

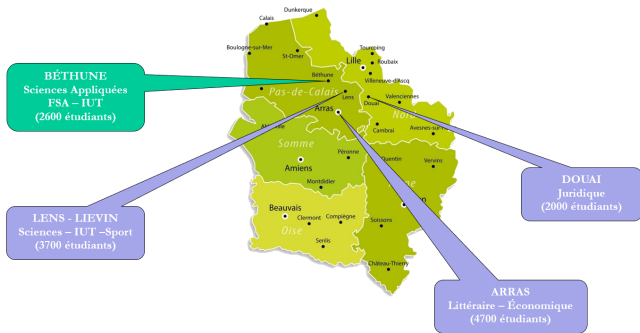
Tutoriel : Introduction à la théorie des fonctions de croyance

David Mercier & Éric Lefèvre

Université d'Artois, LGI2A

Tutoriel PFIA 2026, jeudi 2 juillet 2026

► Université d'Artois : pluridisciplinaire et multipolaire



► LGI2A :

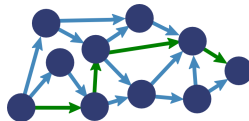
- Création en 2000
- 17 permanents, 1 assistante administrative, 1 ingénieur permanent, 10 doctorants, 5 ATER, 3 ingénieurs et 2 post-docs

- Thématique centrale : Aide à la décision

- Deux thèmes scientifiques :



Décision et Fusion d'Informations



Optimisation des Systèmes complexes

- Deux domaines d'application :



Logistique Durable



Mobilité Intelligente

- ▶ La théorie des **Fonctions de Croyance (FC)** est un **cadre mathématique flexible** qui permet de gérer les **informations imparfaites**.
- ▶ Autres dénominations : la **théorie de Dempster-Shafer** ou la **théorie de l'évidence**.
- ▶ Elle a été introduite par **A. P. Dempster** dans les années 60 lors de ses travaux sur l'inférence statistique et ensuite formalisée par **G. Shafer** en 1976.
- ▶ Cette théorie est utilisée dans de nombreuses applications : IA (apprentissage automatique, systèmes experts), fusion d'informations, traitement d'images,...
- ▶ Le raisonnement évidentiel peut donc être employé dans de **nombreux problèmes**.

Différents types d'imperfection de données

- ▶ **Incertitude** : quantifie la véracité de l'information.
Ex. : "Je pense que John mesure 1.80m"
Ici, l'information est précise mais incertaine
- ▶ **Imprécision** : quantifie la qualité de l'information.
Ex. : "John mesure entre 1.70m et 1.90m"
Ici, l'information est certaine mais imprécise.
- ▶ **Incertitude et imprécision** :
Ex. : "Je pense que John mesure entre 1.70m et 1.90m"
Ici, l'information est à la fois incertaine et imprécise.

Différents types d'imperfection de données

- ▶ **Incertitude** : quantifie la véracité de l'information.
Ex. : "Je pense que John mesure 1.80m"
Ici, l'information est précise mais incertaine
- ▶ **Imprécision** : quantifie la qualité de l'information.
Ex. : "John mesure entre 1.70m et 1.90m"
Ici, l'information est certaine mais imprécise.
- ▶ **Incertitude et imprécision** :
Ex. : "Je pense que John mesure entre 1.70m et 1.90m"
Ici, l'information est à la fois incertaine et imprécise.

- ▶ Incertitude : classiquement traitée avec des probabilités.
- ▶ Imprécision : classiquement traitée avec des ensembles.
- ▶ La théorie des ensembles et la théorie des probabilités sont des cas particuliers de la théorie des fonctions de croyance.

Incertitudes **aléatoires** (hasard)

vs.

Incertitudes **épistémiques** (absence
d'information)



Le Derby d'Epsom (Géricault)

- Considérons une course de chevaux avec 3 chevaux participants : h_1 , h_2 et h_3 .

Expert : "Tous les chevaux ont la même chance de gagner"

Modélisation :

$$p_E(\{h_1\}) = p_E(\{h_2\}) = p_E(\{h_3\}) = \frac{1}{3}$$

Novice : "Je ne connais rien aux courses de chevaux, j'ignore qui va gagner"

Modélisation :

$$p_N(\{h_1\}) = p_N(\{h_2\}) = p_N(\{h_3\}) = \frac{1}{3}$$

- **Problème** : Deux informations différentes sont représentées par la même modélisation.
- Il est nécessaire d'avoir un modèle plus complet.

Représentation de l'information

Combinaison de l'information

Décision

Conclusion

Représentation de l'information

Combinaison de l'information

Décision

Conclusion

Représentation de l'information

Définition - Fonction de masses

- Considérons une variable d'intérêt X prenant ses valeurs dans un ensemble fini d'hypothèses $\Omega = \{\omega_1, \dots, \omega_N\}$ appelé univers ou **cadre de discernement**.
 - Ex. : le cheval qui gagnera la course : $\Omega = \{h_1, h_2, h_3\}$

Représentation de l'information

Définition - Fonction de masses

- ▶ Considérons une variable d'intérêt X prenant ses valeurs dans un ensemble fini d'hypothèses $\Omega = \{\omega_1, \dots, \omega_N\}$ appelé univers ou **cadre de discernement**.
 - ▶ Ex. : le cheval qui gagnera la course : $\Omega = \{h_1, h_2, h_3\}$
- ▶ Une information concernant la valeur ω_0 que peut prendre cette variable peut être représentée en utilisant une **fonction de masse (MF)** m définie comme un mapping $m : 2^\Omega \rightarrow [0, 1]$ telle que :

$$\sum_{A \subseteq \Omega} m(A) = 1.$$

Représentation de l'information

Définition - Fonction de masses

- ▶ Considérons une variable d'intérêt X prenant ses valeurs dans un ensemble fini d'hypothèses $\Omega = \{\omega_1, \dots, \omega_N\}$ appelé univers ou **cadre de discernement**.
 - ▶ Ex. : le cheval qui gagnera la course : $\Omega = \{h_1, h_2, h_3\}$
- ▶ Une information concernant la valeur ω_0 que peut prendre cette variable peut être représentée en utilisant une **fonction de masse (MF)** m définie comme un mapping $m : 2^\Omega \rightarrow [0, 1]$ telle que :

$$\sum_{A \subseteq \Omega} m(A) = 1.$$

- ▶ $m(A)$: **part de croyance allouée à A et qui ne peut pas être attribuée à un sous-ensemble plus petit que A .**

Représentation de l'information

Définition - Fonction de masses

- ▶ Considérons une variable d'intérêt X prenant ses valeurs dans un ensemble fini d'hypothèses $\Omega = \{\omega_1, \dots, \omega_N\}$ appelé univers ou **cadre de discernement**.
 - ▶ Ex. : le cheval qui gagnera la course : $\Omega = \{h_1, h_2, h_3\}$
- ▶ Une information concernant la valeur ω_0 que peut prendre cette variable peut être représentée en utilisant une **fonction de masse (MF)** m définie comme un mapping $m : 2^\Omega \rightarrow [0, 1]$ telle que :

$$\sum_{A \subseteq \Omega} m(A) = 1.$$

- ▶ $m(A)$: **part de croyance** allouée à A et qui ne peut pas être attribuée à un sous-ensemble plus petit que A .
- ▶ Les sous-ensembles A t.q. $m(A) > 0$ sont appelés **éléments focaux** de m .

Représentation de l'information

Fonction de masses - Cas particuliers

- ▶ La fonction de masse m représente l'état de connaissance d'un agent à un instant t .
- ▶ $m(A) = 1 \ \forall A \subseteq \Omega$: fonction de masse catégorique
- ▶ $m(\Omega)$: représente le degré d'ignorance
 - ▶ $m(\Omega) = 1$: ignorance totale (fonction de masse vide)
- ▶ Si tous les éléments focaux de m sont des singletons, m est alors appelée fonction de masse bayésienne. Elle est équivalente à une distribution de probabilité.
- ▶ La fonction de masse peut alors être vue comme :
 - ▶ la généralisation des ensembles
 - ▶ la généralisation d'une distribution de probabilité.



Le Derby d'Epsom (Géricault)

- Considérons une course de chevaux avec 3 chevaux participants : h_1 , h_2 et h_3 .

Expert : "Tous les chevaux ont la même chance de gagner"

Modélisation :

$$m_E(\{h_1\}) = m_E(\{h_2\}) = m_E(\{h_3\}) = \frac{1}{3}$$

Novice : "Je ne connais rien aux courses de chevaux, j'ignore qui va gagner"

Modélisation :

$$m_N(\{h_1, h_2, h_3\}) = 1$$

- ▶ Problème
 - ▶ Un juge sait que :
 - ▶ Big Boss a décidé que M. Jones devait mourir ;
 - ▶ 3 tueurs possibles : Peter, Paul et Mary;
 - ▶ Big Boss désigne à pile ou face le sexe du tueur ;
 - ▶ il n'y a aucune précision sur le choix entre Paul ou Peter dans le cas où c'est un homme qui est choisi ;
 - ▶ M. Jones est tué par l'un des tueurs de Big Boss.
 - ▶ Comment représenter ces connaissances ?
- ▶ Solution à l'aide de fonction de croyance :
 - ▶ Le tueur appartient à $\Omega = \{Peter, Paul, Mary\}$
 - ▶ $m(\{Peter, Paul\}) = 0.5$ et $m(\{Mary\}) = 0.5$

Représentation de l'information

Autres représentations - Fonctions de croyance et de plausibilité

- Une fonction de masse est l'une des manières de représenter l'information. D'autres fonctions existent :

- La **fonction de croyance** Bel définie par :

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad \forall A \subseteq \Omega$$

$Bel(A)$ représente le degré de croyance total supportant le fait que la solution au problème appartienne à A .

- La **fonction de plausibilité** Pl définie par :

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) = 1 - Bel(\bar{A}) \quad \forall A \subseteq \Omega$$

avec $\bar{A} = \Omega \setminus A$

$Pl(A)$ représente la somme totale des croyances qui ne contredisent pas A (qui sont consistantes avec A).

Représentation de l'information

Exemple de calculs des plausibilités et des croyances

Avec $\Omega = \{a, b, c\}$ et $m(\{a\}) = 0.3$, $m(\{b\}) = 0.4$ et $m(\Omega) = 0.3$

- Exemple : calcul de $Bel(\{a, b\})$

$$\begin{aligned} Bel(\{a, b\}) &= \sum_{B \subseteq \{a, b\}} m(B) \\ &= m(\{a\}) + m(\{b\}) + m(\{a, b\}) \\ &= 0.3 + 0.4 + 0 = 0.7 \end{aligned}$$

- Exemple : calcul de $Pl(\{a, b\})$

$$\begin{aligned} Pl(\{a, b\}) &= \sum_{B \cap \{a, b\} \neq \emptyset} m(B) \\ &= m(\{a\}) + m(\{b\}) + m(\{a, b\}) \\ &\quad + m(\{a, c\}) + m(\{b, c\}) + m(\{a, b, c\}) \\ &= 0.3 + 0.4 + 0 + 0 + 0 + 0.3 = 1 \end{aligned}$$

Représentation de l'information

Calcul des plausibilités et des croyances

Avec $\Omega = \{a, b, c\}$ et $m(\{a\}) = 0.3$, $m(\{b\}) = 0.4$ et $m(\Omega) = 0.3$

Ordre binaire		m	Bel	Pl
000	$\emptyset$			
001	$\{a\}$			
010	$\{b\}$			
011	$\{a, b\}$			
100	$\{c\}$			
101	$\{a, c\}$			
110	$\{b, c\}$			
111	$\Omega = \{a, b, c\}$			

Représentation de l'information

Calcul des plausibilités et des croyances

Avec $\Omega = \{a, b, c\}$ et $m(\{a\}) = 0.3$, $m(\{b\}) = 0.4$ et $m(\Omega) = 0.3$

Ordre binaire		m	Bel	Pl
000	$\emptyset$			
001	$\{a\}$	.3		
010	$\{b\}$	.4		
011	$\{a, b\}$			
100	$\{c\}$			
101	$\{a, c\}$			
110	$\{b, c\}$			
111	$\Omega = \{a, b, c\}$	.3		

Représentation de l'information

Calcul des plausibilités et des croyances

Avec $\Omega = \{a, b, c\}$ et $m(\{a\}) = 0.3$, $m(\{b\}) = 0.4$ et $m(\Omega) = 0.3$

Ordre binaire		m	Bel	Pl
000	$\emptyset$			
001	$\{a\}$	.3	.3	
010	$\{b\}$	.4	.4	
011	$\{a, b\}$		.7	
100	$\{c\}$			
101	$\{a, c\}$		.3	
110	$\{b, c\}$		.4	
111	$\Omega = \{a, b, c\}$	.3	1	

Représentation de l'information

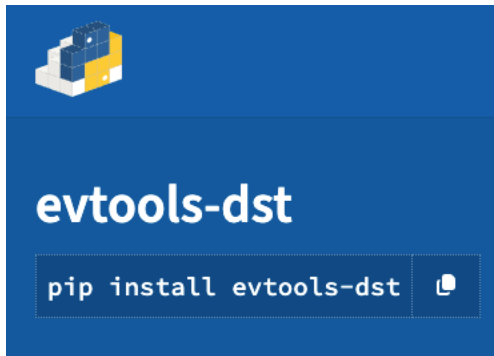
Calcul des plausibilités et des croyances

Avec $\Omega = \{a, b, c\}$ et $m(\{a\}) = 0.3$, $m(\{b\}) = 0.4$ et $m(\Omega) = 0.3$

Ordre binaire		m	Bel	Pl
000	$\emptyset$			
001	$\{a\}$	.3	.3	.6
010	$\{b\}$	.4	.4	.7
011	$\{a, b\}$		.7	1
100	$\{c\}$			.3
101	$\{a, c\}$		.3	.6
110	$\{b, c\}$		.4	.7
111	$\Omega = \{a, b, c\}$	.3	1	1

Représentation de l'information

Evidence Theory Tools — une bibliothèque Python pour travailler avec les fonctions de croyance



Représentation de l'information

Avec evtools

```
>>> from evtools.dsvector import DSVector, Kind
>>> frame = ["a", "b", "c"]
>>> m = DSVector.from_focal(frame, {"a": 0.3, "b": 0.4})
>>> m
DSVector  kind=m  (Basic Belief Assignment)  frame={a, b, c}  3 focal elements
```

Subset	m	
{a}	0.3000	
{b}	0.4000	
{a, b, c}	0.3000	
Total	1.0000	

Représentation de l'information

Avec evtools

```
>>> m.to_bel()
DSVector kind=bel (Belief function) frame={a, b, c} 6 focal elements
```

Subset	bel	
{a}	0.3000	
{b}	0.4000	
{a, b}	0.7000	
{a, c}	0.3000	
{b, c}	0.4000	
{a, b, c}	1.0000	

```
>>> m.to_bel().dense
array([0. , 0.3, 0.4, 0.7, 0. , 0.3, 0.4, 1. ])
```

Représentation de l'information

Propriétés de Bel et Pl

- ▶ $Bel(\emptyset) = Pl(\emptyset) = 0$
- ▶ $Bel(\Omega) \leq 1$ et $Pl(\Omega) \leq 1$ (dans le cas où $m(\emptyset) > 0$)
- ▶ $Bel(A) \leq Pl(A)$
- ▶ $Pl(A) = 1 - Bel(\bar{A})$
- ▶ si m est une fonction de masse bayésienne (c-à-d que les éléments focaux sont des singletons) alors $Bel = Pl$ est une **mesure de probabilité**.

Représentation de l'information

Intervalle $[Bel(A), Pl(A)]$

- ▶ L'**incertitude**, à propos d'une proposition A , est représentée par deux valeurs $Bel(A)$ et $Pl(A)$ (avec $Bel(A) \leq Pl(A)$)
- ▶ L'intervalle $[Bel(A), Pl(A)]$ est **maximal** (c-à-d égale à $[0, 1]$) si m est une **fonction de masse vide** ($m(\Omega) = 1$). Dans ce cas :

$$Bel(A) = 0 \quad \forall A \neq \Omega \quad \text{et} \quad Pl(A) = 1 \quad \forall A \neq \emptyset$$

- ▶ L'intervalle $[Bel(A), Pl(A)]$ est **minimal** ($Bel(A) = Pl(A)$) si m est une fonction de **masse bayésienne**, alors :

$$Bel(A) = Pl(A) = \sum_{\omega \in \Omega} m(\omega)$$

Dans ce cas, Bel et Pl sont des mesures de probabilités.

Représentation de l'information

Exemple avec evtools

```
>>> m = DSVector.from_focal(frame, {"a": 0.3, "b": 0.4, "c": 0.3})
```

DSVector kind=m (Basic Belief Assignment) frame={a, b, c} 3 focal elements

Subset	m	
{a}	0.3000	
{b}	0.4000	
{c}	0.3000	
Total	1.0000	

```
>>> m.dense
array([0. , 0.3, 0.4, 0. , 0.3, 0. , 0. , 0. ])
>>> m.to_bel().dense
array([0. , 0.3, 0.4, 0.7, 0.3, 0.6, 0.7, 1. ])
>>> m.to_pl().dense
array([0. , 0.3, 0.4, 0.7, 0.3, 0.6, 0.7, 1. ])
```

Représentation de l'information

Cas d'une fonction consonante

- ▶ Dans le cas d'**éléments focaux emboîtés** ($A_1 \subset A_2 \subset \dots \subset A_n$), la fonction m est dite **consonante**
- ▶ Dans ce cas, $\forall A \subseteq \Omega, \forall B \subseteq \Omega$:

$$Bel(A \cap B) = \min(Bel(A), Bel(B))$$

et

$$Pl(A \cup B) = \max(Pl(A), Pl(B))$$

ainsi

- ▶ Pl correspond à la **mesure de possibilité**.
- ▶ Bel correspond à la **mesure de nécessité**.

- ▶ Soit une source S , on peut associer à cette source un degré de fiabilité $\beta = (1 - \alpha)$ avec $\alpha \in [0, 1]$
- ▶ La fonction de masse produite par cette source S peut alors être affaiblie :

$$\begin{cases} {}^{\alpha}m(A) &= (1 - \alpha)m(A), \quad \forall A \subset \Omega \\ {}^{\alpha}m(\Omega) &= (1 - \alpha)m(\Omega) + \alpha \end{cases}$$

- ▶ $\alpha = 0$: ${}^{\alpha}m = m$ (non affaiblie)
- ▶ $\alpha = 1$: ${}^{\alpha}m(\Omega) = 1$ (totalement affaiblie)

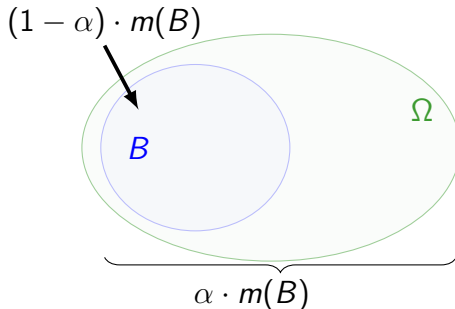
Représentation de l'information

Exemple d'affaiblissement

- ▶ Soit la fonction de masse précédente : $m(\{Peter, Paul\}) = 0.5$ et $m(\{Mary\}) = 0.5$
- ▶ On dispose d'une information complémentaire qui nous dit que les connaissances du juge ne sont pas fiables. Il estime qu'elles sont fiables à 0.8
- ▶ Avec cette information, la fonction de masse m peut être corrigée de la façon suivante :

$$\begin{cases} {}^{\alpha}m(\{Peter, Paul\}) &= 0.8 \times m(\{Peter, Paul\}) = 0.4 \\ {}^{\alpha}m(\{Mary\}) &= 0.8 \times m(\{Mary\}) = 0.4 \\ {}^{\alpha}m(\Omega) &= 0.8 \times m(\Omega) + 0.2 = 0.2 \end{cases}$$

Pour chaque élément focal B de m :



- ▶ Une partie $(1 - \alpha) \cdot m(B)$ reste sur B .
- ▶ Une partie $\alpha \cdot m(B)$ est transférée sur Ω .

Représentation de l'information

Combinaison de l'information

Décision

Conclusion

Combinaison de l'information

Règle de combinaison conjonctive

- ▶ Soit deux fonctions de masse m_1 et m_2 issues de deux sources d'informations **fiables** et **distinctes**.
- ▶ Ces informations peuvent être combinées à l'aide de la **Combinaison conjonctive (CRC)** définie par :

$$m_{1\odot 2}(A) = m_1 \odot m_2(A) = \sum_{B \cap C = A} m_1(B)m_2(C) \quad \forall A \subseteq \Omega$$

Combinaison de l'information

Exemple combinaison conjonctive

Soient deux fonctions de masses m_1 et m_2 définies sur $\Omega = \{a, b, c\}$:

$$\begin{cases} m_1(\{a\}) &= 0.4 \\ m_1(\{a, b\}) &= 0.6 \end{cases} \quad \begin{cases} m_2(\{b, c\}) &= 0.3 \\ m_2(\{a, b, c\}) &= 0.7 \end{cases}$$

On considère que les deux sources qui ont produit ces informations sont fiables et indépendantes.

CRC	$m_2(\{b, c\}) = .3$	$m_2(\{a, b, c\}) = .7$
$m_1(\{a\}) = .4$	$a \cap \{b, c\} = \emptyset$ $.4 \times .3 = .12$	$a \cap \{a, b, c\} = a$ $.4 \times .7 = .28$
$m_1(\{a, b\}) = .6$	$\{a, b\} \cap \{b, c\} = b$ $.6 \times .3 = .18$	$\{a, b\} \cap \{a, b, c\} = \{a, b\}$ $.6 \times .7 = .42$

Le résultat $m_{\odot}$ de la combinaison conjonctive est :

$$\begin{aligned} m_{\odot}(\{a\}) &= 0.4 \times 0.7 = 0.28 \\ m_{\odot}(\{b\}) &= 0.6 \times 0.3 = 0.18 \\ m_{\odot}(\{a, b\}) &= 0.6 \times 0.7 = 0.42 \\ m_{\odot}(\emptyset) &= 0.4 \times 0.3 = 0.12 \end{aligned}$$

Combinaison de l'information

Règle de combinaison de Dempster

- ▶ Si et seulement si m_1 et m_2 sont deux fonctions de masse issues de sources d'information **fiables** et **distinctes** (justification axiomatique de Smets en 2007)
- ▶ **Règle de Dempster** := **CRC normalisée** ($m_{1\oplus 2}(\emptyset) = 0$)
- ▶ On peut utiliser la règle de combinaison de Dempster :

$$\begin{aligned}m_{1\oplus 2}(A) &= (m_1 \oplus m_2)(A) \\&= \frac{1}{1-\kappa} \sum_{B \cap C = A} m_1(B).m_2(C) \\&= \frac{1}{1-\kappa} m_{1\odot 2}(A) \quad \forall A \subseteq \Omega, A \neq \emptyset\end{aligned}$$

avec :

$$\kappa = m_{1\odot 2}(\emptyset) \quad (\text{appelé } \textbf{degré de conflit})$$

Combinaison de l'information

Exemple avec la combinaison de Dempster

Dempster	$m_2(\{b, c\}) = .3$	$m_2(\{a, b, c\}) = .7$
$m_1(\{a\}) = .4$	$a \cap \{b, c\} = \emptyset$ $.4 \times .3 = .12$	$a \cap \{a, b, c\} = a$ $.4 \times .7 = .28$
$m_1(\{a, b\}) = .6$	$\{a, b\} \cap \{b, c\} = b$ $.6 \times .3 = .18$	$\{a, b\} \cap \{a, b, c\} = \{a, b\}$ $.6 \times .7 = .42$

Le résultat $m_{\oplus}$ de la combinaison de Dempster est :

- ▶ $m_{\oplus}(\{a\}) = \frac{0.4 \times 0.7}{1 - 0.12} = 0.318$
- ▶ $m_{\oplus}(\{b\}) = \frac{0.6 \times 0.3}{1 - 0.12} = 0.204$
- ▶ $m_{\oplus}(\{a, b\}) = \frac{.6 \times .7}{1 - 0.12} = 0.478$

- Si les sources sont **distinctes et au moins l'une d'elles est fiable** (sans savoir laquelle), la **règle de combinaison disjonctive (DCR)** peut être appliquée. Elle est définie par :

$$(m_1 \oplus m_2)(A) = \sum_{B \cup C = A} m_1(B) m_2(C) \quad \forall A \subseteq \Omega$$

Combinaison de l'information

Exemple avec la combinaison disjonctive

DCR	$m_2(\{b, c\}) = .3$	$m_2(\{a, b, c\}) = .7$
$m_1(\{a\}) = .4$	$a \cup \{b, c\} = \{a, b, c\}$ $.4 \times .3 = .12$	$a \cup \{a, b, c\} = \{a, b, c\}$ $.4 \times .7 = .28$
$m_1(\{a, b\}) = .6$	$\{a, b\} \cup \{b, c\} = \{a, b, c\}$ $.6 \times .3 = .18$	$\{a, b\} \cup \{a, b, c\} = \{a, b, c\}$ $.6 \times .7 = .42$

Le résultat $m_{\odot}$ de la combinaison disjonctive est :

► $m_{\odot}(\{a, b, c\}) = m_{\odot}(\Omega) = 1$

- ▶ Avec ces règles $\oplus$, $\odot$ et $\oslash$, l'ordre de combinaison des sources ne change pas les résultats
- ▶ $m_1 \odot m_2 = m_2 \odot m_1$ (**Commutativité**) (comme pour $\oplus$ et $\oslash$)
- ▶ $(m_1 \odot m_2) \odot m_3 = m_1 \odot (m_2 \odot m_3)$ (**Associativité**) (comme pour $\oplus$ et $\oslash$)
- ▶ $\oplus$, $\odot$ et $\oslash$ sont des règles non idempotentes
- ▶ L'élément neutre pour la combinaison de Dempster et la combinaison conjonctive est :

$$m(\Omega) = 1$$

- ▶ L'élément absorbant pour la combinaison disjonctive est :

$$m(\Omega) = 1$$

Combinaison de l'information

Remarques

- ▶ Selon un exemple de Zadeh de 1979, la règle de Dempster est considérée comme produisant des résultats contre-intuitifs.
- ▶ Exemple de Zadeh : $\Omega = \{a, b, c\}$, deux experts donnent :
 - ▶ Expert 1: $m_1(\{a\}) = 0.99$, $m_1(\{b\}) = 0.01$ et $m_1(\{c\}) = 0$
 - ▶ Expert 2: $m_2(\{a\}) = 0$, $m_2(\{b\}) = 0.01$ et $m_2(\{c\}) = 0.99$
- ▶ Alors $m_{1 \oplus 2}(\{b\}) = 1$, ce qui est considéré comme contre-intuitif, par certains auteurs, car les deux experts donnent b comme peu probable.
- ▶ Mais :
 - ▶ Les deux experts sont totalement fiables.
 - ▶ Expert 1 indique que c est absolument impossible.
 - ▶ Expert 2 indique que a est absolument impossible.
- ▶ Alors b est la seule possibilité. Nous sommes dans une situation où la seule solution possible donnée par les deux experts est b .
- ▶ La règle de Dempster donne donc des résultats valables si elle est utilisée correctement (c-à-d conformément à ses axiomes).

Combinaison de l'information

Règle de combinaison prudente (cautious rule, Denœux 2008)

- ▶ Les règles $\oplus$ et $\odot$ supposent des sources **distinctes**. Lorsque les sources **ne sont pas distinctes** (informations partiellement partagées, dépendantes), on souhaite une règle **idempotente** : $m \odot m = m$ (pas de double comptage d'une information).
- ▶ Toute fonction de masse **non dogmatique** ($m(\Omega) > 0$) admet une **décomposition canonique** (Smets) en combinaison conjonctive de **fonctions de masse simples généralisées** $A^{w(A)}$:

$$m = \bigodot_{A \subsetneq \Omega} A^{w(A)}, \quad A^{w(A)} : \begin{cases} m(A) = 1 - w(A) \\ m(\Omega) = w(A) \end{cases}$$

où la **fonction de poids** $w(A) \in]0, +\infty[$ se déduit des communalités q :

$$w(A) = \prod_{B \supseteq A} q(B)^{(-1)^{|B|-|A|+1}}.$$

Combinaison de l'information

Règle de combinaison prudente (cautious rule, Denœux 2008)

- La **règle prudente** $\odot$ combine en prenant le **minimum des poids** (principe de moindre engagement) :

$$m_1 \odot m_2 = \bigcap_{A \subseteq \Omega} A^{w_1(A) \wedge w_2(A)}, \quad w_1 \odot_2(A) = \min(w_1(A), w_2(A)).$$

- Propriétés : **commutative**, **associative**, **idempotente** (sa duale disjonctive est la *bold rule*).

Combinaison de l'information

Règle de combinaison prudente - exemple

Soient deux fonctions de masse non dogmatiques sur $\Omega = \{a, b\}$, issues de sources **non distinctes** :

$$\left\{ \begin{array}{lcl} m_1(\{a\}) & = & 0.2 \\ m_1(\{b\}) & = & 0.2 \\ m_1(\Omega) & = & 0.6 \end{array} \right. \quad \left\{ \begin{array}{lcl} m_2(\{a\}) & = & 0.6 \\ m_2(\{b\}) & = & 0.2 \\ m_2(\Omega) & = & 0.2 \end{array} \right.$$

1. **Poids** : $w(\emptyset) = \frac{q(\{a\})q(\{b\})}{q(\emptyset)q(\Omega)}$, $w(\{a\}) = \frac{q(\Omega)}{q(\{a\})}$ et $w(\{b\}) = \frac{q(\Omega)}{q(\{b\})}$

A	$w_1(A)$	$w_2(A)$	$w_1 \otimes w_2(A)$
$\emptyset$	$(.8 \times .8)/(1 \times .6) = 16/15 \approx 1.067$	$(.8 \times .4)/(1 \times .2) = 1.6$	16
$\{a\}$	$.6/.8 = .75$	$.2/.8 = .25$	0.25
$\{b\}$	$.6/.8 = .75$	$.2/.4 = .5$	0.5

Combinaison de l'information

Règle de combinaison prudente - exemple

2. Recombinaison : $m_{\odot} = \emptyset^{16/15} \odot \{a\}^{1/4} \odot \{b\}^{1/2}$:

Avec $q(A) = \prod_{B \not\supseteq A} w(B)$.

A	$\emptyset$	$\{a\}$	$\{b\}$	
$q_{\odot}(A)$	1	$16/15 \times 1/2 = 8/15$	$16/15 \times 1/4 = 4/15$	$16/15$
$m_{\odot}(A)$	$1/3 \approx .333$	$2/5 = .4$	$2/15 \approx .133$	

Représentation de l'information

Combinaison de l'information

Décision

Conclusion

- ▶ Afin de prendre une décision, plusieurs solutions sont offertes.
 - ▶ La première consiste à transformer la fonction m en une **probabilité** appelée **BetP** :

$$BetP(\omega_i) = \sum_{\{A \subseteq \Omega, \omega_i \in A\}} \frac{m(A)}{|A| (1 - m(\emptyset))}, \quad \forall \omega_i \in \Omega$$

- ▶ Exemple : $m(\{Peter, Paul\}) = 0.4$, $m(\{Mary\}) = 0.4$ et $m(\Omega) = 0.2$

$$BetP(Peter) = \frac{m(\{Peter, Paul\})}{2} + \frac{m(\Omega)}{3} = 0.267$$

$$BetP(Paul) = \frac{m(\{Peter, Paul\})}{2} + \frac{m(\Omega)}{3} = 0.267$$

$$BetP(Mary) = m(\{Mary\}) + \frac{m(\Omega)}{3} = 0.467$$

- ▶ Une autre transformation consiste à utiliser la **plausibilité des singletons** (la **fonction de contour** $pl(\omega_i) = Pl(\{\omega_i\})$), puis à la **normaliser** pour obtenir une probabilité **PIP** :

$$PIP(\omega_i) = \frac{Pl(\{\omega_i\})}{\sum_{\omega_j \in \Omega} Pl(\{\omega_j\})}, \quad \forall \omega_i \in \Omega$$

- ▶ Exemple : $m(\{Peter, Paul\}) = 0.4$, $m(\{Mary\}) = 0.4$ et $m(\Omega) = 0.2$

$$Pl(\{Peter\}) = m(\{Peter, Paul\}) + m(\Omega) = 0.6$$

$$Pl(\{Paul\}) = m(\{Peter, Paul\}) + m(\Omega) = 0.6$$

$$Pl(\{Mary\}) = m(\{Mary\}) + m(\Omega) = 0.6$$

La somme vaut $0.6 + 0.6 + 0.6 = 1.8$, d'où :

$$PIP(Peter) = PIP(Paul) = PIP(Mary) = \frac{0.6}{1.8} = 0.333$$

- ▶ On peut alors utiliser les outils classiques de la **théorie de la décision**
 - ▶ Soit $u(\omega, \omega_i)$: l'utilité de choisir ω alors que la solution est ω_i .

$$E_{BetP}(\omega) = \sum_{\omega_i \in \Omega} u(\omega, \omega_i) BetP(\omega_i).$$

- ▶ Décision retenue, dans le cas d'utilité 0 – 1 ($u(\omega_i, \omega_j) = 1$ si $\omega_i = \omega_j$, 0 sinon) :

$$\arg \max_{\omega_i \in \Omega} BetP(\omega_i)$$

- ▶ Il existe d'autres espérances :

- ▶ **Espérance supérieure :**

$$\overline{E}(\omega) = \sum_{A \subseteq \Omega} m(A) \max_{\omega_i \in A} u(\omega, \omega_i)$$

- ▶ **Espérance inférieure :**

$$\underline{E}(\omega) = \sum_{A \subseteq \Omega} m(A) \min_{\omega_i \in A} u(\omega, \omega_i)$$

- ▶ On peut alors prendre des décisions partielles de manière à obtenir un ensemble de solutions non dominées au regard d'une relation de préférence.

► **Dominance forte :**

$\omega_i \succeq_{sd} \omega_j$ si $\underline{E}(\omega_i) \geq \bar{E}(\omega_j)$
(ou $Bel(\{\omega_i\}) \geq Pl(\{\omega_j\})$ dans le cas d'utilité 0 – 1),

► **Dominance faible :**

$\omega_i \succeq_{wd} \omega_j$ si $\underline{E}(\omega_i) \geq \underline{E}(\omega_j)$ et $\bar{E}(\omega_i) \geq \bar{E}(\omega_j)$
(ou $Bel(\{\omega_i\}) \geq Bel(\{\omega_j\})$ et $Pl(\{\omega_i\}) \geq Pl(\{\omega_j\})$ dans le cas d'utilité 0 – 1),

► **Critère de Hurwicz :**

$\omega_i \succeq_{hu}^{\alpha} \omega_j$ si :

$$(1 - \alpha)\underline{E}(\omega_i) + \alpha\bar{E}(\omega_i) \geq (1 - \alpha)\underline{E}(\omega_j) + \alpha\bar{E}(\omega_j).$$

Décision

Exemple de prise de décision partielle avec les critères de dominance forte et de dominance faible

Exemple: avec $\Omega = \{a, b, c\}$, $m(\{a\}) = 0.3$, $m(\{b\}) = 0.4$ et $m(\Omega) = 0.3$

$$Bel = \begin{cases} \{a\} \mapsto 0.3 \\ \{b\} \mapsto 0.4 \\ \{c\} \mapsto 0.0 \end{cases} \quad Pl = \begin{cases} \{a\} \mapsto 0.6 \\ \{b\} \mapsto 0.7 \\ \{c\} \mapsto 0.3 \end{cases}$$

Dom. Forte	Non-dominé
$Bel(\{b\}) = .4 \not\geq Pl(\{a\}) = .6$ et $Bel(\{c\}) = 0 \not\geq Pl(\{a\}) = .6$	a
$Bel(\{a\}) = .3 \not\geq Pl(\{b\}) = .7$ et $Bel(\{c\}) = 0 \not\geq Pl(\{b\}) = .7$	b
$Bel(\{a\}) = .3 \geq Pl(\{c\}) = .3$ (donc $a \succeq_{sd} c$)	-

Conclusion selon la dominance forte : $d = \{a, b\}$.

Relation Dom. Faible	Dominé
$Bel(\{b\}) = 0.4 \geq Bel(\{a\}) = 0.3$ et $Pl(\{b\}) = 0.7 \geq Pl(\{a\}) = 0.6$ (donc $b \succeq_{wd} a$)	a
$Bel(\{b\}) = 0.4 \geq Bel(\{c\}) = 0.0$ et $Pl(\{b\}) = 0.7 \geq Pl(\{c\}) = 0.3$ (donc $b \succeq_{wd} c$)	c



Conclusion selon la dominance faible : $d = \{b\}$.



Représentation de l'information

Combinaison de l'information

Décision

Conclusion

- ▶ La théorie des fonctions de croyance est un **cadre mathématique flexible** permettant de représenter de façon relativement naturelle les **informations imparfaites**.
- ▶ Elle englobe la **théorie des probabilités** et la **théorie des ensembles** qui sont des cas particuliers.
- ▶ Cette théorie est employée dans de nombreuses applications. Au LGI2A : ML, problèmes de tournées de véhicules,...
- ▶ Bien d'autres outils dans cette théorie : sur les cadres de discernement (extension vide, raffinement/grossissement), estimation de la pertinence d'une source,...

Références bibliographiques

- [1] <https://bfasociety.org/#schools> |
https://www.youtube.com/@bfas_channel/playlists
- [2] T. Denœux' talks:
<https://www.hds.utc.fr/~tdenoeux/dokuwiki/en/talks>
- [3] T. Denœux' teaching:
<https://www.hds.utc.fr/~tdenoeux/dokuwiki/en/teaching>
- [4] G. Shafer. *A mathematical theory of evidence*. Princeton, N.J.: Princeton University Press, 1976
- [5] P. Smets. "The application of the matrix calculus to belief functions". In: *International Journal of Approximate Reasoning* 31.1-2 (2002), pp. 1-30
- [6] P. Smets. "Analyzing the Combination of Conflicting Belief Functions". In: *Information Fusion* 8.4 (2007), pp. 387-412
- [7] T. Denœux. "Decision-Making with Belief Functions: a Review". In: *International Journal of Approximate Reasoning* 109 (2019), pp. 7-110