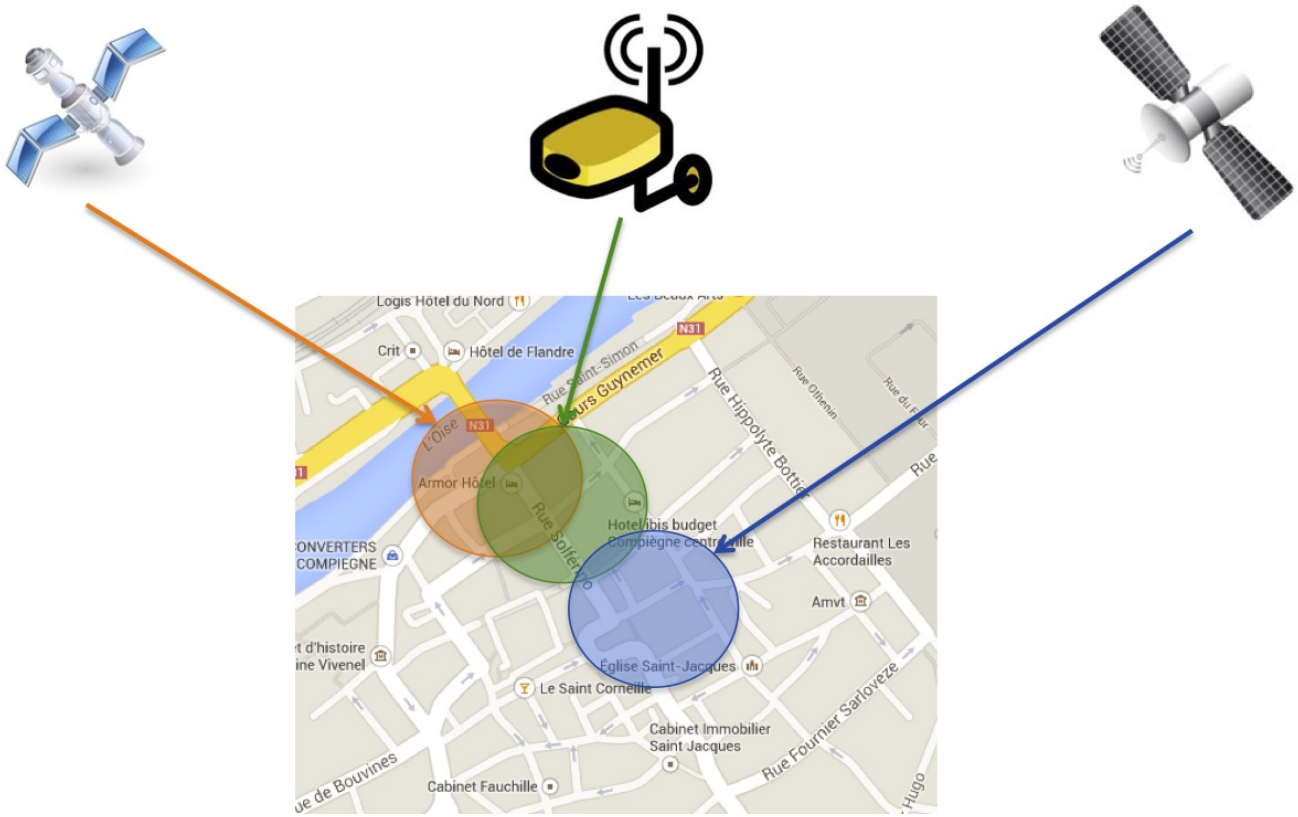
On combination and conflict

Sebastien Destercke

Heuristic and Diagnosis for Complex Systems (HEUDIASYC) laboratory,
Compiegne, France

BFAS School

An illustration of the issue



outline

Goals of the lecture:

- provide warnings about potential misunderstandings
- give some very basic elements on combination
- give answers to "how to measure conflict"?
- go a bit further than basics of combination

References will be given to go further on.

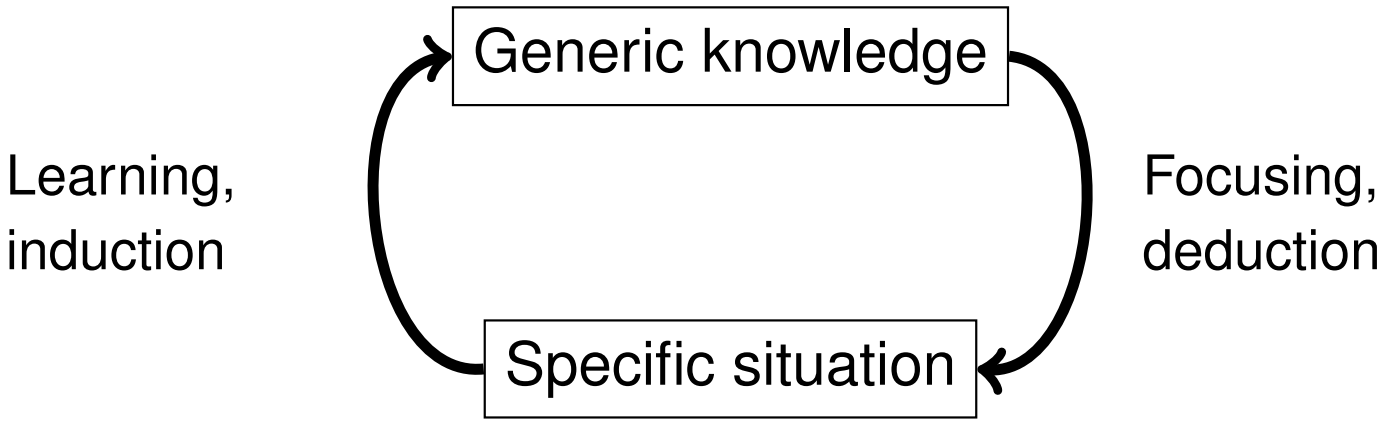
Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Basic definition



- Focusing: particularizing generic knowledge to given situation
- Revising: integrating new information on same level, often with priority to new information
- Combining: integrating multiple information on the same level

Quick reminder

- information modelled by mass function $m = 2^{|\Omega|} \rightarrow [0, 1]$
- $\mathcal{F}$: set of focal elements of m
- induces $Bel(A) = \sum_{E \subseteq A, E \neq \emptyset} m(E)$ and $Pl(A) = \sum_{E \cap A \neq \emptyset} m(E)$
- contour function $pl : \Omega \rightarrow [0, 1]$ has values $pl(\omega) = Pl(\{\omega\})$

Interpretations

- **singular**: $m =$ uncertainty about quantity $\mathcal{W}$ with single true value ω_0 (TBM, evidence theory)
- **statistical**: $m =$ frequencies of imprecise observations (Dempster)
- **robust probabilistic**: Bel/Pl : bounds of ill-known probability

Focusing: example

- Information: means of travel (water, air, road) with frame

$$\Omega = \{Car, Train, Plane, Helicopter, Boat\}$$

$$m(\{C, T\}) = 0.5, \quad m(\{P, H\}) = 0.4, \quad m(\{B\}) = 0.1$$

- Prakash lives in the U.S. $\rightarrow A = \{plane, boat\}$. Focused information is

$$Bel_{||A}(P) = 0, Pl_{||A}(P) = 4/5$$

Because 0.1 ($m(\{B\})$) of 0.5 ($m(\{B\}) + m(\{P, H\})$) mass must remain on Boat.

Focusing: example

- Information: means of travel (water, air, road) with frame

$$\Omega = \{Car, Train, Plane, Helicopter, Boat\}$$

$$m(\{C, T\}) = 0.5, \quad m(\{P, H\}) = 0.4, \quad m(\{B\}) = 0.1$$

- Sebastien lives in the North of France $\rightarrow A = \{car, train\}$. Focused information is

$$Bel_{train||A}(T) = 0, Pl_{||A}(T) = 1$$

Because all mass can be transferred inside/outside of train.

Focusing: some elements

- Conditional knowledge should be **consistent** with initial one
- No sense in performing sequential focusing $\rightarrow$ condition each time on generic knowledge
- Known under the name "regular extension" and "natural extension" in imprecise probabilities

Revising: example

- Information: means of travel (water, air, road) with frame

$$\Omega = \{Car, Train, Plane, Helicopter, Boat\}$$

$$m(\{C, T\}) = 0.5, \quad m(\{P, H\}) = 0.4, \quad m(\{B\}) = 0.1$$

- For some reasons, road vehicles and helicopter are not available to the population, $A = \{plane, boat\}$ and $Bel(A) = 1$. Revised info is

$$Bel_{||A}(P) = Pl_{||A}(P) = 4/5$$

Because Dempster's rule of conditioning results in a probability.

Revising: example

- Information: means of travel (water, air, road) with frame

$$\Omega = \{Car, Train, Plane, Helicopter, Boat\}$$

$$m(\{C, T\}) = 0.5, \quad m(\{P, H\}) = 0.4, \quad m(\{B\}) = 0.1$$

- We now learn that only road vehicles are used, then $\rightarrow B = \{car, train\}$. Since

$$A \cap B = \emptyset,$$

our knowledge become inconsistent $\rightarrow$ eventual need for repair.

Revising: some elements

- Revised model do not need to be fully **consistent** with initial knowledge (we may learn something contradicting what we know)
- It makes sense to revise iteratively, unless information are too inconsistent with each others
- Revised information can be required to be fully consistent with new one $\rightarrow$ Jeffrey's rule and extension, probability kinematics

If you want to know more I

[1] Didier Dubois and Thierry Denœux.
Conditioning in dempster-shafer theory: Prediction vs. revision.
In *Belief Functions: Theory and Applications*, pages 385–392. Springer, 2012.

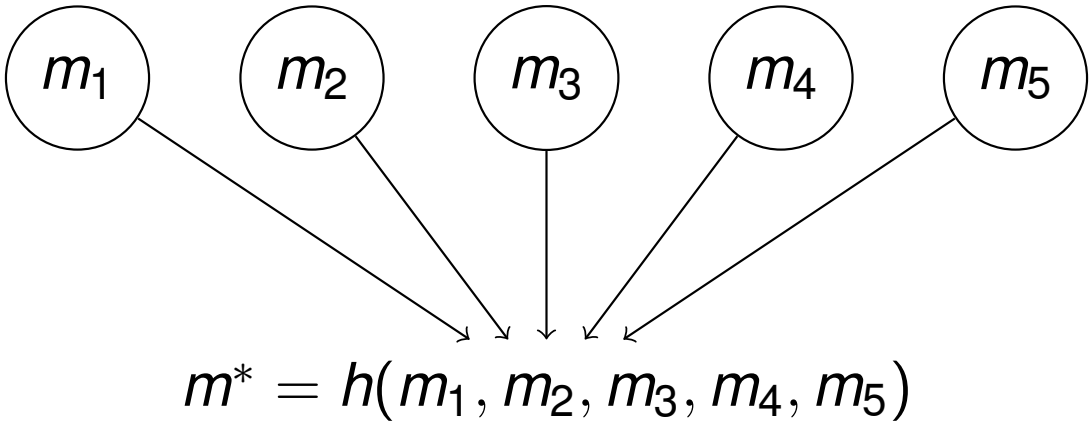
[2] Didier Dubois and Henri Prade.
Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge.
In *Qualitative and quantitative practical reasoning*, pages 96–107. Springer, 1997.

[3] Joseph Y Halpern and Ronald Fagin.
Two views of belief: belief as generalized probability and belief as evidence.
Artificial intelligence, 54(3):275–317, 1992.

[4] Jianbing Ma, Weiru Liu, Didier Dubois, and Henri Prade.
Bridging jeffrey's rule, agm revision and dempster conditioning in the theory of evidence.
International Journal on Artificial Intelligence Tools, 20(04):691–720, 2011.

[5] Enrique Miranda and Ignacio Montes.
Coherent updating of non-additive measures.
International Journal of Approximate Reasoning, 56:159–177, 2015.

Combining or information fusion



- Information on the same level
- No piece of information has priority over the other (a priori)
- Makes sense to combine multiple pieces of information at once
- Rest of the lecture: (mainly) what about h?

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Basics

(Unnormalized) Demspter's rule assume sources to be:

- Completely trustful, relevant, **reliable**, . . . (see David's lecture for differences)
  - Conjunction of focal elements
- Independent
  - Product of masses

If assumptions are not likely to hold, then Dempster's rule **should not be used!**

Conjunctive rule

	$m_1(A_1)$ ...	$m_1(A_i)$	... $m_1(A_m)$
$m_2(B_1)$ $\vdots$			
$m_2(B_j)$		$A_i \cap B_j$	
$\vdots$ $m_2(B_n)$			

- All reliable

Conjunctive rule

	$m_1(A_1)$ ...	$m_1(A_i)$	... $m_1(A_m)$
$m_2(B_1)$ $\vdots$			
$m_2(B_j)$		$m(A_i \cap B_j) =$ $m(A_i)m(B_j)$	
$\vdots$ $m_2(B_n)$			

- All reliable
- Independent

Revisiting Zadeh's (infamous) example

Two doctors, m_1 and m_2, stating about whether a patient has **M**eningitis, **C**oncussion or **B**rain tumor

	$m_1(M) = 0.9$	$m_1(B) = 0.1$
$m_2(C) = 0.9$	$m(\emptyset) = 0.81$	$m(\emptyset) = 0.09$
$m_2(B) = 0.1$	$m(\emptyset) = 0.09$	$m(B) = 0.01$

- If Dempster's assumptions hold, then conclusion (Brain tumor) is legitimate.
- What we can question are whether the assumptions hold, not the rule itself!

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

When?

Dempster's rule does not apply when:

- **Sources are not independent**
- Sources are not reliable

Conjunctive combination without independence

m_1, m_2 and with focal elements $\mathcal{F}_1, \mathcal{F}_2$. The result of conjunctively merging m_1, m_2 is a bba m obtained in 2 steps:

1. Define a joint bba m_{12} s.t.

$$m_1(A) = \sum_{B \in \mathcal{F}_2} m_{12}(A, B) \; \forall A$$

   and likewise for m_2 (**Marginal preservation**).

2. $m_{12}(A, B)$ is allocated to, and only to $A \cap B$ (**Conjunctive allocation**)

$\mathcal{M}_{12}$: set of conjunctively merged bbas m.

To a specific joint m_{12} corresponds a dependence structure.

(Unnormalized) Dempster's rule: $m_{1 \oplus 2}(A, B) = m_1(A) m_2(B)$

Illustration

	$m_1(A_1)$ ...	$m_1(A_i)$	... $m_1(A_m)$	
$m_2(B_1)$ $\vdots$				
$m_2(B_j)$		$m_{12}(A_i, B_j) =$ $m_{12}(A_i \cap B_j)$		$\sum_j m_{12}(A_i, B_j)$ $= m_2(B_j)$
$\vdots$ $m_2(B_n)$				
		$\sum_i m_{12}(A_i, B_j)$ $m_1(A_i)$		

Disjunctive rule extends in the same way (replacing $\cap$ by $\cup$)

When?

Dempster's rule does not apply when:

- Sources are not independent
- **Sources are not reliable**

3 basics combination rule

- Conjunctive ($\cap$): all sources are reliable

$$m_{\cap}(C) = \sum_{A,B:A\cap B=C} m_1(A)m_2(B)$$

- Disjunctive ($\cup$): at least one source is reliable

$$m_{\cup}(C) = \sum_{A,B:A\cup B=C} m_1(A)m_2(B)$$

- Weighted average ($\sum$): sources have reliabilities w_i/counting procedure

$$m_{\sum}(A) = w_1 m_1(A) + w_2 m_2(A)$$

Disjunctive rule

	$m_1(A_1)$ ...	$m_1(A_i)$	... $m_1(A_m)$
$m_2(B_1)$ $\vdots$			
$m_2(B_j)$		$A_i \cup B_j$	
$\vdots$ $m_2(B_n)$			

- At least one reliable

Disjunctive rule

	$m_1(A_1)$ ...	$m_1(A_i)$	... $m_1(A_m)$
$m_2(B_1)$ $\vdots$			
$m_2(B_j)$		$m(A_i \cup B_j) =$ $m(A_i)m(B_j)$	
$\vdots$ $m_2(B_n)$			

- At least one reliable
- Independent

Disjunction example

	$m_1(M) = 0.9$	$m_1(B) = 0.1$
$m_2(C) = 0.9$	$m(M, C) = 0.81$	$m(B, C) = 0.09$
$m_2(B) = 0.1$	$m(B, M) = 0.09$	$m(B) = 0.01$

Average example

$$m^*(M) = 1/2(m_1(M) + m_2(M)) = 0.45$$

$$m^*(C) = 1/2(m_1(C) + m_2(C)) = 0.45$$

$$m^*(B) = 1/2(m_1(B) + m_2(B)) = 0.1$$

Before getting deeper in combination

- If sources are not conflicting (are consistent), then conjunctive approach gives good result (use it!)
- If they are conflicting $\rightarrow$ clues that they may not be reliable $\rightarrow$ need another rule (unless reliability is certain)
- **First step: how to detect that they are conflicting, and why are they conflicting?**
- Second step: how to select an alternative rule?

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Classical measure of conflict

- Inconsistency $\phi(m)$ of m given by $\phi(m) = m(\emptyset)$
- Conflict between m_1 and m_2: $m_{1\oplus 2}(\emptyset)$ resulting from Unnormalized Dempster's rule
- Common critic: $m(\emptyset) \neq 0$ even if $m_1 = m_2$
- Example: two equally ambiguous physicians

	$m_1(M) = 0.5$	$m_1(B) = 0.5$
$m_2(M) = 0.5$	$m(M) = 0.25$	$m(\emptyset) = 0.25$
$m_2(B) = 0.5$	$m(\emptyset) = 0.25$	$m(B) = 0.25$

Two follow-up questions

- Is the case $m_1 = m_2$ problematic, and if yes how to address it?
- Can we justify $m(\emptyset)$ as a conflict measure based on desirable properties/axioms?

Concerning the first question:

1. Lots of authors consider that yes, it is
2. A common proposal is to also measure conflict through distance

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Original proposal

In her 2006 paper, Liu [10] proposes to use two measures:

- the value $m_{1\oplus 2}(\emptyset)$
- the difference $DifBet = \sum_A |BetP_1(A) - BetP_2(A)|$
- say that m_1, m_2 are conflicting if
  - $m_{1\oplus 2}(\emptyset)$ high (e.g. ≥ 0.8)
  - $m_{1\oplus 2}(\emptyset)$ low and *DifBet* high

  and if conflicting, do not apply Dempster's rule

Further consideration

- consider in general distance $d(m_1, m_2)$
- measure conflict by $< m_{1\oplus 2}(\emptyset), d(m_1, m_2) >$
- apply Liu's ideas
- relations between distance studied by Jousselme and Maupin [9]

Example

	$m_1(M) = 0.5$	$m_1(B) = 0.5$
$m_2(M) = 0.5$	$m(M) = 0.25$	$m(\emptyset) = 0.25$
$m_2(B) = 0.5$	$m(\emptyset) = 0.25$	$m(B) = 0.25$

- $m_{1\oplus 2}(\emptyset) = 0.5$
- $DifBet = 0$
- Recommended conclusion: Dempster can be applied

Two follow-up questions

- Is the case $m_1 = m_2$ problematic, and if yes how to address it?
- Can we justify $m(\emptyset)$ as a conflict measure based on desirable properties/axioms?

Let us address the second question by:

1. Study inconsistency from sets, extend it to belief functions
2. Study conflict between sets, extend it to belief functions

As a side result, we will see that $m_1 = m_2$ is not (necessarily) problematic

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Set inconsistency

A quantity $\mathcal{W}$ with value $\omega \in \Omega$

Source information: $\omega \in A$

- A consistent $\Leftrightarrow A \neq \emptyset$
- A inconsistent $\Leftrightarrow A = \emptyset$

Measure of inconsistency

ϕ measures information (A) inconsistency

Property 1: boundedness

ϕ should be bounded $\rightarrow \phi(A) \in [0, 1]$

Property 2: maximal values

ϕ maximal iff information inconsistent, minimum iff information consistent

$$\phi(A) = \begin{cases} 1 & \text{if} \quad A = \emptyset \\ 0 & \text{if} \quad A \neq \emptyset \end{cases}$$

Consistent and inconsistent m

(totally) inconsistent m

m inconsistent if $m(\emptyset) = 1$ (empty mass)

(totally) consistent m

Strong version

m consistent $\Leftrightarrow \cap_{E \in \mathcal{F}} \neq \emptyset$

$\cap_{E \in \mathcal{F}} \neq \emptyset \Leftrightarrow \exists \omega$ s.t. $pl(\omega) = 1$

One state of Ω fully plausible

$\rightarrow$ agree with **singular** int.

Consistent and inconsistent m

(totally) inconsistent m

m inconsistent if $m(\emptyset) = 1$ (empty mass)

(totally) consistent m

Weak version

m consistent $\Leftrightarrow m(\emptyset) = 0$

Usual definition

$\rightarrow$ closer to **statistical/probabilitic** int.

Consistent and inconsistent m

(totally) inconsistent m

m inconsistent if $m(\emptyset) = 1$ (empty mass)

(totally) consistent m

Strong version

m consistent $\Leftrightarrow \cap_{E \in \mathcal{F}} \neq \emptyset$

$\cap_{E \in \mathcal{F}} \neq \emptyset \Leftrightarrow \exists \omega$ s.t. $pl(\omega) = 1$

One state of Ω fully plausible

$\rightarrow$ agree with **singular** int.

Weak version

m consistent $\Leftrightarrow m(\emptyset) = 0$

Usual definition

$\rightarrow$ closer to **statistical/probabilitic** int.

Strong consistency $\Rightarrow$ Weak consistency

N.B.: if m consonant, definitions coincide.

Measure of inconsistency for m

Property 1: boundedness

ϕ should be bounded $\rightarrow \phi(m) \in [0,1]$

Property 2: maximal values

ϕ maximal iff m totally inconsistent, minimum iff m consistent

ϕ depends on consistency definition

Strong version

$$\phi_{pl}(m) = 1 - \max_{\omega \in \Omega} pl(\omega)$$

Contour function based.

Weak version

$$\phi_m(m) = m(\emptyset)$$

Usual definition

Exercice 1

prove that for any m, $\phi_{pl}(m) \geq \phi_m(m)$

Conflict between sets: properties

Property 1: boundedness

$\kappa(A, B)$ should be bounded

Property 2: maximal values

$\kappa(A, B)$ maximal iff $A \cap B$ inconsistent, minimum iff $A \cap B$ consistent

Property 3: symmetry

$\kappa(A, B) = \kappa(B, A)$

Property 4: imprecision monotonicity

if $A \subseteq A'$, then $\kappa(A, B) \geq \kappa(A', B)$

$A = \{\omega_1, \omega_2\}, B = \{\omega_3, \omega_4\} \rightarrow \kappa(A, B) = \phi(A \cap B) = 1$
$A' = \{\omega_1, \omega_2, \omega_3\}, B = \{\omega_3, \omega_4\} \rightarrow \kappa(A', B) = \phi(A' \cap B) = 0$

Conflict between sets: properties

Property 5: ignorance is bliss

if $A = \Omega$, $\kappa(A, B) = \phi(B)$

Property 6: resistance to refinement

$\rho : 2^{|\Omega|} \rightarrow 2^{|\Theta|}$ refinement from Ω to Θ.
Then $\kappa(A, B) = \kappa(\rho(A), \rho(B))$

$A = \{red, blue\}, B = \{red, green\}$, $\kappa(A, B) = 0$

$\rho(A) = \{\text{light red}, \text{dark red}, blue\}, \rho(B) = \{\text{light red}, \text{dark red}, green\}$,
$\kappa(\rho(A), \rho(B)) = 0$

Properties extended

Property 1: boundedness

$\kappa(m_1, m_2)$ should be bounded

Property 2: maximal values

$\kappa(m_1, m_2)$ maximal iff m_1, m_2 totally inconsistent, minimum iff m_1, m_2 totally consistent

Consistent and inconsistent m_1, m_2

(totally) inconsistent m_1, m_2

Let $D_i = \cup_{E \in \mathcal{F}_i}$, then m_1, m_2 totally inconsistent if $D_1 \cap D_2 = \emptyset$

(totally) consistent m_1, m_2

Strong version
m_1, m_2 consistent if

$\cap_{E \in \{\mathcal{F}_1 \cup \mathcal{F}_2\}} E \neq \emptyset \Leftrightarrow \exists \omega$ s.t. for all $m \in \mathcal{M}_{12}$, $pl(\omega) = 1$

Common state fully plausible for both sources

$\rightarrow$ agree with **singular** int.

Consistent and inconsistent m_1, m_2

(totally) inconsistent m_1, m_2

Let $D_i = \cup_{E \in \mathcal{F}_i}$, then m_1, m_2 totally inconsistent if $D_1 \cap D_2 = \emptyset$

(totally) consistent m_1, m_2

Weak version
m_1, m_2 consistent if

for any $A \in \mathcal{F}_1, B \in \mathcal{F}_2$, $A \cap B \neq \emptyset$
$\Leftrightarrow$ for all $m \in \mathcal{M}_{12}$, $m(\emptyset) = 0$

Usual definition.

$\rightarrow$ closer to
statistical/probabilistic int.

Consistent and inconsistent m_1, m_2

(totally) inconsistent m_1, m_2

Let $D_i = \cup_{E \in \mathcal{F}_i}$, then m_1, m_2 totally inconsistent if $D_1 \cap D_2 = \emptyset$

(totally) consistent m_1, m_2

Strong version

m_1, m_2 consistent if

$\cap_{E \in \{\mathcal{F}_1 \cup \mathcal{F}_2\}} E \neq \emptyset \Leftrightarrow \exists \omega$ s.t. for all $m \in \mathcal{M}_{12}$, $pl(\omega) = 1$

Common state fully plausible for both sources

$\rightarrow$ agree with **singular** int.

Weak version

m_1, m_2 consistent if

for any $A \in \mathcal{F}_1, B \in \mathcal{F}_2$, $A \cap B \neq \emptyset$
$\Leftrightarrow$ for all $m \in \mathcal{M}_{12}$, $m(\emptyset) = 0$

Usual definition.

$\rightarrow$ closer to **statistical/probabilistic** int.

Strong consistency $\Rightarrow$ Weak consistency

Inclusion between BF

Here, we consider s-inclusion: if m_1 is a specialization of m_2, we note $m_1 \sqsubseteq_s m_2$

If m_1, m_2 are weight vectors, then bba m_1 is a specialization of bba m_2 if $\exists$ a stochastic matrix S s.t.

- $m_1 = S \cdot m_2$
- $S_{ij} > 0 \Rightarrow A_i \subseteq B_j$
- $m_2(A)$ "flow downs" to subsets of A in m_1

Properties extended (2)

Property 3: symmetry

$\kappa(m_1, m_2) = \kappa(m_2, m_1)$

Property 4: imprecision monotonicity

if $m_1 \sqsubseteq_s m'_1$, then $\kappa(m_1, m_2) \geq \kappa(m'_1, m_2)$

Property 5: ignorance is bliss

if $m_1(\Omega) = 1$, $\kappa(m_1, m_2) = \phi(m_2)$

Property 6: resistance to refinement

$\rho : 2^{|\Omega|} \rightarrow 2^{|\Theta|}$ refinement from Ω to Θ.
Then $\kappa(m_1, m_2) = \kappa(\rho(m_1), \rho(m_2))$

$\rho(m)$ refine every focal element of m with $\rho \rightarrow m(\rho(A))$

Conflict between BF: new property

Property/requirement 7: independence to dependence

Conflict should not be measured according to some unverified (in)dependence assumption between sources

Two solutions if dependence structure ill-known:

- Consider sets of possible dependencies.
- Consider the least-commitment principle.

Measuring conflict

Strong version:

- (In)dependence unknown

$$\kappa_{pl}(m_1, m_2) = [\min_{m_{12} \in \mathcal{M}_{12}} \phi_{pl}(m_{12}), \max_{m_{12} \in \mathcal{M}_{12}} \phi_{pl}(m_{12})]$$
$$= [\min_{m_{12} \in \mathcal{M}_{12}} 1 - \max_{\omega \in \Omega} pl_{12}(\omega), \max_{m_{12} \in \mathcal{M}_{12}} 1 - \max_{\omega \in \Omega} pl_{12}(\omega)];$$

- dependence known $\kappa_{pl}(m_1, m_2) = \phi_{pl}(m_{12}) = 1 - \max_{\omega \in \Omega} pl_{12}(\omega)$

Weak version:

- (In)dependence unknown

$$\kappa_{m}(m_1, m_2) = [\min_{m_{12} \in \mathcal{M}_{12}} \phi_{m}(m_{12}), \max_{m_{12} \in \mathcal{M}_{12}} \phi_{pl}(m_{12})]$$
$$= [\min_{m_{12} \in \mathcal{M}_{12}} m_{12}(\emptyset), \max_{m_{12} \in \mathcal{M}_{12}} m_{12}(\emptyset)];$$

- dependence known $\kappa_{m}(m_1, m_2) = \phi_{pl}(m_{12}) = m_{12}(\emptyset)$

If independence, $\kappa_{m}(m_1, m_2)$ usual measure.

What about $m_1 = m_2$?

If $m_1 = m_2$ and if we apply the least commitment principle, then least conflicting case is given by $m_{12} = m_1 = m_2$ and

- $\min_{m_{12} \in \mathcal{M}_{12}} 1 - \max_{\omega \in \Omega} pl_{12}(\omega) = 1 - \max_{\omega \in \Omega} pl_i(\omega)$
- $\min_{m_{12} \in \mathcal{M}_{12}} m_{12}(\emptyset) = m_i(\emptyset)$

Conflict increases only if we are sure that m_1, m_2 are not completely dependent.

Detecting the source of conflict

	$m_1(M) = 0.5$	$m_1(B) = 0.5$
$m_2(M) = 0.5$	M	$\emptyset$
$m_2(B) = 0.5$	$\emptyset$	B

- $m_{1\oplus 2}(\emptyset) = 0.5$
- $\kappa_m = [0, 1]$
- Conflict highly depends on source dependence ($\rightarrow$ question independency assumption?)

Detecting the source of conflict

	$m_1(C) = 0.9$	$m_1(B) = 0.1$
$m_2(M) = 0.9$	M	$\emptyset$
$m_2(B) = 0.1$	$\emptyset$	B

- $m_{1\oplus 2}(\emptyset) = 0.99$
- $\kappa_m = [0.9, 1]$
- Dependence plays a low role in conflict ($\rightarrow$ question reliability assumption?)

Exercice 2

$$m_1(\{\omega_1,\omega_2\}) = 0.6 \quad m_1(\{\omega_1,\omega_3\}) = 0.4$$
$$m_2(\{\omega_2,\omega_3\}) = 0.5 \quad m_2(\Omega) = 0.5.$$

	$m_1(\{\omega_1,\omega_2\})$	$m_1(\{\omega_1,\omega_3\})$	$\sum_{m_{12}}$
$m_2(\{\omega_2,\omega_3\})$			$= 0.5$
$m_2(\Omega)$			$= 0.5$
$\sum_{m_{12}}$	$= 0.6$	$= 0.4$	

- Are m_1, m_2 weakly/strongly consistent?
- What are the values of κ_m, κ_{pl}?

Exercice 3

Consider $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and the masses

$$m_1(\{\omega_1, \omega_2\}) = 0.6 \quad m_1(\{\omega_2, \omega_4\}) = 0.4$$

$$m_2(\{\omega_2, \omega_3\}) = 0.5 \quad m_2(\{\omega_3\}) = 0.5.$$

- Compute the conflicts κ_m, κ_{pl}
- Does independence appear as main cause of conflict?

Exercice 4

Consider $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and the masses

$$m_1(\{\omega_1, \omega_2\}) = 0.6 \quad m_1(\{\omega_1, \omega_4\}) = 0.4$$

$$m_2(\{\omega_2, \omega_3\}) = 0.5 \quad m_2(\{\omega_3, \omega_4\}) = 0.5.$$

- Compute the conflicts κ_m, κ_{pl}
- Does independence appear as main cause of conflict?

Exercice 5

Given two Bayesian masses m_1, m_2, demonstrate that the lower bounds of κ_m, κ_{pl} are, respectively:

- $1 - \sum_{\omega \in \Omega} \min(pl_1(\omega), pl_2(\omega))$
- $1 - \max_{\omega \in \Omega} \min(pl_1(\omega), pl_2(\omega))$

and can be obtained through the same joint m_{12}.

Other (open) topics of interest

- Can we separate inconsistency between and within a source in a principled way (see works of Daniel)?
- Can we find a characterization/representation theorem of different conflict measures (set of axioms uniquely defining them)?
- How to efficiently compute/approximate bounding conflict measures? In special cases?
- As for specialization and informational orderings, is there a natural partial order of conflicting belief functions? How can we exploit it?

Some references I

[6] Milan Daniel.
Conflicts within and between belief functions.
In *Computational Intelligence for Knowledge-Based Systems Design*, pages 696–705. Springer, 2010.

[7] Milan Daniel.
Properties of plausibility conflict of belief functions.
In *Artificial Intelligence and Soft Computing*, pages 235–246. Springer, 2013.

[8] Sebastien Destercke and Thomas Burger.
Toward an axiomatic definition of conflict between belief functions.
Cybernetics, IEEE Transactions on, 43(2):585–596, 2013.

[9] Anne-Laure Jousselme and Patrick Maupin.
Distances in evidence theory: Comprehensive survey and generalizations.
International Journal of Approximate Reasoning, 53(2):118–145, 2012.

[10] Weiru Liu.
Analyzing the degree of conflict among belief functions.
Artificial Intelligence, 170(11):909–924, 2006.

[11] Arnaud Martin.
About conflict in the theory of belief functions.
In *Belief Functions: Theory and Applications*, pages 161–168. Springer, 2012.

[12] Arnaud Roquel, Sylvie Le Hégarat-Mascle, Isabelle Bloch, and Bastien Vincke.
Decomposition of conflict as a distribution on hypotheses in the framework on belief functions.
International Journal of Approximate Reasoning, 55(5):1129–1146, 2014.

[13] Johan Schubert.
Conflict management in dempster–shafer theory using the degree of falsity.
International Journal of Approximate Reasoning, 52(3):449–460, 2011.

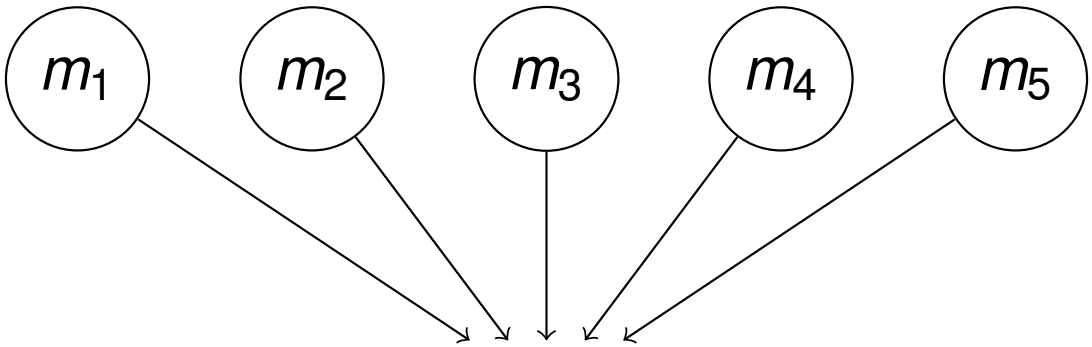
Getting deeper in combination

- If sources are not conflicting (are consistent), then conjunctive approach gives good result
- If they are conflicting $\rightarrow$ clues that they may not be reliable $\rightarrow$ need another rule
- First step: how to detect that they are conflicting, and why are they conflicting?
- **Second step: how to select an alternative rule?**

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Combining or information fusion



How to pick h?

Outline

1 Introduction
- Focusing, Revising, Combining
- A quick revisit of Dempster's rule
- When Dempster's rule does not apply

2 On conflict between belief functions
- Complementing with distances
- Axiomatic approach

3 Combination: selecting a rule
- By properties of combination
- Interpretable rules
- Learnable rules
- Other options

(im)Possibility preservation

Possibility preservation

If sources consider ω plausible, the fusion result should too

- Weak version: if $pl_i(\omega) > 0$ for all i, then $pl^*(\omega) > 0$
- Strong version: if $pl_i(\omega) > 0$ for one i, then $pl^*(\omega) > 0$

Strong version enforces disjunctive or average combination-style rules

Impossibility preservation

If sources consider ω impossible, the fusion result should too

- Weak version: if $pl_i(\omega) = 0$ for all i, then $pl^*(\omega) = 0$
- Strong version: if $pl_i(\omega) = 0$ for one i, then $pl^*(\omega) = 0$

Strong version enforces conjunctive combination-style rule

Impact of uninformative sources

Insensitivity to ignorance

If $m_i(\Omega) = 1$ is a vacuous mass, then

$$h(m_1, \ldots, m_n) = h(m_1, \ldots, m_{i-1}, m_{i+1}, \ldots, m_n)$$

Insensitivity to less informative sources

If $m_i \sqsubset m_j$, then

$$h(m_1, \ldots, m_i, \ldots, m_j, \ldots, m_n) = h(m_1, \ldots, m_i, \ldots, m_n)$$

N.B.: properties unsatisfied by disjunctive rule and average

Equalitarian properties

Commutatitivity

Let $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a permutation, then

$$h(m_1, \ldots, m_n) = h(m_{\sigma(1)}, \ldots, m_{\sigma(n)}).$$

The order of combination should not matter!

(A version of) Fairness

For any m_i, the operation

$$h(m_1, \ldots, m_n) \oplus m_i$$

should be non-vacuous.

$\to$ can be questioned in cases of many sources (isolated sources could be discarded) or of information about sources quality

Extension to n sources

Associativity

We have the property that

$$h(m_1, \ldots, m_n) = h(h(m_1, \ldots, m_{n-1}), m_n), .$$

With commutativity, makes the extension of a binary rule $h(m_1, m_2)$ to n sources straightforward

$\rightarrow$ Computational advantage (perform binary rule n times)

A summary

- Weak version of preservation common sense (strongly desirable)
- Strong versions enforces particular behaviours (disjunctive/conjunctive)
- Uninformativity insensitivity forbids too conservative attitude
- Equalitarian properties essential if no information indicate that sources can be forgotten/privileged
- Associativity useful (for practical reasons), not essential

However, selecting properties usually not sufficient to select a unique rule.

Two common situations

Situation 1: interpretable rule

- Efficiency of rules cannot be assessed
- **Interpretability of the rule important**
- example: few experts, no repeated experiments
- mostly concerns **singular interpretation**

Situation 2: learnable rule

- Efficiency of rules can be assessed
- Interpretability of the rule is a lesser issue
- example: machine learning, many sensors, repeated experiments
- mostly concerns statistical interpretation

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Basic idea

Given a set $\mathcal{A}$ of focal elements $A_1, \ldots, A_n$ with $A_i \in \mathcal{F}_i$:

- All sources are reliable:

$$A^* = A_1 \cap \ldots \cap A_n$$

- At least a source is reliable:

$$A^* = A_1 \cup \ldots \cup A_n$$

- Other solutions by replacing the logical/set operators combining $A_1, \ldots, A_n$ by interpretable combination $\rightarrow$ use equivalence with Boolean formulas

with $m(A^*) = m_1(A_1)m_2(A_2)\ldots m_n(A_n)$ under source independence (assumed in this part)

k-out-of-n reliable

- Assume that k sources among the n are reliable
- Use conjunction on every sub-group of size k, then disjunction between results:

$$A^* = \cup_{J_\ell} \cap_{\substack{A_i \in J_\ell \\ J_\ell \subseteq \mathcal{A} \\ |J_\ell| = k}} A_i$$

- 1-out-of-n: disjunctive rule
- n-out-of-n: conjunctive rule
- Does not obey Fairness property: some source information can be deleted.
- Tends to forget "outliers" and to go for the majority when k closer to n

Maximal coherent subsets

- Identify every subgroup of $\mathcal{A}$ whose intersection non-empty and maximal with this property, call them $M_1, \ldots, M_L$
- Use conjunction on every maximal subgroups, then disjunction between results:

$$A^* = \cup_{M_\ell} \cap_{A_i \in M_\ell} A_i$$

- If all sets A_i agree: conjunction
- If all sets A_i disagree in pairwise way: disjunction
- Ensure fairness property, but can give imprecise results if outliers
- N.B.: when $n = 2$, equivalent to Dubois/Prade rule

Both rules (k/n and MCS) are not associative

Example

- $A_1 = \{\omega_1, \omega_2, \omega_4, \omega_5\}$
- $A_2 = \{\omega_4\}$
- $A_3 = \{\omega_2, \omega_4\}$
- $A_4 = \{\omega_1\}$

3-out-of-4

$$\{\omega_4\} \cup \emptyset \cup \emptyset \cup \emptyset = \{\omega_4\}$$

2-out-of-4

$$\{\omega_1, \omega_2, \omega_4\}$$

MCS

$$(A_1 \cap A_2 \cap A_3) \cup (A_4 \cap A_1) = \{\omega_1, \omega_4\}$$

Example

- $A_1 = \{\omega_1, \omega_2, \omega_4, \omega_5\}$
- $A_2 = \{\omega_4\}$
- $A_3 = \{\omega_2, \omega_4\}$
- $A_4 = \{\omega_1\}$

3-out-of-4

$$\{\omega_4\} \cup \emptyset \cup \emptyset \cup \emptyset = \{\omega_4\}$$

2-out-of-4

$$\{\omega_1, \omega_2, \omega_4\}$$

MCS

$$(A_1 \cap A_2 \cap A_3) \cup (A_4 \cap A_1) = \{\omega_1, \omega_4\}$$

How to select the rule?

Two proposal:

- Take a set $h_1, \ldots, h_H$ of rules such that imprecision will grow

$$h_i(m_1, \ldots, m_n) \sqsubseteq h_{i+1}(m_1, \ldots, m_n)$$

  with each rule, apply them iteratively and stop when consistency high enough (for example, go from n-out-of-n to 1-out-of-n)
- Select some properties desirable in your application/context, and chose a rule fitting all of them (if possible).

Exercice 6

Given $\Omega = \{\omega_1, \omega_2, \omega_3\}$, fill in the following table

A	m_1	m_2	m_3	$\cup$	2-out-of-3	MCS	$\cap$
$\emptyset$	0	0	0				
$\{\omega_1\}$	0.5	0	0				
$\{\omega_1, \omega_2\}$	0	0.2	0				
$\{\omega_3\}$	0	0	0.6				
$\{\omega_1, \omega_3\}$	0	0	0				
Ω	0.5	0.8	0.4				

Two common situations

Situation 1: interpretable rule

- Efficiency of rules cannot be assessed
- Interpretability of the rule important
- example: few experts, no repeated experiments
- mostly concerns singular interpretation

Situation 2: learnable rule

- **Efficiency of rules can be assessed**
- Interpretability of the rule is a lesser issue
- example: machine learning, many sensors, repeated experiments
- mostly concerns **statistical interpretation**

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Procedure

- A set of observed values $\hat{\omega}^1, \ldots, \hat{\omega}^o$
- for each $\hat{\omega}^i$, a set $m_1^i, \ldots, m_n^i$ of observation
- a decision rule $d : \mathcal{M} \rightarrow \Omega$ (pignistic, max plausibility... see Thierry's talk) mapping m to a decision $d(m) \in \Omega$
- from set $\mathcal{H}$ of possible rules, choose

$$h^* = \arg\max_{h \in \mathcal{H}} \sum_i \mathbb{I}_{d(h(m_1^i,\ldots,m_n^i)=\hat{\omega}^i)}$$

- N.B.: this implicitly assumes a 0/1 cost when making mistakes (again, see Thierry's talk for more complex cases)

How to choose $\mathcal{H}$?

- $\mathcal{H}$ should be easy to browse, i.e., based on few parameters
- Maximization optimization problem should be made easy if possible (convex? Linear?)
- In particular, if m_j^i have peculiar forms (possibilities, Bayesian, k-additive, ...), there is a better hope to find efficient methods

Two examples

- Weighted averaging rules (parameters to learn: weights)
- Denoeux T-(co)norm rules based on canonical decomposition (parameters to learn: parameters of the chosen t-norm family)

The case of averaging rule

- Parameters are weights $\mathbf{w} = (w_1, \ldots, w_n)$ such that $\sum_i w_i = 1$ and $w_i > 0$
- Set $\mathcal{H} = \{h_{\mathbf{w}} | w \in [0,1]^n, \sum_i w_i = 1\}$ with

$$h_{\mathbf{w}} = \sum_i w_i m_i$$

- Decision rule d?
  - pignistic: need to do the full combination of $m_1, \ldots, m_n$ each time
  - **maximum of plausibility**: use the fact that plausibility of average = average of plausibilities

Exercice 7: walking dead

A zombie apocalypse has happened, and you must recognize possible threats/supports

The possibilities Ω

- Zombie (Z)
- Friendly Human (F)
- Hostile Human (H)
- Neutral Human (N)

The sources S_i

- Half-broken heat detector (S_1)
- Watch guy 1 (S_2)
- Motion detector (S_3)
- Watch guy 2 (S_4)

Exercice 7: which rule?

Given this table of contour functions, a weighted average and a decision based on maximal plausibility

	$\hat{\omega}^1 = Z$				$\hat{\omega}^2 = H$				$\hat{\omega}^3 = F$			
	Z	F	H	N	Z	F	H	N	Z	F	H	N
S_1	1	0,5	0,5	0,5	1	0,5	0,5	0,5	0,5	1	1	1
S_2	1	0,2	0,8	0,2	0	0,3	1	0,3	0	0,4	1	0,4
S_3	1	0,5	0,5	0,5	0,5	0,7	0,8	0,7	1	0,5	0,5	0,5
S_4	1	1	1	1	0,2	0,2	1	0,5	0,2	1	0,4	0,8
$\mathbf{w}_1 = (0.5, 0.5, 0, 0)$												
$\mathbf{w}_2 = (0, 0, 0.5, 0.5)$												

Choose $h_{\mathbf{w}_1}$ or $h_{\mathbf{w}_2}$? Given the data, can we find a strictly better weight vector?

Outline

1. Introduction
   - Focusing, Revising, Combining
   - A quick revisit of Dempster's rule
   - When Dempster's rule does not apply
2. On conflict between belief functions
   - Complementing with distances
   - Axiomatic approach
3. Combination: selecting a rule
   - By properties of combination
   - Interpretable rules
   - Learnable rules
   - Other options

Other common options to solve conflicting situations:

- Redistribute the amount $m(\emptyset)$ to selected focal elements [27]
- Correct sources according to knowledge we have on them, provided we have/can estimate it (see David's talk) [25, 24]
- Identify the dependence structure that minimize conflict, but keep conjunctive approach (but conflict not always due to dependence) [22, 23]

Other (open) topics of interest

- When a rule is learnable and we have data, for which cases do we obtain an easy-to-solve optimization problem (e.g., quadratic, convex, ...)
- What means for two sources to be dependent, how can we measure such a dependence, especially in the singular case? Some proposals using distances, yet no strong theoretical results or generic approach on this aspect.
- Build a generic benchmark to compare rules or rule selection strategies
- How to adapt combination in general to (very) large domains? [21]
- How to combine in a distributed way, if no central unit? [26]

references I

Combination and properties

[14] Sebastien Destercke, Didier Dubois, and Eric Chojnacki.
Possibilistic information fusion using maximal coherent subsets.
Fuzzy Systems, IEEE Transactions on, 17(1):79–92, 2009.

[15] Didier Dubois, Weiru Liu, Jianbing Ma, and Henri Prade.
Toward a general framework for information fusion.
In *Modeling Decisions for Artificial Intelligence*, pages 37–48. Springer, 2013.

[16] Mehena Loudahi, John Klein, Jean-Marc Vannobel, and Olivier Colot.
New distances between bodies of evidence based on dempsterian specialization matrices and their consistency with the conjunctive combination rule.
International Journal of Approximate Reasoning, 55(5):1093–1112, 2014.

[17] Peter Walley.
The elicitation and aggregation of beliefs.
Technical report, University of Warwick, 1982.

Combination rules

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

[19] Frederic Pichon, Sebastien Destercke, and Thomas Burger.
A consistency-specificity trade-off to select source behavior in information fusion.
Cybernetics, IEEE Transactions on, 45(4):598–609, 2015.

[20] Benjamin Quost, Marie-Hélène Masson, and Thierry Denœux.
Classifier fusion in the dempster–shafer framework using optimized t-norm based combination rules.
International Journal of Approximate Reasoning, 52(3):353–374, 2011.

Other approaches and open topics

references II

[21] Cyrille André, Sylvie Le Hégarat-Mascle, and Roger Reynaud.
Evidential framework for data fusion in a multi-sensor surveillance system.
Engineering Applications of Artificial Intelligence, 43:166–180, 2015.

[22] Andrey Bronevich and Igor Rozenberg.
The choice of generalized dempster–shafer rules for aggregating belief functions.
International Journal of Approximate Reasoning, 56:122–136, 2015.

[23] Marco EGV Cattaneo.
Belief functions combination without the assumption of independence of the information sources.
International Journal of Approximate Reasoning, 52(3):299–315, 2011.

[24] Arnaud Martin, Anne-Laure Jousselme, and Christophe Osswald.
Conflict measure for the discounting operation on belief functions.
In *Information Fusion, 2008 11th International Conference on*, pages 1–8. IEEE, 2008.

[25] Frédéric Pichon, David Mercier, François Delmotte, and Éric Lefèvre.
Truthfulness in contextual information correction.
In *Belief Functions: Theory and Applications*, pages 11–20. Springer, 2014.

[26] Jovan Radak, Bertrand Ducourthial, Veronique Cherfaoui, and Stephane Bonnet.
Detecting road events using distributed data fusion: Experimental evaluation for the icy roads case.

[27] P. Smets.
Analyzing the combination of conflicting belief functions.
Information Fusion, 8:387–412, 2006.

Some final words

When dealing with conflict and combination, important things to consider:

- What are the assumptions I can make?
- What are the properties I absolutely want?
- Do I care about interpretation?
- Do I care (that much) about tractability and/or scalability?