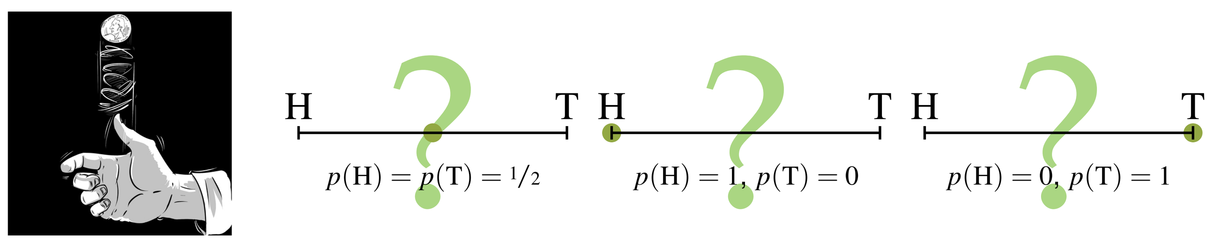
Choice functions as a tool to model uncertainty

Arthur Van Camp and **Gert de Cooman** {Arthur.VanCamp,Gert.deCooman}@UGent.be
SYSTeMS research group, Ghent University, Belgium
UNIVERSITEIT GENT

1. Introduction

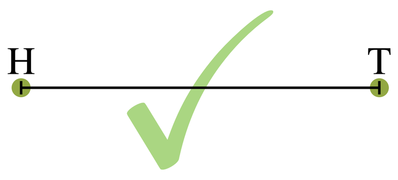
Context A random variable X takes values in a finite possibility space $\mathscr{X}$. An agent has beliefs about X. Such beliefs are commonly expressed using *(probability) mass functions*. However, such mass functions may *not* be *expressive* enough.

Motivating example We flip a coin with *identical sides of unknown type*: either twice heads or twice tails.



$\mathscr{X} = \{\mathrm{H}, \mathrm{T}\}$

There is no mass function that expresses this elementary belief. What we want is a more expressive model that can represent the stated belief, being an *OR* statement. This belief model should resemble the situation depicted on the right.



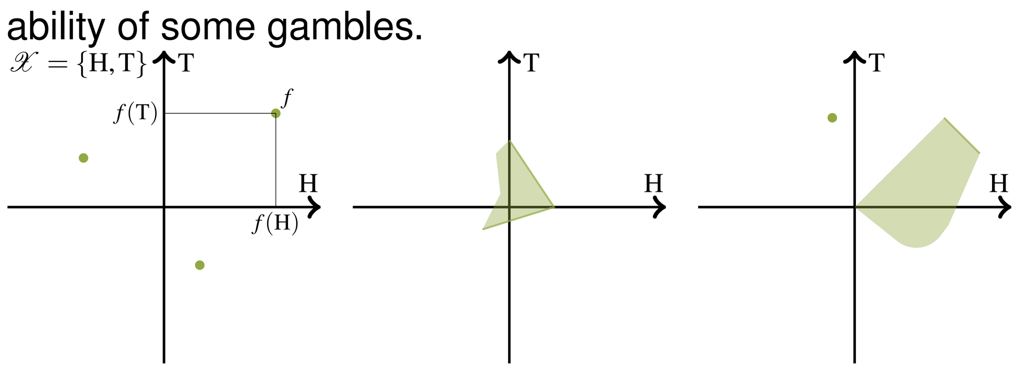
2. Sets of desirable gambles

Gambles A gamble f is a bounded real-valued function on $\mathscr{X}$. We collect all gambles in $\mathscr{L}(\mathscr{X})$. f is interpreted as an *uncertain reward*: if the outcome of X turns out to be x, then the agent receives the *possibly negative* reward $f(x)$.

Sets of desirable gambles The agent gives his assessment by specifying which gambles he strictly prefers to zero—or assesses to be desirable:

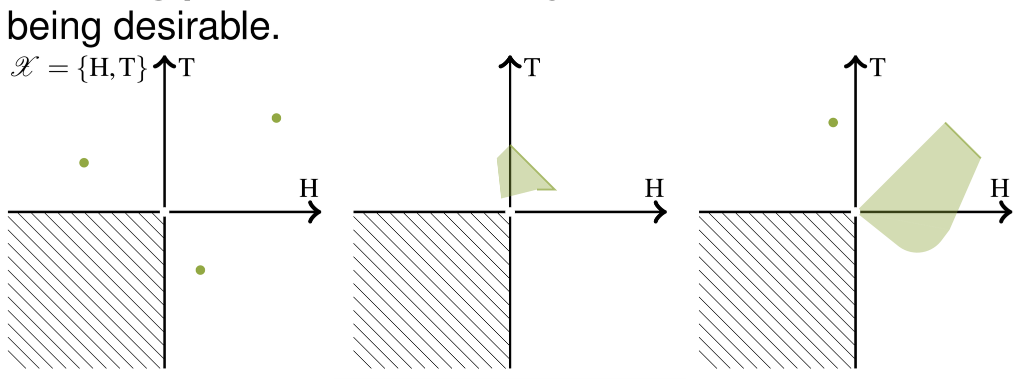
$\mathscr{D} = \{f \in \mathscr{L}(\mathscr{X}) : 0 \prec f\}$ for some strict partial order $\prec$.

Assessment An assessment $\mathscr{A}$ is a statement of desirability of some gambles.



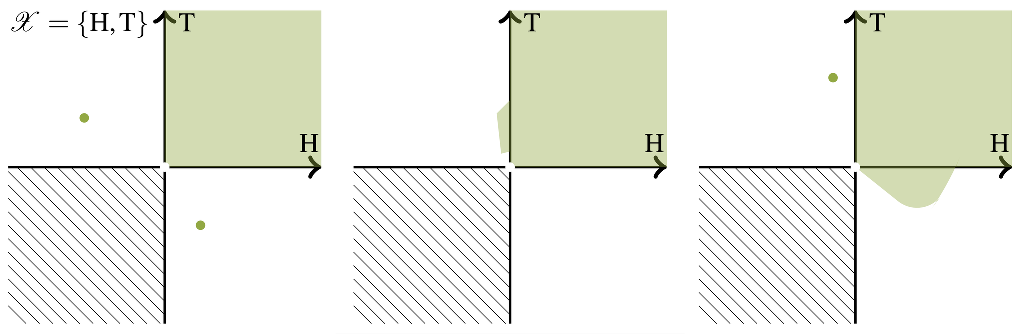
There are four *rationality requirements*.

Avoiding partial loss Some gambles are excluded from being desirable.



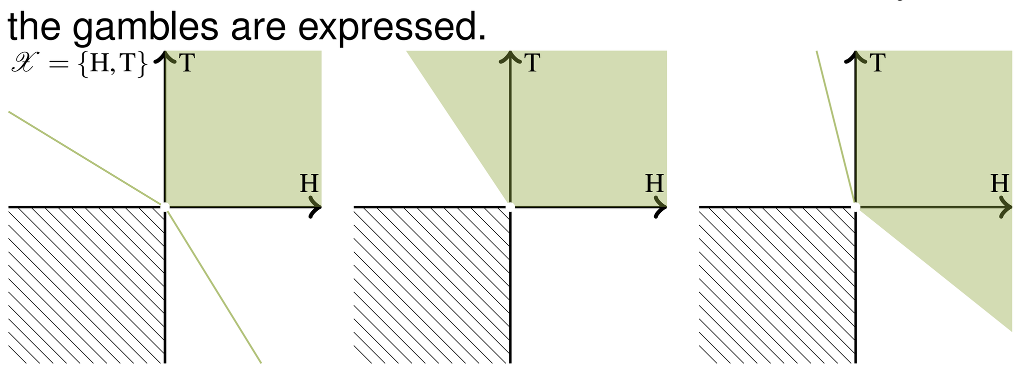
If $f \leq 0$, then $f \notin \mathscr{D}$

Accepting partial gain Some gambles are always desirable.



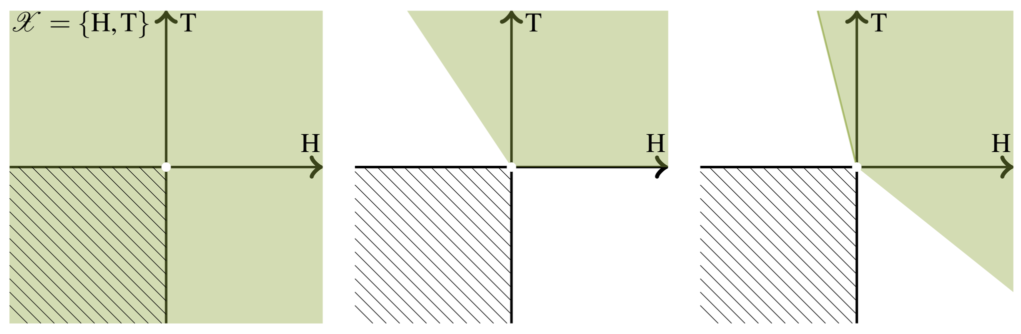
If $f > 0$, then $f \in \mathscr{D}$

Scale invariance It should not matter in which utility scale the gambles are expressed.



If $f \in \mathscr{D}$ and $\lambda \in \mathbb{R}_{>0}$, then $\lambda f \in \mathscr{D}$

Combination Given two desirable gambles, their sum is also desirable.



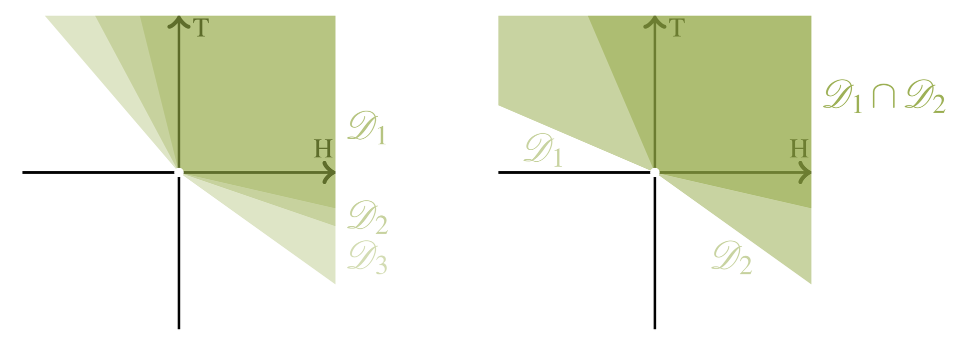
If $f, g \in \mathscr{D}$, then $f + g \in \mathscr{D}$

"Not more informative than" relation Given two sets of desirable gambles $\mathscr{D}_1$ and $\mathscr{D}_2$,

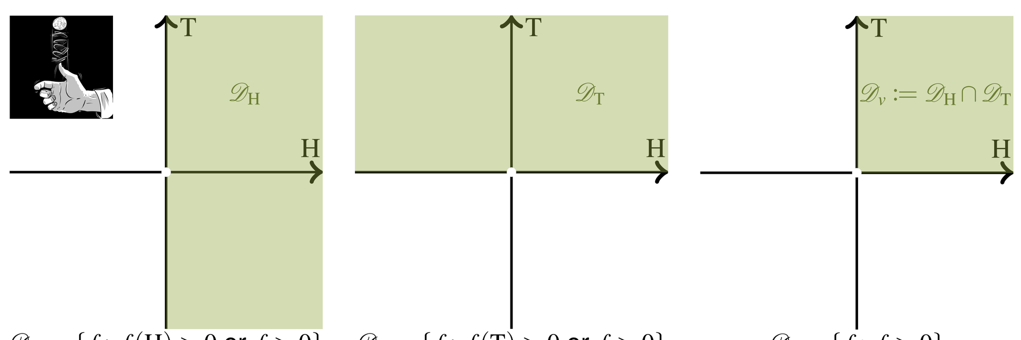
$\mathscr{D}_1$ is not more informative than $\mathscr{D}_2 \Leftrightarrow \mathscr{D}_1 \subseteq \mathscr{D}_2$.

Given a collection $\mathbf{D} = \{\mathscr{D}_1, \mathscr{D}_2, \dots\}$ of coherent sets of desirable gambles, its infimum with respect to the partial order $\subseteq$

$\inf \mathbf{D} = \bigcap \mathbf{D}$ is a coherent set of desirable gambles.

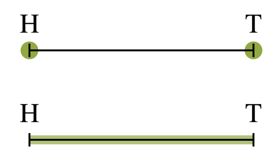


Coin example We want to use coherent sets of desirable gambles to express the belief that the coin has two identical sides.



$\mathscr{D}_\mathrm{H} = \{f : f(\mathrm{H}) > 0 \text{ or } f > 0\}$ $\mathscr{D}_\mathrm{T} = \{f : f(\mathrm{T}) > 0 \text{ or } f > 0\}$ $\mathscr{D}_v = \{f : f > 0\}$

The language of sets of desirable gambles is not powerful enough to distinguish between what we want to represent and the *vacuous*—or minimal informative—belief model $\mathscr{D}_v$: every mass function is possible.



3. Choice functions

Because the "sets of desirable gambles" approach fails in the coin example, we need more than *binary comparisons* of gambles. We collect all the finite non-empty sets of gambles in $\mathscr{Q}$.

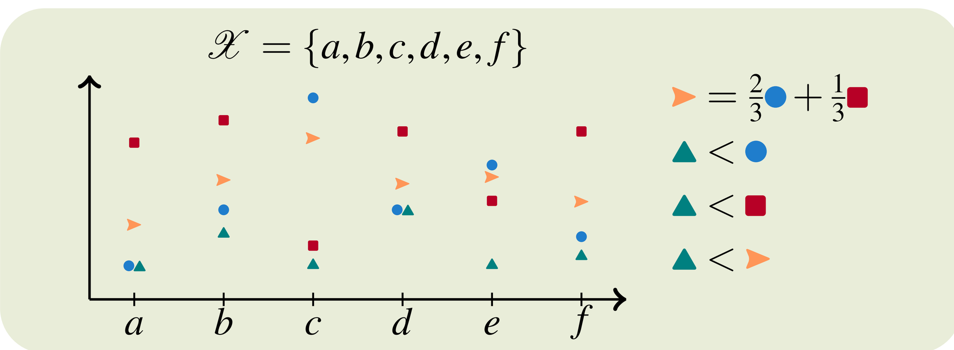
Choice functions A choice function C is a map

$C : \mathscr{Q} \to \mathscr{Q} \cup \{\emptyset\} : O \mapsto C(O)$ such that $C(O) \subseteq O$.

The idea is that C selects the set $C(O)$ of "best" gambles in the option set O. This allows for more than binary comparison between gambles. We define the rejection function R as

$R(O) := O \setminus C(O)$ for all O in $\mathscr{Q}$.

There are seven *rationality requirements*.



Non-emptiness The choice should never be empty.

$C(\{●, ▲, ■, ►\}) \neq \emptyset$ $C(O) \neq \emptyset$ for all O in $\mathscr{Q}$

Non-triviality Dominated gambles should not be chosen.

▲ $\notin C(\{▲, ●\})$, ▲ $\notin C(\{▲, ■\})$, ▲ $\notin C(\{▲, ►\})$
▲ $\in R(\{▲, ●\})$, ▲ $\in R(\{▲, ■\})$, ▲ $\in R(\{▲, ►\})$

If $f < g$, then $\{g\} = C(\{f, g\})$

Sen's property α A rejected gamble should not be promoted to a chosen gamble after adding gambles to the option set.

▲ $\in R(\{▲, ●\})$, ▲ $\in R(\{▲, ■, ►\})$, ▲ $\in R(\{▲, ►, ●\})$

If $O_1 \subseteq R(O_2)$ and $O_2 \subseteq O_3$, then $O_1 \subseteq R(O_3)$

A variant of Aizerman's condition A rejected gamble should not be promoted to a chosen gamble after deleting rejected gambles from the option set.

if $\{▲, ●\} \subseteq R(\{▲, ●, ■\})$, then ▲ $\in R(\{▲, ■\})$

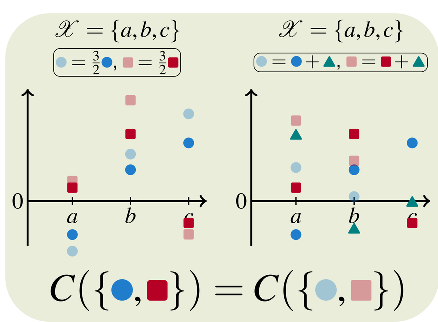
If $O_2 \subseteq R(O_3)$ and $O_1 \subseteq O_2$, then $O_2 \setminus O_1 \subseteq R(O_3 \setminus O_1)$

Convexity We introduce a new axiom: two gambles should not be rejected from their convex hull.

not both ● $\in R(\{●, ■, ►\})$ and ■ $\in R(\{●, ■, ►\})$

If $O \subseteq \mathrm{CH}(\{f, g\})$, then $\{f, g\} \nsubseteq R(O \cup \{f, g\})$

Scaling and independence Scaling all gambles and adding a single gamble is of no importance:



$C(\lambda O + \{f\}) = \lambda C(O) + \{f\}$

for all λ in $\mathbb{R}_{>0}$ and f in $\mathscr{L}(\mathscr{X})$.

"Not more informative than" relation Given two choice functions C_1 and C_2,

C_1 is not more informative than $C_2 \Leftrightarrow C_1(O) \supseteq C_2(O)$

for all O in $\mathscr{Q}$. For a collection $\mathbf{C}$ of coherent choice functions, its infimum is the coherent choice function given by

$(\inf \mathbf{C})(O) = \bigcup \mathbf{C}(O)$ for all O in $\mathscr{Q}$.

if $C_1(\{●, ▲, ■, ►\}) = \{●\}$, $C_2(\{●, ▲, ■, ►\}) = \{■, ►\}$,
then $(\inf\{C_1, C_2\})(\{●, ▲, ■, ►\}) = \{●, ■, ►\}$

Link with sets of desirable gambles Choice functions are essentially non-pairwise comparisons of gambles. Therefore, we can associated a single coherent set of desirable gambles with a coherent choice function C

$\mathscr{D}_C = \{f \in \mathscr{L}(\mathscr{X}) : 0 \in R(\{f, 0\})\}$.

Conversely, given a coherent set of desirable gambles $\mathscr{D}$, there are multiple associated coherent choice functions, and the least informative one is given by

$C_\mathscr{D}(O) = \{f \in O : (\forall g \in O) g - f \notin \mathscr{D}\}$ for all O in $\mathscr{Q}$.

Inference Given a collection $\mathbf{D}$ of coherent sets of desirable gambles,

$\mathscr{D}_{\inf\{C_{\mathscr{D}'} : \mathscr{D}' \in \mathbf{D}\}} = \inf \mathbf{D}$.   **no loss** of information

Given a collection $\mathbf{C}$ of coherent choice functions, for all O in $\mathscr{Q}$,

$C_{\inf\{\mathscr{D}_{C'} : C' \in \mathbf{C}\}}(O) \supseteq (\inf \mathbf{C})(O)$.   **loss** of information

4. Coin example with choice functions

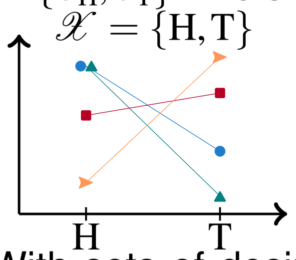
Choice functions are expressive enough to model the belief used in the coin example. We start with $\mathscr{D}_v$, and associate the choice function $C_v := C_{\mathscr{D}_v}$ with it:

$C_v(O) = \{f \in O : (\forall g \in O) f \not< g\}$ for all O in $\mathscr{Q}$,

being the *vacuous*—or minimal informative—choice function.
We now work with $\mathscr{D}_\mathrm{H}$ and $\mathscr{D}_\mathrm{T}$, and associate $C_\mathrm{H} := C_{\mathscr{D}_\mathrm{H}}$ and $C_\mathrm{T} := C_{\mathscr{D}_\mathrm{T}}$.

$C_\mathrm{H}(O) = \arg\max\{f(\mathrm{H}) : f \in O\} \cap C_v(O)$ for all O in $\mathscr{Q}$,

and analogous for C_T.
The most informative choice function that is implied by C_H and C_T is $C_\text{coin} := \inf\{C_\mathrm{H}, C_\mathrm{T}\}$. We show that $C_\text{coin} \neq C_v$.



The vacuous choice function: $C_v(\{●, ▲, ■, ►\}) = \{●, ■, ►\}$
The "coin" choice function: $C_\text{coin}(\{●, ▲, ■, ►\}) = \{●, ►\}$

$C_\text{coin} \neq C_v$

With sets of desirable gambles, we are unable to express the difference between the belief of the example and knowing nothing, whereas choice functions are expressive enough to expose this difference.

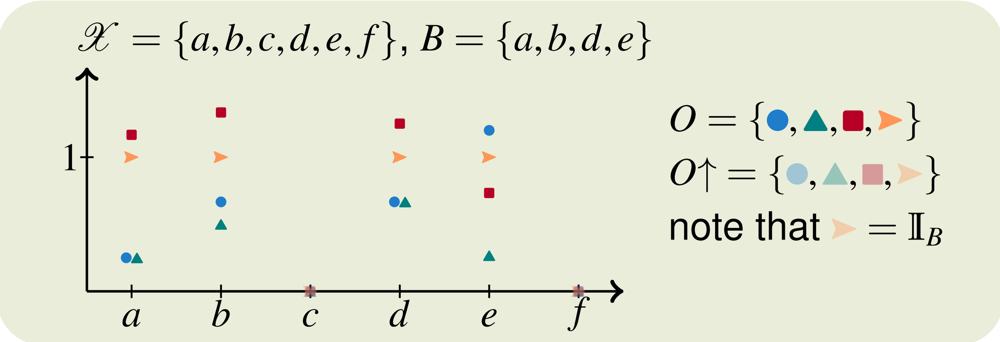
5. Conditioning with choice functions

If we want choice functions to be successful as a general tool to model uncertainty, we need a good *conditioning rule* that coincides with

- the existing conditioning rule for sets of desirable gambles;
- Bayes's rule for mass functions.

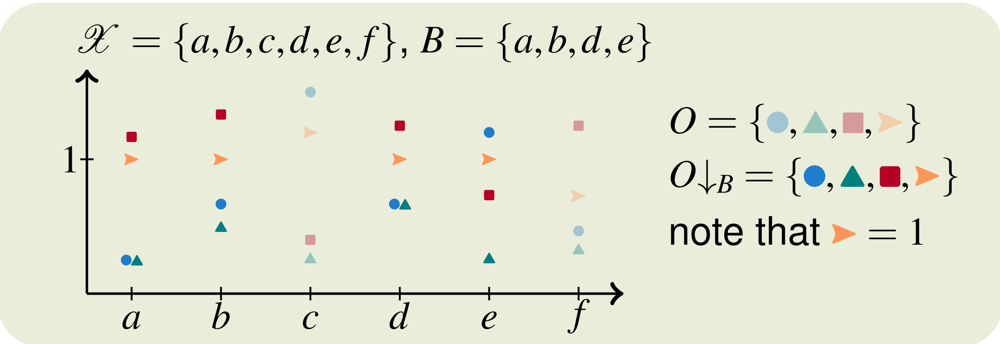
Some definitions With $B \subseteq \mathscr{X}$, $\mathbb{I}_B$ is the *indicator* of B.
With O in $\mathscr{Q}(\mathscr{L}(B))$, define

$O\!\uparrow := \{f \in \mathscr{L}(\mathscr{X}) : f\mathbb{I}_{B^c} = 0 \text{ and } (\exists g \in O) f\mathbb{I}_B = g\mathbb{I}_B\} \in \mathscr{Q}(\mathscr{L}(\mathscr{X}))$.



With O in $\mathscr{Q}(\mathscr{L}(\mathscr{X}))$, define

$O\!\downarrow_B := \{f \in \mathscr{L}(B) : (\exists g \in O) f\mathbb{I}_B = g\mathbb{I}_B\} \in \mathscr{Q}(\mathscr{L}(B))$.



Conditioning rule We start with a coherent choice function C on $\mathscr{Q}(\mathscr{L}(\mathscr{X}))$, and we obtain the *new information* that $B \subseteq \mathscr{X}$ occurs. Conditioning—or updating—is changing the belief model C according to this new information:

C on $\mathscr{Q}(\mathscr{L}(\mathscr{X})) \xrightarrow{\text{new information: } B \text{ occurs}} C\rfloor B$ on $\mathscr{Q}(\mathscr{L}(B))$

We propose the following conditioning rule:

$\boxed{C\rfloor B(O) := C(O\!\uparrow)\!\downarrow_B}$ for all O in $\mathscr{Q}(\mathscr{L}(B))$.

Properties Given a coherent choice function C,

$C\rfloor B$ is coherent $\Leftrightarrow B \neq \emptyset$.

For sets of desirable gambles, the conditioning rule leads to the *usual conditioning rule* for sets of desirable gambles: given a coherent choice function C,

$\mathscr{D}_C\rfloor B = \{f \in \mathscr{L}(B) : f\mathbb{I}_B \in \mathscr{D}_C\}$,

meaning that our conditioning rule **implies Bayes's rule**.