Prejudiced information fusion using belief functions

Didier Dubois (with Francis Faux and Henri Prade)

IRIT, CNRS, University of Toulouse, France

BELIEF school 2019 Sienna

Outline

Introduction

While Dempster introduced upper and lower probabilities induced by a probability space and a multiple-valued map (an ill-known random variable)

G. Shafer has presented his theory of belief functions (BFs) as an approach to the fusion of independent unreliable elementary testimonies by Demspter's rule of combination,

- each testimony represented by a simple support function.

In agreement with early steps in the theory of probability end of XVIIth century (Bernoulli, Hooper, Lambert etc.)

Basic concepts

Basic belief assignment

On the frame of discernment Ω, consider a basic belief assignment or mass function m.

- $m : 2^\Omega \longrightarrow [0,1]$ verifies $\sum_{A \subseteq \Omega} m(A) = 1$.
- Focal set: $A \subseteq \Omega$ such that $m(A) > 0$ (here including $\emptyset$)

Simple support function

- An elementary testimony with strength s in favor of proposition $x \in A \in 2^\Omega$ is represented by a simple support function:
  - $m(A) = s$, for $A \neq \Omega$ (confidence in A),
  - $m(\Omega) = d = 1 - s$, with $0 < d \leq 1$ (non-dogmaticism),

  and is denoted by $m = A^d$ where d expresses a degree of diffidence ($d = \mathrm{dif}(A) = Pl(\overline{A})$).

Separable belief functions

Definition

(Shafer) A belief function is separable iff it is the orthogonal sum of simple support functions.
(Denœux) A belief function is unnormalized-separable iff it is the conjunctive combination of simple support functions

If a BF m is u-separable then

- it is non-dogmatic : $m(\Omega) > 0$
- its focal sets are closed under intersection

$\longrightarrow$ Hence not all BFs are (u-)separable !

Combination rules for simple support functions

- Conjunctive rule $\textcircled{\cap}_{i=1}^{n} A_i^{d_i}$

$$m_{\textcircled{\cap}}(A) = \sum_{\cap_{i \in I} A_i = A} \prod_{i \in I} (1 - d_i) \cdot \prod_{j \notin I} d_j, \forall I \subseteq [n]$$

$(\ m_{\textcircled{\cap}}(\emptyset) = \sum_{I : \cap_{i \in I} A_i = \emptyset} \prod_{i \in I} (1 - d_i) \cdot \prod_{j \notin I} d_j .)$

- Orthogonal sum :

$$m_{\oplus}(A) = \frac{m_{\textcircled{\cap}}(A)}{m_{\textcircled{\cap}}(\emptyset)}, A \neq \emptyset$$

$$m_{\oplus}(\emptyset) = 0.$$

Separable belief functions using diffidence functions

diffidence functions

Consider the separable BF $m = \bigoplus_{i=1}^{n} A_i^{d_i}$. Define the function $\delta(.)$, called a diffidence function:

$$\forall A \subset \Omega, A, \delta(A) = \begin{cases} d_i, & \text{if } \exists i, A = A_i; \\ 1 & \text{otherwise.} \end{cases}$$

Then $m = \bigoplus_{\emptyset \neq A \subset \Omega} A^{\delta(A)}$. where $\delta(A) \leq 1, \forall A \subset \Omega$.

- If the belief function is not dogmatic ($m(\Omega) > 0$) and the A_i's are distinct, this representation is unique.
- For u-separable belief functions $m = \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{\delta(A)}$ and $\delta(\emptyset) > 0$ is allowed

Commonality and diffidence functions

It is known that $Q_m(A) = \sum_{B \supseteq A} m(B)$ and $Q_{m_1 \cap\!\!\!\bigcirc m_2} = Q_{m_1} \cdot Q_{m_2}$.

If $m_i = A_i^{d_i}, i = 1, \ldots, n$, $Q_{m_i}(A) = \begin{cases} 1 \text{ if } A \subseteq A_i \\ d_i \text{ otherwise.} \end{cases}$

If $m = \bigcirc\!\!\!\!\cap_{i=1}^{n} m_i$, then:

$$Q_m(A) = \prod_{i:A \not\subseteq A_i} d_i = \prod_{B:A \not\subseteq B} \delta(B), A \neq \emptyset$$

- $Q_m(\Omega) = \prod_{A \subset \Omega} \delta(A) = \prod_{i=1}^{n} d_i$
- $\{B : \emptyset \not\subseteq B\} = \emptyset$ and $Q_m(\emptyset) = 1$ imply $\prod_{B:\emptyset \not\subseteq B} \delta(B) = 1$

We can define a belief function using the diffidence function

Shafer (1976) proves that this is equivalent to

$$\delta(A) = \prod_{B \supseteq A} Q_m(B)^{(-1)^{|B|-|A|+1}}$$

Completing the definition of diffidence functions

The expression $\delta(A) = \prod_{B \supseteq A} Q(B)^{(-1)^{|B|-|A|+1}}$ makes sense for $A = \Omega$ (Kallel, Le Hegarat-Mascle, et al, 2008):

$$\delta(\Omega) = 1/Q(\Omega) = 1/\prod_{i=1}^{n} d_i$$

- Function δ can be extended to the whole of 2^Ω, even if only sets $A \subset \Omega$ appear in the decomposition formula.
- $\delta(\Omega) > 1$ but this will be also the case for the diffidence weights of other subsets for non-separable belief functions.

From BPA to diffidence and back

Consequence

A diffidence function δ computed from m is such that:

$$\prod_{A\subseteq\Omega}\delta(A)=\delta(\Omega)\cdot\prod_{A\subset\Omega}\delta(A)=\delta(\Omega)\cdot Q(\Omega)=1$$

Commonality function $Q_\delta(A)$ from δ

$$Q_\delta(A)=\prod_{B:A\not\subseteq B}\delta(B)=\frac{1}{\prod_{A\subseteq B}\delta(B)}$$

Retrieving m from δ via the commonality function

$$m(E)=\sum_{E\subseteq A}(-1)^{|B|-|A|}\left(\frac{1}{\prod_{A\subseteq B}\delta(B)}\right)$$

.

Extending the decomposition to all non-dogmatic belief functions

- Ph. Smets [IJCAI 1995] addresses this question by:
  - generalizing simple support functions (GSSFs) allowing diffidence values > 1,
  - writing a BF as the conjunctive combination of GSSFs.
- $\longrightarrow$ All non-dogmatic BFs are the result of such a combination.

Difficulties

masses of GSSFs can be negative !

$\longrightarrow$ Model **difficult to interpret**,

Generalized simple support functions

Smets introduced generalized SSF A^d such that $d \in (0, +\infty)$.
$\rightarrow$ We call d *the diffidence coefficient*,

- $d \rightarrow 0$, little diffidence $\longrightarrow$ high confidence in A,
- $d = 1$ agnosticism $\rightarrow$ total ignorance about A,
- $d = \infty$ full diffidence about A.

It gives a **general definition** of a diffidence function as a mapping:

$$\delta : 2^{\Omega} \rightarrow (0, +\infty)$$

$$\prod_{A \subseteq \Omega} \delta(A) = 1, \quad \delta(\Omega) \geq 1$$

Smets result

Any non dogmatic BFs can be decomposed as

$$m = \bigcirc\!\!\!\!\cap_{\emptyset \neq A \subset \Omega} A^{\delta(A)}$$

where $\forall A \subset \Omega$: $\delta(A) = \prod_{B \supseteq A} Q(B)^{(-1)^{|B|-|A|+1}} \in (0, +\infty)$

$$= \prod_{C \cap A = \emptyset} Q(A \cup C)^{(-1)^{|C|+1}} = \frac{\prod_{C \cap A = \emptyset, |C| \text{ odd}} Q(A \cup C)}{\prod_{C \cap A = \emptyset, |C| \text{ even}} Q(A \cup C)}$$

(the same number of terms at the numerator and the denominator)

Every non-dogmatic mass function m yields a $(0, +\infty)$*-valued diffidence function, but not all such diffidence functions correspond to well-behaved mass functions*

Latent belief structure

Consider a non-dogmatic BPA such that $m = \bigcircledcap_{\emptyset \neq A \subset \Omega} A^{\delta(A)}$, where some A may have $\delta(A) > 1$.
Define two separable BFs m^+, m^- associated to m as follows: $\forall A \subset \Omega$,

- $\delta^+(A) = \min(1, \delta(A))$, and $m^+ = \bigcircledcap_{\emptyset \neq A \subset \Omega} A^{\delta^+(A)}$
- $\delta^-(A) = \min(1, 1/\delta(A))$, and $m^- = \bigcircledcap_{\emptyset \neq A \subset \Omega} A^{\delta^-(A)}$

A non-dogmatic BF can be decomposed in a unique irredundant way as a pair (m^+, m^-), of separable BFs called *latent belief structure*.

It is a bipolar decomposition into positive and negative parts

Example

Two overlapping focal sets on a 4-element frame

Let $\Omega = \{a, b, c, d\}$

Table: Decomposition with focal sets: *ab*, *ac*, *a* and Ω

A	m	Q
$\{a\}$	α	1
$\{a, b\}$	β	$Q(ab) = 1 - \alpha - \gamma$
$\{a, c\}$	γ	$Q(ac) = 1 - \alpha - \beta$
Ω	$1 - \alpha - \beta - \gamma$	$Q(\Omega) = 1 - \alpha - \beta - \gamma$
other subsets	0	one of the above

Example

Two overlapping focal sets on a 4-element frame

Let $\Omega = \{a, b, c, d\}$

Table: Decomposition with focal sets: *ab*, *ac*, *a* and Ω

A	m	δ
$\{a\}$	α	$\delta(\{a\}) = \frac{(1-\alpha-\gamma)(1-\alpha-\beta)}{1-\alpha-\beta-\gamma}$
$\{a, b\}$	β	$\delta(ab) = \frac{1-\alpha-\beta-\gamma}{1-\alpha-\gamma} \leq 1$
$\{a, c\}$	γ	$\delta(ac) = \frac{1-\alpha-\beta-\gamma}{1-\alpha-\beta} \leq 1$
Ω	$1-\alpha-\beta-\gamma$	$\delta(\Omega) = \frac{1}{1-\alpha-\beta-\gamma} \geq 1$
other subsets	0	1

$$m = \{ab\}^{\delta(ab)} \bigcirc\!\!\!\!\cap \{ac\}^{\delta(ac)} \bigcirc\!\!\!\!\cap \{a\}^{\delta(a)}$$

Example: Two overlapping focal sets on a 4-element frame

- consider the special case $\beta = \gamma = m(ab) = m(ac)$
- Condition of separability : $\delta(a) \leq 1$
  $\longrightarrow\ m(a)^2 + m(a)(-1+2\beta) + \beta^2 \leq 0$

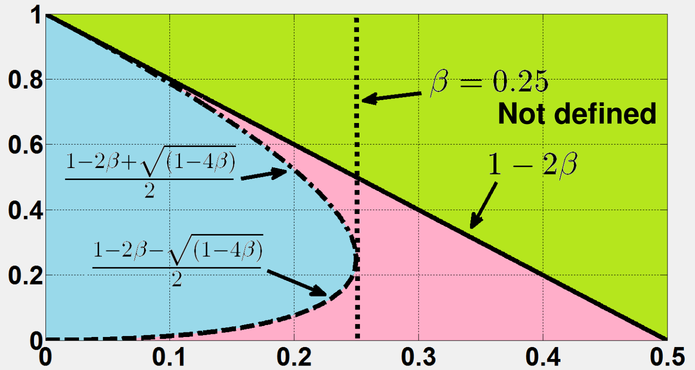

Figure: $m(a)$ in terms of $\beta = \gamma$ if $\delta(a) = 1$.

Example: Two overlapping focal sets on a 4-element frame

- consider the special case $\beta = \gamma = m(ab) = m(ac)$
- Condition of separability : $\delta(a) \leq 1$
  $\longrightarrow \ m(a)^2 + m(a)(-1 + 2\beta) + \beta^2 \leq 0$

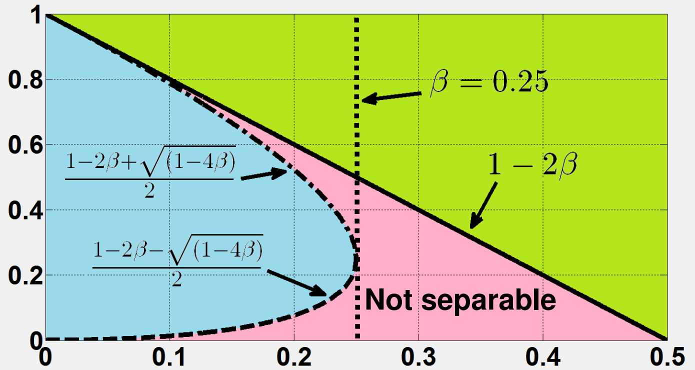

Figure: $m(a)$ in terms of $\beta = \gamma$ if $\delta(a) = 1$.

Example: Two overlapping focal sets on a 4-element frame

- consider the special case $\beta = \gamma = m(ab) = m(ac)$
- Condition of separability : $\delta(a) \leq 1$
  $\longrightarrow \; m(a)^2 + m(a)(-1 + 2\beta) + \beta^2 \leq 0$

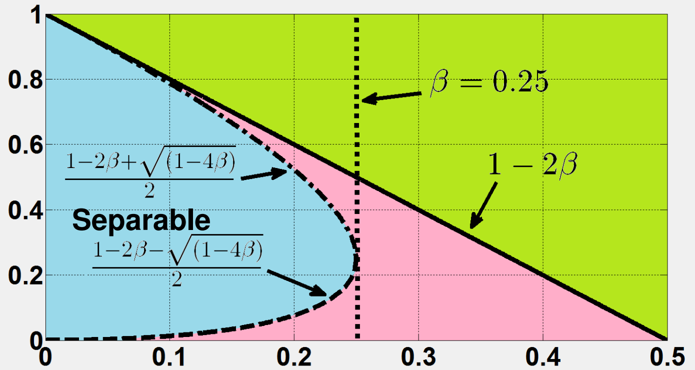

Figure: $m(a)$ in terms of $\beta = \gamma$ if $\delta(a) = 1$.

Example: Two overlapping focal sets on a 4-element frame

- Let $\beta = \gamma = m(ab) = m(ac)$

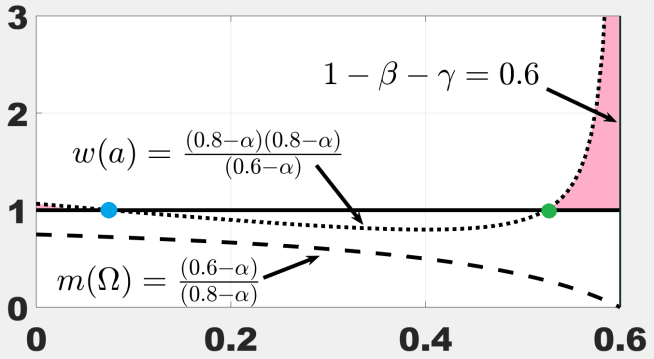

Figure: Diffidence weights in terms of $m(a)$

It may sound strange that there are two separability thresholds !

The retraction operation

Given two BPA's m_1 and m_2 the retraction operation $\oslash$ is defined such that $m = m_1 \oslash m_2$ is the solution (when it exists) of $m \cap\!\!\!\!\bigcirc\, m_2 = m_1$.

In terms of commonalities : $Q_m(A) = \frac{Q_{m_1}(A)}{Q_{m_2}(A)}, \forall A \subseteq \Omega$

- The retraction operation is the inverse of the conjunctive rule of combination
- For non-dogmatic belief functions $Q(A) \geq Q(\Omega) > 0$ so the retraction is well-defined

Warning: Q_m obtained by retraction may fail to be a commonality function, and m may fail to be valued in $[0, 1]$.

The retraction operation on latent belief structures

Proposition

If $m_i = \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{\delta_i(A)}$, $i = 1, 2$ then $m_1 \bigcirc\!\!\!\!\!\not\cap\, m_2 = \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{\frac{\delta_1(A)}{\delta_2(A)}}$

Consequences:

- $m_1 \bigcirc\!\!\!\!\!\not\cap\, m_2 = m_1 \bigcirc\!\!\!\!\cap\, m'_2$ where m'_2 has diffidence function $\delta'_2 = 1/\delta_2$.
- If a non dogmatic BF m has a latent belief structure $({\color{blue}m^+}, {\color{red}m^-})$, then we have that $m = m^+ \bigcirc\!\!\!\!\!\not\cap\, m^-$.
- Let $\mathcal{T} = \{A : \delta(A) < 1\}$ and $\mathcal{P} = \{B : \delta(B) > 1\}$. Then we have

$$
\begin{aligned}
m &= (\bigcirc\!\!\!\!\cap_{A \in \mathcal{T}} A^{\delta(A)}) \bigcirc\!\!\!\!\!\not\cap\, {\color{red}(\bigcirc\!\!\!\!\cap_{B \in \mathcal{P}} B^{1/\delta(B)})} \\
&= (\bigcirc\!\!\!\!\cap_{A \in \mathcal{T}} A^{\delta(A)}) \bigcirc\!\!\!\!\cap\, {\color{red}(\bigcirc\!\!\!\!\cap_{B \in \mathcal{P}} B^{\delta(B)})}
\end{aligned}
$$

Interpretation

A non-dogmatic belief function m with latent belief structure (m^+, m^-) can be viewed as the result of a complex merging of elementary pieces of uncertain evidence modelled by SSFs:

- A set $\mathcal{T}$ of testimonies A with low diffidence $\delta(A) < 1$ that are merged by the conjunctive rule.
- A set $\mathcal{P}$ of propositions B against which there is a prejudice of strength $\delta(B) > 1$ (strong diffidence).

each $B \in \mathcal{P}$ is the result of the fusion of testimonies in $\mathcal{T}$.

The role of $\mathcal{P}$ is to reduce the masses allocated to the combination of testimonies.

In the previous example $\mathcal{T} = \{\{a, b\}, \{a, c\}\}$, $\mathcal{P}$ contains only $\{a\}$ if $\delta(a) > 1$.

Prejudiced information fusion

Idea

The agent possessing a prejudice against believing A is ready to doubt about the truth of A whenever receiving a testimony claiming that A is true.

Retraction: the simple case

The retraction $A^v \oslash B^w$ of an SSF B^w, $w < 1$ from an SSF A^v, $v < 1$ yields the diffidence function $\delta(\cdot)$;

$$\textit{if } B \neq A,\ \delta(E) = \begin{cases} v & \text{if } E = A \\ 1/w > 1 & \text{if } E = B \text{ NOT A BF} \\ 1 & \text{otherwise.} \end{cases}$$

$$\textit{if } A = B,\ \delta(A) = v/w \text{ and } 1 \text{ otherwise.}$$

Retraction as a form of belief change

- An agent possesses a prior belief that A is false : $\overline{A}^d, d < 1$.
- He receives a testimony that A is true, with reliability $m(A) = 1 - w : A^w$

In the retraction process, the two pieces of information do not play the same role: the prior belief $\overline{A}^d$ is interpreted as a prejudice against A with diffidence $1/d > 1$

- The resulting information is $A^w \oslash A^d = A^w \cap A^{1/d} = A^{w/d}$
- Clearly after retraction, $m_{\oslash}(A) < m(A)$ (attenuation) if $w < d$, possibly resulting in ignorance ($w = d$)
- If $w > d$ (prior belief is stronger), one may interpret the remaining prejudice $A^{w/d}$ as a weaker posterior belief $\overline{A}^{d/w}$

Linda Problem after [Kahneman, Tversky 1983]

Two testimonies:

1. One (B^v) claiming that Linda is a banker
2. Another one (A^w) that she is a philanthropist.

- Dempster's rule allocates a mass $(1 - v)(1 - w)$ to the conjunction $A \cap B$: the fact that she is a philanthropist banker.
- Suppose the prejudiced receiver doesn't believe in philanthropist bankers, a zero mass is finally attached to $A \cap B$ in the resulting epistemic state of the receiver.

Linda Problem with prejudice

Two testimonies:

1. $\mathcal{T} = \{A, B\}$ and $m^+ = B^v \textcircled{\cap} A^w$: merging testimonies that say Linda is a Banker and a Philantropist
2. $\mathcal{P} = \{A \cap B\}$ and $m^- = (B \cap A)^u$ there is a prejudice of strength $1/u$ against philantropist bankers

Fusion process:

- Allocates a belief degree $(1 - v)(1 - w)$ to the conjunction $A \cap B$ (she is a philanthropist banker).
- Erode, possibly erase, this belief by retracting the SSF $(A \cap B)^u$ from the result of the fusion of testimonies: $m = (B^v \textcircled{\cap} A^w) \textcircled{\not\cap} (B \cap A)^u$.
- $\longrightarrow m(A \cap B) = 1 - \frac{(v+w-vw)}{u}$, that is all the lesser as the **prejudice** is strong (u < 1).

Retraction versus combination with the opposite information

Consider the Linda problem: Instead of a retraction $(B^v \circledcap A^w) \not\circledcap (B \cap A)^u$, consider combining the fused testimonies $(B^v \circledcap A^w)$ with the prior belief $(\overline{B \cap A})^u$:

Table: Symmetric combination with prior knowledge: $(B^v \circledcap A^w) \circledcap (\overline{B \cap A})^u$

$\downarrow (B^v \circledcap A^w) \circledcap (\overline{A \cap B})^u \rightarrow$	$\overline{A \cap B} : 1 - u$	$\Omega : u$
$B : (1 - v)w$	$B \cap \overline{A} : (1 - u)(1 - v)w$	$B : u(1 - v)w$
$A : v(1 - w)$	$A \cap \overline{B} : (1 - u)(1 - w)v$	$A : uv(1 - w)$
$A \cap B : (1 - v)(1 - w)$	$\emptyset : (1 - v)(1 - w)(1 - u)$	$A \cap B : (1 - v)(1 - w)u$
$\Omega : vw$	$\overline{A \cap B} : vw(1 - u)$	$\Omega : uvw$

If u, v, w are small, we get a strong contradiction, and a very confusing result: we do not get rid of $A \cap B$, while retraction can delete it, leaving only focal sets A, B.

Here, prior knowledge and input information play the same role

Retraction versus combination with the opposite information

In the retraction case:

Table: Retraction of prior knowledge: $(B^v \bigcirc\!\!\!\!\cap A^w) \bigcirc\!\!\!\!\not\cap (B \cap A)^u$

$\downarrow (B^v \bigcirc\!\!\!\!\cap A^w) \bigcirc\!\!\!\!\cap (A \cap B)^{1/u} \rightarrow$	$A \cap B : 1 - 1/u$	$\Omega : 1/u$
$B : (1-v)w$	$A \cap B : (1 - 1/u)(1-v)w$	$B : (1-v)w/u$
$A : v(1-w)$	$A \cap B : (1 - 1/u)(1-w)v$	$A : v(1-w)/u$
$A \cap B : (1-v)(1-w)$	$A \cap B : (1-v)(1-w)(1-1/u)$	$A \cap B : (1-v)(1-w)/u$
$\Omega : vw$	$A \cap B : vw(1 - 1/u)$	$\Omega : vw/u$

We get
$m(A \cap B) = 1 - 1/u + (1-v)(1-w)/u = 1 - (v + w - vw)/u$,
no mass on $B \cap \overline{A}, A \cap \overline{B}, \emptyset, \overline{A \cap B}$.
Here, prior knowledge and input information DO NOT play the same role

Retracting an SSF from a belief function

Retracting an SSF B^y from m
= conjunctive combination $m' = m \cap B^z$ with $z = 1/y > 1$.

$$\forall E \subseteq \Omega, m'(E) = zm(E) + (1-z) \sum_{A: A \cap B = E} m(A)$$
$$= zm(E) + (1-z) m(E|_{nn} B)$$

Proposition

Let a non-dogmatic BPA m with focal sets family $\mathcal{F}$.

- If $B \notin \mathcal{F}$,
- or if $B \in \mathcal{F}$ such that B neither contains nor is contained in another focal set $\neq \Omega$,

then, if $y < 1$, $m' = m \varnothing B^y$ is not a belief function.

Retracting an SSF from a separable belief function

If m is separable,
let $\mathcal{T} = \{A_1, \ldots, A_n\}$ and $\mathcal{F} = \{A_I = \cap_{i \in I} A_i, I \subseteq [n]\}$.
Suppose we retract A_I

Proposition

Let m be a separable BPA with focal sets family $\mathcal{F}$. Retracting focal set A_J from m :

- decreases its mass and may delete it.
- affects and may delete all focal sets $A_I \subset A_J$ $(J \subset I)$ as well,
- the focal sets $A_I \not\subset A_J$ retain a positive mass $m'(E_I) \leq 1$

Retraction versus discounting

Discounting

- Discounting procedure reduces the masses bearing on all $A \neq \Omega$ with a factor $\alpha \in [0, 1]$ and increases the weight on Ω accordingly,

$$m_\alpha(A) = \begin{cases} \alpha \cdot m(A), & \text{if } A \neq \Omega, \\ (1-\alpha) + \alpha \cdot m(\Omega), & \text{otherwise.} \end{cases}$$

- For SSFs: $A^v : m_\alpha(A) = A^{(1-\alpha)+\alpha.v} = A^v \oslash\!\!\!\bigcirc A^{1/w}$ provided that $0 < \alpha = \frac{1-wv}{1-v} < 1$

- We can discount a single focal set via retraction
- Discounting affects all focal sets different from Ω to the same extent.

The orthogonal rule as conjunctive combination + retraction

Claim (Denœux, 2009)

Dempster's rule of combination comes down to retracting the empty set from the result of the conjunctive combination rule.

Proposition Let $m_{\bigcirc\!\!\!\!\cap} = m_1 \bigcirc\!\!\!\!\cap\, m_2$ obtained as the conjunctive combination of two non-dogmatic belief functions m_1 and m_2 such that $m_{\bigcirc\!\!\!\!\cap}(\emptyset) > 0$.
Then their orthogonal sum $m_\oplus = m_1 \oplus m_2$ is of the form

$$m_{\bigcirc\!\!\!\!\cap} \;\bigcirc\!\!\!\!\not\cap\; \emptyset^d$$

where $d = 1 - m_{\bigcirc\!\!\!\!\cap}(\emptyset)$.
Normalisation is a form of retraction (prejudice against the contradiction).

Prejudiced information fusion

- It is natural to consider that information we receive from the outside is challenged by our prior information taking the form of stereotypes, or prejudices that one is often unaware of.
- So we can consider that any BF comes from merging unreliable elementary testimonies, with prejudices that weaken the weights pertaining to the conjunction of information items coming from sources.
- The receiver is reluctant to consider the result of such conjunctions valid.

The diffidence ordering (Denœux)

$$m_1 \sqsubseteq_w m_2 \iff \delta_1(A) \leq \delta_2(A), \forall A \subseteq \Omega$$

(m_1 is at least as informed as m_2)

- stronger than all other information orderings (especially specialization)
- stronger than Dempsterian specialisation
  $m_1 \sqsubseteq_d m_2$ iff there exists a BPA m such that $m_1 = m \circledcap m_2$.
- $m_1 \sqsubseteq_w m_2$ iff there exists a separable BPA m such that $m_1 = m \circledcap m_2$.
- $m_1 \sqsubseteq_x m_2$ and $m_3 \sqsubseteq_x m_4$ imply $m_1 \circledcap m_3 \sqsubseteq_x m_2 \circledcap m_4$, for $x = d, w$.

Meaning of the diffidence ordering

- If m_1 and m_2 are separable, $m_i = \bigcap_{A_i \in \mathcal{T}_i} A_i^{\delta_i(A)}$, where $\delta_i(A) < 1$ if $A \in \mathcal{T}_i$. Then

$$m_1 \sqsubseteq_w m_2 \iff \mathcal{T}_2 \subseteq \mathcal{T}_1 \text{ and } \delta_1(A) \leq \delta_2(A), \forall A \in \mathcal{T}_1$$

The BPA m_2 results from merging a subset of the pieces of information whose merging yields m_1, and grants less confidence in these pieces of information than m_1

- As a consequence
  - Even if $A \subset B$, $A^d \not\sqsubseteq_w B^{d'}$ even if $d \leq d'$ because $\{A\} \not\subseteq \{B\}$
  - At the limit, $A \subset B$ does not imply $A \sqsubset_w B$, while $A \sqsubset_d B$. A is not more w-specific than B: A paradox ?

Meaning of the diffidence ordering: general case

Non dogmatic m_i's with latent belief structure $L_i = (m_i^+, m_i^-)$ where $m_i^+ = \bigcirc\!\!\!\!\cap_{A_i \in \mathcal{T}_i} A_i^{\delta_i^+(A)}$ and $m_i^- = \bigcirc\!\!\!\!\cap_{B_i \in \mathcal{P}_i} B_i^{\delta_i^-(B)}$.

$$m_1 \sqsubseteq_w m_2 \iff \begin{cases} \mathcal{T}_2 \subseteq \mathcal{T}_1 \\ \delta_1^+(A) \leq \delta_2^+(A), \forall A \in \mathcal{T}_1 \\ \text{but also } \mathcal{P}_1 \subseteq \mathcal{P}_2 \\ \delta_1^-(A) \geq \delta_2^-(A), \forall A \in \mathcal{P}_2 \end{cases}$$

m_1 results from more testimonies and less prejudices than m_2, testimonies common to m_1 and m_2 are less reliable for m_2, and prejudices common to m_1 and m_2 are less strong for m_1.

Meaning of the diffidence ordering: general case

$$m_1 \sqsubseteq_w m_2 \iff \begin{cases} \mathcal{T}_2 \subseteq \mathcal{T}_1 \\ \delta_1^+(A) \leq \delta_2^+(A), \forall A \in \mathcal{T}_1 \\ \text{but also } \mathcal{P}_1 \subseteq \mathcal{P}_2 \\ \delta_1^-(A) \geq \delta_2^-(A), \forall A \in \mathcal{P}_2 \end{cases}$$

1. on $\mathcal{T}_2$: $\delta_1 = \delta_1^+ \leq \delta_2 = \delta_2^+ \leq 1$
2. on $\mathcal{T}_1 \setminus (\mathcal{T}_2 \cup \mathcal{P}_1)$: $\delta_1 = \delta_1^+ \leq \delta_2 = 1$: (always true)
3. on $\mathcal{T}_1 \cap \mathcal{P}_2$: $\delta_1 < 1 < \delta_2$ (always true)
4. on $\mathcal{P}_2 \setminus (\mathcal{T}_1 \cup \mathcal{P}_1)$: $1 \leq \delta_2$ (always true)
5. on $\mathcal{P}_1$: $1 < \delta_1 \leq \delta_2$, i.e., $1 > \delta_1^- \geq \delta_2^-$

The idempotent diffidence-based combination rule

- Conjunctive rule $\Rightarrow$ independent sources of information.
- Idempotent rules: cautious when ill-known dependencies.
- The latter are not easy to express with BPA's belief or plausibility functions.

Diffidence-based combination rules (Denœux (2008))

$$\delta = \delta_1 \odot \delta_2$$

for some suitable monotonic operation on $(0, +\infty)$.

Idempotent rule: $m = m_1 \owedge m_2$ such that $\delta = \min(\delta_1, \delta_2)$

- $\min(\delta_1, \delta_2)$ yields the least informative belief function such that $m \sqsubseteq_w m_1$ and $m \sqsubseteq_w m_2$.
- In contrast, the conjunctive rule of combination is such that $\odot$ = product instead of min.

The idempotent diffidence-based combination rule

The idempotent rule sometimes coincides with the product-based conjunctive rule.

proposition

Consider two separable BPA's m_1 and m_2 with respective testimony sets $\mathcal{T}_i, i = 1, 2$. Then $m_1 \owedge m_2 = m_1 \bigcirc\!\!\!\!\cap\, m_2$ if and only if $\mathcal{T}_1 \cap \mathcal{T}_2 = \emptyset$

$m_1 \owedge m_2$ assumes that sources delivering the same information are dependent, while sources delivering distinct pieces of information are independent

The idempotent diffidence-based combination rule

General case ($\mathcal{T}_1 \cap \mathcal{T}_2 \neq \emptyset$) under the separability assumption, let

- $m_{1\backslash 2} = \textcircled{\cap}_{A \in \mathcal{T}_1 \backslash \mathcal{T}_2} A^{\delta_1(A)}$
- $m_{2\backslash 1} = \textcircled{\cap}_{A \in \mathcal{T}_2 \backslash \mathcal{T}_1} A^{\delta_2(A)}$
- $m_{12} = \textcircled{\cap}_{A \in \mathcal{T}_1 \cap \mathcal{T}_2} A^{\min(\delta_1(A), \delta_2(A))}$ (the idempotent part)

$$m_1 \textcircled{\wedge} m_2 = m_{1\backslash 2} \textcircled{\cap} m_{12} \textcircled{\cap} m_{2\backslash 1}.$$

The consonant case

- If m is consonant, it is defined by a possibility distribution $\pi(\omega_i) = \sum_{\omega_i \in E} m(E) = \pi_i$
- Comes down to the merging of consonant pieces of information $E_i^{\delta_i}$ with $E_1 \subset E_2 \subset \cdots \subset E_k$ (separable)
- if $\omega \in E_i \setminus E_{i-1}, i > 1, \pi(\omega) = \pi_i = \prod_{j=1}^{i-1} \delta_j$

If $\pi^1 = 1 > \pi^2 > \ldots \pi^{k+1} > 0$ defined as above, the diffidence function of the corresponding necessity measure can be expressed as

$$\delta(A) = \begin{cases} \dfrac{\pi^{i+1}}{\pi^i} \text{ if } A = E_i, i = 1, \ldots, k \\ 1 \text{ otherwise.} \end{cases}$$

The consonant case diffidence ordering vs. specificity

If π_1 and π_2 have nested focal sets $\mathcal{F}_1$, and $\mathcal{F}_2$:

$$\pi_1 \sqsubseteq_w \pi_2 \text{ iff } \mathcal{F}_2 \subseteq \mathcal{F}_1 \text{ and } \frac{\pi_1^{i+1}}{\pi_1^i} \leq \frac{\pi_2^{i+1}}{\pi_2^i}, \forall E_i \in \mathcal{F}_1.$$

- This is different from the specificity ordering $\pi_1 \leq \pi_2$.
- The diffidence ordering is not natural if possibility degrees are interpreted as upper probability bounds.

The consonant case : Likelihood view

- Shafer (1976) assumes that likelihood functions $P(\mathcal{D}|\omega)$, based on dataset $\mathcal{D}$ are proportional to contour functions of consonant belief functions, obtained by renormalizing the former so that its maximal value is 1.
- but comparing likelihood functions pointwisely makes no sense as they are defined up to multiplicative constants

We can only compare *likelihood ratios* relative to results of distinct data sets $\mathcal{D}_1$ and $\mathcal{D}_2$, Letting $\pi_i(\omega) = c_i P(\mathcal{D}_i|\omega)$ where c_i is a value such that $\max_{\omega\in\Omega} \pi_i(\omega) = 1$:

$$\frac{P(\mathcal{D}_1|\omega)}{P(\mathcal{D}_1|\omega')} \leq \frac{P(\mathcal{D}_2|\omega)}{P(\mathcal{D}_2|\omega')} \iff \pi_1 \sqsubseteq_w \pi_2$$

provided that $P(\mathcal{D}_1|\omega)$ and $P(\mathcal{D}_2|\omega)$ induce the same orderings.

Conclusion

- We revisit the decomposition of a BFs by Smets showing that it can be viewed as the merging of uncertain testimonies, of prejudices that resist against the results of their partial conjunctions.
- This is an approach to BF based on diffidence functions and the merging of pieces of evidence, as opposed to the approach based on upper and lower probability.
- We have studied the retraction operation as a belief change tool
- we have shown the meaning of information ordering and idempotent combination based on diffidence function

Future research

- Study the general retraction operation $m_1 \oslash m_2$.
- Consider other decompositions of belief functions, dropping the independence assumptions like in Pichon (2019).

References

- Denœux, T., 2008. Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. Artificial Intelligence 172, 234-264.
- Dubois, D., Faux, F., Prade, H., 2018. Prejudiced information fusion using belief functions, in Proc. BELIEF 2018, Compiegne, Springer, pp 77-85
- Pichon, F., 2018. Canonical decomposition of belief functions based on Teugels representation of the multivariate bernoulli distribution. Inf. Sci. 428, 76-104.
- Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton Univ. Press.
- Smets, P., 1995. The canonical decomposition of a weighted belief, in: Proc. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI), Montreal, Aug. 20-25, pp. 1896-1901.