Le modèle des croyances transférables

Nouvelles règles pour la combinaison d'informations non distinctes

Thierry Denœux[1]

[1]Université de Technologie de Compiègne
HEUDIASYC (UMR CNRS 6599)

Logique Floue et Applications, 2006

Philippe Smets (1938-2005)

Plan de la présentation

Plan de la présentation

Qu'est-ce que le MCT ?

Situation par rapport aux autres théories de l'incertain

- Un modèle de représentation des connaissances partielles (incertaines, imprécises), développé par Philippe Smets.
- Modèles « concurrents » :
  - Théorie Bayésienne des probabilités
  - Théorie des possibilités
  - Théorie des probabilités imprécises
  - ...
- Spécificité : utilisation du formalisme mathématique de la théorie des fonctions de croyance (FCs) (Dempster 1967, Shafer 1976).

Spécificités du MCT

1. Un modèle subjectiviste et non probabiliste

- Interprétation d'une fonction de croyance comme représentant l'opinion d'un agent rationnel relativement à une certaine question, sans référence à une probabilité ou une famille de probabilités sous-jacentes ;
- Autres interprétations des fonctions de croyance :
  - Modèle de Dempster, théorie des Hints (Kohlas et Monney) ;
  - Probabilités imprécises ;
  - Théorie des ensembles aléatoires ;
- Ces interprétations sont équivalentes pour ce qui concerne la partie statique du modèle, mais elles diffèrent dans la partie dynamique (prise en compte de nouvelles informations).

Spécificités du MCT

2. Notion de probabilité pignistique

- Hypothèse : raisonnement dans l'incertain et décision sont deux tâches cognitives de natures différentes, qui peuvent être traitées à deux niveaux distincts :
  - niveau crédal : représentation et manipulation des croyances à l'aide de fonctions de croyance ;
  - niveau pignistique : prise de décision selon le principe de maximisation de l'utilité espérée, relativement à une mesure de probabilité.
- Le passage du niveau crédal au niveau pignistique est effectué au moment de la prise de décision à l'aide la transformation pignistique.
- Le MCT ne se distingue pas du modèle Bayésien dans la phase décisionnelle. Il échappe donc à certains arguments sur l'incohérence de décisions ne reposant pas sur une probabilité (paris hollandais).

Spécificités du MCT

3. Distinction entre monde ouvert/clos

- Hypothèse du monde clos : toutes les réponses possibles à la question considérée ont été recensées dans le cadre de discernement.
- Hypothèse du monde ouvert : il peut exister certaines possibilités non recensées initialement.
- Sous l'hypothèse du monde ouvert, une masse peut être allouée à l'ensemble vide. Interprétation : croyance dans le fait que la réponse à la question se trouve hors du cadre de discernement.
- Plus généralement, la normalisation n'est généralement pas utilisée dans le MCT, la masse $m(\emptyset)$ étant porteuse d'une information sur le conflit entre les sources d'information.

La combinaison des informations dans le MCT

Insuffisance du MCT par rapport à d'autres modèles

- Problème fondamental dans de nombreuses applications (fusion multi-capteurs, combinaison d'avis d'experts) ;
- Théorie des FCs appliquée avec succès en fusion d'informations, mais...
- Cette théorie peut sembler « pauvre » de ce point de vue par rapport à d'autres approches comme la théorie des possibilités :
  - Deux opérateurs principaux : règle de Dempster (normalisée ou non) et somme disjonctive ;
  - La règle de Dempster peut avoir un comportement indésirable en cas de conflit important entre les sources ;
  - Ces opérateurs supposent que les sources combinées sont distinctes.

La combinaison des informations dans le MCT
Approches du problème

- Gestion du conflit :
  - Grand nombre de travaux.
  - Proposition de règles de combinaison plus robustes que la règle de Dempster en présence de conflit.
- Caractère distinct des sources combinées :
  - Peu de travaux ;
  - Peu de solutions actuellement en présence de sources non distinctes, lorsque l'interaction entre les sources ne peut être modélisée.
  - De nouvelles solutions à ce problème seront proposées dans cet exposé.

Plan de la présentation

Fonction de masse
Définition

Soit Ω un ensemble fini de réponses à une certaine question Q (cadre de discernement).

Définition (Fonction de masse)

Une fonction de masse de croyance sur Ω *est une application* $m : 2^\Omega \to [0,1]$ *t.q.*

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

Tout $A \subseteq \Omega$ *t.q.* $m(A) > 0$ *est appelé ensemble focal (EF) de* m*.*

Fonction de masse
Interprétation

- La fonction de masse m représente
  - l'état de connaissance d'un agent rationel Ag à un certain instant t, relativement à Q ;
  - par extension, un élément d'évidence qui induit un tel état de connaissance.
- $m(A)$: part de croyance allouée à A (et à aucun sous-ensemble strict).
- $m(\Omega)$: degré d'ignorance.
- $m(\emptyset)$: degré de conflit. Sous l'hypothèse du monde ouvert, degré de croyance dans l'hypothèse selon laquelle la réponse à la question Q se trouve en dehors de Ω.

Représentation matricielle d'une fonction de masse

- Etant donné un ordre sur 2^Ω, une fonction de masse m peut être représentée par un vecteur $\mathbf{m} = (m(A_1), \ldots, m(A_n))^t \in \mathbb{R}^n$ avec $n = 2^{|\Omega|}$.
- Ordre « naturel » : ordre binaire.

	c	b	a	
$\emptyset$	0	0	0	0
$\{a\}$	0	0	1	1
$\{b\}$	0	1	0	2
$\{a, b\}$	0	1	1	3
$\{c\}$	1	0	0	4
$\{a, c\}$	1	0	1	5
$\{b, c\}$	1	1	1	6
$\{a, b, c\}$	1	1	1	7

Représentation matricielle d'une fonction de masse

- Etant donné un ordre sur 2^Ω, une fonction de masse m peut être représentée par un vecteur $\mathbf{m} = (m(A_1), \ldots, m(A_n))^t \in \mathbb{R}^n$ avec $n = 2^{|\Omega|}$.
- Ordre « naturel » : ordre binaire.

	c	b	a	
$\emptyset$	0	0	0	0
$\{a\}$	0	0	1	1
$\{b\}$	0	1	0	2
$\{a,b\}$	0	1	1	3
$\{c\}$	1	0	0	4
$\{a,c\}$	1	0	1	5
$\{b,c\}$	1	1	1	6
$\{a,b,c\}$	1	1	1	7

Fonctions associées

Fonctions de croyance et d'implicabilité

Définition (Fonction de croyance)

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B), \quad \forall A \subseteq \Omega$$

Interprétation de $bel(A)$: degré de croyance en A.

Définition (Fonction d'implicabilité)

$$b(A) = bel(A) + m(\emptyset), \quad \forall A \subseteq \Omega$$

Fonctions associées

Fonctions de croyance et d'implicabilité

Définition (Fonction de croyance)

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B), \quad \forall A \subseteq \Omega$$

Interprétation de $bel(A)$: degré de croyance en A.

Définition (Fonction d'implicabilité)

$$b(A) = bel(A) + m(\emptyset), \quad \forall A \subseteq \Omega$$

Fonctions associées

Plausibilité et communalité

Définition (Fonction de plausibilité)

$$\textit{Plausibilité} : pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \quad \forall A \subseteq \Omega$$

Propriété : $pl(A) = bel(\Omega) - bel(\overline{A})$.

Définition (Fonction de communalité)

$$\textit{Communalité} : q(A) = \sum_{B \supseteq A} m(B), \quad \forall A \subseteq \Omega$$

Fonctions associées

Plausibilité et communalité

Définition (Fonction de plausibilité)

$$\textit{Plausibilité : } pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \quad \forall A \subseteq \Omega$$

Propriété : $pl(A) = bel(\Omega) - bel(\overline{A})$.

Définition (Fonction de communalité)

$$\textit{Communalité : } q(A) = \sum_{B \supseteq A} m(B), \quad \forall A \subseteq \Omega$$

Equivalence des représentations

- Les fonctions *bel*, *b*, *pl*, *q*, *m* sont en correspondance biunivoque $\longrightarrow$ représentations équivalentes.
- Vectoriellement, on passe d'une représentation à une autre par des opérateurs linéaires.
- Exemples

$$\mathbf{q} = \mathbf{M}_{m\rightarrow q}\mathbf{m} \Leftrightarrow \mathbf{m} = \mathbf{M}_{q\rightarrow m}\mathbf{q}$$

avec $\mathbf{M}_{q\rightarrow m} = \mathbf{M}_{m\rightarrow q}^{-1}$.

Révision des croyances

Impact d'un élément d'évidence sur une fonction de masse m

- Soit m une fonction de masse représentant un état de connaissance sur Q à t.
- A $t' > t$ l'agent prend connaissance d'un élément d'évidence Ev.
- L'agrégation de m et de Ev produit une nouvelle fonction de masse m'.
- On peut toujours écrire $\mathbf{m}' = \mathbf{M} \cdot \mathbf{m}$, où $\mathbf{M}$ est une matrice stochastique, a priori fonction de m et de Ev.
- Cas particuliers : spécialisation, généralisation.

Révision conjonctive
Matrice de spécialisation

- Lorsque la révision a pour effet de transformer une fonction de masse m en une fonction de masse m' plus spécifique (riche, informative, précise), la révision est dite conjonctive.
- Processus : transfert de masses $m(B)$ à chaque sous-ensemble $A \subseteq B$, en proportion $S(A, B)$. Formellement $\mathbf{m}' = \mathbf{S} \cdot \mathbf{m}$ où $\mathbf{S}$ est une matrice de spécialisation.

Définition (Matrice de spécialisation)

Une matrice de spécialisation $\mathbf{S} = [\mathbf{S}(A, B)]$, $A, B \subseteq \Omega$ *est une matrice stochastique vérifiant* $\mathbf{S}(A, B) = 0 \; \forall A \not\subseteq B$.

- Si $\mathbf{m}' = \mathbf{S} \cdot \mathbf{m}$, on dit que m' est une spécialisation de m et on note $m' \sqsubseteq_s m$ (ordre partiel).

Révision conjonctive
Matrice de spécialisation

- Lorsque la révision a pour effet de transformer une fonction de masse m en une fonction de masse m' plus spécifique (riche, informative, précise), la révision est dite conjonctive.
- Processus : transfert de masses $m(B)$ à chaque sous-ensemble $A \subseteq B$, en proportion $S(A, B)$. Formellement $\mathbf{m}' = \mathbf{S} \cdot \mathbf{m}$ où $\mathbf{S}$ est une matrice de spécialisation.

Définition (Matrice de spécialisation)

Une matrice de spécialisation $\mathbf{S} = [\mathbf{S}(A, B)]$, $A, B \subseteq \Omega$ *est une matrice stochastique vérifiant* $\mathbf{S}(A, B) = 0 \; \forall A \not\subseteq B$.

- Si $\mathbf{m}' = \mathbf{S} \cdot \mathbf{m}$, on dit que m' est une spécialisation de m et on note $m' \sqsubseteq_s m$ (ordre partiel).

La règle de conditionnement de Dempster

- Soit m ma fonction de masse représentant notre état de connaissance à t.
- A $t' > t$, on apprend que la réponse à la question se trouve dans $C \subseteq \Omega$.
- Un nouvelle fonction de masse m' est obtenue en transférant toute masse $m(B)$ à $B \cap C$.
- La transformation de m en m' peut être représentée par une matrice de spécialisation $\mathbf{S}_C$ définie par

$$\mathbf{S}_C(A, B) = \begin{cases} 1 & \text{si } B \cap C = A, \\ 0 & \text{sinon.} \end{cases}$$

- $\mathbf{S}_C$ correspond au conditionnement de m par C. On note $\mathbf{m}[C] = \mathbf{S}_C \cdot \mathbf{m}$.

Interprétation de $pl(A)$ et $q(A)$

- Plausibilité :
  - $pl(A) = bel[A](A) = \max_B bel[B](A)$
  - degré maximal de croyance susceptible d'être alloué à A après conditionnement.
- Communalité :
  - $q(A) = m[A](A)$
  - degré d'ignorance après conditionnement par A.

La règle de combinaison de Dempster
Définition

Définition

Soit m une fonction de masse. La matrice de spécialisation $\mathbf{S}_m$ *définie par*

$$\mathbf{S}_m(A, B) = m[B](A), \quad \forall A, B \subseteq \Omega$$

est appelée *matrice de spécialisation dempstérienne* *associée à m.*

Définition (Règle de combinaison de Dempster)

$$m_1 \textcircled{\cap} m_2 = \mathbf{S}_{m_2} \cdot \mathbf{m}_1$$

La règle de combinaison de Dempster
Propriétés

- Commutativité, associativité ;
- Elément neutre : m_Ω ;
- Généralisation du conditionnement :

$$m \textcircled{\cap} m_C = m[C],$$

où m_C est la fonction de masse catégorique vérifiant $m_C(C) = 1$.

- Les valeurs propres de la matrice de spécialisation dempstérienne $\mathbf{S}_m$ sont les communalités $q(A)$, $A \subseteq \Omega$. Conséquence :

$$q_{1 \textcircled{\cap} 2} = q_1 \cdot q_2.$$

Justification de la règle de Dempster

Klawonn et Smets (1992)

Soit $m_1 * m_2$ la fonction de masse résultant de l'impact d'un élément d'évidence m_2 sur une fonction de masse m_1. Si :

1. L'impact de m_2 sur m_1 est une spécialisation (combinaison conjonctive) ;
2. Cet impact est un conditionnement lorsque m_2 est catégorique ($m_2(C) = 1$ pour un $C \subseteq \Omega$).
3. m_1 et m_2 sont distinctes ($\mathbf{S}_{m_2}$ ne dépend pas de m_1),
4. L'ordre dans lequel sont pris en compte différents éléments d'évidence est indifférent,

alors $\mathbf{S}_{m_2}$ doit être dempstérienne, et par conséquent $* = \bigcirc\!\!\!\!\cap$.

Révision disjonctive
Matrice de généralisation

- Lorsque la révision a pour effet de transformer une fonction de masse m en une fonction de masse m' moins spécifique, la révision est dite disjonctive.
- Processus : transfert de masses $m(B)$ à chaque surensemble $A \supseteq B$, en proportion $G(A, B)$. Formellement $\mathbf{m}' = \mathbf{G} \cdot \mathbf{m}$ où $\mathbf{G}$ est une matrice de généralisation.

Définition (Matrice de généralisation)

Une matrice de généralisation $\mathbf{G} = [\mathbf{G}(A, B)]$, $A, B \subseteq \Omega$ *est une matrice stochastique vérifiant* $\mathbf{G}(A, B) = 0 \; \forall A \not\supseteq B$.

- Si $\mathbf{m}' = \mathbf{G} \cdot \mathbf{m}$, on dit que m' est une généralisation de $m \Leftrightarrow m$ et une spécialisation de m'.

Révision disjonctive

Matrice de généralisation

- Lorsque la révision a pour effet de transformer une fonction de masse m en une fonction de masse m' moins spécifique, la révision est dite disjonctive.
- Processus : transfert de masses $m(B)$ à chaque surensemble $A \supseteq B$, en proportion $G(A, B)$. Formellement $\mathbf{m}' = \mathbf{G} \cdot \mathbf{m}$ où $\mathbf{G}$ est une matrice de généralisation.

Définition (Matrice de généralisation)

Une matrice de généralisation $\mathbf{G} = [\mathbf{G}(A, B)]$, $A, B \subseteq \Omega$ *est une matrice stochastique vérifiant* $\mathbf{G}(A, B) = 0 \; \forall A \not\supseteq B$.

- Si $\mathbf{m}' = \mathbf{G} \cdot \mathbf{m}$, on dit que m' est une généralisation de $m \Leftrightarrow m$ et une spécialisation de m'.

La règle de combinaison disjonctive
Définition

Définition

Soit m une fonction de masse. La matrice de généralisation $\mathbf{G}_m$ *définie par*

$$\mathbf{G}_m(A,B) = \sum_{C : B \cup C = A} m(C), \quad \forall A, B \subseteq \Omega$$

est appelée matrice de généralisation dempstérienne associée à m.

Définition (Règle de combinaison disjonctive)

$$m_1 \textcircled{\cup} m_2 = \mathbf{G}_{m_2} \cdot \mathbf{m}_1$$

La règle de combinaison disjonctive
Propriétés

- Commutativité, associativité ;
- Elément neutre $m_\emptyset$.
- $b_{1 \cup\!\!\!\!\bigcirc 2} = b_1 \cdot b_2$.

Négation et lois de De Morgan

- Complémentaire (négation) de m :

$$\overline{m}(A) = m(\overline{A}), \quad \forall A \subseteq \Omega.$$

- Propriété :

$$\overline{b}(A) = q(\overline{A}), \quad \forall A \subseteq \Omega.$$

- Lois de De Morgan pour $\bigcirc\!\!\!\!\cap$ et $\bigcirc\!\!\!\!\cup$:

$$\begin{aligned}
\overline{m_1 \bigcirc\!\!\!\!\cup\, m_2} &= \overline{m_1} \bigcirc\!\!\!\!\cap\, \overline{m_2}, \\
\overline{m_1 \bigcirc\!\!\!\!\cap\, m_2} &= \overline{m_1} \bigcirc\!\!\!\!\cup\, \overline{m_2},
\end{aligned}$$

Plan de la présentation

1. Le modèle des croyances transférables
   - Généralités
   - Concepts de base
   - Décomposition canonique conjonctive
2. Nouvelles règles de combinaison
   - Règle conjonctive prudente
   - Règle disjonctive hardie
   - Règles prudentes et hardies généralisées

Fonction de masse simple

Définition et notation

Définition (Fonction de masse simple)

Une fonction de masse est dite simple si elle est de la forme

$$\begin{aligned} m(A) &= 1 - w \\ m(\Omega) &= w, \end{aligned}$$

avec $w \in [0,1]$ *et* $A \subseteq \Omega$. *Notation :* $m = A^w$.

- Propriété : $A^{w_1} \cap\!\!\!\!\bigcirc A^{w_2} = A^{w_1 w_2}$.
- Cas particuliers :
  - Fonction de masse vide : A^1 avec A quelconque.
  - Fonction de masse catégorique A^0.
- Peut-on décomposer une fonction de masse quelconque en produit par $\cap\!\!\!\!\bigcirc$ de fonctions de masse simples ?

Fonction de masse simple
Définition et notation

Définition (Fonction de masse simple)

Une fonction de masse est dite simple si elle est de la forme

$$\begin{aligned} m(A) &= 1 - w \\ m(\Omega) &= w, \end{aligned}$$

avec $w \in [0,1]$ *et* $A \subseteq \Omega$. *Notation :* $m = A^w$.

- Propriété : $A^{w_1} \cap\!\!\!\!\bigcirc A^{w_2} = A^{w_1 w_2}$.
- Cas particuliers :
  - Fonction de masse vide : A^1 avec A quelconque.
  - Fonction de masse catégorique A^0.
- Peut-on décomposer une fonction de masse quelconque en produit par $\cap\!\!\!\!\bigcirc$ de fonctions de masse simples ?

Fonction de masse séparable

Définition (Shafer, 1976)

Une fonction de masse est dite séparable si elle admet une décomposition de la forme :

$$m = A_1^{w_1} \bigcirc\!\!\!\!\cap \ldots \bigcirc\!\!\!\!\cap A_n^{w_n}$$

avec $w_i \in [0, 1]$, $i = 1, \ldots, n$.

Toutes les fonctions de masse ne sont pas séparables !

Décomposition canonique d'une fonction de masse séparable

Théorème (Shafer, 1976)

Si m est non dogmatique ($m(\Omega) > 0$) et si les A_i sont tous distincts, la décomposition est unique (décomposition canonique). ON peut écrite :

$$m = \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w(A)},$$

avec $w(A) \in]0, 1]$ pour tout $A \subset \Omega$.

- Toute fonction de masse séparable non dogmatique peut donc être représentée de manière unique par une fonction $w : 2^{\Omega} \rightarrow]0, 1]$.
- Généralisation à une fonction de masse non séparable ?

Généralisation

Fonction de masse simple généralisée

Définition

On appelle fonction de masse simple généralisée toute fonction $\mu : 2^\Omega \rightarrow \mathbb{R}$ *t.q.*

$$\begin{aligned} \mu(A) &= 1 - w, \\ \mu(\Omega) &= w, \\ \mu(B) &= 0 \quad \forall B \in 2^\Omega \setminus \{A, \Omega\}, \end{aligned}$$

pour un $A \neq \Omega$ *et un* $w \in [0, +\infty)$*. Notation :* $\mu = A^w$*.*

Généralisation

Fonction de masse simple généralisée : interprétation

- Si $w \leq 1$, μ est une fonction de masse simple.
- Si $w > 1$, μ n'est pas une fonction de masse $\rightarrow$ fonction de masse simple inverse.
- Interprétation : modélise un état de connaissance caractérisé par une défiance à l'encontre de l'hypothèse A. Il faut acquérir un élément d'évidence en faveur de A pour atteindre un état neutre :

$$A^w \bigcirc\!\!\!\!\cap\, A^{1/w} = A^1.$$

Généralisation

Décomposition canonique d'une fonction de masse non dogmatique

Théorème (Smets, 1995)

Toute fonction de masse non dogmatique peut être décomposée de manière unique en produit par $\textcircled{\cap}$ *de fonctions de masse simples généralisées :*

$$m = \textcircled{\cap}_{A \subset \Omega} A^{w(A)},$$

avec $w(A) \in]0, +\infty[$ *pour tout* $A \subset \Omega$.

- m est séparable ssi $w(A) \leq 1$ pour tout A.
- La fonction $w : 2^{\Omega} \rightarrow]0, +\infty[$ est une nouvelle représentation de m (fonction de pondération conjonctive).

Généralisation

Décomposition canonique d'une fonction de masse non dogmatique

Théorème (Smets, 1995)

Toute fonction de masse non dogmatique peut être décomposée de manière unique en produit par $\textcircled{\cap}$ *de fonctions de masse simples généralisées :*

$$m = \textcircled{\cap}_{A \subset \Omega} A^{w(A)},$$

avec $w(A) \in]0, +\infty[$ *pour tout* $A \subset \Omega$*.*

- m est séparable ssi $w(A) \leq 1$ pour tout A.
- La fonction $w : 2^{\Omega} \to]0, +\infty[$ est une nouvelle représentation de m (fonction de pondération conjonctive).

Fonction de pondération conjonctive
Calcul

- Calcul de w à partir de q :

$$\ln w(A) = -\sum_{B \supseteq A} (-1)^{|B|-|A|} \ln q(B), \quad \forall A \subset \Omega.$$

- Similarité avec

$$m(A) = \sum_{B \supseteq A} (-1)^{|B|-|A|} q(B), \quad \forall A \subseteq \Omega$$

- Toute procédure de transformation de q en m peut donc être utilisée pour transformer $\ln q$ en $\ln w$. En particulier :

$$\ln \mathbf{w} = \mathbf{M}_{q \to m} \ln \mathbf{q}.$$

Exemple

A	$m(A)$	$q(A)$	$w(A)$
$\emptyset$	0	1	1
$\{a\}$	0	0.5	1
$\{b\}$	0	1	7/4
$\{a,b\}$	0.3	0.5	2/5
$\{c\}$	0	0.7	1
$\{a,c\}$	0	0.2	1
$\{b,c\}$	0.5	0.7	2/7
Ω	0.2	0.2	

On a donc :

$$m = \{a,b\}^{2/5} \textcircled{\cap} \{b,c\}^{2/7} \textcircled{\cap} \{b\}^{7/4}$$

Propriété de w

Propriété

On a

$$\begin{aligned} m_1 \bigcirc\!\!\!\!\cap\, m_2 &= \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A)} \right) \bigcirc\!\!\!\!\cap \left(\bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_2(A)} \right) \\ &= \bigcirc\!\!\!\!\cap_{A \subset \Omega} A^{w_1(A) w_2(A)}. \end{aligned}$$

d'où

$$w_{1 \bigcirc\!\!\!\!\cap 2} = w_1 \cdot w_2.$$

Nouvelles règles de combinaison
Motivations

- La règle de Dempster $\textcircled{\cap}$ et la règle disjonctive duale $\textcircled{\cup}$ supposent que les fonctions de croyance combinées sont issues de sources distinctes.
- Définition formelle : $\mathbf{m}_{12} = \mathbf{M}(m_2) \cdot \mathbf{m}_1$ pour une matrice de spécialisation ou de généralisation $\mathbf{M}(m_2)$.
- Signification intuitive : la même information ne doit pas être comptée deux fois.
- Exemple : des opinions d'experts différents se basant sur une expérience commune ne sont pas des sources distinctes.

Nouvelles règles de combinaison
Nécessité de l'idempotence

- Une règle utilisable pour combiner des sources non distinctes doit être idempotente : $m * m = m$.
- Seul opérateur idempotent en théorie des fonctions de croyance : moyenne, mais pas associatif.
- Introduction et justification de 2 nouvelles règles commutatives, associatives et idempotentes :
  - règle conjonctive prudente $\owedge$;
  - règle disjonctive hardie $\ovee$.

	conjonctif	disjonctif
distinct	$\cap\!\!\!\bigcirc$	$\cup\!\!\!\bigcirc$
non distinct	$\owedge$	$\ovee$

Plan de la présentation

Principe d'engagement minimal

Définition (Principe d'engagement minimal)

Lorque plusieurs fonctions de croyance sont compatibles avec un ensemble de contraintes, la moins informative doit être choisie.

- Nécessité de définir une ordre partiel sur les fonctions de croyance : « m_1 plus informative (riche) que m_2 ».
- Une telle relation doit logiquement correspondre à l'inclusion lorsque m_1 et m_2 sont catégoriques.

Relations d'informativité relative

- $m_1 \sqsubseteq_s m_2$ ssi $\mathbf{m}_1 = \mathbf{S} \cdot \mathbf{m}_2$, $\mathbf{S}$ matrice de spécialisation.
- Spécialisation dempsterienne : $m_1 \sqsubseteq_d m_2$ ssi $\mathbf{m}_1 = \mathbf{S}_m \cdot \mathbf{m}_2$, $\mathbf{S}_m$ matrice de spécialisation dempstérienne (on a alors $m_1 = m \bigcirc\!\!\!\!\cap\, m_2$).
- $m_1 \sqsubseteq_{pl} m_2$ ssi $pl_1(A) \leq pl_2(A), \quad \forall A \subseteq \Omega$
- $m_1 \sqsubseteq_q m_2$ ssi $q_1(A) \leq q_2(A), \quad \forall A \subseteq \Omega$
- Propriétés :
  - $m_1 \sqsubseteq_d m_2 \Rightarrow m_1 \sqsubseteq_s m_2 \Rightarrow \begin{cases} m_1 \sqsubseteq_{pl} m_2 \\ m_1 \sqsubseteq_q m_2, \end{cases}$
  - La fonction de masse vide m_Ω est l'unique plus grand élément de la relation $\sqsubseteq_x$ pour $x \in \{pl, q, s, d\}$

Une nouvelle relation d'informativité

- Posons : $m_1 \sqsubseteq_w m_2$ ssi $w_1(A) \leq w_2(A)$, pour tout $A \subset \Omega$ (m_1 et m_2 non dogmatiques).
- Signification : $m_1 = m \bigcirc\!\!\!\!\cap\, m_2$ pour une fonction séparable m.
- Propriétés :
  - $m_1 \sqsubseteq_w m_2 \Rightarrow m_1 \sqsubseteq_d m_2 \Rightarrow m_1 \sqsubseteq_s m_2 \Rightarrow \begin{cases} m_1 \sqsubseteq_{pl} m_2 \\ m_1 \sqsubseteq_q m_2, \end{cases}$
  - m_Ω est le seul élément maximal de $\sqsubseteq_w$:

$$m_\Omega \sqsubseteq_w m \Rightarrow m = m_\Omega.$$

La règle conjonctive prudente
Principe

- Un agent reçoit deux fonctions de masse m_1 et m_2 issues de sources fiables.
- Soit m_{12} la fct de masse représentant son état de connaissance après réception de m_1 et m_2.
- m_{12} doit être plus informative que m_1 et m_2.
  Formellement : $m_{12} \in \mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$, pour un $x \in \{w, d, s, pl, q\}$, avec $\mathcal{S}_x(m)$= ens. des fcts de masse plus riches que m au sens de $\sqsubseteq_x$.
- Principe d'engagement minimal : on choisit dans $\mathcal{S}_x(m_1) \cap \mathcal{S}_x(m_2)$ la fonction de masse la moins riche au sens de $\sqsubseteq_x$ (si elle existe).

La règle conjonctive prudente
Définition

Théorème

Soient m_1 and m_2 deux fcts de masse non dogmatiques. L'élément le moins riche de $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ au sens de $\sqsubseteq_w$ existe et est unique. Il est défini par

$$w_{1 \textcircled{\wedge} 2}(A) = w_1(A) \wedge w_2(A), \quad \forall A \subset \Omega.$$

Définition (Règle conjonctive prudente)

$$m_1 \textcircled{\wedge} m_2 = \textcircled{\cap}_{A \subset \Omega} A^{w_1(A) \wedge w_2(A)}.$$

La règle conjonctive prudente

Définition

Théorème

Soient m_1 and m_2 deux fcts de masse non dogmatiques. L'élément le moins riche de $\mathcal{S}_w(m_1) \cap \mathcal{S}_w(m_2)$ au sens de $\sqsubseteq_w$ existe et est unique. Il est défini par

$$w_{1 \textcircled{\wedge} 2}(A) = w_1(A) \wedge w_2(A), \quad \forall A \subset \Omega.$$

Définition (Règle conjonctive prudente)

$$m_1 \textcircled{\wedge} m_2 = \textcircled{\cap}_{A \subset \Omega} A^{w_1(A) \wedge w_2(A)}.$$

Exemple

A	$m_1(A)$	$w_1(A)$	$m_2(A)$	$w_2(A)$	$w_{1 \wedge 2}(A)$	$m_{1 \wedge 2}(A)$
$\emptyset$	0	1	0	1	1	0
$\{a\}$	0	1	0	1	1	0
$\{b\}$	0	7/4	0.3	0.7	0.7	0.6
$\{a, b\}$	0.3	2/5	0	1	2/5	0.12
$\{c\}$	0	1	0	1	1	0
$\{a, c\}$	0	1	0	1	1	0
$\{b, c\}$	0.5	2/7	0.4	3/7	2/7	0.2
Ω	0.2		0.3			0.08

Propriétés

Commutativité : $\forall m_1, m_2,\ m_1 \circledwedge m_2 = m_2 \circledwedge m_1$

Associativité : $\forall m_1, m_2, m_3,$

$$m_1 \circledwedge (m_2 \circledwedge m_3) = (m_1 \circledwedge m_2) \circledwedge m_3$$

Idempotence : $\forall m,\ m \circledwedge m = m$

Propriétés (suite)

Distributivité de $\bigcirc\!\!\!\!\cap$ par rapport à $\bigcirc\!\!\!\!\wedge$:

$$(m_1 \bigcirc\!\!\!\!\cap\, m_2) \bigcirc\!\!\!\!\wedge\, (m_1 \bigcirc\!\!\!\!\cap\, m_3) = m_1 \bigcirc\!\!\!\!\cap\, (m_2 \bigcirc\!\!\!\!\wedge\, m_3), \forall m_1, m_2, m_3.$$

L'élément d'évidence m_1 n'est compté qu'une fois !

Absence d'élément neutre, mais $m_\Omega \bigcirc\!\!\!\!\wedge\, m = m$ ssi m est séparable.

Relation avec $\bigcirc\!\!\!\!\cap$: Si m_1 et m_2 sont séparables, alors

$$m_1 \bigcirc\!\!\!\!\cap\, m_2 \sqsubseteq_w m_1 \bigcirc\!\!\!\!\wedge\, m_2$$

Plan de la présentation

Recherche d'une règle disjonctive

- L'agent reçoit deux masses m_1 et m_2 en provenance de deux sources dont l'une au moins est fiable.
- La fonction de masse résultante doit être moins riche que m_1 et m_2. Formellement, $m_{12} \in \mathcal{G}_x(m_1) \cap \mathcal{G}_x(m_2)$, pour un $x \in \{w, d, s, pl, q\}$, avec $\mathcal{G}_x(m)$= ens. des fcts de masse moins riches que m au sens de $\sqsubseteq_x$.
- Principe d'information maximum : on choisit dans $\mathcal{G}_x(m_1) \cap \mathcal{G}_x(m_2)$ la fonction de masse la plus riche au sens de $\sqsubseteq_x$ (si elle existe).

Recherche d'une règle disjonctive

Problème du choix d'une relation d'informativité

- Avec $x = w$, cette approche conduit à prendre la fct de masse m_{12} définie par $w_{12} = w_1 \vee w_2$.
- Problème : $w_1 \vee w_2$ ne correspond pas toujours à une fonction de croyance.
- Il faut trouver la bonne relation d'ordre...

Décomposition canonique disjonctive

Fonction de masse simple complémentaire

Définition

On appelle fonction de masse simple complémentaire toute fonction $\nu : 2^\Omega \rightarrow \mathbb{R}$ *t.q.*

$$\begin{aligned} \nu(A) &= 1 - v, \\ \nu(\emptyset) &= v, \\ \nu(B) &= 0 \quad \forall B \in 2^\Omega \setminus \{A, \emptyset\}, \end{aligned}$$

pour un $A \neq \emptyset$ *et un* $v \in [0, +\infty)$. *On note* $\nu = A_v$.

Décomposition canonique disjonctive

Fonction de pondération disjonctive

Théorème

Toute fonction de masse m sous-normale (t.q. $m(\emptyset) > 0$) peut être décomposée de manière unique comme somme disjonctive de fcts de masse simples complémentaires $A_{v(A)}$:

$$m = \bigcirc\!\!\!\!\cup_{A \neq \emptyset} A_{v(A)}.$$

Définition

La fonction $v : 2^\Omega \setminus \{\emptyset\} \to (0, +\infty)$ est appelée fonction de pondération disjonctive.

Fonction de pondération disjonctive

Propriétés

- Dualité avec w : $v(A) = \overline{w}(\overline{A}), \forall A \neq \emptyset$ (idem $b(A) = \overline{q}(\overline{A})$).
- Calcul à partir de b :

$$\ln v(A) = -\sum_{B \subseteq A} (-1)^{|A|-|B|} \ln b(B).$$

Similarité avec

$$m(A) = \sum_{B \subseteq A} (-1)^{|A|-|B|} b(B), \quad \forall A \subseteq \Omega.$$

- Combinaison disjonctive :

$$v_{1 \bigcirc\!\!\!\!\cup\, 2} = v_1 \cdot v_2.$$

Relation d'ordre induite par v

- Soient deux fcts de masse sous-normales m_1 et m_2.
- Si $v_1(A) \geq v_2(A)$, pour tout $A \neq \emptyset$, alors on a $m_2 = m_1 \bigcirc\!\!\!\!\!\cup\, m$ pour une fonction de masse m t.q. $v = v_2/v_1 \rightarrow m_1$ est une spécialisation de m_2.
- Définition :

$$m_1 \sqsubseteq_v m_2 \text{ ssi } v_1(A) \geq v_2(A), \quad \forall A \neq \emptyset.$$

- Propriété :

$$m_1 \sqsubseteq_v m_2 \Rightarrow m_1 \sqsubseteq_s m_2$$

(La réciproque est fausse).

La règle disjonctive hardie

Théorème

Soient m_1 et m_2 deux fonctions de masse sous-normales. La fct de masse la plus riche dans $\mathcal{G}_v(m_1) \cap \mathcal{G}_v(m_2)$ existe et est unique. Elle est définie par :

$$v_{1 \circledvee 2}(A) = v_1(A) \wedge v_2(A), \quad A \in 2^\Omega \setminus \emptyset.$$

Définition (Règle disjonctive hardie)

$$m_1 \circledvee m_2 = \bigcirc\!\!\!\!\cup_{A \neq \emptyset} A_{v_1(A) \wedge v_2(A)}.$$

Exemple

A	$m_1(A)$	$v_1(A)$	$m_2(A)$	$v_2(A)$	$v_{1 \circledvee 2}(A)$	$m_{1 \circledvee 2}(A)$
$\emptyset$	0.1	1	0.1	1	1	0.0060
$\{a\}$	0	1	0	1	1	0
$\{b\}$	0	1	0.5	1/6	1/6	0.0298
$\{a, b\}$	0.3	1/4	0	1	1/4	0.1071
$\{c\}$	0	1	0	1	1	0
$\{a, c\}$	0	1	0	1	1	0
$\{b, c\}$	0.6	1/7	0.4	0.6	1/7	0.2143
Ω	0	14/5	0	1	1	0.6429

Propriétés de la règle $\circledvee$

Commutativité : $\forall m_1, m_2,\ m_1 \circledvee m_2 = m_2 \circledvee m_1$

Associativité : $\forall m_1, m_2, m_3,\ m_1 \circledvee (m_2 \circledvee m_3) = (m_1 \circledvee m_2) \circledvee m_3$

Idempotence : $\forall m,\ m \circledvee m = m$;

Distributivité de $\circledcup$ par rapport à $\circledvee$: $\forall m_1, m_2, m_3,$

$$(m_1 \circledcup m_2) \circledvee (m_1 \circledcup m_3) = m_1 \circledcup (m_2 \circledvee m_3).$$

Propriétés de la règle $\textcircled{\vee}$

Relation avec $\textcircled{\cup}$ et $\textcircled{\vee}$

- Si m_1 et m_2 sont deux fcts de masse sous-normales t.q. $\overline{m}_1$ et $\overline{m}_2$ sont séparables, alors

$$m_1 \textcircled{\vee} m_2 \sqsubseteq_v m_1 \textcircled{\cup} m_2.$$

- Lois de De Morgan :

$$\begin{aligned} \overline{m_1 \textcircled{\vee} m_2} &= \overline{m}_1 \textcircled{\wedge} \overline{m}_2 \\ \overline{m_1 \textcircled{\wedge} m_2} &= \overline{m}_1 \textcircled{\vee} \overline{m}_2 \end{aligned}$$

Plan de la présentation

Généralisation des règles $\textcircled{\wedge}$ et $\textcircled{\vee}$

Résumé

	produit	minimum	$*$
poids conjonctifs w	$\textcircled{\cap}$	$\textcircled{\wedge}$	?
poids disjonctifs v	$\textcircled{\cup}$	$\textcircled{\vee}$	?

- Soit $*$ un opérateur binaire à valeurs dans $\mathbb{R}_+$. On définit :

$$m_1 \circledast_w m_2 = \textcircled{\cap}_{A \subset \Omega} A^{w_1(A) * w_2(A)}$$

$$m_1 \circledast_v m_2 = \textcircled{\cup}_{A \neq \emptyset} A_{v_1(A) * v_2(A)}.$$

- Propriétés de $*$?

Propriétés attendues de $\circledast_w$

- R_1 : Quelles que soient m_1 and m_2 non dogmatiques, $m_1 \circledast_w m_2$ doit être non dogmatique.
- R_2 : Commutativité et associativité de $\circledast_w$.
- R_3 : Quelles que soient m_1, m_2 and m_3 non dogmatiques :

$$m_1 \sqsubseteq_w m_2 \Rightarrow m_1 \circledast_w m_3 \sqsubseteq_w m_2 \circledast_w m_3.$$

Conditions suffisantes sur $*$

Théorème

Soit $*$ un opérateur binaire de $\mathbb{R}_+^2$ dans $\mathbb{R}_+$ vérifiant les conditions :

- *Positivité :* $\forall x, y > 0, \quad x * y = 0 \Rightarrow x = 0$ *or* $y = 0$;
- *Dominance par le minimum :*
$$x * y \leq x \wedge y, \quad \forall x, y > 0;$$
- *Commutativité, associativité ;*
- *Monotonicité :* $y \leq z \Rightarrow x * y \leq x * z, \quad \forall x, y, z > 0.$

Alors $\circledast_w$ vérifie les conditions R_1, R_2 et R_3.

Une famille d'opérateurs basés sur les t-normes

Théorème

Soit $\top$ *une t-norme positive et g une fonction strictement croissante de* $[0, +\infty)$ *dans* $[0, 1]$ *telle que g(0)=0. Alors l'opérateur operator* $*_{\top,g}$ *defini par*

$$x *_{\top,g} y = g^{-1}\left[g(x) \top g(y)\right], \quad \forall x, y \geq 0$$

vérifie les conditions précédentes.

Soient $\owedge_{\top,g}$ et $\ovee_{\top,g}$ les opérateurs conjonctif et disjonctif correspond à un choix de $\top$ et de g.

Conséquences

- Il est possible de définir une infinité d'opérateurs conjonctifs $\circledast_{\top,g}$ et disjonctifs $\circledast_{\bot,g}$ vérifiant les propriétés R_1, R_2 et R_3 (donc commutatifs et associatifs).
- Les règles $\textcircled{\wedge}$ et $\textcircled{\vee}$ sont retrouvées comme cas particuliers en prenant $\top = \wedge$ et g quelconque.
- Les règles $\textcircled{\cap}$ et $\textcircled{\cup}$ peuvent être rerouvées si on se restreint à des fonctions de masse particulières (séparables ou de complémentaire séparable).
- Famille d'opérateurs contenant $\textcircled{\cap}$ et $\textcircled{\wedge}$, $\textcircled{\cup}$ et $\textcircled{\vee}$?
- Possibilité d'identification d'un opérateur parmi une famille paramétrée à partir d'exemples.

Conclusion

- Deux règles duales commutatives, associatives et idempotentes :
  - règle conjonctive prudente $w_{1 \textcircled{\wedge} 2} = w_1 \wedge w_2$;
  - la règle disjonctive hardie $v_{1 \textcircled{\vee} 2} = v_1 \wedge v_2$.
- Règles dérivées du Principe d'Engagement Minimal, pour des relations d'ordres particulières.
- Une infinité d'opérateurs commutatifs et associatifs (mais non idempotents) peuvent être construits à partir de normes triangulaires quelconques : même flexibilité que dans la théorie des possibilités.
- L'interprétation et l'intérêt pratique de ces opérateurs généralisés restent à étudier.

Références I

Ph. Smets.
The canonical decomposition of a weighted belief.
In *Int. Joint Conf. on Artificial Intelligence*, pages 1896–1901, San Mateo, Ca, 1995. Morgan Kaufman.

T. Denœux.
The cautious rule of combination for belief functions and some extensions.
In *Proceedings of the 9th International Conference on Information Fusion*, Florence (Italy), July 2006.