Information fusion problems

Sebastien Destercke

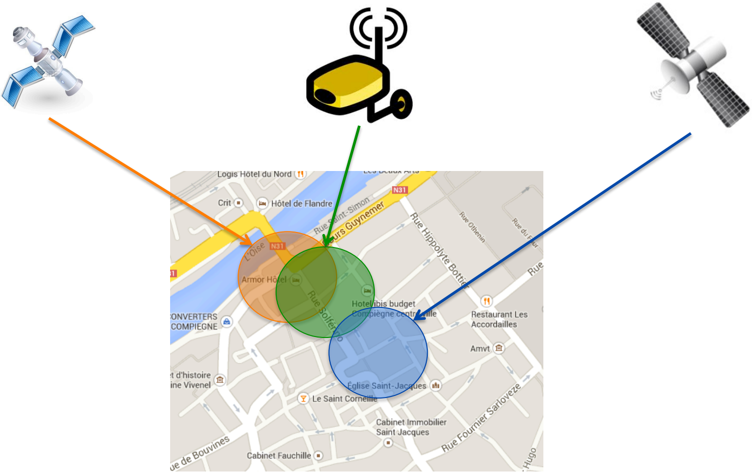
Heuristic and Diagnosis for Complex Systems (HEUDIASYC) laboratory,
Compiegne, France

Workshop on data Reconciliation

Information fusion vs other approaches

- Data reconciliation: reconcile data with model
  $\rightarrow$ you trust more the model than the data
- Regression/Learning: reconcile model with data
  $\rightarrow$ you trust more the data than the model
- Information fusion: reconcile multiple source of informations
  $\rightarrow$ you trust your sources... to some extent

An illustration of the issue



Definition and goals

Combine partial information $E_1, \ldots, E_n$ on quantity X given by n sources:

$$f(E_1, \ldots, E_n) = E^*$$

- X assumed to have a **true**, yet unknown value
- **Goal** of information fusion: how to pick f to
  - Gain information from $E_1, \ldots, E_n$
  - Increase the reliability (trust) in my final result

Modelling (set) information

Set E_1 associated to

- its indicator function

$$E_1(x) = \begin{cases} 1, & x \in E_1 \\ 0, & x \notin E_1 \end{cases}$$

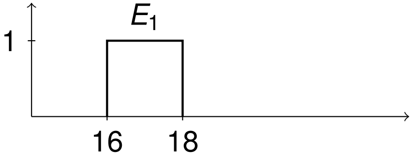
- two uncertainty measures Π, N such that

$$\Pi(A) = \sup_{x \in A} E_1(x)$$

$$N(A) = 1 - \Pi(A^c)$$

$$N(A) \leq \Pi(A)$$

$$E_1 = [16, 18]$$



Measuring uncertainty from sets

3 (extreme) situations

- **Certainty truth** in A

$$E \subseteq A$$

$$\Pi(A) = N(A) = 1$$

- **Ignorance** about A

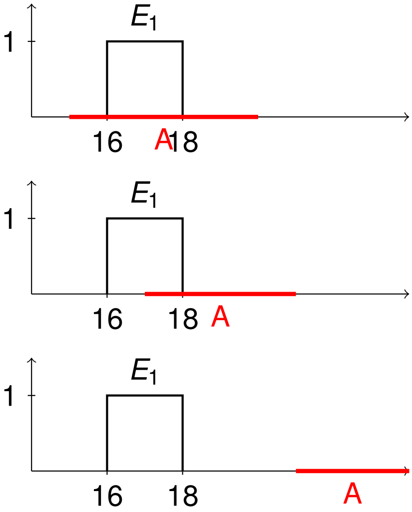
$$A \cap E \neq \emptyset \text{ and } A^c \cap E \neq \emptyset$$

$$\Pi(A) = 1 \text{ and } N(A) = 0$$

- Certainty A **false**

$$E \cap A = \emptyset$$

$$\Pi(A) = N(A) = 0$$



Some issues (+/- talk outline)

- **Issue 1: how to choose f (how to combine)?**
- Issue 2: how to deal with conflict
  - with no information about sources nor further assumptions
  - with added assumptions
- Applying it in practice: a real-world example

!We concentrate on possibility theory, but guidelines true/applied in all other uncertainty theories!

Basic principle 1

Commutativity: all sources equal

In absence of information about them, all sources should be treated equally

If σ is a permutation of $\{1, \ldots, n\}$, then

$$f(E_{\sigma(1)}, \ldots, E_{\sigma(n)}) = f(E_1, \ldots, E_n)$$

$$f(E_1, E_2, E_3) = f(E_1, E_3, E_2) = f(E_2, E_1, E_3) = \cdots$$

Basic principle 2

Zero preservation

An element considered impossible by all sources should remain impossible

If $E_i(x) = 0$ for any E_i, then

$$f(E_1, \ldots, E_n)(x) = 0$$

Finding consensus $\neq$ Information fusion

- Two person stating their preference in terms of temperature X in the room

$$E_1 = 10 \text{ and } E_2 = 20$$

$\rightarrow X = 15$ is an acceptable answer

- Two thermometers measuring the temperature X of the room

$$E_1 = 10.0 \text{ and } E_2 = 20.0$$

$\rightarrow X = 15$ is **not** an acceptable answer (even considering significant digit)

Basic principle 3

Possibility preservation

An element considered possible by all sources should remain possible

If $E_i(x) = 1$ for all E_i, then

$$f(E_1, \dots, E_n)(x) = 1$$

Conjunctive principle

Strong zero preservation

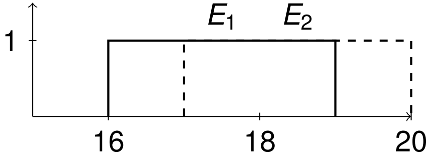
If one source consider an element impossible, it should be impossible

If $E_i(x) = 0$ for any E_i, then

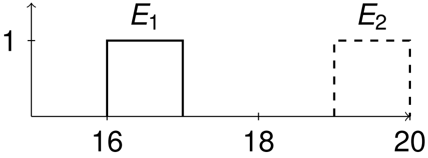
$$f(E_1, \ldots, E_n)(x) = 0$$

$\rightarrow f(E_1, \ldots, E_n)(x) = \min(E_1(x), \ldots, E_n(x)) = \bigcap E_i$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$



Conjunctive principle

Strong zero preservation

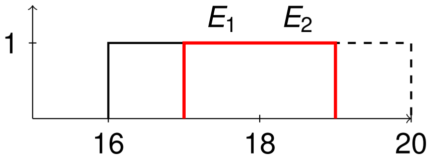
If one source consider an element impossible, it should be impossible

If $E_i(x) = 0$ for any E_i, then

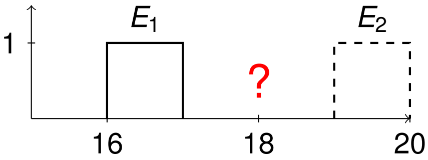
$$f(E_1, \dots, E_n)(x) = 0$$

$\rightarrow f(E_1, \dots, E_n)(x) = \min(E_1(x), \dots, E_n(x)) = \bigcap E_i$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$



Disjunctive principle

Strong possibility preservation

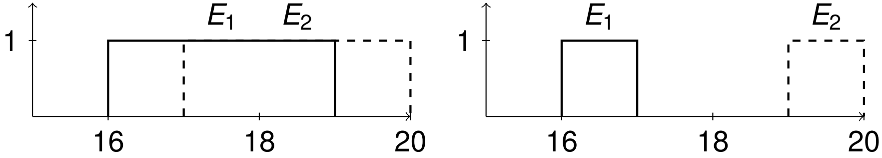
If one source consider an element possible, it should be possible

If $E_i(x) = 1$ for any E_i, then

$$f(E_1, \dots, E_n)(x) = 1$$

$\rightarrow f(E_1, \dots, E_n)(x) = \max(E_1(x), \dots, E_n(x)) = \bigcup E_i$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$ $\quad$ $E_1 = [16, 17]$ and $E_2 = [19, 20]$



Disjunctive principle

Strong possibility preservation

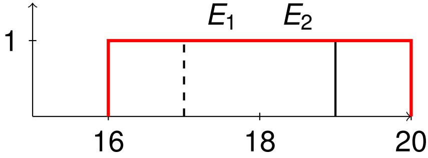
If one source consider an element possible, it should be possible

If $E_i(x) = 1$ for any E_i, then

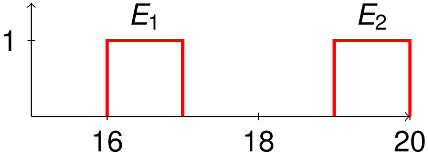
$$f(E_1, \dots, E_n)(x) = 1$$

$\rightarrow f(E_1, \dots, E_n)(x) = \max(E_1(x), \dots, E_n(x)) = \bigcup E_i$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$



The two goals revisited

Goal: how to pick f to

- Gain information from $E_1, \ldots, E_n$
- Increase the reliability (trust) in my final result

Conjunctive approach

- Achieve both goals if no conflict
- Otherwise, need to deal with conflict

Disjunctive approach

- Provide very trustful and never conflicting results
- But very uninformative ($\simeq$ poor practical use)

Some issues (+/- talk outline)

- Issue 1: how to choose *f* (how to combine)?
- **Issue 2: how to deal with conflict**
  - with no information about sources nor further assumptions
  - with added assumptions
- Applying it in practice: a real-world example

!We concentrate on possibility theory, but guidelines true/applied in all other uncertainty theories!

Possibility theory as gradual intervals

Sets: either conflict ($\bigcap E_i = \emptyset$) or not $\rightarrow$ poor flexibility to manage it

Information E_1 associated to

- a possibility function

$$E_1(x) \in [0,1]$$

  with $E_1(x)$ the possibility degree of x

- two uncertainty measures Π, N

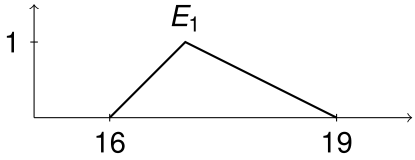
$$\Pi(A) = \sup_{x \in A} E_1(x)$$

$$N(A) = 1 - \Pi(A^c)$$

Typical information: reference value e_1^* + support E_1

$$e_1^* = 17$$

$$E_1 = [16, 19]$$

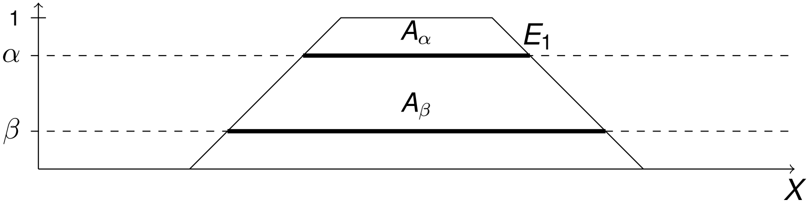


A nice characteristic: Alpha-cut

Definition

$$A_\alpha = \{x \in X : E_1(x) > \alpha\}$$

- $N(A_\alpha) = 1 - \alpha$
- If $\beta \leq \alpha$, $A_\alpha \subseteq A_\beta$

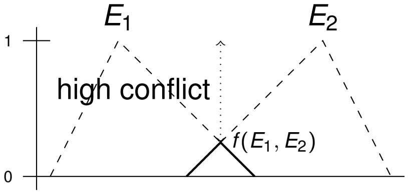


$\Rightarrow$ Nested confidence intervals (useful for elicitation/statistical modelling)

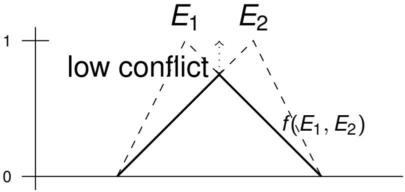
Conjunction and conflict

$$f(E_1, \ldots, E_n) = \min(E_1, \ldots, E_n) = E^*$$

Gradual conflict: $1 - \max_{x \in X} E^*(x)$



Poorly reliable → questionable



Pretty reliable → acceptable

Conflict and no assumption

Agreement principle / no source forgetting

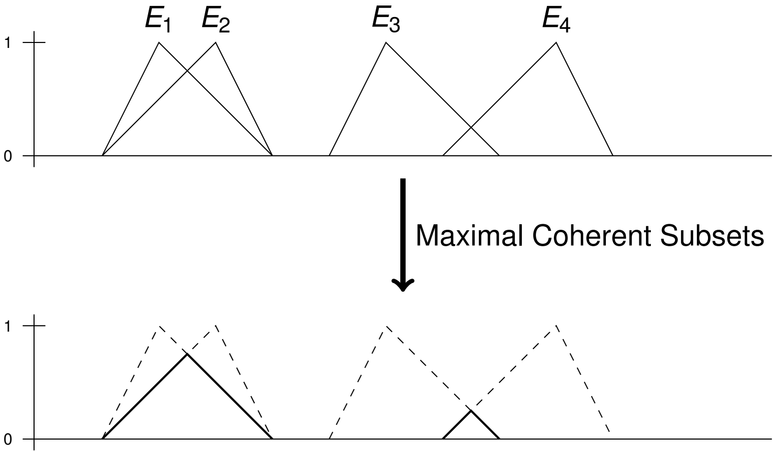
For any E_i, we should have

$$E_i \cap f(E_1, \ldots, E_n) \neq \emptyset$$

Maximal coherent subset principle

1. Conjunction between consistent sub-groups of sources
2. Disjunction between those conjunctions

Maximal coherent subsets: example



In case of conflict, fairness to all sources still provide fairly imprecise results

Adding assumptions

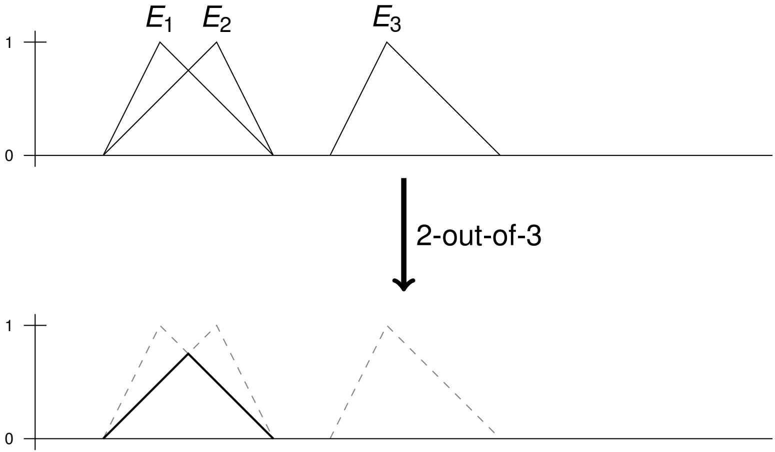
All sources treated equally, but allow forgetting

$\rightarrow$ assume k-out-of-n are reliable: take disjunctions of conjunctions of groups of k sources

- 1-out-of-n: disjunction
- n-out-of-n: conjunction

Similar to a majority principle $\rightarrow$ usually ok for sensors and similar situations, may be questioned in other situations (e.g., expert opinion combination)

k-out-of-n: example



If n not too small and $k \simeq n$, reasonable (and clear assumption)

Qualitative reliability information

Sources no longer equal $\rightarrow$ use fact that some are better than others

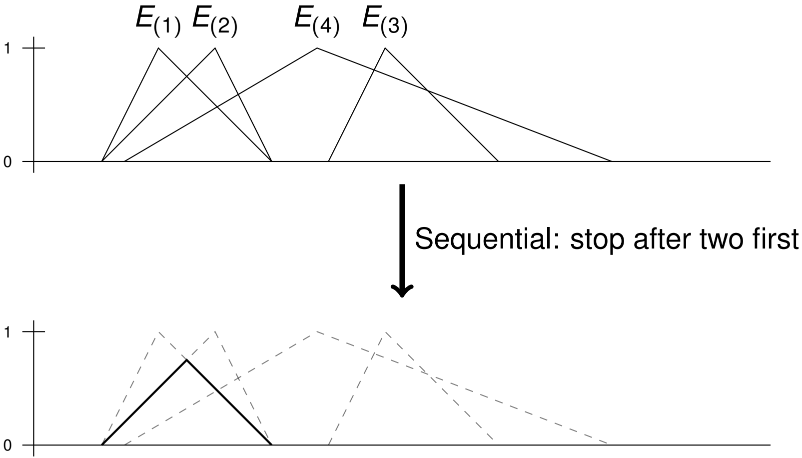
Sequential fusion

- An idea of which sources are more reliable than others

$$E_{(1)} \succ E_{(2)} \succ \ldots \succ E_{(n)}$$

- Successive conjunctive merging, until conflict too important or result empty

sequential: example



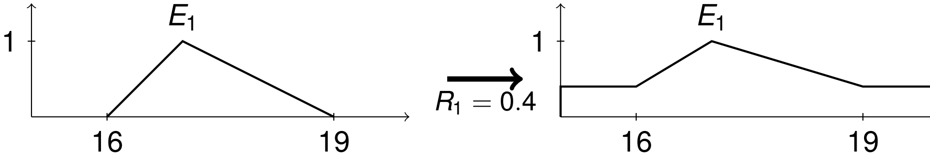
Only require qualitative information → measurements not necessary

Quantitative reliability information

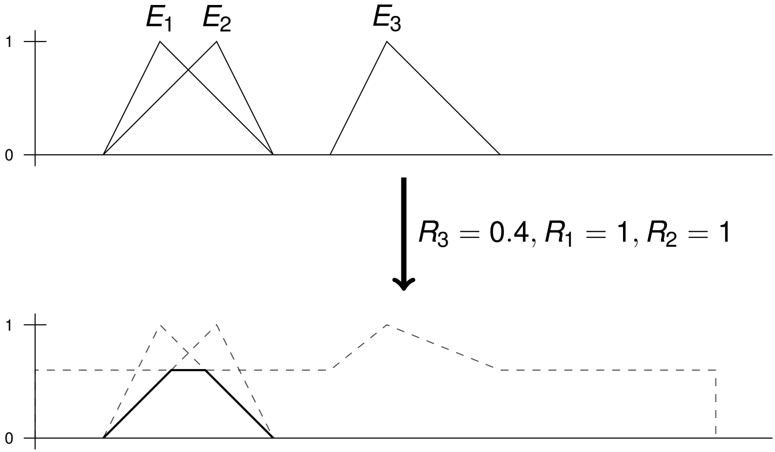
Modifying a source information

- We can evaluate that the reliability $R_i \in [0,1]$ of a source, with $R_i = 1$ complete reliability, $R_i = 0$ complete unreliability
- Modify source E_i into E_i' according to reliability:

$$E_i'(x) = R_i \cdot E_i(x) + (1 - R_i) \cdot 1$$



Reliability "rates": example

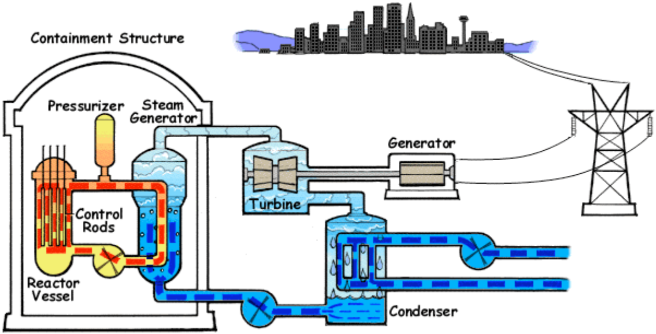


Some issues (+/- talk outline)

- Issue 1: how to choose f (how to combine)?
- Issue 2: how to deal with conflict
  - with no information about sources nor further assumptions
  - with added assumptions
- **Applying it in practice: a real-world example**

!We concentrate on possibility theory, but guidelines true/applied in all other uncertainty theories!

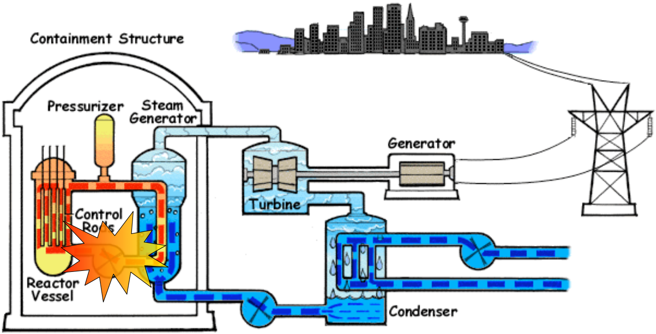
BEMUSE: an international exercice



Cooling circuit break → need to assess uncertainties about peak temperature (to be sure that fusion is avoided)

→ ten participants to a benchmark exercise

Introduction



Cooling circuit break → need to assess uncertainties about peak temperature (to be sure that fusion is avoided)

→ ten participants to a benchmark exercise

BEMUSE: the (raw) results

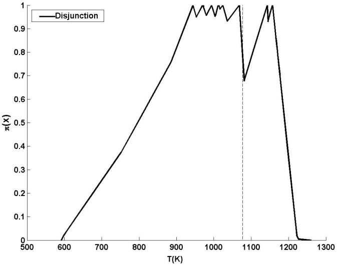
	1PCT (K)			**2PCT (K)**			T_{inj} (s)			T_q (s)		
	Low	Ref	Up	Low	Ref	Up	Low	Ref	Up	Low	Ref	Up
CEA	919	1107	1255	674	993	1176	14.8	16.2	16.8	30	69.7	98
GRS	969	1058	1107	955	1143	1171	14	15.6	17.6	62.9	80.5	103.3
IRSN	872	1069	1233	805	1014	1152	15.8	16.8	17.3	41.9	50	120
KAERI	759	1040	1217	598	1024	1197	12.7	13.5	16.6	60.9	73.2	100
KINS	626	1063	1097	608	1068	1108	13.1	13.8	13.8	47.7	66.9	100
NRI1	913	1058	1208	845	1012	1167	13.7	14.7	17.7	51.5	66.9	87.5
NRI2	903	1041	1165	628	970	1177	12.8	15.3	17.8	47.4	62.7	82.6
PSI	961	1026	1100	887	972	1014	15.2	15.6	16.2	55.1	78.5	88.4
UNIPI	992	1099	1197	708	944	1118	8.0	16.0	23.5	41.4	62.0	81.5
UPC	1103	1177	1249	989	1157	1222	12	13.5	16.5	56.5	63.5	66.5
Exp. Val.		1062			1077			16.8			64.9	

The two basic strategies

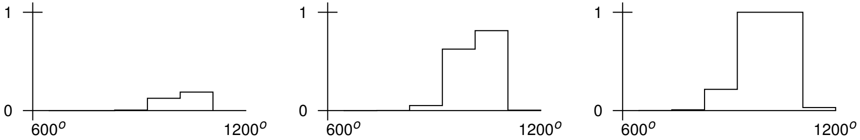
Conjunction

conflict > 0.9

Disjunction

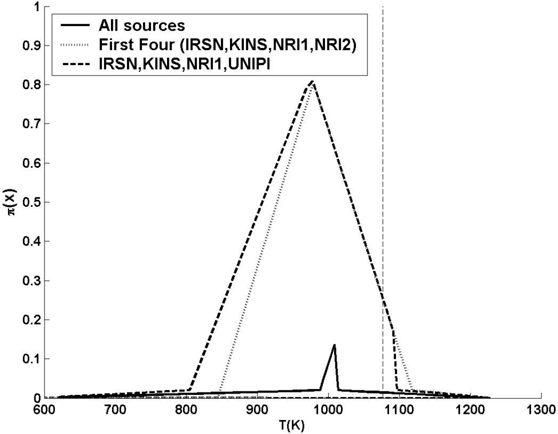


K-out-of-n strategy



10-out-of-10 result   9-out-of-10 result   8-out-of-10 result

Reliabilities from previous performances



As a conclusion

- Merging information from sources to find **true** value of a quantity
- Some basic principles applicable in all cases
- If sources agree, conjunction $\rightarrow$ problem solved
- If not
  - disjunction: truthful but imprecise
  - can be improved without further assumptions (Max. Coh. Sub.), but not by much
  - various strategies to modulate conflict quantity and information gain

Generality

True for most uncertainty theories:

- Interval analysis (scope of action limited)
- Possibility/fuzzy theory
- Dempster-Shafer (Evidence) theory and belief functions
- Convex sets of probabilities (imprecise probability)
- ...

Harder (and less natural) for probability theory

- Conjunction empty unless equality of distributions
- Disjunction not a probability but a set of probabilities

Other fusion problem not considered here

- Decentralized information fusion:
  - one source may not see all other sources
  - information items may not all come at the same time
- Dependency: measuring and integrating dependency information between sources
- Heterogeneity of space description: dealing with sources giving information at different levels or using different descriptions

Some selected references I

[1] A. Appriou, A. Ayoun, S. Benferhat, P. Besnard, I. Bloch, L. CHolvy, R. Cooke, F. Cuppens, D. Dubois, H. Fargier, M. Grabisch, A. Hunter, R. Kruse, J. Lang, S. Moral, H. Prade, A. Safiotti, P. Smets, and C. Sossai.
Fusion : General concepts and characteristics.
I. J. of Intelligent Systems, 16(10):1107–1134, October 2001.

[2] R.M. Cooke.
Experts in uncertainty.
Oxford University Press, Oxford, UK, 1991.

[3] S. Destercke, D. Dubois, and E. Chojnacki.
Possibilistic information fusion using maximal coherent subsets.
IEEE Trans. on Fuzzy Systems (in press), 2008.

Some selected references II

[4] Didier Dubois, Weiru Liu, Jianbing Ma, and Henri Prade.
Toward a general framework for information fusion.
In *MDAI*, pages 37–48, 2013.

[5] S. Moral and J. Sagrado.
Aggregation of imprecise probabilities.
In B. BouchonMeunier, editor, *Aggregation and Fusion of Imperfect Information*, pages 162–188. Physica-Verlag, Heidelberg, 1997.

[6] Frédéric Pichon, Sébastien Destercke, and Thomas Burger.
Selecting source behavior in information fusion on the basis of consistency and specificity.
In *ECSQARU*, pages 473–484, 2013.

Some selected references III

[7] S.A. Sandri, D. Dubois, and H.W. Kalfsbeek.
Elicitation, assessment and pooling of expert judgments using possibility theory.
IEEE Trans. on Fuzzy Systems, 3(3):313–335, August 1995.

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

[9] P. Walley.
Measures of uncertainty in expert systems.
Artifical Intelligence, 83:1–58, 1996.

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