Information correction and fusion using belief functions

Frédéric Pichon

Laboratory of Computer Engineering and Automation of Artois (LGI2A)
Université d'Artois, Béthune, France

BFAS School
Siena, Italy
October 28, 2019

Information correction and fusion ...

- Problem: to extract truthful and precise knowledge about a quantity of interest, from information coming from various sources.
- Applications: computer vision, robotics, machine learning...
- Old problem: origin of probability theory, where formalizing and merging partially reliable testimonies was a concern.
- Requires meta-knowledge on the sources, i.e., knowledge about their quality (typically, their reliability).
- Called information correction when there is a single information source and information fusion when there are several sources.

... using belief functions

- Related to the issue of uncertainty modeling.
- Uncertainty theories: probability, possibility, belief function, imprecise probability theories.
- Central role in belief function theory (BFT):
  1. [Shafer, 1976]: BFT as an approach for representing and merging partially reliable and elementary testimonies;
  2. Numerous theoretical contributions on information fusion;
  3. BFT used in applications for merging information.

... using belief functions

- Related to the issue of uncertainty modeling.
- Uncertainty theories: probability, possibility, belief function, imprecise probability theories.
- Central role in belief function theory (BFT):
  1. [Shafer, 1976]: BFT as an approach for representing and merging partially reliable and elementary testimonies;
  2. Numerous theoretical contributions on information fusion;
  3. BFT used in applications for merging information.

→ This lecture: some recent results in line with 1 – 3 , based on a modeling of source quality, reinforcing the relevance of BFT for information correction and fusion.

Contents of this lecture

- A general approach to information correction and fusion using belief functions
- A prism to understand some important belief function correction and fusion schemes
- An interpretation of belief functions ($\sim$ [Shafer, 1976] revisited)
- Means to tackle correction and fusion problems in practice

Contents of this lecture

- A general approach to information correction and fusion using belief functions
- A prism to understand some important belief function correction and fusion schemes
- An interpretation of belief functions ($\sim$ [Shafer, 1976] revisited)
- Means to tackle correction and fusion problems in practice

Not in this lecture:

- An exhaustive review of all combination rules
- A discussion on conflit measurement (see Anne-Laure's lecture)
- A discussion on rule properties (see, e.g., Sébastien's lecture at the 2015 BFAS school)
- Implementation aspects (see Arnaud's lecture)

Outline

Outline

Reliability

- Classically, to interpret information items provided by sources (sensor, human, ...), assumptions are made about their reliability (relevance), where a reliable source is a source providing useful information regarding the quantity of interest.
- Examples :
  - A broken watch is useless to try and find the time it is since there is no way to know whether the supplied information is correct or not: this source is not reliable for the time;
  - My six-year-old child is ignorant about the name of the latest Nobel Peace Prize laureate: he is not reliable for this question (in contrast to the source nobelprize.org).
- Basic idea : a piece of information received from a reliable source is considered valid, whereas it is useless if the source is not reliable.

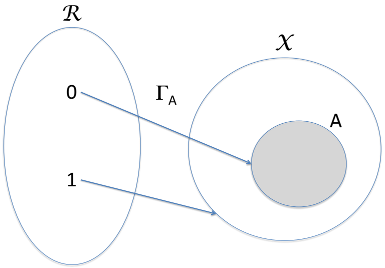
Formalization

- Let X be a variable of interest taking values in a finite set $\mathcal{X} = \{x_1, \ldots, x_n\}$ (frame of discernment), and whose actual value is unknown
- Assume a source $\mathfrak{s}$ telling that $X \in A \subseteq \mathcal{X}$
  - If $\mathfrak{s}$ is not reliable, we replace $X \in A$ by $X \in \mathcal{X}$
  - If $\mathfrak{s}$ is reliable, we keep $X \in A$

Formalization

- Let X be a variable of interest taking values in a finite set $\mathcal{X} = \{x_1, \ldots, x_n\}$ (frame of discernment), and whose actual value is unknown
- Assume a source $\mathfrak{s}$ telling that $X \in A \subseteq \mathcal{X}$
  - If $\mathfrak{s}$ is not reliable, we replace $X \in A$ by $X \in \mathcal{X}$
  - If $\mathfrak{s}$ is reliable, we keep $X \in A$
- Let R be the variable denoting its reliability, defined on $\mathcal{R} = \{0, 1\}$ where 0 means that $\mathfrak{s}$ is reliable and 1 means not reliable.
- The interpretation of the testimony according to the reliability may be encoded by $\Gamma_A : \mathcal{R} \rightarrow 2^{\mathcal{X}}$ such that

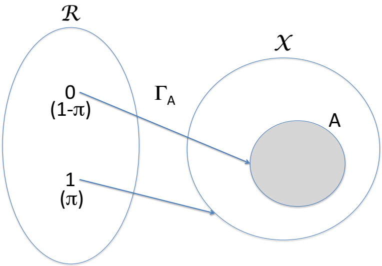
$$\Gamma_A(0) = A,$$
$$\Gamma_A(1) = \mathcal{X}.$$



Uncertain reliability

- Assume now $\mathfrak{s}$ is not reliable with probability $P^{\mathcal{R}}(R=1)=\pi$ (and reliable with probability $P^{\mathcal{R}}(R=0)=1-\pi$) with $\pi \in [0,1]$
- What can then be inferred about X?
- π should be transferred to $\Gamma_A(1)=\mathcal{X}$, $1-\pi$ to $\Gamma_A(0)=A$, and thus our knowledge about X is represented by a mass function (MF) on $\mathcal{X}$ such that

$$\begin{aligned} m(A) &= 1-\pi, \\ m(\mathcal{X}) &= \pi \end{aligned}$$

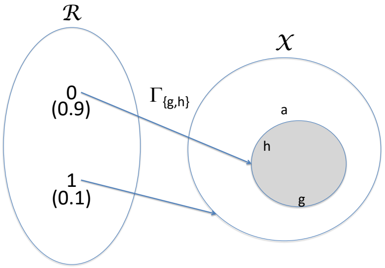


- $m(A)$: probability of knowing that $X \in A$ and nothing more, given the available evidence.
- m is a so-called simple mass function (SMF), since it has two focal sets including $\mathcal{X}$. It is more simply denoted by A^{π}.
- Other useful notation for m: $m[P^{\mathcal{R}}, A]$

Example

- Assume a sensor $\mathfrak{s}$ in charge of recognizing the type X of an aircraft which can be airplane (a), glider (g), or helicopter (h), i.e., $\mathcal{X} = \{a, g, h\}$.
- $\mathfrak{s}$ tells it is a glider or a helicopter, i.e., $X \in A = \{g, h\}$.
- The probability that the sensor is not reliable is 0.1, i.e., $\pi = 0.1$.
- Hence, our knowledge about X is represented by the SMF $\{g, h\}^{0.1}$

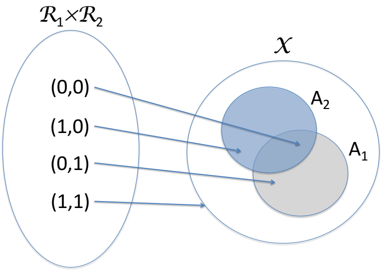
$$
\begin{aligned}
m(\{g, h\}) &= 0.9 \\
m(\mathcal{X}) &= 0.1
\end{aligned}
$$



Outline

Two information sources

- Assume now two sources $\mathfrak{s}_1$ and $\mathfrak{s}_2$ providing information $X \in A_1$ and $X \in A_2$, respectively.
- Let $\Gamma_{A_i} : \mathcal{R}_i \rightarrow 2^{\mathcal{X}}$ represent the interpretation of information A_i from $\mathfrak{s}_i$ given its reliability R_i defined on $\mathcal{R}_i = \{0, 1\}$.
- If they are in the state
  - $(R_1 = 0, R_2 = 0)$, then $X \in \Gamma_{A_1}(0) \cap \Gamma_{A_2}(0) = A_1 \cap A_2$
  - $(R_1 = 1, R_2 = 0)$, then $X \in \Gamma_{A_1}(1) \cap \Gamma_{A_2}(0) = \mathcal{X} \cap A_2 = A_2$
  - $(R_1 = 0, R_2 = 1)$, then $X \in \Gamma_{A_1}(0) \cap \Gamma_{A_2}(1) = A_1 \cap \mathcal{X} = A_1$
  - $(R_1 = 1, R_2 = 1)$, then $X \in \Gamma_{A_1}(1) \cap \Gamma_{A_2}(1) = \mathcal{X} \cap \mathcal{X} = \mathcal{X}$



Notation

- When the sources provide information $\mathbf{A} = (A_1, A_2)$ and are in the state $\mathbf{r} = (r_1, r_2) \in \boldsymbol{\mathcal{R}} := \mathcal{R}_1 \times \mathcal{R}_2$, we should deduce

$$X \in \Gamma_{\mathbf{A}}(\mathbf{r}) := \Gamma_{A_1}(r_1) \cap \Gamma_{A_2}(r_2)$$

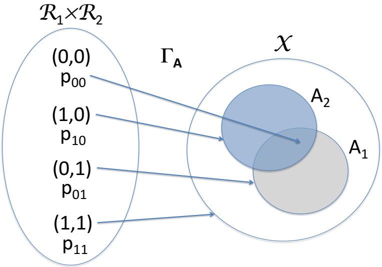
- $\Gamma_{\mathbf{A}} : \boldsymbol{\mathcal{R}} \to 2^{\mathcal{X}}$
- Example

$$\begin{aligned} \Gamma_{\mathbf{A}}(0,1) &= \Gamma_{A_1}(0) \cap \Gamma_{A_2}(1) \\ &= A_1 \cap \mathcal{X} \\ &= A_1 \end{aligned}$$

Uncertain reliabilities

- Assume now the sources are in state $\mathbf{r} = (r_1, r_2)$ with probability $P^{\mathcal{R}}(R_1 = r_1, R_2 = r_2) = p_{\mathbf{r}}$
- $p_{\mathbf{r}}$ should be transferred to $\Gamma_{\mathbf{A}}(\mathbf{r})$.
- Our knowledge about X can then be represented by

$$m(B) = \sum_{\mathbf{r}:\Gamma_{\mathbf{A}}(\mathbf{r})=B} p_{\mathbf{r}}.$$

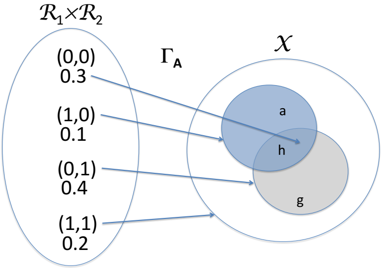


- Notation: $m[P^{\mathcal{R}}, \mathbf{A}]$

Example

- Two sensors $\mathfrak{s}_1$ and $\mathfrak{s}_2$ for the type X of an aircraft
- $\mathfrak{s}_1$ tells $X \in A_1 = \overline{\{a\}} = \{g, h\}$
- $\mathfrak{s}_2$ tells $X \in A_2 = \overline{\{g\}} = \{a, h\}$
- We have

$$\begin{aligned}
\Gamma_{\mathbf{A}}(0,0) &= A_1 \cap A_2 = \{h\} \\
\Gamma_{\mathbf{A}}(1,0) &= A_2 = \{a, h\} \\
\Gamma_{\mathbf{A}}(0,1) &= A_1 = \{g, h\} \\
\Gamma_{\mathbf{A}}(1,1) &= \mathcal{X}
\end{aligned}$$



- Induced knowledge about X:

$$m(\{h\}) = 0.3, m(\{a, h\}) = 0.1, m(\{g, h\}) = 0.4, m(\mathcal{X}) = 0.2$$

Decomposition of meta-knowledge

- $P^{\mathcal{R}}$ is a bivariate Bernoulli distribution
- It is characterized by

$$\begin{aligned}
\pi_i &:= \mathbb{E}[R_i] = P^{\mathcal{R}_i}(R_i = 1), \quad i = 1, 2, \\
\sigma &:= \mathbb{E}\left[(R_1 - \pi_1)(R_2 - \pi_2)\right] = \mathbb{E}\left[R_1 R_2\right] - \mathbb{E}[R_1]\mathbb{E}[R_2] \\
&= P^{\mathcal{R}}(R_1 = 1, R_2 = 1) - P^{\mathcal{R}_1}(R_1 = 1)P^{\mathcal{R}_2}(R_2 = 1)
\end{aligned}$$

- We have

$$\begin{aligned}
P^{\mathcal{R}}(R_1 = 0, R_2 = 0) &= \overline{\pi_1} \cdot \overline{\pi_2} + \sigma \\
P^{\mathcal{R}}(R_1 = 1, R_2 = 0) &= \pi_1 \cdot \overline{\pi_2} - \sigma \\
P^{\mathcal{R}}(R_1 = 0, R_2 = 1) &= \overline{\pi_1} \cdot \pi_2 - \sigma \\
P^{\mathcal{R}}(R_1 = 1, R_2 = 1) &= \pi_1 \cdot \pi_2 + \sigma
\end{aligned}$$

with $\overline{\pi_i} = 1 - \pi_i$

Example

- Knowledge on the reliabilities of the sensors $\mathfrak{s}_1$ and $\mathfrak{s}_2$:

$$\left.\begin{array}{l} P^{\mathcal{R}}(R_1 = 0, R_2 = 0) = 0.3 \\ P^{\mathcal{R}}(R_1 = 1, R_2 = 0) = 0.1 \\ P^{\mathcal{R}}(R_1 = 0, R_2 = 1) = 0.4 \\ P^{\mathcal{R}}(R_1 = 1, R_2 = 1) = 0.2 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{l} \pi_1 = 0.3 \\ \pi_2 = 0.6 \\ \sigma = 0.02 \end{array}\right.$$

Independent reliabilities

SMF-based expression

- R_1 and R_2 independent $\Leftrightarrow \sigma = 0$
- In this case

$$\begin{aligned} m[P^{\mathcal{R}}, \mathbf{A}] &= m[P^{\mathcal{R}_1}, A_1] \textcircled{\cap} m[P^{\mathcal{R}_2}, A_2] \\ &= A_1^{\pi_1} \textcircled{\cap} A_2^{\pi_2} \end{aligned}$$

- Reminder: unnormalized Dempster's rule (conjunctive rule)

$$(m_1 \textcircled{\cap} m_2)(A) = \sum_{B \cap C = A} m_1(B) m_2(C), \quad \forall A \subseteq \mathcal{X}.$$

Dependent reliabilities

SMF-based expression

- More generally, i.e., for any dependency σ, $m[P^{\mathcal{R}}, \mathbf{A}]$ can always be expressed as a conjunctive combination of $A_1^{\pi_1}$ and $A_2^{\pi_2}$ having some dependency...
- "Reminder": conjunctive combination $m_\cap$ of m_1 and m_2 having some known dependency
  1. A joint MF $jm : 2^{\mathcal{X}} \times 2^{\mathcal{X}} \to [0,1]$ is built, having m_1 and m_2 as marginals and encoding their mutual dependence
  2. Each joint mass $jm(B, C)$ is allocated to $B \cap C$:

$$m_\cap(A) = \sum_{B \cap C = A} jm(B, C)$$

Dependent reliabilities

SMF-based expression

- Let $m_i = A_i^{\pi_i}$. Any *jm* having $A_1^{\pi_1}$ and $A_2^{\pi_2}$ as marginals can always be written as

$$\begin{aligned} jm(A_1, A_2) &= \overline{\pi_1} \cdot \overline{\pi_2} + \sigma \\ jm(\mathcal{X}, A_2) &= \pi_1 \cdot \overline{\pi_2} - \sigma \\ jm(A_1, \mathcal{X}) &= \overline{\pi_1} \cdot \pi_2 - \sigma \\ jm(\mathcal{X}, \mathcal{X}) &= \pi_1 \cdot \pi_2 + \sigma \end{aligned}$$

  for some σ.

- Conjunctive combination of $A_1^{\pi_1}$ and $A_2^{\pi_2}$ with dependence structure represented by *jm*, is completely determined by σ.

→ Parameterized conjunctive rule $\circledcap_\sigma$ for two SMF, with parameter σ representing the dependence structure, such that

$$\circledcap_\sigma(A_1^{\pi_1}, A_2^{\pi_2}) := m_\cap$$

- For $\sigma = 0$, $\circledcap_\sigma \Leftrightarrow \circledcap$

Dependent reliabilities

SMF-based expression

- For any dependency σ between the source reliabilities, we have

$$\begin{aligned} m[P^{\mathcal{R}}, \mathbf{A}] &= \circledcap_{\sigma}(m[P^{\mathcal{R}_1}, A_1], m[P^{\mathcal{R}_2}, A_2]) \\ &= \circledcap_{\sigma}(A_1^{\pi_1}, A_2^{\pi_2}) \end{aligned}$$

- Example:
  - Sensor $\mathfrak{s}_1$ not reliable with probability $\pi_1 = 0.3$
  - Sensor $\mathfrak{s}_2$ not reliable with probability $\pi_2 = 0.6$
  - Dependence between their reliability: $\sigma = 0.02$
  - Induced knowledge on $\mathcal{X}$ from the information $\mathbf{A} = (\{g, h\}, \{a, h\})$ provided by the sensors satisfies

$$m[P^{\mathcal{R}}, \mathbf{A}] = \circledcap_{(0.02)} \left(\{g, h\}^{0.3}, \{a, h\}^{0.6}\right)$$

Cautious rule for SMF

- Let $A_1^{\pi_1}$ and $A_2^{\pi_2}$ be two non-independent SMF.
- How to combine them ?
- Cautious conjunctive combination $\owedge$: select the least committed (according to some informational ordering) MF among those that are at least as committed as $A_1^{\pi_1}$ and $A_2^{\pi_2}$.
- Solution based on the w-ordering yields

$$A_1^{\pi_1} \owedge A_2^{\pi_2} = \begin{cases} A_1^{\pi_1 \wedge \pi_2} & \text{if } A_1 = A_2 \\ A_1^{\pi_1} \textcircled{\scriptsize$\cap$} A_2^{\pi_2} & \text{if } A_1 \neq A_2 \end{cases}$$

Cautious rule revisited

- We have

$$A_1^{\pi_1} \owedge A_2^{\pi_2} = \Cap_\sigma(A_1^{\pi_1}, A_2^{\pi_2})$$

  with

$$\sigma = \begin{cases} \pi_1 \wedge \pi_2 - \pi_1\pi_2 & \text{if } A_1 = A_2 \\ 0 & \text{if } A_1 \neq A_2 \end{cases}$$

- Partially reliable sources analysis:
  - $\mathfrak{s}_i$ not reliable with probability π_i and telling $X \in A_i$
  - $\mathfrak{s}_1$ and $\mathfrak{s}_2$ have dependent reliabilities if they support the same subset (actually, perfect dependence between them not being reliable) and independent reliabilities otherwise.

Outline

K information sources

- Assume sources $\mathfrak{s}_i$, $i = 1, \ldots, K$, providing $\mathbf{A} = (A_1, \ldots, A_K)$.
- When the sources are in the state $\mathbf{r} = (r_1, \ldots, r_K) \in \boldsymbol{\mathcal{R}} := \times_{i=1}^{K} \mathcal{R}_i$, we should deduce

$$X \in \Gamma_{\mathbf{A}}(\mathbf{r}) := \bigcap_{i=1}^{K} \Gamma_{A_i}(r_i)$$

- Example: $K = 3$

$$\begin{aligned} \Gamma_{\mathbf{A}}(0,0,1) &= \Gamma_{A_1}(0) \cap \Gamma_{A_2}(0) \cap \Gamma_{A_3}(1) \\ &= A_1 \cap A_2 \cap \mathcal{X} \\ &= A_1 \cap A_2 \end{aligned}$$

Uncertain reliabilities

- Assume state $\mathbf{r}$ is allocated probability $p_{\mathbf{r}}$:

$$P^{\mathcal{R}}(R_1 = r_1, \ldots, R_K = r_K) = p_{\mathbf{r}}$$

- Knowledge about X is then represented by

$$m[P^{\mathcal{R}}, \mathbf{A}](B) = \sum_{\mathbf{r}:\Gamma_{\mathbf{A}}(\mathbf{r})=B} p_{\mathbf{r}}$$

$\rightarrow$ Any set of partially reliable and elementary testimonies is represented by a (unique) MF

Separable mass function

- Independence of all the R_i

$$\begin{aligned} m[P^{\mathcal{R}}, \mathbf{A}] &= \textcircled{\cap}_{i=1}^{K} m[P^{\mathcal{R}_i}, A_i] \\ &= \textcircled{\cap}_{i=1}^{K} A_i^{\pi_i} \end{aligned}$$

with $\pi_i := P^{\mathcal{R}_i}(R_i = 1)$.

$\rightarrow$ Choosing $\mathbf{A}$ s.t. $A_i \neq A_j$, $1 \leq i < j \leq K$, we obtain a partially reliable sources analysis of separable MF[1], which form an important class of MF (often encountered in practice)

[1] MF that can be written as a conjunctive combination of independent SMF supporting different subsets

Separable mass function

- Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $A^{\mathbf{r}}$ denote the subset of $\mathcal{X}$ such that $x_i \in A^{\mathbf{r}}$ if $r_i = 1$ and $x_i \notin A^{\mathbf{r}}$ if $r_i = 0$, for $\mathbf{r} = (r_1, \ldots, r_K)$.
- Example: $\mathcal{X} = \{x_1, x_2, x_3\}$ then $A^{001} = \{x_3\}$

Separable mass function

- Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $A^{\mathbf{r}}$ denote the subset of $\mathcal{X}$ such that $x_i \in A^{\mathbf{r}}$ if $r_i = 1$ and $x_i \notin A^{\mathbf{r}}$ if $r_i = 0$, for $\mathbf{r} = (r_1, \ldots, r_K)$.
- Example: $\mathcal{X} = \{x_1, x_2, x_3\}$ then $A^{001} = \{x_3\}$

Proposition

Let m be a MF on $\mathcal{X}$, $K = |\mathcal{X}|$, $A_i = \overline{\{x_i\}}$, and $p_{\mathbf{r}} = m(A^{\mathbf{r}})$. We have

$$m[P^{\mathcal{R}}, \mathbf{A}] = m$$

Proof: Follows from $\Gamma_{\mathbf{A}}(\mathbf{r}) = A^{\mathbf{r}}$.

Separable mass function

- Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $A^{\mathbf{r}}$ denote the subset of $\mathcal{X}$ such that $x_i \in A^{\mathbf{r}}$ if $r_i = 1$ and $x_i \notin A^{\mathbf{r}}$ if $r_i = 0$, for $\mathbf{r} = (r_1, \ldots, r_K)$.
- Example: $\mathcal{X} = \{x_1, x_2, x_3\}$ then $A^{001} = \{x_3\}$

Proposition

Let m be a MF on $\mathcal{X}$, $K = |\mathcal{X}|$, $A_i = \overline{\{x_i\}}$, and $p_{\mathbf{r}} = m(A^{\mathbf{r}})$. We have

$$m[P^{\mathcal{R}}, \mathbf{A}] = m$$

Proof: Follows from $\Gamma_{\mathbf{A}}(\mathbf{r}) = A^{\mathbf{r}}$.

- Recall: Any set of partially reliable and elementary testimonies is represented by a (unique) MF.

Separable mass function

- Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $A^{\mathbf{r}}$ denote the subset of $\mathcal{X}$ such that $x_i \in A^{\mathbf{r}}$ if $r_i = 1$ and $x_i \notin A^{\mathbf{r}}$ if $r_i = 0$, for $\mathbf{r} = (r_1, \ldots, r_K)$.
- Example: $\mathcal{X} = \{x_1, x_2, x_3\}$ then $A^{001} = \{x_3\}$

Proposition

Let m be a MF on $\mathcal{X}$, $K = |\mathcal{X}|$, $A_i = \overline{\{x_i\}}$, and $p_{\mathbf{r}} = m(A^{\mathbf{r}})$. We have

$$m[P^{\mathcal{R}}, \mathbf{A}] = m$$

Proof: Follows from $\Gamma_{\mathbf{A}}(\mathbf{r}) = A^{\mathbf{r}}$.

- Recall: Any set of partially reliable and elementary testimonies is represented by a (unique) MF.

$\rightarrow$ Any mass function represents (at least) a set of partially reliable and elementary testimonies

Example

- Let m be the FM on $\mathcal{X} = \{a, g, h\}$ defined by

$$m(\{a, g\}) = 0.5, m(\{h\}) = 0.2, m(\{g, h\}) = 0.3$$

- Consider $K = |\mathcal{X}| = 3$ sources $\mathfrak{s}_1$, $\mathfrak{s}_2$ and $\mathfrak{s}_3$, providing respectively information

$$\begin{aligned} X \in A_1 &= \overline{\{a\}} = \{g, h\} \\ X \in A_2 &= \overline{\{g\}} = \{a, h\} \\ X \in A_3 &= \overline{\{h\}} = \{a, g\} \end{aligned}$$

Example

- Consider meta-knowledge $P^{\mathcal{R}}$ on the sources such that

$\mathbf{r}$	$A^{\mathbf{r}}$	$m(A^{\mathbf{r}}) = p_{\mathbf{r}}$
$(0,0,0)$	$\emptyset$	0
$(1,0,0)$	$\{a\}$	0
$(0,1,0)$	$\{g\}$	0
$(1,1,0)$	$\{a,g\}$	0.5
$(0,0,1)$	$\{h\}$	0.2
$(1,0,1)$	$\{a,h\}$	0
$(0,1,1)$	$\{g,h\}$	0.3
$(1,1,1)$	$\{a,g,h\}$	0

- Then, e.g., $p_{001} = P^{\mathcal{R}}(R_1 = 0, R_2 = 0, R_3 = 1) = m(A^{001}) = 0.2$ is allocated to $\Gamma_{\mathbf{A}}(0,0,1) = A_1 \cap A_2 \cap \mathcal{X} = \overline{\{a\}} \cap \overline{\{g\}} = \{h\} = A^{001}$.
- Since this happens for all $\mathbf{r}$, we have $m[P^{\mathcal{R}}, \mathbf{A}] = m$.

Decomposition of meta-knowledge

- $P^{\mathcal{R}}$ is a multivariate Bernoulli distribution
- It is characterized by

$$\pi_i = \mathbb{E}[R_i]$$

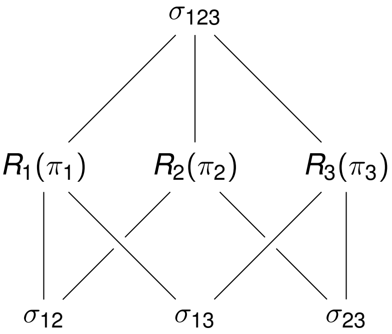
$$\sigma_{\mathbf{r}} = \mathbb{E}\left[\prod_{i=1}^{K}(R_i - \pi_i)^{r_i}\right]$$

with $\mathbf{r} = (r_1, \ldots, r_K)$ such that $\sum_{i=1}^{K} r_i > 1$

- There are $2^K - K - 1$ central moments $\sigma_{\mathbf{r}}$. They represent the dependencies between any subset (of at least two) of all the R_i.
- Notation: $\boldsymbol{\sigma}$ vector whose elements are the dependencies $\sigma_{\mathbf{r}}$

Example

$K = 3$



where

$$\sigma_{12} := \sigma_{110} = \mathbb{E}\left[(R_1 - \pi_1)(R_2 - \pi_2)\right]$$

and $\sigma_{13} := \sigma_{101}$, $\sigma_{23} := \sigma_{011}$, $\sigma_{123} := \sigma_{111}$

$$\boldsymbol{\sigma} = (\sigma_{12}, \sigma_{13}, \sigma_{23}, \sigma_{123})$$

Example

- Knowledge on the reliabilities of the sources $\mathfrak{s}_1$, $\mathfrak{s}_2$ and $\mathfrak{s}_3$:

$$\left.\begin{array}{l} P^{\mathcal{R}}(R_1 = 0, R_2 = 0, R_3 = 0) = 0 \\ P^{\mathcal{R}}(R_1 = 1, R_2 = 0, R_3 = 0) = 0 \\ P^{\mathcal{R}}(R_1 = 0, R_2 = 1, R_3 = 0) = 0 \\ P^{\mathcal{R}}(R_1 = 1, R_2 = 1, R_3 = 0) = 0.5 \\ P^{\mathcal{R}}(R_1 = 0, R_2 = 0, R_3 = 1) = 0.2 \\ P^{\mathcal{R}}(R_1 = 1, R_2 = 0, R_3 = 1) = 0 \\ P^{\mathcal{R}}(R_1 = 0, R_2 = 1, R_3 = 1) = 0.3 \\ P^{\mathcal{R}}(R_1 = 1, R_2 = 1, R_3 = 1) = 0 \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{r} \pi_1 = 0.5 \\ \pi_2 = 0.8 \\ \pi_3 = 0.5 \\ \sigma_{12} = 0.1 \\ \sigma_{13} = -0.25 \\ \sigma_{23} = -0.1 \\ \sigma_{123} = 0 \end{array}\right.$$

Example

- $P^{\mathcal{R}}$ is recovered from π_i and $\sigma_{\mathbf{r}}$ as follows:

$$
\begin{aligned}
P^{\mathcal{R}}(R_1=0,R_2=0,R_3=0) &= \overline{\pi_1}\,\overline{\pi_2}\,\overline{\pi_3}+\overline{\pi_3}\sigma_{12}+\overline{\pi_2}\sigma_{13}+\overline{\pi_1}\sigma_{23}-\sigma_{123}\\
P^{\mathcal{R}}(R_1=1,R_2=0,R_3=0) &= \pi_1\overline{\pi_2}\,\overline{\pi_3}-\overline{\pi_3}\sigma_{12}-\overline{\pi_2}\sigma_{13}+\pi_1\sigma_{23}+\sigma_{123}\\
P^{\mathcal{R}}(R_1=0,R_2=1,R_3=0) &= \overline{\pi_1}\pi_2\overline{\pi_3}-\overline{\pi_3}\sigma_{12}+\pi_2\sigma_{13}-\overline{\pi_1}\sigma_{23}+\sigma_{123}\\
P^{\mathcal{R}}(R_1=1,R_2=1,R_3=0)) &= \pi_1\pi_2\overline{\pi_3}+\overline{\pi_3}\sigma_{12}-\pi_2\sigma_{13}-\pi_1\sigma_{23}-\sigma_{123}\\
P^{\mathcal{R}}(R_1=0,R_2=0,R_3=1) &= \overline{\pi_1}\,\overline{\pi_2}\pi_3+\pi_3\sigma_{12}-\overline{\pi_2}\sigma_{13}-\overline{\pi_1}\sigma_{23}+\sigma_{123}\\
P^{\mathcal{R}}(R_1=1,R_2=0,R_3=1) &= \pi_1\overline{\pi_2}\pi_3-\pi_3\sigma_{12}+\overline{\pi_2}\sigma_{13}-\pi_1\sigma_{23}-\sigma_{123}\\
P^{\mathcal{R}}(R_1=0,R_2=1,R_3=1) &= \overline{\pi_1}\pi_2\pi_3-\pi_3\sigma_{12}-\pi_2\sigma_{13}+\overline{\pi_1}\sigma_{23}-\sigma_{123}\\
P^{\mathcal{R}}(R_1=1,R_2=1,R_3=1) &= \pi_1\pi_2\pi_3+\pi_3\sigma_{12}+\pi_2\sigma_{13}+\pi_1\sigma_{23}+\sigma_{123}
\end{aligned}
$$

- Remark: simple matrix-based expressions exist to switch from one representation to the other.

Decomposition of a mass function

Based on the previous proposition as well as the decomposition of meta-knowledge, any MF m on $\mathcal{X}$ is induced by the following basic components:

1. A set of $|\mathcal{X}|$ sources $\mathfrak{S} = \{\mathfrak{s}_1, \ldots, \mathfrak{s}_{|\mathcal{X}|}\}$, with $\mathfrak{s}_i$ providing information $X \in \overline{\{x_i\}}$;
2. Probabilistic knowledge on their reliability characterized by :
   - For each $\mathfrak{s}_i$, a (marginal) probability π_i of being not reliable;
   - For each $S_\mathbf{r} \subseteq \mathfrak{S}$, knowledge about the dependency between their reliabilities in the form of the central moment $\sigma_\mathbf{r}$.

Remark: we have $\pi_i = pl(\{x_i\})$.

Example

- m defined by $m(\{a,g\}) = 0.5, m(\{h\}) = 0.2, m(\{g,h\}) = 0.3$.
- This MF is induced by
  1. Considering $|\mathcal{X}| = 3$ sources $\mathfrak{s}_1$, $\mathfrak{s}_2$ and $\mathfrak{s}_3$, providing information

$$\begin{aligned} X \in A_1 &= \overline{\{a\}} \\ X \in A_2 &= \overline{\{g\}} \\ X \in A_3 &= \overline{\{h\}} \end{aligned}$$

  2. And by assuming that
     - $\mathfrak{s}_1$, $\mathfrak{s}_2$ and $\mathfrak{s}_3$ are not reliable with respective probabilities $\pi_1 = 0.5$, $\pi_2 = 0.8$ and $\pi_3 = 0.5$
     - there is a dependence $\sigma_{12} = 0.1$ between the reliabilities of $\mathfrak{s}_1$ and $\mathfrak{s}_2$, $\sigma_{13} = -0.25$ between $\mathfrak{s}_1$ and $\mathfrak{s}_3$, $\sigma_{23} = -0.1$ between $\mathfrak{s}_2$ and $\mathfrak{s}_3$, and $\sigma_{123} = 0$ between $\mathfrak{s}_1$, $\mathfrak{s}_2$ and $\mathfrak{s}_3$.

SMF-based expression of $m[P^{\mathcal{R}}, \mathbf{A}]$

- Let $m_\cap$ denote the conjunctive combination of SMF $A_i^{\pi_i}$, $1 \leq i \leq K$, with dependence structure represented by *jm*
- *jm* can always be written in a particular (familiar) form such that the dependencies it encodes are completely determined by $2^K - K - 1$ parameters

SMF-based expression of $m[P^{\mathcal{R}}, \mathbf{A}]$

- Let $m_{\cap}$ denote the conjunctive combination of SMF $A_i^{\pi_i}$, $1 \leq i \leq K$, with dependence structure represented by jm
- jm can always be written in a particular (familiar) form such that the dependencies it encodes are completely determined by $2^K - K - 1$ parameters
- Example: $K = 3$, four parameters noted $\sigma_{12}, \sigma_{13}, \sigma_{23}, \sigma_{123}$

$$
\begin{aligned}
jm(A_1, A_2, A_3) &= \overline{\pi_1}\,\overline{\pi_2}\,\overline{\pi_3} + \overline{\pi_3}\sigma_{12} + \overline{\pi_2}\sigma_{13} + \overline{\pi_1}\sigma_{23} - \sigma_{123}\\
jm(\mathcal{X}, A_2, A_3) &= \pi_1\overline{\pi_2}\,\overline{\pi_3} - \overline{\pi_3}\sigma_{12} - \overline{\pi_2}\sigma_{13} + \pi_1\sigma_{23} + \sigma_{123}\\
jm(A_1, A_2, A_3) &= \overline{\pi_1}\pi_2\overline{\pi_3} - \overline{\pi_3}\sigma_{12} + \pi_2\sigma_{13} - \overline{\pi_1}\sigma_{23} + \sigma_{123}\\
jm(\mathcal{X}, \mathcal{X}, A_3) &= \pi_1\pi_2\overline{\pi_3} + \overline{\pi_3}\sigma_{12} - \pi_2\sigma_{13} - \pi_1\sigma_{23} - \sigma_{123}\\
jm(A_1, A_2, \mathcal{X}) &= \overline{\pi_1}\,\overline{\pi_2}\pi_3 + \pi_3\sigma_{12} - \overline{\pi_2}\sigma_{13} - \overline{\pi_1}\sigma_{23} + \sigma_{123}\\
jm(\mathcal{X}, A_2, \mathcal{X}) &= \pi_1\overline{\pi_2}\pi_3 - \pi_3\sigma_{12} + \overline{\pi_2}\sigma_{13} - \pi_1\sigma_{23} - \sigma_{123}\\
jm(A_1, \mathcal{X}, \mathcal{X}) &= \overline{\pi_1}\pi_2\pi_3 - \pi_3\sigma_{12} - \pi_2\sigma_{13} + \overline{\pi_1}\sigma_{23} - \sigma_{123}\\
jm(\mathcal{X}, \mathcal{X}, \mathcal{X}) &= \pi_1\pi_2\pi_3 + \pi_3\sigma_{12} + \pi_2\sigma_{13} + \pi_1\sigma_{23} + \sigma_{123}
\end{aligned}
$$

SMF-based expression of $m[P^{\mathcal{R}}, \mathbf{A}]$

- Parameterized conjunctive rule $\circledcap_{\boldsymbol{\sigma}}$ for K SMF, with $\boldsymbol{\sigma}$ the vector whose elements are the dependency parameters $\sigma_{\mathbf{r}}$, such that

$$\circledcap_{\boldsymbol{\sigma}}(A_1^{\pi_1}, \ldots, A_K^{\pi_K}) := m_{\cap}$$

SMF-based expression of $m[P^{\mathcal{R}}, \mathbf{A}]$

- Parameterized conjunctive rule $\bigcirc\!\!\!\!\cap_{\boldsymbol{\sigma}}$ for K SMF, with $\boldsymbol{\sigma}$ the vector whose elements are the dependency parameters $\sigma_{\mathbf{r}}$, such that

$$\bigcirc\!\!\!\!\cap_{\boldsymbol{\sigma}}(A_1^{\pi_1}, \ldots, A_K^{\pi_K}) := m_\cap$$

Theorem

For any dependency $\boldsymbol{\sigma}$ between the source reliabilities, we have

$$\begin{aligned} m[P^{\mathcal{R}}, \mathbf{A}] &= \bigcirc\!\!\!\!\cap_{\boldsymbol{\sigma}}(m[P^{\mathcal{R}_1}, A_1], \ldots, m[P^{\mathcal{R}_K}, A_K]) \\ &= \bigcirc\!\!\!\!\cap_{\boldsymbol{\sigma}}(A_1^{\pi_1}, \ldots, A_K^{\pi_K}) \end{aligned}$$

Conjunctive canonical decomposition of a MF

Theorem

Any MF m satisfies

$$m = \circledcap_{\boldsymbol{\sigma}} \left(\overline{\{x_1\}}^{pl(x_1)}, \ldots, \overline{\{x_n\}}^{pl(x_n)} \right),$$

with $\boldsymbol{\sigma}$ the vector of dependencies between the reliabilities of the sources underlying m.

Conjunctive canonical decomposition of a MF

Theorem

Any MF m satisfies

$$m = \circledcap_{\boldsymbol{\sigma}} \left(\overline{\{x_1\}}^{pl(x_1)}, \ldots, \overline{\{x_n\}}^{pl(x_n)} \right),$$

with $\boldsymbol{\sigma}$ the vector of dependencies between the reliabilities of the sources underlying m.

- Example: m defined by

$$m(\{a,g\}) = 0.5, m(\{h\}) = 0.2, m(\{g,h\}) = 0.3$$

satisfies

$$m = \circledcap_{(0.1,-0.25,-0.1,0)} \left(\overline{\{a\}}^{0.5}, \overline{\{g\}}^{0.8}, \overline{\{h\}}^{0.5} \right)$$

Conclusions

- Any mass function can be seen as the result of the
  - interpretation of a set of partially reliable and elementary testimonies
  - conjunctive combination of SMF (focused on $\overline{\{x_i\}}$) having some dependencies.
- In the spirit of [Shafer, 1976] and [Smets, 1995] interpretations of belief functions (they considered only independent SMF to try and recover the entire space of mass functions, see Didier's lecture)
- Combining conjunctively SMF focused on $\overline{\{x_i\}}$ is found in some important results: generalized Bayesian theorem, BFT analysis of binary logistic regression [Denoeux, 2019].

Outline

The case of a single source

- Assume now the source $\mathfrak{s}$ provides an uncertain testimony about X in the form of a MF $m_{\mathfrak{s}}$.
- If $\mathfrak{s}$ is reliable, then each $m_{\mathfrak{s}}(A)$ should be transferred to $\Gamma_A(0) = A$.
- If $\mathfrak{s}$ is not reliable, then each $m_{\mathfrak{s}}(A)$ should be transferred to $\Gamma_A(1) = \mathcal{X}$.
- Assuming $\mathfrak{s}$ to be not reliable with probability $P^{\mathcal{R}}(R = 1) = \pi$ then yields the following MF on $\mathcal{X}$:

$$m[P^{\mathcal{R}}, m_{\mathfrak{s}}] = (1 - \pi) \cdot m_{\mathfrak{s}} + \pi \cdot m_{\mathcal{X}}$$

  with $m_{\mathcal{X}}$ the vacuous MF ($m_{\mathcal{X}}(\mathcal{X}) = 1$).
- This is known as the discounting of $m_{\mathfrak{s}}$ with discount rate π (basic information correction mechanism in BFT).

Example

- Uncertain testimony:

$$m_{\mathfrak{s}}(\{a,g\}) = 0.8, m_{\mathfrak{s}}(\{h\}) = 0.2$$

- Uncertain reliability:

$$P^{\mathcal{R}}(R=0) = 0.7, P^{\mathcal{R}}(R=1) = 0.3$$

Example

- Uncertain testimony:

$$m_{\mathfrak{s}}(\{a,g\}) = 0.8, m_{\mathfrak{s}}(\{h\}) = 0.2$$

- Uncertain reliability:

$$P^{\mathcal{R}}(R=0) = 0.7, P^{\mathcal{R}}(R=1) = 0.3$$

$m_{\mathfrak{s}} \setminus P^{\mathcal{R}}$	$R=0$ 0.7	$R=1$ 0.3
$\{a,g\}$ 0.8	$\{a,g\}$ 0.56	$\mathcal{X}$ 0.24
$\{h\}$ 0.2	$\{h\}$ 0.14	$\mathcal{X}$ 0.06

Example

- Uncertain testimony:

$$m_{\mathfrak{s}}(\{a,g\}) = 0.8, m_{\mathfrak{s}}(\{h\}) = 0.2$$

- Uncertain reliability:

$$P^{\mathcal{R}}(R=0) = 0.7, P^{\mathcal{R}}(R=1) = 0.3$$

$m_{\mathfrak{s}} \setminus P^{\mathcal{R}}$	$R=0$ 0.7	$R=1$ 0.3
$\{a,g\}$ 0.8	$\{a,g\}$ 0.56	$\mathcal{X}$ 0.24
$\{h\}$ 0.2	$\{h\}$ 0.14	$\mathcal{X}$ 0.06

$\rightarrow$ $m[P^{\mathcal{R}}, m_{\mathfrak{s}}](\{a,g\}) = 0.56$, $m[P^{\mathcal{R}}, m_{\mathfrak{s}}](\{h\}) = 0.14$, $m[P^{\mathcal{R}}, m_{\mathfrak{s}}](\mathcal{X}) = 0.30$

The case of multiple sources

- Assume sources $\mathfrak{s}_1, \ldots, \mathfrak{s}_K$ provide uncertain testimonies $\mathbf{m} = (m_1, \ldots, m_K)$.
- Assume they are independent: interpreting $m_i(A_i)$ as the probability that $\mathfrak{s}_i$ supplies $X \in A_i$, then the probability, denoted $m(\mathbf{A})$, that they supply $\mathbf{A} = (A_1, \ldots, A_K)$ is $\prod_{i=1}^{K} m_i(A_i)$.
- If they are in state $\mathbf{r} \in \mathcal{R}$, then $m(\mathbf{A})$ should be transferred to $\Gamma_{\mathbf{A}}(\mathbf{r})$
- Assuming they are in state $\mathbf{r}$ with probability $p_{\mathbf{r}}$ then yields the following MF on $\mathcal{X}$:

$$m[P^{\mathcal{R}}, \mathbf{m}](B) = \sum_{\mathbf{r}, \mathbf{A} : \Gamma_{\mathbf{A}}(\mathbf{r}) = B} p_{\mathbf{r}} \cdot m(\mathbf{A})$$

Particular cases

$m[P^{\mathcal{R}}, \mathbf{m}]$ reduces to

- $\textcircled{\cap}_{i=1}^{K} m_i$ if $p_{\mathbf{r}} = 1$ for $\mathbf{r} = (0, 0, \ldots, 0, 0)$, i.e., all sources are reliable

$\rightarrow$ conjunctive rule

- $\textcircled{\cap}_{i=1}^{K} m[P^{\mathcal{R}_i}, m_i]$ if $p_{\mathbf{r}} = \prod_{i=1}^{K} P^{\mathcal{R}_i}(R_i = r_i)$, i.e., the sources have independent reliabilities

$\rightarrow$ discount and combine approach

- $\sum_{i=1}^{K} w_i \cdot m_i$ if $p_{\mathbf{r}} = w_i$ for $\mathbf{r}$ such that $r_i = 1$ and $r_j = 0$, for all $j \neq i$, i.e., source $\mathfrak{s}_i$ is reliable and the other are not with probability w_i

$\rightarrow$ weighted average

Outline

Truthfulness

- Reliability of a source includes another dimension besides its relevance: its truthfulness.
- A source is said truthful if it actually supplies the information it possesses.
- Lack of truthfulness can take several forms, and can be intentional or accidental.
- For instance, a sensor that has a systematic bias (such as a watch that has not been calibrated to winter time), is a kind of (unintentional) lack of truthfulness.
- Let us first consider the crudest form, where a source is non truthful if it tells the contrary of what it knows.

Outline

Formalization

- Assume a source $\mathfrak{s}$ supplying information item $X \in A$
  - If $\mathfrak{s}$ is not relevant, we replace $X \in A$ by $X \in \mathcal{X}$
  - If $\mathfrak{s}$ is relevant,
    - either it is truthful, in which case we keep $X \in A$
    - or it lies, in which case we replace $X \in A$ by $X \in \overline{A}$

Formalization

- Assume a source $\mathfrak{s}$ supplying information item $X \in A$
  - If $\mathfrak{s}$ is not relevant, we replace $X \in A$ by $X \in \mathcal{X}$
  - If $\mathfrak{s}$ is relevant,
    - either it is truthful, in which case we keep $X \in A$
    - or it lies, in which case we replace $X \in A$ by $X \in \overline{A}$
- Relevance R defined on $\mathcal{R} = \{\text{rel}, \neg\text{rel}\}$
- Truthfulness T defined on $\mathcal{T} = \{\text{tru}, \neg\text{tru}\}$
- Let $\mathcal{RT} := \mathcal{R} \times \mathcal{T}$
- The interpretation of the testimony according to the relevance and truthfulness may be encoded by $\Lambda_A : \mathcal{RT} \to 2^{\mathcal{X}}$ such that

$$
\begin{aligned}
\Lambda_A(\text{rel}, \text{tru}) &= A \\
\Lambda_A(\text{rel}, \neg\text{tru}) &= \overline{A} \\
\Lambda_A(\neg\text{rel}, \text{tru}) &= \mathcal{X} \\
\Lambda_A(\neg\text{rel}, \neg\text{tru}) &= \mathcal{X}
\end{aligned}
$$

Example

- Sensor $\mathfrak{s}$ tells $X \in A = \overline{\{a\}} = \{g, h\}$.
- It is assumed to be relevant and non truthful, i.e., in state (rel, ¬tru)
- Knowledge about X:

$$X \in \Lambda_{\overline{\{a\}}}(\text{rel}, \neg\text{tru}) = \{a\}$$

Uncertain meta-knowledge and testimony

- Assume now $\mathfrak{s}$ is in state $(r, t) \in \mathcal{RT}$ with probability $p_{rt} = P^{\mathcal{RT}}(R = r, T = t)$ (and still $P^{\mathcal{R}}(R = \neg\text{rel}) = \pi$)
- In addition, $\mathfrak{s}$ provides the uncertain testimony $m_{\mathfrak{s}}$.
- This yields the following knowledge on $\mathcal{X}$

$$m[P^{\mathcal{RT}}, m_{\mathfrak{s}}] \quad = \quad p_{\text{rel,tru}} \cdot m_{\mathfrak{s}} + p_{\text{rel},\neg\text{tru}} \cdot \overline{m}_{\mathfrak{s}} + \pi m_{\mathcal{X}}$$

with $\overline{m}_{\mathfrak{s}}$ the negation of $m_{\mathfrak{s}}$ ($\overline{m}_{\mathfrak{s}}(A) = m_{\mathfrak{s}}(\overline{A})$, for all $A \subseteq \mathcal{X}$)

Particular cases

$m[P^{\mathcal{RT}}, m_{\mathfrak{s}}]$ reduces to

- $m[P^{\mathcal{R}}, m_{\mathfrak{s}}]$ if $P^{\mathcal{T}}(T = \text{tru}) = 1$, i.e., if $\mathfrak{s}$ is partially relevant and totally truthful

$\rightarrow$ discounting

- $P^{\mathcal{T}}(T = \text{tru}) \cdot m_{\mathfrak{s}} + P^{\mathcal{T}}(T = \neg\text{tru}) \cdot \overline{m}_{\mathfrak{s}}$ if $P^{\mathcal{R}}(R = \text{rel}) = 1$, i.e., if $\mathfrak{s}$ is totally relevant and partially truthful

$\rightarrow$ negating

The case of multiple sources

- Assume sources $\mathfrak{s}_i$, $i = 1\ldots, K$, supplying $\mathbf{A} = (A_1, \ldots, A_K)$.
- Let $\Lambda_{A_i} : \mathcal{RT}_i \rightarrow 2^{\mathcal{X}}$ represent the interpretation of $X \in A_i$ given the reliability R_i and truthfulness T_i of $\mathfrak{s}_i$
- When the sources are in the state

$$\mathbf{rt} = (rt_1, \ldots, rt_K) \in \boldsymbol{\mathcal{RT}} := \times_{i=1}^{K} \mathcal{RT}_i$$

we must conclude

$$X \in \Lambda_{\mathbf{A}}(\mathbf{rt}) := \bigcap_{i=1}^{K} \Lambda_{A_i}(rt_i)$$

- Example: K=2

$$\begin{aligned}\Lambda_{\mathbf{A}}(\mathrm{rel}_1, \mathrm{tru}_1, \mathrm{rel}_2, \neg\mathrm{tru}_2) &= \Lambda_{A_1}(\mathrm{rel}_1, \mathrm{tru}_1) \cap \Lambda_{A_2}(\mathrm{rel}_2, \neg\mathrm{tru}_2) \\ &= A_1 \cap \overline{A_2}\end{aligned}$$

Non-elementary behavior assumptions

- Non-elementary assumptions $\mathbf{RT} \subseteq \mathcal{RT}$ on the relevance and truthfulness of the sources can also be considered.
- We have

$$\Lambda_{\mathbf{A}}(\mathbf{RT}) = \bigcup_{\mathbf{rt} \in \mathbf{RT}} \Lambda_{\mathbf{A}}(\mathbf{rt})$$

- Example: $\mathbf{RT} = \{(\mathrm{rel}_1, \mathrm{tru}_1, \mathrm{rel}_2, \neg\mathrm{tru}_2), (\mathrm{rel}_1, \neg\mathrm{tru}_1, \mathrm{rel}_2, \mathrm{tru}_2)\}$ ($\mathfrak{s}_1$ and $\mathfrak{s}_2$ relevant and exactly one of them is truthful)

$$\begin{aligned} \Lambda_{\mathbf{A}}(\mathbf{RT}) &= \Lambda_{\mathbf{A}}(\mathrm{rel}_1, \mathrm{tru}_1, \mathrm{rel}_2, \neg\mathrm{tru}_2) \cup \Lambda_{\mathbf{A}}(\mathrm{rel}_1, \neg\mathrm{tru}_1, \mathrm{rel}_2, \mathrm{tru}_2) \\ &= (A_1 \cap \overline{A_2}) \cup (\overline{A_1} \cap A_2) \\ &= A_1 \Delta A_2 \text{ (exclusive or)} \end{aligned}$$

$\rightarrow$ All connectives of Boolean logic can be reinterpreted in terms of source behavior assumptions wrt relevance and truthfulness

- $\otimes_{\mathbf{RT}}$ Boolean connective associated to **RT**.
- Different assumptions may induce the same connective.

Uncertain meta-knowledge and testimonies

- Sources $\mathfrak{s}_1, \ldots, \mathfrak{s}_K$ provide $\mathbf{m} = (m_1, \ldots, m_K)$ and are assumed to be independent.
- Uncertain meta-knowledge in the form of a MF $m^{\mathcal{RT}}$:

$$
\begin{aligned}
m[m^{\mathcal{RT}}, \mathbf{m}](B) &= \sum_{\mathbf{RT}, \mathbf{A} : \Lambda_{\mathbf{A}}(\mathbf{RT}) = B} m^{\mathcal{RT}}(\mathbf{RT}) \cdot m(\mathbf{A}) \\
&= \sum_{\otimes, \mathbf{A} : \otimes(\mathbf{A}) = B} p(\otimes) \cdot m(\mathbf{A})
\end{aligned}
$$

with

$$
p(\otimes) = \sum_{\otimes_{\mathbf{RT}} = \otimes} m^{\mathcal{RT}}(\mathbf{RT})
$$

Particular cases

- Suppose $m^{\mathcal{RT}}$ is such that $m^{\mathcal{RT}}(\mathbf{RT}) = 1$ for some $\mathbf{RT} \subseteq \mathcal{RT}$ and $K = 2$. Then, $m[m^{\mathcal{RT}}, \mathbf{m}]$ reduces to
  - $m_1 \textcircled{\cap} m_2$ for **RT**= $\mathfrak{s}_1$ and $\mathfrak{s}_2$ relevant and truthful
  - $m_1 \textcircled{\cup} m_2$ for **RT**= $\mathfrak{s}_1$ and $\mathfrak{s}_2$ relevant and at least one of them is truthful (disjunctive rule, def. of $\textcircled{\cap}$ with $\cap$ replaced by $\cup$)
  - $m_1 \underline{\textcircled{\cup}} m_2$ for **RT**= $\mathfrak{s}_1$ and $\mathfrak{s}_2$ relevant and exaclty one of them is truthful (exclusive disjunctive rule, $\cap$ replaced by Δ)
  - $m_1 \underline{\textcircled{\cap}} m_2$ for **RT**= $\mathfrak{s}_1$ and $\mathfrak{s}_2$ relevant and $\mathfrak{s}_1$ is truthful if and only if $\mathfrak{s}_2$ is so too (equivalence rule, $\cap$ replaced by $\leftrightarrow$)
- More generally, all rules relying on Boolean connectives are particular cases. For instance, the rule extending q-relaxation from interval analysis is recovered for **RT**= $(K - q)$-out-of-K sources relevant and all truthful (ranges from $\textcircled{\cap}$ to $\textcircled{\cup}$).

Outline

Contextual and polarized lack of truthfulness

- $\neg$tru : one must deduce the contrary of what $\mathfrak{s}$ tells for each $x_i \in \mathcal{X}$ and whatever the polarity of the clause used by $\mathfrak{s}$ regarding x_i.

Contextual and polarized lack of truthfulness

- $\neg$tru : one must deduce the contrary of what $\mathfrak{s}$ tells for each $x_i \in \mathcal{X}$ and whatever the polarity of the clause used by $\mathfrak{s}$ regarding x_i.
- $\mathfrak{s}$ non truthful only for some $x_i \in \mathcal{X}$, and maybe even only for the positive or negative clauses regarding x_i.

Contextual and polarized lack of truthfulness

- $\neg$tru : one must deduce the contrary of what $\mathfrak{s}$ tells for each $x_i \in \mathcal{X}$ and whatever the polarity of the clause used by $\mathfrak{s}$ regarding x_i.
- $\mathfrak{s}$ non truthful only for some $x_i \in \mathcal{X}$, and maybe even only for the positive or negative clauses regarding x_i.
- Example : Sensor $\mathfrak{s}$ is
  - non truthful when it tells that a is not a possible value for X,
  - non truthful when it tells that g is a possible value for X,
  - and truthful in all other cases, e.g., truthful when it tells that a is a possible value for X
- Sensor $\mathfrak{s}$ tells $X \in A = \{g, h\}$, i.e., a is not a possible value and g and h are possible values for X
- We deduce (assuming $\mathfrak{s}$ relevant): $X \in \{a, h\}$

Contextual and polarized lack of truthfulness

- $\neg$tru : one must deduce the contrary of what $\mathfrak{s}$ tells for each $x_i \in \mathcal{X}$ and whatever the polarity of the clause used by $\mathfrak{s}$ regarding x_i.
- $\mathfrak{s}$ non truthful only for some $x_i \in \mathcal{X}$, and maybe even only for the positive or negative clauses regarding x_i.
- Example : Sensor $\mathfrak{s}$ is
  - non truthful when it tells that a is not a possible value for X,
  - non truthful when it tells that g is a possible value for X,
  - and truthful in all other cases, e.g., truthful when it tells that a is a possible value for X
- Sensor $\mathfrak{s}$ tells $X \in A = \{g, h\}$, i.e., a is not a possible value and g and h are possible values for X
- We deduce (assuming $\mathfrak{s}$ relevant): $X \in \{a, h\}$

$\rightarrow$ By considering source states based on this refined form of lack of truthfulness, we can recover contextual discounting and the α-junctions, and contextualize negating (see appendix)

Outline

Beyond relevance and truthfulness

- Knowledge about the source quality may be different from knowing their relevance and truthfulness
- The provided information by a source may also bear on another variable Y, related to X.

$\rightarrow$ An approach to account for general source quality (behaviour) assumptions

$$\begin{array}{rcl} \mathcal{R}, \mathcal{RT} & \rightsquigarrow & \mathcal{H} = \{h^1, \ldots, h^N\} \\ X \in A \subseteq \mathcal{X} & \rightsquigarrow & Y \in A \subseteq \mathcal{Y} \end{array}$$

- If the source is in state $h \in \mathcal{H}$, we should deduce $X \in B \subseteq \mathcal{X}$ from information item $Y \in A \subseteq \mathcal{Y}$.
- For all $A \subseteq \mathcal{Y}$, $\Pi_A : \mathcal{H} \rightarrow 2^{\mathcal{X}}$ such that

$$\Pi_A(h) = B$$

Example

- X with possible values in $\mathcal{X} = \{a, g, h\}$
- Sensor $\mathfrak{s}$ does not know the type airplane, i.e., $\mathcal{Y} = \{g, h\}$.
- It uses either the shape or the material of the aircraft
  - If $\mathfrak{s}$ uses the shape, then when it tells
    - glider, we can deduce airplane or glider
    - helicopter, we keep this piece of information
  - If $\mathfrak{s}$ uses the material, then when it tells
    - glider, we keep this piece of information
    - helicopter, we replace by helicopter or airplane

Example

- X with possible values in $\mathcal{X} = \{a, g, h\}$
- Sensor $\mathfrak{s}$ does not know the type airplane, i.e., $\mathcal{Y} = \{g, h\}$.
- It uses either the shape or the material of the aircraft
  - If $\mathfrak{s}$ uses the shape, then when it tells
    - glider, we can deduce airplane or glider
    - helicopter, we keep this piece of information
  - If $\mathfrak{s}$ uses the material, then when it tells
    - glider, we keep this piece of information
    - helicopter, we replace by helicopter or airplane
- $\mathcal{H} = \{\text{shape}, \text{material}\}$

$$\begin{aligned} \Pi_g(\text{shape}) &= \{a, g\} \\ \Pi_h(\text{shape}) &= \{h\} \\ \Pi_g(\text{material}) &= \{g\} \\ \Pi_h(\text{material}) &= \{a, h\} \end{aligned}$$

Example

- We are interested by the number $X \in \mathcal{X} = \{x_1, \ldots, x_n\}$=$\{1, \ldots, n\}$ of aircrafts in a particular area.
- Information about X comes from a source $\mathfrak{s}$, which can be reliable, approximately reliable or non reliable.
- If $\mathfrak{s}$ is approximately reliable, the information item it supplies must be expanded using the lowest and highest closest values.

Example

- We are interested by the number $X \in \mathcal{X} = \{x_1, \ldots, x_n\}$=$\{1, \ldots, n\}$ of aircrafts in a particular area.
- Information about X comes from a source $\mathfrak{s}$, which can be reliable, approximately reliable or non reliable.
- If $\mathfrak{s}$ is approximately reliable, the information item it supplies must be expanded using the lowest and highest closest values.
- $\mathcal{H} = \{\text{rel}, \text{ap-rel}, \neg\text{rel}\}$
- For any $A_{i,j} \subseteq \mathcal{X}$, with $A_{i,j} = \{x_i, \ldots, x_j\}$, $1 \leq i \leq j \leq n$

$$
\begin{aligned}
\Pi_{A_{i,j}}(\text{rel}) &= A_{i,j} \\
\Pi_{A_{i,j}}(\text{ap-rel}) &= \{x_{i-1}\} \cup A_{i,j} \cup \{x_{j+1}\} \\
\Pi_{A_{i,j}}(\neg\text{rel}) &= \mathcal{X}
\end{aligned}
$$

with $x_0 = x_{n+1} = \emptyset$.

Uncertain meta-knowledge and testimonies

- Single information source

$$m[m^{\mathcal{H}}, m^{\mathcal{Y}}_{\mathfrak{S}}](B) = \sum_{H,A:\Pi_A(H)=B} m^{\mathcal{H}}(H) \cdot m^{\mathcal{Y}}_{\mathfrak{S}}(A)$$

Behaviour-based correction (BBC)

Uncertain meta-knowledge and testimonies

- Single information source

$$m[m^{\mathcal{H}}, m^{\mathcal{Y}}_{\mathfrak{S}}](B) \;=\; \sum_{H,A:\Pi_A(H)=B} m^{\mathcal{H}}(H) \cdot m^{\mathcal{Y}}_{\mathfrak{S}}(A)$$

Behaviour-based correction (BBC)

- Multiple information sources: $\boldsymbol{\mathcal{H}} := \times_{i=1}^{K} \mathcal{H}_i$

$$m[m^{\boldsymbol{\mathcal{H}}}, \mathbf{m}](B) \;=\; \sum_{\mathbf{H},\mathbf{A}:\Pi_{\mathbf{A}}(\mathbf{H})=B} m^{\boldsymbol{\mathcal{H}}}(\mathbf{H}) \cdot m(\mathbf{A})$$

with $m(\mathbf{A}) = \prod_{i=1}^{K} m_i^{\mathcal{Y}}(A_i)$

Behaviour-based fusion (BBF)

Operations on product spaces

BBC and BBF can be recovered using the following standard operations of BFT :

- Marginalization $\downarrow$

$$m^{\mathcal{X}\times\mathcal{Y}\downarrow\mathcal{X}}(A) = \sum_{\{B\subseteq\mathcal{X}\times\mathcal{Y}, (B\downarrow X)=A\}} m^{\mathcal{X}\times\mathcal{Y}}(B), \quad \forall A \subseteq \mathcal{X},$$

- Conjunctive rule on product spaces

$$m_1^{\mathcal{X}} \textcircled{\cap} m_2^{\mathcal{Y}} = m_1^{\mathcal{X}\uparrow\mathcal{X}\times\mathcal{Y}} \textcircled{\cap} m_2^{\mathcal{Y}\uparrow\mathcal{X}\times\mathcal{Y}}.$$

with $\uparrow$ (vacuous extension) defined as

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

BBC

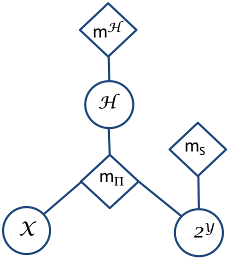
- Mappings Π_A, $A \subseteq \mathcal{Y}$, define a relation between spaces $\mathcal{H}$, $2^{\mathcal{Y}}$ and $\mathcal{X}$, which can be represented by MF m_Π on $\mathcal{H} \times 2^{\mathcal{Y}} \times \mathcal{X}$ s.t.

$$m_\Pi \left[\bigcup_{h \in \mathcal{H}, A \in 2^{\mathcal{Y}}} (\{h\} \times \{A\} \times \Pi_A(h)) \right] = 1$$

BBC

- Mappings Π_A, $A \subseteq \mathcal{Y}$, define a relation between spaces $\mathcal{H}$, $2^{\mathcal{Y}}$ and $\mathcal{X}$, which can be represented by MF m_Π on $\mathcal{H} \times 2^{\mathcal{Y}} \times \mathcal{X}$ s.t.

$$m_\Pi \left[\bigcup_{h \in \mathcal{H}, A \in 2^{\mathcal{Y}}} (\{h\} \times \{A\} \times \Pi_A(h)) \right] = 1$$

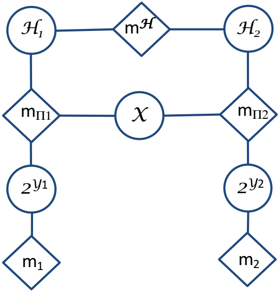


Lemma

$$m[m^{\mathcal{H}}, m_{\mathfrak{S}}^{\mathcal{Y}}] = (m_{\mathfrak{S}} \circledcap m_\Pi \circledcap m^{\mathcal{H}})^{\downarrow \mathcal{X}}$$

with $m_{\mathfrak{S}}$ on $2^{\mathcal{Y}}$ s.t. $m_{\mathfrak{S}}(\{A\}) = m_{\mathfrak{S}}^{\mathcal{Y}}(A)$

BBF



Lemma

$$m[m^{\mathcal{H}}, \mathbf{m}] \quad = \quad \left(\bigcap_{i=1}^{K} (m_i \bigcap m_{\Pi i}) \bigcap m^{\mathcal{H}} \right)^{\downarrow \mathcal{X}}$$

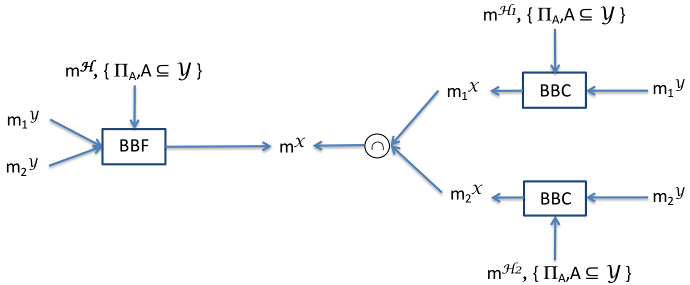
Independent behaviours (meta-independence)

Theorem

If $m^{\mathcal{H}} = \textcircled{\cap}_{i=1}^{K} m^{\mathcal{H}_i}$ *then*

$$m[m^{\mathcal{H}}, \mathbf{m}] = \textcircled{\cap}_{i=1}^{K} m[m^{\mathcal{H}_i}, m_i]$$

Proof: Uses local computation (see Prakash's lecture).



Outline

Typology of approaches

- The model allows to interpret pieces of information given meta-knowledge on the emitting sources.
- It does not however indicate which meta-knowledge to use.

$\rightarrow$ Means to select meta-knowledge

- Two possible situations:
  1. One has some prior information (learning data, expert knowledge) on the sources
  2. The only available information are the pieces of information received
- Typically, in both cases, a set $\mathcal{S}$ of candidate assumptions (meta-knowledge) is considered, and some sensible strategy is used to pick an assumption in this set.

Outline

Consistency and specificity

- Only $\mathbf{m} = (m_1, \ldots, m_K)$ available.

→ Selection of meta-knowledge based on the two primary features sought regarding knowledge about X: consistency and specificity

Consistency and specificity

- Only $\mathbf{m} = (m_1, \ldots, m_K)$ available.

→ Selection of meta-knowledge based on the two primary features sought regarding knowledge about X: consistency and specificity

- 3 sources about $X \in \mathcal{X} = \{a, g, h\}$ supplying $\mathbf{A} = (A_1, A_2, A_3)$ s.t.

$$A_1 = \{a\}, A_2 = \{a, g\}, A_3 = \{g, h\}$$

- Assumption $\mathbf{R}_1$ = "all sources are reliable" yields

$$X \in \Gamma_{\mathbf{A}}(\mathbf{R}_1) = A_1 \cap A_2 \cap A_3 = \emptyset$$

  i.e. an inconsistent result, and thus cannot hold.

- In contrast, the assumption $\mathbf{R}_3$ = "at least one of the sources is reliable" yields

$$X \in \Gamma_{\mathbf{A}}(\mathbf{R}_3) = A_1 \cup A_2 \cup A_3 = \mathcal{X}$$

  and is thus plausible (it does not yield a contradiction). However, it is useless as it is not informative at all.

Meta-knowledge selection strategy

- The intermediate assumption $\mathbf{R}_2 =$ "at least two of the sources are reliable" yields

$$X \in \Gamma_{\mathbf{A}}(\mathbf{R}_2) = (A_1 \cap A_2) \cup (A_1 \cap A_3) \cup (A_2 \cap A_3) = \{a, g\}$$

  $\mathbf{R}_2$ is plausible (the result is consistent) and informative (or, at least, more informative than $\mathbf{R}_3$).

Meta-knowledge selection strategy

- The intermediate assumption $\mathbf{R}_2$ = "at least two of the sources are reliable" yields

$$X \in \Gamma_{\mathbf{A}}(\mathbf{R}_2) = (A_1 \cap A_2) \cup (A_1 \cap A_3) \cup (A_2 \cap A_3) = \{a, g\}$$

  $\mathbf{R}_2$ is plausible (the result is consistent) and informative (or, at least, more informative than $\mathbf{R}_3$).
- Here, $\mathbf{R}_2$ is preferable, but for other $\mathbf{A}$, it could be $\mathbf{R}_1$ or $\mathbf{R}_3$ due to

$$\Gamma_{\mathbf{A}}(\mathbf{R}_1) \subseteq \Gamma_{\mathbf{A}}(\mathbf{R}_2) \subseteq \Gamma_{\mathbf{A}}(\mathbf{R}_3), \quad \forall \mathbf{A}$$

  $\mathbf{R}_{i+1}$ will always yield a result that is on the hand at least as consistent as that of $\mathbf{R}_i$, but also on the other hand as most as specific as that of $\mathbf{R}_i$.

$\rightarrow$ Consistency and specificity are antagonists goals

- Sensible strategy for a given $\mathbf{A}$: test iteratively each $\mathbf{R}_i$ and select the first one which yields a consistent result (it will then be the most specific and consistent possible result).

Extension to uncertain meta-knowledge and testimonies

- In general, meta-knowledge and supplied information are uncertain, i.e., we have $\mathbf{m}^{\mathcal{H}}$ and $\mathbf{m} = (m_1, \ldots, m_K)$, and thus their interpretation is the MF $m[m^{\mathcal{H}}, \mathbf{m}]$ (assuming independent sources).
- Need extensions to MF of consistency and specificity in order to compare pieces of meta-knowledge:
  - consistency of a MF m: $\phi(m) = \max_{x \in \mathcal{X}} pl(x)$.
  - specificity: $m_1 \sqsubseteq m_2$ with $\sqsubseteq$ the specialization

Proposition

Let $m_1^{\mathcal{H}}$ and $m_2^{\mathcal{H}}$ be two assumptions.

$$m[m_1^{\mathcal{H}}, \mathbf{m}] \sqsubseteq m[m_2^{\mathcal{H}}, \mathbf{m}], \forall \mathbf{m} \Rightarrow \phi(m[m_1^{\mathcal{H}}, \mathbf{m}]) \leq \phi(m[m_2^{\mathcal{H}}, \mathbf{m}]), \forall \mathbf{m}$$

$\rightarrow$ Consistency and specificity are at odds !

General meta-knowledge selection strategy

Strategy

1. Define a set $\mathcal{S} = \{m_1^{\mathcal{H}}, ..., m_M^{\mathcal{H}}\}$:
   - $m[m_j^{\mathcal{H}}, \mathbf{m}] \sqsubseteq m[m_{j+1}^{\mathcal{H}}, \mathbf{m}], \forall \mathbf{m}$;
   - $m_1^{\mathcal{H}}$ corresponds to the conjunctive rule.
2. Test iteratively each $m_j^{\mathcal{H}}$ until $\phi(\mathbf{m}[m_j^{\mathcal{H}}, \mathbf{m}]) \geq \tau$.

- Practical instances of $\mathcal{S}$:
  - $m_j^{\mathcal{H}}$: $K - j + 1$ out of K reliable sources.
  - $m_j^{\mathcal{H}}$: sources with independent reliabilities, source i reliable with probability p_i^j such that $p_i^j \geq p_i^{j+1}$ (increasing discount and combine, often used for conflict management)
  - $m_j^{\mathcal{H}}$: meta-knowledge corresponding to the α-conjunctions for some $\alpha = \alpha_j$ such that $\alpha_j \geq \alpha_{j+1}$.

Application

Nuclear reactor safety

- Project BEMUSE of the Nuclear Energy Agency.
- $K = 10$ sources (CEA, IRSN,...) providing uncertain estimates of parameter values of a nuclear power plant.
- Costly data and complex phenomena involved $\rightarrow$ no reliable means to know the source reliabilities.
- Chose $\mathcal{S}$ with $m_j^{\mathcal{H}} = K - j + 1$ out of K reliable sources.
- PCT2 parameter with domain $\mathcal{X} = \{x_1, \ldots, x_6\}$, $\mathbf{m} := (m_1, \ldots, m_{10})$.
  - $\phi(m[m_1^{\mathcal{H}}, \mathbf{m}]) = 0.19$ (all sources reliable)
  - $\phi(m[m_2^{\mathcal{H}}, \mathbf{m}]) = 0.81$ (9 out of 10 reliable)
  - $\phi(m[m_3^{\mathcal{H}}, \mathbf{m}]) = 1$ (8 out of 10 reliable)
  - Values x_4 and x_5 are definitely more plausible.

$\rightarrow$ Results that are consistent, informative and readable by the end-user.

Outline

General setting

- Consider a system which outputs for a given object o, a guess about the actual value x^* of some feature $X \in \mathcal{X}$ of o.
- To produce this output, the system uses internally some information correction or fusion, characterized by some $m^{\mathcal{H}} \in \mathcal{S}$.
- Output for object o may thus be noted $f(o; m^{\mathcal{H}})$.

General setting

- Consider a system which outputs for a given object o, a guess about the actual value x^* of some feature $X \in \mathcal{X}$ of o.
- To produce this output, the system uses internally some information correction or fusion, characterized by some $m^{\mathcal{H}} \in \mathcal{S}$.
- Output for object o may thus be noted $f(o; m^{\mathcal{H}})$.
- Assume a set of ℓ objects for which the true value of X is known, i.e., $\{x_i^*\}_{i=1}^{\ell}$ is available.
- Assume outputs $\{f(o_i; m^{\mathcal{H}})\}_{i=1}^{\ell}$ may be obtained for any $m^{\mathcal{H}} \in \mathcal{S}$.

Loss minimization

- The $\hat{m}^{\mathcal{H}}$ to be used to produce the output for a new object may then be the chosen as the one in $\mathcal{S}$ minimizing the average loss

$$J(m^{\mathcal{H}}) = \frac{1}{n}\sum_{i=1}^{\ell} \mathcal{L}(f(o_i; m^{\mathcal{H}}), x_i^*)$$

  for some loss function $\mathcal{L}(f(o; m^{\mathcal{H}}), x^*)$
- Typically, $f(o; m^{\mathcal{H}})$ is a MF on $\mathcal{X}$, which is transformed into a probability measure $P_o^{\mathcal{X}}$, and the squared error (SE) or cross-entropy (CE) loss is used:

$$\begin{aligned}\mathcal{L}_{SE}(f(o; m^{\mathcal{H}}), x^*) &= \sum_{x\in\mathcal{X}} (1_{x^*}(x) - p_o(x))^2 \\ \mathcal{L}_{CE}(f(o; m^{\mathcal{H}}), x^*) &= -\sum_{x\in\mathcal{X}} 1_{x^*}(x) \log p_o(x)\end{aligned}$$

- Remark: more or less complex optimisation problem to solve depending on $\mathcal{S}$ and $\mathcal{L}$

Application

Classifier correction [Elouedi et al., 2004]

- X is the class of an object.
- The system is a classifier whose outputs are corrected with meta-knowledge $m^{\mathcal{H}} = P^{\mathcal{R}}$ (discounting) with

$$P^{\mathcal{R}} \in \mathcal{S} = \{P^{\mathcal{R}} | \pi \in [0, 1]\}$$

- The classifier output for a given object o is a mass function m_o.
- The system output is thus

$$f(o; m^{\mathcal{H}}) = m[P^{\mathcal{R}}, m_o]$$

- Loss function : pignistic probability transformation with SE.

Application

Illustrative example

- Classifier outputs m_{o_i} for 4 objects with actual values x_i^* in $\mathcal{X} = \{a, g, h\}$.

	g	h	$\{a,h\}$	$\{g,h\}$	$\mathcal{X}$	x_i^*
m_{o_1}	0	0.5	0	0.3	0.2	a
m_{o_2}	0.5	0.2	0	0	0.3	g
m_{o_3}	0.4	0	0.6	0	0	a
m_{o_4}	0	0	0.6	0.4	0	h

- Meta-knowledge minimizing the average loss: $\hat{\pi} = 0.66$

Summary

- Interpretation of BFT as a theory of partially reliable and elementary pieces of information
  - Any set of such pieces of information is represented by a unique MF
  - To any MF can be associated uniquely such a set.
- Beyond reliability, information correction and fusion given knowledge on other aspects of source quality, such as truthfulness.
- Numerous and important correction and fusion approaches can be read using this prism.
- Means to determine knowledge on source quality in practice, with and without prior information on the sources.

Open topics of interest

- Exploitation of the $\bigcirc\!\!\!\!\cap_\sigma$ rule for SMF and the associated decomposition of a MF into (in)dependent SMF
  - Cautious combination
  - Refining of approaches based on conjunctive combination of independent SMF, such as GBT, E-KNN, DS analysis of GLR classifiers, contextual reinforcement.
- Interpretation of other correction and fusion approaches.
- Selection of meta-knowledge: refine arguments for the
  - Choice of $\mathcal{S}$ (include dependence between the sources)
  - Choice of $\mathcal{L}$ (including for the case of partially known true values)
- Conflict measurement: decomposition, measure selection for a given situation (properties, learning), refine with measures from logic, links with distances

References

G. Shafer.
A Mathematical Theory of Evidence.
Princeton University Press, Princeton, NJ, 1976.

P. Smets.
The canonical decomposition of a weighted belief.
In Proc. of IJCAI 1995, p.1896–1901. Morgan Kaufmann, 1995.

Z. Elouedi, K. Mellouli, and P. Smets.
Assessing Sensor Reliability for Multisensor Data Fusion Within the Transferable Belief Model.
IEEE Transactions on Systems, Man and Cybernetics, Part B, 34(1):782–787, 2004.

F. Pichon, D. Dubois, and T. Denœux.
Relevance and truthfulness in information correction and fusion.
Int. Journal of Approximate Reasoning, 53(2):159–175, 2012.

F. Pichon, S. Destercke, and T. Burger.
A consistency-specificity trade-off to select source behavior in information fusion.
IEEE Transactions on Cybernetics, 45(4):598–609, 2015.

F. Pichon.
Canonical decomposition of belief functions based on Teugels' representation of the multivariate Bernoulli distribution.
Information Sciences, 428:76–104, 2018.

T. Denœux.
Logistic regression, neural networks and Dempster-Shafer theory: a new perspective.
Knowledge-Based Systems, 176:54–67, 2019.

Additional bibliography

The following bibliography contains:

- Some other relevant BFT-based references on modeling and selecting assumptions on sources
- Some other interesting references, and in particular some more application-oriented papers, where correction/fusion is not the main topic but plays an important part.

Modeling assumptions on sources I

P. Smets.
Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem.
Int. Journal of Approximate Reasoning, 9(1):1–35, 1993.

P. Smets.
The α-junctions: combination operators applicable to belief functions.
In Proc. of ECSQARU-FAPR 1997, LNCS, vol. 1244, p131–153. Springer, 1997.

R. Haenni.
Uncover Dempster's rule where it is hidden.
In Proc. of FUSION 2006, 2006.

R. Haenni, and S. Hartmann.
Modeling partially reliable information sources: a general approach based on Dempster-Shafer theory.
Information Fusion, 7(4):361–379, 2006.

P. Smets.
Analyzing the combination of conflicting belief functions.
Information Fusion, 8:397–412, 2007.

T. Denœux.
Conjunctive and disjunctive combination of belief functions induced by non-distinct bodies of evidence.
Artificial Intelligence 172:234–264, 2008.

A. Kallel, and S. Le Hégarat-Mascle.
Combination of partially nondistinct beliefs : the cautious-adaptive rule.
Int. Journal of Approximate Reasoning, 50(7):1000–1021, 2009.

Modeling assumptions on sources II

D. Mercier, T. Denœux, and M.-H. Masson.
Belief function correction mechanisms.
In Foundations of Reasoning under Uncertainty, vol. 249, Studies in Fuzziness and Soft Computing, p. 203–222. Springer-Verlag, 2010.

M. E. G. V. Cattaneo.
Belief functions combination without the assumption of independence of the information sources.
Int. Journal of Approximate Reasoning, 52(3):299–315, 2011.

S. Destercke, and D. Dubois.
Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory.
Information Sciences, 181(18):3925–3945, 2011.

F. Pichon.
On the α-conjunctions for combining belief functions.
In Proc. of BELIEF 2012, Advances in Soft Computing, vol. 164, p. 285–292. Springer, 2012.

D. Dubois, W. Liu, J. Ma, and H.Prade.
The basic principles of uncertain information fusion. An organised review of merging rules in different representation frameworks.
Information Fusion, 32:12–39, 2016.

D. Mercier, F. Pichon, and É. Lefèvre.
Corrigendum to "Belief functions contextual discounting and canonical decompositions"
Int. Journal of Approximate Reasoning, 70:137–139, 2016.

Modeling assumptions on sources III

J. Klein, S. Destercke, O. Colot.
Idempotent conjunctive and disjunctive combination of belief functions by distance minimization
Int. Journal of Approximate Reasoning 92:32–48, 2018.

F. Pichon, D. Dubois, and T. Denœux.
Quality of information sources in information fusion.
In Information Quality in Information Fusion and Decision Making, p. 31–49, Springer, 2019.

Selecting assumptions on sources I

A. Martin, A.-L. Jousselme, and C. Osswald.
Conflict measure for the discounting operation on belief functions.
In Proc. of FUSION 2018, p. 1003–1010, 2008.

J. Klein, and O. Colot.
Automatic discounting rate computation using a dissent criterion.
In Proc. of BELIEF 2010, p. 1–6, 2010.

Z. Elouedi, E. Lefèvre, and D. Mercier.
Discountings of a belief function using a confusion matrix.
In Proc. of ICTAI 2010, p. 287–294, 2010.

J. Schubert.
Conflict management in Dempster-Shafer theory using the degree of falsity.
Int. Journal of Approximate Reasoning, 52(3):449–460, 2011.

Y. Yang, D. Han, and C. Han.
Discounted combination of unreliable evidence using degree of disagreement.
Int. Journal of Approximate Reasoning, 54(8):1197–1216, 2013.

S. Destercke, P. Buche, and B. Charnomordic.
Evaluating data reliability : an evidential answer with application to a web-enabled data warehouse.
IEEE Transactions on Knowledge and Data Engineering, 25(1):92–105, 2013.

É. Lefèvre, F. Pichon, D. Mercier, Z. Elouedi, and B. Quost.
Estimation de sincérité et pertinence à partir de matrices de confusion pour la correction de fonctions de croyance.
In Proc. of LFA 2014, p. 115–122, Cépaduès, 2014.

Selecting assumptions on sources II

F. Pichon, C. Labreuche, B. Duqueroie, and T. Delavallade.
Multidimensional Approach to Reliability Evaluation of Information Sources.
Dans Information evaluation, p. 129–156, Wiley-ISTE, 2014.

M. Chebbah, A. Martin, and B. Ben Yaghlane.
Combining partially independent belief functions.
Decision Support Systems, 73:37–46, 2015.

F. Pichon, D. Mercier, É. Lefèvre, and F. Delmotte.
Proposition and learning of some belief function contextual correction mechanisms.
Int. Journal of Approximate Reasoning, 72:4–42, 2016.

S. Mutmainah, S. Hachour, F. Pichon, D. Mercier.
On learning evidential contextual corrections from soft labels using a measure of discrepancy between contour functions.
In Proc. of SUM 2019, to appear.

Applications

T. Denœux.
A k-nearest neighbor classification rule based on Dempster-Shafer theory.
IEEE Transactions on Systems, Man and Cybernetics, 25(5):804–213, 1995.

S. Le Hégarat-Mascle, and R. Seltz.
Automatic change detection by evidential fusion of change indices.
Remote Sensing of Environment, 91:390–404, 2004.

A. Roquel, S. Le Hégarat-Mascle, I. Bloch, and B. Vincke.
Decomposition of conflict as a distribution on hypotheses in the framework on belief functions.
Int. Journal of Approximate Reasoning, 55:1129–1146, 2014.

M. Kurdej, J. Moras, V. Cherfaoui, and P. Bonnifait.
Map-Aided Evidential Grids for Driving Scene Understanding.
IEEE Intelligent Transportation Systems Magazine, 7(1):30–41, 2015.

J. Radak, B. Ducourthial, V. Cherfaoui, and S. Bonnet.
Detecting road events using distributed data fusion: Experimental evaluation for the icy roads case.
IEEE Transactions on Intelligent Transportation Systems, 17(1):184–194, 2016.

N. Pellicanò, S. Le Hégarat-Mascle, and E. Aldea.
2CoBel: A scalable belief function representation for 2D discernment frames.
Int. Journal of Approximate Reasoning, 103:320–342, 2018.

T. Denœux, O. Kanjanatarakul, and S. Sriboonchitta.
A New Evidential K-Nearest Neighbor Rule based on Contextual Discounting with Partially Supervised learning.
Int. Journal of Approximate Reasoning, 113:287–302, 2019.

Contextual and polarized lack of truthfulness

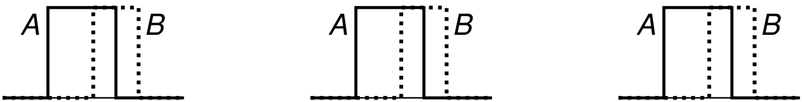
- $\neg$tru : one must deduce the contrary of what $\mathfrak{s}$ tells for each $x_i \in \mathcal{X}$ and whatever the polarity of the clause used by $\mathfrak{s}$ regarding x_i.
- $\mathfrak{s}$ non truthful only for some $x_i \in \mathcal{X}$, and maybe even only for the positive or negative clauses regarding x_i.
- Example : Sensor $\mathfrak{s}$ is
  - non truthful when it tells that a is not a possible value for X (negatively non truthful for a)
  - and non truthful when it tells that g is a possible value for X (positively non truthful for g)
  - and truthful in all other cases, e.g., truthful when it tells that a is a possible value for X (positively truthful for a).
- Sensor $\mathfrak{s}$ tells $X \in A = \{g, h\}$, i.e., a is not a possible value and g and h are possible values for X
- We deduce (assuming $\mathfrak{s}$ relevant): $X \in \{a, h\}$

Contextual and polarized lack of truthfulness

- Three interesting states (contextual lies):
  - n_B : negatively non truthfull for $x_i \in B$;
  - p_B : positively non truthful for $x_i \in \overline{B}$;
  - ℓ_B : non truthful for $x_i \in \overline{B}$.
- Let $\tilde{\mathcal{T}} = \{n_B, p_B, \ell_B | B \subseteq \mathcal{X}\}$ and $\tilde{\Lambda}_A : \mathcal{R}\tilde{\mathcal{T}} \rightarrow 2^X$ represent the interpretations of testimony $X \in A$ given the possible states in $\mathcal{R}\tilde{\mathcal{T}}$ of the source
- $\tilde{\Lambda}_A$ extends Λ_A, e.g., $\tilde{\Lambda}_A(\ell_\emptyset) = \Lambda_A(\neg \text{tru})$ and $\tilde{\Lambda}_A(\ell_\mathcal{X}) = \Lambda_A(\text{tru})$ (assuming relevance)

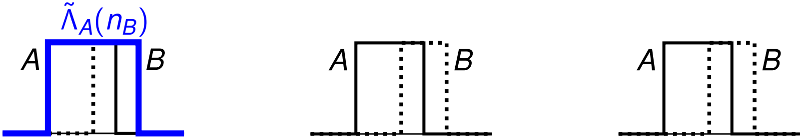
Contextual and polarized lack of truthfulness

- Three interesting states (contextual lies):
  - n_B : negatively non truthfull for $x_i \in B$;
  - p_B : positively non truthful for $x_i \in \overline{B}$;
  - ℓ_B : non truthful for $x_i \in \overline{B}$.
- Let $\tilde{\mathcal{T}} = \{n_B, p_B, \ell_B | B \subseteq \mathcal{X}\}$ and $\tilde{\Lambda}_A : \mathcal{R}\tilde{\mathcal{T}} \to 2^X$ represent the interpretations of testimony $X \in A$ given the possible states in $\mathcal{R}\tilde{\mathcal{T}}$ of the source
- $\tilde{\Lambda}_A$ extends Λ_A, e.g., $\tilde{\Lambda}_A(\ell_\emptyset) = \Lambda_A(\neg\text{tru})$ and $\tilde{\Lambda}_A(\ell_\mathcal{X}) = \Lambda_A(\text{tru})$ (assuming relevance)



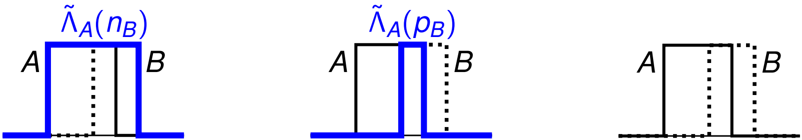
Contextual and polarized lack of truthfulness

- Three interesting states (contextual lies):
  - n_B : negatively non truthfull for $x_i \in B$;
  - p_B : positively non truthful for $x_i \in \overline{B}$;
  - ℓ_B : non truthful for $x_i \in \overline{B}$.
- Let $\tilde{\mathcal{T}} = \{n_B, p_B, \ell_B | B \subseteq \mathcal{X}\}$ and $\tilde{\Lambda}_A : \mathcal{R}\tilde{\mathcal{T}} \rightarrow 2^X$ represent the interpretations of testimony $X \in A$ given the possible states in $\mathcal{R}\tilde{\mathcal{T}}$ of the source
- $\tilde{\Lambda}_A$ extends Λ_A, e.g., $\tilde{\Lambda}_A(\ell_\emptyset) = \Lambda_A(\neg\text{tru})$ and $\tilde{\Lambda}_A(\ell_\mathcal{X}) = \Lambda_A(\text{tru})$ (assuming relevance)



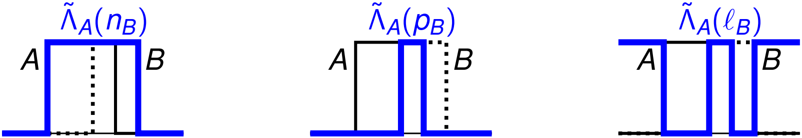
Contextual and polarized lack of truthfulness

- Three interesting states (contextual lies):
  - n_B : negatively non truthfull for $x_i \in B$;
  - p_B : positively non truthful for $x_i \in \overline{B}$;
  - ℓ_B : non truthful for $x_i \in \overline{B}$.
- Let $\tilde{\mathcal{T}} = \{n_B, p_B, \ell_B | B \subseteq \mathcal{X}\}$ and $\tilde{\Lambda}_A : \mathcal{R}\tilde{\mathcal{T}} \rightarrow 2^X$ represent the interpretations of testimony $X \in A$ given the possible states in $\mathcal{R}\tilde{\mathcal{T}}$ of the source
- $\tilde{\Lambda}_A$ extends Λ_A, e.g., $\tilde{\Lambda}_A(\ell_\emptyset) = \Lambda_A(\neg\text{tru})$ and $\tilde{\Lambda}_A(\ell_\mathcal{X}) = \Lambda_A(\text{tru})$ (assuming relevance)



Contextual and polarized lack of truthfulness

- Three interesting states (contextual lies):
  - n_B : negatively non truthfull for $x_i \in B$;
  - p_B : positively non truthful for $x_i \in \overline{B}$;
  - ℓ_B : non truthful for $x_i \in \overline{B}$.
- Let $\tilde{\mathcal{T}} = \{n_B, p_B, \ell_B | B \subseteq \mathcal{X}\}$ and $\tilde{\Lambda}_A : \mathcal{R}\tilde{\mathcal{T}} \to 2^X$ represent the interpretations of testimony $X \in A$ given the possible states in $\mathcal{R}\tilde{\mathcal{T}}$ of the source
- $\tilde{\Lambda}_A$ extends Λ_A, e.g., $\tilde{\Lambda}_A(\ell_\emptyset) = \Lambda_A(\neg\text{tru})$ and $\tilde{\Lambda}_A(\ell_\mathcal{X}) = \Lambda_A(\text{tru})$ (assuming relevance)



Uncertain meta-knowledge and testimonies

- Single information source

$$m[m^{\mathcal{R}\tilde{\mathcal{T}}}, m_{\mathfrak{s}}](B) = \sum_{R\tilde{T}, A : \tilde{\Lambda}_A(R\tilde{T}) = B} m^{\mathcal{R}\tilde{\mathcal{T}}}(R\tilde{T}) \cdot m_{\mathfrak{s}}(A)$$

- Multiple information sources

$$m[m^{\boldsymbol{\mathcal{R}}\tilde{\boldsymbol{\mathcal{T}}}}, \mathbf{m}](B) = \sum_{\mathbf{R}\tilde{\mathbf{T}}, \mathbf{A} : \tilde{\Lambda}_{\mathbf{A}}(\mathbf{R}\tilde{\mathbf{T}}) = B} m^{\boldsymbol{\mathcal{R}}\tilde{\boldsymbol{\mathcal{T}}}}(\mathbf{R}\tilde{\mathbf{T}}) \cdot m(\mathbf{A})$$

Particular cases

- Let $\mathcal{B} = \{B_1, \ldots, B_N\} \subseteq 2^X$. Consider iterative corrections (series of agents) of testimony $m_{\mathfrak{S}}$ provided by agent 1 with respective assumptions "preceding agent i is relevant, and is truthful with probability β_{B_i} and with probability $1 - \beta_{B_i}$ commits lie
  - n_{B_i}": $m_{\mathfrak{S}} \textcircled{\cup}_{B_i \in \mathcal{B}} m_{B_i}$ with $m_{B_i}(\emptyset) = \beta_{B_i}, m_{B_i}(B_i) = 1 - \beta_{B_i}$, called contextual discounting (it can also be obtained as a single correction $m[m^{\mathcal{R}\tilde{\mathcal{T}}}, m_{\mathfrak{S}}]$ with $m^{\mathcal{R}\tilde{\mathcal{T}}}$ the $\textcircled{\cup}$-combination of the preceding assumptions)
  - ℓ_{B_i}": $m_{\mathfrak{S}} \textcircled{\underline{\cap}}_{B_i \in \mathcal{B}} B_i^{\beta_{B_i}}$, contextual negating
  - p_{B_i}": $m_{\mathfrak{S}} \textcircled{\cap}_{B_i \in \mathcal{B}} B_i^{\beta_{B_i}}$, contextual reinforcement
- Remarks:
  - These correction mechanisms generalize their non-contextual versions for specific $\mathcal{B}$ such that $|\mathcal{B}| = 1$, hence their names.
  - An alternative interpretation exists for contextual discounting when $\mathcal{B}$ is a partition of $\mathcal{X}$ (see Thierry's lecture).

Example

Contextual discounting

- Suppose a sensor $\mathfrak{s}$ supplies information $X \in A = \{g\}$
- We know that $\mathfrak{s}$ is relevant and that at least one of the following independent pieces of meta-knowledge holds:
  - $\mathfrak{s}$ commits lie $n_{\{a,g\}}$ with probability 0.2
  - $\mathfrak{s}$ commits lie $n_{\{g,h\}}$ with probability 0.3
- Our knowledge on $\mathcal{X}$ is then obtained by

$$m_{\mathfrak{s}}(\{g\}) = 1 \ \text{\textcircled{$\cup$}} \left\{ \begin{array}{r} m_{\{a,g\}}(\{a,g\}) = 0.2 \\ m_{\{a,g\}}(\emptyset) = 0.8 \end{array} \right\} \text{\textcircled{$\cup$}} \left\{ \begin{array}{r} m_{\{g,h\}}(\{g,h\}) = 0.3 \\ m_{\{g,h\}}(\emptyset) = 0.7 \end{array} \right\}$$

which yields

$$m(\{g\}) = 0.56, m(\{a,g\}) = 0.14, m(\{g,h\}) = 0.24, m(\mathcal{X}) = 0.06$$

Particular cases

- Consider the following meta-knowledge about two sources $\mathfrak{s}_1$ and $\mathfrak{s}_2$ supplying information m_1 and m_2:
  - They are both relevant
  - And they are either both truthful or commit the same contextual lie ℓ_B with probability $\alpha^{|B|}(1-\alpha)^{|\overline{B}|}$, for some $\alpha \in [0,1]$
- Then

$$m[m^{\mathcal{R}\tilde{\mathcal{T}}}, \mathbf{m}](A) = \sum_{(A_1 \cap A_2) \cup (\overline{A_1} \cap \overline{A_2} \cap B) = A} m_1(A_1)\, m_2(A_2)\, m_\alpha(B)$$

where $m_\alpha(B) = \alpha^{|\overline{B}|}(1-\alpha)^{|B|}$

$\rightarrow$ α-conjunctions $\textcircled{\cap}^\alpha$ [Smets, 1997]: family of the associative, commutative and linear combination rules having the vacuous mass function as neutral element (family depending on a parameter $\alpha \in [0,1]$, such that $\textcircled{\cap}^1 = \textcircled{\cap}$ and $\textcircled{\cap}^0 = \underline{\textcircled{\cap}}$).