On belief function corrections

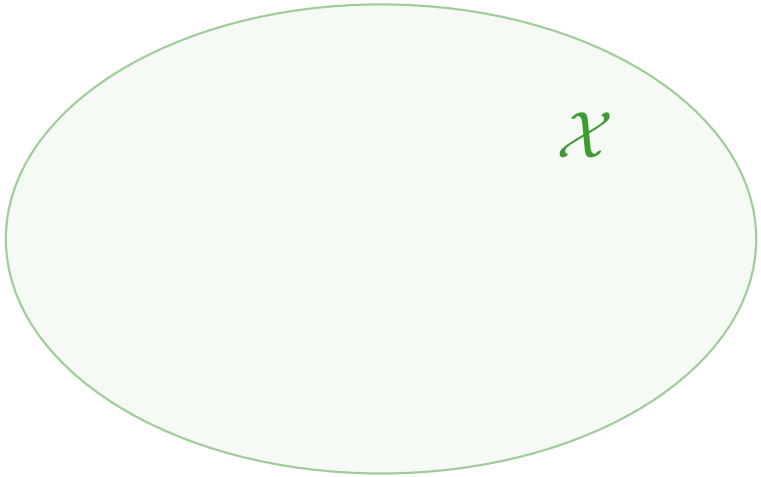
D. Mercier

University of Artois, EA 3926 LGI2A, Béthune, France

2015 BFAS School, Stella Plage, Friday, October 2nd, 2015

What do we mean by correction of a belief function (BF)?

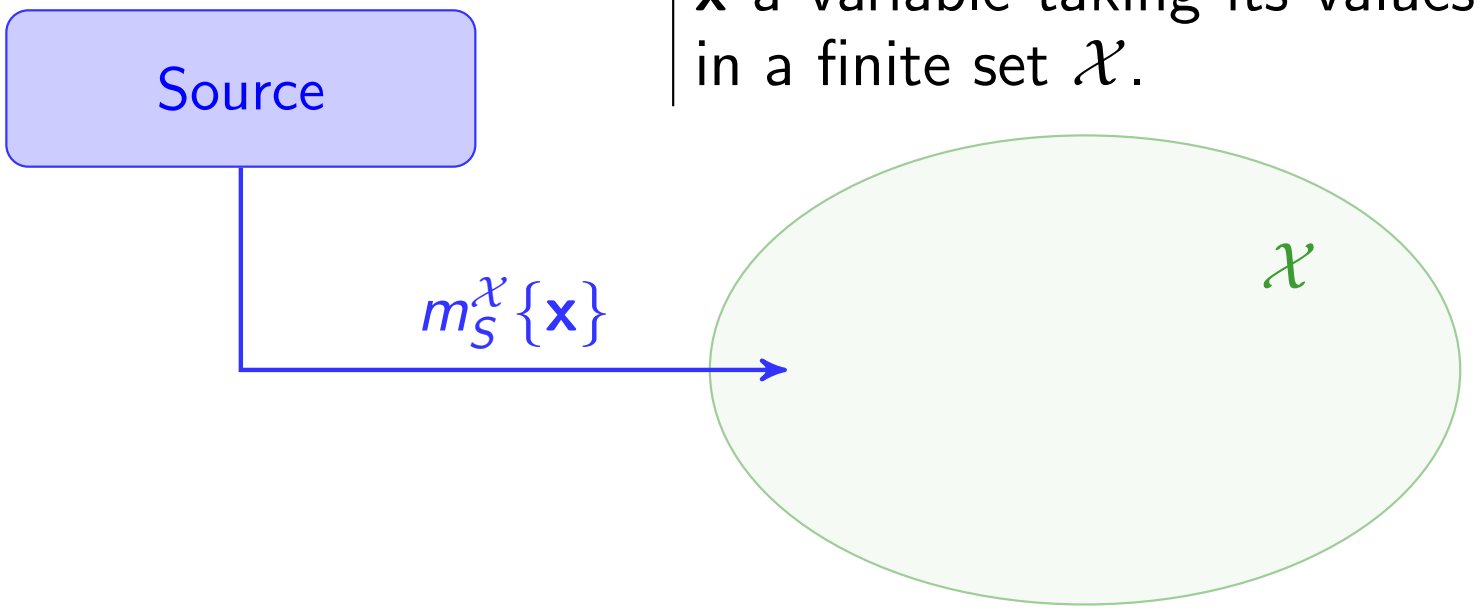
x a variable taking its values in a finite set $\mathcal{X}$.



What do we mean by correction of a belief function (BF)?

A source provides a piece of information on the actual value of **x**

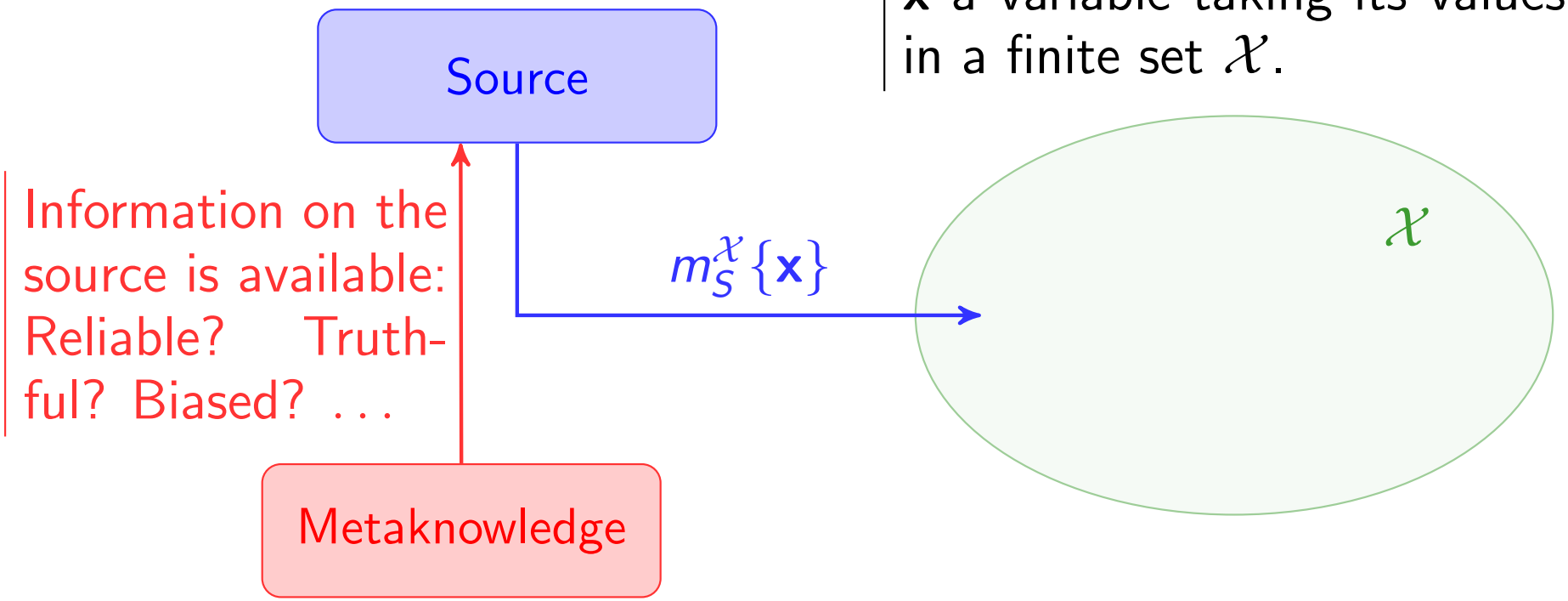
x a variable taking its values in a finite set $\mathcal{X}$.



What do we mean by correction of a belief function (BF)?

A source provides a piece of information on the actual value of **x**

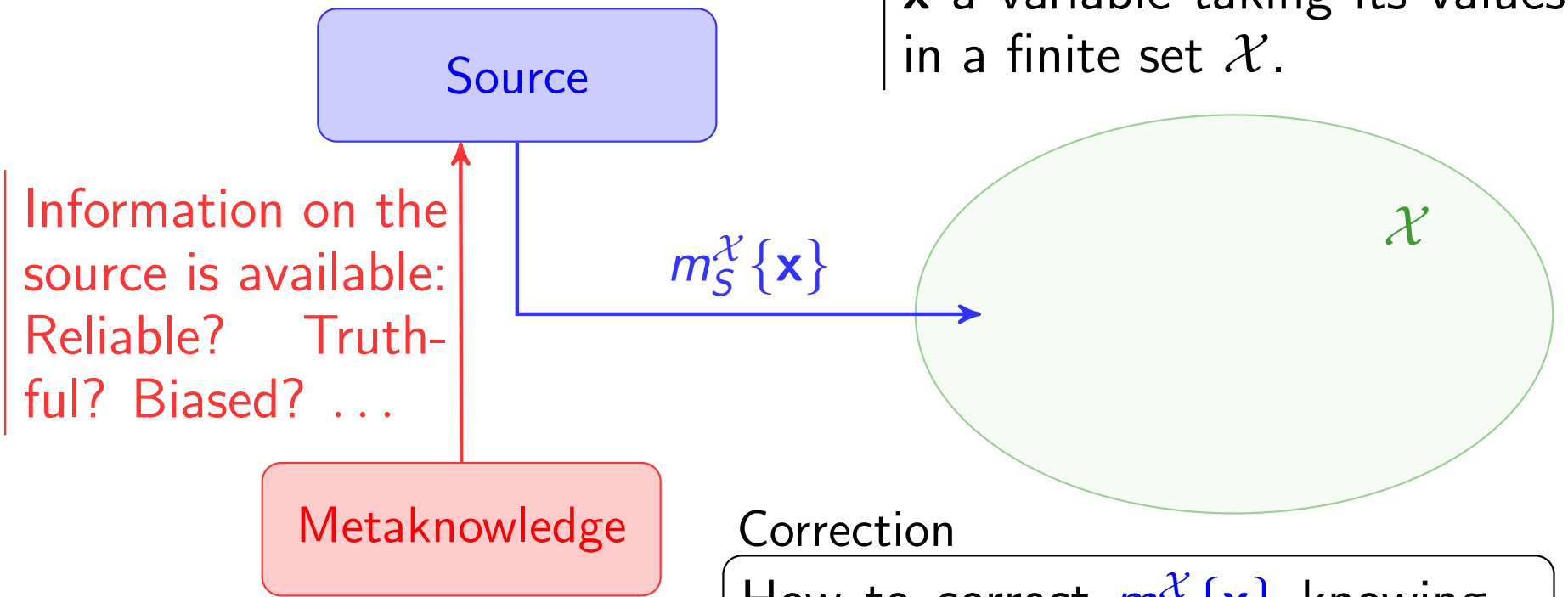
x a variable taking its values in a finite set $\mathcal{X}$.



What do we mean by correction of a belief function (BF)?

A source provides a piece of information on the actual value of **x**

x a variable taking its values in a finite set $\mathcal{X}$.

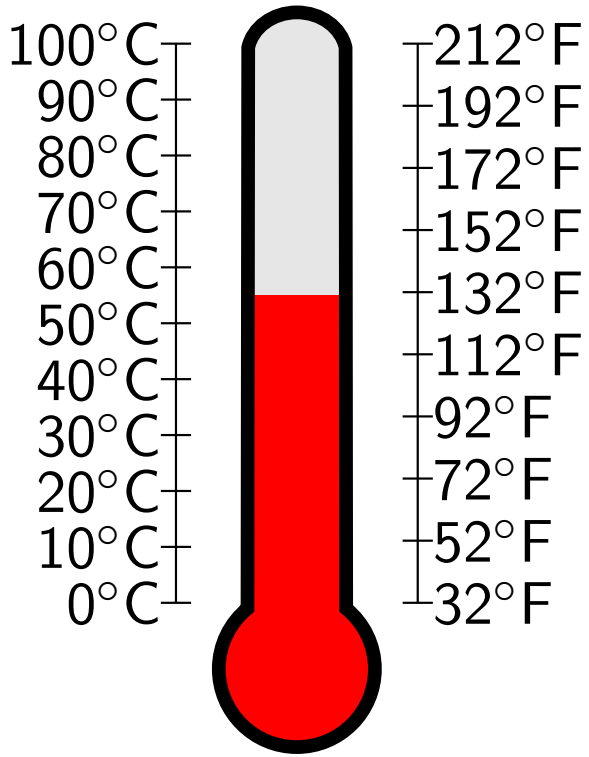


Correction

How to correct $m_S^{\mathcal{X}}\{\mathbf{x}\}$ knowing this metaknowledge?

Examples of different kinds of metaknowledge

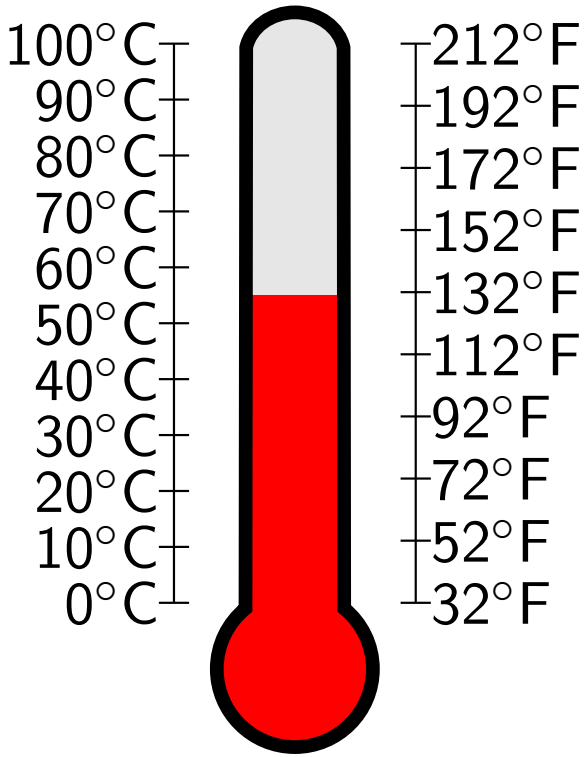
A thermometer gives us a temperature 55°C.



Examples of different kinds of metaknowledge

A thermometer gives us a temperature 55°C.

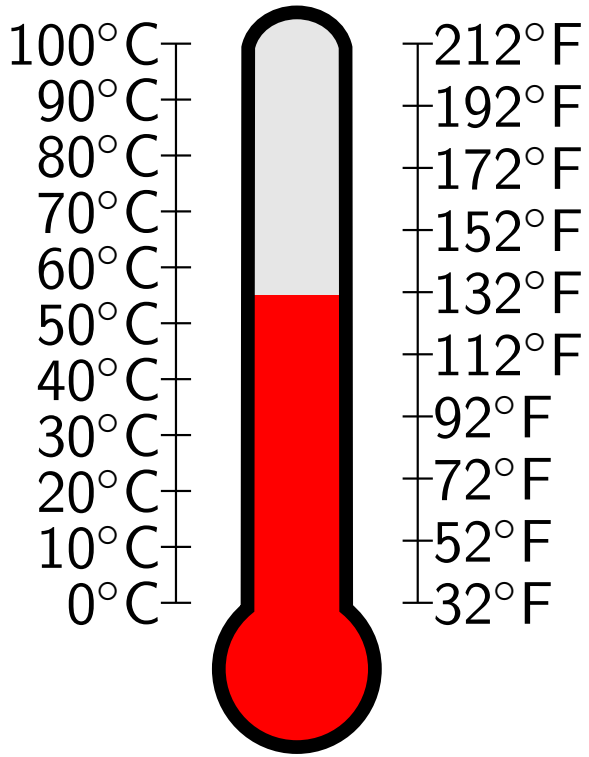
- **Metaknowledge 1 (global):** doubt about its reliability



Examples of different kinds of metaknowledge

A thermometer gives us a temperature 55°C.

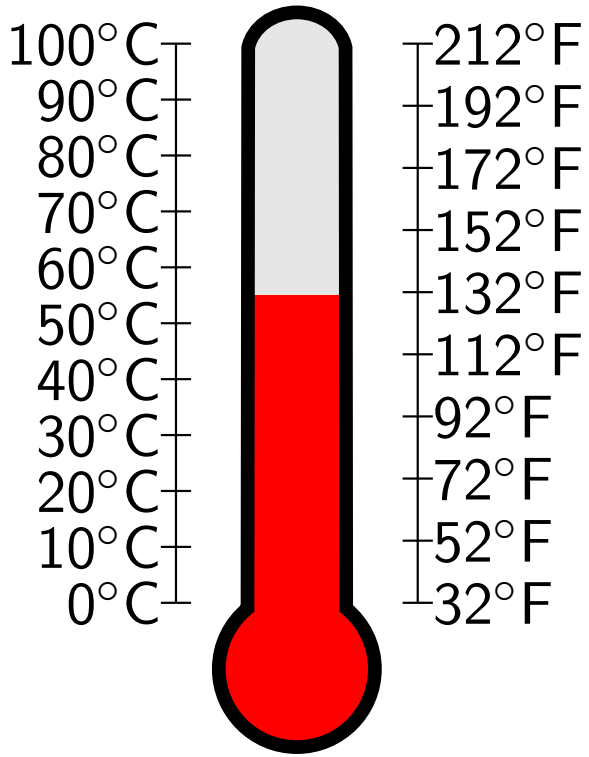
- **Metaknowledge 1 (global):** doubt about its reliability
- **Metaknowledge 2 (contextual):** reliable in the range $[0°C, 100°C]$, unreliable outside this range.



Examples of different kinds of metaknowledge

A thermometer gives us a temperature 55°C.

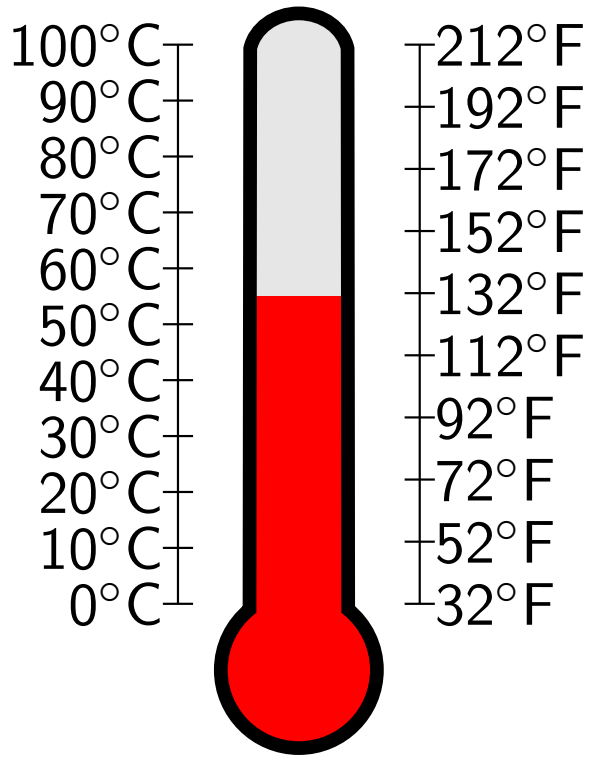
- **Metaknowledge 1 (global):** doubt about its reliability
- **Metaknowledge 2 (contextual):** reliable in the range $[0°C, 100°C]$, unreliable outside this range.
- **Metaknowledge 3 (shift):** If it gives t°C: it means in fact that the temperature is between $t - 1°C$ and $t + 4°C$



Examples of different kinds of metaknowledge

A thermometer gives us a temperature 55°C.

- **Metaknowledge 1 (global):** doubt about its reliability
- **Metaknowledge 2 (contextual):** reliable in the range $[0°C, 100°C]$, unreliable outside this range.
- **Metaknowledge 3 (shift):** If it gives t°C: it means in fact that the temperature is between $t - 1°C$ and $t + 4°C$



How to make such corrections when information is represented by belief functions?

Main objective of the lecture

- Give an overview of correction models with their justifications / derivations (where do they come from?, how they are derived?).
- Show how to automatically learn some of them from labelled data (which can also help to build belief functions).

Outline

Discounting: a detailed presentation

A parametrized family of correction mechanisms

Pichon et al.'s Behaviour Based Correction (BBC)

Contextual discounting (CD) and reinforcement (CR)

Learning CD and CR from labelled data

Outline

Discounting: a detailed presentation

A parametrized family of correction mechanisms

Pichon et al.'s Behaviour Based Correction (BBC)

Contextual discounting (CD) and reinforcement (CR)

Learning CD and CR from labelled data

Discounting

A simple correction example. Introduced by Shafer in 1976.

Discounting of a mass function (MF) m is defined by (Shafer,1976):

$$\begin{cases} {}^{\alpha}m(A) &= (1-\alpha)m(A), \quad \forall A \subset \mathcal{X}, \\ {}^{\alpha}m(\mathcal{X}) &= (1-\alpha)m(\mathcal{X}) + \alpha \ , \end{cases}$$

where $\alpha \in [0,1]$ is the discount rate.

Discounting

A simple correction example. Introduced by Shafer in 1976.

Discounting of a mass function (MF) m is defined by (Shafer,1976):

$$\begin{cases} {}^{\alpha}m(A) &= (1-\alpha)m(A), \quad \forall A \subset \mathcal{X}, \\ {}^{\alpha}m(\mathcal{X}) &= (1-\alpha)m(\mathcal{X}) + \alpha \ , \end{cases}$$

where $\alpha \in [0,1]$ is the discount rate.

Example:

- $\mathcal{X} = \{x_1, x_2, x_3\}$
- $m(\{x_1\}) = .2$, $m(\{x_2\}) = .4$ and $m(\{x_1, x_2\}) = .4$
- With discount rate $\alpha = .2$:

$$\begin{cases} {}^{\alpha}m(\{x_1\}) &= .8 \times .2 &= .16 \\ {}^{\alpha}m(\{x_2\}) &= .8 \times .4 &= .32 \\ {}^{\alpha}m(\{x_1, x_2\}) &= .8 \times .4 &= .32 \\ {}^{\alpha}m(\mathcal{X}) &= .8 \times .0 + .2 &= .20 \end{cases}$$

Discounting

Matrix representation (Smets, 2002)

Discounting ${}^{\alpha}m$ of m is a generalization of m (${}^{\alpha}m \sqsupseteq_s m$):

$$ {}^{\alpha}m(A) = \sum_{B \subseteq \mathcal{X}} {}^{\alpha}G(A,B)m(B) \ , $$

with ${}^{\alpha}\mathbf{G}$ a generalisation matrix defined by:

$$ {}^{\alpha}G(A,B) = \begin{cases} 1-\alpha & \text{if } A = B \neq \mathcal{X}, \\ \alpha & \text{if } A = \mathcal{X} \text{ and } B \subset A, \\ 1 & \text{if } A = B = \mathcal{X} \\ 0 & \text{otherwise.} \end{cases} $$

$$ {}^{\alpha}\mathbf{G} = \begin{pmatrix} 1-\alpha & 0 & \dots & 0 & 0 \\ 0 & 1-\alpha & \dots & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & & 0 & 0 \\ 0 & 0 & \dots & 1-\alpha & 0 \\ \alpha & \alpha & \dots & \alpha & 1 \end{pmatrix} $$

Discounting

Matrix representation: example

With $\alpha = .2$, $\beta = 1 - \alpha = .8$ and $\mathcal{X} = \{x_1, x_2, x_3\}$:

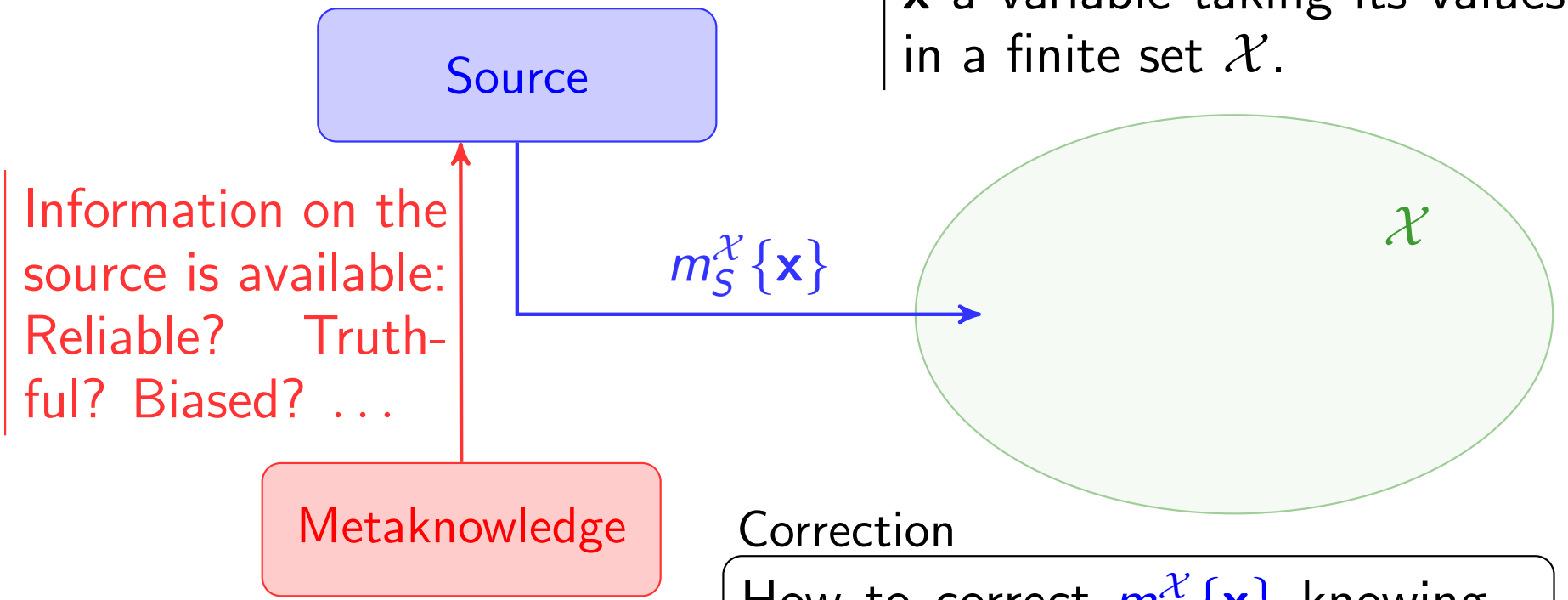
$$
\begin{array}{c} {}^{\alpha}m \\ \emptyset \\ \{x_1\} \\ \{x_2\} \\ \{x_1, x_2\} \\ \{x_3\} \\ \{x_1, x_3\} \\ \{x_2, x_3\} \\ \{x_1, x_2, x_3\} \end{array}
\begin{pmatrix} .0 \\ .16 \\ .32 \\ .32 \\ .0 \\ .0 \\ .0 \\ .20 \end{pmatrix}
=
\begin{array}{c} {}^{\alpha}G(A,B) \\ \\ \\ \\ \\ \\ \\ \\ \end{array}
\begin{pmatrix}
\beta & & & & & & & \\
 & \beta & & & & & & \\
 & & \beta & & & & & \\
 & & & \beta & & & & \\
 & & & & \beta & & & \\
 & & & & & \beta & & \\
 & & & & & & \beta & \\
\alpha & \alpha & \alpha & \alpha & \alpha & \alpha & \alpha & 1
\end{pmatrix}
\cdot
\begin{array}{c} m \\ \\ \\ \\ \\ \\ \\ \\ \end{array}
\begin{pmatrix} .0 \\ .2 \\ .4 \\ .4 \\ .0 \\ .0 \\ .0 \\ .0 \end{pmatrix}
$$

Discounting

Eh, wait a minute... What is the link between discounting and the slide illustrating belief function corrections?

A source provides a piece of information on the actual value of **x**

x a variable taking its values in a finite set $\mathcal{X}$.



Correction

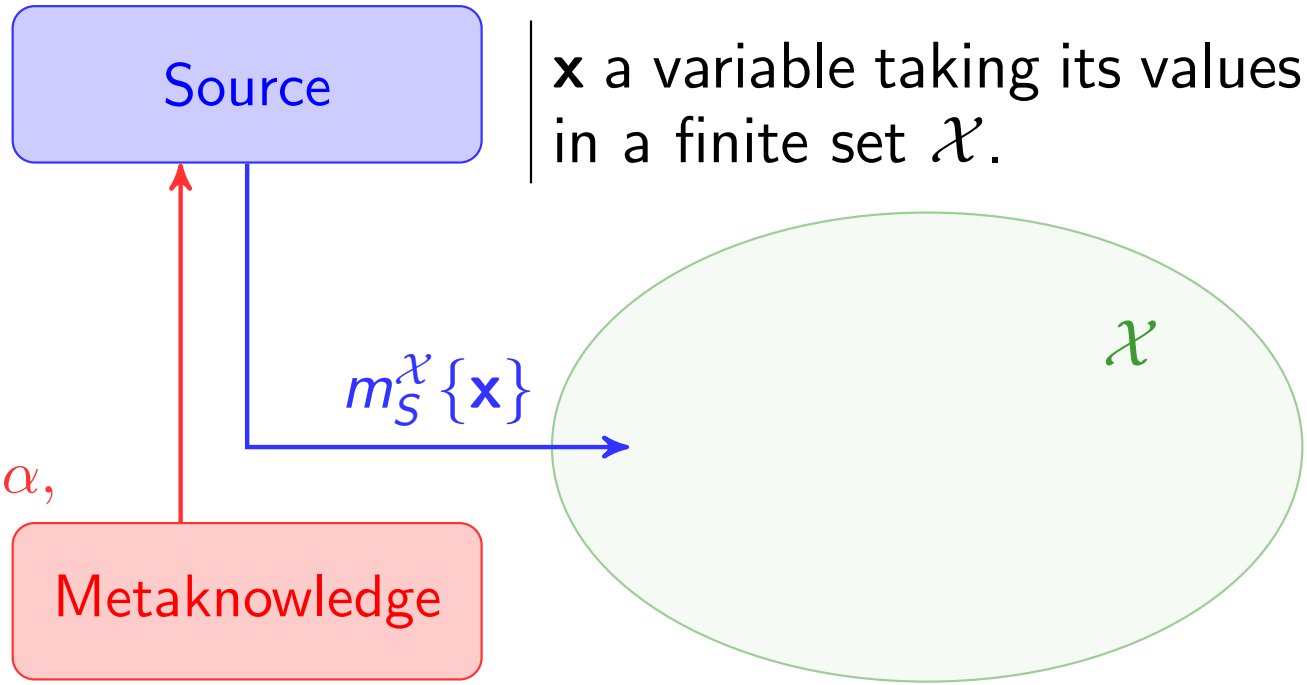
How to correct $m_S^{\mathcal{X}}\{\mathbf{x}\}$ knowing this metaknowledge?

Discounting

Answer: discounting can be derived from the following Model

(Proof given by Smets in 1993)

The source is reliable (R) or not (NR).
$\mathcal{R} = \{R, NR\}$.
$m^{\mathcal{X}}[\{R\}] = m_S^{\mathcal{X}}$.
$m^{\mathcal{X}}[\{NR\}] = m_{\mathcal{X}}$.
$\begin{cases} m^{\mathcal{R}}(\{R\}) = 1-\alpha, \\ m^{\mathcal{R}}(\mathcal{R}) = \alpha\ . \end{cases}$



x a variable taking its values in a finite set $\mathcal{X}$.

Discounting

Answer: discounting can be derived from the following Model

(Proof given by Smets in 1993)

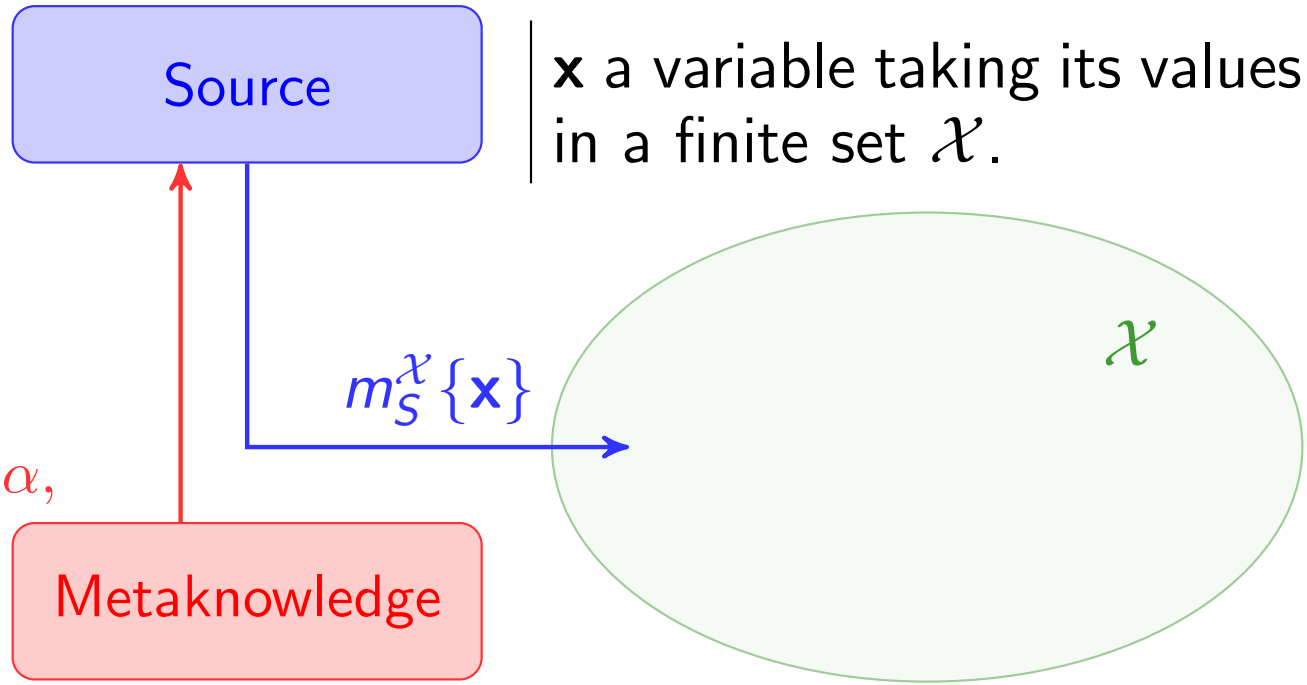
The source is reliable (R) or not (NR).

$\mathcal{R} = \{R, NR\}$.

$m^{\mathcal{X}}[\{R\}] = m_S^{\mathcal{X}}$.

$m^{\mathcal{X}}[\{NR\}] = m_{\mathcal{X}}$.

$\begin{cases} m^{\mathcal{R}}(\{R\}) = 1-\alpha, \\ m^{\mathcal{R}}(\mathcal{R}) = \alpha\ . \end{cases}$



Discounting

${}^{\alpha}m$ is then obtained from $m^{\mathcal{R}}$ and $m_S^{\mathcal{X}}$: ${}^{\alpha}m = \left(m^{\mathcal{X}}[\{NR\}]^{\Uparrow\mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{X}}[\{R\}]^{\Uparrow\mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}}\right)^{\downarrow\mathcal{X}}$.

Discounting

Let us do the computations :)

Discounting

Let us do the computations :)

But first, let us give some recalls of the notions of:

- conditioning and deconditioning,
- marginalisation and vacuous extension,

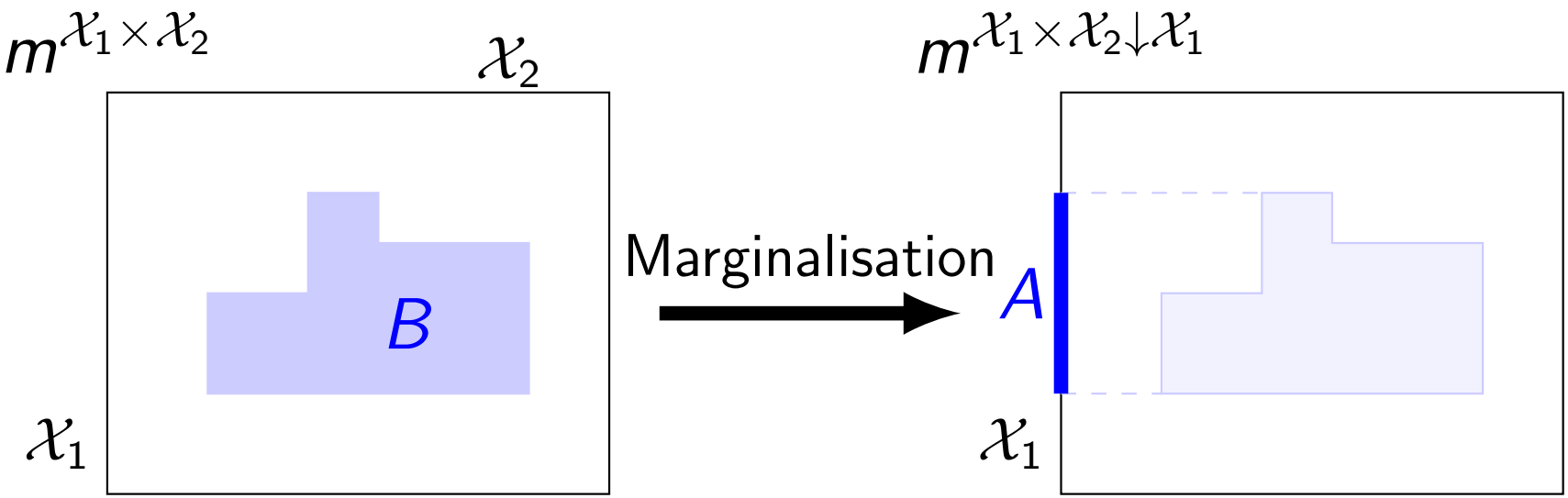
on a product space $\mathcal{X}_1 \times \mathcal{X}_2$.

Marginalisation in the case of a product space (recalls)

- MF $m^{\mathcal{X}_1 \times \mathcal{X}_2}$ can be marginalised on $\mathcal{X}_1$ by transferring each mass $m^{\mathcal{X}_1 \times \mathcal{X}_2}(B)$, $B \subseteq \mathcal{X}_1 \times \mathcal{X}_2$, to the projection of B on $\mathcal{X}_1$:

$$m^{\mathcal{X}_1 \times \mathcal{X}_2 \downarrow \mathcal{X}_1}(A) = \sum_{\{B \subseteq \mathcal{X}_1 \times \mathcal{X}_2 \mid proj(B \downarrow \mathcal{X}_1) = A\}} m^{\mathcal{X}_1 \times \mathcal{X}_2}(B) \ , \forall A \subseteq \mathcal{X}_1 \ .$$

- Illustration

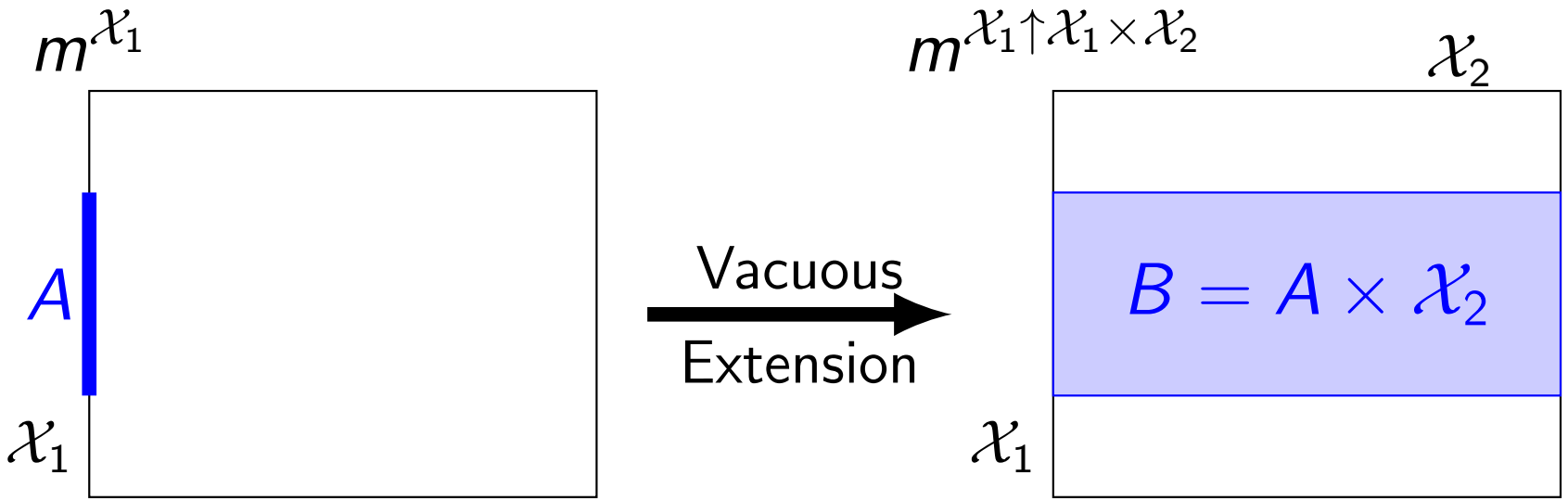


Vacuous extension in the case of a product space (recalls)

- Vacuous extension of MF $m^{\mathcal{X}_1}$ on $\mathcal{X}_1 \times \mathcal{X}_2$ is defined by (s-least committed solution, cf Lecture 2 of T. Denœux):

$$m^{\mathcal{X}_1 \uparrow \mathcal{X}_1 \times \mathcal{X}_2}(B) = \begin{cases} m_1^{\mathcal{X}}(A) & \text{if } B = A \times \mathcal{X}_2,\ A \subseteq \mathcal{X}_1 \\ 0 & \text{otherwise.} \end{cases}$$

- Illustration

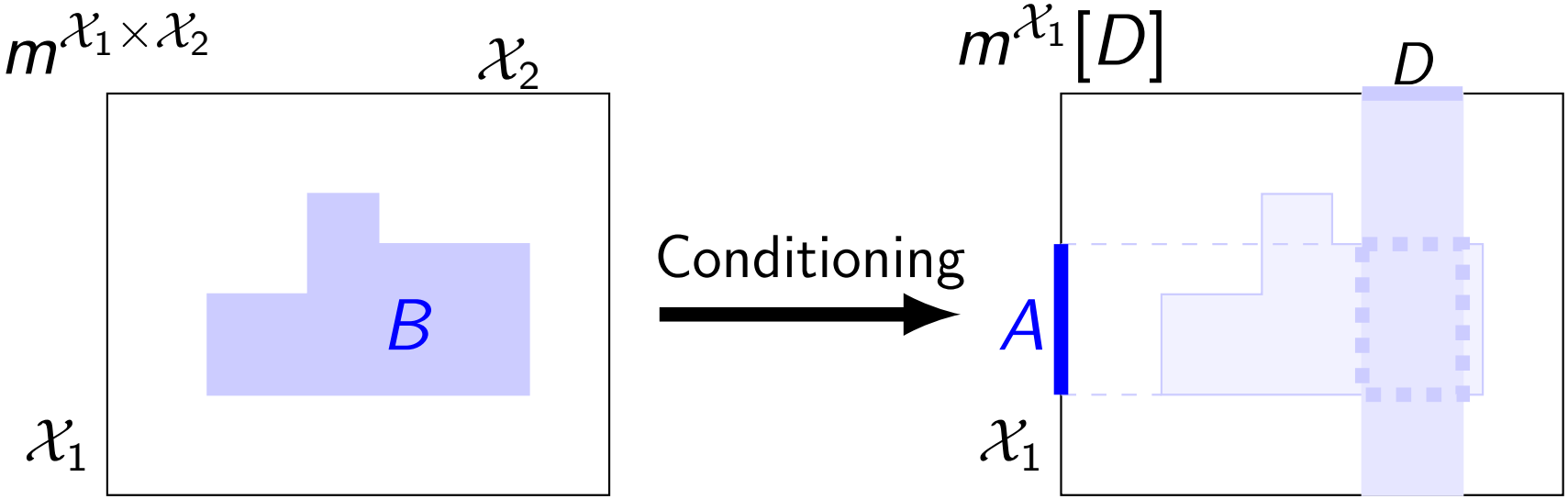


Conditioning in the case of a product space (recalls)

- With $D \subseteq \mathcal{X}_2$, the conditioning of a MF $m^{\mathcal{X}_1 \times \mathcal{X}_2}$ is noted $m^{\mathcal{X}_1}[D]$ and defined by:

$$m^{\mathcal{X}_1}[D] = \left(m^{\mathcal{X}_1 \times \mathcal{X}_2} \textcircled{\cap} m_D^{\mathcal{X}_2 \uparrow \mathcal{X}_1 \times \mathcal{X}_2} \right)^{\downarrow \mathcal{X}_1} .$$

- Illustration

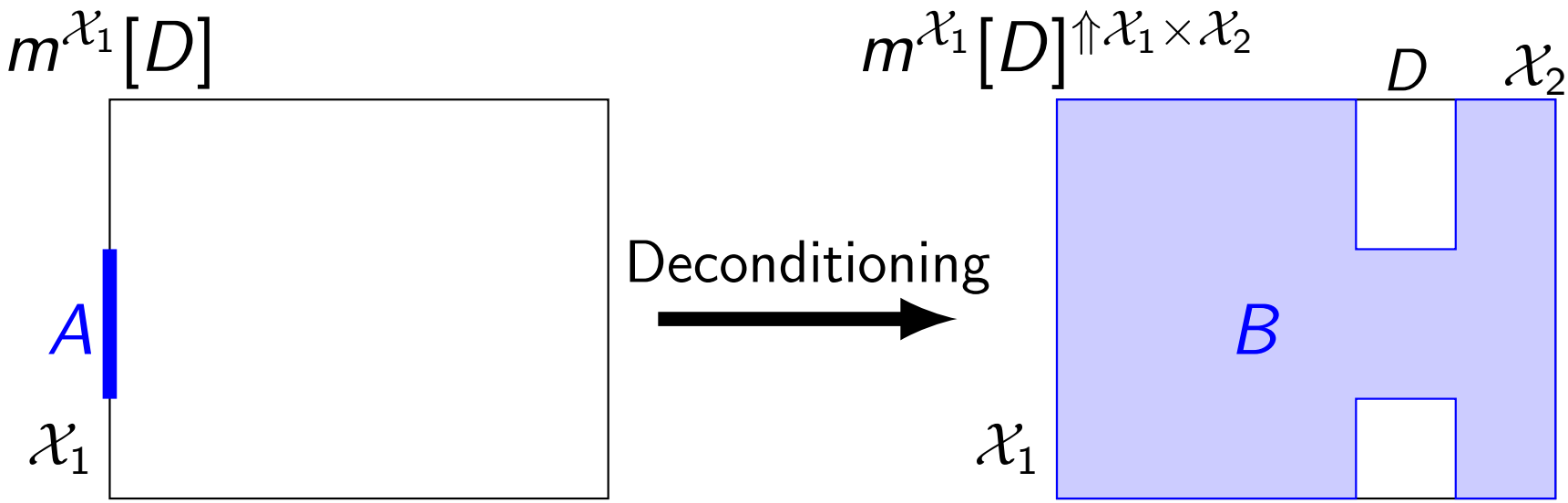


Deconditioning in the case of a product space (recalls)

- Deconditioning of a MF $m^{\mathcal{X}_1}[D]$ on $\mathcal{X}_1 \times \mathcal{X}_2$ is defined by (s-least committed solution):

$$m^{\mathcal{X}_1}[D]^{\Uparrow \mathcal{X}_1 \times \mathcal{X}_2}(A \times D \cup \mathcal{X}_1 \times \overline{D}) = m_1^{\mathcal{X}}[D](A), \quad \forall A \subseteq \mathcal{X}_1 \ .$$

- Illustration



- We want to compute:

$$\left(m^{\mathcal{X}}[\{NR\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow \mathcal{X}\times\mathcal{R}}\right)^{\downarrow \mathcal{X}} \quad (1)$$

with: $\mathcal{R} = \{R, NR\}$, $m^{\mathcal{X}}[\{R\}] = m^{\mathcal{X}}_S$, $m^{\mathcal{X}}[\{NR\}] = m_{\mathcal{X}}$ and

$$\begin{cases} m^{\mathcal{R}}(\{R\}) &= 1-\alpha, \\ m^{\mathcal{R}}(\mathcal{R}) &= \alpha\ . \end{cases}$$

- We want to compute:

$$\left(m^{\mathcal{X}}[\{NR\}]^{\Uparrow\mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{X}}[\{R\}]^{\Uparrow\mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}}\right)^{\downarrow\mathcal{X}} \quad (1)$$

with: $\mathcal{R} = \{R, NR\}$, $m^{\mathcal{X}}[\{R\}] = m^{\mathcal{X}}_S$, $m^{\mathcal{X}}[\{NR\}] = m_{\mathcal{X}}$ and
$\begin{cases} m^{\mathcal{R}}(\{R\}) &= 1-\alpha, \\ m^{\mathcal{R}}(\mathcal{R}) &= \alpha\ . \end{cases}$

- As $m^{\mathcal{X}}[\{NR\}]$ is the vacuous MF on $\mathcal{X}$, $m^{\mathcal{X}}[\{NR\}]^{\Uparrow\mathcal{X}\times\mathcal{R}}$ is the vacuous MF on $\mathcal{X}\times\mathcal{R}$, and (1) is equivalent to:

$$\left(m^{\mathcal{X}}[\{R\}]^{\Uparrow\mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}}\right)^{\downarrow\mathcal{X}} \quad (2)$$

- The deconditioning of $m^{\mathcal{X}}[\{R\}]$, is given for all $A \subseteq \mathcal{X}$ by:

$$m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X} \times \mathcal{R}}(A \times \{R\} \cup \mathcal{X} \times \{NR\}) = m_S^{\mathcal{X}}(A) \ .$$

- The deconditioning of $m^{\mathcal{X}}[\{R\}]$, is given for all $A \subseteq \mathcal{X}$ by:

$$m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}}(A \times \{R\} \cup \mathcal{X} \times \{NR\}) = m_S^{\mathcal{X}}(A) \ .$$

- The vacuous extension of $m^{\mathcal{R}}$ on $\mathcal{X} \times \mathcal{R}$ is:

$$\left\{ \begin{array}{lcl} m^{\mathcal{R}\uparrow \mathcal{X}\times\mathcal{R}}(\mathcal{X} \times \{R\}) & = & 1-\alpha \\ m^{\mathcal{R}\uparrow \mathcal{X}\times\mathcal{R}}(\mathcal{X} \times \mathcal{R}) & = & \alpha \ . \end{array} \right.$$

- The deconditioning of $m^{\mathcal{X}}[\{R\}]$, is given for all $A \subseteq \mathcal{X}$ by:

$$m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}}(A \times \{R\} \cup \mathcal{X} \times \{NR\}) = m_S^{\mathcal{X}}(A) \ .$$

- The vacuous extension of $m^{\mathcal{R}}$ on $\mathcal{X} \times \mathcal{R}$ is:

$$\begin{cases} m^{\mathcal{R}\uparrow \mathcal{X}\times\mathcal{R}}(\mathcal{X} \times \{R\}) &= 1-\alpha \\ m^{\mathcal{R}\uparrow \mathcal{X}\times\mathcal{R}}(\mathcal{X} \times \mathcal{R}) &= \alpha \ . \end{cases}$$

- For all $A \subseteq \mathcal{X}$, intersections of focal elements (FEs) are given by:

$$\begin{cases} (A \times \{R\} \cup \mathcal{X} \times \{NR\}) \cap (\mathcal{X} \times \{R\}) &= A \times \{R\} \\ (A \times \{R\} \cup \mathcal{X} \times \{NR\}) \cap (\mathcal{X} \times \mathcal{R}\}) &= A \times \{R\} \cup \mathcal{X} \times \{NR\} \end{cases}$$

- So, FEs of $m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}}$ are:

$$\begin{array}{lll} A \times \{R\} & \text{with mass} & (1-\alpha)\cdot m_S(A) \\ A \times \{R\} \cup \mathcal{X} \times \{NR\} & \text{with mass} & \alpha \cdot m_S(A) \end{array}$$

for all $A \subseteq \mathcal{X}$.

- So, FEs of $m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}}$ are:

$$\begin{array}{lll} A\times\{R\} & \text{with mass} & (1-\alpha)\cdot m_S(A) \\ A\times\{R\}\cup\mathcal{X}\times\{NR\} & \text{with mass} & \alpha\cdot m_S(A) \end{array}$$

  for all $A\subseteq\mathcal{X}$.

- FEs of $(m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}}$ are then:

$$\begin{array}{lll} A\subset\mathcal{X} & \text{with mass} & (1-\alpha)\cdot m_S(A) \\ \mathcal{X} & \text{with mass} & (1-\alpha)\cdot m_S(\mathcal{X}) + \sum_{A\subseteq\mathcal{X}} \alpha\cdot m_S(A) \\ & & = (1-\alpha)\cdot m_S(\mathcal{X}) + \alpha \ . \end{array}$$

- So, FEs of $m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}}$ are:

$$\begin{array}{lll} A\times\{R\} & \text{with mass} & (1-\alpha)\cdot m_S(A) \\ A\times\{R\}\cup\mathcal{X}\times\{NR\} & \text{with mass} & \alpha\cdot m_S(A) \end{array}$$

  for all $A\subseteq\mathcal{X}$.

- FEs of $(m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}}$ are then:

$$\begin{array}{lll} A\subset\mathcal{X} & \text{with mass} & (1-\alpha)\cdot m_S(A) \\ \mathcal{X} & \text{with mass} & (1-\alpha)\cdot m_S(\mathcal{X}) + \sum_{A\subseteq\mathcal{X}} \alpha\cdot m_S(A) \\ & & = (1-\alpha)\cdot m_S(\mathcal{X}) + \alpha \ . \end{array}$$

$$(m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}} = {}^{\alpha}m_S$$

Discounting

Remarks

- With $\left\{ \begin{array}{lcl} m^{\mathcal{R}}(\{R\}) & = & 1-\alpha, \\ m^{\mathcal{R}}(\{NR\}) & = & \alpha\ , \end{array} \right.$ we also obtain:

$$(m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}} = {}^{\alpha}m_S\ .$$

- With $\left\{ \begin{array}{lcl} m^{\mathcal{R}}(\{NR\}) & = & 1-\alpha, \\ m^{\mathcal{R}}(\mathcal{R}) & = & \alpha\ , \end{array} \right.$ we obtain:

$$(m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap} m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}} = m_{\mathcal{X}}\ .$$

Discounting

Simpler expression

- Discounting of a MF m can be expressed, with $\alpha \in [0, 1]$, by:

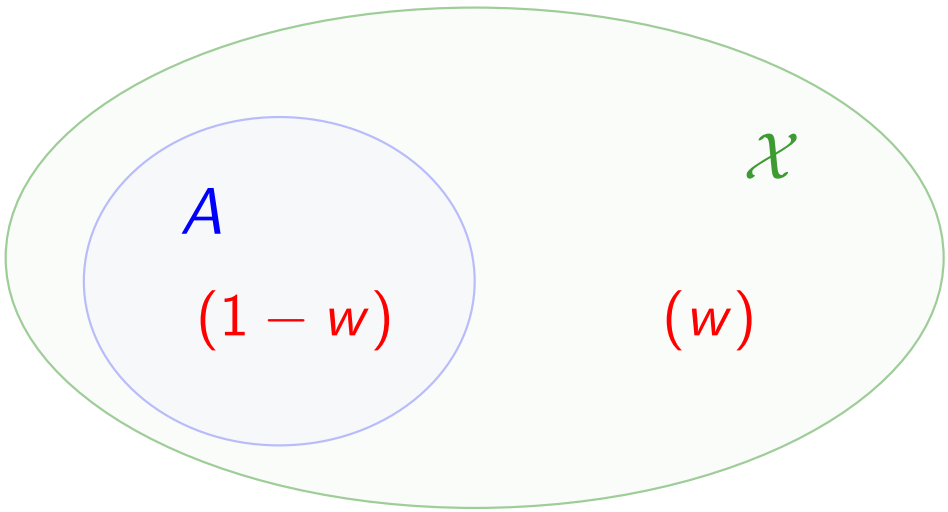
$$\begin{cases} {}^{\alpha}m(A) &= (1-\alpha)m(A), \quad \forall A \subset \mathcal{X}, \\ {}^{\alpha}m(\mathcal{X}) &= (1-\alpha)m(\mathcal{X}) + \alpha \ , \end{cases}$$

- or, using categorical (or logical) MF $m_{\mathcal{X}}$ ($m_{\mathcal{X}}(\mathcal{X}) = 1$):

$${}^{\alpha}m = (1-\alpha)m + \alpha m_{\mathcal{X}} \ .$$

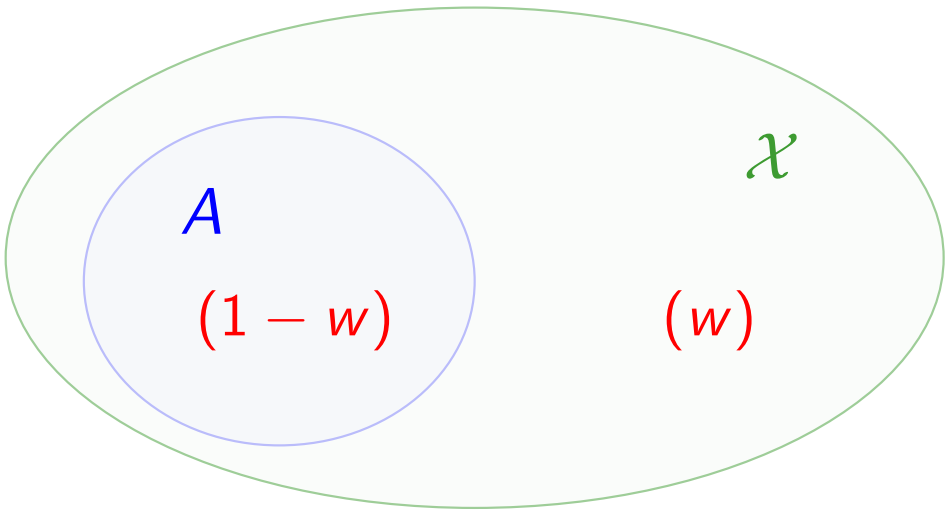
Simple and negative MF (recalls)

- A MF m defined by $m(\mathcal{X}) = w$ and $m(A) = 1 - w$, with $w \in [0, 1]$ and $A \subset \mathcal{X}$, can be conveniently noted A^w. It is called a simple MF.



Simple and negative MF (recalls)

- A MF m defined by $m(\mathcal{X}) = w$ and $m(A) = 1 - w$, with $w \in [0, 1]$ and $A \subset \mathcal{X}$, can be conveniently noted A^w. It is called a simple MF.



- A MF m such that $m(\emptyset) = v$ and $m(A) = 1 - v$, with $v \in [0, 1]$, $A \subseteq \mathcal{X}$, $A \neq \emptyset$, can be conveniently noted A_v. It is called a negative simple MF.

Negation of a MF (recalls)

- The negation $\overline{m}$ of a MF m is defined by $\overline{m}(A) = m(\overline{A})$ for all $A \subset \mathcal{X}$.
- we have:

$$\overline{A^w} = \overline{\left\{ \begin{array}{lcl} \mathcal{X} & \mapsto & w \\ A & \mapsto & 1-w \end{array} \right.} = \left\{ \begin{array}{lcl} \overline{\mathcal{X}} = \emptyset & \mapsto & w \\ \overline{A} & \mapsto & 1-w \end{array} \right. = \overline{A}_w \ .$$

Discounting

Expressions

- Discounting of a MF m can be expressed, with $\alpha \in [0,1]$, by:

$$\begin{cases} {}^{\alpha}m(A) &= (1-\alpha)m(A), \quad \forall A \subset \mathcal{X}, \\ {}^{\alpha}m(\mathcal{X}) &= (1-\alpha)m(\mathcal{X}) + \alpha \ , \end{cases}$$

- or, using categorical (or logical) MF $m_{\mathcal{X}}$ ($m_{\mathcal{X}}(\mathcal{X}) = 1$):

$${}^{\alpha}m = (1-\alpha)m + \alpha m_{\mathcal{X}} \ ,$$

- or, using negative simple MF:

$${}^{\alpha}m = m \textcircled{\cup} \mathcal{X}_{1-\alpha} = m \textcircled{\cup} \begin{cases} \emptyset &\mapsto 1-\alpha \\ \mathcal{X} &\mapsto \alpha \ . \end{cases}$$

References 1

[1] G. Shafer. *A mathematical theory of evidence*. Princeton, N.J.: Princeton University Press, 1976 (See Page 252)

[2] P. Smets. "Belief functions: the disjunctive rule of combination and the generalized bayesian theorem". In: *International Journal of Approximate Reasoning* 9 (1993), pp. 1–35 (See Section 5.7)

[3] D. Mercier, B. Quost, and T. Denœux. "Refined modeling of sensor reliability in the belief function framework using contextual discounting". In: *Information Fusion* 9.2 (Apr. 2008), pp. 246–258 (See Section 2.5)

[4] P. Smets. "The application of the matrix calculus to belief functions". In: *International Journal of Approximate Reasoning* 31.1–2 (2002), pp. 1–30

Outline

Discounting: a detailed presentation

A parametrized family of correction mechanisms

Pichon et al.'s Behaviour Based Correction (BBC)

Contextual discounting (CD) and reinforcement (CR)

Learning CD and CR from labelled data

- Discounting case: the source can be in two states: reliable (R_1) or not reliable (R_2), $\mathcal{R} = \{R_1, R_2\}$, where R_1 (resp. R_2) means that $m[\{R_1\}] = m_S$ (resp. $m[\{R_2\}] = m_{\mathcal{X}}$), and $m^{\mathcal{R}}(\{R_1\}) = 1 - \alpha$ and $m^{\mathcal{R}}(\{R_2\}) = \alpha$.

- Discounting case: the source can be in two states: reliable (R_1) or not reliable (R_2), $\mathcal{R} = \{R_1, R_2\}$, where R_1 (resp. R_2) means that $m[\{R_1\}] = m_S$ (resp. $m[\{R_2\}] = m_{\mathcal{X}}$), and $m^{\mathcal{R}}(\{R_1\}) = 1 - \alpha$ and $m^{\mathcal{R}}(\{R_2\}) = \alpha$.
- **Idea**: Let us consider a more general situation where the source can be in n states R_i, $i \in I = \{1, \ldots, n\}$, $\mathcal{R} = \cup_{i \in I}\{R_i\}$, the meaning of each state being specified by $m[\{R_i\}] = f_i(m_S) = m_i$, $\forall i \in I$. Moreover: $m^{\mathcal{R}}(\{R_i\}) = \nu_i \in [0, 1]$, $\forall i \in I$ (so $\sum_i \nu_i = 1$)

- Discounting case: the source can be in two states: reliable (R_1) or not reliable (R_2), $\mathcal{R} = \{R_1, R_2\}$, where R_1 (resp. R_2) means that $m[\{R_1\}] = m_S$ (resp. $m[\{R_2\}] = m_{\mathcal{X}}$), and $m^{\mathcal{R}}(\{R_1\}) = 1 - \alpha$ and $m^{\mathcal{R}}(\{R_2\}) = \alpha$.
- **Idea**: Let us consider a more general situation where the source can be in n states R_i, $i \in I = \{1, \ldots, n\}$, $\mathcal{R} = \cup_{i \in I}\{R_i\}$, the meaning of each state being specified by $m[\{R_i\}] = f_i(m_S) = m_i$, $\forall i \in I$. Moreover: $m^{\mathcal{R}}(\{R_i\}) = \nu_i \in [0, 1]$, $\forall i \in I$ (so $\sum_i \nu_i = 1$)
- All these pieces of information can be combined as follows:

$$(\textcircled{\cap}_{i=1}^{n} m^{\mathcal{X}}[\{R_i\}]^{\Uparrow \mathcal{X} \times \mathcal{R}} \textcircled{\cap} m^{\mathcal{R} \uparrow \mathcal{X} \times \mathcal{R}})^{\downarrow \mathcal{X}}$$

A parametrized family of correction mechanisms

Let us do the computations

A parametrized family of correction mechanisms

Let us do the computations (or not)

. . . Let as an exercise . . .

. . . Let as an exercise . . .

Result is given by:

$$(\circledcap_{i=1}^{n} m^{\mathcal{X}}[\{R_i\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \circledcap m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}} = \sum_i \nu_i m_i \ .$$

A parametrized family of correction mechanisms

Let us do the computations (or not)

... Let as an exercise ...

Result is given by:

$$(\circledcap_{i=1}^{n} m^{\mathcal{X}}[\{R_i\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \circledcap m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}} = \sum_i \nu_i m_i \ .$$

Discounting case:

$$(\circledcap_{i=1}^{2} m^{\mathcal{X}}[\{R_i\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \circledcap m^{\mathcal{R}\uparrow\mathcal{X}\times\mathcal{R}})^{\downarrow\mathcal{X}} = (1-\alpha)m_S + \alpha m_{\mathcal{X}} \ .$$

De-discounting (Denœux and Smets, 2006) is also retrieved

- De-discounting = Discounting reverse
- We receive a MF ${}^{\alpha}m$ discounted with a rate $\alpha < 1$.
- If α is known, m can be computed:

$$m = \frac{{}^{\alpha}m - \alpha\, m_{\mathcal{X}}}{1-\alpha}\ . \tag{3}$$

- Else any value α in $[0, {}^{\alpha}m(\mathcal{X})]$ is possible, indeed ${}^{\alpha}m(\mathcal{X})$ is the maximum value of α such that (3) leads to a MF. De-discounting with $\alpha = {}^{\alpha}m(\mathcal{X})$ yields to the following MF (totally de-discounted):

$${}^{tdd}m(A) = \begin{cases} \frac{m(A)}{1-m(\mathcal{X})} & \forall A \subset \mathcal{X}\ , \\ 0 & \text{otherwise.} \end{cases}$$

De-discounting (Denœux and Smets, 2006) is also retrieved

- De-discounting = Discounting reverse
- We receive a MF ${}^{\alpha}m$ discounted with a rate $\alpha < 1$.
- If α is known, m can be computed:

$$m = \frac{{}^{\alpha}m - \alpha\, m_{\mathcal{X}}}{1-\alpha}\ . \tag{3}$$

- Else any value α in $[0, {}^{\alpha}m(\mathcal{X})]$ is possible, indeed ${}^{\alpha}m(\mathcal{X})$ is the maximum value of α such that (3) leads to a MF. De-discounting with $\alpha = {}^{\alpha}m(\mathcal{X})$ yields to the following MF (totally de-discounted):

$${}^{tdd}m(A) = \begin{cases} \frac{m(A)}{1-m(\mathcal{X})} & \forall A \subset \mathcal{X}\ , \\ 0 & \text{otherwise.} \end{cases}$$

- We can show that $\nu m + (1-\nu)^{tdd}m$, with $\nu \in [0,1]$, is equivalent to (3) with $\nu = \frac{m(\mathcal{X})-\alpha}{(1-\alpha)m(\mathcal{X})}$, $\alpha \in [0, m(\mathcal{X})]$ (Exercice).

Illustration of the interest of one this mechanism in a car-to-car communication application (Bou Farah's PhD thesis, 2014) 1/2

- Vehicles exchange messages about events happening on the road.
- Information about each event **e** is represented by a MF $m^{\mathcal{X}}$ with $\mathcal{X} = \{\exists, \nexists\}$ in each message ($\exists$ meaning event **e** exists, $\nexists$ meaning event **e** does not exist).

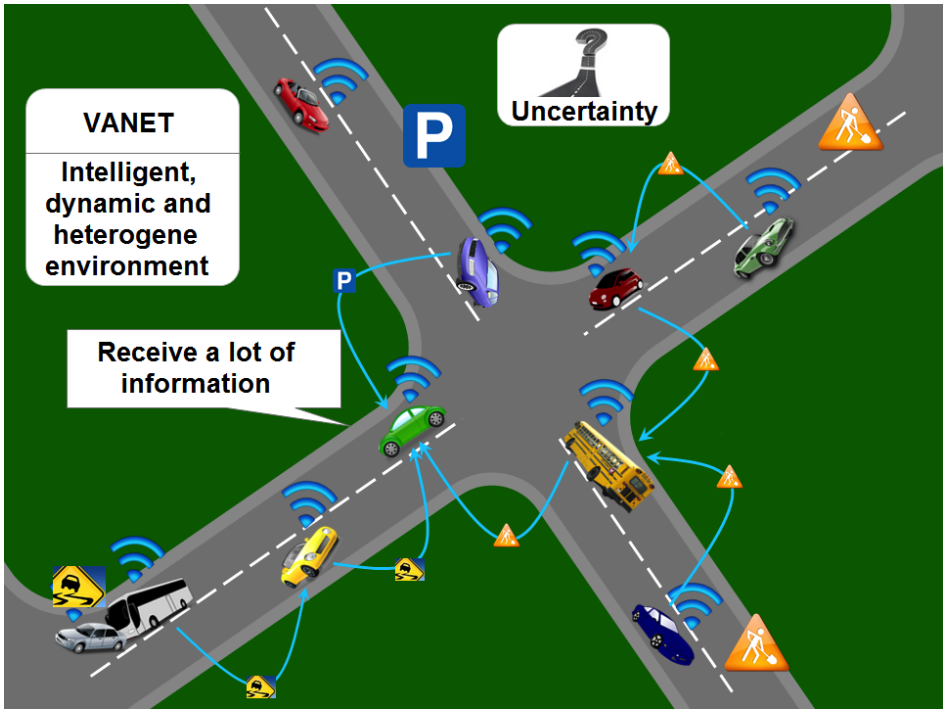


Illustration of the interest of one this mechanism in a car-to-car communication application (Bou Farah's PhD thesis, 2014) 2/2

Two strategies for messages ageing:

1. Either discount MFs $m^{\mathcal{X}}$: $(1-\alpha)m^{X}+\alpha m_{\mathcal{X}}$, with $\alpha \in [0,1]$ (over time, we do not know if the event is present or not).

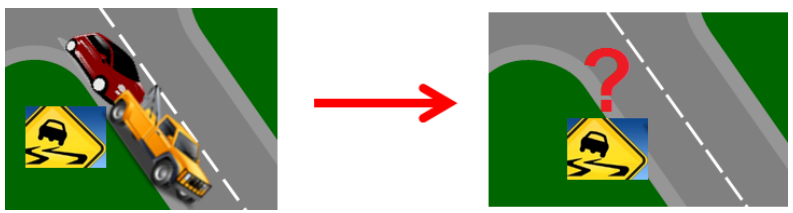
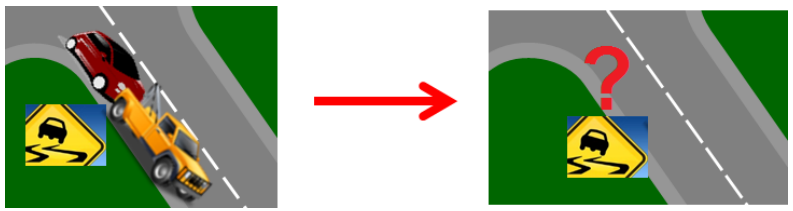


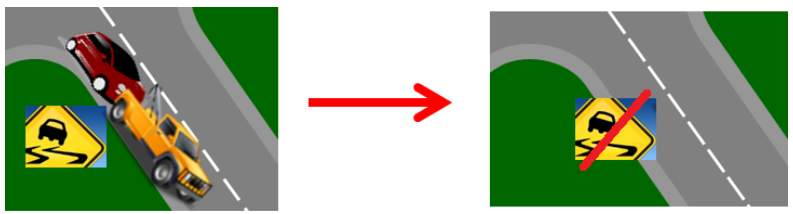
Illustration of the interest of one this mechanism in a car-to-car communication application (Bou Farah's PhD thesis, 2014) 2/2

Two strategies for messages ageing:

1. Either discount MFs $m^{\mathcal{X}}$: $(1-\alpha)m^{X}+\alpha m_{\mathcal{X}}$, with $\alpha \in [0,1]$ (over time, we do not know if the event is present or not).



2. Or use the following mechanisms $(1-\nu)m^{X}+\nu m_{\{\nexists\}}$, with $\nu \in [0,1]$ (over time we think the event is going to disappear).



References 2

[5] D. Mercier, T. Denœux, and M.-H. Masson. "Belief function correction mechanisms". In: *Studies in Fuzziness and Soft Computing, B. Bouchon-Meunier et al. (Eds.)* 249 (Jan. 2010), pp. 203–222

[6] T. Denœux and P. Smets. "Classification using Belief Functions: the Relationship between the Case-based and Model-based Approaches". In: *IEEE transactions on Systems, Man and Cybernetics B* 36.6 (2006), pp. 1395–1406 (See Sections 2.4)

[7] M. Bou Farah et al. "A high-level application using belief functions for exchanging and managing uncertain events on the road in vehicular ad-hoc networks". In: *Annals of telecommunications, special issue on belief functions and uncertainty management in networks and telecommunication* 69 (3-4 Apr. 2014), pp. 185–199 (See Sections 2.2.2 and 4.3)

[8] D. Mercier et al. "Decision fusion for postal address recognition using belief functions". In: *Expert Systems with Applications* 36 (3 2009), pp. 5643–5653 (See Section 3.3)

Outline

Discounting: a detailed presentation

A parametrized family of correction mechanisms

Pichon et al.'s Behaviour Based Correction (BBC)

Contextual discounting (CD) and reinforcement (CR)

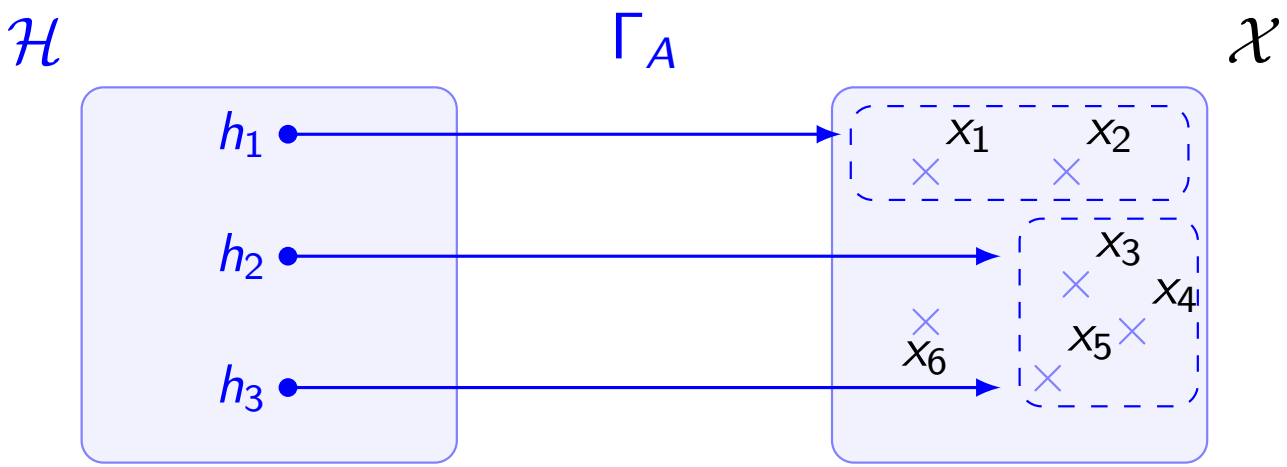
Learning CD and CR from labelled data

Pichon et al.'s Behaviour Based Correction (BBC)

Pichon et al., ijar 2012

In this model:

- a source provides a MF $m^{\mathcal{Y}}$,
- the state (or configuration) h in which stands the source, is described by a MF $m^{\mathcal{H}}$,
- and for all $A \subseteq \mathcal{Y}$ a function Γ_A defined from $\mathcal{H}$ to $2^{\mathcal{X}}$ which indicates how to interpret the piece of information $y \in A \subseteq \mathcal{Y}$ for each state $h \in \mathcal{H}$.



Pichon et al.'s Behaviour Based Correction (BBC)

Thermometer example

- Consider a thermometer that can be:
  - reliable: what it outputs is correct.
  - biased (approximately reliable): if it provides a temperature t, the true one is in $\{t-1, t, t+1, t+2, t+3, t+4\}$.
  - unreliable: what it outputs is incorrect. Our state of knowledge remains ignorance.

Pichon et al.'s Behaviour Based Correction (BBC)

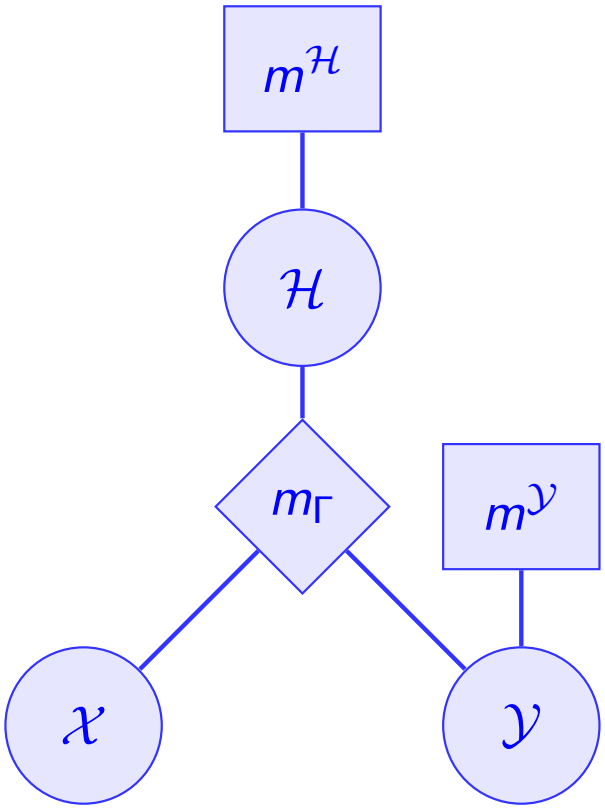
Thermometer example

- Consider a thermometer that can be:
  - reliable: what it outputs is correct.
  - biased (approximately reliable): if it provides a temperature t, the true one is in $\{t-1, t, t+1, t+2, t+3, t+4\}$.
  - unreliable: what it outputs is incorrect. Our state of knowledge remains ignorance.
- Model:
  - $\mathcal{Y} = \mathcal{X}$ the set of temperatures,
  - $\mathcal{H} = \{\mathit{reliable}, \mathit{biased}, \mathit{unreliable}\}$,
  - and Γ_A is defined for all $A \subseteq \mathcal{X}$ by:

$$\begin{array}{lcl} \Gamma_A(\mathit{reliable}) & = & A \ , \\ \Gamma_A(\mathit{biased}) & = & A \cup_{t \in A} \{t-1, t+1, t+2, t+3, t+4\} \ , \\ \Gamma_A(\mathit{unreliable}) & = & \mathcal{X} \ . \end{array}$$

Pichon et al.'s Behaviour Based Correction (BBC)

BBC evidential networks (See Lecture 6 - P. Shenoy)



With m_Γ defined by:

$$m_\Gamma^{\mathcal{H}\times\mathcal{X}\times\mathcal{Y}}\left(\cup_{h\in\mathcal{H},A\subseteq\mathcal{Y}}\{h\}\times A\times\Gamma_A(h)\right)=1.$$

Pichon et al.'s Behaviour Based Correction (BBC)

Result

- To have the information on $\mathcal{X}$, pieces of information $m^{\mathcal{Y}}, m^{\mathcal{H}}, \{\Gamma_A, A \subseteq \mathcal{Y}\} = m^{\mathcal{Y}}, m^{\mathcal{H}}, m_{\Gamma}^{\mathcal{H}\times\mathcal{X}\times\mathcal{Y}}$ can be combined as follows:

$$\left(\left(m^{\mathcal{Y}\uparrow\mathcal{H}\times\mathcal{X}\times\mathcal{Y}} \textcircled{\cap} m_{\Gamma}^{\mathcal{H}\times\mathcal{X}\times\mathcal{Y}}\right)^{\downarrow\mathcal{H}\times\mathcal{X}} \textcircled{\cap} m^{\mathcal{H}\uparrow\mathcal{H}\times\mathcal{X}}\right)^{\downarrow\mathcal{X}}.$$

Pichon et al.'s Behaviour Based Correction (BBC)

Result

- To have the information on $\mathcal{X}$, pieces of information $m^{\mathcal{Y}}, m^{\mathcal{H}}, \{\Gamma_A,\, A \subseteq \mathcal{Y}\} = m^{\mathcal{Y}}, m^{\mathcal{H}}, m_{\Gamma}^{\mathcal{H}\times\mathcal{X}\times\mathcal{Y}}$ can be combined as follows:

$$\left(\left(m^{\mathcal{Y}\uparrow\mathcal{H}\times\mathcal{X}\times\mathcal{Y}} \textcircled{\cap} m_{\Gamma}^{\mathcal{H}\times\mathcal{X}\times\mathcal{Y}}\right)^{\downarrow\mathcal{H}\times\mathcal{X}} \textcircled{\cap} m^{\mathcal{H}\uparrow\mathcal{H}\times\mathcal{X}}\right)^{\downarrow\mathcal{X}} .$$

- Result is given for all $C \subseteq \mathcal{X}$ by:

$$m^{\mathcal{X}}(C) = \sum_{H\subseteq\mathcal{H}} m^{\mathcal{H}}(H) \sum_{B:\Gamma_B(H)=C} m^{\mathcal{Y}}(B),$$

with $\Gamma_B(H) = \cup_{h\in H}\Gamma_B(h)$.

Pichon et al.'s Behaviour Based Correction (BBC)

Result

- To have the information on $\mathcal{X}$, pieces of information $m^{\mathcal{Y}}, m^{\mathcal{H}}, \{\Gamma_A,\ A \subseteq \mathcal{Y}\} = m^{\mathcal{Y}}, m^{\mathcal{H}}, m_{\Gamma}^{\mathcal{H}\times\mathcal{X}\times\mathcal{Y}}$ can be combined as follows:

$$\left(\left(m^{\mathcal{Y}\uparrow\mathcal{H}\times\mathcal{X}\times\mathcal{Y}} \textcircled{\cap} m_{\Gamma}^{\mathcal{H}\times\mathcal{X}\times\mathcal{Y}}\right)^{\downarrow\mathcal{H}\times\mathcal{X}} \textcircled{\cap} m^{\mathcal{H}\uparrow\mathcal{H}\times\mathcal{X}}\right)^{\downarrow\mathcal{X}} .$$

- Result is given for all $C \subseteq \mathcal{X}$ by:

$$m^{\mathcal{X}}(C) = \sum_{H\subseteq\mathcal{H}} m^{\mathcal{H}}(H) \sum_{B:\Gamma_B(H)=C} m^{\mathcal{Y}}(B),$$

with $\Gamma_B(H) = \cup_{h\in H}\Gamma_B(h)$.

As an exercise: Do the computations ;)

Pichon et al.'s Behaviour Based Correction (BBC)

Particular case: Discounting

Discounting can be retrieved by choosing for example:

- $\mathcal{Y} = \mathcal{X}$,
- $\mathcal{H} = \{h_1, h_2\}$, with $m^{\mathcal{H}}(\{h_1\}) = \beta = 1 - \alpha$ and $m^{\mathcal{H}}(\{h_2\}) = \alpha$,
- and Γ_A is defined for all $A \subseteq \mathcal{X}$ by:

$$\begin{aligned} \Gamma_A(h_1) &= A\ , \\ \Gamma_A(h_2) &= \mathcal{X}\ . \end{aligned}$$

Pichon et al.'s Behaviour Based Correction (BBC)

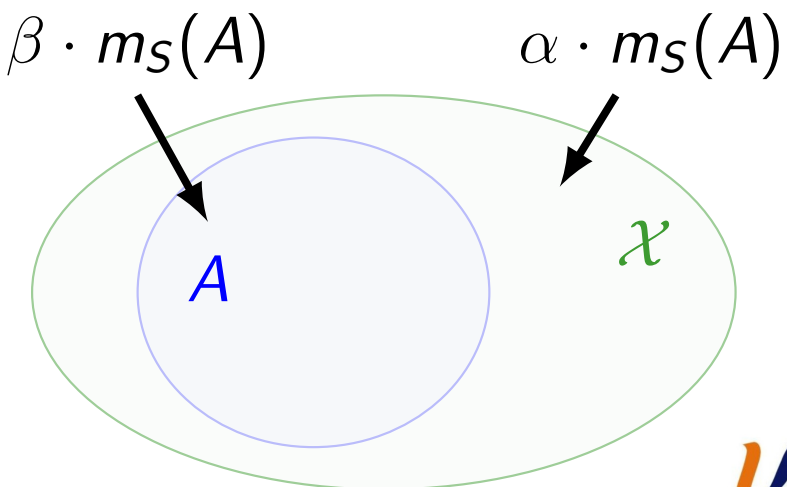
Particular case: Discounting

Discounting can be retrieved by choosing for example:

- $\mathcal{Y} = \mathcal{X}$,
- $\mathcal{H} = \{h_1, h_2\}$, with $m^{\mathcal{H}}(\{h_1\}) = \beta = 1 - \alpha$ and $m^{\mathcal{H}}(\{h_2\}) = \alpha$,
- and Γ_A is defined for all $A \subseteq \mathcal{X}$ by:

$$\begin{aligned} \Gamma_A(h_1) &= A\ , \\ \Gamma_A(h_2) &= \mathcal{X}\ . \end{aligned}$$

Discounting: for each FE A of m_S, a part β remains on A and a part α is transferred to $\mathcal{X}$.



References 3

[9] F. Pichon, D. Dubois, and T. Denœux. "Relevance and truthfulness in information correction and fusion". In: *International Journal of Approximate Reasoning* 53.2 (2012), pp. 159–175

Outline

Discounting: a detailed presentation

A parametrized family of correction mechanisms

Pichon et al.'s Behaviour Based Correction (BBC)

Contextual discounting (CD) and reinforcement (CR)

Learning CD and CR from labelled data

Contextual discounting (CD)

Introduction

- **Idea** (Mercier et al., 2008): the reliability of the source may depend on the true value of the variable of interest **x**.
- Example 1: A thermometer reliable if the temperature is in the range $[0°C, 100°C]$, unreliable otherwise.
- Example 2: A cardiologist may be more reliable to diagnose cardiac disease than other kinds of disease.

Contextual discounting (CD)

Introduction

- **Idea** (Mercier et al., 2008): the reliability of the source may depend on the true value of the variable of interest **x**.
- Example 1: A thermometer reliable if the temperature is in the range $[0°C, 100°C]$, unreliable otherwise.
- Example 2: A cardiologist may be more reliable to diagnose cardiac disease than other kinds of disease.
- Model:
  - Let $\mathcal{A}$ be a partition of $\mathcal{X}$ representing different contexts,
  - The source may be reliable (R) or not (NR), $\mathcal{R} = \{R, NR\}$, $m^{\mathcal{X}}[\{R\}] = m_S^{\mathcal{X}}$, $m^{\mathcal{X}}[\{NR\}] = m_{\mathcal{X}}$.
  - For all contexts $A_k \in \mathcal{A}$: $m^{\mathcal{R}}[A_k](\{R\}) = \beta_k$ and $m^{\mathcal{R}}[A_k](\mathcal{R}) = \alpha_k$.

Contextual discounting (CD)

Introduction

- **Idea** (Mercier et al., 2008): the reliability of the source may depend on the true value of the variable of interest **x**.
- Example 1: A thermometer reliable if the temperature is in the range $[0°C, 100°C]$, unreliable otherwise.
- Example 2: A cardiologist may be more reliable to diagnose cardiac disease than other kinds of disease.
- Model:
  - Let $\mathcal{A}$ be a partition of $\mathcal{X}$ representing different contexts,
  - The source may be reliable (R) or not (NR), $\mathcal{R} = \{R, NR\}$, $m^{\mathcal{X}}[\{R\}] = m_S^{\mathcal{X}}$, $m^{\mathcal{X}}[\{NR\}] = m_{\mathcal{X}}$.
  - For all contexts $A_k \in \mathcal{A}$: $m^{\mathcal{R}}[A_k](\{R\}) = \beta_k$ and $m^{\mathcal{R}}[A_k](\mathcal{R}) = \alpha_k$.
  - Remark: if $\mathcal{A} = \{\mathcal{X}\}$ we retrieve the discounting model: $m^{\mathcal{R}}[\mathcal{X}](\{R\}) = m^{\mathcal{R}}(\{R\}) = \beta$ and $m^{\mathcal{R}}[\mathcal{X}](\mathcal{R}) = m^{\mathcal{R}}(\mathcal{R}) = \alpha$.

Contextual discounting (CD)

Derivation

- Available evidence can be synthesized by:

$$\left(m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap}_{A_k\in\mathcal{A}} m^{\mathcal{R}}[A_k]^{\Uparrow \mathcal{X}\times\mathcal{R}}\right)^{\downarrow\mathcal{X}}$$

Contextual discounting (CD)

Derivation

- Available evidence can be synthesized by:

$$\left(m^{\mathcal{X}}[\{R\}]^{\Uparrow \mathcal{X}\times\mathcal{R}} \textcircled{\cap}_{A_k\in\mathcal{A}} m^{\mathcal{R}}[A_k]^{\Uparrow \mathcal{X}\times\mathcal{R}}\right)^{\downarrow\mathcal{X}}$$

- Result is given by:

$$m_S \textcircled{\cup}_{A\in\mathcal{A}} A_{\beta_A}$$

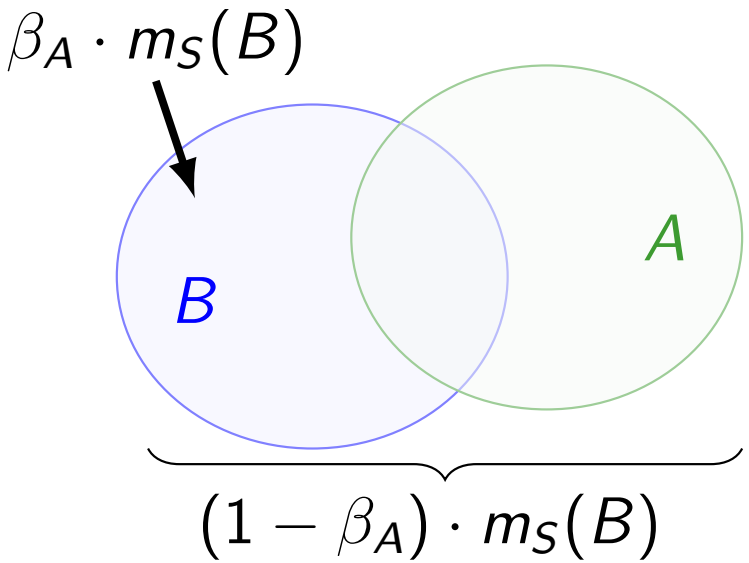
- Discounting is retrieved when $\mathcal{A} = \{\mathcal{X}\}$ ($\mathcal{A}$ contains one context $\mathcal{X}$):

$$m_S \textcircled{\cup} \mathcal{X}_\beta$$

Contextual discounting (CD)

CD results in terms of masses transfers

For each focal element B of m_S, for each context $A \in \mathcal{A}$:



- A portion $\beta_A \cdot m_S(B)$ remains on B.
- A portion $(1-\beta_A) \cdot m_S(B)$ is transferred to $B \cup A$.

Contextual discounting (CD)

Justification with BBC (Pichon et al. 2014)

- CD with one context $A \subseteq \mathcal{X}$ (*i.e.* $\mathcal{A} = \{A\}$) is equivalent to BBC with:
  - $\mathcal{Y} = \mathcal{X}$, $\mathcal{H} = \{h_1, h_2\}$ with $m^{\mathcal{H}}(\{h_1\}) = \beta_A$ and $m^{\mathcal{H}}(\{h_2\}) = \alpha_A$.
  - Interpretation of the two states h_1 and h_2 is given by $\Gamma_B(h_1) = B$ and $\Gamma_B(h_2) = B \cup A$, $\forall B \subseteq \mathcal{X}$.
- With both models, the same correction defined by:

$$f_{m^{\mathcal{H}}}(m) = m \textcircled{\cup} A_{\beta_A}$$

is obtained.

Contextual discounting (CD)

Generalisation with BBC (Pichon et al. 2014)

- For any set of contexts $\mathcal{A}$ (no partition is required), we have:

$$\alpha_{\cup} m = m \textcircled{\cup}_{A \in \mathcal{A}} A_{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m^{\mathcal{H}}_{A,\cup}})(m), \tag{4}$$

where:

- $\mathcal{H} = \{h_A | A \subseteq \mathcal{X}\}$
- Interpretations are given by $\Gamma_B(h_A) = B \cup A$, $\forall B \subseteq \mathcal{X}$;
- For all $A \in \mathcal{A}$, MFs $m^{\mathcal{H}}_{A,\cup}$ are defined by:

$$\begin{cases} m^{\mathcal{H}}_{A,\cup}(\{h_\emptyset\}) &= \beta_A \\ m^{\mathcal{H}}_{A,\cup}(\{h_A\}) &= \alpha_A \end{cases}$$

Contextual discounting (CD)

Example 1/3

With $\mathcal{X} = \{x_1, x_2, x_3\}$, $\mathcal{A} = \{\{x_1\}\}$, generalisation matrix ${}^{\alpha\cup}G$ associated with ${}^{\alpha\cup}m = m\textcircled{\cup}\ \{x_1\}_{\beta_1} = m\textcircled{\cup} \begin{cases} \emptyset & \mapsto & \beta_1 \\ \{x_1\} & \mapsto & \alpha_1 \end{cases}$ is given by:

$$\begin{pmatrix} \beta_1 & & & & & & & \\ \alpha_1 & 1 & & & & & & \\ & & \beta_1 & & & & & \\ & & \alpha_1 & 1 & & & & \\ & & & & \beta_1 & & & \\ & & & & \alpha_1 & 1 & & \\ & & & & & & \beta_1 & \\ & & & & & & \alpha_1 & 1 \end{pmatrix} \begin{matrix} \emptyset \\ \{x_1\} \\ \{x_2\} \\ \{x_1, x_2\} \\ \{x_3\} \\ \{x_1, x_3\} \\ \{x_2, x_3\} \\ \{x_1, x_2, x_3\} \end{matrix}$$

Contextual discounting (CD)

Example 2/3

Discounting is retrieved with $\mathcal{X} = \{x_1, x_2, x_3\}$, $\mathcal{A} = \{\mathcal{X}\}$: generalisation matrix ${}^{\alpha_\cup}G$ associated with ${}^{\alpha_\cup}m = m\textcircled{\cup}\ \{\mathcal{X}\}_\beta = m\textcircled{\cup} \left\{ \begin{array}{lcl} \emptyset & \mapsto & \beta \\ \mathcal{X} & \mapsto & \alpha \end{array} \right.$ is given by:

$$\left(\begin{array}{cccccccc} \beta & & & & & & & \\ & \beta & & & & & & \\ & & \beta & & & & & \\ & & & \beta & & & & \\ & & & & \beta & & & \\ & & & & & \beta & & \\ & & & & & & \beta & \\ \alpha & \alpha & \alpha & \alpha & \alpha & \alpha & \alpha & 1 \end{array} \right) \begin{array}{c} \emptyset \\ \{x_1\} \\ \{x_2\} \\ \{x_1, x_2\} \\ \{x_3\} \\ \{x_1, x_3\} \\ \{x_2, x_3\} \\ \{x_1, x_2, x_3\} \end{array}$$

Contextual discounting (CD)

Example 3/3

With $\mathcal{X} = \{x_1, x_2, x_3\}$, $\mathcal{A} = \{\{x_1\}, \{x_2\}, \{x_3\}\}$, generalisation matrix ${}^{\alpha\cup}G$ associated with ${}^{\alpha\cup}m = m\circledcup \{x_1\}_{\beta_1}\circledcup\{x_2\}_{\beta_2}\circledcup\{x_3\}_{\beta_3}$ is given by:

$$\left(\begin{array}{cccccccc}
\beta_1\beta_2\beta_3 & & & & & & & \\
\alpha_1\beta_2\beta_3 & \beta_2\beta_3 & & & & & & \\
\beta_1\alpha_2\beta_3 & & \beta_1\beta_3 & & & & & \\
\alpha_1\alpha_2\beta_3 & \alpha_2\beta_3 & \alpha_1\beta_3 & \beta_3 & & & & \\
\beta_1\beta_2\alpha_3 & & & & \beta_1\beta_2 & & & \\
\alpha_1\beta_2\alpha_3 & \beta_2\alpha_3 & & & \alpha_1\beta_2 & \beta_2 & & \\
\beta_1\alpha_2\alpha_3 & & \beta_1\alpha_3 & & \beta_1\alpha_2 & & \beta_1 & \\
\alpha_1\alpha_2\alpha_3 & \alpha_2\alpha_3 & \alpha_1\alpha_3 & \alpha_3 & \alpha_1\alpha_2 & \alpha_2 & \alpha_1 & 1
\end{array}\right)
\begin{array}{c}
\emptyset \\ \{x_1\} \\ \{x_2\} \\ \{x_1, x_2\} \\ \{x_3\} \\ \{x_1, x_3\} \\ \{x_2, x_3\} \\ \{x_1, x_2, x_3\}
\end{array}$$

References 4

[3] D. Mercier, B. Quost, and T. Denœux. "Refined modeling of sensor reliability in the belief function framework using contextual discounting". In: *Information Fusion* 9.2 (Apr. 2008), pp. 246–258

[9] F. Pichon, D. Dubois, and T. Denœux. "Relevance and truthfulness in information correction and fusion". In: *International Journal of Approximate Reasoning* 53.2 (2012), pp. 159–175

[10] F. Pichon et al. "Truthfulness in contextual information correction". In: *Proceedings of the 3rd International Conference on Belief Functions, BELIEF 2014*. Ed. by F. Cuzzolin. Vol. 8764. Lecture Notes in Computer Science. Oxford, United Kingdom: Springer International Publishing, Sept. 2014, pp. 11–20

Contextual reinforcement (CR)

Definition and results after applying CR (masses transfers)

- With $\mathcal{A}$ a set of subsets (contexts) of $\mathcal{X}$, CR of a MF m_S is the MF m defined, with $\beta_A \in [0,1]$, $\forall A \in \mathcal{A}$, by:

$$^{\boldsymbol{\alpha}\cap} m = m_S \textcircled{\cap}_{A\in\mathcal{A}} A^{\beta_A}$$

Contextual reinforcement (CR)

Definition and results after applying CR (masses transfers)

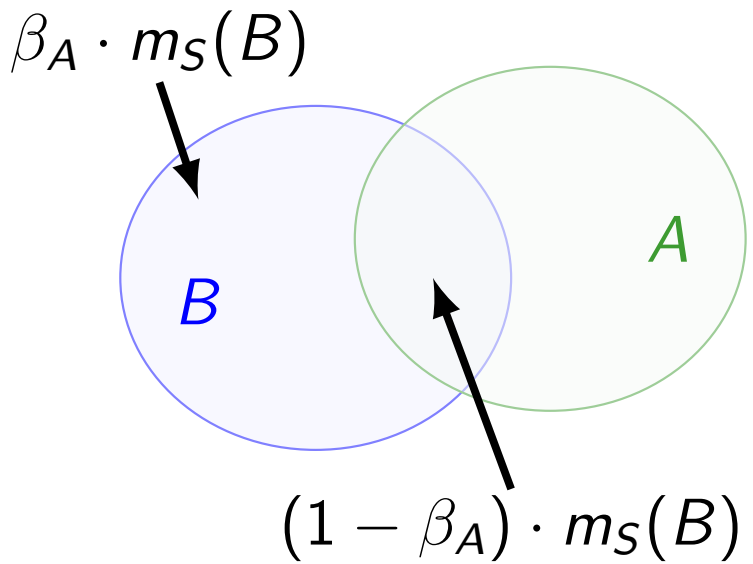
- With $\mathcal{A}$ a set of subsets (contexts) of $\mathcal{X}$, CR of a MF m_S is the MF m defined, with $\beta_A \in [0,1]$, $\forall A \in \mathcal{A}$, by:

$$^{\alpha\cap} m = m_S \textcircled{\cap}_{A\in\mathcal{A}} A^{\beta_A}$$

- For each focal element B of m_S, for each context $A \in \mathcal{A}$:

A part $\beta_A \cdot m_S(B)$ remains on B.

A part $(1-\beta_A)\cdot m_S(B)$ is transferred to $B \cap A$.



Contextual reinforcement (CR)

Justification (Pichon et al. 2014)

- For any set of contexts $\mathcal{A}$ (no partition is required), we have:

$$\alpha_{\cap} m = m \textcircled{\cap}_{A \in \mathcal{A}} A^{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m^{\mathcal{H}}_{A,\cap}})(m), \tag{5}$$

where:

- $\mathcal{H} = \{h_A | A \subseteq \mathcal{X}\}$
- Interpretations are given by $\Gamma_B(h_A) = B \cap A$, $\forall B \subseteq \mathcal{X}$;
- For all $A \in \mathcal{A}$, MFs $m^{\mathcal{H}}_{A,\cap}$ are defined by:

$$\begin{cases} m^{\mathcal{H}}_{A,\cap}(\{h_\emptyset\}) &= \beta_A \\ m^{\mathcal{H}}_{A,\cap}(\{h_A\}) &= \alpha_A \end{cases}$$

Contextual reinforcement (CR)

Example

With $\mathcal{X} = \{x_1, x_2, x_3\}$, $\mathcal{A} = \{\{x_1\}\}$, specialisation matrix ${}^{\alpha\cap}S$ associated with ${}^{\alpha\cap}m = m_{\textcircled{\cap}}\ \{x_1\}^{\beta_1} = m_{\textcircled{\cap}} \left\{ \begin{array}{ccc} \mathcal{X} & \mapsto & \beta_1 \\ \{x_1\} & \mapsto & \alpha_1 \end{array} \right.$ is given by:

$$\left(\begin{array}{ccccccc} 1 & & \alpha_1 & & \alpha_1 & & \alpha_1 \\ & 1 & & \alpha_1 & & \alpha_1 & & \alpha_1 \\ & & \beta_1 & & & & \\ & & & \beta_1 & & & \\ & & & & \beta_1 & & \\ & & & & & \beta_1 & \\ & & & & & & \beta_1 \\ & & & & & & & \beta_1 \end{array} \right) \begin{array}{c} \emptyset \\ \{x_1\} \\ \{x_2\} \\ \{x_1, x_2\} \\ \{x_3\} \\ \{x_1, x_3\} \\ \{x_2, x_3\} \\ \{x_1, x_2, x_3\} \end{array}$$

Outline

Discounting: a detailed presentation

A parametrized family of correction mechanisms

Pichon et al.'s Behaviour Based Correction (BBC)

Contextual discounting (CD) and reinforcement (CR)

Learning CD and CR from labelled data

Learning method

- From data composed of n objects o_i, $i \in \{1, \ldots, n\}$, for each of which we have:
  1. the true class of o_i belonging to $\mathcal{X} = \{x_1, \ldots, x_K\}$, and
  2. the MF $m_S\{o_i\}$ output by S regarding the class of o_i,
- From a chosen measure of discrepancy between beliefs and ground truth, we propose to obtain:
  1. the CD of m_S which optimizes this measure of discrepancy.
  2. the CR of m_S which optimizes this measure of discrepancy.

Example of labelled data

Inspired from (Elouedi, 2004)

- Data - 4 objects (o_1, o_2, o_3 and o_4) - 2 sensors S_1 and S_2:

		$\{a\}$	$\{h\}$	$\{r\}$	$\{a,h\}$	$\{a,r\}$	$\{h,r\}$	$\mathcal{X}$	Ground truth
S_1	$m_{S_1}\{o_1\}$	0	0	0.5	0	0	0.3	0.2	a
	$m_{S_1}\{o_2\}$	0	0.5	0.2	0	0	0	0.3	h
	$m_{S_1}\{o_3\}$	0	0.4	0	0	0.6	0	0	a
	$m_{S_1}\{o_4\}$	0	0	0	0	0.6	0.4	0	r
S_2	$m_{S_2}\{o_1\}$	0	0	0	0.7	0	0	0.3	a
	$m_{S_2}\{o_2\}$	0.3	0	0	0.4	0	0	0.3	h
	$m_{S_2}\{o_3\}$	0.2	0	0	0	0	0.6	0.2	a
	$m_{S_2}\{o_4\}$	0	0	0	0	0	1	0	r

Chosen measure of discrepancy

- Chosen measure of discrepancy between the corrected information and the ground truth:

$$E_{pl}(\beta) = \sum_{i=1}^{n} \sum_{k=1}^{K} (pl\{o_i\}(\{x_k\}) - \delta_{i,k})^2 \ ,$$

- With:
  - $pl\{o_i\}$: the plausibility function obtained from a contextual correction of m_S (CD or CR) with vector $\beta \in [0,1]^{|\mathcal{A}|}$.
  - $\delta_{i,k}$: binary variable indicating the class of o_i:

$$\begin{cases} \delta_{i,k} = 1 & \text{if object } o_i \text{ belongs to the class } x_k, \\ \delta_{i,k} = 0 & \text{otherwise.} \end{cases}$$

Advantages of E_{pl} measure

1. It yields to a least square optimization procedure (easy and quick to solve).
2. It allows us to have an easy understanding of CD and CR impacts on the measure.
3. It is at least as justified as other measures.

Learning CD and CR (Results)

Mercier et al. 2015

- With CD:
  1. Plausibilities on singletons are given, $\forall o_i$, by:

$$pl\{o_i\}(\{x\}) = 1 - (1 - pls\{o_i\}(\{x\})) \prod_{A \in \mathcal{A}, x \in A} \beta_A \ .$$

  2. Minimization of E_{pl} is obtained using the vector $\boldsymbol{\beta} = (\beta_{\{x_k\}}, k \in \{1, \ldots, K\})$ (i.e. with K contexts $\{x_k\}$, $k \in \{1, \ldots, K\}$).
- With CR:
  1. Plausibilities on singletons are given, $\forall o_i$, by:

$$pl\{o_i\}(\{x\}) = pls\{o_i\}(\{x\}) \prod_{A \in \mathcal{A}, x \notin A} \beta_A \ .$$

  2. Minimization of E_{pl} is obtained using the vector $\boldsymbol{\beta} = (\beta_{\overline{\{x_k\}}}, k \in \{1, \ldots, K\})$ (i.e. with K contexts $\overline{\{x_k\}}$, $k \in \{1, \ldots, K\}$).

CD and CR respective correction capacities regarding plausibilities on singletons

- With CD:
  - As $pl(\{x_k\}) = 1 - (1 - pl_S(\{x_k\}))\beta_{\{x_k\}}$, $\forall x_k \in \mathcal{X}$, with $\beta_{\{x_k\}}$ varying in $[0, 1]$.
  - So $pl(\{x_k\})$ can take any values in $[pl_S(\{x\}), 1]$.
  - The value on each singleton $pl_S(\{x\})$ can be shifted as close to 1 as required (i.e. weakened as required).

CD and CR respective correction capacities regarding plausibilities on singletons

- With CD:
  - As $pl(\{x_k\}) = 1 - (1 - pl_S(\{x_k\}))\beta_{\{x_k\}}$, $\forall x_k \in \mathcal{X}$, with $\beta_{\{x_k\}}$ varying in $[0, 1]$.
  - So $pl(\{x_k\})$ can take any values in $[pl_S(\{x\}), 1]$.
  - The value on each singleton $pl_S(\{x\})$ can be shifted as close to 1 as required (i.e. weakened as required).
- With CR:
  - As $pl(\{x_k\}) = pl_S(\{x_k\})\beta_{\overline{\{x_k\}}}$, $\forall x_k \in \mathcal{X}$, with $\beta_{\overline{\{x_k\}}}$ varying in $[0, 1]$.
  - So $pl(\{x_k\})$ can take any values in $[0, pl_S(\{x_k\})]$.
  - The value on each singleton $pl_S(\{x\})$ can be carried as close to 0 as necessary (i.e. strengthened as necessary).

CD and CR respective correction capacities regarding plausibilities on singletons

- With CD:
  - As $pl(\{x_k\}) = 1 - (1 - pl_S(\{x_k\}))\beta_{\{x_k\}}$, $\forall x_k \in \mathcal{X}$, with $\beta_{\{x_k\}}$ varying in $[0, 1]$.
  - So $pl(\{x_k\})$ can take any values in $[pl_S(\{x\}), 1]$.
  - The value on each singleton $pl_S(\{x\})$ can be shifted as close to 1 as required (i.e. weakened as required).
- With CR:
  - As $pl(\{x_k\}) = pl_S(\{x_k\})\beta_{\overline{\{x_k\}}}$, $\forall x_k \in \mathcal{X}$, with $\beta_{\overline{\{x_k\}}}$ varying in $[0, 1]$.
  - So $pl(\{x_k\})$ can take any values in $[0, pl_S(\{x_k\})]$.
  - The value on each singleton $pl_S(\{x\})$ can be carried as close to 0 as necessary (i.e. strengthened as necessary).
- Summary:
  - With CD, pl on singletons can be $\nearrow$ to 1 as required.
  - With CR, pl on singletons can be $\searrow$ to 0 as required.

Illustrative example

Inspired from (Elouedi, 2004) (Mercier, 2008)

- Data - 4 objects (o_1, o_2, o_3 and o_4) - 2 sensors S_1 and S_2:

		$\{a\}$	$\{h\}$	$\{r\}$	$\{a,h\}$	$\{a,r\}$	$\{h,r\}$	$\mathcal{X}$	Ground truth
S_1	$m_{S_1}\{o_1\}$	0	0	0.5	0	0	0.3	0.2	a
	$m_{S_1}\{o_2\}$	0	0.5	0.2	0	0	0	0.3	h
	$m_{S_1}\{o_3\}$	0	0.4	0	0	0.6	0	0	a
	$m_{S_1}\{o_4\}$	0	0	0	0	0.6	0.4	0	r
S_2	$m_{S_2}\{o_1\}$	0	0	0	0.7	0	0	0.3	a
	$m_{S_2}\{o_2\}$	0.3	0	0	0.4	0	0	0.3	h
	$m_{S_2}\{o_3\}$	0.2	0	0	0	0	0.6	0.2	a
	$m_{S_2}\{o_4\}$	0	0	0	0	0	1	0	r

- Results:

Contextual correction	Sensor 1	Sensor 2
CD	$\beta = (0.76, 1.00, 1.00)$ $E_{pl}(\beta) = 3.39$	$\beta = (0.74, 1.00, 1.00)$ $E_{pl}(\beta) = 4.81$
CR	$\beta = (0.94, 0.66, 0.38)$ $E_{pl}(\beta) = 2.33$	$\beta = (0.65, 0.22, 0.55)$ $E_{pl}(\beta) = 2.39$

- Comments:
  - For CD:
    - Both sensors are reliable to detect objects h and r ($\beta_{\{h\}} = \beta_{\{r\}} = 1$), and have some problems with object a.
    - There is no need to increase plausibilities on h and r, whereas $pl_S(\{a\})$ should be increased (a being not well recognized).
  - For CR:
    - More differences between sensors
    - For sensor 1: $pl_S(\{a\})$ should (almost) not be decreased, whereas $pl_S(\{h\})$ and even more $pl_S(\{r\})$ have to be decreased.
  - It is possible to obtain lower values for E_{pl} using CR instead of using CD.

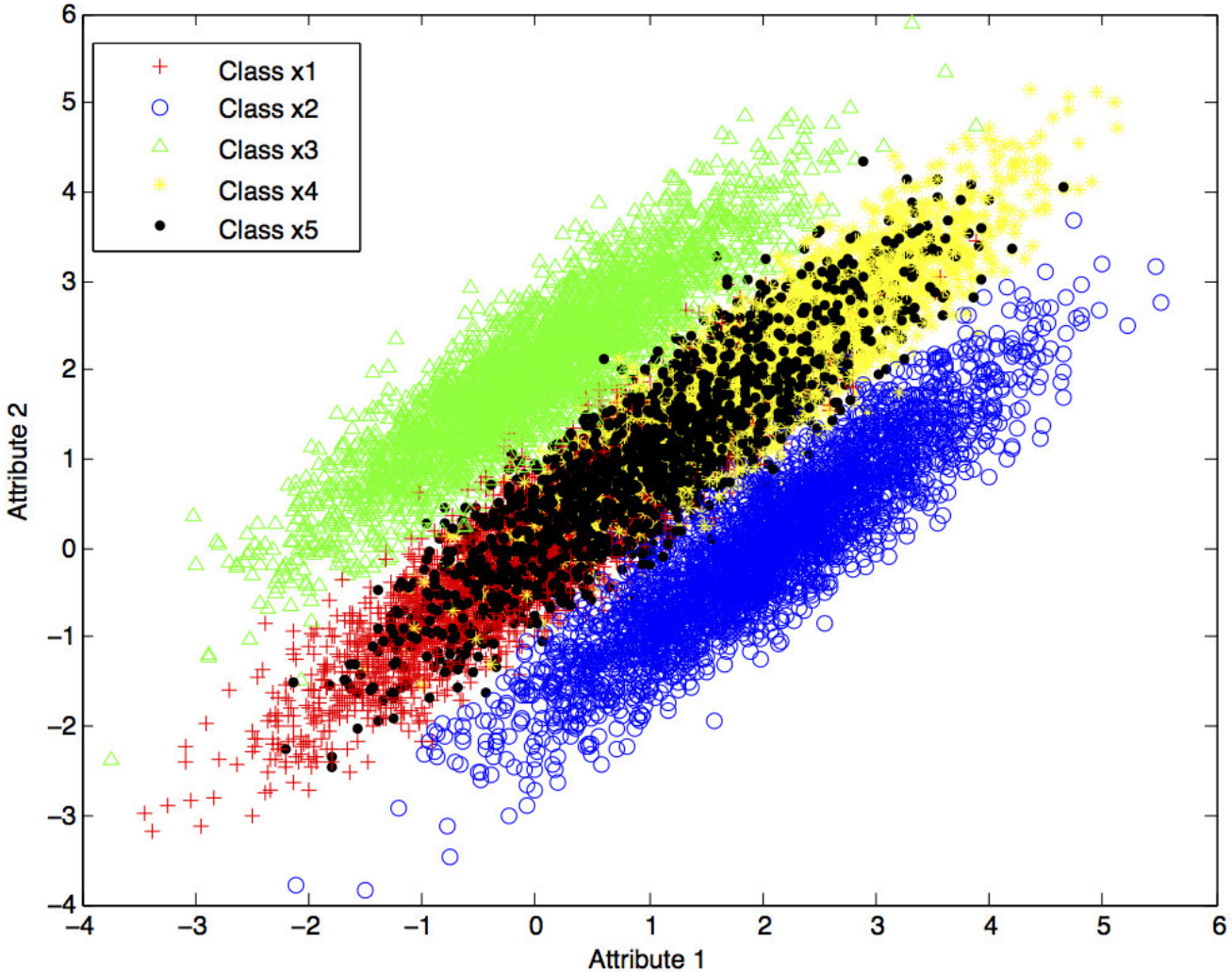
- **Goal**: we want to correct the information output by an evidential classifier using CD or CR.
- The evidential k-nearest neighbour classifier (*ev-knn*) introduced by Denœux (1995) is chosen with $k = 3$.
- 5-class classification problem with data generated from 5 bivariate normal distributions with respective means $\mu_{x_1} = (0, 0)$, $\mu_{x_2} = (2, 0)$, $\mu_{x_3} = (0, 2)$, $\mu_{x_4} = (2, 2)$, $\mu_{x_5} = (1, 1)$ and common variance matrix

$$\Sigma = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}.$$

- For each class $x \in \mathcal{X} = \{x_1, x_2, x_3, x_4, x_5\}$, 1000 instances have been generated. The total amount of data is then composed of 5000 instances.

An experiment in classification

Illustrations of the 5000 instances

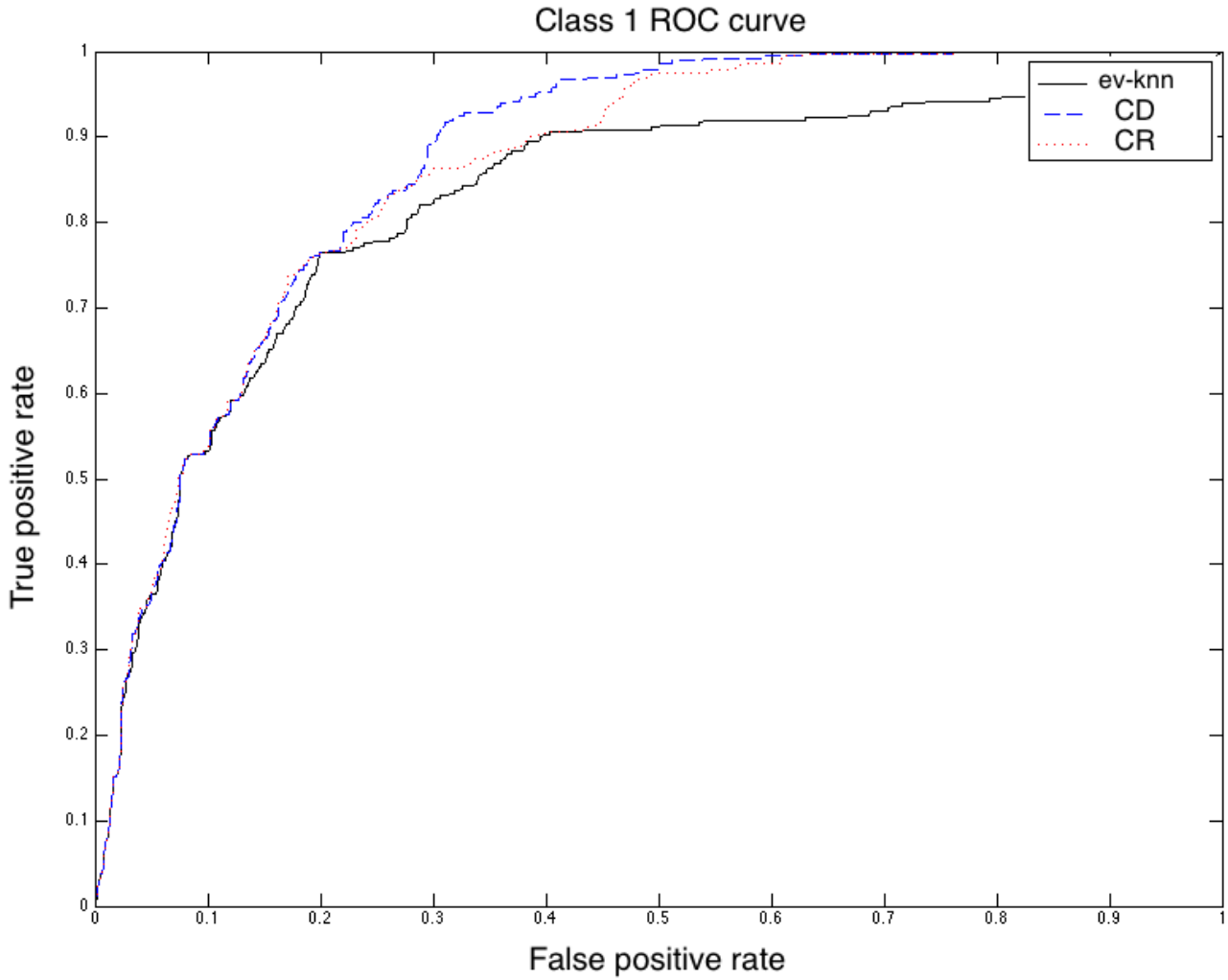


The 5000 instances have been divided into three parts:

1. The first third of the data constitutes the training set of ev-knn.
2. The second third of the data constitutes the training set for CD and CR.
3. The last third of the data constitutes the test set.

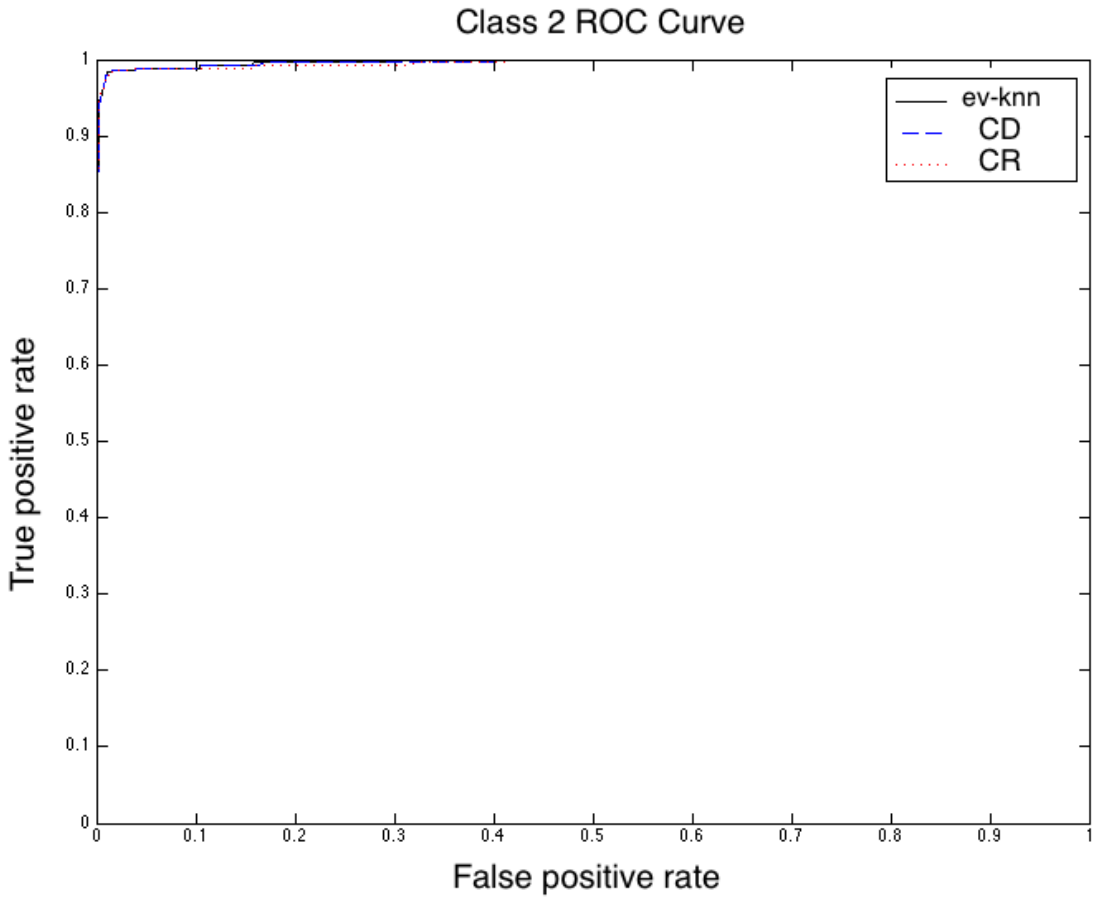
An experiment in classification

Results: Class 1 ROC Curve



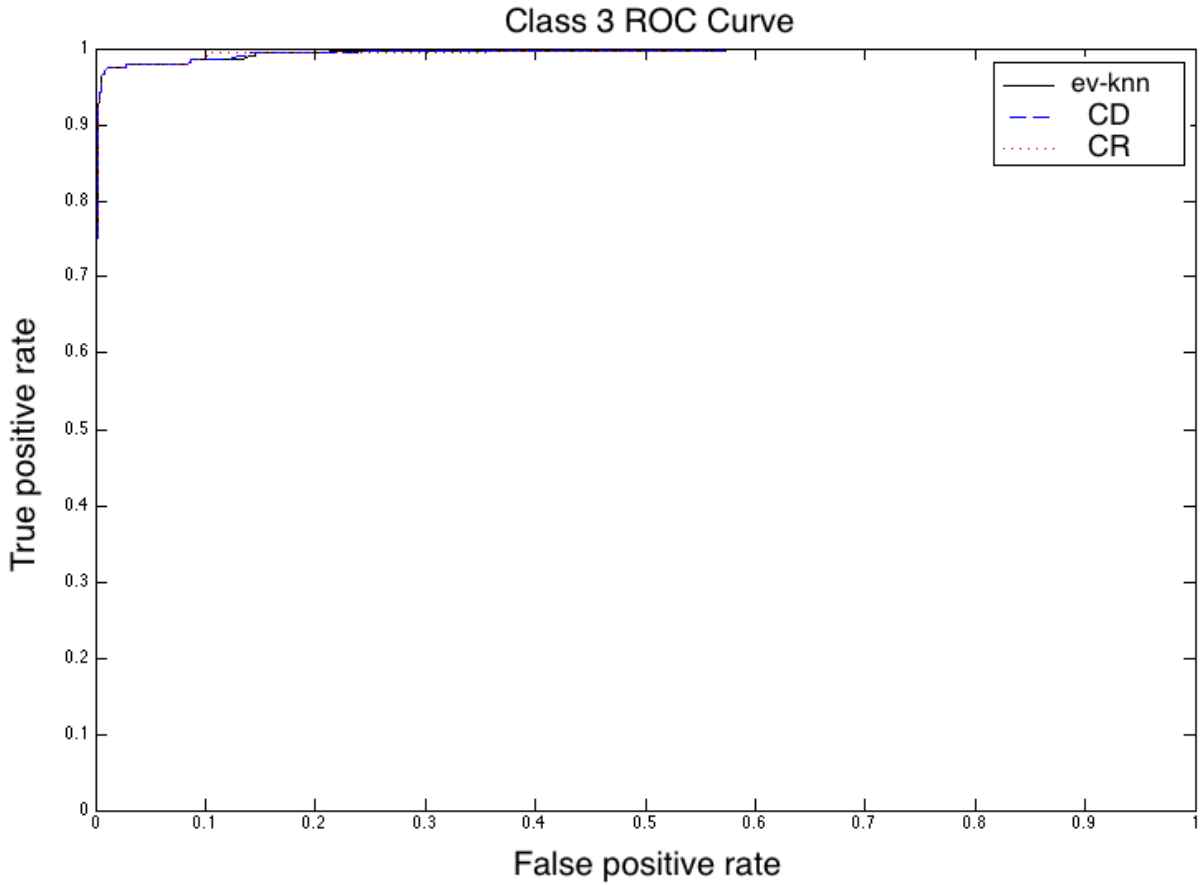
An experiment in classification

Results: Class 2 ROC Curve



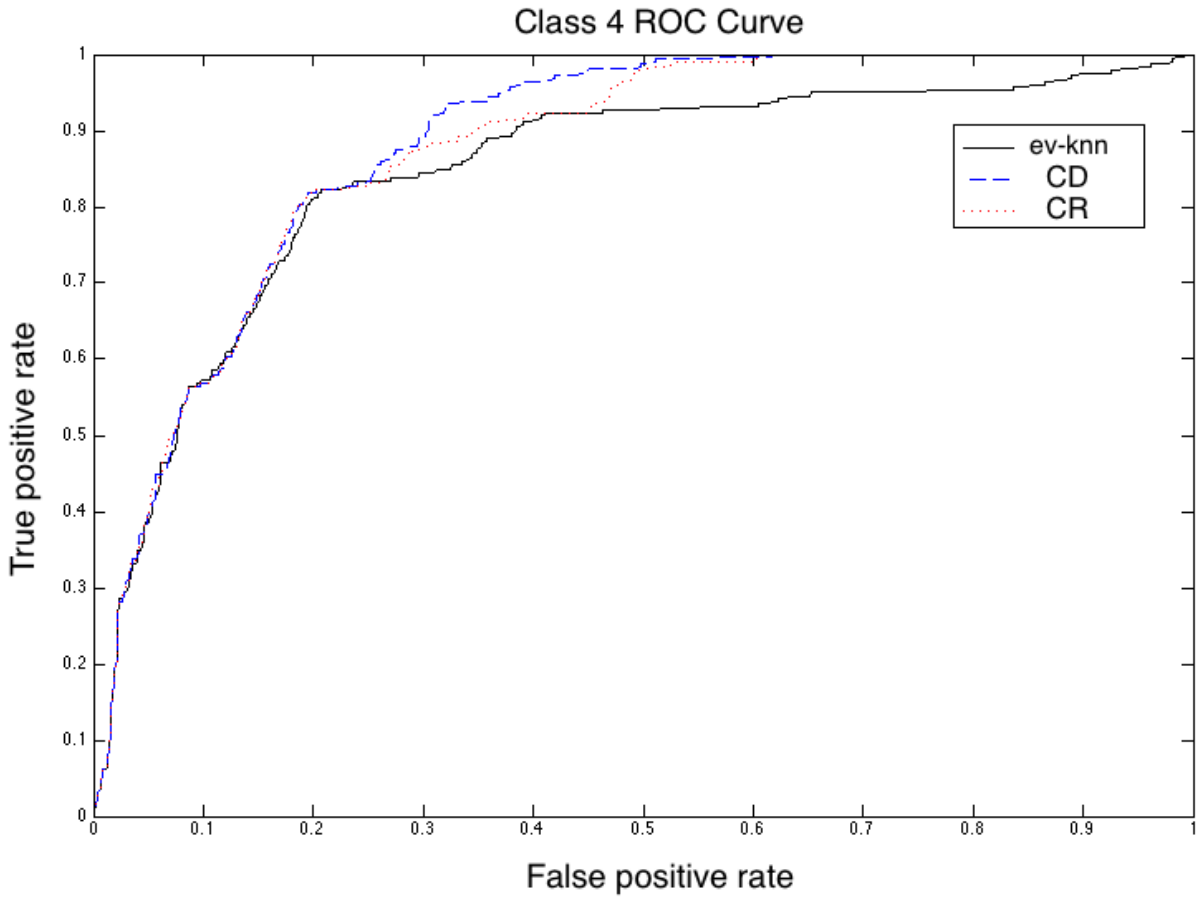
An experiment in classification

Results: Class 3 ROC Curve



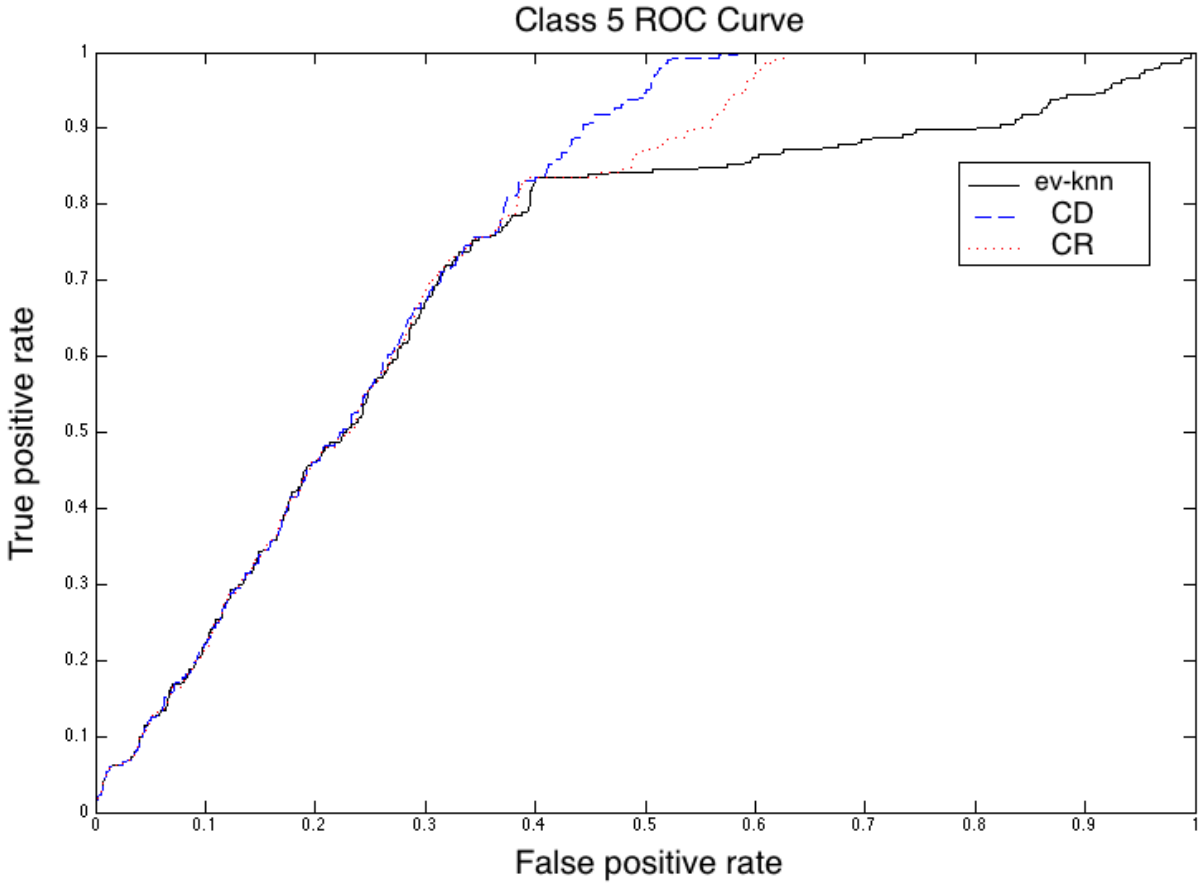
An experiment in classification

Results: Class 4 ROC Curve



An experiment in classification

Results: Class 5 ROC Curve



References 5

[10] F. Pichon et al. "Truthfulness in contextual information correction". In: *Proceedings of the 3rd International Conference on Belief Functions, BELIEF 2014*. Ed. by F. Cuzzolin. Vol. 8764. Lecture Notes in Computer Science. Oxford, United Kingdom: Springer International Publishing, Sept. 2014, pp. 11–20

[11] D. Mercier et al. "Learning Contextual Discounting and Contextual Reinforcement from Labelled Data". In: *Proceedings of the 13th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU 2015*. Ed. by S. Destercke and T. Denœux. Vol. 9161. Lecture Notes in Computer Science. Compiègne, France: Springer International Publishing, July 2015, pp. 472–481

Thank you for your attention.