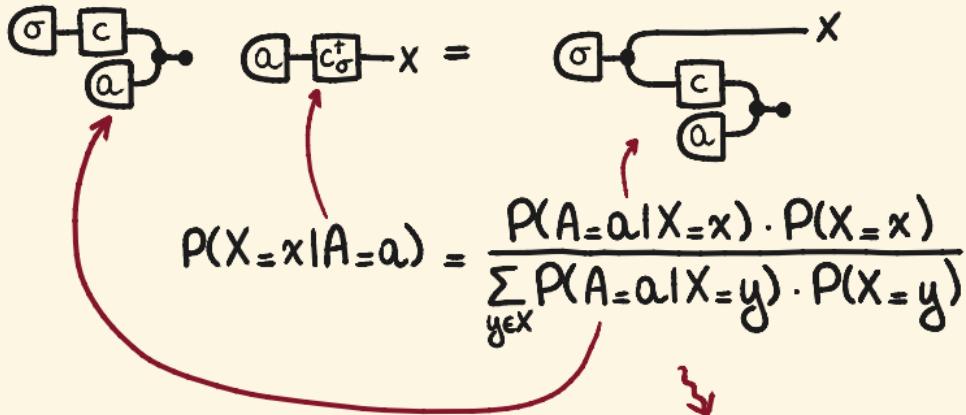

$$= \text{---}$$


$$= \text{---}$$

**COMPARATOR**

# SYNTHETIC BAYES THEOREM

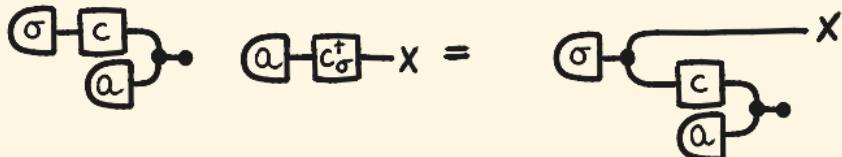
A deterministic observation  $a: I \rightarrow A$  from a prior  $\sigma: I \rightarrow X$  through a channel  $c: X \rightarrow A$  determines an update proportional to the Bayes inversion  $c_\sigma^+$  evaluated on  $a$ .



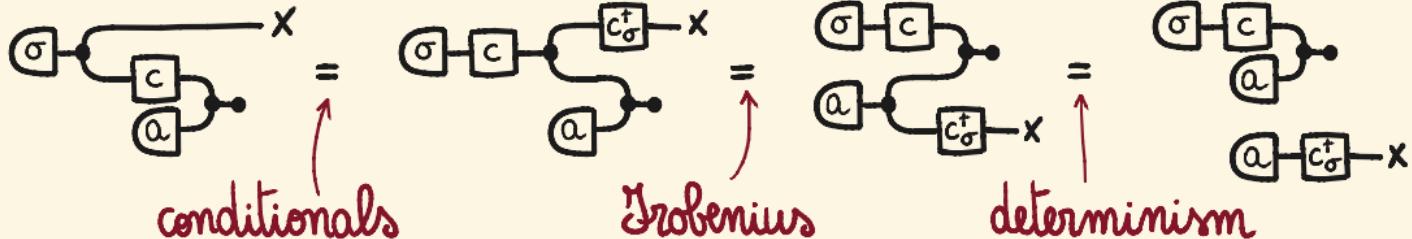
classical formula  
for Bayes theorem

# 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  $c_\sigma^+$  evaluated on  $a$ .



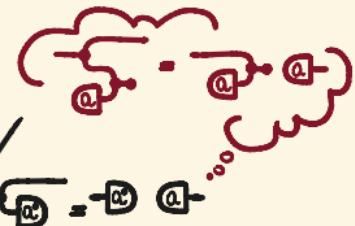
PROOF



□

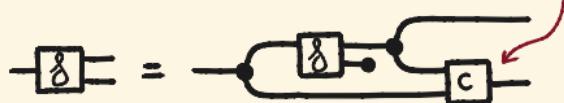
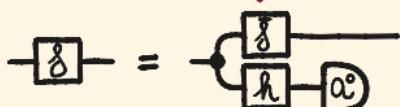
# PROCESSES WITH EXACT OBSERVATIONS

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



$$\text{exOb}(\mathcal{C}) = (\mathcal{C} + \{\mathbf{A} - \square\mid \square \rightarrow \mathbf{A} \text{ deterministic}\}) /$$

conditionals and normalisations are computed in  $\mathcal{C}$   
normalisation of  $\mathbf{g}$  conditional of  $\bar{\mathbf{g}}$



# SUMMARY

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

$$\text{Diagram showing the equivalence of two circuit expressions:} \\ \text{Left: } \sigma \xrightarrow{\text{c}} a \quad a \xrightarrow{\text{c}^+} x \\ \text{Right: } \sigma \xrightarrow{\text{c}} a \quad a \xrightarrow{x} x \\ \text{Equivalence: } \sigma \xrightarrow{\text{c}} a = \sigma \xrightarrow{\text{c}} a \quad a \xrightarrow{x} x$$

Synthetic Bayes theorem

They are copy-discard categories with conditionals and comparators.

$$\begin{array}{ccc} \text{Conditionals:} & \text{Comparators:} & \text{Discard:} \\ \text{Copy: } \text{---} & \text{---} & \text{---} \\ \text{Discard: } \text{---} & \text{---} & \text{---} \\ \text{Conditional: } \text{---} & \text{---} & \text{---} \\ \text{Combiner: } \text{---} & \text{---} & \text{---} \end{array}$$

Arrows indicate the relationships between the components:

- A red bracket groups the first two columns (Conditionals and Comparators) under the heading "copy-discard".
- A red bracket groups the last two columns (Comparators and Discard) under the heading "conditionals and comparators".
- A red arrow points from the "copy-discard" group to the "Conditional" component.
- A red arrow points from the "copy-discard" group to the "Combiner" component.
- A red arrow points from the "conditionals and comparators" group to the "Discard" component.

# 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€ .

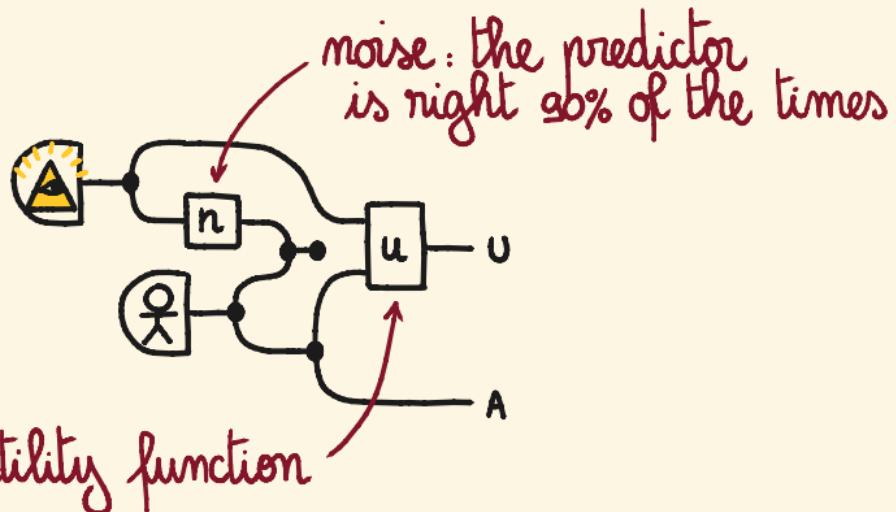
⇒ I will one-box

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?"

