

Stata Fest

27 April 2023

EVIDENTIAL DECISION THEORY VIA PARTIAL MARKOV CATEGORIES

Elena Di Stefano and Mario Román

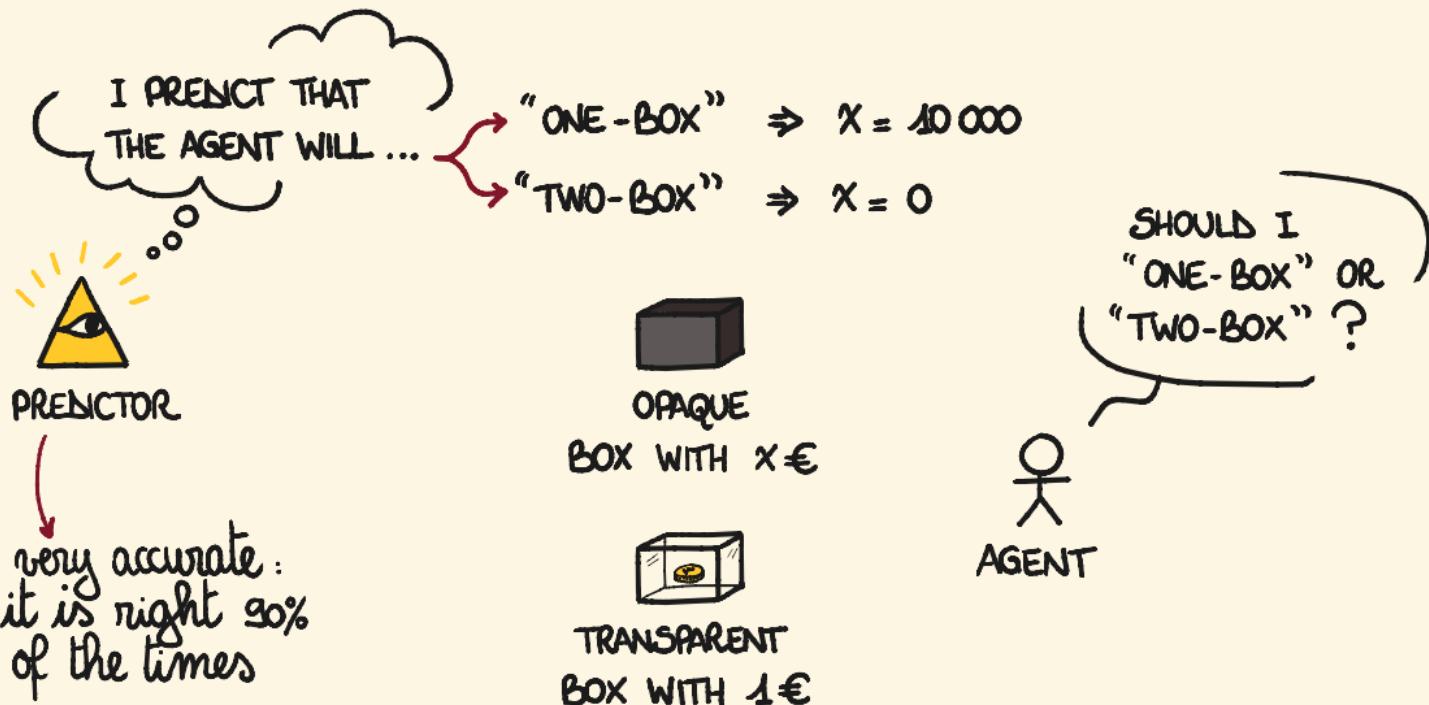
Tallinn University of Technology



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?”

Whatever the predictor did,
I get 1€ extra if I two-box
⇒ I will two-box



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

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

MOST LIKELY



EVIDENTIAL VS CAUSAL DECISION THEORY



CAUSAL
DECISION
THEORIST



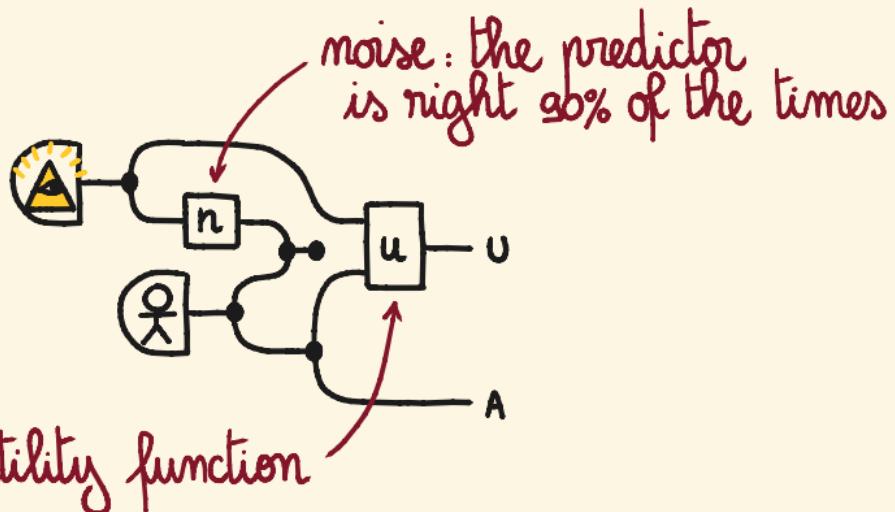
EVIDENTIAL
DECISION
THEORIST

EXPECTED
UTILITY

$$\begin{aligned} & 0.9 \times 1 \text{ €} \\ & + 0.1 \times 10\,001 \text{ €} \\ & = 1\,001 \text{ €} \end{aligned}$$

$$\begin{aligned} & 0.9 \times 10\,000 \text{ €} \\ & + 0.1 \times 0 \text{ €} \\ & = 9\,000 \text{ €} \end{aligned}$$

A DIAGRAM FOR NEWCOMB



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

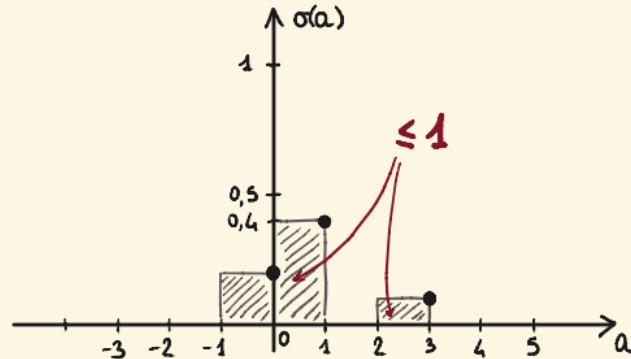
SUBDISTRIBUTIONS

A subdistribution σ on A is a distribution on $A+1$:

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

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

ex $A = \mathbb{N}, \sigma =$



Subdistributions give a monad $\mathcal{D}_{\leq 1}: \text{Set} \rightarrow \text{Set}$
(there's a distributive law $\mathcal{D}(-)+1 \rightarrow \mathcal{D}(-+1)$)

SUBDISTRIBUTIONS

A morphism $x \dashv \vdash A$ in $\text{Kl}\mathcal{D}_{\leq 1}$ is a function $X \rightarrow \mathcal{D}_{\leq 1}(A)$

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

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

composition is:

$$x \dashv \vdash g \dashv \vdash B \quad (b|x) := \sum_{a \in A} f(a|x) \cdot g(b|a)$$

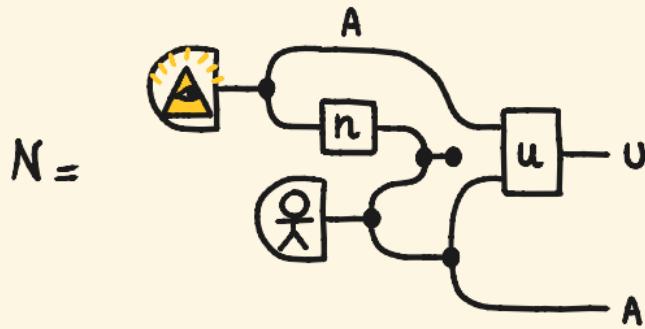
$$x \dashv \vdash g \dashv \vdash B \quad (\perp|x) := f(\perp|x) + \sum_{a \in A} f(a|x) \cdot g(\perp|a)$$

Equality checks are lifted from partial functions:

$$\begin{array}{c} A \\ \dashv \vdash \\ A \end{array} \quad (a|a_1, a_2) := \begin{cases} 1 & a_1 = a_2 = a \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{array}{c} A \\ \dashv \vdash \\ A \end{array} \quad (\perp|a_1, a_2) := \begin{cases} 0 & a_1 = a_2 \\ 1 & a_1 \neq a_2 \end{cases}$$

READING NEWCOMB'S DIAGRAM



$$N(v, a | *) = \sum_{p \in A} \Delta(p | *) \cdot n(a | p) \cdot \xi(a | *) \cdot u(v | p, a)$$

$$N(\perp | *) = \sum_{a \neq a' \in A} \sum_{p \in A} \Delta(p | *) \cdot n(a' | p) \cdot \xi(a | *)$$

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

1. stochastic processes
2. partiality & constraints

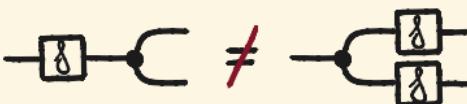
OUTLINE

- 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

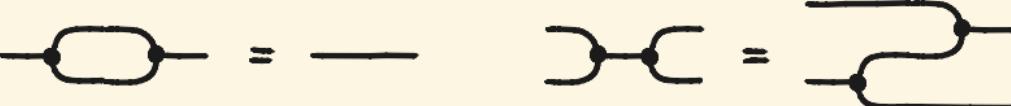
COPY - DISCARD STRUCTURE

$$\text{---} \quad \text{---} = \quad \text{---} \quad \text{---} \neq \quad \text{---} \quad \text{---} \neq \quad \text{---}$$


CONDITIONALS

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


PARTIAL FROBENIUS STRUCTURE

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


(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

$$\text{---} \quad \text{---} = \quad \text{---} \quad \text{---} \neq \quad \text{---} \quad \text{---} \neq \quad \text{---}$$

CONDITIONALS

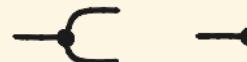
$$\text{---} \quad = \quad \text{---} \quad \text{---}$$

PARTIAL FROBENIUS STRUCTURE

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

COPY-DISCARD CATEGORIES

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



COCOMMUTATIVE COMONOID

A diagram showing three configurations of three lines meeting at a central point, connected by an equals sign. The first configuration has lines entering from the left and exiting to the right. The second configuration has lines entering from the left and exiting to the right, with a twist. The third configuration has lines entering from the left and exiting to the right, with a twist.

A diagram showing a configuration of three lines meeting at a central point, connected by an equals sign, followed by a single line segment.

A diagram showing a configuration of three lines meeting at a central point, connected by an equals sign, followed by a single line segment.

UNIFORMITY

A diagram showing a configuration of four lines (labeled x and y) entering and exiting, connected by an equals sign, followed by a configuration where x and y enter through separate lines and exit together.

A diagram showing a configuration of two lines (labeled x and y) entering, connected by an equals sign, followed by two separate lines (x and y).

NO NATURALITY REQUIRED

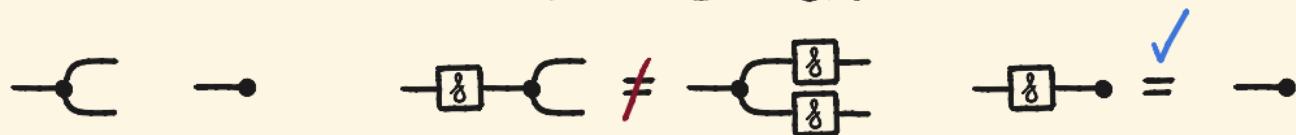
A diagram showing two configurations involving a square box labeled delta, connected by a not-equals sign.

A diagram showing two configurations involving a square box labeled delta, connected by a not-equals sign.

MARKOV CATEGORIES & CONDITIONALS

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

COPY - DISCARD STRUCTURE



CONDITIONALS



MARGINALS IN MARKOV CATEGORIES

Marginals in Markov categories are as expected :

$$x - \boxed{m} - A = x - \boxed{\delta} - \begin{matrix} A \\ B \end{matrix}$$

PROOF

$$\begin{aligned} & \text{---} \boxed{\delta} \text{ ---} \\ = & \text{---} \bullet \text{---} \boxed{m} \text{---} \bullet \text{---} \boxed{c} \text{---} \bullet \text{---} \\ = & \text{---} \bullet \text{---} \boxed{m} \text{---} \bullet \text{---} \end{aligned}$$

$$\begin{aligned} & \text{conditionals :} \\ \rightsquigarrow & \text{---} \boxed{\delta} \text{ ---} = \text{---} \bullet \text{---} \boxed{m} \text{---} \bullet \text{---} \boxed{c} \text{---} \bullet \text{---} \\ & \rightsquigarrow \text{totality} \end{aligned}$$

□

DROPPING TOTALITY

We want to keep the nice marginals of Markov categories.

Should we ask conditionals to be total ? X NO

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

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

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

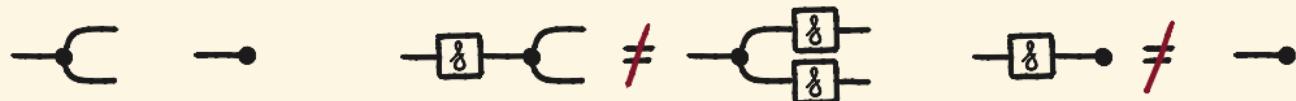
QUASI-TOTAL MORPHISM (in a copy-discard category)

$$\begin{array}{c} \text{Diagram: } \text{copy} \rightarrow \text{discard} \\ \text{Left: } \text{copy} \rightarrow \text{copy} \times \text{copy} \\ \text{Right: } \text{copy} \times \text{copy} \rightarrow \text{copy} \\ \text{Equation: } \text{copy} \rightarrow \text{copy} \times \text{copy} = \text{copy} \times \text{copy} \rightarrow \text{copy} \\ \text{Conclusion: failure is deterministic} \\ \text{Domain of definition} \end{array}$$

PARTIAL MARKOV CATEGORIES

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

COPY - DISCARD STRUCTURE



CONDITIONALS

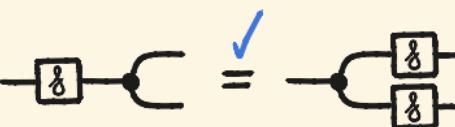


quasi-totality

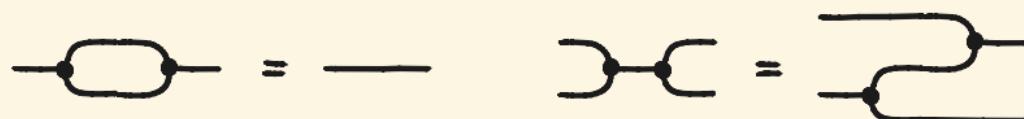
CONSTRAINTS VIA PARTIAL FROBENIUS

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

COPY - DISCARD STRUCTURE

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


PARTIAL FROBENIUS STRUCTURE

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


↑
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

$$\text{---} \bullet \text{---} = \text{---} \boxed{\delta} \text{---} \bullet \text{---} \neq \text{---} \bullet \text{---} \boxed{\delta} \text{---} \bullet \text{---} \neq \text{---} \bullet \text{---}$$

CONDITIONALS

$$\text{---} \boxed{\delta} \text{---} = \text{---} \bullet \text{---} \boxed{m} \text{---} \bullet \text{---} \boxed{c} \text{---} \quad \text{---} \bullet \text{---} \boxed{c} \text{---} \bullet \text{---} = \text{---} \bullet \text{---} \boxed{c} \text{---} \bullet \text{---}$$

PARTIAL FROBENIUS STRUCTURE

$$\text{---} \bullet \text{---} = \text{---} \quad \text{---} \bullet \text{---} = \text{---} \quad \text{---} \bullet \text{---} = \text{---} \quad \text{---} \bullet \text{---} = \text{---}$$

COMPARATOR

EXAMPLES : PARTIAL STOCHASTIC PROCESSES

A partial stochastic process is a stochastic process
that may fail.

↳ Maybe monad

your favourite Markov category
with conditionals

PROPOSITION

Partial stochastic processes form a partial Markov category.
{ 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
- $\text{Kl}(\mathcal{C}\mathbf{iry}_S(\cdot + 1))$ ↳ subdistributions on standard Borel spaces

OUTLINE

- motivation: Evidential Decision Theory

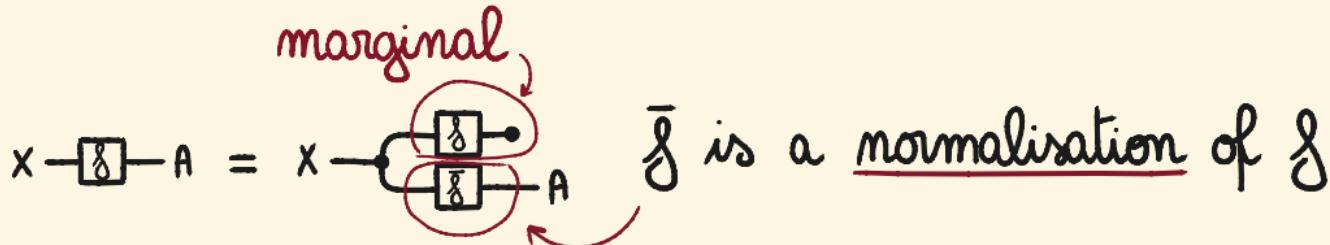
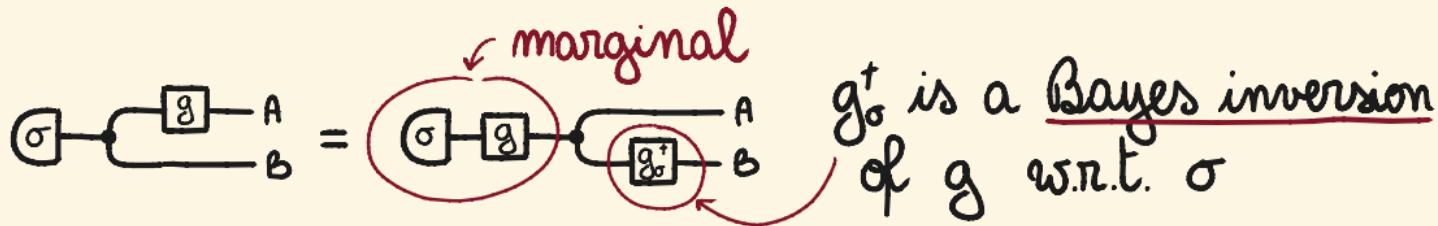
- (discrete) partial Markov categories

- Bayes, observations & updates

]

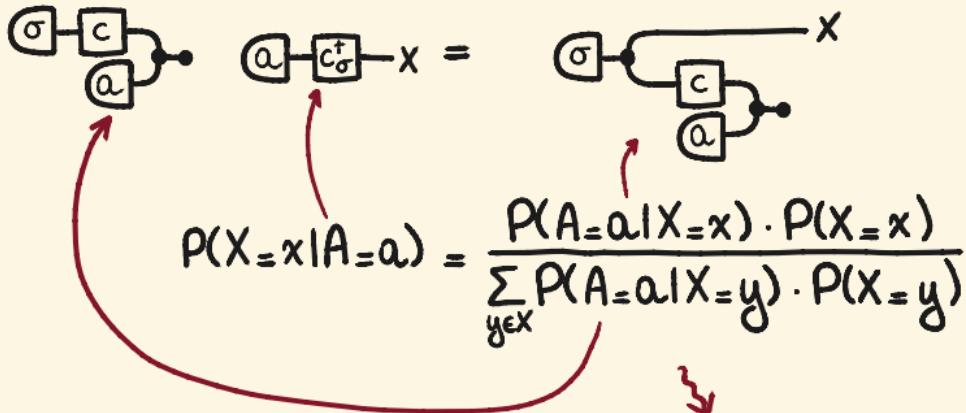
BAYES INVERSION & NORMALISATION

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



SYNTHETIC BAYES THEOREM

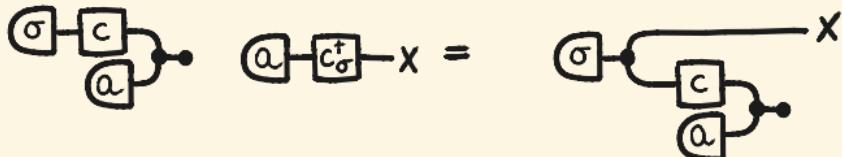
A deterministic observation $a: I \rightarrow A$ from a prior $\sigma: I \rightarrow X$ through a channel $c: X \rightarrow A$ determines an update proportional to the Bayes inversion c_σ^+ evaluated on a .



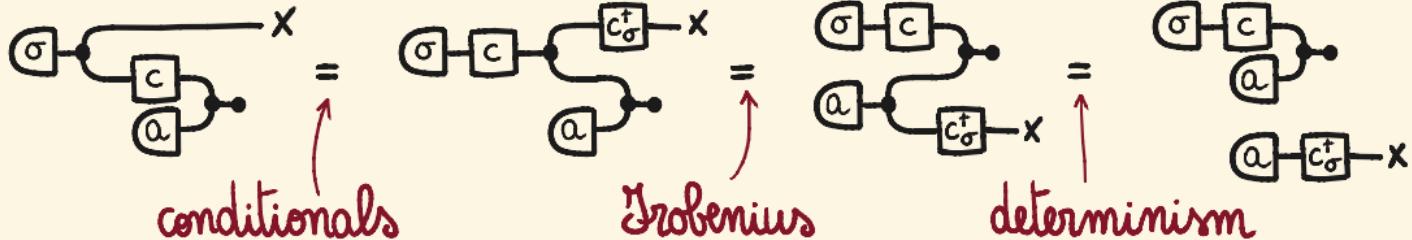
classical formula
for Bayes theorem

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 .

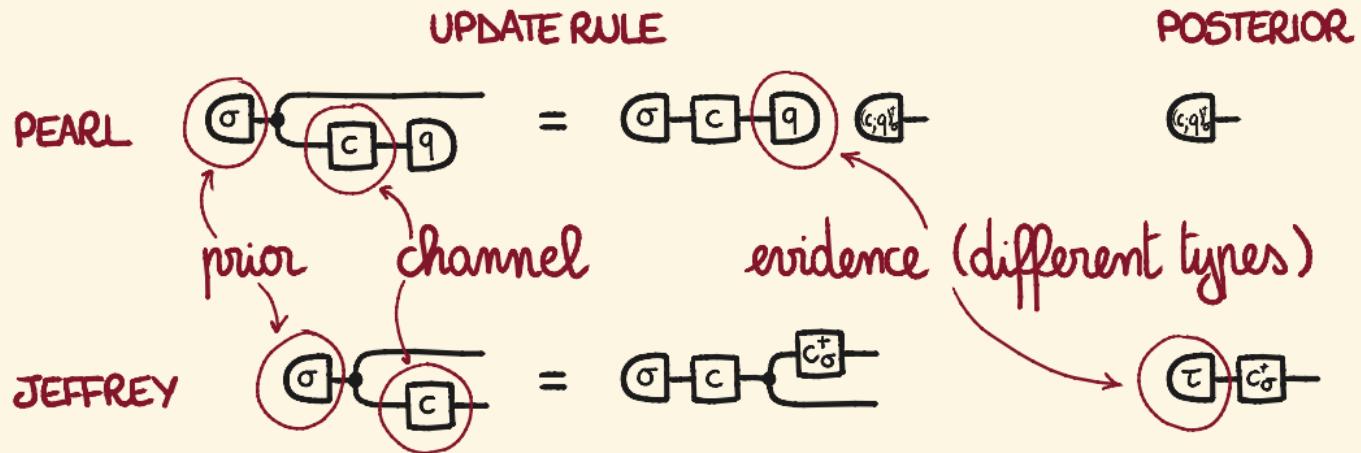


PROOF



□

PEARL'S VS JEFFREY'S UPDATES



Pearl's update on  coincides with Jeffrey's update on 

[Jacobs 2019]

PROCESSES WITH EXACT OBSERVATIONS

We construct a partial Markov category $\text{exOb}(\mathcal{C})$ on top of a Markov category \mathcal{C} with conditionals by freely adding, for every deterministic state @-A in \mathcal{C} , a costate $A-\text{@}^\circ$ and quotienting by $\overline{-\text{@}^\circ} = -\text{@}^\circ \text{@-}$.

$$\overline{\text{@}} = \overline{\text{@}} \cdot \overline{\text{@-}}$$

$$\text{exOb}(\mathcal{C}) = (\mathcal{C} + \{A-\text{@}^\circ \mid \text{@-A deterministic}\}) / \overline{-\text{@}^\circ} = -\text{@}^\circ \text{@-}$$

↑ embeds faithfully into $(\mathcal{C} + \mathbb{I}) / \begin{matrix} \text{partial} \\ \text{Frobenius} \end{matrix}$

COMPUTING PROCESSES WITH EXACT OBSERVATIONS

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

$$\boxed{\delta} = \begin{array}{c} \text{normalisation of } \delta \\ \xrightarrow{\quad} \end{array} \begin{array}{c} \text{graph} \\ \text{with} \\ \text{nodes} \end{array}$$

The diagram illustrates the process of normalisation. On the left is a box containing the symbol δ . An equals sign follows, and to its right is a box containing the text "normalisation of δ ". An arrow points from this text box to a graph on the right. The graph consists of three rectangular nodes arranged vertically. The top node contains the letter g , the middle node contains the letter h , and the bottom node contains the letters a° . Arrows connect the nodes: one from g to h , another from h to a° , and a third from g directly to a° .

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

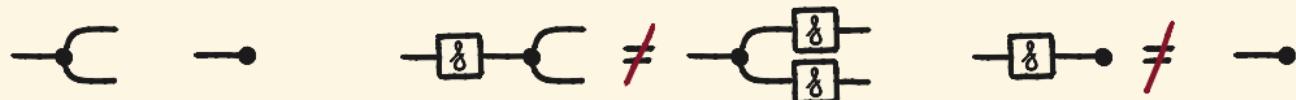
$$\boxed{\delta} = \begin{array}{c} \text{conditional of } \delta \\ \xrightarrow{\quad} \end{array} \begin{array}{c} \text{graph} \\ \text{with} \\ \text{nodes} \end{array}$$

The diagram illustrates a conditional construction. On the left is a box containing the symbol δ . An equals sign follows, and to its right is a box containing the text "conditional of δ ". An arrow points from this text box to a graph on the right. The graph consists of three rectangular nodes arranged vertically. The top node contains the letter g , the middle node contains the letter h , and the bottom node contains the letter c . Arrows connect the nodes: one from g to h , another from h to c , and a third from g directly to c .

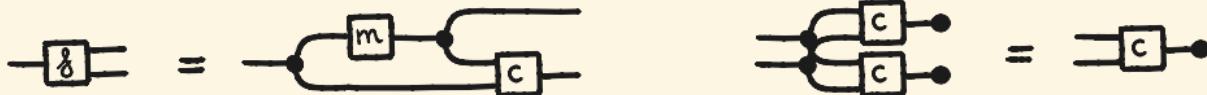
that can be computed by conditioning in \mathcal{C} .

SUMMARY : DISCRETE PARTIAL MARKOV CATS

COPY - DISCARD STRUCTURE



CONDITIONALS



PARTIAL FROBENIUS STRUCTURE



THANKS FOR LISTENING !