Information Fusion and Decision-Making Support with Belief Functions

Jean Dezert[1] & Deqiang Han[2]

[1]The French Aerospace Lab (ONERA), Palaiseau, France.
jdezert@gmail.com

[2]Xi'an Jiaotong University, Xi'an, China.
deqhan@gmail.com

July 10th, 2017

Part I

Information Fusion with Belief Functions

Outline of Part 1

Short historical overview

1933 - Probability Theory

- studied by Blaise Pascal in 1634
- Objective, i.e. frequency interpretation $P(A) = \frac{\text{# of possible outcomes for event } A}{\text{# of possible outcomes for space } S}$
- Geometric interpretation: $P(A) = \frac{\text{Geometric measure of set } A}{\text{Geometric measure of space } S}$
- Long run freq. interpretation (Von Mises): $P(A) = \lim_{N\to\infty} \frac{\text{# of possible outcomes of event } A}{N \text{ (total # of trials)}}$
- Subjective interpretation (De Finetti): $P(A)$ as subjective degree of belief in A
- Axiomatic framework based on measure theory (Kolmogorov 1933)
- Game-theoretic framework (Vovk & Shafer 2001)

1976 - Dempster-Shafer Theory (DST)

- introduction of Belief Functions (BF) by Shafer based on Dempster's works (1967)

1978 - Theory of possibilities

- introduced by Zadeh, Dubois & Prade.
- Fuzzy sets are interpreted as possibility distributions

1991 - Theory of Imprecise Probabilities

- introduced by Walley to deal with 2nd order probabilities

2003 - Dezert-Smarandache Theory (DSmT)

- new theoretical framework and methods to work with belief functions

Limitations of probabilities

- Deal generally with information drawn from generic knowledge based on population of items, laws of physics, or common sense
- Capture only one aspect of the uncertainty (the randomness, i.e. the variability through repeated measurements)
- Do not account for incomplete knowledge (epistemic uncertainty)
- Cannot distinguish between uncertainty due to variability, and uncertainty due to lack of knowledge

Variability is related with precisely observed random observations

Incompleteness is related with missing and partial information

Consider a random variable W taking its value $w \in [1, 2]$

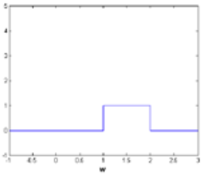
Suppose ignorance modeling is done with uniform distribution on $[1, 2]$ based on the insufficient reason principle

Cumulative distribution function (cdf) of W

$$W \sim u([1,2]) \Leftrightarrow P(W \leqslant w) = \begin{cases} 0 & \text{if } w < 1 \\ w - 1 & \text{if } 1 \leqslant w \leqslant 2 \\ 1 & \text{if } w > 2 \end{cases}$$

Proba density function (pdf) of W

$$p_W(w) \triangleq \frac{\partial}{\partial w} P(W \leqslant w) = \begin{cases} 1 & \text{if } w \in [1,2] \\ 0 & \text{if } w \notin [1,2] \end{cases}$$



On modeling ignorance with probabilities (2)

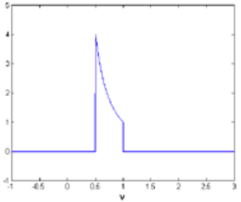
Take $V = 1/W$ with $W \sim u([1,2])$, then $v \in [1/2, 1]$

Cumulative distribution function (cdf) of V

$$P(V \leqslant v) = P(\frac{1}{W} \leqslant v) = P(W \geqslant \frac{1}{v}) = 1 - P(W < \frac{1}{v}) = \begin{cases} 1 & \text{if } \frac{1}{v} < 1 \\ 2 - \frac{1}{v} & \text{if } \frac{1}{v} \in [1,2] \\ 0 & \text{if } \frac{1}{v} > 2 \end{cases}$$

Proba density function (pdf) of V

$$p_V(v) \triangleq \frac{\partial}{\partial v} P(V \leqslant v) = \begin{cases} \frac{1}{v^2} & \text{if } v \in [1/2, 1] \\ 0 & \text{if } w \notin [1/2, 1] \end{cases}$$



V is **not uniformly distributed** on $[1/2, 1]$.This is not very satisfactory to model ignorance because full ignorance on W should not provide information on $1/W$.

The matter in this problem is the choice of random variable $W \in [1,2]$ or $V = 1/W \in [1/2, 1]$ and the particular choice of underlying probability distribution to model ignorance. Probability Theory cannot help efficiently for the choice of a priori distribution under epistemic uncertainty (lack of knowledge).

Basics of the theory of belief functions

Belief is the state of mind in which one thinks something to be true.

History

- introduced by **Glenn Shafer** in 1976 [Shafer 1976]
- also known as **Dempster-Shafer Theory** (DST) in the literature
  http://www.glennshafer.com/books/amte.html

Main references

G. Shafer, A mathematical theory of evidence, 1976.

R. Yager, L. Liu, Classic Works of the Dempster-Shafer Theory of Belief Functions, 2008.

Paradigm shift

Beliefs often are related with singular event or evidence, and are **not necessarily related with statistical data** and generic knowledge.

Frame of discernment and power set

Frame of discernment (FoD)

The set of all possible solutions of the problem under concern is called the FoD. Typically noted

$$\Theta = \{\theta_i, i = 1, \dots, n\}$$

Criminal investigation example (list of suspects)

$$\Theta = \{\theta_1 = \text{Peter}, \theta_2 = \text{Paul}, \theta_3 = \text{Mary}\}$$

Shafer's model of FoD

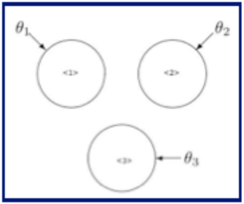
Θ is a finite set, with all elements exclusive two by two.

Power set of Θ is the set of all subsets of Θ (empty set $\emptyset$ included) noted

$$2^\Theta \triangleq \{X | X \subseteq \Theta\}$$

# of elements of the power set : $|2^\Theta| = 2^{|\Theta|}$

$$\Theta = \{\theta_1, \theta_2, \theta_3\} \Rightarrow$$



Impossibility

partial ignorances

full ignorance

$$2^{\Theta} = \{\emptyset, \theta_1, \theta_2, \theta_3, \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cup \theta_2 \cup \theta_3\}$$

$$|2^{\Theta}| = 2^3 = 8$$

Any subset A of the FoD Θ corresponds to the proposition

$$P_\theta(A) \triangleq \text{The true value of } \theta \text{ is in the subset } A \text{ of } \Theta$$

Equivalence between set operators and logical operators

Operations	Subsets	Propositions
Intersection/conjunction	$A \cap B$	$\mathcal{P}_\theta(A) \wedge \mathcal{P}_\theta(B)$
Union/disjunction	$A \cup B$	$\mathcal{P}_\theta(A) \vee \mathcal{P}_\theta(B)$
Inclusion/implication	$A \subset B$	$\mathcal{P}_\theta(A) \Rightarrow \mathcal{P}_\theta(B)$
Complementation/negation	$A = c_\Theta(B)$	$\mathcal{P}_\theta(A) = \neg\mathcal{P}_\theta(B)$

Mass function (i.e. BBA)

A source of evidence (SoE) about θ is represented by a BBA (or mass function) $m^{\Theta}(\cdot) : 2^{\Theta} \mapsto [0,1]$ such that[1]

$$m^{\Theta}(\emptyset) = 0 \qquad \text{and} \qquad \sum_{A \in 2^{\Theta}} m^{\Theta}(A) = 1$$

(1) $\Rightarrow$ no positive mass is committed to impossible event.
(2) $\Rightarrow$ a mass function is normalized to one.

Focal element (**FE**) of $m(\cdot)$

$A \subseteq \Theta$ is a Focal Element (FE) of $m(\cdot)$ if $m(A) > 0$

$$\mathcal{F}(m) \triangleq \{A \in 2^{\Theta} | m(A) > 0\}$$

Core of $m(\cdot)$

$$\mathcal{C}(m) \triangleq \bigcup_{A \in \mathcal{F}(m)} A$$

[1] For notation simplicity $m^{\Theta}(\cdot)$ will be noted $m(\cdot)$ if there is no confusion.

Special BBAs

Let's take the FoD $\Theta = \{A, B, C\}$ as example.

Categorical mass function: $m(\cdot)$ has a unique focal element different from Θ

- $m(A) = 1$ and $m(X) = 0$ for any $X \in 2^{\Theta}$ such that $X \neq A$
- $m(A \cup C) = 1$ and $m(X) = 0$ for any $X \in 2^{\Theta}$ such that $X \neq A \cup C$

Consonant mass function: if FE of $m(.)$ are nested, $A_1 \subset A_2 \ldots \subset \Theta$

- $m(A) = 0.6$, $m(A \cup C) = 0.1$ and $m(A \cup B \cup C) = 0.3$

Dogmatic mass function: if $m(\Theta) = 0$

Certain mass function: if $m(X) = 1$ for some singleton $X \in 2^{\Theta}$

Simple support mass function: if $m(A) = r$ and $m(\Theta) = 1 - r$ for some $A \in 2^{\Theta}$

Bayesian belief mass: FE are only singletons of 2^{Θ} ($\sim$ proba pmf)

- $m(A) = 0.6$, $m(B) = 0.4$
- $m(A) = 1/3$, $m(B) = 1/3$ and $m(C) = 1/3$

Vacuous belief assignment (VBA): It represents the full ignorant (uninformative) SoE

$$m_v(\Theta) = 1 \quad \text{and} \quad m_v(A) = 0, \ \forall A \neq \Theta$$

Belief in A: Total degree of support of A by the source of evidence

$$\text{mtobel}: \quad Bel(A) \triangleq \sum_{B\in 2^\Theta | B\subseteq A} m(B) = Pl(\Theta) - Pl(\bar{A}) = 1 - Pl(\bar{A})$$

Plausibility of A: Total degree of non contradiction of A by the SoE

$$\text{mtopl}: \quad Pl(A) \triangleq \sum_{B\in 2^\Theta | B\cap A\neq\emptyset} m(B) = Bel(\Theta) - Bel(\bar{A}) = 1 - Bel(\bar{A})$$

where $\bar{A} \triangleq \Theta - A$ is the complement of A in Θ.

Belief interval, and uncertainty on A: $\begin{cases} BI(A) \triangleq [Bel(A), Pl(A)] \\ U(A) \triangleq Pl(A) - Bel(A) \end{cases}$

Property: $\forall A \in 2^\Theta$, $Bel(A) \leqslant Pl(A)$

Interpretation: $Bel(A)$ and $Pl(A)$ are usually interpreted as lower and upper bound of the unknown probability $P(A)$ of A, and $\forall A \subseteq \Theta$

$$0 \leqslant Bel(A) \leqslant P(A) \leqslant Pl(A) \leqslant 1$$

$$BI(A) = P(A) = Pl(A) \quad \text{if } m(\cdot) \text{ is a Bayesian BBA}$$

Example of **Bayesian** BBA

$$\Theta = \{A, B, C\}$$

$m(.)$	$Bel(\cdot)$	$Pl(\cdot)$
$m(\emptyset) = 0$	0	0
$m(A) = 0.1$	0.1	0.1
$m(B) = 0.3$	0.3	0.3
$m(C) = 0.6$	0.6	0.6
$m(A \cup B) = 0$	$0.4 = 0.1 + 0.3$	$0.4 = 0.1 + 0.3$
$m(A \cup C) = 0$	$0.7 = 0.1 + 0.6$	$0.7 = 0.1 + 0.6$
$m(B \cup C) = 0$	$0.9 = 0.3 + 0.6$	$0.9 = 0.3 + 0.6$
$m(A \cup B \cup C) = 0$	$1 = 0.1 + 0.3 + 0.6$	$1 = 0.1 + 0.3 + 0.6$

$$Bel(Y) \triangleq \sum_{X \in 2^{\Theta} | X \subseteq Y} m(X) \quad \text{and} \quad Pl(Y) \triangleq \sum_{X \in 2^{\Theta} | X \cap Y \neq \emptyset} m(X)$$

Example of **Non Bayesian** BBA

$$\Theta = \{A, B, C\}$$

$m(.)$	$Bel(\cdot)$	$Pl(\cdot)$
$m(\emptyset) = 0$	0	0
$m(A) = 0.01$	0.01	$0.65 = Bel(A) + 0.04 + 0.2 + 0.4$
$m(B) = 0.02$	0.02	$0.76 = Bel(B) + 0.04 + 0.3 + 0.4$
$m(C) = 0.03$	0.03	$0.93 = Bel(C) + 0.2 + 0.3 + 0.4$
$m(A \cup B) = 0.04$	$0.07 = 0.01 + 0.02 + 0.04$	$0.97 = Bel(A \cup B) + 0.2 + 0.3 + 0.4$
$m(A \cup C) = 0.2$	$0.24 = 0.01 + 0.03 + 0.2$	$0.98 = Bel(A \cup C) + 0.04 + 0.3 + 0.4$
$m(B \cup C) = 0.3$	$0.35 = 0.02 + 0.03 + 0.3$	$0.99 = Bel(B \cup C) + 0.04 + 0.2 + 0.4$
$m(A \cup B \cup C) = 0.4$	1	1

$$Bel(Y) \triangleq \sum_{X \in 2^\Theta | X \subseteq Y} m(X) \quad \text{and} \quad Pl(Y) \triangleq \sum_{X \in 2^\Theta | X \cap Y \neq \emptyset} m(X)$$

$\mathrm{Bel}(\cdot): 2^{\Theta} \mapsto [0,1]$ is a monotone capacity function which satisfies

$$\mathrm{Bel}(\emptyset) = 0 \quad \text{and} \quad \mathrm{Bel}(\Theta) = 1$$

and $\forall k \geqslant 2$ and for any collection $A_1, \ldots, A_k$ in 2^{Θ} the inequality

$$\mathrm{Bel}\Big(\bigcup_{i=1}^{k} A_i\Big) \geqslant \sum_{\emptyset \neq I \subset \{1,\ldots,k\}} (-1)^{|I|+1} \mathrm{Bel}\Big(\bigcap_{i \in I} A_i\Big)$$

Properties of Bel

- Sub-additivity: $\mathrm{Bel}(A) + \mathrm{Bel}(B) \leqslant \mathrm{Bel}(A \cup B)$, in particular $\mathrm{Bel}(A) + \mathrm{Bel}(\bar{A}) \leqslant 1$
- Monotonicity: $A \subseteq B \Rightarrow \mathrm{Bel}(A) \leqslant \mathrm{Bel}(B)$

Properties of Pl

- Super-additivity: $\mathrm{Pl}(A) + \mathrm{Pl}(B) \geqslant \mathrm{Pl}(A \cup B)$, in particular, $\mathrm{Pl}(A) + \mathrm{Pl}(\bar{A}) \geqslant 1$
- Monotonicity: $A \subseteq B \Rightarrow \mathrm{Pl}(A) \leqslant \mathrm{Pl}(B)$

Dempster construction of belief functions by multivalued mapping

Fundamental Dempster's idea [Dempster 1967]

Belief (lower proba) and Plausibility (upper proba) construction come from a **multivalued mapping** as follows

- Start with a random variable X with set of possible values in $\mathcal{X} = \{x_j, \ldots, x_m\}$ with known probabilities $p_j = P(X = x_j)$
- Choose a frame of discernment $\Theta = \{\theta_1, \ldots, \theta_n\}$ for the variable θ
- Learn a (multivalued) mapping $\Gamma : \mathcal{X} \mapsto 2^\Theta$ with the meaning: if $X = x_i$, then $\theta \in A$, where $A = \Gamma(x_i) \in 2^\Theta$
- The belief (lower proba) and plausibility (upper proba) that $\theta \in A$ are given by

$$P_*(A) = Bel(A) = Bel(\theta \in A) = P(\{x \in \mathcal{X} | \Gamma(x) \neq \emptyset, \Gamma(x) \subseteq A\})$$

$$P^*(A) = Pl(A) = Pl(\theta \in A) = P(\{x \in \mathcal{X} | \Gamma(x) \cap A \neq \emptyset\})$$

see examples on the next slide

Smets TBM proposal [Smets 1990, Smets Kennes 1994]

Smets proposed his Transferable Belief Model (TBM) to justify belief functions axiomatically with no need of underlying probabilistic multivalued mapping $\Gamma(\cdot)$.

Example **Testimony report from sometimes reliable witness**

Paul has been killed and Police asks a witness W: Did you see Mary killing Paul? Witness answer is **Yes**

$\mathcal{X} = \{x_1 = W \text{ is reliable}, x_2 = W \text{ is not reliable}\}$, and assume $\begin{cases} P(x_1) = 0.4 \\ P(x_2) = 0.6 \end{cases}$

FoD $\Theta = \{\theta_1 = \text{Mary is guilty}, \theta_2 = \text{Mary is not guilty}\}$

Multivalued mapping

$$\Gamma(x_1 = W \text{ is reliable}) = \theta_1 \quad \Rightarrow \text{Mary is guilty}$$
$$\Gamma(x_2 = W \text{ is not reliable}) = \{\theta_1, \theta_2\} = \Theta \quad \Rightarrow \text{We don't know}$$

Belief values

$$Bel(\theta_1) = P(\{x|\Gamma(x) \subseteq \theta_1\}) = P(x_1 = \text{reliable}) = 0.4$$
$$Bel(\theta_2) = P(\{x|\Gamma(x) \subseteq \theta_2\}) = 0$$
$$Bel(\theta_1 \cup \theta_2) = P(\{x|\Gamma(x) \subseteq \theta_1 \cup \theta_2\}) = P(\{x_1, x_2\}) = P(x_1) + P(x_2) = 1$$

Plausibility values

$$Pl(\theta_1) = P(\{x|\Gamma(x) \cap \theta_1 \neq \emptyset\}) = P(\{x_1, x_2\}) = P(x_1) + P(x_2) = 1$$
$$Pl(\theta_2) = P(\{x|\Gamma(x) \cap \theta_2 \neq \emptyset\}) = P(x_2) = 0.6$$
$$Pl(\theta_1 \cup \theta_2) = P(\{x|\Gamma(x) \cap (\theta_1 \cup \theta_2) \neq \emptyset\}) = P(\{x_1, x_2\}) = P(x_1) + P(x_2) = 1$$

Example Testimony report from more or less precise witness

$\mathcal{X} = \{x_1 = W \text{ is precise}, x_2 = W \text{ is approximate}, x_3 = W \text{ is not reliable}\}$
and assume $P(x_1) = 0.3$, $P(x_2) = 0.1$ and $P(x_3) = 0.6$

FoD: $\Theta = \{\theta_1 = \text{Mary}, \theta_2 = \text{Peter}, \theta_3 = \text{John}\}$

Paul has been killed and Police asks a witness W: Who did you see killing Paul?
Witness answer is **Mary**

Multivalued mapping:

$$\Gamma(x_1 = W \text{ is precise}) = \theta_1 \quad \Rightarrow \text{Mary killed Paul}$$
$$\Gamma(x_2 = W \text{ is approximate}) = \{\theta_1, \theta_2\} \quad \Rightarrow \text{Mary or Peter killed Paul}$$
$$\Gamma(x_3 = W \text{ is not reliable}) = \{\theta_1, \theta_2, \theta_3\} = \Theta \quad \Rightarrow \text{We don't know}$$

Belief values

$$Bel(\theta_1) = P(\{x|\Gamma(x) \subseteq \theta_1\}) = P(x_1 = W \text{ is precise}) = 0.4$$
$$Bel(\theta_2) = P(\{x|\Gamma(x) \subseteq \theta_2\}) = 0$$
$$Bel(\theta_3) = P(\{x|\Gamma(x) \subseteq \theta_3\}) = 0$$
$$Bel(\theta_1 \cup \theta_2) = P(\{x|\Gamma(x) \subseteq \theta_1 \cup \theta_2\}) = P(\{x_1, x_2\}) = P(x_1) + P(x_2) = 0.4$$
$$Bel(\Theta) = P(x|\Gamma(x) \subseteq \Theta) = P(\{x_1, x_2, x_3\}) = P(x_1) + P(x_2) + P(x_3) = 1$$

Möbius inversion formula [Kennes 1992]

To any $Bel(\cdot)$ functions corresponds a unique mass function $m(\cdot)$ given by

$$\text{beltom}: \quad \forall A \in 2^{\Theta}, \quad m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Bel(B)$$

To any $Pl(\cdot)$ functions corresponds a unique mass function $m(\cdot)$ given by

$$\text{pltom}: \quad \forall A \in 2^{\Theta}, \quad m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} (1 - Pl(\bar{B}))$$

$m(\cdot)$, $Bel(\cdot)$ and $Pl(\cdot)$ are one-to-one and are **equivalent representations** of a SoE.

Useful for computation of belief functions in fusion rules

Implicability function

$$\text{mtob}: \quad b(A) \triangleq \sum_{B\in 2^\Theta | B\subseteq A} m(B) = Bel(A) + m(\emptyset)$$

$$\text{btom}: \quad m(A) = \sum_{B\in 2^\Theta | B\subseteq A} (-1)^{|A|-|B|} b(B)$$

Commonality function

$$\text{mtoq}: \quad q(A) \triangleq \sum_{B\in 2^\Theta | B\supseteq A} m(B)$$

$$\text{qtom}: \quad m(A) = \sum_{B\in 2^\Theta | B\supseteq A} (-1)^{|A|-|B|} q(B)$$

All one-to-one transformations between Bel, b, Pl, q and m are listed in [Smets 2002]

Discounting sources of evidence

Shafer's reliability discounting rule [Shafer 1976]

To be used if one has a **good estimation** of the reliability factor $\alpha \in [0,1]$ of the SoE based on past experiments and ground truth.

$$\begin{cases} m^{\alpha}(A) \triangleq \alpha \cdot m(A) & \forall A \neq \Theta \\ m^{\alpha}(\Theta) \triangleq \alpha \cdot m(\Theta) + (1-\alpha) \end{cases}$$

$\alpha = 1$ means "the SoE is 100% reliable" $\Rightarrow m^{\alpha=1}(\cdot) = m(\cdot)$ (the BBA is unchanged)
$\alpha = 0$ means "the SoE is 100% unreliable" $\Rightarrow m^{\alpha=0}(\cdot) = m_v(\cdot)$ (the BBA is changed to vacuous BBA)

If a source is totally unreliable $(\alpha = 0)$, it can be combined with the other BBAs **if and only if** the fusion rule preserves the neutral impact of vacuous BBA, otherwise this source must be discarded (i.e. removed of the set of BBAs to fuse)

More refined discounting rules exist

- **Contextual discounting** [Mercier et al. 2005, Mercier et al. 2006]

Importance discounting of a BBA

Proposed in [Smarandache Dezert Tacnet 2010] to take into account the importance of a SoE in the fusion process (see later).

Importance discounting rule

The importance factor of the SoE is modeled by $\beta \in [0, 1]$, and discounted BBA by

$$\begin{cases} m^{\beta}(A) \triangleq \beta \cdot m(A) & \forall A \neq \emptyset \\ m^{\beta}(\emptyset) \triangleq \beta \cdot m(\emptyset) + (1 - \beta) \end{cases}$$

$\beta = 1$ means "the SoE is 100% important" $\Rightarrow m^{\beta=1}(\cdot) = m(\cdot)$
$\beta = 0$ means "the SoE is not important at all" $\Rightarrow m^{\beta=0}(\emptyset) = 1$

If a source is not important at all $(\beta = 0)$, this source must be discarded (i.e. removed of the set of BBAs to fuse)

Note: Important discounted BBA $m^{\beta\neq1}(\cdot)$ is improper (i.e. not regular) since $m^{\beta\neq1}(\emptyset) > 0$. It is however necessary to distinguish importance discounting from reliability discounting in the fusion of sources. This discounting is useful in Multi-Criteria Decision-Making Support problems involving BF (see Part II).

Dempster-Shafer rule of combination

Dempster-Shafer fusion rule [Dempster 1967, Shafer 1976]

Let m_1 and m_2 be mass functions over the **same** frame Θ provided by two **distinct** SoE[2]. DS fusion rule $m_1 \oplus m_2$ is defined by $m_{12}^{DS}(\emptyset) = 0$, and $\forall X \in 2^\Theta$

$$m_{12}^{DS}(X) = [m_1 \oplus m_2](X) \triangleq \frac{m_{12}(X)}{1 - m_{12}(\emptyset)}$$

where $m_{12}(\cdot)$ is the conjunctive rule[3] defined $\forall X \in 2^\Theta$ by

$$m_{12}(X) \triangleq \sum_{X_1, X_2 \in 2^\Theta | X_1 \cap X_2 = X} m_1(X_1) m_2(X_2)$$

Degree of conflict between m_1 and m_2

$$K_{12} \triangleq m_{12}(\emptyset) = \sum_{X_1, X_2 \in 2^\Theta | X_1 \cap X_2 = \emptyset} m_1(X_1) m_2(X_2)$$

DS formula can be used if $m_{12}(\emptyset) < 1$, i.e. the SoE are not in total conflict
DS formula extents directly for the combination of $n > 2$ distinct SoE.

DS rule = Normalized Conjunctive rule

[2] assumed both reliable with same importance.
[3] We also use notation $m_{12}^{Conj}(.)$ to identify it more precisely if needed.

DS rule is not idempotent in general : if m is not categorical then $m \oplus m \neq m$

Advantages

- **Commutativity**: $m_1 \oplus m_2 = m_2 \oplus m_1$
- **Associativity**: One can do the fusion sequentially in any order

$$m_1 \oplus m_2 \oplus m_3 \oplus \ldots m_n = (((m_1 \oplus m_2) \oplus m_3) \oplus \ldots) \oplus m_n$$

- **Neutrality of VBA**: Full ignorant SoE does not impact the fusion result

$$m \oplus m_v = m$$

- *Some similarity* with Bayes rule for conditioning by a certain set $m_Z(Z) = 1$

$$m(X|Z) = [m \oplus m_Z](X) \Rightarrow \begin{cases} Bel(X|Z) = \frac{Bel(X \cup \bar{Z}) - Bel(\bar{Z})}{1 - Bel(\bar{Z})} \\ Pl(X|Z) = \frac{Pl(X \cap Z)}{Pl(Z)} \end{cases}$$

Drawbacks

- **Very complex** in the worst case when $\mathcal{F}(m_1) = \mathcal{F}(m_2) = 2^\Theta - \{\emptyset\}$ for large FoD
- **Counter-intuitive** results in an infinite number of cases even if the conflict is low!

The validity of DS rule and DST has been disputed by many authors including [Zadeh 1979, Lemmer 1985, Voorbraak 1988, Gelman 2006, Dezert Tchamova 2011, Brodzik Enders 2011, Dezert Wang Tchamova 2012, Tchamova Dezert 2012, Dezert Tchamova Han Tacnet 2013, Dezert Tchamova 2014, Heendeni et al. 2016]

Try to work with simpler FoD (by coarsening) and approximate BBAs (less FE, etc)

Sampling technique to approximate DS result [Wilson 1991, Dambreville 2009]

The estimate $\hat{m}_{12}^{DS}(\cdot)$ of $m_{12}^{DS}(\cdot)$ can be obtained by the sampling process using N samples as follows

1. Repeat from $n = 1, \dots, N$
   - draw $Y_1 \in 2^\Theta$ from BBA m_1, and Y_2 from m_2
   - if $Y_1 \cap Y_2 = \emptyset$, set X_n = rejected
   - otherwise, set $X_n = Y_1 \cap Y_2$
2. Compute the rejection rate

$$\hat{z} = \frac{1}{N} \sum_{n=1,\dots,N} I[X_n = \text{rejected}]$$

3. For any $X \in 2^\Theta$, approximate $m_{12}^{DS}(X)$ by

$$\hat{m}_{12}^{DS}(X) = \frac{1}{N(1-\hat{z})} \sum_{n=1,\dots,N} I[X_n = X] \approx m_{12}^{DS}(X)$$

where $I[X_n = X]$ is Kronecker delta function, i.e $I[X_n = X] = \delta(X_n, X) = 1$ if $X_n = X$, and zero otherwise.

Zadeh's example [Zadeh 1979]

Medical diagnosis problem

$$\Theta = \{M = \text{Meningitis}, C = \text{Concussion}, T = \text{Tumor}\}$$

Bayesian BBA in high conflict

Two independent doctors provides the following reports for a patient as follows

$$\begin{array}{lll} m_1(M) = 1 - \epsilon_1 & m_1(C) = 0 & m_1(T) = \epsilon_1 \\ m_2(M) = 0 & m_2(C) = 1 - \epsilon_2 & m_2(T) = \epsilon_2 \end{array}$$

The conflict is $K_{12} = m_{12}(\emptyset) = (1-\epsilon_1)(1-\epsilon_2) + (1-\epsilon_1)\epsilon_2 + \epsilon_2(1-\epsilon_1) = 1 - \epsilon_1\epsilon_2$
Suppose doctors are in hight conflict, say $\epsilon_1 = \epsilon_2 = 0.1$ and so $K_{12} = 1 - 0.01 = 0.99$

$$m_{12}^{DS}(T) = \frac{m_1(T)m_2(T)}{1 - K_{12}} = \frac{\epsilon_1\epsilon_2}{1-(1-\epsilon_1\epsilon_2)} = \frac{\epsilon_1\epsilon_2}{\epsilon_1\epsilon_2} = 1$$

DS fusion results says that patient suffers of Tumor which is counter-intuitive, because both doctors agree that there is a little chance that it is a tumor.

- DS rule provides same results whatever the values $\epsilon_1 > 0$ and $\epsilon_2 > 0$ are !
- DS rule provides coherent result only when $\epsilon_1 = \epsilon_2 = 1$ (i.e. non conflict case)

Proponents of DS rule have strongly disputed this example ... but more interesting examples exist.

Zadeh's example with Low Conflict

Bayesian BBA in **low** conflict

Two independent doctors provides the following reports for a patient as follows

$$\begin{array}{lll} m_1(M) = 0.01 & m_1(C) = 0 & m_1(T) = 0.99 \\ m_2(M) = 0 & m_2(C) = 0.01 & m_2(T) = 0.99 \end{array}$$

The doctors are in very low conflict because

$$K_{12} = 1 - \epsilon_1\epsilon_2 = 1 - 0.9801 = 0.0199$$

Applying DS rule yields

$$m_{12}^{DS}(T) = \frac{m_1(T)m_2(T)}{1 - K_{12}} = \frac{\epsilon_1\epsilon_2}{\epsilon_1\epsilon_2} = 1$$

DS fusion result gives complete support for the diagnosis of a brain tumor, i.e. patient suffers of Tumor for sure, which both doctors believed very likely.

DS result is **counter-intuitive** and one rather expects $m(T) < 1$ because the existence of non-zero belief masses for other diagnoses implies less than complete support for the brain tumor diagnosis, because conflict is non null.

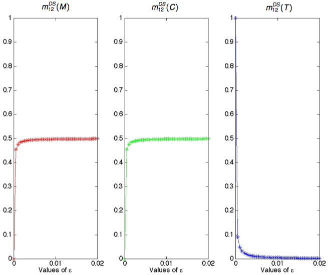
Numerical robustness issue for DS rule

Consider Zadeh's example and change a bit the inputs as follows

$$\begin{array}{lll} m_1(M) = 0.99 - \epsilon & m_1(C) = \epsilon & m_1(T) = 0.01 \\ m_2(M) = \epsilon & m_2(C) = 0.99 - \epsilon & m_2(T) = 0.01 \end{array}$$

if $\epsilon = 0$, $m_{12}^{DS}(M) = 0 \quad m_{12}^{DS}(C) = 0 \quad m_{12}^{DS}(T) = 1$
if $\epsilon = 0.0005$, $m_{12}^{DS}(M) = 0.4541 \quad m_{12}^{DS}(C) = 0.4541 \quad m_{12}^{DS}(T) = 0.0918$

When ϵ changes, one gets



DS rule is not robust to slight input changes.

A more interesting example [Dezert Tchamova 2011, Dezert Wang Tchamova 2012]

Dezert-Tchamova example (2011)

Non-Bayesian BBA $\Theta = \{A, B, C\}$, with $m_1 \neq m_2 \neq m_v$

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	b_1
C	0	$1-b_1-b_2$
$A \cup B \cup C$	0	b_2

Conjunctive rule

$$m_{12}(A) = m_1(A)m_2(A \cup B) + m_1(A)m_2(A \cup B \cup C) = a(b_1 + b_2)$$
$$m_{12}(A \cup B) = m_1(A \cup B)m_2(A \cup B) + m_1(A \cup B)m_2(A \cup B \cup C) = (1-a)(b_1 + b_2)$$

Degree of conflict: $\Rightarrow$ Independent of m_1 !!!

$$\begin{aligned} K_{12} = m_{12}(\emptyset) &= m_1(A)m_2(C) + m_1(A \cup B)m_2(C) \\ &= a(1-b_1-b_2) + (1-a)(1-b_1-b_2) = 1 - b_1 - b_2 \end{aligned}$$

Note: K_{12} can be chosen as low or as high as we want.

Dezert-Tchamova example (cont'd)

Applying DS rule gives

$$m_{12}^{DS}(A) = \frac{m_{12}(A)}{1-K_{12}} = \frac{a(b_1+b_2)}{b_1+b_2} = a = m_1(A)$$
$$m_{12}^{DS}(A\cup B) = \frac{m_{12}(A\cup B)}{1-K_{12}} = \frac{(1-a)(b_1+b_2)}{b_1+b_2} = 1-a = m_1(A\cup B)$$

Remarks

- $m_{12}^{DS}(\cdot) = [m_1 \oplus m_2](\cdot) = m_1(\cdot)$, even if $m_2 \neq m_v$ and $K_{12} > 0$
- Informative source m_2 does not impact DS result !
- Dictatorial power of DS rule !
- The level of conflict does not matter at all !
- Cast serious doubts on normalization step used in DS rule

DS rule result is very counter-intuitive in such Non-Bayesian example (even with low conflict!)

$\Rightarrow$ Need for better rule of combination (better behavior and numerical robustness)
$\Rightarrow$ Logical contradiction in foundations of DST [Dezert Tchamova 2014]

Naive Bayes fusion rule $\Rightarrow$ one assumes $P(Z_1 \cap Z_2|X) = P(Z_1|X)P(Z_2|X)$

$$P(X|Z_1 \cap Z_2) = \frac{P(Z_1 \cap Z_2 \cap X)}{P(Z_1 \cap Z_2)} = \frac{P(Z_1 \cap Z_2|X)P(X)}{P(Z_1 \cap Z_2)} = \frac{P(Z_1|X)P(Z_2|X)P(X)}{\sum_{i=1}^{N} P(Z_1|X=x_i)P(Z_2|X=x_i)P(X=x_i)}$$

DS rule is not a generalization of Bayes rule because it is incompatible with Bayes rule when the prior is not uniform, nor vacuous [Dezert Tchamova Han Tacnet 2013]

Example $\Theta_X \triangleq \{x_1, x_2, x_3\}$ with Shafer's model

Prior pmf

$$\begin{cases} m_0(x_1) = P(X = x_1) = 0.6 \\ m_0(x_2) = P(X = x_2) = 0.3 \\ m_0(x_3) = P(X = x_3) = 0.1 \end{cases}$$

$P(X|Z_1)$

$$\begin{cases} m_1(x_1) = P(X = x_1|Z_1) = 0.2 \\ m_1(x_2) = P(X = x_2|Z_1) = 0.3 \\ m_1(x_3) = P(X = x_3|Z_1) = 0.5 \end{cases}$$

$P(X|Z_2)$

$$\begin{cases} m_2(x_1) = P(X = x_1|Z_2) = 0.5 \\ m_2(x_2) = P(X = x_2|Z_2) = 0.1 \\ m_2(x_3) = P(X = x_3|Z_2) = 0.4 \end{cases}$$

Fusion with Bayes rule

$$\begin{cases} P(x_1|Z_1 \cap Z_2) & = \frac{0.2 \cdot 0.5/0.6}{2.2667} = \frac{0.1667}{2.2667} \approx 0.0735 \\ P(x_2|Z_1 \cap Z_2) & = \frac{0.3 \cdot 0.1/0.3}{2.2667} = \frac{0.1000}{2.2667} \approx 0.0441 \\ P(x_3|Z_1 \cap Z_2) & = \frac{0.5 \cdot 0.4/0.1}{2.2667} = \frac{2.0000}{2.2667} \approx 0.8824 \end{cases}$$

$\neq$

Fusion with DS rule

$$\begin{cases} m_{012}^{DS}(x_1) & = \frac{0.2 \cdot 0.5 \cdot 0.6}{1 - 0.9110} = \frac{0.060}{0.089} \approx 0.6742 \\ m_{012}^{DS}(x_2) & = \frac{0.3 \cdot 0.1 \cdot 0.3}{1 - 0.9110} = \frac{0.009}{0.089} \approx 0.1011 \\ m_{012}^{DS}(x_3) & = \frac{0.5 \cdot 0.4 \cdot 0.1}{1 - 0.9110} = \frac{0.020}{0.089} \approx 0.2247 \end{cases}$$

DS rule is **compatible** with (naive) Bayes rule **only if the prior is uniform or vacuous**

1. due to different reliability of the SoE (based on statistical criteria)
2. due to the possible subjectivity and bias of the SoE because they can have their own interpretation of elements of the FoD
3. due to the final interest of experts/SoE which can be different/antagonist when they report their assessment on a given problem . . .
4. due to serious flaw in DST foundations (logical contradiction)

Classical Attempts to prevent problems with DS rule

- apply ad-hoc thresholding techniques on the degree of conflict level to accept, or reject, DS result
- modify BBAs of SoE by discounting techniques
- identify the bad SoE and don't use it in the fusion
- mix the previous strategies

. . . but DS rule results can still be problematic $\Rightarrow$ **switch for better rules**

This is what DSmT proposes (see later) . . .

Other rules of combination

Conjunctive rule It keeps only the items of information asserted by both sources

$$m_{12}^{Conj}(X) = [m_1 \circledcirc m_2](X) \triangleq \sum_{X_1, X_2 \in 2^{\Theta} | X_1 \cap X_2 = X} m_1(X_1) m_2(X_2)$$

Defended by Smets in his Transferable Belief Model (TBM) [Smets 1990]

- Commutative, associative, not idempotent, numerically robust
- Neutrality of VBA $\Rightarrow m \circledcirc m_v = m$

Implemented with Fast Möbius Transform by the product of commonalily numbers [Smets 2002]

$$\begin{cases} \mathbf{m}_1 \\ \mathbf{m}_2 \end{cases} \rightarrow \begin{cases} \mathbf{q}_1 = \mathrm{mtoq}(\mathbf{m}_1) \\ \mathbf{q}_2 = \mathrm{mtoq}(\mathbf{m}_1) \end{cases} \rightarrow \mathbf{q}_{12} = \mathbf{q}_1 . * \mathbf{q}_2 \rightarrow \mathbf{m}_{12}^{Conj} = \mathrm{qtom}(\mathbf{q}_{12})$$

This rule is problematic because $\emptyset$ is an absorbing element for this rule

- Fast tendency to get $m_{12\ldots n}^{Conj}(\emptyset) = 1$ when fusing many BBAs (directly or sequentially) which makes the result quickly useless
- ambiguous interpretation of the empty set

Simple example of behavior of conjunctive rule

Independent sensor (or expert) reports expressed by BBAs are fused sequentially with the conjunctive rule in the TBM framework

$\Theta = \{A, B, C\}$ with Shafer model for the FoD

- Time 1: $m_1(A) = 0.4$, $m_1(B) = 0$, $m_1(C) = 0.6$
- Time 2: $m_2(A) = 0.7$, $m_1(B) = 0.3$, $m_1(C) = 0$
  - **TBM Conjunctive rule** $m_1 \bigcirc\!\!\!\!\cap\, m_2$**:** $m_{12}^{Conj}(A) = 0.28$, $m_{12}^{Conj}(\emptyset) = 0.72$
  - **DS rule** $m_1 \oplus m_2$**:** $m_{12}^{DS}(A) = 1$
- Time 3: $m_3(A) = 0$, $m_1(B) = 0.8$, $m_1(C) = 0.2$
  - **TBM Conjunctive rule** $(m_1 \bigcirc\!\!\!\!\cap\, m_2) \bigcirc\!\!\!\!\cap\, m_3$**:** $m_{123}^{Conj}(\emptyset) = 1$
  - **DS rule** $(m_1 \oplus m_2) \oplus m_3$**:** Not applicable (total conflict between m_3 and m_{12}^{Conj})
- Time 4, 5, $\ldots$ k: if taking into account new evidential reports, one gets
  - **TBM Conjunctive rule** $((m_1 \bigcirc\!\!\!\!\cap\, m_2) \bigcirc\!\!\!\!\cap\, m_3) \ldots \bigcirc\!\!\!\!\cap\, m_k$**:** $m_{12\ldots k}^{Conj}(\emptyset) = 1$
  - **DS rule** $((m_1 \oplus m_2) \oplus m_3) \ldots m_k$**:** Not applicable (total conflict from Time 3)

$\Rightarrow$ Very quickly the conjunctive rule does not respond to new evidential reports in the fusion process!

Disjunctive rule It keeps all items of information provided by the sources

$$m_{12}^{\mathrm{Disj}}(X) = [m_1 \bigcirc\!\!\!\!\cup\, m_2](X) \triangleq \sum_{X_1, X_2 \in 2^\Theta | X_1 \cup X_2 = X} m_1(X_1) m_2(X_2)$$

- Commutative, associative, numerically robust

This rule is problematic because Θ (full ignorance) is an absorbing element for this rule

- Absorptive impact of VBA $\Rightarrow m \bigcirc\!\!\!\!\cup\, m_v = m_v$
- Fast tendency to get $m_{12\ldots n}^{\mathrm{Disj}}(\Theta) = 1$ when fusing many BBAs (directly or sequentially) which makes the result quickly useless

Implemented with Fast Möbius Transform by the product of implicability numbers [Smets 2002]

$$\begin{cases} \mathbf{m}_1 \\ \mathbf{m}_2 \end{cases} \rightarrow \begin{cases} \mathbf{b}_1 = \mathrm{mtob}(\mathbf{m}_1) \\ \mathbf{b}_2 = \mathrm{mtob}(\mathbf{m}_1) \end{cases} \rightarrow \mathbf{b}_{12} = \mathbf{b}_1 .* \mathbf{b}_2 \rightarrow \mathbf{m}_{12}^{\mathrm{Disj}} = \mathrm{btom}(\mathbf{b}_{12})$$

This fusion rule is usually used when some SoR are unreliable but we don't know which one.

Zhang rule [Zhang 1994]

modified version of DS rule including a degree of intersection between focal elements

$$m_{12}^{Z}(\emptyset) = 0 \quad \text{and} \quad m_{12}^{Z}(X) = \frac{1}{K} \sum_{X_1, X_2 \in 2^{\Theta} | X_1 \cap X_2 = X} \frac{|X_1 \cap X_2|}{|X_1| \cdot |X_2|} m_1(X_1) m_2(X_2)$$

- Commutative, not associative, not idempotent, not numerically robust

Yager rule [Yager 1987]

Transfer the total conflicting mass $m_{12}(\emptyset)$ to full ignorance Θ

$$m_{12}^{Y}(\emptyset) = 0 \quad \text{and} \quad m_{12}^{Y}(X) = \begin{cases} m_{12}^{Conj}(X), \forall X \in 2^{\Theta} \setminus \{\emptyset, \Theta\} \\ m_{12}^{Conj}(\Theta) + m_{12}^{Conj}(\emptyset), \text{for } X = \Theta \end{cases}$$

- Commutative, quasi-associative, not idempotent, neutrality of VBA
- increasing of ignorance

These rules can be directly extended for the fusion of $n > 2$ SoE

Dubois and Prade rule [Dubois Prade 1988]

Transfer every partial conflicting mass to its corresponding partial ignorance

$$m_{12}^{DP}(\emptyset) = 0 \quad \text{and} \quad m_{12}^{DP}(A) = m_{12}^{Conj}(A) + \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset \\ X_1 \cup X_2 = A}} m_1(X_1) m_2(X_2)$$

- Commutative, not associative, not idempotent
- increasing of ignorance

Florea Adaptive Combination Rule (ACR) [Florea et al. 2006]

An adaptive balance between conjunctive and disjunctive rules depending on the degree of conflict (extended in [Florea et al. 2009, Li et al. 2017])

$$m_{12}^{ACR}(\emptyset) = 0 \quad \text{and} \quad m_{12}^{ACR}(A) = \frac{1 - K_{12}}{1 - K_{12} + K_{12}^2} m_{12}^{Conj}(A) + \frac{K_{12}}{1 - K_{12} + K_{12}^2} m_{12}^{Disj}(A)$$

- Commutative, not associative, not idempotent
- Neutral impact of VBA

These rules can be directly extended for the fusion of $n > 2$ SoE

Going beyond DST with DSmT

Why going beyond DST

Dempster-Shafer Theory of belief functions is very interesting because it proposes

- Important paradigm shift for modeling epistemic uncertainty
- New appealing mathematical formalism of (quantitative) belief functions
- A combination rule for combining belief functions (DS rule) with nice properties

...but BF and DST have never been fully accepted by a part of scientific community and statisticians mainly because

- Independency between SoE has never been well established once for all
- Doubts on the validity of DS rule (normalization is controversial)
- Lack of good experimental protocol to validate DST and DS rule
- Different disputed semantic interpretations of DST and DS rule

What we have proved [Dezert Tchamova 2014]

1. the dictatorial power of DS rule to fuse equi-reliable sources of evidence.
2. the conflict (high or low) can be totally ignored through DS rule.
3. the problem of validity of DST is not due to conflict level, but the absolute truth Shafer's interpretation of propositions evaluated by SoE
4. there exists a logical contradiction in the foundations of DST

Our recommendation

BF are mathematically appealing and well defined, but use DS rule at your own risks, even in low conflicting situations.

Developed by Dezert and Smarandache in 2003–2015

- DSmT follows Shafer's paradigm of belief functions for modeling epistemic uncertainty.
- DSmT extends the belief function framework to work
  - with different models for the frame
  - with possibly imprecise quantitative belief functions
  - with qualitative belief functions expressed as labels
  - with new decision-making methods
- proposes new efficient (complicate) rules of combination, and conditioning.

Main references $\Rightarrow$ Four Free e-Books on DSmT [DSmT books]

http://www.onera.fr/fr/staff/jean-dezert
http://www.smarandache.com/DSmT.htm
http://fs.gallup.unm.edu/DSmT.htm

Shafer's interpretation

A reliable source of evidence provides an absolute truth from partial knowledge, observations, experience, etc.

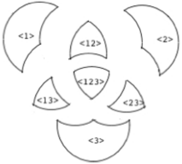
Dezert-Smarandache interpretation

A reliable source of evidence provides only a relative truth from partial knowledge, observations, experience, etc.

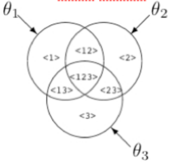
This new interpretation proposed in DSmT makes difference in the way to process belief functions.

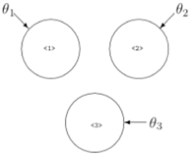
Super-power set | **Hyper-power set (Dedekind's lattice)** | **power set**

$$|2^{\Theta_{ref}} = S^{\Theta} \triangleq (\Theta, \cup, \cap, c(.))| > |D^{\Theta} = (\Theta, \cup, \cap)| > |2^{\Theta} = (\Theta, \cup)|$$

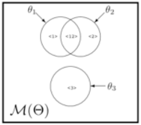


Free DSm model





hybrid DSm model



Free DSm model + Integrity constraint

General notation

The Fusion Space for the problem under concern is denoted G^{Θ}

G^{Θ} represents either 2^{Θ}, D^{Θ} or $S^{\Theta} \equiv 2^{\Theta_{refined}}$

Method of generation of hyper-power set D^Θ for $\Theta = \{\theta_1, \dots, \theta_n\}$

1. $\emptyset, \theta_1, \dots, \theta_n \in D^\Theta$
2. $\forall A, B \in D^\Theta, (A \cup B) \in D^\Theta, (A \cap B) \in D^\Theta$
3. No other elements belong to D^Θ, except those obtained by using rules 1 or 2

Hyper-power set D^Θ reduces to classical power set 2^Θ if Shafer's model for Θ holds (when all elements are mutually exclusive)

The cardinality of hyper-power sets $|D^\Theta|$ follows Dedekind's numbers sequence when cardinality $|\Theta|$ of the FoD Θ increases

Example $\Theta = \{\theta_1, \theta_2, \theta_3\} \Rightarrow |\Theta| = 3$, $|2^\Theta| = 8$ and $|D^\Theta| = 19$

$\alpha_0 \triangleq \emptyset$	$\alpha_4 \triangleq \theta_2 \cap \theta_3$	$\alpha_8 \triangleq (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$	$\alpha_{12} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3$	$\alpha_{16} \triangleq \theta_1 \cup \theta_3$
$\alpha_1 \triangleq \theta_1 \cap \theta_2 \cap \theta_3$	$\alpha_5 \triangleq (\theta_1 \cup \theta_2) \cap \theta_3$	$\alpha_9 \triangleq \theta_1$	$\alpha_{13} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2$	$\alpha_{17} \triangleq \theta_2 \cup \theta_3$
$\alpha_2 \triangleq \theta_1 \cap \theta_2$	$\alpha_6 \triangleq (\theta_1 \cup \theta_3) \cap \theta_2$	$\alpha_{10} \triangleq \theta_2$	$\alpha_{14} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1$	$\alpha_{18} \triangleq \theta_1 \cup \theta_2 \cup \theta_3$
$\alpha_3 \triangleq \theta_1 \cap \theta_3$	$\alpha_7 \triangleq (\theta_2 \cup \theta_3) \cap \theta_1$	$\alpha_{11} \triangleq \theta_3$	$\alpha_{15} \triangleq \theta_1 \cup \theta_2$	

Generalized Belief Functions (GBF)

Same definitions as Shafer's ones (when $G^\Theta = 2^\Theta$), except the Fusion Space can be now $G^\Theta = D^\Theta$, or $G^\Theta = D^\Theta$

Mass of belief in A: Degree of support precisely committed to A by the SoE

A source of evidence (SoE) about θ is represented by a generalized mass function $m^\Theta(\cdot) : G^\Theta \mapsto [0, 1]$ such that

$$m^\Theta(\emptyset) = 0 \quad \text{and} \quad \sum_{A \in G^\Theta} m^\Theta(A) = 1$$

Belief in A

$$Bel(A) \triangleq \sum_{B \in G^\Theta | B \subseteq A} m(B)$$

Plausibility of A

$$Pl(A) \triangleq \sum_{B \in G^\Theta | B \cap A \neq \emptyset} m(B)$$

Let us consider the simplest FoD defined by $\Theta = \{A, B\}$

- Working with $G^\Theta = 2^\Theta$ (power set and Shafer's model of FoD)

$$m(A) + m(B) + m(A \cup B) = 1$$

- Working with $G^\Theta = D^\Theta$ (hyper-power set and DSm free model)

$$m(A) + m(B) + m(A \cup B) + m(A \cap B) = 1$$

- Working with $G^\Theta = S^\Theta$ (super-power set)

$$m(A) + m(B) + m(A \cup B) + m(A \cap B) + m(\bar{A}) + m(\bar{B}) + m(\bar{A} \cup \bar{B}) = 1$$

Note: For simplicity of presentation, in the sequel **we will ONLY work with power-set**, that is $G^\Theta = 2^\Theta$.

PCR rules of combination

Principle of PCR rules

1. Apply the conjunctive rule
2. Identify and calculate all conflicting masses
3. Redistribute the (total or partial) conflicting masses proportionally on non-empty sets according to the integrity constraints one has for the FoD

PCR can be done in many ways [DSmT books] (Vol. 2).

Main PCR rules

- PCR rule #5 (PCR5) proposed by Smarandache & Dezert [DSmT books] (Vol. 2)
- PCR rule #6 (PCR6) proposed by Martin & Osswald [DSmT books] (Vol. 2)

PCR5=PCR6 for combining 2 SoE, but PCR5≠PCR6 when fusing more than 2 SoE
PCR6 is better than PCR5 because it is consistent with frequentist proba estimation

PCR5/6 formula for the combination of 2 BBAs

$m_{12}^{PCR5/6}(\emptyset) = 0$ and $\forall X \neq \emptyset \in 2^\Theta$

$$m_{12}^{PCR5/6}(X) = m_{12}^{Conj}(X) + \sum_{\substack{Y \in 2^\Theta \\ X \cap Y = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)}\right]$$

For general PCR5 and PCR6 formulas to fuse $s > 2$ BBAs, see [DSmT books], Vol. 2
For PCR rules with Zhang's degree of intersection, see [Smarandache Dezert 2015]

Sampling technique to approximate PCR5/6 result [Dambreville 2009]

The estimate $\hat{m}_{12}^{PCR5/6}(\cdot)$ of $m_{12}^{PCR5/6}(\cdot)$ can be obtained by the sampling process using N samples as follows

1. Repeat from $n = 1, \ldots, N$
   - draw $Y_1 \in 2^{\Theta}$ from BBA m_1, and Y_2 from m_2
   - if $Y_1 \cap Y_2 \neq \emptyset$, set $X_n = Y_1 \cap Y_2$
   - otherwise, do
     1. compute $u_1 = \frac{m_1(Y_1)}{m_1(Y_1)+m_2(Y_2)}$
     2. generate random number u uniformly distributed on $[0,1]$
     3. if $u < u_1$, set $X_n = Y_1$, otherwise set $X_n = Y_2$
2. For any $X \in 2^{\Theta}$, approximate $m_{12}^{DS}(X)$ by

$$\hat{m}_{12}^{PCR5/6}(X) = \frac{1}{N} \sum_{n=1,\ldots,N} I[X_n = X] \approx m_{12}^{PCR5/6}(X)$$

Advantages

- They exploit separately information entailed in all partial conflicts contrary to what is done in most fusion rules (except DP rule)
- They do not increase the uncertainty in the fusion of BBAs more than justified
- They work with any level of conflict between sources
- They are numerically robust to input changes
- They transfer the partial conflicting masses to the elements involved in the partial conflict proportionally to masses of **only elements involved in the partial conflict**. For instance, if $A \cap B = \emptyset$ and $m_1(A)m_2(B) > 0$ then $m_1(A)m_2(B)$ will be redistributed back only to A and B and proportionally to $m_1(A)$ and $m_2(B)$

Drawbacks

- They are commutative, not idempotent and not associative (quasi-associative only)
- Non associativity implies that the **fusion order does matter** and it impacts the fusion result. Therefore the PCR fusion must be applied globally (not sequentially) to get the best result.
- Very complicate to implement for combining altogether $S > 2$ SoE

Good news: some toolboxes implementing PCR rules are available (see later)
Basic Matlab codes for PCR5/6 rules are given in [Smarandache Dezert Tacnet 2010]

$$m_{12}^{PCR5/6}(X) = m_{12}^{Conj}(X) + \sum_{\substack{Y\in 2^\Theta \\ X\cap Y=\emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X)+m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X)+m_1(Y)}\right]$$

Very simple example $\Theta = \{A, B\}$

	A	B	$A\cup B$
$m_1(.)$	0.6	0.3	0.1
$m_2(.)$	0.2	0.3	0.5
$m_{12}(.)$	0.44	0.27	0.05

$$m_{12}(A\cap B=\emptyset) = m_1(A)m_2(B) + m_1(B)m_2(A) = 0.18 + 0.06 = 0.24$$

$$x_1/0.6 = y_1/0.3 = (x_1+y_1)/(0.6+0.3) = 0.18/0.9 = 0.2 \longrightarrow \begin{cases} x_1 = 0.6\cdot 0.2 = 0.12 \\ y_1 = 0.3\cdot 0.2 = 0.06 \end{cases}$$

$$x_2/0.2 = y_2/0.3 = (x_2+y_2)/(0.2+0.3) = 0.06/0.5 = 0.12 \longrightarrow \begin{cases} x_2 = 0.2\cdot 0.12 = 0.024 \\ y_2 = 0.3\cdot 0.12 = 0.036 \end{cases}$$

PCR5/6 result

$$\begin{cases} m_{12}^{PCR5/6}(A) = \boxed{0.44} + \boxed{0.12} + \boxed{0.024} = 0.584 \\ m_{12}^{PCR5/6}(B) = \boxed{0.27} + \boxed{0.06} + \boxed{0.036} = 0.366 \\ m_{12}^{PCR5/6}(A\cup B) = \boxed{0.05} + 0 = 0.05 \end{cases}$$

DS result

$$\begin{cases} m_{12}^{DS}(A) \approx 0.579 \\ m_{12}^{DS}(B) \approx 0.355 \\ m_{12}^{DS}(A\cup B) \approx 0.066 \end{cases}$$

One sees that the mass committed to ignorance with PCR5/6 is lower than with DST

Difference between PCR5 and PCR6

Very simple example $\Theta = \{A, B\}$

$m_1(A) = \boxed{0.6} \quad m_1(B) = 0.3 \quad m_1(A \cup B) = 0.1$

$m_2(A) = 0.2 \quad m_2(B) = \boxed{0.3} \quad m_2(A \cup B) = 0.5$

$m_3(A) = 0.7 \quad m_3(B) = \boxed{0.1} \quad m_3(A \cup B) = 0.2$

Let's consider the partial conflicting mass.

$m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$

With PCR5, one takes

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B)m_3(B)}$$

$$\frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{0.03} = \frac{0.018}{0.6 + 0.03} \approx 0.02857 \longrightarrow \begin{cases} x_A^{PCR5} = 0.60 \cdot 0.02857 \approx 0.01714 \\ x_B^{PCR5} = 0.03 \cdot 0.02857 \approx 0.00086 \end{cases}$$

With PCR6, one takes

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_B^{PCR6}}{m_2(B) + m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + (m_2(B) + m_3(B))}$$

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_B^{PCR6}}{0.3 + 0.1} = \frac{0.018}{0.6 + (0.3 + 0.1)} = 0.018 \longrightarrow \begin{cases} x_A^{PCR6} = 0.6 \cdot 0.018 = 0.0108 \\ x_B^{PCR6} = (0.3 + 0.1) \cdot 0.018 = 0.0072 \end{cases}$$

PCR6 result is more stable than PCR5 result for decision making, and PCR6 is consistent with frequentist proba estimate.

Theorem [Smarandache Dezert 2013]: When $S \geqslant 2$ SoE provide binary BBAs on 2^Θ whose total conflicting mass is 1, then PCR6 rule coincides with the averaging rule.

Random coin flip experiment $\Theta = \{\mathrm{H} = \text{Head}, \mathrm{T} = \text{Tail}\}$

Observations sequence: $Obs = \{H, H, T, H, T, H, H, T\} \Rightarrow n(H) = 5$ and $n(T) = 3$

Probas: $\hat{P}(H|Obs) = \frac{n(H)}{n} = \frac{5}{8} = m_{12\ldots8}^{Aver}(H)$ and $\hat{P}(T|Obs) = \frac{n(T)}{n} = \frac{3}{8} = m_{12\ldots8}^{Aver}(T)$

bba's \ Focal elem.	H	T
$m_1(.)$	1	0
$m_2(.)$	1	0
$m_3(.)$	0	1
$m_4(.)$	1	0
$m_5(.)$	0	1
$m_6(.)$	1	0
$m_7(.)$	1	0
$m_8(.)$	0	1

- DS rule does not work (conflict=1)
- PCR6 works because Theorem applies

$$m_{12\ldots8}^{PCR6}(H) = \frac{5}{8} \quad \text{and} \quad m_{12\ldots8}^{PCR6}(H) = \frac{3}{8}$$

- PCR5 does not work efficiently

$$m_{12\ldots8}^{PCR5}(H) = m_{12\ldots8}^{PCR5}(T) = 0.5$$

because

$$\frac{x_H}{1 \cdot 1 \cdot 1 \cdot 1 \cdot 1} = \frac{x_T}{1 \cdot 1 \cdot 1}$$
$$= \frac{m_{12\ldots8}(\emptyset)}{(1 \cdot 1 \cdot 1 \cdot 1 \cdot 1) + (1 \cdot 1 \cdot 1)} = \frac{1}{2}$$

Zadeh example with PCR5/6

$$\Theta = \{M = \text{Meningitis}, C = \text{Concussion}, T = \text{Tumor}\}$$

$$\begin{array}{lll} m_1(M) = 1-\epsilon_1 & m_1(C) = 0 & m_1(T) = \epsilon_1 \\ m_2(M) = 0 & m_2(C) = 1-\epsilon_2 & m_2(T) = \epsilon_2 \end{array}$$

$$K_{12} = m_{12}(\emptyset) = (1-\epsilon_1)(1-\epsilon_2) + (1-\epsilon_1)\epsilon_2 + \epsilon_2(1-\epsilon_1) = 1 - \epsilon_1\epsilon_2$$

$$m_{12}^{PCR5/6}(M) = \frac{(1-\epsilon_1)(1-\epsilon_2)}{(1-\epsilon_1)+(1-\epsilon_2)}(1-\epsilon_1) + \frac{(1-\epsilon_1)\epsilon_2}{(1-\epsilon_1)+\epsilon_2}(1-\epsilon_1)$$

$$m_{12}^{PCR5/6}(C) = \frac{(1-\epsilon_1)(1-\epsilon_2)}{(1-\epsilon_1)+(1-\epsilon_2)}(1-\epsilon_2) + \frac{\epsilon_1(1-\epsilon_2)}{\epsilon_1+(1-\epsilon_2)}(1-\epsilon_2)$$

$$m_{12}^{PCR5/6}(T) = \epsilon_1\epsilon_2 + \frac{(1-\epsilon_1)\epsilon_2}{(1-\epsilon_1)+\epsilon_2}\epsilon_2 + \frac{\epsilon_1(1-\epsilon_2)}{\epsilon_1+(1-\epsilon_2)}\epsilon_1$$

Bayesian BBA in **high conflict** $\epsilon_1 = \epsilon_2 = 0.1 \Rightarrow K_{12} = 1 - (0.1 \cdot 0.1) = 0.99$

$m_{12}^{DS}(T) = 1$ but $m_{12}^{PCR5/6}(M) = 0.486$ $m_{12}^{PCR5/6}(C) = 0.486$ $m_{12}^{PCR5/6}(T) = \mathbf{0.028}$

Bayesian BBA in **low conflict** $\epsilon_1 = \epsilon_2 = 0.99 \Rightarrow K_{12} = 1 - (0.99 \cdot 0.99) = 0.0199$

$m_{12}^{DS}(T) = 1$ but $m_{12}^{PCR5/6}(M) \approx 0.00015$ $m_{12}^{PCR5/6}(C) \approx 0.00015$ $m_{12}^{PCR5/6}(T) \approx \mathbf{0.9997}$

Non-Bayesian BBA $\Theta = \{A, B, C\}$, with $m_1 \neq m_2 \neq m_v$

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	b_1
C	0	$1-b_1-b_2$
$A \cup B \cup C$	0	b_2

$$K_{12} = m_{12}(\emptyset) = a(1-b_1-b_2) + (1-a)(1-b_1-b_2) = 1-b_1-b_2 > 0$$

DS result: $m_{12}^{DS}(A) = m_1(A) = a$ and $m_{12}^{DS}(A \cup B) = m_1(A \cup B) = 1-a$ which means that m_2 has no impact in DS fusion result even if the SoE are in (strong or low) conflict

PCR5/6 result:

$$m_{12}^{PCR5/6}(A) = a(b_1+b_2) + \frac{a(1-b_1-b_2)}{a+(1-b_1-b_2)} \cdot a$$

$$m_{12}^{PCR5/6}(A \cup B) = (1-a)(b_1+b_2) + \frac{(1-a)(1-b_1-b_2)}{(1-a)+(1-b_1-b_2)} \cdot (1-a)$$

$$m_{12}^{PCR5/6}(C) = \frac{a(1-b_1-b_2)}{a+(1-b_1-b_2)} \cdot (1-b_1-b_2) + \frac{(1-a)(1-b_1-b_2)}{(1-a)+(1-b_1-b_2)} \cdot (1-b_1-b_2)$$

One sees that $m_{12}^{PCR5/6} \neq m_{12}^{DS} \Rightarrow$ the source m_2 **has an impact in the fusion result**

Independent sensor (or expert) reports expressed by BBAs are fused sequentially with the conjunctive rule in Smets TBM framework

$\Theta = \{A, B, C\}$ with Shafer's model of the FoD

- Time 1: $m_1(A) = 0.4$, $m_1(B) = 0$, $m_1(C) = 0.6$
- Time 2: $m_2(A) = 0.7$, $m_1(B) = 0.3$, $m_1(C) = 0$

  TBM Conjunctive rule: $m_{12}^{Conj}(A) = 0.28$, $m_{12}^{Conj}(\emptyset) = 0.72$

  DS rule: $m_{12}^{DS}(A) = 1$

  PCR5/6 rule: $m_{12}^{PCR5/6}(A) = 0.574725$, $m_{12}^{PCR5/6}(B) = 0.111429$, $m_{12}^{PCR5/6}(C) = 0.313846$

- Time 3: $m_3(A) = 0$, $m_1(B) = 0.8$, $m_1(C) = 0.2$

  TBM Conjunctive rule: $m_{123}^{Conj}(\emptyset) = 1$

  DS rule is not applicable because of total conflict between m_3 and m_{12}^{Conj}

  PCR5/6 rule: $m_{(12)3}^{PCR5/6}(A) = 0.277490$, $m_{(12)3}^{PCR5/6}(B) = 0.545010$, $m_{(12)3}^{PCR5/6}(C) = 0.177500$

- Time 4, 5, ... k: if new evidential reports are available, one will get

  TBM Conjunctive rule is not responding because $m_{12\ldots k}^{Conj}(\emptyset) = 1$

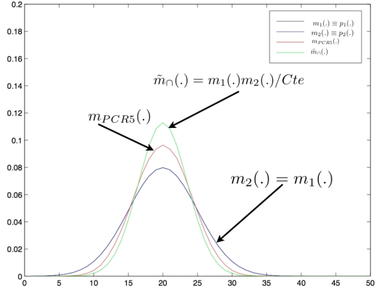
  DS rule is not applicable because of total conflict from Time 3

  PCR5/6 rule: is still responding to new evidential reports coming

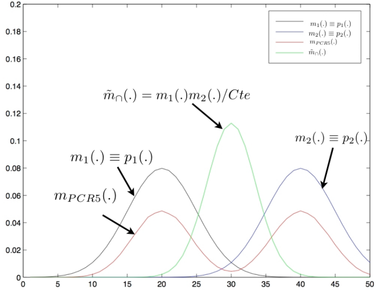
Naive Bayes fusion: $p_{12}^{Bayes}(x) = \frac{p(x|z_1)p(x|z_2)}{p(x)} \propto p_1(x)p_2(x)$ when $p(x)$ is uniform pdf

We extend PCR5 to work on a continuous frame with pdf as follows

PCR5/6 fusion: $p_{12}^{PCR5/6}(x) \triangleq p_1(x) \int_\Theta \frac{p_1(x)p_2(y)}{p_1(x)+p_2(y)} dy + \int_\Theta \frac{p_2(x)p_1(y)}{p_2(x)+p_1(y)} dy$



Fusion of Gaussian pdf – $p_1(\cdot) = p_2(\cdot)$



Fusion of Gaussian pdf – $p_1(\cdot) \neq p_2(\cdot)$

PCR5/6 rule allows to keep the modes of pdf through the fusion process

Application $\Rightarrow$ Particle Filtering for target tracking [Kirchner et al. 2007]

Approximation of a BBA by probability measures

Many methods exist, we only present the most popular – see [DSmT books] (Vol. 3)

Simplest method
Take the mass of each element of Θ and normalize, but it does not take into account partial ignorances

$$P_m(A) = \frac{m(A)}{\sum_{B\in\Theta} m(B)}$$

Method based on plausibility [Cobb Shenoy 2006]
Take the plausibility of each element of Θ and normalize, but it is inconsistent with belief interval

$$P_{Pl}(A) = \frac{Pl(A)}{\sum_{B\in\Theta} Pl(B)}$$

Pignistic probability [Smets 1990]
Redistribute the mass of partial ignorances **equally** to singletons included in them
$\Rightarrow$ higher entropy obtained with $BetP(\cdot)$

$$BetP(A) = \sum_{X\in 2^\Theta} \frac{|X\cap A|}{|A|} m(X)$$

DSmP probability [Dezert Smarandache 2008]
Redistribute mass of partial ignorances **proportionally to masses of singletons** included in them. $\epsilon > 0$ is a small parameter to prevent division by zero in some cases.
$\Rightarrow$ smaller entropy obtained with $DSmP(\cdot)$

$$DSmP_\epsilon(A) = \sum_{Y\in 2^\Theta} \frac{\sum\limits_{\substack{Z\subseteq A\cap Y\\ |Z|=1}} m(Z) + \epsilon|A\cap Y|}{\sum\limits_{\substack{Z\subseteq Y\\ |Z|=1}} m(Z) + \epsilon|Y|} m(Y)$$

$P_{Pl}(.)$ is inconsistent with belief interval! Consider $\Theta = \{A, B, C\}$, and the BBA

$$\begin{cases} m(A) = 0.2 \\ m(B \cup C) = 0.8 \end{cases} \Rightarrow \begin{cases} [Bel(A), Pl(A)] = [0.2, 0.2] \\ [Bel(B), Pl(B)] = [0, 0.8] \\ [Bel(C), Pl(C)] = [0, 0.8] \end{cases} \Rightarrow \begin{cases} P_{Pl}(A) = \frac{0.2}{0.2+0.8+0.8} \approx 0.112 < Bel(A) \\ P_{Pl}(B) = \frac{0.8}{0.2+0.8+0.8} \approx 0.444 \\ P_{Pl}(C) = \frac{0.8}{0.2+0.8+0.8} \approx 0.444 \end{cases}$$

Note: inconsistency also occurs with $P_{Bel}(.)$

Simple example for BetP and DSmP calculation

Consider $\Theta = \{A, B\}$, and $m(A) = 0.3$, $m(B) = 0.1$, $m(A \cup B) = 0.6$

$$\begin{cases} BetP(A) = m(A) + \frac{1}{2} m(A \cup B) = 0.3 + (0.6/2) = 0.6 \\ BetP(B) = m(B) + \frac{1}{2} m(A \cup B) = 0.1 + (0.6/2) = 0.4 \end{cases}$$

With DSmP the masses of singletons are used as a priori information to make the redistribution of the mass of ignorance (reinforcement principle)

$$\begin{cases} DSmP_{\epsilon=0.001}(A) = m(A) + \frac{m(A)+\epsilon}{m(A)+m(B)+2\epsilon} \cdot m(A \cup B) = 0.7492 \\ DSmP_{\epsilon=0.001}(B) = m(B) + \frac{m(B)+\epsilon}{m(A)+m(B)+2\epsilon} \cdot m(A \cup B) = 0.2508 \end{cases}$$

Shannon entropy (measure of randomness): $H(P) = -\sum_i p_i \log p_i$

$$H(DSmP) = 0.8125 \text{ bits} < H(BetP) = 0.9710 \text{ bits}$$

Thus, decision-making is made easier with DSmP because randomness is reduced

Distances between two BBAs

A strict distance metric $d : (x, y) \in \mathcal{S} \times \mathcal{S} \mapsto d(x, y) \in \mathbb{R}$ must satisfy

1. Nonnegativity: $d(x, y) \geqslant 0$;
2. Nondegeneracy: $d(x, y) = 0 \Leftrightarrow x = y$;
3. Symmetry: $d(x, y) = d(y, x)$;
4. Triangle inequality: $d(x, y) + d(y, z) \geqslant d(x, z), \forall z \in \mathcal{S}$.

References on distances : [Jousselme Maupin 2012, Han Dezert Yang 2017]

- **Tessem distance** [Tessem 1993] $\Rightarrow$ Not a strict distance metric

$$d_T(m_1, m_2) \triangleq \max_{A \subseteq \Theta}\{|BetP_1(A) - BetP_2(A)|\}$$

- **Jousselme distance** [Jousselme Grenier Bossé 2001]

$$d_J(m_1, m_2) \triangleq \sqrt{0.5 \cdot (m_1 - m_2)^T \mathbf{Jac}\,(m_1 - m_2)}$$

where the elements $Jac(A, B)$ of Jaccard's weighting matrix $\mathbf{Jac}$ are defined by

$$\mathbf{Jac}(A, B) = |A \cap B|/|A \cup B|$$

$\Rightarrow$ proved to be a strict distance metric in [Bouchard Jousselme Doré 2013]

The belief interval of $A \in 2^\Theta$ is defined as $BI(A) \triangleq [Bel(A), Pl(A)]$

- **Euclidean belief interval based distance** [Han Dezert Yang 2014]

$$d_{BI}^{E}(m_1, m_2) \triangleq \sqrt{\frac{1}{2^{|\Theta|-1}} \cdot \sum_{A \in 2^\Theta} d^I(BI_1(A), BI_2(A))^2}$$

$\Rightarrow$ proved to be a strict distance metric in [Han Dezert Yang 2014]

- **Chebyshev belief interval based distance** [Han Dezert Yang 2014]

$$d_{BI}^{C}(m_1, m_2) \triangleq \max_{A \in 2^\Theta} \left\{ d^I \left(BI_1(A), BI_2(A) \right) \right\}$$

$\Rightarrow$ proved to be a strict distance metric in [Han Dezert Yang 2014]

d^I is Wasserstein distance of interval numbers

$$d^I([a_1, b_1], [a_2, b_2]) = \sqrt{\left[\frac{a_1 + b_1}{2} - \frac{a_2 + b_2}{2}\right]^2 + \frac{1}{3}\left[\frac{b_1 - a_1}{2} - \frac{b_2 - a_2}{2}\right]^2}$$

Simple example [Han Dezert Yang 2014] $\Theta = \{\theta_1, \theta_2, \theta_3\}$

$$m_1(\theta_1) = m_1(\theta_2) = m_1(\theta_3) = 1/3$$
$$m_2(\theta_1) = m_2(\theta_2) = m_2(\theta_3) = 0.1, m_2(\Theta) = 0.7$$
$$m_3(\theta_1) = m_3(\theta_2) = 0.1, m_3(\theta_3) = 0.8$$

Results

distances	d_T	d_J	d_{BI}^E	d_{BI}^C
$d(m_1, m_2)$	0	0.4041	0.2858	0.2333
$d(m_1, m_3)$	0.4667	0.4041	0.4041	0.4667

- **Using Jousselme distance**

  The result is not very reasonable because m_2 makes no preference for choice, whereas m_3 prefers the 3rd element θ_3.
- **Using Tessem pseudo-distance**

  The result is not intuitively acceptable because m_1 is different of m_2 but $d_T(m_1, m_2) = 0$
- **Using belief interval distances d_{BI}^E or d_{BI}^C**

  The results make more sense because $d_{BI}(m_1, m_2) < d_{BI}(m_1, m_3)$

Measures of uncertainty of a belief function

How to characterize a BBA to measure the level of uncertainty it contains?

→ see the excellent survey in [Jousselme et al. 2006], with remarks in [Klir Lewis 2008]

Simplest approach

Approximate $m(\cdot)$ in a probability measure $P(\cdot)$ and use Shannon entropy $H(P)$

- it measures (approximately) the randomness in the BBA but not the imprecision (ambiguities), and many probabilistic transformations are possible
- some information is lost in the transformation $m(\cdot) \to P(\cdot)$
- these measures do not well measure uncertainty, see [Klir Lewis 2008]

Example: Ambiguity measure (or Pignistic Entropy) [Jousselme et al. 2006]

$$AM(m) \triangleq - \sum_{\theta \subseteq \Theta} BetP(\theta) \log_2(BetP(\theta))$$

Measures of discord of a belief function (entropy-alike measures)

1. Confusion [Höhle1982] $\quad Conf(m) \triangleq - \sum_{A \subseteq \Theta} m(A) \log_2(Bel(A))$
2. Dissonance [Yager 1983] $\quad Diss(m) \triangleq - \sum_{A \subseteq \Theta} m(A) \log_2(Pl(A))$
3. Discord [Klir Ramer 1990] $\quad Disc(m) \triangleq - \sum_{A \subseteq \Theta} m(A) \log_2(1 - \sum_{B \subseteq \Theta} m(B) \frac{|B-A|}{|B|})$
4. Strife [Klir Parviz 1992] $\quad Stri(m) \triangleq - \sum_{A \subseteq \Theta} m(A) \log_2(1 - \sum_{B \subseteq \Theta} m(B) \frac{|A-B|}{|A|})$

Measures of non-specificity of a belief function

Non-specificity (or ambiguity) means that some focal elements of $m(\cdot)$ are disjunctions of elements of the FoD Θ

- Non-specificity [Dubois Prade 1985, Ramer 1987]
  $NS(m) \triangleq \sum_{A\subseteq\Theta} m(A) \log_2 |A|$
  - generalization of Hartley measure of a set
  - if $m(\cdot)$ is Bayesian, $NS(m) = 0$ (the min value)
  - if $m(\cdot)$ is vacuous, $NS(m) = \log_2 |\Theta|$ (the max value)

Measures of total uncertainty of a belief function

- Aggregated uncertainty [Harmanec Klir 1994]

$$AU(m) \triangleq \max[-\sum_{\theta\in\Theta} P(\theta) \log_2 P(\theta)] \quad \text{s.t.} \quad \begin{cases} P(\theta) \in [0,1], \forall\theta\in\Theta \\ \sum_{\theta\in\Theta} P(\theta) = 1 \\ Bel(A) \leqslant \sum_{\theta\in A} P(\theta) \leqslant Pl(A), \forall A \subseteq \Theta \end{cases}$$

  $AU(m)$ is the max of Shannon entropies (upper entropy) of all probability measures $P(\cdot)$ compatible with $m(\cdot)$. It is interesting because [Abellan et al. 2008]
  - it captures both non-specificity and discord
  - it offers a probability consistency and set consistency
  - value range, monotonicity, sub-additivity and additivity for the joint BBA in Cartesian space

A new measure of total uncertainty of a belief function [Yang Han 2016]

→ based on belief Intervals which includes both the randomness and the imprecision (non-specificity)

Basic idea Given a belief interval $[Bel(A), Pl(A)]$, if this interval is farther from the most uncertain case represented by $[0, 1]$, then A has smaller uncertainty; if the belief interval of A is nearer to $[0, 1]$, then A has larger uncertainty.

Total uncertainty measure

$$TU(m) \triangleq 1 - \frac{\sqrt{3}}{|\Theta|} \sum_{\theta_i \in \Theta} d^I([Bel(\theta_1), Pl(\theta_i)], [0, 1])$$

where d^I is Wasserstein distance of interval numbers

$$d^I([a_1, b_1], [a_2, b_2]) = \sqrt{\left[\frac{a_1 + b_1}{2} - \frac{a_2 + b_2}{2}\right]^2 + \frac{1}{3}\left[\frac{b_1 - a_1}{2} - \frac{b_2 - a_2}{2}\right]^2}$$

$d^I([Bel(\theta_i), Pl(\theta_i)], [0, 1])$ reaches the bounds $1/\sqrt{3}$ when $[Bel(\theta_i), Pl(\theta_i)] = [0, 0]$ and $[Bel(\theta_i), Pl(\theta_i)] = [1, 1]$. Therefore, the normalization factor is $\frac{1}{d^I([0,0],[0,1])} = \frac{1}{d^I([1,1],[0,1])} = \sqrt{3}$

$$TU(m) \triangleq 1 - \frac{\sqrt{3}}{|\Theta|} \sum_{\theta_i \in \Theta} d^I([Bel(\theta_1), Pl(\theta_i)], [0, 1])$$

Properties of TU measure of total uncertainty [Yang Han 2016]

- $TU(m) \in [0, 1]$
- if $m(\cdot)$ is vacuous, $m(\Theta) = 1$, then $TU(m) = 1$

$$\forall \theta_i \in \Theta, [Bel(\theta_i), pl(\theta_i)] = [0, 1] \Rightarrow d^I([Bel(\theta_i), Pl(\theta_i)], [0, 1]) = 0 \Rightarrow TU(m) = 1$$

- if $m(\cdot)$ is categorical, $m(\theta_i) = 1$ for some $\theta_i \in \Theta$, then $TU(m) = 0$

$$\begin{cases} \text{for } \theta_i, [Bel(\theta_i), Pl(\theta_i)] = [1, 1] \Rightarrow d^I([1, 1], [0, 1]) = 1/\sqrt{3} \\ \forall \theta_j \neq \theta_i, [Bel(\theta_j), Pl(\theta_j)] = [0, 0] \Rightarrow d^I([0, 0], [0, 1]) = 1/\sqrt{3} \end{cases} \Rightarrow TU(m) = 0$$

- $TU(m)$ satisfies monotonicity, that is

$$\text{if } \forall A \subseteq \Theta, [Bel_1(A), Pl_1(A)] \subseteq [Bel_2(A), Pl_2(A)] \text{ then } TU(m_1) \leqslant TU(m_2)$$

Consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$ with the following BBA

$$m(\theta_1) = 0.3, \quad m(\theta_2 \cup \theta_3) = 0.5, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.2$$

Then
$$\begin{cases} m(\theta_1) = 0.3 \\ m(\theta_2 \cup \theta_3) = 0.5 \\ m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.2 \end{cases} \Rightarrow \begin{cases} [Bel(\theta_1), Pl(\theta_1)] = [0.3, 0.5] \\ [Bel(\theta_2), Pl(\theta_2)] = [0, 0.7] \\ [Bel(\theta_3), Pl(\theta_3)] = [0, 0.7] \end{cases}$$

The Wasserstein distances are

$$\begin{cases} d^I([Bel(\theta_1), Pl(\theta_1)], [0,1]) = \sqrt{\left[\frac{0.3+0.5}{2} - \frac{0+1}{2}\right]^2 + \frac{1}{3}\left[\frac{0.5-0.3}{2} - \frac{1-0}{2}\right]^2} = 0.2517 \\ d^I([Bel(\theta_2), Pl(\theta_2)], [0,1]) = \sqrt{\left[\frac{0+0.7}{2} - \frac{0+1}{2}\right]^2 + \frac{1}{3}\left[\frac{0.7-0}{2} - \frac{1-0}{2}\right]^2} = 0.1732 \\ d^I([Bel(\theta_3), Pl(\theta_3)], [0,1]) = \sqrt{\left[\frac{0+0.7}{2} - \frac{0+1}{2}\right]^2 + \frac{1}{3}\left[\frac{0.7-0}{2} - \frac{1-0}{2}\right]^2} = 0.1732 \end{cases}$$

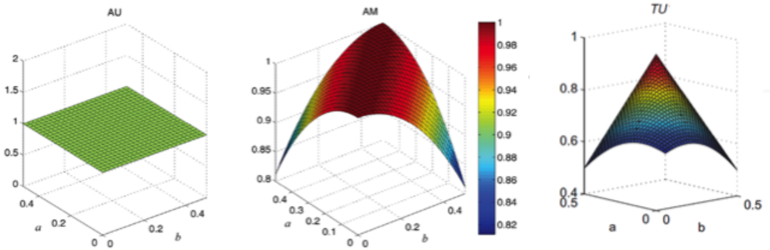
because Wasserstein distance between intervals $[a_1, b_1]$ and $[a_2, b_2]$ is defined by

$$d^I([a_1, b_1], [a_2, b_2]) = \sqrt{\left[\frac{a_1 + b_1}{2} - \frac{a_2 + b_2}{2}\right]^2 + \frac{1}{3}\left[\frac{b_1 - a_1}{2} - \frac{b_2 - a_2}{2}\right]^2}$$

Therefore, $TU(m) = 1 - \frac{\sqrt{3}}{3}(0.2517 + 0.1732 + 0.1732) = 0.6547$

More examples with applications in [Yang Han 2016]

Suppose that the FOD $\Theta = \{\theta_1, \theta_2\}$. A BBA over Θ is $m(\{\theta_1\}) = a$, $m(\{\theta_2\}) = b$, $m(\{\Theta\}) = 1 - a - b$, where $a, b \in [0, 0.5]$



$$Bel(\{\theta_1\}) = a \leqslant P(\theta_1) \leqslant 1 - b = Pl(\{\theta_1\});$$
$$Bel(\{\theta_2\}) = b \leqslant P(\theta_2) \leqslant 1 - a = Pl(\{\theta_2\});$$
$$Bel(\Theta) = 1 - a - b \leqslant P(\theta_1) \leqslant 1 = Pl(\Theta);$$

AM reaches its maximum when *a*=*b*, because when *a*=*b*, the pignistic probability is uniformly distributed. Counter-intuitive!

AU tries to find a p.m.f. with the maximum Shannon entropy. The uniformly distributed $P(\theta_1)=P(\theta_2)=0.5$ is always picked up, since it always satisfies the constraints above

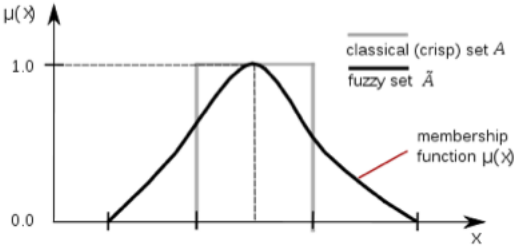
$$m_1(\{\theta_1\}) = m_1(\{\theta_2\}) = 0.5$$
$$m_2(\{\theta_1\}) = m_2(\{\theta_2\}) = 0.25, m_2(\Theta) = 0.5$$

BBA construction from FMF

How to construct a BBA from a Fuzzy Membership Function (FMF)?

Fuzzy sets and fuzzy membersip function

Definition: A fuzzy set, denoted by $A \subseteq \Theta$, is defined by a fuzzy membership function (FMF) $\mu_A(\theta) : \Theta \mapsto [0, 1]$, which quantifies the grade of membership of element θ of the fuzzy set A.



The FMF is a generalization of the characteristic function in classical set and can take its values in the interval $[0, 1]$.

Relationship between FMF and BBA

Theorem: If $\Theta_+ = \{\theta_1, \ldots, \theta_n\}$ is countable, the necessary and sufficient condition for $\mu(\cdot)$ to be a plausibility function is:

$$\sum_{i=1}^{n} \mu(\theta_i) \geqslant 1$$

The necessary and sufficient condition for $\mu(\cdot)$ to be a belief function is

$$\sum_{i=1}^{n} \mu(\theta_i) \leqslant 1$$

where Θ_+ defined as the set $\Theta_+ = \{\theta | \mu(\theta) > 0\}$

Proposition: Any membership function $\mu(\theta)$, defining on Θ a fuzzy set, can be viewed as the restriction to singletons θ either of a plausibility measure $\mu(\theta) = Pl(\theta)$, or a belief function $\mu(\theta) = Bel(\theta)$.

According to the above, the transformation of BBA into FMF can be obtained.
What about the reverse direction?

Multi-answer problems when transforming a FMF into a BBA

Suppose that $\sum_{\theta\in\Theta}\mu(\theta)\geqslant 1$ with $|\Theta|=n$, then the FMF is equivalentt to the one-point plausibility. For the frame of discernment, there may exist at most $2^{|\Theta|}-1$ subsets which are not empty. That is

$$
\begin{aligned}
&A_i=\{\theta_i\}\subseteq\Theta, \qquad m(A_i)\geqslant 0,\ i=1,2,\ldots,n\\
&A_{ij}=\{\theta_i,\theta_j\}\subseteq\Theta, \qquad m(A_{ij})\geqslant 0,\ i\leqslant j,\ i,j=1,2,\ldots,n\\
&A_{ijk}=\{\theta_i,\theta_j,\theta_k\}\subseteq\Theta, \qquad m(A_{ijk})\geqslant 0,\ i\leqslant j\leqslant k\\
&\quad\vdots\\
&A_{12\ldots n}=\Theta, \qquad m(\Theta)\geqslant 0
\end{aligned}
$$

The problem consists of $n+1$ linear equations given by

$$
\begin{cases}
m(A_1)+\sum_j m(A_{1j})+\sum_{j,k} m(A_{1jk})+\ldots+m(\Theta)=\mu(\theta_1)\\
\vdots\\
m(A_n)+\sum_j m(A_{nj})+\sum_{j,k} m(A_{njk})+\ldots+m(\Theta)=\mu(\theta_n)\\
\sum_i m(A_i)+\sum_{i,j} m(A_{ij})+\sum_{i,j,k} m(A_{ijk})+\ldots+m(\Theta)=1
\end{cases}
$$

The 2^n-1 focal elements' mass values are unknown variables to find, but we have only $n+1$ linear equations $\Rightarrow$ solution to build a BBA from a FMF is **not unique**

Transformation of FMF into a BBA [Han 2016]

Given $\mu(\theta_i) \in [0,1]$, $\forall \theta_i \in \Theta$, $i = 1, \ldots, n$, if $\sum_{i=1}^{n} \mu(\theta_i) \geqslant 1$, the FMF is equivalent to the plausibility for one-point (singleton). A BBA can be obtained by solving the following maximization problem.

Find the BBA $m(\cdot)$ such that

$$\underset{m}{\text{Max}} \left\{ J(m) = -\sum_{i=1}^{n} \Big(\sum_{B, \forall \theta_i \in B \subseteq \Theta} \frac{m(B)}{|B|} \log_2 \Big(\sum_{B, \forall \theta_i \in B \subseteq \Theta} \frac{m(B)}{|B|} \Big) \Big) \right\}$$

$$\text{s.t.} \begin{cases} \sum\limits_{B \in \mathcal{P}(\Theta)} m(B) = 1 \\ \sum\limits_{B, \forall \{\theta_i\} \cap B \neq \emptyset} m(B) = \mu(\theta_i), \forall \theta_i \in \Theta \\ 0 \leq m(B) \leq 1 \end{cases}$$

where $\mathcal{P}(\Theta) = 2^{\Theta}$ (i.e. the power set of the FoD Θ)

Example - part 1 [Han 2016]

Sequence number	Focal element
A_1	$\{\theta_3\}$
A_2	$\{\theta_2\}$
A_3	$\{\theta_1\}$
A_4	$\{\theta_3, \theta_2\}$
A_5	$\{\theta_3, \theta_1\}$
A_6	$\{\theta_2, \theta_1\}$
A_7	$\{\theta_3, \theta_2, \theta_1\}$

$$\mathrm{BetP}_m(\theta_1) = \frac{m(A_3)}{1} + \frac{m(A_5)}{2} + \frac{m(A_6)}{2} + \frac{m(A_7)}{3};$$

$$\mathrm{BetP}_m(\theta_2) = \frac{m(A_2)}{1} + \frac{m(A_4)}{2} + \frac{m(A_6)}{2} + \frac{m(A_7)}{3};$$

$$\mathrm{BetP}_m(\theta_3) = \frac{m(A_1)}{1} + \frac{m(A_4)}{2} + \frac{m(A_5)}{2} + \frac{m(A_7)}{3}.$$

Objective Function

$$\begin{aligned}
\mathrm{AM}(m) = \\
-\Big\{ & \left(\frac{m(A_3)}{1} + \frac{m(A_5)}{2} + \frac{m(A_6)}{2} + \frac{m(A_7)}{3}\right) \cdot \log_2\left(\frac{m(A_3)}{1} + \frac{m(A_5)}{2} + \frac{m(A_6)}{2} + \frac{m(A_7)}{3}\right) \\
& + \left(\frac{m(A_2)}{1} + \frac{m(A_4)}{2} + \frac{m(A_6)}{2} + \frac{m(A_7)}{3}\right) \cdot \log_2\left(\frac{m(A_3)}{1} + \frac{m(A_5)}{2} + \frac{m(A_6)}{2} + \frac{m(A_7)}{3}\right) \\
& + \left(\frac{m(A_1)}{1} + \frac{m(A_4)}{2} + \frac{m(A_5)}{2} + \frac{m(A_7)}{3}\right) \cdot \log_2\left(\frac{m(A_1)}{1} + \frac{m(A_4)}{2} + \frac{m(A_5)}{2} + \frac{m(A_7)}{3}\right) \Big\}
\end{aligned}$$

Example - part 2 [Han 2016]

For example, the cardinality of the FOD is 3 and the corresponding FMF is $\mu(\theta_1) = 1, \mu(\theta_2) = 0.6, \mu(\theta_3) = 0.2$. Obviously, there exists $\sum_{i=1,\ldots,n} \mu(\theta_i) \geq 1$. All the constraints are as follows (see the details of focal elements in Table 1):

Constraints

$$\begin{cases} 1 = \mu(\theta_1) = m(A_3) + m(A_5) + m(A_6) + m(A_7); \\ 0.6 = \mu(\theta_2) = m(A_2) + m(A_4) + m(A_6) + m(A_7); \\ 0.2 = \mu(\theta_3) = m(A_1) + m(A_4) + m(A_5) + m(A_7); \\ \sum_{i=1}^{7} m(A_i) = 1; \\ 0 \leq m(A_i) \leq 1. \end{cases}$$

Objective Function

$$\mathrm{AM}(m) =$$

$$-\Big\{\big(\tfrac{m(A_3)}{1} + \tfrac{m(A_5)}{2} + \tfrac{m(A_6)}{2} + \tfrac{m(A_7)}{3}\big) \cdot \log_2\big(\tfrac{m(A_3)}{1} + \tfrac{m(A_5)}{2} + \tfrac{m(A_6)}{2} + \tfrac{m(A_7)}{3}\big)$$

$$+\big(\tfrac{m(A_2)}{1} + \tfrac{m(A_4)}{2} + \tfrac{m(A_6)}{2} + \tfrac{m(A_7)}{3}\big) \cdot \log_2\big(\tfrac{m(A_3)}{1} + \tfrac{m(A_5)}{2} + \tfrac{m(A_6)}{2} + \tfrac{m(A_7)}{3}\big)$$

$$+\big(\tfrac{m(A_1)}{1} + \tfrac{m(A_4)}{2} + \tfrac{m(A_5)}{2} + \tfrac{m(A_7)}{3}\big) \cdot \log_2\big(\tfrac{m(A_1)}{1} + \tfrac{m(A_4)}{2} + \tfrac{m(A_5)}{2} + \tfrac{m(A_7)}{3}\big)\Big\}$$

$$m(A_3) = m(\{\theta_1\}) = 0.2;$$
$$m(A_5) = m(\{\theta_3, \theta_1\}) = 0.2;$$
$$m(A_6) = m(\{\theta_2, \theta_1\}) = 0.6.$$

Unique BBA obtained

Working with admissible imprecise BBA

Operation on sets of numbers [Dezert Smarandache 2006, DSmT books], Vol. 2

- **Addition**: $\mathcal{X}_1 \boxplus \mathcal{X}_2 = \mathcal{X}_2 \boxplus \mathcal{X}_1 \triangleq \{x | x = x_1 + x_2, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\}$
- **Multiplication**: $\mathcal{X}_1 \boxdot \mathcal{X}_2 = \mathcal{X}_2 \boxdot \mathcal{X}_1 \triangleq \{x | x = x_1 \cdot x_2, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\}$
- **Division**: defined for case where $0 \notin \mathcal{X}_2$, $\inf(\mathcal{X}_2) \neq 0$, $\sup(\mathcal{X}_2) \neq 0$

$$\mathcal{X}_1 \boxdiv \mathcal{X}_2 \triangleq \{x | x = x_1 \div x_2, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\}$$

Imprecise BBA

- Imprecise BBA is a BBA whose each mass of FE is an interval of numbers.
  Example: $\Theta = \{\theta_1, \theta_2\}$, $m^{Imp}(\theta_1) = [0.2, 0.3]$, $m^{Imp}(\theta_2) = (0.4, 0.5)$ → improper
  Because $m(\emptyset) = 0$, then $m^{Imp}(\emptyset) = [0, 0]$ (degenerate interval)
- General imprecise BBA is a BBA whose each mass of FE is a disjunction of intervals and sets of numbers
  Example: $m^{Imp}(\theta_1) = [0.1, 0.2] \cup \{0.3\}$, $m^{Imp}(\theta_2) = \{0.4, 0.6\} \cup (0.1, 0.2]$

Imprecise admissible BBA

$m^{Imp}(\cdot)$ is **admissible** if $\forall A \in \mathcal{F}(m^{Imp}), \exists m(A) \in m^{Imp}(A)$, s.t. $\sum_{A \in \mathcal{F}(m^{Imp})} m(A) = 1$

Example:

$$\begin{cases} m^{Imp}(\theta_1) = [0.1, 0.2] \cup \{0.3\} \\ m^{Imp}(\theta_2) = (0.4, 0.6) \cup [0.7, 0.8] \end{cases} \rightarrow \exists \begin{cases} m(\theta_1) = 0.3 \in m^{Imp}(\theta_1) \\ m(\theta_2) = 0.7 \in m^{Imp}(\theta_2) \end{cases} \rightarrow \text{s.t. } m(\theta_1) + m(\theta_2) = 1$$

Working with imprecise admissible BBA needs operators on sets of numbers

$\Theta = \{\theta_1, \theta_2\}$ with Shafer model for the FoD

BBA \FE	θ_1	θ_2
$m_1^{Imp}(\cdot)$	$[0.2, 0.3]$	$[0.6, 0.8]$
$m_2^{Imp}(\cdot)$	$[0, 4, 0.7]$	$[0.5, 0.6]$

Conjunctive rule gives $m_{12}(\theta_1) = [0.08, 0.21]$ and $m_{12}(\theta_2) = [0.30, 0.48]$

$$\begin{aligned} K_{12} = m_{12}(\emptyset) &= [m_1^{Imp}(\theta_1) \boxdot m_2^{Imp}(\theta_2)] \boxplus [m_1^{Imp}(\theta_2) \boxdot m_2^{Imp}(\theta_1)] \\ &= ([0.2, 0.3] \boxdot [0.5, 0.6]) \boxplus ([0.4, 0.7] \boxdot [0.6, 0.8]) = [0.34, 0.74] \end{aligned}$$

PCR5/6 rule gives [DSmT books], Vol. 2, pp. 52-53

$$\frac{x_{\theta_1}}{[0.2,0.3]} = \frac{x_{\theta_2}}{[0.5,0.6]} = \frac{[0.2,0.3] \boxdot [0.5,0.6]}{[0.2,0.3] \boxplus [0.5,0.6]} = \frac{[0.10,0.18]}{[0.7,0.9]} \Rightarrow \begin{cases} x_{\theta_1} \approx [0.022, 0.077] \\ x_{\theta_2} \approx [0.055, 0.154] \end{cases}$$

$$\frac{y_{\theta_1}}{[0.4,0.7]} = \frac{y_{\theta_2}}{[0.6,0.8]} = \frac{[0.4,0.7] \boxdot [0.6,0.8]}{[0.4,0.7] \boxplus [0.6,0.8]} = \frac{[0.24,0.56]}{[1,1.5]} \Rightarrow \begin{cases} y_{\theta_1} \approx [0.064, 0.392] \\ y_{\theta_2} \approx [0.096, 0.448] \end{cases}$$

Therefore $m_{12}^{PCR5/6}(\emptyset) = [0, 0]$ and

$$m_{12}^{PCR5/6}(\theta_1) = m_{12}(\theta_1) \boxplus x_{\theta_1} \boxplus y_{\theta_1} \approx [0.166, 0.679]$$

$$m_{12}^{PCR5/6}(\theta_2) = m_{12}(\theta_2) \boxplus x_{\theta_2} \boxplus y_{\theta_2} \approx [0.451, 1]$$

Compute divisions at the end to get tightest bounds. Use Interval Arithmetic toolboxes.

Working with qualitative BBA

Linguistic labels $L = \{L_{\min} = L_0, L_1, \ldots, L_n, L_{\max} = L_{n+1}\}$ with $L_0 < L_1 < \ldots L_{n+1}$

Assuming linguistically equidistant labels of L, we make an isomorphism between $L = \{L_0, L_1, L_2, \ldots, L_{n+1}\}$ and $\{0 = \frac{0}{n+1}, \frac{1}{n+1}, \frac{2}{n+1}, \ldots, 1 = \frac{n+1}{n+1}\}$

Operators on linguistic labels [DSmT books] (Vol. 2, Chap. 10) & [Martin et al. 2008]

q -addition and subtraction

$$L_i + L_j = \begin{cases} L_{i+j} & \text{if } i+j < n+1 \\ L_{n+1} & \text{if } i+j \geqslant n+1 \end{cases}$$

$$L_i - L_j = \begin{cases} L_{i-j} & \text{if } i \geqslant j \\ -L_{j-i} & \text{if } i < j \end{cases}$$

with a scalar

$$\begin{cases} L_i + r = r + L_i = L_{[i+r(n+1)]} \\ L_i - r = L_{[i-r(n+1)]} \end{cases}$$

q -multiplication and division

$L_i \cdot L_j = L_{[(i \cdot j)/(n+1)]}$ with $[x]$ = closest integer to x

$$L_i / L_{j \neq 0} = \begin{cases} L_{[(i/j)\cdot(n+1)]} & \text{if } [(i/j)\cdot(n+1)] < n+1 \\ L_{n+1} & \text{otherwise} \end{cases}$$

by a scalar

$$a \cdot L_i = \frac{a \cdot i}{n+1} \approx \begin{cases} L_{[a \cdot i]} & \text{if } [a \cdot i] \geqslant 0 \\ L_{-[a \cdot i]} & \text{otherwise} \end{cases}$$

No matter how many operations on labels we have, the most accurate result is obtained if we do only one approximation, and that one should be just at the very end.

Linguistic labels $L = \{L_{\text{min}} = L_0, L_1, \dots, L_n, L_{\text{max}} = L_{n+1}\}$ with $L_0 \prec L_1 \prec \dots L_{n+1}$

We can also work with refined labels (labels having non integer index) to get more exact results [DSmT books], Vol. 3, Chap. 2

Basic idea: Use real index of label to be more precise, for instance $L_{\frac{3}{2}} = L_{1.5}$ to express a label between L_1 and L_2

Operations with refined linguistic labels

- q-addition of refined labels

$$L_a + L_b = L_{a+b}$$

- q-multiplication of refined labels

$$L_a \cdot L_b = L_{a \cdot b/(n+1)}$$

- q-division of refined labels (if $b \neq 0$)

$$L_a \div L_b = L_{(a/b)(n+1)}$$

More operations presented in [DSmT books], Vol. 3, Chap. 2

Example $L = \{L_0, L_1, L_2, L_3, L_4, L_5\} \Leftrightarrow \{0, L_1 \equiv 0.2, L_2 \equiv 0.4, L_3 \equiv 0.6, L_4 \equiv 0.8, 1\}$

Product using labels: $L_2 \cdot L_3 = L_{[(2.3)/5]} = L_{[6/5]} = L_{[1.2]} = L_1$
Product using numbers: $0.4 \cdot 0.6 = 0.24 \approx 0.2 = L_1$

Product using labels: $L_3 \cdot L_3 = L_{[(3.3)/5]} = L_{[9/5]} = L_{[1.8]} = L_2$
Product using numbers: $0.6 \cdot 0.6 = 0.36 \approx 0.4 = L_2$

Qualitative BBA $qm(.) : 2^\Theta \mapsto L = \{L_0, L_1, \ldots, L_n, L_{n+1}\}$

Quasi-normalization conditions

$$qm(\emptyset) = L_0 \quad \text{and} \quad \sum_{X \in 2^\Theta} qm(X) = \sum_k L_{i_k} = L_{n+1}$$

Qualitative rules of combination

- All previous rules of combinations (as well as BBA transformations) can be done with qualitative BBA thanks to operators on linguistic labels [Martin et al. 2008].
- Extension for working with imprecise qualitative BBAs is proposed in [Li Dai Dezert Smarandache 2010]

Example drawn from [Martin et al. 2008]

$L = \{L_0, L_1 = \text{very poor}, L_2 = \text{poor}, L_3 = \text{good}, L_4 = \text{very good}, L_5 = \text{very very good}, L_6\}$
$\Leftrightarrow \{0, 1/6 \approx 0.166, 2/6 \approx 0.333, 3/6 = 0.5, 4/6 \approx 0.666, 5/6 \approx 0.833, 1\}$

$\Theta = \{A, B\}$ satisfying Shafer's model, and the two qualitative normalized BBAs

$$qm_1(A) = L_1, qm_1(B) = L_3, qm_1(A \cup B) = L_2$$
$$qm_2(A) = L_4, qm_2(B) = L_1, qm_2(A \cup B) = L_1$$

Conjunctive rule (with refined labels calculus)

$$\begin{aligned}
K_{12} = qm_{12}(\emptyset) &= qm_1(A)qm_2(B) + qm_1(B)qm_2(A) \\
&= L_1L_1 + L_3L_4 = L_{\frac{1\cdot 1}{6}} + L_{\frac{3\cdot 4}{6}} = L_{\frac{1+12}{6}} = L_{\frac{13}{6}} = L_{2.166} \approx L_2 \\
qm_{12}(A) &= qm_1(A)qm_2(A) + qm_1(A)qm_2(A \cup B) + qm_2(A)qm_1(A \cup B) \\
&= L_1L_4 + L_1L_1 + L_4L_2 = L_{\frac{1\cdot 4}{6}} + L_{\frac{1\cdot 1}{6}} + L_{\frac{4\cdot 2}{6}} = L_{\frac{4+1+8}{6}} = L_{\frac{13}{6}} = L_{2.166} \approx L_2 \\
qm_{12}(B) &= qm_1(B)qm_2(B) + qm_1(B)qm_2(A \cup B) + qm_2(B)qm_1(A \cup B) \\
&= L_3L_1 + L_3L_1 + L_1L_2 = L_{\frac{3\cdot 1}{6}} + L_{\frac{3\cdot 1}{6}} + L_{\frac{1\cdot 2}{6}} = L_{\frac{3+3+2}{6}} = L_{\frac{8}{6}} = L_{1.333} \approx L_1 \\
qm_{12}(A \cup B) &= qm_1(A \cup B)qm_2(A \cup B) = L_2L_1 = L_{\frac{2\cdot 1}{6}} = L_{\frac{2}{6}} = L_{0.333} \approx L_0
\end{aligned}$$

With refined labels, qm_{12} **is normalized**: $L_{\frac{13}{6}} + L_{\frac{13}{6}} + L_{\frac{8}{6}} + L_{\frac{2}{6}} = L_{\frac{36}{6}} = L_6 = L_{max}$
With approximate labels, qm_{12} is not normalized: $L_2 + L_2 + L_1 + L_0 = L_5 \neq L_6 = L_{max}$

$$qm_1(A) = L_1, qm_1(B) = L_3, qm_1(A \cup B) = L_2$$
$$qm_2(A) = L_4, qm_2(B) = L_1, qm_2(A \cup B) = L_1$$

PCR5/6 rule (with refined labels calculus)

Partial conflict $qm_1(A)qm_2(B) = L_1L_1 = L_{\frac{1\cdot 1}{6}} = L_{\frac{1}{6}}$ goes back to A and to B with

$$\frac{x_A}{L_1} = \frac{x_B}{L_1} = \frac{L_1L_1}{L_1 + L_1} = \frac{L_{\frac{1\cdot 1}{6}}}{L_2} = L_{(\frac{1}{6}\div 2)\cdot 6} = L_{\frac{1}{2}} \Rightarrow \begin{cases} x_A = L_1L_{\frac{1}{2}} = L_{(1\cdot\frac{1}{2})/6} = L_{\frac{1}{12}} \approx L_{0.083} \\ x_B = L_1L_{\frac{1}{2}} = L_{(1\cdot\frac{1}{2})/6} = L_{\frac{1}{12}} \approx L_{0.083} \end{cases}$$

Partial conflict $qm_2(A)qm_1(B) = L_4L_3 = L_{\frac{4\cdot 3}{6}} = L_{\frac{12}{6}}$ goes back to A and to B with

$$\frac{y_A}{L_4} = \frac{y_B}{L_3} = \frac{L_4L_3}{L_4 + L_3} = \frac{L_{\frac{12}{6}}}{L_7} = L_{(\frac{12}{6}\div 7)\cdot 6} = L_{\frac{12}{7}} \Rightarrow \begin{cases} y_A = L_4L_{\frac{12}{7}} = L_{(4\cdot\frac{12}{7})/6} = L_{\frac{8}{7}} \approx L_{1.142} \\ y_B = L_3L_{\frac{12}{7}} = L_{(3\cdot\frac{12}{7})/6} = L_{\frac{6}{7}} \approx L_{0.857} \end{cases}$$

Finally, one gets $qm^{PCR5/6}(\emptyset) = L_0$ and

$$qm^{PCR5/6}(A) = qm_{12}(A) + x_A + y_A = L_{\frac{13}{6}} + L_{\frac{1}{12}} + L_{\frac{8}{7}} = L_{\frac{285}{84}} \approx L_{3.392} \approx L_3$$
$$qm^{PCR5/6}(B) = qm_{12}(B) + x_B + y_B = L_{\frac{8}{6}} + L_{\frac{1}{12}} + L_{\frac{6}{7}} = L_{\frac{191}{84}} \approx L_{2.273} \approx L_2$$
$$qm^{PCR5/6}(A \cup B) = qm_{12}(A \cup B) = L_{\frac{2}{6}} = L_{\frac{28}{84}} \approx L_{0.333} \approx L_0$$

With refined labels one has $L_0 + L_{\frac{285}{84}} + L_{\frac{191}{84}} + L_{\frac{28}{84}} = L_{\frac{504}{84}} = L_6 = L_{max}$

Part II

Decision-Making Support with Belief Functions

Outline of Part 2

Classical decision-making methods with belief functions

Decision-making problem (DMP) FoD $\Theta = \{\theta_1, \dots, \theta_n\}$ = set of possible solutions

Knowing a BBA $m(\cdot)$ over 2^Θ, how should I make my decision δ based on $m(\cdot)$?

In the classical DMP, we restrict $\delta \in \Theta$, i.e. the best decision $\hat{\theta}$ is a singleton of 2^Θ.

Classical DM methods

- Pessimistic Decision-Making attitude: Maximum of belief strategy

$$m(\cdot) \rightarrow Bel(\cdot) \quad \text{and} \quad \delta = \hat{\theta} = \arg\max_{\theta_i \in \Theta} Bel(\theta_i)$$

- Optimistic Decision-Making attitude: Maximum of plausibility strategy

$$m(\cdot) \rightarrow Pl(\cdot) \quad \text{and} \quad \delta = \hat{\theta} = \arg\max_{\theta_i \in \Theta} Pl(\theta_i)$$

- Compromise Decision-Making attitude: Maximum of probability strategy

$$m(\cdot) \rightarrow P(\cdot) \quad \text{and} \quad \delta = \hat{\theta} = \arg\max_{\theta_i \in \Theta} P(\theta_i)$$

where $P(\cdot) \in [Bel(\cdot), Pl(\cdot)]$ is a (subjective) proba measure approximated from the BBA $m(\cdot)$, typically obtained with a lossy transformation like BetP, or DSmP

Decision-making based on distances [Han Dezert Yang 2014, Dezert et al. 2016]

A better theoretical approach for decision-making is to use a strict distance metric $d(\cdot,\cdot)$ between two BBAs and to make the decision by

$$\delta = \hat{X} = \arg\min_{X \in \mathcal{X}} d(m, m_X)$$

$\mathcal{X} = \{\text{admissible} X, X \in 2^\Theta\}$ is the set of possible admissible decisions we consider. For instance, if δ must be a singleton, then $\mathcal{X} = \Theta = \{\theta_1, \ldots, \theta_n\}$.

m_X is the BBA focused on X defined by $m_X(Y) = 0$ if $Y \neq X$, and $m_X(Y) = 1$ if $Y = X$

Few strict distance metrics are possible

- Jousselme distance: $d_J(m_1, m_2) \triangleq \sqrt{0.5 \cdot (m_1 - m_2)^T \mathbf{Jac}\, (m_1 - m_2)}$
- Euclidean d_{BI} distance: $d^E_{BI}(m_1, m_2) \triangleq \sqrt{\frac{1}{2^{|\Theta|-1}} \cdot \sum_{A \in 2^\Theta} d^I(BI_1(A), BI_2(A))^2}$
- Chebyshev d_{BI} distance: $d^C_{BI}(m_1, m_2) \triangleq \max_{A \in 2^\Theta} \{d^I(BI_1(A), BI_2(A))\}$

In practice, we recommend to use $d^E_{BI}(m_1, m_2)$ [Han Dezert Yang 2017]

Quality of the decision $$q(\hat{X}) = 1 - \frac{d_{BI}(m, m_{\hat{X}})}{\sum_{X \in \mathcal{X}} d_{BI}(m, m_X)} \in [0, 1]$$

Higher is $q(\hat{X})$ more trustable is the decision $\delta = \hat{X}$

General mono-criteria decision-making problem

General mono-criteria decision-making problem

How to make a decision among several possible choices, based on some contexts ?

Problem modeling

$q \geqslant 2$ alternatives (choices) $\mathcal{A} = \{A_1, \dots, A_q\}$
$n \geqslant 1$ states of nature (contexts) $\mathcal{S} = \{S_1, \dots, S_n\}$

$$\mathbf{C} \triangleq \begin{array}{cc} & \begin{array}{ccccc} S_1 & \dots & S_j & \dots & S_n \end{array} \\ \begin{array}{c} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{array} & \begin{bmatrix} C_{11} & \dots & C_{1j} & \dots & C_{1n} \\ & & \vdots & & \\ C_{i1} & \dots & C_{ij} & \dots & S_{in} \\ & & \vdots & & \\ C_{q1} & \dots & C_{qj} & \dots & C_{qn} \end{bmatrix} \end{array}$$

$\mathbf{C}$ is the benefit (payoff) matrix of the problem under consideration

Investment company example

An investment company wants to invest a given amount of money in the best option $A^* \in \mathcal{A} = \{A_1, A_2, A_3\}$, where A_1 = car company, A_2 = food company, and A_3 = computer company. Several scenarios (states of nature) S_i are identified depending on national economical situation predictions, which provide the elements of the payoff matrix $\mathbf{C}$ according to a given criteria (growth analysis criterion by example).

Several decision-making frameworks are possible

- **Decision under certainty**
  If we know the **true** state of nature is S_j, take as decision $\delta = A^*$ with
  $$A^* = A_{i*} \quad \text{with} \quad i^* = \arg\max_i \{C_{ij}\}$$
- **Decision under risk**
  If we know **all** probabilities $p_j = P(S_j)$ of the states of nature, compute the expected benefit $E[C_i] = \sum_j p_j C_{ij}$ of each A_i and take as decision $\delta = A^*$ with
  $$A^* = A_{i*} \quad \text{with} \quad i^* = \arg\max_i \{E[C_i]\}$$
- **Decision under ignorance**
  If **we don't know the probabilities** $p_j = P(S_j)$ of the states of nature, use OWA (**O**rdered **W**eighted **A**veraging) approach [Yager 1988], or Cautious-OWA [Tacnet Dezert 2011], or Fuzzy-Cautious-OWA [Han Dezert Tacnet Han 2012]
- **Decision under uncertainty**
  If we have **only a BBA** over the states of the nature $\mathcal{S} = \{S_1, \ldots, S_n\}$ defined on the power set $2^{\mathcal{S}}$, we can use Yager extended OWA approach.

$$\mathbf{C} \triangleq \begin{array}{c} \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{array} \begin{array}{c} \begin{array}{ccccc} S_1,p_1 & \dots & S_j,p_j & \dots & S_n,p_n \end{array} \\ \begin{bmatrix} C_{11} & \dots & C_{1j} & \dots & C_{1n} \\ & & \vdots & & \\ C_{i1} & \dots & C_{ij} & \dots & C_{in} \\ & & \vdots & & \\ C_{q1} & \dots & C_{qj} & \dots & C_{qn} \end{bmatrix} \end{array} \Rightarrow E[\mathbf{C}] = \begin{bmatrix} E[C_1] = \sum_j p_j C_{1j} \\ \vdots \\ E[C_i] = \sum_j p_j C_{ij} \\ \vdots \\ E[C_q] = \sum_j p_j C_{qj} \end{bmatrix}$$

Decision: A^* is the chosen alternative corresponding to highest expected benefit.

Example

$$\mathbf{C} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccc} S_1, p_1 = 1/4 & S_2, p_2 = 1/4 & S_3, p_3 = 1/2 \end{array} \\ \begin{bmatrix} 16 & 12 & 20 \\ 32 & 4 & 6 \\ 12 & 20 & 4 \\ 40 & 4 & 8 \end{bmatrix} \end{array} \Rightarrow \begin{bmatrix} E[C_1] = (1/4)16 + (1/4)12 + (1/2)20 = 17 \\ E[C_2] = (1/4)32 + (1/4)4 + (1/2)6 = 12 \\ E[C_3] = (1/4)12 + (1/4)20 + (1/2)4 = 10 \\ E[C_4] = (1/4)40 + (1/4)4 + (1/2)8 = 15 \end{bmatrix}$$

Sorting the expected benefits by their decreasing values gives the ranking

$$A_1 \succ A_4 \succ A_2 \succ A_4$$

The decision to take is $A^* = A_1$

The probabilities $p_j = P(S_j)$ of the states of the nature are unknown

$$\begin{array}{cc} & \begin{array}{ccccc} S_1, p_1 = ? & \dots & S_j, p_j = ? & \dots & S_n, p_n = ? \end{array} \\ \mathbf{C} \triangleq \begin{array}{c} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{array} & \begin{bmatrix} C_{11} & \dots & C_{1j} & \dots & C_{1n} \\ & & \vdots & & \\ C_{i1} & \dots & C_{ij} & \dots & C_{in} \\ & & \vdots & & \\ C_{q1} & \dots & C_{qj} & \dots & C_{qn} \end{bmatrix} \end{array}$$

OWA (**O**rdered **W**eighted **A**veraging) method [Yager 1988]

1. Decisional attitude: choose the set of n weights $\mathbf{w} = [w_1 \ \dots \ w_n]$ with $\sum_j w_j = 1$
   - **Optimistic** (max benefit): $\mathbf{w} = [1\ 0 \dots 0]$
   - **Hurwicz** (a balance between min and max): $\mathbf{w} = [\alpha\ 0 \dots 0\ (1-\alpha)]$, typically $\alpha = 1/2$
   - **Normative** (equi weights): $\mathbf{w} = [\frac{1}{n} \dots \frac{1}{n}]$
   - **Pessimistic** (min benefit): $\mathbf{w} = [0 \dots 0\ 1]$
2. Evaluation: compute the weighted average of ordered benefits for each alternative

$$V_i = OWA(C_{i1}, \dots, C_{in}) = \sum_{j=1}^{n} w_j b_{ij}$$

where b_{ij} is the j-th element/benefit among $\{C_{i1}, \dots, C_{in}\}$ and $\mathbf{b}_i = [b_{i1}\ b_{i2}\ \dots\ b_{in}]$ is the reordering of i-th row by decreasing values

3. Decision: take $\delta = A^*$ with

$$A^* = A_{i*} \quad \text{with} \quad i^* = \arg\max_i \{V_i\}$$

Example of decision under ignorance with OWA

The probabilities $p_j = P(S_j)$ of the states of the nature are unknown

$$
\mathbf{C} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \end{array}
\begin{array}{cccc} S_1, p_1 = ? & S_2, p_2 = ? & S_3, p_3 = ? & S_4, p_4 = ? \\ \end{array}
\begin{bmatrix} 10 & 0 & 20 & 30 \\ 1 & 10 & 20 & 30 \\ 30 & 10 & 2 & 5 \end{bmatrix}
$$

- OWA result with **optimistic** attitude $\mathbf{w} = [1\ 0\ 0\ 0] \rightarrow$ we take the max by row

$$
\begin{cases} V_1 = OWA(10,0,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 0]' = 30 \\ V_2 = OWA(1,10,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 1]' = 30 \\ V_3 = OWA(30,10,2,5) = \mathbf{w} \cdot [30\ 10\ 5\ 2]' = 30 \end{cases} \Rightarrow \text{No best choice exists}
$$

- OWA result with **Hurwicz** attitude with $\alpha = 0.5 \Rightarrow \mathbf{w} = [(1/2)\ 0\ 0\ (1/2)]$

$$
\begin{cases} V_1 = OWA(10,0,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 0]' = (30/2) + (0/2) = 15 \\ V_2 = OWA(1,10,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 1]' = (30/2) + (1/2) = 15.5 \\ V_3 = OWA(30,10,2,5) = \mathbf{w} \cdot [30\ 10\ 5\ 2]' = (30/2) + (2/2) = 16 \end{cases} \Rightarrow A_3 \text{ is the best choice}
$$

- OWA result with **normative** attitude $\mathbf{w} = [(1/4)\ (1/4)\ (1/4)\ (1/4)]$

$$
\begin{cases} V_1 = OWA(10,0,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 0]' = 60/4 = 15 \\ V_2 = OWA(1,10,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 1]' = 61/4 \\ V_3 = OWA(30,10,2,5) = \mathbf{w} \cdot [30\ 10\ 5\ 2]' = 47/4 \end{cases} \Rightarrow A_2 \text{ is the best choice}
$$

- OWA result with **pessimistic** attitude $\mathbf{w} = [0\ 0\ 0\ 1] \rightarrow$ we take the min by row

$$
\begin{cases} V_1 = OWA(10,0,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 0]' = 0 \\ V_2 = OWA(1,10,20,30) = \mathbf{w} \cdot [30\ 20\ 10\ 1]' = 1 \\ V_3 = OWA(30,10,2,5) = \mathbf{w} \cdot [30\ 10\ 5\ 2]' = 2 \end{cases} \Rightarrow A_3 \text{ is the best choice}
$$

Decision under uncertainty using OWA

Probas $p_j = P(S_j)$ of the states S_j **are unknown**, but we know a BBA $m(\cdot) : 2^{\mathcal{S}} \mapsto [0, 1]$

$$\mathbf{C} = [\mathbf{c}_1 \dots \mathbf{c}_j \dots \mathbf{c}_n] \triangleq \begin{array}{c c} & \begin{array}{ccccc} S_1, p_1 = ? & \dots & S_j, p_j = ? & \dots & S_n, p_n = ? \end{array} \\ \begin{array}{c} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{array} & \begin{bmatrix} C_{11} & \dots & C_{1j} & \dots & C_{1n} \\ & & \vdots & & \\ C_{i1} & \dots & C_{ij} & \dots & C_{in} \\ & & \vdots & & \\ C_{q1} & \dots & C_{qj} & \dots & C_{qn} \end{bmatrix} \end{array}$$

Method 1: Approximate $m(\cdot)$ by a proba measure $\Rightarrow$ decison-making under risk

Method 2: Extended OWA method [Yager 1988]

1. Decisional attitude: choose the decisional attitude (optimistic,pessimistic, etc)
2. Apply OWA on each sub-matrix $\mathbf{C}(X_k)$ of benefits associated with the focal element X_k, $k = 1, \dots, r$ of $m(\cdot)$ to get valuations $V_i(X_k)$, $i = 1, \dots, q$

$$\mathbf{C}(X_k) = [\mathbf{c}_j | S_j \subseteq X_k]$$

3. Compute the generalized expected benefits for $i = 1, \dots, q$

$$E[C_i] = \sum_{k=1}^{r} m(X_k) V_i(X_k)$$

4. Decision: take the decision $\delta = A^* = A_{i*}$ with $i^* = \arg\max_i \{E[C_i]\}$

Probas $p_j = P(S_j)$ of the states S_j **are unknown**, but we know a BBA $m(\cdot) : 2^{\mathcal{S}} \mapsto [0, 1]$

$$\mathbf{C} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1, p_1 = ? & S_2, p_2 = ? & S_3, p_3 = ? & S_4, p_4 = ? & S_5, p_5 = ? \end{array} \\ \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} \end{array}$$

The uncertainty is modeled by a BBA with 3 focal elements as follows

BBA\FE	$X_1 \triangleq S_1 \cup S_3 \cup S_4$	$X_2 \triangleq S_2 \cup S_5$	$X_3 \triangleq S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5$
$m(\cdot)$	0.6	0.3	0.1

Construction of benefit sub-matrices for each focal element of $m(\cdot)$

$$\mathbf{C}(X_1) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 7 & 12 & 13 \\ 12 & 5 & 11 \\ 9 & 3 & 10 \\ 6 & 11 & 15 \end{bmatrix} \end{array} \quad \mathbf{C}(X_2) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{cc} S_2 & S_5 \end{array} \\ \begin{bmatrix} 5 & 6 \\ 10 & 2 \\ 13 & 9 \\ 9 & 4 \end{bmatrix} \end{array} \quad \mathbf{C}(X_3) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} \end{array}$$

Example of decision under uncertainty using OWA (2)

Using **pessimistic** decisional attitude

- Apply OWA for each sub-matrix $\mathbf{C}(X_3)$, $k = 1, 2, 3$

$$\mathbf{C}(X_1) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 7 & 12 & 13 \\ 12 & 5 & 11 \\ 9 & 3 & 10 \\ 6 & 11 & 15 \end{bmatrix} \end{array} \Rightarrow \begin{cases} V_1(X_1) = OWA(7, 12, 13) = [0\ 0\ 1] \cdot [13\ 12\ 7]' = 7 \\ V_2(X_1) = OWA(12, 5, 11) = [0\ 0\ 1] \cdot [12\ 11\ 5]' = 5 \\ V_3(X_1) = OWA(9, 3, 10) = [0\ 0\ 1] \cdot [10\ 9\ 3]' = 3 \\ V_4(X_1) = OWA(6, 11, 15) = [0\ 0\ 1] \cdot [15\ 11\ 6]' = 6 \end{cases}$$

$$\mathbf{C}(X_2) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{cc} s_2 & s_5 \end{array} \\ \begin{bmatrix} 5 & 6 \\ 10 & 2 \\ 13 & 9 \\ 9 & 4 \end{bmatrix} \end{array} \Rightarrow \begin{cases} V_1(X_2) = OWA(5, 6) = [0\ 1] \cdot [6\ 5]' = 5 \\ V_2(X_2) = OWA(10, 2) = [0\ 1] \cdot [10\ 2]' = 2 \\ V_3(X_2) = OWA(13, 9) = [0\ 1] \cdot [13\ 9]' = 9 \\ V_4(X_2) = OWA(9, 4) = [0\ 1] \cdot [9\ 4]' = 4 \end{cases}$$

$$\mathbf{C}(X_3) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccccc} s_1 & s_2 & s_3 & s_4 & s_5 \end{array} \\ \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} \end{array} \Rightarrow \begin{cases} V_1(X_3) = OWA(7, 5, 12, 13, 6) = [0\ 0\ 0\ 0\ 1] \cdot [13\ 12\ 7\ 6\ 5]' = 5 \\ V_2(X_3) = OWA(12, 10, 5, 11, 2) = [0\ 0\ 0\ 0\ 1] \cdot [12\ 11\ 10\ 5\ 2]' = 2 \\ V_3(X_3) = OWA(9, 13, 3, 10, 9) = [0\ 0\ 0\ 0\ 1] \cdot [13\ 10\ 9\ 9\ 3]' = 3 \\ V_4(X_3) = OWA(6, 9, 11, 15, 4) = [0\ 0\ 0\ 0\ 1] \cdot [15\ 11\ 9\ 6\ 4]' = 4 \end{cases}$$

- Compute generalized expected benefits $E[C_i] = \sum_k m(X_k) V_i(X_k)$ with $m(X_1) = 0.6$, $m(X_2) = 0.3$ and $m(X_3) = 0.1$

$$\begin{aligned} E[C_1] &= 0.6 \cdot 7 + 0.3 \cdot 5 + 0.1 \cdot 5 = 6.2 \\ E[C_2] &= 0.6 \cdot 5 + 0.3 \cdot 2 + 0.1 \cdot 2 = 3.8 \\ E[C_3] &= 0.6 \cdot 3 + 0.3 \cdot 9 + 0.1 \cdot 3 = 4.8 \\ E[C_4] &= 0.6 \cdot 6 + 0.3 \cdot 4 + 0.1 \cdot 4 = 5.2 \end{aligned}$$

- Take final decision with alternative having highest expected benefit $\rightarrow$ $A^* = A_1$

Using **optimistic** decisional attitude

- Apply OWA for each sub-matrix $\mathbf{C}(X_3)$, $k = 1, 2, 3$

$$\mathbf{C}(X_1) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 7 & 12 & 13 \\ 12 & 5 & 11 \\ 9 & 3 & 10 \\ 6 & 11 & 15 \end{bmatrix} \end{array} \Rightarrow \begin{cases} V_1(X_1) = OWA(7,12,13) = [1\ 0\ 0] \cdot [13\ 12\ 7]' = 13 \\ V_2(X_1) = OWA(12,5,11) = [1\ 0\ 0] \cdot [12\ 11\ 5]' = 12 \\ V_3(X_1) = OWA(9,3,10) = [1\ 0\ 0] \cdot [10\ 9\ 3]' = 10 \\ V_4(X_1) = OWA(6,11,15) = [1\ 0\ 0] \cdot [15\ 11\ 6]' = 15 \end{cases}$$

$$\mathbf{C}(X_2) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{cc} s_2 & s_5 \end{array} \\ \begin{bmatrix} 5 & 6 \\ 10 & 2 \\ 13 & 9 \\ 9 & 4 \end{bmatrix} \end{array} \Rightarrow \begin{cases} V_1(X_2) = OWA(5,6) = [1\ 0] \cdot [6\ 5]' = 6 \\ V_2(X_2) = OWA(10,2) = [1\ 0] \cdot [10\ 2]' = 10 \\ V_3(X_2) = OWA(13,9) = [1\ 0] \cdot [13\ 9]' = 13 \\ V_4(X_2) = OWA(9,4) = [1\ 0] \cdot [9\ 4]' = 9 \end{cases}$$

$$\mathbf{C}(X_3) = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \begin{array}{ccccc} s_1 & s_2 & s_3 & s_4 & s_5 \end{array} \\ \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} \end{array} \Rightarrow \begin{cases} V_1(X_3) = OWA(7,5,12,13,6) = [1\ 0\ 0\ 0\ 0] \cdot [13\ 12\ 7\ 6\ 5]' = 13 \\ V_2(X_3) = OWA(12,10,5,11,2) = [1\ 0\ 0\ 0\ 0] \cdot [12\ 11\ 10\ 5\ 2]' = 12 \\ V_3(X_3) = OWA(9,13,3,10,9) = [1\ 0\ 0\ 0\ 0] \cdot [13\ 10\ 9\ 9\ 3]' = 13 \\ V_4(X_3) = OWA(6,9,11,15,4) = [1\ 0\ 0\ 0\ 0] \cdot [15\ 11\ 9\ 6\ 4]' = 15 \end{cases}$$

- Compute generalized expected benefits $E[C_i] = \sum_k m(X_k)V_i(X_k)$ with $m(X_1) = 0.6$, $m(X_2) = 0.3$ and $m(X_3) = 0.1$

$$E[C_1] = 0.6 \cdot 13 + 0.3 \cdot 6 + 0.1 \cdot 13 = 10.9$$
$$E[C_2] = 0.6 \cdot 12 + 0.3 \cdot 10 + 0.1 \cdot 12 = 11.4$$
$$E[C_3] = 0.6 \cdot 10 + 0.3 \cdot 13 + 0.1 \cdot 13 = 11.2$$
$$E[C_4] = 0.6 \cdot 15 + 0.3 \cdot 9 + 0.1 \cdot 15 = 13.2$$

- Take final decision with alternative having highest expected benefit $\rightarrow$ $A^* = A_4$

Advantage of OWA

Very simple to apply

Limitation of OWA

The result strongly depends on the decisional attitude chosen when applying OWA
How to avoid this? → complicate methods exist to select weights (using entropy)

Improvements of OWA

Use jointly the two most extreme decisional attitudes (pessimistic and optimistic) to be more cautious, which can be done as follows

1. Applying OWA using Hurwicz attitude by taking $\alpha = 1/2$
   → a balance only between min and max benefit values
2. Applying modified OWA based on belief functions
   → we use all benefit values between min and max

   - Cautious OWA (COWA) [Tacnet Dezert 2011]

     Pessimistic and optimistic generalized expected benefits allow to build belief intervals, and to get BBAs that are combined with PCR6 to get combined BBA from which the final decision is taken.

   - Fuzzy Cautious OWA (FCOWA) [Han Dezert Tacnet Han 2012]

     A version of COWA more efficient and more simple to implement

Cautious OWA for decision under ignorance or uncertainty

At first, apply OWA with pessimistic and optimistic attitudes to get bounds $[E^{min}[C_i], E^{max}[C_i]]$ of expected benefits of each alternative A_i

Main steps of Cautious OWA (COWA) [Tacnet Dezert 2011]

1. Normalization of exp. benefits intervals ($\div$ by max value) to get intervals in $[0, 1]$
2. Conversion of each interval in a BBA $m_i(A_i)$, $m_i(\bar{A}_i)$, $m_i(A_i \cup \bar{A}_i)$
3. Fusion of the q BBAs $m_i(\cdot)$, $i = 1, \dots, q$ (by PCR6) to get the combined BBA $m(\cdot)$
4. Final decision drawn from $m(\cdot)$ by a chosen decision rule, for example by max BetP, max DSmP, or by min d_{BI})

Drawbacks of COWA

- High computational complexity of the combination (highly dependent on the number q of alternatives)
- In COWA, each expected interval is used as a SoE. However these intervals are jointly obtained which introduces a correlation between the sources, and which is harmful for the combination of BBAs.

Overcoming the drawbacks of COWA

$\rightarrow$ Use Fuzzy-COWA approach, which is more efficient and simpler

Let consider the previous example with

BBA\FE	$X_1 \triangleq S_1 \cup S_3 \cup S_4$	$X_2 \triangleq S_2 \cup S_5$	$X_3 \triangleq S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5$
$m(\cdot)$	0.6	0.3	0.1

and the benefit matrix

$$
\mathbf{C} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array}
\begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} \end{array}
\Rightarrow
\begin{array}{l}
[E^{min}[C_1] = 6.2, E^{max}[C_1] = 10.9] \\
[E^{min}[C_2] = 3.8, E^{max}[C_2] = 11.4] \\
[E^{min}[C_3] = 4.8, E^{max}[C_3] = 11.2] \\
[E^{min}[C_4] = 5.2, E^{max}[C_4] = 13.2]
\end{array}
\Rightarrow
\begin{array}{l}
[6.2/13.2, 10.9/13.2] \approx [0.47, 0.82] \\
[3.8/13.2, 11.4/13.2] \approx [0.29, 0.86] \\
[4.8/13.2, 11.2/13.2] \approx [0.36, 0.85] \\
[5.2/13.2, 13.2/13.2] \approx [0.39, 1.00]
\end{array}
$$

BBA construction from interval $[a, b] \subseteq [0, 1]$

$$
\begin{cases}
m_i(A_i) = a, \\
m_i(\bar{A}_i) = 1 - b \\
m_i(A_i \cup \bar{A}_i) = m_i(\Theta) = b - a
\end{cases}
$$

Alternatives A_i	$m_i(A_i)$	$m_i(\bar{A}_i)$	$m_i(A_i \cup \bar{A}_i)$
A_1	0.47	0.18	0.35
A_2	0.29	0.14	0.57
A_3	0.36	0.15	0.49
A_4	0.39	0	0.61

Fusion of BBAs (here with PCR5)

Focal Element	$m_{PCR5}(.)$
A_1	0.2488
A_2	0.1142
A_3	0.1600
A_4	0.1865
$A_1 \cup A_4$	0.0045
$A_2 \cup A_4$	0.0094
$A_1 \cup A_2 \cup A_4$	0.0236
$A_3 \cup A_4$	0.0075
$A_1 \cup A_3 \cup A_4$	0.0198
$A_2 \cup A_3 \cup A_4$	0.0374
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.1883

Final decision (by max of Bel, BetP, DSmP or Pl)

A_i	$Bel(A_i)$	$BetP(A_i)$	$DSmP(A_i)$	$Pl(A_i)$
A_1	0.2488	0.3126	0.3364	0.4850
A_2	0.1142	0.1863	0.1623	0.3729
A_3	0.1600	0.2299	0.2242	0.4130
A_4	0.1865	0.2712	0.2771	0.4521

Final decision is $A^* = A_1$

At first, apply OWA with pessimistic and optimistic attitudes to get bounds $[E^{min}[C_i], E^{max}[C_i]]$ of expected benefits of each alternative A_i

Main steps of Fuzzy Cautious OWA (FCOWA) [Han Dezert Tacnet Han 2012]

1. Normalize each column $E^{min}[C]$ and $E^{max}[C]$ **separately** to obtain $E^{Fuzzy}(C)$
2. Conversion of the two normalized columns, i.e. two Fuzzy Membership Functions (FMF), into **two pessimistic and optimistic BBAs** $m_{Pess}(\cdot)$ and $m_{Opti}(\cdot)$
3. Fusion of $m_{Pess}(\cdot)$ and $m_{Opti}(\cdot)$ to get a combined BBA $m(\cdot)$
4. Final decision drawn from $m(\cdot)$ by a chosen decision rule, for example by max of BetP, DSmP, or by min of d_{BI}

Advantages of FCOWA

- only 2 BBAs are involved in the combination $\Rightarrow$ only one fusion step is needed
- the BBAs in FCOWA (built by using alpha-cuts) are consonant support (FE are nested), which brings less computational complexity than with COWA
- good performances of FCOWA w.r.t. COWA
- good robustness of FCOWA to scoring errors w.r.t. COWA

Physical meaning

In FCOWA, the 2 SoE are pessimistic OWA and optimistic OWA. The combination result can be regarded as a tradeoff between these two attitudes.

Differences between COWA and FCOWA

The differences between COWA and FCOWA (on previous example):

Difference in normalization

COWA-ER

$$E^{Imp}[C] = \begin{bmatrix} [6.2/13.2; 10.9/13.2] \\ [3.8/13.2; 11.4/13.2] \\ [4.8/13.2; 11.2/13.2] \\ [5.2/13.2; 13.2/13.2] \end{bmatrix} \approx \begin{bmatrix} [0.47; 0.82] \\ [0.29; 0.86] \\ [0.36; 0.85] \\ [0.39; 1.00] \end{bmatrix}$$

FCOWA-ER

$$E^{Fuzzy}[C] = \begin{bmatrix} [6.2/6.2; 10.9/13.2] \\ [3.8/6.2; 11.4/13.2] \\ [4.8/6.2; 11.2/13.2] \\ [5.2/6.2; 13.2/13.2] \end{bmatrix} \approx \begin{array}{c} \text{FMF1 FMF2} \\ \begin{bmatrix} [1.0000; 0.8258] \\ [0.6129; 0.8636] \\ [0.7742; 0.8485] \\ [0.8387; 1.0000] \end{bmatrix} \end{array}$$

Difference in BBA modeling

COWA-ER

$$\begin{cases} m_i(A_i) = a, \\ m_i(\bar{A}_i) = 1 - b \\ m_i(A_i \cup \bar{A}_i) = m_i(\Theta) = b - a \end{cases}$$

FCOWA-ER

$$\begin{cases} B_j = \{A_i \in \Theta | \mu(A_i) \geq \alpha_j\} \\ m(B_j) = \frac{\alpha_j - \alpha_{j-1}}{\alpha_M} \end{cases}$$ **alpha-cut**

COWA

Step 1

$$E^{Imp}[C] = \begin{bmatrix} [6.2/13.2; 10.9/13.2] \\ [3.8/13.2; 11.4/13.2] \\ [4.8/13.2; 11.2/13.2] \\ [5.2/13.2; 13.2/13.2] \end{bmatrix} \approx \begin{bmatrix} [0.47; 0.82] \\ [0.29; 0.86] \\ [0.36; 0.85] \\ [0.39; 1.00] \end{bmatrix}$$

Step 2

$$\begin{cases} m_i(A_i) = a, \\ m_i(\bar{A}_i) = 1 - b \\ m_i(A_i \cup \bar{A}_i) = m_i(\Theta) = b - a \end{cases}$$

Steps 3&4

Combination + decision

FCOWA

Step 1

$$E^{Fuzzy}[C] = \begin{bmatrix} [6.2/6.2; 10.9/13.2] \\ [3.8/6.2; 11.4/13.2] \\ [4.8/6.2; 11.2/13.2] \\ [5.2/6.2; 13.2/13.2] \end{bmatrix} \approx \begin{array}{c} \quad \mu_1(.) \quad\;\; \mu_2(.) \\ \begin{bmatrix} [1.0000; 0.8258] \\ [0.6129; 0.8636] \\ [0.7742; 0.8485] \\ [0.8387; 1.0000] \end{bmatrix} \end{array}$$

Step 2

alpha-cut

$$\begin{cases} B_j = \{A_i \in \Theta | \mu(A_i) \geq \alpha_j\} \\ m(B_j) = \frac{\alpha_j - \alpha_{j-1}}{\alpha_M} \end{cases}$$

Focal Element	$m_{Pess}(.)$	Focal Element	$m_{Opti}(.)$
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.6129	$A_1 \cup A_2 \cup A_3 \cup A_4$	0.8257
$A_1 \cup A_3 \cup A_4$	0.1613	$A_2 \cup A_3 \cup A_4$	0.0227
$A_1 \cup A_4$	0.0645	$A_2 \cup A_4$	0.0152
A_1	0.1613	A_4	0.1364

Steps 3&4

Combination + decision

Example: To compute $m_{Pess}(A_1 \cup A_4) = 0.0645$, we sort $\mu_1(.)$ in increasing order:

$$\mu_1(A_2) = 0.6129 = \alpha_1, \mu_1(A_3) = 0.7742 = \alpha_2, \mu_1(A_4) = 0.8387 = \alpha_3, \mu_1(A_1) = 1 = \alpha_4$$

Focal element $B_3 = \{A_i | \mu_1(A_i) \geq \alpha_3 = 0.8387\} = \{A_1 \cup A_4\}$, because only $\mu_1(A_1) \geq \alpha_3$ and $\mu_1(A_4) \geq \alpha_3$. Therefore

$$m_{Pess}(B_3) = m_{Pess}(A_1 \cup A_4) = \frac{\alpha_3 - \alpha_2}{\alpha_M} = \frac{0.8387 - 0.7742}{1} = 0.0645$$

Example: decision under ignorance with COWA and FCOWA

$$C = \begin{array}{c} S_1, S_2, S_3, S_4, S_5 \\ \begin{bmatrix} 12 & 11 & 10 & 120 & 7 \\ 9 & 10 & 6 & 110 & 3 \\ 7 & 13 & 5 & 100 & 6 \\ 6 & 2 & 3 & \underline{150} & 4 \end{bmatrix} \end{array} \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \qquad E[C] = \begin{bmatrix} [7, & 120] \\ [3, & 110] \\ [5, & 100] \\ [2, & 150] \end{bmatrix}$$

Rank-level Fusion

	A_1	A_2	A_3	A_4
S_1	1	2	3	4
S_2	2	3	1	4
S_3	1	2	3	4
S_4	2	3	4	1
S_5	1	4	2	3

COWA

$$E^{Imp}[C] = \begin{bmatrix} [0.0467, & 0.8000] \\ [0.0200, & 0.7333] \\ [0.0333, & 0.6667] \\ [0.0133, & 1.0000] \end{bmatrix}$$

Alternatives A_i	$m_i(A_i)$	$m_i(\bar{A}_i)$	$m_i(A_i \cup \bar{A}_i)$
A_1	0.0467	0.2000	0.7533
A_2	0.0200	0.2667	0.7133
A_3	0.0333	0.3333	0.6334
A_4	0.0133	0	0.9867

Focal Element	$BetP(.)$
A_1	0.2625
A_2	0.2152
A_3	0.2038
A_4	0.3185

FCOWA

$$E^{Fuzzy}[C] = \begin{bmatrix} [1.0000, & 0.8000] \\ [0.4286, & 0.7333] \\ [0.7143, & 0.6667] \\ [0.2857, & 1.0000] \end{bmatrix}$$

Focal Element	$m_{Pess}(.)$	Focal Element	$m_{Opti}(.)$
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.2857	$A_1 \cup A_2 \cup A_3 \cup A_4$	0.6667
$A_1 \cup A_2 \cup A_3$	0.1429	$A_1 \cup A_3 \cup A_4$	0.0667
$A_1 \cup A_3$	0.2857	$A_3 \cup A_4$	0.0667
A_1	0.2857	A_1	0.1999

Focal Element	$BetP(.)$
A_1	0.5500
A_2	0.1056
A_3	0.2037
A_4	0.1407

The FCOWA method provides a decision $A^* = A_1$ which is consistent with what we obtain by rank-level fusion, contrariwise to what gives COWA

- no general proof of this good behavior of FCOWA has been proved so far
- impact of the normalization method on FCOWA performances not available yet

Methods for Multi-Criteria Decision-Making support

How to make a choice among several alternatives based on different criteria?

Problem modeling 1 $\Rightarrow$ using pairwise comparison matrices $\rightarrow$ **AHP methods**
We consider a set of criteria $C_1, \ldots, C_N$ with preferences of importance established from a pairwise comparison matrix (PCM) $\mathbf{M}$. For each criteria C_j, a set of preferences of the alternatives is established from a given **pairwise comparison matrix** $\mathbf{M}_j$.

Problem modeling 2 $\Rightarrow$ using directly the score matrix $\rightarrow$ **TOPSIS methods**

- A set of $M \geqslant 2$ alternatives $\mathcal{A} \triangleq \{A_1, \ldots, A_M\}$
- A set of $N > 1$ Criteria $\mathbf{C} \triangleq \{C_1, \ldots, C_N\}$
- A set of $N > 1$ criteria **importance weights** $W = \{w_1, \ldots, w_N\}$, with $w_j \in [0, 1]$ and $\sum_j w_j = 1$

$$\mathbf{S} \triangleq \begin{array}{c} \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_M \end{array} \begin{array}{c} \begin{array}{ccccc} C_1, w_1 & \ldots & C_j, w_j & \ldots & C_N, w_N \end{array} \\ \begin{bmatrix} S_{11} & \ldots & S_{1j} & \ldots & S_{1N} \\ & & \vdots & & \\ S_{i1} & \ldots & S_{ij} & \ldots & S_{iN} \\ & & \vdots & & \\ S_{M1} & \ldots & S_{Mj} & \ldots & S_{MN} \end{bmatrix} \end{array}$$

$\mathbf{S}$ is the score matrix of the MCDM problem under consideration

Car example: How to buy a car based on some criteria (i.e. cost, safety, etc.)?

Important remarks

- All methods developed so far suffer from rank reversal problem [Wang Luo 2009], which means that the rank is changed by adding (or deleting) an alternative in the problem. We consider rank reversal as very serious drawback.
- Most of existing methods require score normalization at first, except for ERV (Estimator Ranking Vector) method [Yin et al. 2013]. Normalization has been identified as one of the origins of rank reversal problem.
- There is no MCDM method which makes consensus among users, . . . but some are very popular
  - AHP (Analytic Hierarchy Process) method is very popular in operational research community but not exempt of problems
  - TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) method is very popular **but the choice of normalization is disputed**

What we present

- AHP method and its extension DSm-AHP using belief functions [Saaty 1980, Dezert et al. 2010, Dezert Tacnet 2011]
- a new Belief-Function-based TOPSIS method called BF-TOPSIS to solve classical and non-classical MCDM problems [Dezert Han Yin 2016, Carladous et al. 2016]

Methods for Multi-Criteria Decision-Making support

- AHP and DSm-AHP methods

AHP = **A**nalytic **H**ierarchy **P**rocess

AHP is a Multi-Criteria Decision-Making method developed by Thomas Saaty in 1980's based on the derivation of priority from preferences.

Main steps of classical AHP method [Saaty 1980]

1. The multiple criteria $C_1, \dots, C_N$ are ordered in a hierarchy of importance characterized by $\mathbf{w} = [w_1 \ \dots \ w_N]$ such that $\sum_{j=1}^{N} w_j = 1$, obtained either through a given pairwise comparison matrix (PCM), or given directly.
2. For each criterion C_j, $j = 1, \dots, N$, a set of preferences $\mathbf{w}(C_j)$ of the choice of alternatives is established from given pairwise comparison matrix $\mathbf{M}(C_j)$
3. Combine by the weighted arithmetic mean these preferences to get the global ranking of the alternatives
4. Final decision-making is based on the result of step 3 by selecting the most preferred alternative

Normalized Perron-Frobenius (NPF) eigen vectors (i.e. the eigen vector associated to largest eigen value) of Pairwise Comparison Matrices **are the keys of AHP** method.

Car selection example Θ = set of cars = $\{A, B, C, D\}$

We consider 3 criteria: C_1=Gasoil economy, C_2=Reliability, and C_3=style.

- Establishing importance of criteria from PCM (using NPF eigen vector)

$$\mathbf{M} = [M_{ij}] = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \end{array} \begin{array}{c} \begin{array}{ccc} C_1 & C_2 & C_3 \end{array} \\ \begin{bmatrix} 1/1 & 1/3 & 4/1 \\ 3/1 & 1/1 & 5/1 \\ 1/4 & 1/5 & 1/1 \end{bmatrix} \end{array} \quad \Rightarrow \quad \mathbf{w} = \begin{bmatrix} 0.2797 \\ 0.6267 \\ 0.0936 \end{bmatrix} \quad \Rightarrow \quad C_2 \succ C_1 \succ C_3$$

$M_{21} = 3/1$ means C_2 is 3 times as important as C_1

$M_{23} = 5/1$ means C_2 is 5 times as important as C_3

- Similarly, based on the given comparison matrices $\mathbf{M}(C_j)$ we get $\mathbf{w}(C_j)$

  For example, suppose we obtain from some PCM $\mathbf{M}(C_1)$, $\mathbf{M}(C_2)$ and $\mathbf{M}(C_3)$

$$\mathbf{W} \triangleq [\mathbf{w}(C_1)\ \mathbf{w}(C_2)\ \mathbf{w}(C_3)] = \begin{array}{c} \\ \text{Car } A \\ \text{Car B} \\ \text{Car C} \\ \text{Car D} \end{array} \begin{array}{c} \begin{array}{ccc} \mathbf{w}(C_1) & \mathbf{w}(C_2) & \mathbf{w}(C_3) \end{array} \\ \begin{bmatrix} 0.2500 & 0.4733 & 0.1129 \\ 0.1304 & 0.0611 & 0.4435 \\ 0.5109 & 0.1832 & 0.0565 \\ 0.1087 & 0.2824 & 0.3871 \end{bmatrix} \end{array}$$

- Combination (by weighted arithmetic mean) to get final ranking vector $\mathbf{r}$

$$\mathbf{r} = \mathbf{W} \times \mathbf{w} = \begin{bmatrix} 0.2500 & 0.4733 & 0.1129 \\ 0.1304 & 0.0611 & 0.4435 \\ 0.5109 & 0.1832 & 0.0565 \\ 0.1087 & 0.2824 & 0.3871 \end{bmatrix} \times \begin{bmatrix} 0.2797 \\ 0.6267 \\ 0.0936 \end{bmatrix} = \begin{array}{c} \\ \text{Car } A \\ \text{Car B} \\ \text{Car C} \\ \text{Car D} \end{array} \begin{array}{c} \mathbf{r} \\ \begin{bmatrix} 0.3771 \\ 0.1163 \\ 0.2630 \\ 0.2436 \end{bmatrix} \end{array} \Rightarrow A \succ C \succ D \succ B$$

- Final decision based on $\mathbf{r}$ vector: δ = Car A

Advantages

- quite easy to implement (toolboxes exist for eigen vector computation)
- easy to use
- pairwise comparison matrices are convenient for preference elicitation for experts

Limitations

- rank reversal problem
- does not take into account for uncertainties in the ranking process

Extension of AHP with DST [Beynon 2002]

- DS-AHP extends AHP using belief functions and Dempster-Shafer (DS) rule
- . . . but DS rule is questionable, and the importance discounting is not efficient

Extension of AHP with DSmT [Dezert et al. 2010, Dezert Tacnet 2011]

- DSm-AHP proposes a better rule of combination (PCR6)
- DSm-AHP proposes a more interesting importance discounting technique

DSm-AHP is an extension of Analytic Hierarchy Process (AHP) with using, PCR rules of combination, and the new importance discounting technique to take into account uncertainty in the ranking process

Main steps of DSm-AHP method [Dezert et al. 2010, Dezert Tacnet 2011]

1. Construction of uncertain comparison matrices. Take as BBA, the normalized Perron-Frobenius vector of each pairwise comparison matrix
2. Use PCR6 rule, to combine BBAs to get a final priority ranking vector $\mathbf{r}$
3. Make final decision by a chosen classical decision rule (i.e. max of Bel, max of Pl, max of BetP, max of DSmP, or min of d_{BI})

Advantages of DSm-AHP method

- better efficient rule of combination
- distinction between Shafer's reliability discounting and importance discounting

Drawbacks of DSm-AHP method

- rank reversal can occur with DS-AHP and DSm-AHP
- complicate to implement because of PCR6 general formula
- cannot work with many criteria and alternatives because of its too high complexity
- Solution → use BBA approximation techniques, and PCR6 rule sequentially

Reliability discounting versus importance discounting

Reliability discounting, $\alpha \in [0,1]$		**Importance discounting**, $\beta \in [0,1]$
$\begin{cases} m^{\alpha}(A) = \alpha \cdot m(A) & \forall A \neq \Theta \\ m^{\alpha}(\Theta) = \alpha \cdot m(\Theta) + (1-\alpha) \end{cases}$	$\neq$	$\begin{cases} m^{\beta}(A) = \beta \cdot m(A) & \forall A \neq \emptyset \\ m^{\beta}(\emptyset) = \beta \cdot m(\emptyset) + (1-\beta) \end{cases}$

$\alpha = 1 \Leftrightarrow$ SoE is 100% reliable — $\beta = 1 \Leftrightarrow$ SoE is 100% important

$\alpha = 0 \Leftrightarrow$ SoE is 100% unreliable — $\beta = 0 \Leftrightarrow$ SoE is not important

PCR5/6 fusion[4] of importance discounted BBAs (if $\beta_1 = \beta_2 = 0$, $m_{12}^{PCR5/6}(\Theta) \triangleq 1$)

$$m_{12}^{PCR5/6_{\emptyset}}(X) = m_{12}^{Conj,\beta_1\beta_2}(X) + \sum_{\substack{Y \in 2^{\Theta} \\ X \cap Y = \emptyset}} \left[\frac{m_1^{\beta_1}(X)^2 m_2^{\beta_2}(Y)}{m_1^{\beta_1}(X) + m_2^{\beta_2}(Y)} + \frac{m_2^{\beta_2}(X)^2 m_1^{\beta_1}(Y)}{m_2^{\beta_2}(X) + m_1^{\beta_1}(Y)}\right]$$

Because $m_{12}^{PCR5/6_{\emptyset}}(\emptyset) > 0$, a classical normalization applies, that is $m_{12}^{PCR5/6}(\emptyset) = 0$, and $m_{12}^{PCR5/6}(X) = m_{12}^{PCR5/6_{\emptyset}}(X)/[1 - m_{12}^{PCR5/6_{\emptyset}}(\emptyset)]$ for $X \neq \emptyset$

Note: Dempster-Shafer rule **does not react** to importance discounting

[4] use general PCR6 formula for combining more than two BBAs, see [DSmT books], Vols. 2 & 3

Car example $\{\text{Cars}\} = \Theta = \{A, B, C\}$, $\{\text{Criteria}\} = \{C_1 \triangleq \text{Economy}, C_2 \triangleq \text{Reliability}\}$

Suppose the two given pairwise comparisons matrices $\mathbf{M}(C_1)$ and $\mathbf{M}(C_2)$ are

$$\mathbf{M}(C_1) = \begin{array}{c} \\ A \\ B \cup C \\ \Theta \end{array} \begin{array}{c} \begin{array}{ccc} A & B \cup C & \Theta \end{array} \\ \begin{bmatrix} 1 & ? \to 0 & 1/3 \\ ? \to 0 & 1 & 2 \\ 3 & 1/2 & 1 \end{bmatrix} \end{array} \Rightarrow \mathbf{w}(C_1) \approx \begin{bmatrix} 0.0889 \\ 0.5337 \\ 0.3774 \end{bmatrix} = \begin{bmatrix} m_1(A) \\ m_1(B \cup C) \\ m_1(\Theta) \end{bmatrix}$$

$$\mathbf{M}(C_2) = \begin{array}{c} \\ A \\ B \\ A \cup C \\ B \cup C \end{array} \begin{array}{c} \begin{array}{cccc} A & B & A \cup C & B \cup C \end{array} \\ \begin{bmatrix} 1 & 2 & 4 & 3 \\ 1/2 & 1 & 1/2 & 1/5 \\ 1/4 & 2 & 1 & ? \to 0 \\ 1/3 & 5 & ? \to 0 & 1 \end{bmatrix} \end{array} \Rightarrow \mathbf{w}(C_2) \approx \begin{bmatrix} 0.5002 \\ 0.1208 \\ 0.1222 \\ 0.2568 \end{bmatrix} = \begin{bmatrix} m_2(A) \\ m_2(B) \\ m_2(A \cup C) \\ m_2(B \cup C) \end{bmatrix}$$

Suppose the two criteria have **same full importances**, i.e. $\beta_1 = 1$ and $\beta_2 = 1$

We apply directly PCR6 rule

FE of 2^Θ	$m_1(\cdot)$	$m_2(\cdot)$	$m_{12}^{PCR5/6}(\cdot)$
$\emptyset$	0	0	0
A	0.0889	0.5002	0.3837
B	0	0	0.1162
$A \cup B$	0	0.1208	0
C	0	0	0.0652
$A \cup C$	0	0.1222	0.0461
$B \cup C$	0.5337	0.2568	0.3887
$A \cup B \cup C$	0.3774	0	0

We take the final decision according a chosen decision rule from $m_{12}^{PCR5/6}(\cdot)$

FE of 2^Θ	$Bel(\cdot)$	$BetP(\cdot)$	$Pl(\cdot)$
A	0.3837	0.4068	0.4298
B	0.1162	0.3105	0.5049
C	0.0652	0.2826	0.5000

Car example again with **different importances** $\beta_1 = 0.25$ and $\beta_2 = 0.75$

With DSm-AHP

We apply importance discounting to derive $m_1^{\beta_1}(\cdot)$ and $m_2^{\beta_2}(\cdot)$, apply PCR5/6 rule to get $m_{12}^{PCR5/6\emptyset}(\cdot)$ and normalize to get $m_{12}^{PCR5/6}(\cdot)$

FE of 2^Θ	$m_1(\cdot)$	$m_2(\cdot)$	$m_1^{\beta_1}(.)$	$m_2^{\beta_2}(\cdot)$	$m_{12}^{PCR5/6\emptyset}(\cdot)$	$m_{12}^{PCR5/6}(\cdot)$
$\emptyset$	0	0	0.7500	0.2500	0.6558	0
A	0.0889	0.5002	0.0222	0.3751	0.1794	0.5213
B	0	0	0	0	0.0121	0.0351
$A \cup B$	0	0.1208	0	0.0906	0.0159	0.0461
C	0	0	0	0	0.0122	0.0355
$A \cup C$	0	0.1222	0	0.0917	0.0161	0.0469
$B \cup C$	0.5337	0.2568	0.1334	0.1926	0.1020	0.2963
$A \cup B \cup C$	0.3774	0	0.0944	0	0.0065	0.0188

With classic AHP (by simple componentwise weighted averaging)

$$m_{12}^{AHP}(\cdot) = [m_1(\cdot)\ m_2(\cdot)] \times \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0.0889 & 0.5002 \\ 0 & 0 \\ 0 & 0.1208 \\ 0 & 0 \\ 0 & 0.1222 \\ 0.5337 & 0.2568 \\ 0.3774 & 0 \end{bmatrix} \times \begin{bmatrix}0.25\\ 0.75\end{bmatrix} \approx \begin{bmatrix} 0 \\ 0.3974 \\ 0 \\ 0.0906 \\ 0 \\ 0.0916 \\ 0.3260 \\ 0.0944 \end{bmatrix}$$

Car example again with **different importances** $\beta_1 = 0.25$ and $\beta_2 = 0.75$

$\Rightarrow$ DSm-AHP reduces the uncertainty of the result $U(X) = Pl(X) - Bel(X)$

Decision drawn from classical AHP using $m_{12}^{AHP}(\cdot) \rightarrow \delta = A$

FE of 2^Θ	$Bel(\cdot)$	$BetP(\cdot)$	$Pl(\cdot)$	$U(\cdot)$
A	0.3974	0.5200	0.6741	0.2767
B	0	0.2398	0.5110	0.5110
C	0	0.2403	0.5121	0.5121

Decision drawn from DSm-AHP using $m_{12}^{PCR5/6}(\cdot) \rightarrow \delta = A$

FE of 2^Θ	$Bel(\cdot)$	$BetP(\cdot)$	$Pl(\cdot)$	$U(\cdot)$
A	0.5213	0.5741	0.6331	0.1118
B	0.0351	0.2126	0.3963	0.3612
C	0.0355	0.2134	0.3974	0.3619

In this example AHP and DSm-AHP provide the same decision, but DSm-AHP offers a better precision (less uncertainty) on the result

Methods for Multi-Criteria Decision-Making support

- TOPSIS and BF-TOPSIS methods

TOPSIS = **T**echnique for **O**rder **P**reference by **S**imilarity to **I**deal **S**olution

Classical TOPSIS method [Hwang Yoon 1981]

1. Build the normalized score matrix $\mathbf{R} = [R_{ij}] = [S_{ij}/\sqrt{\sum_i S_{ij}^2}]$
2. Calculate the weighted normalized decision matrix $\mathbf{D} = [w_j \cdot R_{ij}]$
3. Determine the positive (best) ideal solution A^{best} by **taking the best/max value** in each column of $\mathbf{D}$
4. Determine the negative (worst) ideal solution A^{worst} by **taking the worst/min value** in each column of $\mathbf{D}$
5. Compute L2-distances $d(A_i, A^{best})$ of A_i, (i=1,...,M) to A^{best}, and $d(A_i, A^{worst})$ of A_i to A^{worst}
6. Calculate the relative closeness of A_i to best ideal solution A^{best} by

$$C(A_i, A^{best}) \triangleq \frac{d(A_i, A^{worst})}{d(A_i, A^{worst}) + d(A_i, A^{best})}$$

   When $C(A_i, A^{best}) = 1$, its means that $A_i = A^{best}$ because $d(A_i, A^{best}) = 0$

   When $C(A_i, A^{best}) = 0$, its means that $A_i = A^{worst}$ because $d(A_i, A^{worst}) = 0$

7. Rank alternatives A_i according to $C(A_i, A^{best})$ in descending order, and select the highest preferred solution

A very simple example for TOPSIS

$$\mathbf{S} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \end{array} \begin{array}{cc} C_1, w_1 = 1/2 & C_2, w_2 = 1/2 \\ \left[\begin{array}{cc} 6 & 2 \\ 3 & 5 \\ 4 & 4 \end{array}\right] \end{array}$$

1. Step 1 & 2 (normalization & columns weighting):

$$\mathbf{R} = [S_{ij}/\sqrt{\sum_i S_{ij}^2}] \Rightarrow \mathbf{R} = \begin{array}{cc} C_1, 1/2 & C_2, 1/2 \\ \left[\begin{array}{cc} 0.7682 & 0.2981 \\ 0.3841 & 0.7454 \\ 0.5121 & 0.5963 \end{array}\right] \end{array} \Rightarrow \mathbf{D} = \begin{bmatrix} 0.3841 & 0.1491 \\ 0.1921 & 0.3727 \\ 0.2561 & 0.2981 \end{bmatrix}$$

2. Step 3 & 4 (best and worst solutions) $A^{best} = [0.3841\ 0.3727]$, $A^{worst} = [0.1921\ 0.1491]$
3. Step 5 (L_2-distance of A_i to A^{best} and to A^{worst}):

$$\begin{array}{l} A_1 = [0.3841\ 0.1491] \\ A_2 = [0.1921\ 0.3727] \\ A_3 = [0.2561\ 0.2981] \end{array} \begin{array}{cc} A^{best} = [0.3841\ 0.3727] & A^{worst} = [0.1921\ 0.1491] \\ \left[\begin{array}{ll} d(A_1, A^{best}) = 0.2236 & d(A_1, A^{worst}) = 0.1921 \\ d(A_2, A^{best}) = 0.1921 & d(A_2, A^{worst}) = 0.2236 \\ d(A_3, A^{best}) = 0.1482 & d(A_3, A^{worst}) = 0.1622 \end{array}\right] \end{array}$$

4. Step 6 (relative closeness of A_i to A^{best}): $C(A_i, A^{best}) \triangleq \frac{d(A_i,A^{worst})}{d(A_i,A^{worst})+d(A_i,A^{best})}$

$$C(A_1, A^{best}) = 0.4620 \qquad C(A_2, A^{best}) = 0.5380 \qquad C(A_3, A^{best}) = 0.5227$$

5. Step 7 (ranking by decreasing order of $C(A_i, A^{best})$): $A_2 \succ A_3 \succ A_1$

Based on TOPSIS, the decision δ to make is $\delta = A_2$

BF-TOPSIS method for MCDM

BF-TOPSIS is a TOPSIS-alike method based on belief functions [Dezert Han Yin 2016]

Advantages of BF-TOPSIS

- no need for ad-hoc choice of scores normalization
- relatively simple to implement
- more robust to rank reversal phenomena (although not exempt)

Main problem to overcome

Working with belief functions requires the construction of BBAs. How to build efficiently BBAs from the score values

Solution → see next slides

Four BF-TOPSIS methods available with different complexity

1. BF-TOPSIS1: smallest complexity
2. BF-TOPSIS2: medium complexity
3. BF-TOPSIS3: high complexity (because of PCR6 fusion rule)
4. BF-TOPSIS4: high complexity (because of ZPCR6 fusion rule)

BF-TOPSIS for working with imprecise scores presented in [Dezert Han Tacnet 2017]

- Positive support of A_i based on all scores values of a criteria C_j

$$\mathrm{Sup}_j(A_i) \triangleq \sum_{k\in\{1,\dots M\}|S_{kj}\leqslant S_{ij}} |S_{ij} - S_{kj}|$$

$\mathrm{Sup}_j(A_i)$ measures how much A_i is better (higher) than other alternatives

- Negative support of A_i based on all scores values of a criteria C_j

$$\mathrm{Inf}_j(A_i) \triangleq -\sum_{k\in\{1,\dots M\}|S_{kj}\geqslant S_{ij}} |S_{ij} - S_{kj}|$$

$\mathrm{Inf}_j(A_i)$ measures how much A_i is worse (lower) than other alternatives

Important inequality see proof in [Dezert Han Yin 2016]

$$\frac{\mathrm{Sup}_j(A_i)}{A^j_{\text{max}}} \leqslant 1 - \frac{\mathrm{Inf}_j(A_i)}{A^j_{\text{min}}}$$

iff $A^j_{\text{max}} \triangleq \max_i \mathrm{Sup}_j(A_i)$ and $A^j_{\text{min}} \triangleq \min_i \mathrm{Inf}_j(A_i)$ are different from zero.

Reminder $\frac{Sup_j(A_i)}{A^j_{max}} \leqslant 1 - \frac{Inf_j(A_i)}{A^j_{min}}$

Belief function modeling

$$Bel_{ij}(A_i) \triangleq \frac{Sup_j(A_i)}{A^j_{max}} \quad \text{and} \quad Bel_{ij}(\bar{A}_i) \triangleq \frac{Inf_j(A_i)}{A^j_{min}}$$

If $A^j_{max} = 0$, we set $Bel_{ij}(X_i) = 0$
If $A^j_{min} = 0$, we set $Pl_{ij}(A_i) = 1$ so that $Bel_{ij}(\bar{A}_i) = 0$

By construction, $0 \leqslant Bel_{ij}(A_i) \leqslant (Pl_{ij}(A_i) = 1 - Bel_{ij}(\bar{A}_i)) \leqslant 1$

BBA construction from Belief Interval

From $[Bel_{ij}(A_i), Pl_{ij}(A_i)]$, one gets the $M \times N$ BBAs matrix $\mathbf{M} = [m_{ij}(\cdot)]$ by taking

$$m_{ij}(A_i) = Bel_{ij}(A_i)$$
$$m_{ij}(\bar{A}_i) = Bel_{ij}(\bar{A}_i) = 1 - Pl_{ij}(A_i)$$
$$m_{ij}(A_i \cup \bar{A}_i) = Pl_{ij}(A_i) - Bel_{ij}(A_i)$$

Advantages of this BBA construction

1. if all S_{ij} are the same for a given column, we get $\forall A_i$, $Sup_j(A_i) = Inf_j(A_i) = 0$ and therefore $m_{ij}(A_i \cup \bar{A}_i) = 1$ which is the vacuous BBA, which makes sense.
2. it is invariant to the bias and scaling effects of score values. Indeed, if S_{ij} are replaced by $S'_{ij} = a \cdot S_{ij} + b$, with a scale factor $a > 0$ and a bias $b \in \mathbb{R}$, then $m_{ij}(\cdot)$ and $m'_{ij}(\cdot)$ remain equal.
3. if a numerical value S_{ij} is missing or indeterminate, then we use the vacuous belief assignment $m_{ij}(A_i \cup \bar{A}_i) = 1$.
4. We can also discount the BBA $m_{ij}(\cdot)$ by a reliability factor using the classical Shafer's discounting method if one wants to express some doubts on the reliability of $m_{ij}(\cdot)$.

In summary

From $[S_{ij}]$, we know how to build the matrix $\mathbf{M} = [(m_{ij}(A_i), m_{ij}(\bar{A}_i), m_{ij}(A_i \cup \bar{A}_i))]$

How to use these BBAs to rank A_i to make a decision? $\rightarrow$ BF-TOPSIS methods

Steps of BF-TOPSIS1 [Dezert Han Yin 2016]

1. From $\mathbf{S}$, compute BBAs $m_{ij}(A_i)$ $m_{ij}(\bar{A}_i)$, and $m_{ij}(A_i \cup \bar{A}_i)$
2. Set $m_{ij}^{\text{best}}(A_i) \triangleq 1$, and $m_{ij}^{\text{worst}}(\bar{A}_i) \triangleq 1$ and compute distances $d_{BI}^{E}(m_{ij}, m_{ij}^{\text{best}})$ and $d_{BI}^{E}(m_{ij}, m_{ij}^{\text{worst}})$ to ideal solutions.
3. Compute the weighted average distances of A_i to ideal solutions

$$d^{\text{best}}(A_i) \triangleq \sum_{j=1}^{N} w_j \cdot d_{BI}^{E}(m_{ij}, m_{ij}^{\text{best}})$$

$$d^{\text{worst}}(A_i) \triangleq \sum_{j=1}^{N} w_j \cdot d_{BI}^{E}(m_{ij}, m_{ij}^{\text{worst}})$$

4. Compute the relative closeness of A_i with respect to ideal best solution A^{best}

$$C(A_i, A^{\text{best}}) \triangleq \frac{d^{\text{worst}}(A_i)}{d^{\text{worst}}(A_i) + d^{\text{best}}(A_i)}$$

5. Rank A_i by $C(A_i, A^{\text{best}})$ in descending order.

Steps of BF-TOPSIS2 [Dezert Han Yin 2016]

1. From $\mathbb{S}$, compute BBAs $m_{ij}(A_i)$ $m_{ij}(\bar{A}_i)$, and $m_{ij}(A_i \cup \bar{A}_i)$
2. Set $m_{ij}^{\text{best}}(A_i) \triangleq 1$, and $m_{ij}^{\text{worst}}(\bar{A}_i) \triangleq 1$ and compute distances $d_{BI}^{E}(m_{ij}, m_{ij}^{\text{best}})$ and $d_{BI}^{E}(m_{ij}, m_{ij}^{\text{worst}})$ to ideal solutions.
3. For each criteria C_j, compute the relative closeness of A_i to best ideal solution A^{best} by

$$C_j(A_i, A^{\text{best}}) \triangleq \frac{d_{BI}^{E}(m_{ij}, m_{ij}^{\text{worst}})}{d_{BI}^{E}(m_{ij}, m_{ij}^{\text{worst}}) + d_{BI}^{E}(m_{ij}, m_{ij}^{\text{best}})}$$

4. Compute the weighted average of $C_j(A_i, A^{\text{best}})$ by

$$C(A_i, A^{\text{best}}) \triangleq \sum_{j=1}^{N} w_j \cdot C_j(A_i, A^{\text{best}})$$

5. Rank A_i by $C(A_i, A^{\text{best}})$ in descending order.

Steps of BF-TOPSIS3 [Dezert Han Yin 2016]

1. Compute BBAs $m_{ij}(A_i)$, $m_{ij}(\bar{A}_i)$ and $m_{ij}(A_i \cup \bar{A}_i)$ and apply importance discounting of each BBA with weight w_j, see [Smarandache Dezert Tacnet 2010]
2. For each A_i, fuse the discounted BBAs with PCR6 to get BBAs $m_i^{\text{PCR6}}(\cdot)$
3. Set $m_i^{\text{best}}(A_i) \triangleq 1$, and $m_i^{\text{worst}}(\bar{A}_i) \triangleq 1$. Compute distances

$$d^{\text{best}}(A_i) \triangleq d_{BI}^{E}(m_i^{\text{PCR6}}, m_i^{\text{best}})$$

$$d^{\text{worst}}(A_i) \triangleq d_{BI}^{E}(m_i^{\text{PCR6}}, m_i^{\text{worst}})$$

4. Compute the relative closeness of A_i, $i = 1, \ldots, M$, with respect to ideal best solution A^{best}

$$C(A_i, A^{\text{best}}) \triangleq \frac{d^{\text{worst}}(A_i)}{d^{\text{worst}}(A_i) + d^{\text{best}}(A_i)}$$

5. Rank A_i by $C(A_i, A^{\text{best}})$ in descending order.

BF-TOPSIS4 method

Same as BF-TOPSIS3, but PCR6 rule is replaced by ZPCR6 rule (i.e. PCR6 rule including Zhang's degree of intersection) [Smarandache Dezert 2015]

BF-TOPSIS methods **are consistent** with direct ranking in mono-criteria case

Example (Mono-criteria) Preference order → greater value is better

$$
\mathbf{S} \triangleq \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \end{array}
\overset{\displaystyle C_1}{\begin{bmatrix} 10 \\ 20 \\ -5 \\ 0 \\ 100 \\ -11 \\ 0 \end{bmatrix}}
\Rightarrow \mathbf{M} \triangleq
\begin{array}{c|ccc}
 & m_{i1}(A_i) & m_{i1}(\bar{A}_i) & m_{i1}(A_i \cup \bar{A}_i) \\
\hline
A_1 & 0.0955 & 0.5236 & 0.3809 \\
A_2 & 0.1809 & 0.4188 & 0.4003 \\
A_3 & 0.0102 & 0.8115 & 0.1783 \\
A_4 & 0.0273 & 0.6806 & 0.2921 \\
A_5 & 1.0000 & 0 & 0 \\
A_6 & 0 & 1.0000 & 0 \\
A_7 & 0.0273 & 0.6806 & 0.2921
\end{array}
\Rightarrow
\begin{array}{c|c}
 & C(A_i, A^{\text{best}}) \\
\hline
A_1 & 0.1130 \\
A_2 & 0.1948 \\
A_3 & 0.0257 \\
A_4 & 0.0485 \\
A_5 & 1.0000 \\
A_6 & 0 \\
A_7 & 0.0485
\end{array}
$$

Results

Ranking methods	**Preferences order**
By direct ranking	$A_5 \succ A_2 \succ A_1 \succ (A_4 \sim A_7) \succ A_3 \succ A_6$
By BF-TOPSIS	$A_5 \succ A_2 \succ A_1 \succ (A_4 \sim A_7) \succ A_3 \succ A_6$
By DS fusion	$A_5 \succ (A_1 \sim A_2 \sim A_3 \sim A_4 \sim A_6 \sim A_7)$
By PCR6 fusion	$A_5 \succ A_2 \succ A_1 \succ A_4 \succ (A_3 \sim A_6 \sim A_7)$

Rankings resulting of DS and PCR6 fusion of the BBAs do not match with direct ranking even in mono criteria case because of strong dependencies between BBAs in their construction.

In this example, we have $Score(A_5) >> Score(A_2)$

$$\mathbf{S} \triangleq \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \end{array} \begin{array}{c} C_1 \\ \begin{bmatrix} 10 \\ 20 \\ -5 \\ 0 \\ 100 \\ -11 \\ 0 \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \end{array} \begin{array}{c} C(A_i, A^{\text{best}}) \\ \begin{bmatrix} 0.1130 \\ 0.1948 \\ 0.0257 \\ 0.0485 \\ 1.0000 \\ 0 \\ 0.0485 \end{bmatrix} \end{array} \Rightarrow A_5 \succ A_2 \succ A_1 \succ (A_4 \sim A_7) \succ A_3 \succ A_6$$

Let's modify the example with $Score(A_5) \sim Score(A_2)$

$$\mathbf{S} \triangleq \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \end{array} \begin{array}{c} C_1 \\ \begin{bmatrix} 10 \\ 20 \\ -5 \\ 0 \\ 21 \\ -11 \\ 0 \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \end{array} \begin{array}{c} C(A_i, A^{\text{best}}) \\ \begin{bmatrix} 0.5072 \\ 0.9472 \\ 0.0675 \\ 0.1584 \\ 1.0000 \\ 0 \\ 0.1584 \end{bmatrix} \end{array} \Rightarrow A_5 \succ A_2 \succ A_1 \succ (A_4 \sim A_7) \succ A_3 \succ A_6$$

We see that A_2 is very close to ideal best solution, even if final result is unchanged.

When all scores are the same

$\Rightarrow$ all BBAs are the same and equal to the vacuous BBA

$\Rightarrow$ all closeness measures to best ideal solution are equal

$$\mathbf{S} \triangleq \begin{array}{c} \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_M \end{array} \begin{array}{c} C_1 \\ \begin{bmatrix} s \\ \vdots \\ s \\ \vdots \\ s \end{bmatrix} \end{array} \Rightarrow \mathbf{M} \triangleq \begin{array}{c} \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_M \end{array} \begin{array}{c} m_{i1}(A_i \cup \bar{A}_i) \\ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_M \end{array} \begin{array}{c} C(A_i, A^{\text{best}}) \\ \begin{bmatrix} c \\ \vdots \\ c \\ \vdots \\ c \end{bmatrix} \end{array}$$

Conclusion: No specific choice can be drawn, which is perfectly normal.

Multi-Criteria example [Wang Luo 2009]

We consider 5 alternatives, and 4 criteria

$$\mathbf{S} \triangleq \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{array} \begin{array}{c} \begin{array}{cccc} C_1, \frac{1}{6} & C_2, \frac{1}{3} & C_3, \frac{1}{3} & C_4, \frac{1}{6} \end{array} \\ \begin{bmatrix} 36 & 42 & 43 & 70 \\ 25 & 50 & 45 & 80 \\ 28 & 45 & 50 & 75 \\ 24 & 40 & 47 & 100 \\ 30 & 30 & 45 & 80 \end{bmatrix} \end{array}$$

Rank reversal with TOPSIS

Set of alternatives	TOPSIS
$\{A_1, A_2, A_3\}$	$A_3 \succ A_2 \succ A_1$
$\{A_1, A_2, A_3, A_4\}$	$A_2 \succ A_3 \succ A_1 \succ A_4$
$\{A_1, A_2, A_3, A_4, A_5\}$	$A_3 \succ A_2 \succ A_4 \succ A_1 \succ A_5$
	Rank reversal

Rank reversal with BF-TOPSIS

Set of alternatives	BF-TOPSIS1 & BF-TOPSIS2	BF-TOPSIS3 & BF-TOPSIS4
$\{A_1, A_2, A_3\}$	$A_2 \succ A_3 \succ A_1$	$A_3 \succ A_2 \succ A_1$
$\{A_1, A_2, A_3, A_4\}$	$A_3 \succ A_2 \succ A_4 \succ A_1$	$A_3 \succ A_2 \succ A_4 \succ A_1$
$\{A_1, A_2, A_3, A_4, A_5\}$	$A_3 \succ A_2 \succ A_4 \succ A_1 \succ A_5$	$A_3 \succ A_2 \succ A_4 \succ A_1 \succ A_5$
	Rank reversal	No rank reversal

Car selection example

How to buy a car among 4 possible choices, and based on 5 different criteria with weights $w_1 = 5/17$, $w_2 = 4/17$, $w_3 = 4/17$, $w_4 = 1/17$, and $w_5 = 3/17$

- C_1 = price (in €); the least is the best
- C_2 = fuel consumption (in L/km); the least is the best
- C_3 = CO_2 emission (in g/km); the least is the best
- C_4 = fuel tank volume (in L); the biggest is the best
- C_5 = trunk volume (in L); the biggest is the best

Building the score matrix from http://www.choisir-sa-voiture.com

A_1 = TOYOTA YARIS 69 VVT-i Tendance
A_2 = SUZUKI SWIFT MY15 1.2 VVT So'City
A_3 = VOLKSWAGEN POLO 1.0 60 Confortline
A_4 = OPEL CORSA 1.4 Turbo 100 ch Start/Stop Edition

$$\mathbf{S} \triangleq \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} C_1, \frac{5}{17} \quad C_2, \frac{4}{17} \quad C_3, \frac{4}{17} \quad C_4, \frac{1}{17} \quad C_5, \frac{3}{17} \\ \begin{bmatrix} 15000 & 4.3 & 99 & 42 & 73 \\ 15290 & 5.0 & 116 & 42 & 892 \\ 15350 & 5.0 & 114 & 45 & 952 \\ 15490 & 5.3 & 123 & 45 & 1120 \end{bmatrix} \end{array}$$

A_1 **is the expected best choice** because the 3 most important criteria meet their best values for car A_1.

With classical TOPSIS $A_4 \succ A_1 \succ A_3 \succ A_2$ (counter-intuitive)

With all BF-TOPSIS methods $A_1 \succ A_3 \succ A_2 \succ A_4$ (which fits with what we expect)

Best student prize example

How to give the best student prize awards among 4 students $\{A_1, A_2, A_3, A_4\}$, and based on 10 different criteria with equal importance ($w_j = 1/10$, $j = 1, \ldots, 10$) ?

	A_1	A_2	A_3	A_4
$C_1 \triangleq$ Math	90	80	70	60
$C_2 \triangleq$ Arts	90	80	70	60
$C_3 \triangleq$ English	90	80	70	60
$C_4 \triangleq$ Geography	90	80	70	60
$C_5 \triangleq$ Physics	90	80	70	75
$C_6 \triangleq$ Music	90	80	70	95
$C_7 \triangleq$ History	80	90	70	85
$C_8 \triangleq$ Chemistry	80	90	70	85
$C_9 \triangleq$ Biology	80	90	70	85
$C_{10} \triangleq$ Long jump	3.5m	3.7m	4.0m	3.6m

BF-TOPSIS results

	Considering 3 students $\{A_1, A_2, A_3\}$ only		Considering the 4 students	
Methods	Ranking vectors	Preferences orders	Ranking vectors	Preferences orders
ERV	[0.748, 0.636, 0.188]	$A_1 \succ A_2 \succ A_3$	[0.620, 0.636, 0.248, 0.386]	$A_2 \succ A_1 \succ A_4 \succ A_3$
BF-TOPSIS1	[0.729, 0.594, 0.100]	$A_1 \succ A_2 \succ A_3$	[0.675, 0.620, 0.195, 0.320]	$A_1 \succ A_2 \succ A_4 \succ A_3$
BF-TOPSIS2	[0.731, 0.597, 0.100]	$A_1 \succ A_2 \succ A_3$	[0.677, 0.622, 0.194, 0.319]	$A_1 \succ A_2 \succ A_4 \succ A_3$
BF-TOPSIS3	[0.803, 0.736, 0.100]	$A_1 \succ A_2 \succ A_3$	[0.766, 0.775, 0.158, 0.288]	$A_2 \succ A_1 \succ A_4 \succ A_3$
BF-TOPSIS4	[0.803, 0.736, 0.100]	$A_1 \succ A_2 \succ A_3$	[0.766, 0.775, 0.158, 0.288]	$A_2 \succ A_1 \succ A_4 \succ A_3$

ERV, BFTOPSIS3, and BFTOPSIS4 exhibit rank reversal

BFTOPSIS1and BFTOPSIS2 work fine here (no rank reversal)

Non classical MCDM problem

How to make a choice in $\mathcal{A}$ from multi-criteria scores expressed on power-set of $\mathcal{A}$?

$$
\mathbf{S} \triangleq
\begin{array}{c c}
X_i \in 2^{\mathcal{A}} & \begin{array}{ccccc} C_1, w_1 & \dots & C_j, w_j & \dots & C_N, w_N \end{array} \\
\begin{array}{l} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_M \\ \vdots \\ A_1 \cup A_2 \\ \vdots \\ A_1 \cup \ldots \cup A_i \cup \ldots \cup A_M \end{array} &
\begin{bmatrix}
S_{11} & \dots & S_{1j} & \dots & S_{1N} \\
 & & \vdots & & \\
S_{i1} & \dots & S_{ij} & \dots & S_{iN} \\
 & & \vdots & & \\
S_{M1} & \dots & S_{Mj} & \dots & S_{MN} \\
 & & \vdots & & \\
S_{(M+1)1} & \dots & S_{(M+1)j} & \dots & S_{(M+1)N} \\
 & & \vdots & & \\
S_{(2^M-1)1} & \dots & S_{(2^M-1)j} & \dots & S_{(2^M-1)N}
\end{bmatrix}
\end{array}
$$

See [Dezert Han Tacnet Carladous Yin 2016, Carladous 2017] for details

BBA construction for non classical MCDM

How to build $m(.) : 2^{\mathcal{A} \triangleq \{A_1, A_2, \ldots, A_M\}} \mapsto [0,1]$ from scores $\mathbf{S} \triangleq [S_{ij}]$?

Direct extension of BBA construction [Dezert Han Tacnet Carladous Yin 2016]

- Positive support of $X_i \in 2^{\mathcal{A}}$ based on all scores values of a criteria C_j

$$\mathrm{Sup}_j(X_i) \triangleq \sum_{Y \in 2^{\mathcal{A}} | S_j(Y) \leqslant S_j(X_i)} |S_j(X_i) - S_j(Y)|$$

$\mathrm{Sup}_j(X_i)$ measures how much X_i is better (higher) than other Y of $2^{\mathcal{A}}$

- Negative support of $X_i \in 2^{\mathcal{A}}$ based on all scores values of a criteria C_j

$$\mathrm{Inf}_j(X_i) \triangleq - \sum_{Y \in 2^{\mathcal{A}} | S_j(Y) \geqslant S_j(X_i)} |S_j(X_i) - S_j(Y)|$$

$\mathrm{Inf}_j(X_i)$ measures how much X_i is worse (lower) than other Y of $2^{\mathcal{A}}$

Belief function modeling

$$0 \leqslant \frac{\mathrm{Sup}_j(X_i)}{X^j_{\max}} \leqslant 1 - \frac{\mathrm{Inf}_j(X_i)}{X^j_{\min}} \leqslant 1 \Rightarrow \begin{cases} \mathrm{Bel}_{ij}(X_i) \triangleq \frac{\mathrm{Sup}_j(X_i)}{X^j_{\max}}, & \text{with } X^j_{\max} = \max_i \mathrm{Sup}_j(X_i) \\ \mathrm{Bel}_{ij}(\bar{X}_i) \triangleq \frac{\mathrm{Inf}_j(X_i)}{X^j_{\min}}, & \text{with } X^j_{\min} = \min_i \mathrm{Inf}_j(X_i) \end{cases}$$

Example 1

Five students $A_1, \dots, A_5$ have to be ranked based on two criteria

- C_1 = long jump performance
- C_2 = collected funds for an animal protection project

The scores are given as follows

$$
\mathbf{S} = \begin{array}{l} X_i \in 2^{\mathbf{A}} \\ A_1 \\ A_3 \\ A_4 \\ A_5 \\ A_1 \cup A_2 \\ A_3 \cup A_4 \end{array}
\begin{array}{c} \begin{array}{cc} C_1, w_1 & C_2, w_2 \end{array} \\ \begin{bmatrix} 3.7\,\text{m} & \emptyset \\ 3.6\,\text{m} & \emptyset \\ 3.8\,\text{m} & \emptyset \\ 3.7\,\text{m} & 640\,€ \\ \emptyset & 600\,€ \\ \emptyset & 650\,€ \end{bmatrix} \end{array}
$$

Difficulties:

- Scores are given in different units and different scales
- Some scores values can be missing
- Criteria C_j do not have same weights of importance w_j (in general)

Step 1: BBA matrix construction

$$
\mathbf{S} = \begin{array}{c} \text{FE} \in 2^{\mathcal{A}} \\ A_1 \\ A_3 \\ A_4 \\ A_5 \\ A_1 \cup A_2 \\ A_3 \cup A_4 \end{array}
\begin{array}{c} \begin{array}{cc} C_1, w_1 & C_2, w_2 \end{array} \\ \begin{bmatrix} 3.7\,\text{m} & \emptyset \\ 3.6\,\text{m} & \emptyset \\ 3.8\,\text{m} & \emptyset \\ 3.7\,\text{m} & 640\,€ \\ \emptyset & 600\,€ \\ \emptyset & 650\,€ \end{bmatrix} \end{array}
\Rightarrow \mathbf{M} = \begin{array}{c} \begin{array}{cc} C_1, w_1 & C_2, w_2 \end{array} \\ \begin{bmatrix} (0.25, 0.25, 0.50) & (0, 0, 1) \\ (0, 1, 0) & (0, 0, 1) \\ (1, 0, 0) & (0, 0, 1) \\ (0.25, 0.25, 0.50) & (0.6667, 0.1111, 0.2222) \\ (0, 0, 1) & (0, 1, 0) \\ (0, 0, 1) & (1, 0, 0) \end{bmatrix} \end{array}
$$

Step 2: distances to ideal best and worst solutions

Focal elem.	$d_{BI}(m_{i1}, m^{best})$	$d_{BI}(m_{i1}, m^{worst})$	$d_{BI}(m_{i2}, m^{best})$	$d_{BI}(m_{i2}, m^{worst})$
A_1	0.6016	0.2652	0.7906	0.2041
A_3	0.8416	0	0.7906	0.2041
A_4	0	0.8416	0.7906	0.2041
A_5	0.6016	0.2652	0.2674	0.5791
$A_1 \cup A_2$	0.5401	0.3536	0.6770	0
$A_3 \cup A_4$	0.5401	0.3536	0	0.6770

Steps 3-5: weighted distances with $w_1 = 1/3$ and $w_2 = 2/3$, closeness and ranking

Focal elem.	$d^{\text{best}}(X_i)$	$d^{\text{worst}}(X_i)$	$C(X_i, X^{\text{best}})$	Ranking
A_1	0.7276	0.2245	0.2358	4
A_3	0.8076	0.1361	0.1442	6
A_4	0.5270	0.4166	0.4415	3
A_5	0.3788	0.4745	0.5561	2
$A_1 \cup A_2$	0.6314	0.1179	0.1573	5
$A_3 \cup A_4$	0.1800	0.5692	0.7597	1

A more concrete example of non classical MCDM problem

Application: Protecting housing areas against torrential floods

Presented in [Dezert Han Tacnet Carladous Yin 2016, Carladous 2017]

List of alternatives (possible actions to take)

- A_1= maintenance of check dams' series
- A_2= no maintenance, but build a sediment trap upstream
- A_3= make individual protections to limit damage on buildings

List of criteria

- C_1 (in €) = investment cost (in negative values)
- C_2 (in €) = risk reduction in 50 years between the current situation and expected situation with the chosen action
- C_3 (in {1,2,. . . ,10}) = impact on environment
- C_4 (in m^2) = the land-use areas needed in privates

Score matrix (the higher is the score, the better is the proposition)

$$
\mathbf{S} = \begin{array}{l} \\ A_1 \\ A_2 \\ A_3 \\ A_1 \cup A_2 \\ A_1 \cup A_3 \\ A_2 \cup A_3 \\ A_1 \cup A_2 \cup A_3 \end{array}
\begin{array}{c} \begin{array}{cccc} C_1, w_1 = 0.33 & C_2, w_2 = 0.33 & C_3, w_3 = 0.20 & C_4, w_4 = 0.14 \end{array} \\
\begin{bmatrix}
-150000 & 100000 & 10 & 0 \\
-500000 & 200000 & 2 & -20000 \\
-550000 & 250000 & 10 & -5000 \\
-650000 & 230000 & 2 & -20000 \\
-700000 & 250000 & 10 & -5000 \\
-1050000 & 250000 & 2 & -25000 \\
-1200000 & 250000 & 2 & -25000
\end{bmatrix} \end{array}
$$

Note: the scores are not cumulative in the same way for each criterion. For C_1 and C_4, the score of the disjunction of two alternatives is the sum of individual scores whereas it is not the case for C_2 and C_3.

Here we apply BFTOPSIS1 method for its simplicity

Step 1: BBA construction from score matrix $\mathbf{S}$

$$\mathbf{M} = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_1 \cup A_2 \\ A_1 \cup A_3 \\ A_2 \cup A_3 \\ A_1 \cup A_2 \cup A_3 \end{array} \begin{array}{c} \begin{array}{cccc} C_1, w_1 & C_2, w_2 & C_3, w_3 & C_4, w_4 \end{array} \\ \begin{bmatrix} (1,0,0) & (0,1,0) & (1,0,0) & (1,0,0) \\ (0.44,0.10,0.46) & (0.45,0.28,0.27) & (0,1,0) & (0.10,0.67,0.23) \\ (0.37,0.13,0.50) & (1,0,0) & (1,0,0) & (0.70,0.07,0.23) \\ (0.27,0.21,0.52) & (0.73,0.10,0.17) & (0,1,0) & (0.10,0.67,0.23) \\ (0.23,0.26,0.51) & (1,0,0) & (1,0,0) & (0.70,0.07,0.23) \\ (0.04,0.75,0.21) & (1,0,0) & (0,1,0) & (0,1,0) \\ (0,1,0) & (1,0,0) & (0,1,0) & (0,1,0) \end{bmatrix} \end{array}$$

Steps 2-5: weighted distances with $w_1 = w_2 = 0.33$, $w_3 = 0.20$, $w_4 = 0.14$, closeness and ranking

Focal elem. X_i	$d^{\text{best}}(X_i)$	$d^{\text{worst}}(X_i)$	$C(X_i, X^{\text{best}})$	Ranking
A_1	0.3012	0.6116	0.6700	3
A_2	0.5668	0.3677	0.3935	6
A_3	0.1830	0.7483	0.8035	2
$A_1 \cup A_2$	0.4476	0.4901	0.5226	4
$A_1 \cup A_3$	0.1555	0.7775	0.8333	1
$A_2 \cup A_3$	0.5562	0.3614	0.3938	5
$A_1 \cup A_2 \cup A_3$	0.8328	0.2694	0.2444	7

Final ranking: best action(s) to take

$$(A_1 \cup A_3) \succ A_3 \succ A_1 \succ (A_1 \cup A_2) \succ (A_2 \cup A_3) \succ A_2 \succ (A_1 \cup A_2 \cup A_3)$$

Toolboxes for working with belief functions

- To start working with BF, we recommend Smets TBM MatLab codes that include many useful efficient functions based on Fast Möbius Transforms

  http://iridia.ulb.ac.be/~psmets/

- Main toolboxes for working with BF can be found from Belief Functions and Applications Society (www.bfasociety.org) wiki webpage at

  http://bfaswiki.iut-lannion.fr/wiki/index.php/Toolboxes

- Explanations for implementation of generalized belief functions can be found in

  A. Martin, Implementing general belief function framework with a practical codification for low complexity, in [DSmT books], Vol. 3, Chap 7, 2009.

- Implementation of fusion rules by sampling techniques (java package)
  http://refereefunction.fredericdambreville.com

Part III

Bibliography and biographies

Useful References

- J. Abellan, A. Masegosa, Requirements for total uncertainty measures in DempsterÐShafer theory of evidence, Int. J. Gen.Syst., Vol. 37(6), 2008.
- M. Beynon, DS/AHP method: A mathematical analysis, including an understanding of uncertainty, Eur. J. of Oper. Res., Vol. 140, 2002.
- M. Bouchard, A.-L. Jousselme, P.-E. Doré, A Proof for the Positive Definiteness of the Jaccard Index Matrix, IJAR, Vol. 54, 2013.
- A. Brodzik, R. Enders, A case of combination of evidence in the Dempster-Shafer theory inconsistent with evaluation of probabilities, 2011. `https://arxiv.org/pdf/1107.0082.pdf`
- S. Carladous, J.-M. Tacnet, J. Dezert, D. Han, M. Batton-Hubert, Applying ER-MCDA and BF-TOPSIS to Decide on Effectiveness of Torrent Protection, Proc. of Belief 2016.
- S. Carladous, Approche intégrée d'aide à la décision basée sur la propagation de l'imperfection de l'information - Application à l'efficacité des mesures de protection torrentielle, Ph.D. Thesis, Lyon Univ., France, April 2017.
- B. Cobb, P. Shenoy, On the plausibility transformation method for translating belief function models to probability models, IJAR, Vol. 41, 2006.
- F. Dambreville, definition of evadence fusion rules based on referee functions, in [DSmT books], Vol. 3, Chap. 6, 2009.
- A.-P. Dempster, Upper and lower probabilities induced by a multivalued mapping, A. of Math. Stat., Vol. 38, 1967.
- J. Dezert, F. Smarandache, Introduction to the Fusion of Quantitative and Qualitative Beliefs, in Information & Security J., Vol. 20, 2006.
- J. Dezert, F. Smarandache, A new probabilistic transformation of belief mass assignment, Proc. of Fusion 2008.
- J. Dezert, J.-M. Tacnet, M. Batton-Hubert, F. Smarandache, Multi-criteria decision making based on DSmT/AHP, In Proc. of Int. Workshop on Belief Functions, Brest, France, 2010.

Useful References

J. Dezert, J.-M. Tacnet, Evidential Reasoning for Multi-Criteria Analysis based on DSmT-AHP, ISAHP 2011, Italy, June 2011.

J. Dezert, A. Tchamova, On the behavior of Dempster's rule of combination, Spring school on BFTA, Autrans, France, April 2011. http://hal.archives-ouvertes.fr/hal-00577983/

J. Dezert, P. Wang, A. Tchamova, On The Validity of Dempster-Shafer Theory, Proc. of Fusion 2012.

J. Dezert, D. Han, Z. Liu Z., J.-M. Tacnet, Hierarchical DSmP transformation for decision-making under uncertainty, Proc. of Fusion 2012.

J. Dezert, D. Han, Z. Liu Z., J.-M. Tacnet, Hierarchical proportional redistribution for bba approximation, Proc. of Belief 2012.

J. Dezert, A. Tchamova, D. Han, J.-M. Tacnet, Why Dempster's fusion rule is not a generalization of Bayes fusion rule, Proc. of Fusion 2013.

J. Dezert, A. Tchamova, On the validity of Dempster's fusion rule and its interpretation as a generalization of Bayesian fusion rule, Int. J. of Intell. Syst., Vol. 29, March 2014.

J. Dezert, D. Han, H. Yin, A new belief function based approach for multi-criteria decision-making support, in Proc. of Fusion 2016.

J. Dezert, D. Han, J.-M. Tacnet, S. Carladous, Y. Yang, Decision-Making with Belief Interval Distance, Proc. of Belief 2016.

J. Dezert, D. Han, J.-M. Tacnet, S. Carladous, H. Yin, The BF-TOPSIS approach for solving non-classical MCDM problems, Proc. of Belief 2016.

J. Dezert, D. Han, J.-M. Tacnet, Multi-Criteria decision-Making with imprecise scores and BF-TOPSIS, in Proc. of Fusion 2017.

D. Dubois, H. Prade, A note on measures of specificity for fuzzy sets, Int. J. Gen. Syst., Vol. 10(4), 1985.

D. Dubois, H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, Computational Intelligence, Vol. 4, 1988

Useful References

- D. Dubois, H. Prade, Properties of measures of information in evidence and possibility theories, Fuzzy Sets Syst., Vol. 100, 1999.
- M. Florea, J. Dezert, P. Valin, F. Smarandache, A.-L. Jousselme, Adaptative combination rule and proportional conflict redistribution rule for information fusion, in: Proc. of COGISÕ06, Paris, 2006.
- M. Florea, A.-L. Jousselme, E. Bossé, D. Grenier, Robust combination rules for evidence theory, Information Fusion, Vol. 10, 2009.
- A. Gelman, The boxer, the wrestler, and the coin flip: a paradox of robust Bayesian inference and belief functions, American Statistician, Vol. 60, 2006.
- J.H. Halpern, R. Fagin, Two views of belief . . . , Artificial Intelligence, Vol. 54, 1992.
- D. Han, J. Dezert, J.-M. Tacnet, C. Han, A Fuzzy-Cautious OWA Approach with Evidential Reasoning, in Proc. Of Fusion 2012.
- D. Han, J. Dezert, Y. Yang, New Distance Measures of Evidence based on Belief Intervals, Proc. of Belief 2014, Oxford, 2014.
- D. Han, Novel approaches for the transformation of fuzzy membership function into basic probability assignment based on uncertainty optimization, Int. J. of Uncertainty Fuzziness and Knowledge Based Systems, Vol. 94, 2016.
- D. Han, J. Dezert, Y. Yang, Belief interval Based Distances Measures in the Theory of Belief Functions, IEEE Trans. on SMC, 2017.
- D. Harmanec, G.J. Klir, Measuring total uncertainty in Dempster-Shafer theory: a novel approach, Int. J. Gen. Syst.; Vol. 22(4), 1994.
- J.Heendeni et al., A Generalization of Bayesian Inference in the Dempster-Shafer Belief Theoretic Framework, in Proc. Fusion 2016.
- U. Höhle, Entropy with respect to plausibility measures, Proc. of 12th Int Symp. on Multiple-Valued Logic, 1982.
- C.L. Hwang, K. Yoon, Multiple Attribute Decision Making Methods and Applications, Springer, 1981

Useful References

- L. Jaulin, M. Kieffer, O. Didrit, E. Walter, Applied Interval Analysis, Springer, 2001.
- A.-L. Jousselme, D. Grenier, E. Bossé, A New Distance between Two Bodies of Evidence, Information Fusion, Vol. 2, 2001
- A.-L. Jousselme, C. Liu, D. Grenier, E. Bossé, Measuring ambiguity in the evidence theory, IEEE Trans. on SMC, Part A, Vol. 36(5), 2006.
- A.-L. Jousselme, P. Maupin, Distances in Evidence Theory: Comprehensive Survey and Generalizations, IJAR, Vol. 5, 2012.
- R. Kennes, Computational aspects of the Möbius transformation of graphs, IEEE SMC, Vol. 22,1992.
- A. Kirchner, F. Dambreville, F. Celeste, J. Dezert J., F. Smarandache, Application of probabilistic PCR5 fusion rule for multisensor target tracking, Proc. of Fusion 2007.
- G. Klir, A. Ramer, Uncertainty in the DempsterÐShafer theory: a critical re-examination, Int. J. Gen. Syst., Vol. 18(2), 1990.
- G. Klir, B. Parviz, A note on the measure of discord, Proc. of 8th UAI Conf., 1992.
- G. Klir, H.W. Lewis III, Remarks on ”Measuring ambiguity in the evidence theory”, IEEE Trans. on SMC, Part A, Vol. 38(4), 2008.
- J. Kohlas, S. Moral, Handbook of defeasible reasoning and uncertainty management systems, Springer, 2000.
- J. Lemmer, Confidence factors, empiricism and the Dempster-Shafer theory of evidence, in Proc. of UAI-85, pp. 160Ð176, 1985
- X. Li X., X. Dai, J. Dezert, F. Smarandache, Fusion of imprecise qualitative information, Applied Intel., Vol. 33, 2010.
- F. Li et al., Adaptive and robust evidence theory with applications in prediction of floor water inrush in coal mine, Trans. of the Inst. of Measurement and Control, Vol. 39 (4), 2017.

Useful References

A. Martin, C. Osswald, J. Dezert, F. Smarandache, General combination rules for qualitative and quantitative beliefs, JAIF, Vol.3, 2008.

D. Mercier, B. Quost, T. Denœux, Contextual discounting of belief functions, in Proc. of ECSQARU 2005.

D. Mercier, T. Denœux, M.-H. Masson, Refined sensor tuning in the belief function framework using contextual discounting, Proc. IPMU'2006.

A. Ramer, Uniqueness of information measure in the theory of evidence, Fuzzy Sets Syst., Vol. 24(2), 1987.

T. Saaty, The Analytical Hierarchy Process, McGraw Hill, 1980.

K. Sentz, S. Ferson, Combination of Evidence in Dempster-Shafer Theory, SANDIA Tech. rep. no. SAND2002-0835, 2002

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

F. Smarandache, An in-depth look at quantitative information fusion rules, in [DSmT books], Vol. 2, 2006.

F. Smarandache, J. Dezert, On the consistency of PCR6 with the averaging rule and its application to probability estimation, Proc. of Fusion 2013.

F. Smarandache, V. Kroumov, J. Dezert, Examples where the conjunctive and Dempster's rules are insensitive, Proc. of 2013 Int. Conf. on Advanced Mechatronic Syst., 2013.

F. Smarandache, J. Dezert (Eds), Advances and applications of DSmT for information fusion, Vols. 1-4, 2004, 2006, 2009 & 2015. `http://www.onera.fr/staff/jean-dezert?page=2`

F. Smarandache, J. Dezert, J.-M. Tacnet, Fusion of sources of evidence with different importances and reliabilities, Proc. of Fusion 2010.

F. Smarandache, J. Dezert, Modified PCR Rules of Combination with Degrees of Intersections, in Proc. of Fusion 2015.

Useful References

- P. Smets, The Combination of Evidence in the Transferable Belief Model, IEEE Trans. on PAMI, Vol. 12, 1990.
- P. Smets, R. Kennes, The Transferable Belief Model, Artificial Intel., Vol. 66, 1994.
- P. Smets, Matrix Calculus for Belief Functions, IJAR, Vol. 31, 2002.
- J.-M. Tacnet, J. Dezert, Cautious OWA and Evidential Reasoning for Decision Making under Uncertainty, in Proc. Of Fusion 2011.
- A. Tchamova, J. Dezert J., On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory, Proc. of IEEE IS'2012 Conf.
- B. Tessem, Approximations for Efficient Computation in the Theory of Evidence, Artificial Intel., Vol. 61, 1993.
- F. Voorbraak, On the justification of Dempster's rule of combination, Tech. Rep. 42, Dept. of Philosophy, Utrecht Univ., 1988.
- P. Wang, A defect in Dempster-Shafer theory, Proc. of 10th Conf. on Uncertainty in AI, 1994.
- Y.-M. Wang, Y. Luo, On rank reversal in decision analysis, Math. and Computer Modelling, Vol. 49, 2009.
- I. Wasserman, Comments on Shafer's "Perspectives on the theory and practice of belief functions", IJAR, Vol.6, 1992.
- N. Wilson, A Monte-Carlo algorithm for Dempster-Shafer belief, in Proc. of UAI'91, 1991.
- N. Wilson, Algorithms for Dempster-Shafer Theory, in [Kohlas Moral 2000], 2000.
- R. Yager, Entropy and specificity in a mathematical theory of evidence, Int. J. Gen. Syst., Vol. 9 (4), 1983.
- R. Yager, On the Dempster-Shafer Framework and New Combination Rules, Information Sciences, Vol. 41, 1987.

Useful References

- R. Yager, On ordered weighted averaging operators in multi-criteria decision making, IEEE Trans. on SMC, 1988.
- R. Yager, J. Kacprzyk, M. Fedrizzi, Advances in the Dempster-Shafer Theory of Evidence, John Wiley & Sons, New York, 1994.
- R. Yager, L. Liu, Classic Works of the Dempster-Shafer Theory of Belief Functions, Springer, 2008.
- Y. Yang, D. Han, A new distance-based total uncertainty measure in the theory of belief functions, Knowledge-Based Systems, Vol. 94, 2016.
- H. Yin, J. Lan, X.-R. Li, Measures for ranking estimation performance based on single or multiple performance metrics, Proc. of Fusion 2013.
- L. Zadeh, On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, 1979.
- L. Zhang, Representation, independence, and combination of evidence in the Dempster-Shafer theory, in [Yager et al. 1994]

Jean Dezert was born in France on August 25, 1962. He got his Ph.D. from Paris XI Univ., Orsay, in 1990 in Automatic Control and Signal Processing. During 1986-1990 he was with the Syst. Dept. at ONERA and did research in multi-sensor multi-target tracking (MS-MTT). During 1991-1992, he was post-doc at ESE dept., UConn, USA under supervision of Prof. Bar-Shalom. During 1992-1993 he was teaching assistant in EE Dept, Orléans Univ., France. Since 1993, he is Senior Research Scientist at ONERA. His research interests include estimation, information fusion, reasoning under uncertainty, and multi-criteria decision-making support. He has organized Fusion conference in Paris in 2000 and has been TPC member of Fusion 2000-2017 conferences. He served as ISIF 2016 President. Dr. Dezert published more than 150 papers in conferences and journals, and edited four books on Dezert-Smarandache Theory (DSmT) with Prof. Smarandache.

Web page: `http://www.onera.fr/staff/jean-dezert`

Deqiang Han Short Biography

Deqiang Han, Ph.D., Professor
Center for Information Eng. Science
Xi'an Jiaotong University
Xi'an, China 710049
Email: deqhan@mail.xjtu.edu.cn

Deqiang Han was born in Xi'an, China, in 1980. He received the bachelor's degree in Communication and Control Engineering and the master's and Ph.D. degrees in Control Science and Engineering from Xi'an Jiaotong University, Xi'an, in 2001, 2004, and 2008, respectively. He is currently a Professor with Xi'an Jiaotong University. His current research interests include evidence theory, information fusion, and pattern classification. Dr. Han is an International Society of Information Fusion (ISIF) Member and Technical Program Committee Member for the 1st–7th Chinese Conference on Information Fusion in 2009–2015, and 2017. He is a member of organization committee for FUSION 2017 held in Xi'an, China.