Université des Sciences et de la Technologie
Houari Boumediene

Faculté d'Electronique et d'Informatique
Département d'Informatique

La théorie des fonctions de croyances et ses applications

Mme Khellaf-Haned Hadja Faiza
Melle Zebbouchi Karima
Mr Debbieche Mounaim

Avril 2012

Motivations

- Gestion de l'incertitude dans les systèmes à base de connaissances .
- Deux sources possibles d'incertitude (connaissance, ou données) :

1. la connaissance (du domaine) peut être **incomplète, ou vague :**

   - théories imparfaites du domaine physique ;

   - concepts vagues ou flous ;

2. les données peuvent être imprécises, non fiables, ou vagues.

Approches de l'incertitude

On peut distinguer:

1. Les approches probabilistes (réseaux bayésiens).

 - Difficulté de prise en compte de certains aspects « humains » de la connaissance,

 - pouvant être assez lourd.

2. les approches de la logique floue (logique floue, logique des possibilités).

3. des théories basées sur les notions de « croyances » et de « preuves » qui veulent simuler le raisonnement humain (la *théorie de Dempster-Shafer) ;*

Plan

1-Généralités
2- Exemple introductif
3-Modélisation des connaissances
4-Estimation des fonctions de masse
5-Combinaison des connaissances
6-Prise de décision
7-Domaines d'application
8- Conclusion

1-Généralités

- la théorie de l'évidence, de croyance ou la théorie du raisonnement plausible est une théorie mathématiques basée sur la notion de « preuve », introduite par Arthur Dempster en 1968 et développée par Glenn Shafer en 1976.

- C'est une généralisation de l'inférence Bayésienne au traitement de l'incertain, elle permet de manipuler des événements non nécessairement exclusifs. Cette capacité lui confère l'avantage de pouvoir représenter explicitement l'incertitude sur un événement.

- Un indice corrobore souvent sur un ensemble d'hypothèses, sans distinguer entre ces hypothèses (*attribuer des valeurs aux sous-ensembles, et pas seulement aux singletons).*

- *Elle permet de* distinguer ce en quoi on « croit », c'est-à-dire ce que les indices indiquent, et ce qui nous semble « plausible », c'est-à-dire ce qui ne contredit pas les indices (*au lieu* d'avoir une probabilité, on a un *intervalle de valeurs).*

- Elle prend en compte ce qui reste inconnu et elle représente parfaitement ce qui est déjà connu.

2- Exemple introductif

- On suppose qu'une personne dispose, pour se déplacer, de plusieurs possibilités :

- une voiture, *a ;*
- un vélo, *v ;*
- ses pieds, *p ;*
- le bus, *b*.

- On veut déterminer quel moyen de transport elle utilisera demain pour se rendre au travail.

- L'ensemble des possibilités (appelé *univers ou cadre de discernement ou hypothèses) est* $\Theta = \{a, a, v, p, b\}$

2- Exemple introductif

- Initialement, tout est plausible, mais on ne croit en rien (en dehors d'être dans Θ).

- La fonction de croyance, Bel : $P(\Theta) \rightarrow [0, 1]$, pourrait s'écrire :

$$Bel(E) = \begin{cases} 1 \text{ si } E = \Theta \\ 0 \text{ sinon} \end{cases}$$

- La fonction de plausibilité, Pl : $P(\Theta) \rightarrow [0, 1]$, pourrait s'écrire :

$$Pl(E) = \begin{cases} 0 \text{ si } E = \varnothing \\ 1 \text{ sinon} \end{cases}$$

- $Pl(E) = 1 - Bel(\neg E)$: E est plausible si on ne croit pas en sa négation.

2- Exemple introductif (masses initiales)

- La théorie de Dempster-Shafer manipule en fait une *fonction de masse m qui résume les deux fonctions.*

- *La fonction de masse (ou* « affectation de probabilité de base ») attribue à chaque sous-ensemble de Θ le « poids » de preuves corroborant ce sous-ensemble (et spécifiquement lui, pas ses éléments ou un sur-ensemble).

- Ainsi, initialement :

$$m(E) = \begin{cases} 0 \text{ si } E \neq \Theta \\ 1 \text{ si } E = \Theta \end{cases}$$

- la valeur de 1 pour *E est faite pour « compenser », la somme des valeurs de la* fonction de masse devant faire 1.

2- Exemple introductif (masses initiales)

On définit alors :

$$\text{Bel}(X) = \sum_{A \subseteq X} m(A)$$

La croyance en X *est donc la somme des masses des sous-ensembles de* X,

$$\text{Pl}(X) = \sum_{A \cap X \neq \emptyset} m(A)$$

la plausibilité est la somme des masses « compatibles » avec X.

2- Exemple introductif (premier indice)

-La météo prévoit (a priori) de la pluie demain. On peut estimer ses chances d'avoir raison de 80 %. Dans ce cas notre personne n'ira ni à pieds ni en vélo.
-Après modification , on pourrait avoir comme nouvelle fonction de masse :

$$m_1(\{a, b\}) = 0,8$$
$$m_1(\Theta) = 0,2$$

-On note que s'il ne pleut pas, ça ne veut pas dire que la personne ira à pieds ou en vélo, donc on n'attribue pas de masse à $\{p, v\}$.

- *D'où, par exemple :*

	Bel	Pl		Bel	Pl
$\{a\}$	0	1	$\{b, p, v\}$	0	1
$\{a, v\}$	0	1	$\{b, p\}$	0	1
$\{a, b\}$	0,8	1	$\{p, v\}$	0	0,2
$\{a, b, v\}$	0,8	1	$\{p\}$	0	0,2
$\{a, b, p, v\}$	1	1	$\emptyset$	0	0

2- Exemple introductif (deuxième indice)

• La personne n'habite pas à Alger même Ça diminue ses chances d'aller à pieds ou en bus (indépendamment du lieu de travail).
Sans combinaison, on aurait pas exemple :

$$m_2(\{a, v\}) = 0,6$$

$$m_2(\Theta) = 0,4$$

• La nouvelle fonction de masse avec combinaison devrait avoir cet aspect :
- *a seul est renforcé par les deux indices :* $m(\{a\}) > 0$
- $\{a, v\}$ *et* $\{a, b\}$ *gardent une masse, mais « transfèrent » une partie de celle-ci vers* $\{a\}$.
- Θ garde le reste.

• Concrètement, le calcul est le suivant :

$$m(\{a\}) = m_1(\{a, b\}).m_2(\{a, v\}) = 0,48$$

$$m(\{a, b\}) = m_1(\{a, b\}).m_2(\Theta) = 0,32$$

$$m(\{a, v\}) = m_1(\Theta).m_2(\{a, v\}) = 0,12$$

$$m(\Theta) = m_1(\Theta).m_2(\Theta) = 0,08$$

2- Exemple introductif (deuxième indice)

Le tableau de croyance contient alors :

	Bel	Pl		Bel	Pl
{a}	0,48	1	*{b, p, v}*	0	0,52
{a, v}	0,6	1	*{b, p}*	0	0,4
{a, b}	0,8	1	*{p, v}*	0	0,2
{a, b, p}	0,8	1	*{v}*	0	0,2
{a,b,v}	0,92	1	{p}	0	0,08
{a, b, p, v}	1	1	$\varnothing$	0	0

On constate la conservation des fonctions pour *{a, b} et {a, v} par rapport respectivement* à m_1 *et* m_2.

La formule appliquée au-dessus ressemblerait à :

$$m(E) = \sum_{X \cap Y = E} m_1(X).m_2(Y)$$

2- Exemple introductif (troisième indice)

• Indice : la distance entre son domicile et son travail est de moins de 5 km. On suppose que ça encourage la marche.
Par exemple :

$$m_3(\{p\}) = 0,25$$
$$m_3(\Theta) = 0,75$$

• Si on appliquait la proposition précédente, on obtiendrait :

$$m(\{p\}) = m_1(\Theta).m_3(\{p\}) = 0,05$$
$$m(\{a, b\}) = m_1(\{a, b\}).m_3(\Theta) = 0,6$$
$$m(\Theta) = m_1(\Theta).m_3(\Theta) = 0,15$$
$$m(\emptyset) = m_1(\{a, b\}).m_3(\{p\}) = 0,2$$

• $m(\emptyset) > 0$ *indique que les indices sont contradictoires. Néanmoins, cela n'a pas de sens de* donner un croyance dans aucune hypothèse.
Il faut donc ramener $m(\emptyset)$ *à* 0.

2- Exemple introductif (normalisation)

- On le fait en normalisant les autres valeurs, c’est-à-dire en les divisant par $1 - m(\emptyset)$.

Le but est que la somme des valeurs de m soit toujours à 1.

$$m(\{p\}) = m_1(\Theta).m_3(\{p\})/0,8 = 0,0625$$
$$m(\{a, b\}) = m1(\{a, b\}).3_2(\Theta)/0,8 = 0,75$$
$$m(\Theta) = m_1(\Theta).m_3(\Theta)/0,8 = 0,1875$$
$$m(\emptyset) = 0 \text{ (par définition)}$$

	Bel	**Pl**		**Bel**	**Pl**
{a}	0	0,9375	*{b, p, v}*	0,6251	1
{p}	0,0625	0,25	*{b, p}*	0,75	0,9375
{v}	0	0,1875	*{a,b,v}*	0,8125	1
{a, b}	0,75	0,9375	*{p,v}*	0,0625	0,25
{a,p}	0,0625	1	{b,p}	0	0,9375
{a, b, p, v}	1	1	$\emptyset$	0	0

2- Exemple introductif (combinaison)

Quel que soit l'ordre de combinaison des 3 indices (Les 3 masses), on obtient le même résultat :

$$m(\{a\}) = 0,468$$
$$m(\{a, b\}) = 0,311$$
$$m(\{a, v\}) = 0,117$$
$$m(\{p\}) = 0,026$$
$$m(\Theta) = 0,078$$

Pour pouvoir prendre une décision adéquate à partir d'informations issues de plusieurs sources , en utilisant la théorie de l'évidence, il faut passer par les étapes suivantes :

1-Modélisation

2-Estimation

3-Combinaison

4-Décision

3- Modélisation des connaissances

1- Cadre de discernement

- le cadre de discernement ou cadre d'intérêt = $\{H_1,...,H_n\}$ pour le problème traité est composé de n hypothèses exhaustives et exclusives, dont l'une d'elles est susceptible d'être la solution du problème posé.

- le raisonnement porte sur l'ensemble des parties 2^{Θ} appelé également « power-set », qui est l'ensemble de tous les sous-ensembles de Θ.

- Tous les éléments de ce « power-set » correspondent à une proposition disjonctive (union de « singletons ») du cadre de discernement.

3- Modélisation des connaissanc

- Propriétés du cadre de discernement

- Ensemble des hypothèse **exhaustif**: la solution est l'une des hypothèses H_i de Θ (en logique propositionnelle, désigne la proposition certaine).

<u>Remarque</u> : Il existe certains types de problèmes où nous rencontrons une difficulté à recenser toutes les hypothèses.
Ici le cadre est non exhaustif , l'ensemble est dit « open world »

- Hypothèses sont mutuellement **exclu** $H_i \cap H_j = \emptyset \; \forall i \neq j$ **)**

2-Fonctions de masse élémentaires

- L'expression d'un degré de confiance, appelé croyance, pour chacun des sous ensembles de Θ est effectué à travers l'affectation d'une masse élémentaire (m) définie de 2^{Θ}sur [0,1] de la façon suivante :

$$m(\emptyset) = 0$$

et

$$\Sigma\, m(A) = 1 \text{ où } A \subseteq \Theta.$$

Interprétation

- m(.) représente la masse d'évidence associée à une source d'information sur un certain nombre d'hypothèses ou de propositions du cadre de discernement

Interprétation (suite)

- Aucune croyance n'est affectée à l'ensemble vide (Ø).
- La somme de toutes les fonctions de masse attribuées aux différents sous-ensembles d'une hypothèse devra être égale à 1.
- m(A) correspond à la masse de croyance « directe » affectée à l'hypothèse ou proposition A, c'est-à-dire au degré de confiance accordé à A selon les sources d'information (appelés également sources d'évidences).
- Les sous ensembles ***A*** tels que ***m(A)*** >0 sont appelés éléments focaux de m.
- Il faut noter que si l'on considère deux sources d'information ou évidences S_1 et S_2, $m(A_1)$ et $m(A_2)$ peuvent être tous les deux nuls sans pour autant que l'on obtienne $m(A_1 U A_2) = 0$.

Propriétés

- Si $\boldsymbol{m}$ **(Θ) = 1**, alors la source *S* est dans *l'ignorance totale*.

- Si $\boldsymbol{m}$ **(*A*) = 1** (***A*** une disjonction d'éléments de **Θ),** alors la source ***S*** a une connaissance imprécise ; elle ne croit qu'en *A*.

- Si $\boldsymbol{m}$ **(*C*)** **=** 1 (***C*** est un élément singleton de **Θ**), alors la source *S* a une connaissance précise, elle croit pleinement en ***C***.

- Si $\boldsymbol{m}$ **(*A*)** = ***s*** et $\boldsymbol{m}$ **(Θ)** = 1-***s***, alors la source *S* a une connaissance incertaine et imprécise ; elle croit en partie en ***A***, mais rien de plus.

3-Fonctions de croyance

- La fonction de croyance (Bel) ou de crédibilité (Cr) est dérivée des fonctions de masse.

Ainsi, *Cr* (A) totalise toutes les raisons de soutenir l'hypothèse A.

- La mesure de la croyance (Cr) s'effectue sur le cadre de discernement, mais il est possible de la représenter par rapport à une fonction de masse comme suit :

$$Cr(A) = \sum_{B \subset A} m(B).$$

Note:

- Une fonction de croyance nulle représente ainsi l'ignorance totale sur le problème considéré.

4-Fonctions de plausibilité

- La plausibilité (Pl) est également dérivée des fonctions de masse. La fonction de plausibilité *Pl* (A) exprime l'intensité avec laquelle on devrait soutenir l'hypothèse A si tous les faits non connus allaient dans le sens de cette hypothèse. Elle peut être définie de deux manières différentes :

  - Soit par la fonction de croyance de l'évènement contraire à A, notée A^c

$$Pl(A) = 1 - Cr(A^c)$$

  - Soit par rapport à la fonction de masse qui correspond à la définition originale de Dempster :

$$Pl(A) = \sum_{A \cap B \neq \Phi} m(B).$$

3- Modélisation des connaissan

Remarque

• *Cr*(A) et *Pl*(A) ont une valeur comprise dans l'intervalle [0,1].

• A partir de ces deux valeurs, on peut déterminer une probabilité inconnue nommée

« intervalle de croyance (Ic) »
ou **« intervalle de confiance »**
qui correspond à :

$$Cr(A) \leq Ic(A) \leq Pl(A).$$

4- Estimation des fonctions de masses

Bien que c'est une étape difficile à mettre en œuvre, plusieurs approches ont été développées pour l'estimation des fonctions de masse nous citons :

- L'approche basée sur la vraisemblance,
- L'approche basée sur la distance.

1- Approche basée sur la vraisemblance

Nous supposons connues $f(x|\theta i)$ les densités de probabilité conditionnellement aux classes.
En ayant observé x, la fonction de vraisemblance $L(\theta q|x)$ est une fonction de Θ dans[0,+ ∞] définie par $L(\theta q \mid x) = f(x|\theta q)$ pour tout $q \in \{1, \ldots . Q\}$

A partir de L, **Shafer** a proposé de construire une fonction de croyance sur Θ définie par sa fonction de plausibilité suivante :

$$Pl(A) = \frac{max\ \theta q \in A[L(\theta q|x)]}{MAXq\ [\ L(\theta q|x)]}\ \ \forall A \subseteq \theta.$$

4- Estimation des fonctions de masses

→Une autre méthode a été proposée par **Appriou** basée sur la construction de Q fonction de croyance $m_q(.)$.

→L'idée consiste à prendre en compte de manière séparée chaque classe et à évaluer le degré de croyance affecté à chacune d'entre elles.

→ Les éléments focaux de cette fonction de croyance m_q sont des singletons.

→ Appriou a obtenu deux modèles :

Modèle 1

$$mq(\{\theta q\}) = \alpha q \frac{R.L(\theta q|x)}{1+R.L(\theta q|x)}$$

$$m_q(\overline{\theta q}) = \alpha q \frac{1}{1+R.L(\theta q|x)}$$

$$m_q(\Theta) = 1-\alpha q$$

Modèle 2

Le plus utilisé

$$m_q(\{\theta q\}) = 0$$

$$m_q(\overline{\theta q}) = \alpha_q(1-R.L(\theta q|x))$$

$$m_q(\Theta) = 1-\alpha_q(1-R.L(\theta q|x)).$$

4- Estimation des fonctions de masses

Notes

- Dans ces équations, αq est un coefficient qui peut être utilisé pour modéliser une information complémentaire comme par exemple la fiabilité d'un capteur , et R est une constante de normalisation qui est choisie, dans la plage

$$]0, (\max(L(\theta q|x)))^{-1}]$$

- En pratique , les performances de ces deux méthodes sont équivalentes.

- Cependant Appriou recommande l'utilisation du deuxième modèle qui a l'avantage d'être consistent avec le « théorème de Bayes généralisé » proposé par Smets.

4- Estimation des fonctions de masses

2- Approche basée sur la distance

- Une approche tout à fait différente a été développée par **Denoeux.** Dans cette méthode, une fonction de croyance est directement construite par le biais des vecteurs d'apprentissage situés au voisinage du vecteur à classer.

- Si k plus proches voisins (au sens de la distance euclidienne par exemple) sont considérés, nous obtenons ainsi k fonctions de croyance qui seront alors combinées par l'opérateur de Dempster.

- La méthode initiale a ensuite été améliorée en optimisant les paramètres .

4- Estimation des fonctions de masses

- Cette version utilise un ensemble de prototype dont chacun est considéré comme un corpus d'évidence qui influence la croyance dans l'appartenance de x à une classe .une fonction de croyance m^i associe à chaque prototype i et ainsi définie.

$$\forall q \in \{1,\dots,Q\}$$

$$m^i(\{\Theta_k\}) = \alpha^i\phi^i(d^i),$$

$$m^i(\Theta) = 1 - \alpha^i\phi^i(d^i),$$

$$m^i(A) = 0 \quad \forall A \in 2^\Theta \setminus \{\{\theta_q\},\Theta\}.$$

- Où d^i est la distance Euclidienne au i-eme prototype, α^i est un paramètre associé à chaque prototype i et $\phi i(.)$ est une fonction décroissante définie par :

$$\phi i(d^i) = \exp\left[-\gamma^i (d^i)^2\right].$$

- Dans cette expression, γ^i est un paramètre positif associé à chaque prototype i. Les fonctions de croyance m^i pour chaque prototype sont ensuite combinées avec la règle de Dempster.

5- Combinaison des connaissances

Combinaison

Dempster

Conjonctive

$$m(A) = m1 \wedge m2 = \sum_{\substack{B \cap C = A \\ B, C \subseteq \Theta}} m1(B)m2(C) \quad \forall A \subseteq \Theta.$$

Smets

Disjonctive

$$m(A) = m1 \vee m2 = \sum_{\substack{B \cup C = A \\ B, C \subseteq \Theta}} m1(B)m2(C) \quad \forall A \subseteq \Theta.$$

Opérateurs conjonctifs et disjonctifs

Soient m1 et m2 deux fonctions de masse sur issues de 2 sources.

Un opérateur de combinaison fusionne m1 et m2 en une nouvelle fonction de masse m12 = m1 m2 (nouvel état de connaissance).

L'opérateur est:

- conjonctif si m12 est plus informative que m1 et m2.
- disjonctif si m12 est moins informative que m1 et m2.

- Le choix d'un opérateur conjonctif suppose que les sources soient toutes deux fiables.

- la combinaison disjonctive : stratégie prudente (l'une au moins des deux sources est fiable).

5- Combinaison des connaissances

- Dempster introduit une règle permettant de combiner deux fonctions de croyances, appelée *règle de combinaison de Dempster*.

- Cette règle est une opération associative et commutative qui associe à un couple de fonctions de croyance définies sur le même espace une **nouvelle** fonction de croyance sur **S**.

5-Combinaison des connaissances

Notes :

- Shafer ne considère que des fonctions de croyance normalisées, c'est-à-dire que m(***Φ***) = 0.

- La particularité de la théorie de l'évidence est qu'elle permet de quantifier le conflit entre des sources.

- soient deux sources 1 et 2, la valeur quantifiant le conflit peut être évaluée comme suit:

$$K = m(\Phi) = \sum_{B \cap C = \Phi} m1(B).m2(C).$$

- K est dit conflit conjonctif.

5- Combinaison des connaissances

- Les causes du conflit peuvent être multiples :
  - mauvais fonctionnement du capteur,
  - mauvaise définition de l'ensemble des mondes possibles,
  - mauvaise définition des fonctions de croyance...

- Dans la théorie de l'évidence, Dempster préconise de re normaliser, comme suit, la distribution de masse obtenue après combinaison :

$$m(A) = m1 \wedge m2 = \frac{1}{1-k} \sum_{\substack{B \cap C = A \\ B, C \subseteq \Theta}} m1(B)m2(C) \quad \forall A \subseteq \Theta.$$

5-Combinaison des connaissances

Exemple : Combinaison d'informations avec la règle conjonctive.

- Soit un décideur qui est en mesure d'utiliser trois capteurs d'informations statistiquement indépendantes et de même fiabilité.

- Le capteur 1 donne l'information suivante :

$$\{m(\theta 1) = 0.60; m(\theta 2) = 0.20; m(It) = 0.20\}.$$

- Le capteur 2 donne l'information suivante :

$$\{m(\theta 1) = 0.20; m(\theta 2) = 0.20; m(It) = 0.60\}.$$

- Le capteur 3 donne quant à lui l'information suivante :

$$m(\theta 1) = 0.50; m(\theta 2) = 0.40; m(It) = 0.10\}.$$

5-Combinaison des connaissances

Table1

	$m1(A)$	$m2(A)$	$m3(A)$
$\{\theta 1\}$	0.6	0.2	0.5
$\{\theta 2\}$	0.2	0.2	0.0
$\{\theta 3\}$	0.0	0.0	0.4
$\{It\}$	0.2	0.6	0.1

La première étape consiste à faire le produit orthogonal entre les informations du capteur1 avec ceux du capteur 2.
Selon l'équation de combinaison on obtient la table 2.

Table2

	$m1(\theta 1)$ 0.6	$m1(\theta 2)$ 0.2	$m1(It)$ 0.2
$m2(\theta 1)$ 0.2	$\theta 1$ 0.12	$\theta 1 \cap \theta 2 = \emptyset$ 0.04	$\theta 1$ 0.04
$m2(\theta 2)$ 0.2	$\theta 1 \cap \theta 2 = \emptyset$ 0.12	$\theta 2$ 0.04	$\theta 2$ 0.04
$m2(It)$ 0.6	$\theta 1$ 0.36	$\theta 2$ 0.12	It 0.12

5-Combinaison des connaissances

La seconde étape consiste à combiner les résultats de la combinaison du capteur 1 et du capteur 2 avec les informations du capteur 3. Il en résulte la table 3.

Table3

	$m12(\theta1)$ 0.52	$m12(\theta2)$ 0.20	$m12(\theta1 \cap \theta2) = \emptyset$ 0.16	$m12(It)$ 0.12
$m3(\theta1)$	$\theta1$ 0.26	$\theta1 \cap \theta2 = \emptyset$ 0.10	$\emptyset$ 0.08	$\theta1$ 0.06
$m3(\theta3)$	$\theta1 \cap \theta3 = \emptyset$ 0.208	$\theta2 \cap \theta3 = \emptyset$ 0.08	$\emptyset$ 0.064	$\theta3$ 0.048
$m3(It)$	$\theta1$ 0.052	$\theta2$ 0.02	$\emptyset$ 0.016	It 0.012

Une étape finale de regroupement donne la table finale 4. On y retrouve l'objet dominant $\theta1$ qui est l'identité choisie par le décideur s'il se base sur la masse la plus élevée et que l'on rejette $\emptyset$.

Table4

	$\theta1$	$\theta2$	$\theta3$	It	$Le\ reste$
$m123$	**0.372**	0.02	0.048	0.012	0.548

5-Combinaison des connaissances

La table 5 représente la combinaison avec redistribution de la masse conflictuelle suivant la règle de Dempster.

Table5

	$\theta 1$	$\theta 2$	$\theta 3$	$\theta 1 \cap \theta 2$	$\theta 1 \cap \theta 3$	$\theta 2 \cap \theta 3$	$\theta 1 \cap \theta 2 \cap \theta 3$	It	$\emptyset$
$m123$	**0.823**	0.044	0.106	0.000	0.000	0.000	0.000	0.027	0.000

6- Prise de décision

La prise de décision dans la théorie de Dempster et Shafer est une étape difficile qui a fait et fait toujours l'objet de nombreux débats.

Elle est obtenue par la maximisation d'un critère :

- Maximum de croyance

le critère de maximum de croyance consiste à choisir d_i dans l'observation x si :

$$Cr\,(d_i)\,(x) = \max_{1 \leq k \leq n} Cr\,(d_k)\,(x)$$

Note

→ *Cette méthode est connue pour son caractère pessimiste.*

- Maximum de plausibilité

- Cette méthode consiste à choisir un singleton d_i qui maximise la plausibilité ainsi pour l'observation *x*, *nous choisirons* d_i si :

$$PL\,(d_i)\,(x) = \frac{MAX}{1 \leq K \leq n} PL\,(d_i)\,(x).$$

Note

→*L'inconvénient de cette méthode est qu'elle est considérée comme très optimiste car prendre le maximum de plausibilité sur les singletons, revient à déterminer la décision donnant le maximum de chance à chacun des singletons.*

- Maximum pignistique

- la probabilité pignistique, ainsi définie par Smets, est une mesure de probabilité subjective qui se trouverait entre la mesure de crédibilité et de plausibilité . Elle est exprimée de la façon suivante :

$$\forall \theta_i \in \Theta, P\{\theta_i\} = \sum_{B \subseteq \Theta | \theta_i \in B} \frac{1}{|B|} m(B)$$

6-Prise de décision

Exemple

Soit Θ={A,B,C}
m (Ø) = 0
m(A) = 0.30
m(B) = 0.10
m(C) = 0.20
m(AU B) = 0.10
m(AU C) = 0.20
m(BU C) = 0.05
m(AU BU C) = 0.05
Σ m. = 1

Prise de décision pessimiste (Cr)

Cr(A) = m(A) = **0,30**
Cr(B) = m(B) = **0,10**
Cr(C) = m(C) = **0,20**

Prise de décision optimiste (Pl)

Pl(A) = m(A) + m(AU B) + m(AU C) + m(A U BU C) = **0,65**
Pl(B) = m(B) + m(AU B) + m(BU C) + m(AU BU C) = **0,30**
Pl(C) = m(C) + m(AU C) + m(BU C) + m(AU BU C) = **0,50**

Prise de décision pignistique (Cr)

P(A) = m(A) +1/2m(AU B) +1/2 m(AU C) +1/3 m(AU BU C) = **0,466**
P(B) = m(B) +1/2 m(AU B) +1/2 m(BU C) +1/3 m(AU BU C) = **0,191**
P(C) = m(C) +1/2 m(AU C) +1/2 m(BU C) +1/3 m(AU BU C) = **0,341**

7- Domaines d’application

La théorie des croyances peut être appliquée dans un grand nombre de situations ce qui a poussé à l'introduction de fonction de croyance dans toute sorte de domaines.
Nous reprenons ci-dessous quelque situations dans lesquelles la théorie des croyances est intéressante :

- En traitement d'images (exemple les images médicales).
- En traitement des donnés météorologique .
- En traitement de la parole.
- Pour les données temporelles.
- Pour la classification .

7-Domaines d’application

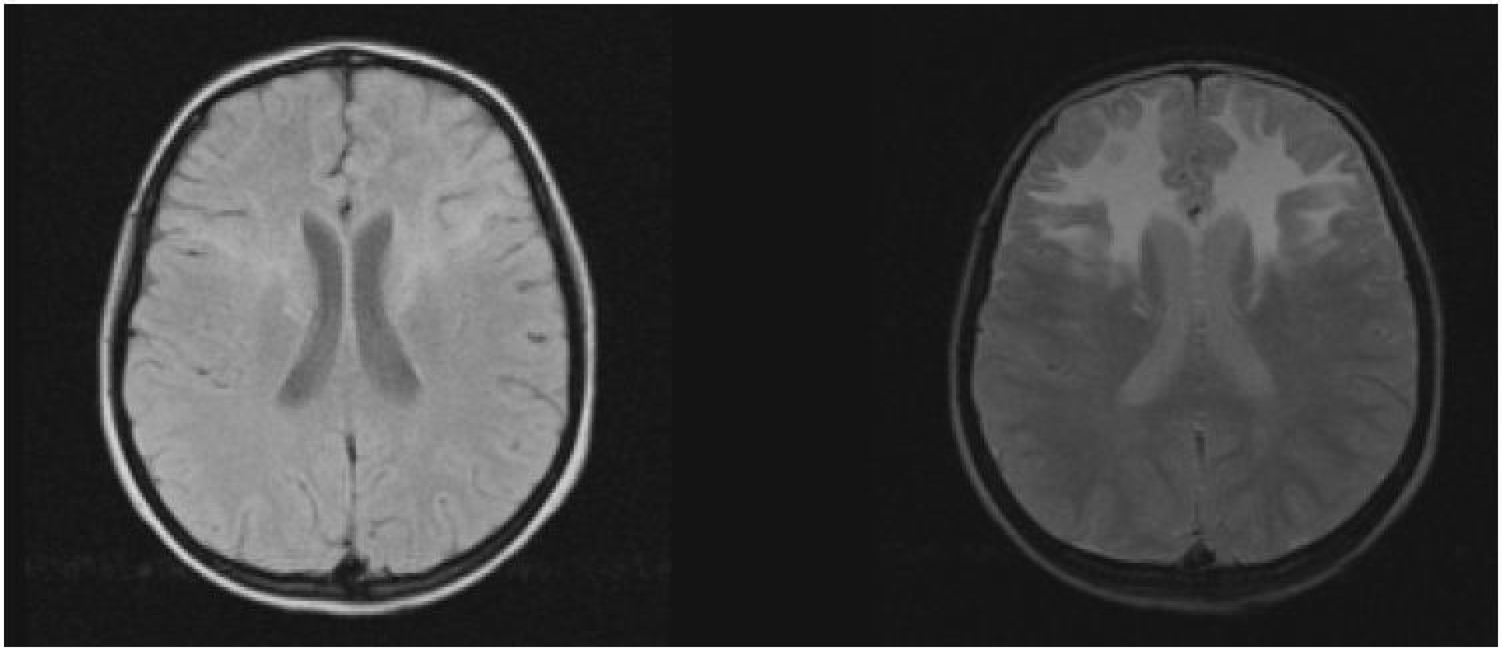
En traitement d’images

• la théorie de Dempster-Shafer a souvent été abordée en traitement d'images par exemple dans le cas d'images médicales pour but de segmenter les images gautier99 pour des images IRM ou de détecter des anomalies (pour des IRM).

• Cette théorie exprime la grande flexibilité des modélisations possibles en traitement d’image, prenant en compte à la fois l'incertitude et l'imprécision, l'ignorance partielle ou globale, la fiabilité des sources , la capacité de chaque source à fournir des informations fiables ou non sur chaque classe, des informations a priori qui ne sont pas forcement représentables par des probabilités, etc.

7-Domaines d’application

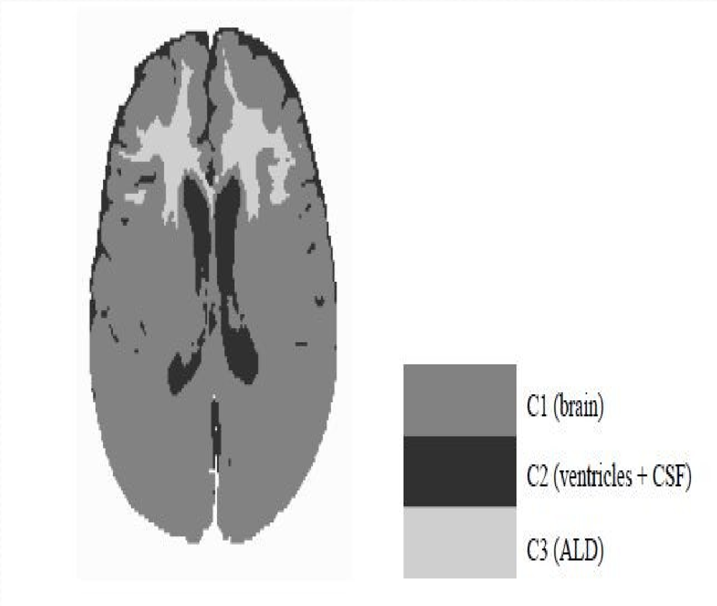
Exemple en imagerie médicale

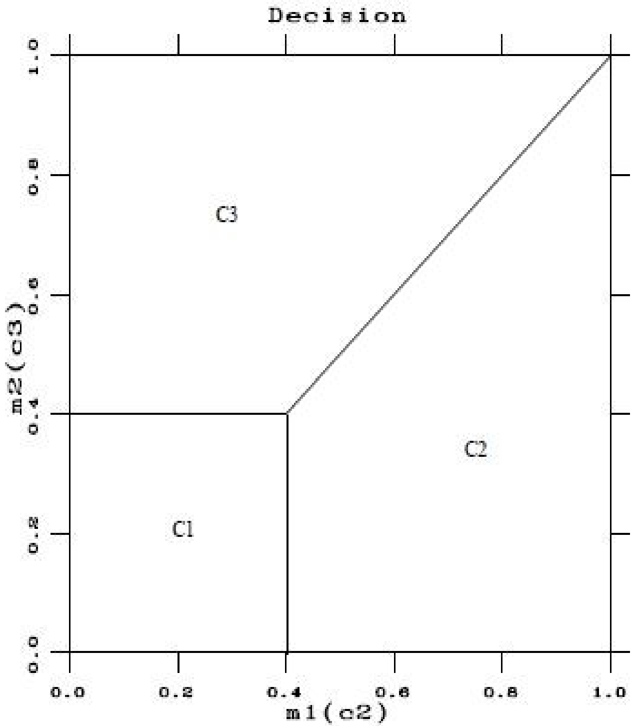


Exemple d'une coupe IRM du cerveau acquise avec deux temps d‘échos.

La zone pathologique correspond aux zones les plus blanches dans le haut de l'image.

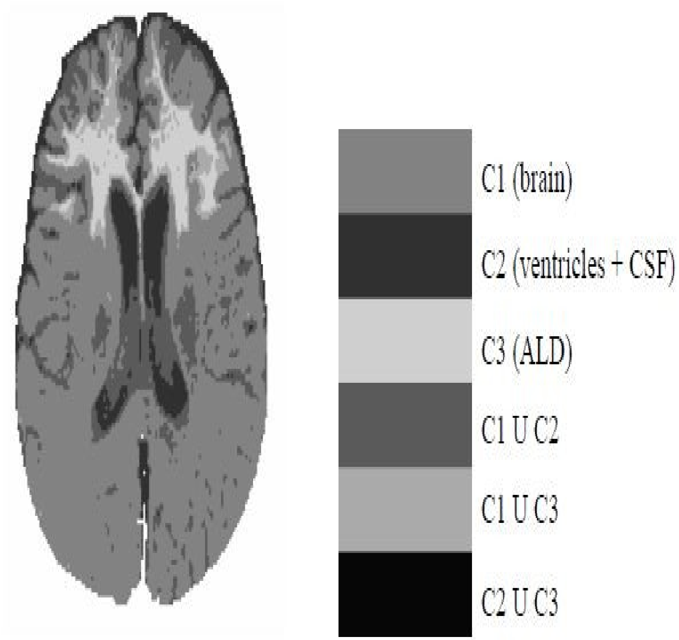
Exemple en imagerie médicale

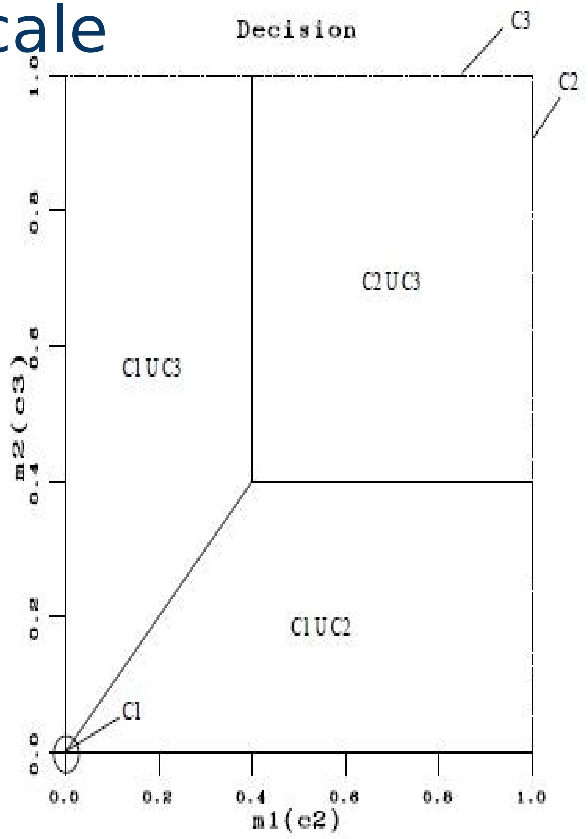




Les différentes zones de décision en fonction des valeurs de m1(C2) et m2(C3) et image de décision, en prenant le maximum de croyance sur les hypothèses simples.

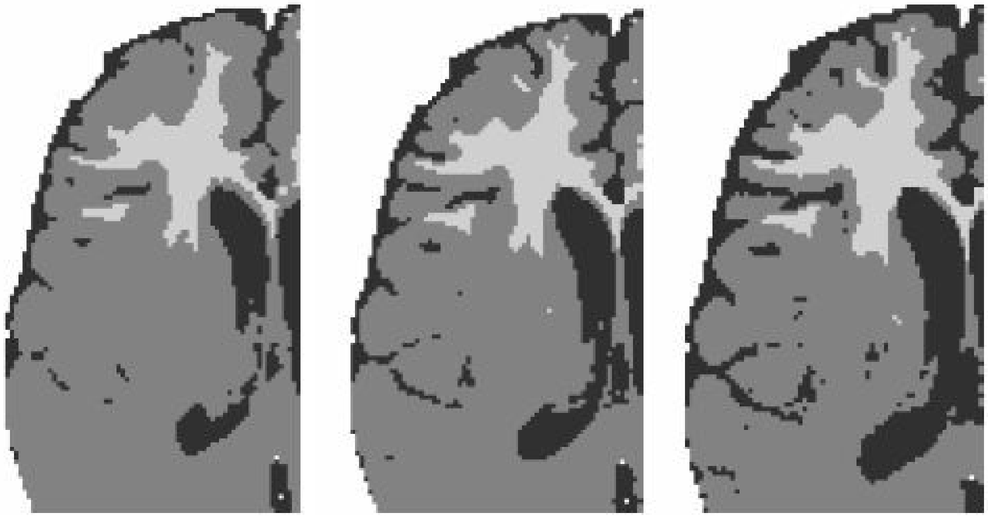
Exemple en imagerie médicale





Les différentes zones de décision en fonction des valeurs de m1(C2) et m2(C3) et image de décision, en prenant le maximum de croyance sur toutes les hypothèses sauf D.

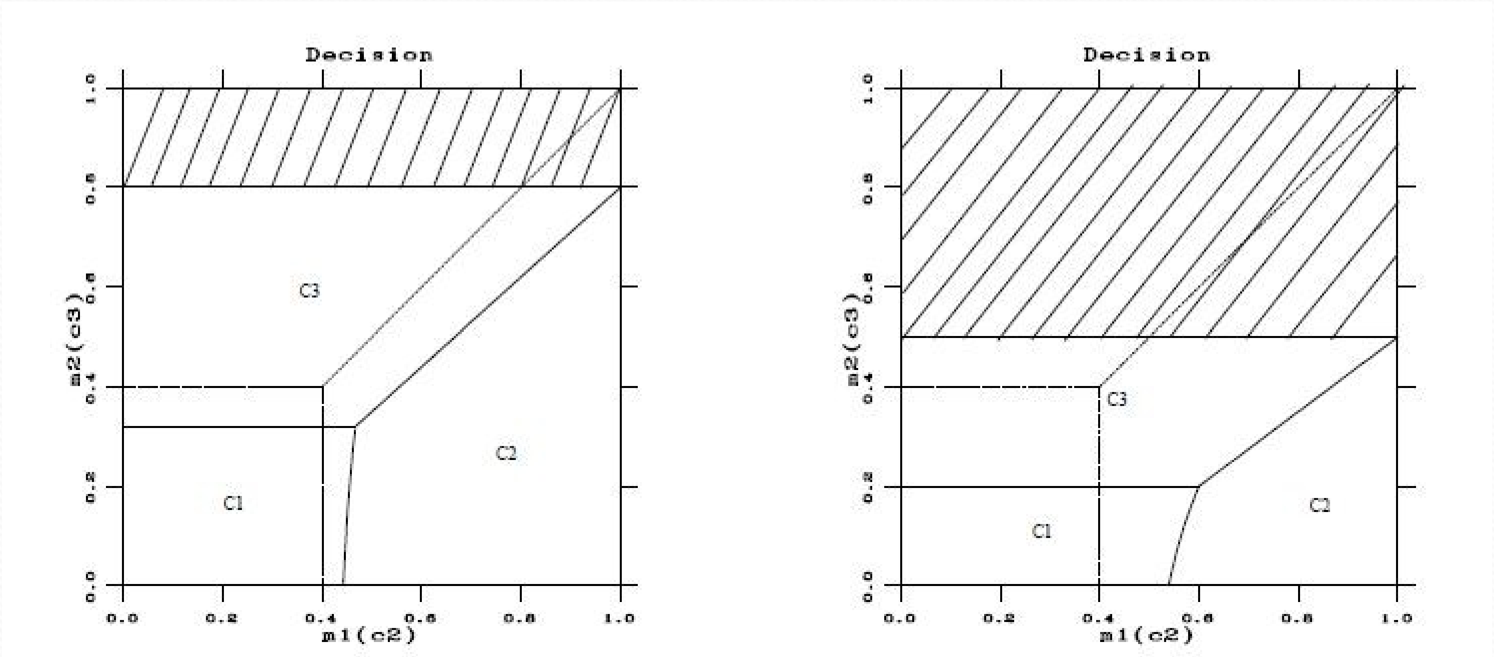
Exemple en imagerie médicale



Décision par le maximum de croyance sur les hypothèses simples,
pour un poids croissant sur m2(C3) et m1(C2)

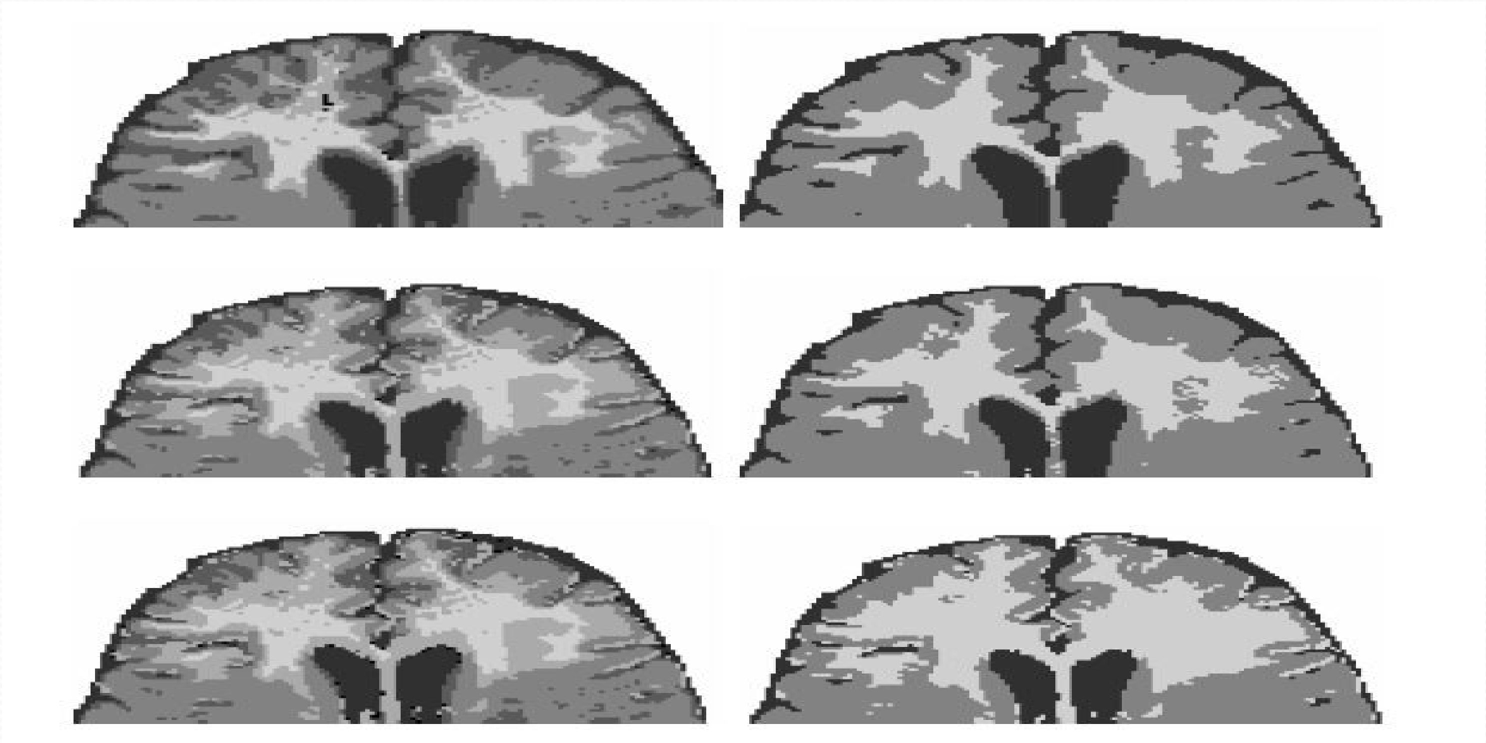
7-Domaines d’application

Exemple en imagerie médicale



Zones de décision en fonction des valeurs de m1(C2) et m2(C3) en introduisant
une masse croissante sur C1 [C3 pour la deuxième image (m2(C3) doit être inferieur a 1 - m2(C1 ∪ C3)).

Exemple en imagerie médicale



Décision par le maximum de croyance sur toutes les hypothèses sauf D
(première colonne) et sur toutes les hypothèses simples
(deuxième colonne)

7-Domaines d’application

Images satellitaires

la théorie des fonctions de croyances peut être utilisée dans le cas des images satellitaires multi capteurs et multi dates pour la détection et la cartographie des changements d’occupation des terres.

7-Domaines d’application

- En traitement des donnés météorologique

- La théorie des croyances de Dempster-Shafer a été utilisée pour la fusion des informations (imparfaites) issues des capteurs de mesures de pollution de manière à prédire l’apparition d’un pic à l’ozone

Application Samasafia

Station
d'EL
HAMMA

les
paramètres
existants:

- SO2
- NO2
- O3
- PM10

Station
Du 1er MAI

les
paramètres
existants:

- SO2
- NO2
- PM10

Station
de BEN
AKNOUN

les
paramètres
existants:

- SO2
- NO2

Station
De BAB EL
OUED

les
paramètres
existants:

- SO2
- NO2
- O3
- PM10

Application Samasafia

- **Paramètres considérés**
- CO04 : Monoxyde de carbone
- HT04 : Hydrocarbure totaux
- NO04 : Monoxyde d’azote
- N2O4 : Dioxyde d’azote
- NX04 : Oxyde d’azote
- PS04 : Poussières
- TE04 : Température
- HU04 : Humidité
- V V04:Vitesse du vent
- DV04 : Direction du vent
-

Application Samasafia

La forme des données

A	B	C	D	E	F	G	H	I	J	K
	CO04	HT04	NO04	N204	NX04	PS04	TE04	HU04	VV04	DV04
	ppm	ppm	ug/m3	ug/m3	ug/m3	microg/m3	degres	%	m/s	degre
01/01/2008 01:00	0,0	3394	2	8	8		10,2	76	1,2	
01/01/2008 02:00	0,0	3392	2	8	9		10,6	75	1,3	
01/01/2008 03:00	0,0	3350	2	8	8		10,7	76	1,2	
01/01/2008 04:00	0,0	3361	2	8	8		10,6	76	1,4	
01/01/2008 05:00	0,0	3414	2	7	8		10,6	74	1,7	
01/01/2008 06:00	0,0	3371	2	8	8		10,3	77	2,5	
01/01/2008 07:00	0,0	3348	2	10	10		10,3	76	2,5	
01/01/2008 08:00	0,8	3389	2	8	8		10,0	76	2,5	
01/01/2008 09:00	0,0	3390	2	8	9		10,2	74	2,4	
01/01/2008 10:00	0,3	3525	8	10	15		11,0	71	1,2	
01/01/2008 11:00	0,8	3436	13	15	24		12,7	66	0,6	
01/01/2008 12:00	0,0	3387	4	8	11		12,6	66	1,2	
01/01/2008 13:00	0,0	3358	4	8	10		11,4	70	1,5	
01/01/2008 14:00	0,0	3339	3	10	10		11,2	70	1,4	
01/01/2008 15:00	0,0	3322	3	8	10		12,7	61	1,4	
01/01/2008 16:00	0,0	3314	3	10	10		13,2	58	0,8	
01/01/2008 17:00	0,0	3385	5	9	12		12,7	59	0,8	
01/01/2008 18:00	1,6	3684	13	12	21		11,3	65	0,6	
01/01/2008 19:00	6,1	4068	26	12	36		10,7	69	1,2	
01/01/2008 20:00	7,0	3930	30	12	38		10,3	70	1,2	
01/01/2008 21:00	5,2	3778	26	12	35		9,8	72	1,0	

Application Samasafia

Station 1 → m1(c1),m1(c2),..........,m1(c10)

Station 2 → m2(c1),m2(c2),..........,m2(c10)

Station 3 → m3(c1),m3(c2),..........,m3(c10)

Station 4 → m4(c1),m4(c2),..........,m4(c10)

Ci c’est la classe qualitative de la qualité de l’aire ordonnée selon l'ordre suivant : excellent , très bon , bon , assez bon, moyen, médiocre , très médiocre ,mauvais , très mauvais ,exécrable .

L’exemple de combinaison entre 2 station :

	m1(c1) 0.25	m1(c2) 0.25	m1(c3) 0.25	m1(c10) 0.25
m2(c1) 0.30	C1 0.25*0.30	0.25*0.30	0.25*0.30	0.25*0.30
m2(c2) 0.20	0.20*0.25	C2 0.20*0.25	0.20*0.25	0.20*0.25
m2(c3) 0.10	0.10*0.25	0.10*0.25	C3 0.10*0.25	0.10*0.25
m2(c10) 0.4	0.4*0.25	0.4*0.25	0.4*0.25	C10 0.4*0.25

Application Samasafia

Résultat de la combinaison orthogonal

m12(C1)	m12(C2)	m12(C3)	m12(C1∩ C2)	m12(C1∩ C3)	m12(C2 ∩C3)	m12(C1∩C2∩ C3)
0.68	0.12	0.05	0.03	0.02	0.05	0.05

la décision sera fait selon la plus grande valeur

8- Conclusion

Les avantages de la méthode des fonctions de croyances :

- Une gestion des imprécision et incertitudes relatives aux différentes sources
- Une gestion directe des valeurs manquantes.
- Une robustesse vis-à-vis des mesures aberrantes.
- la richesse de l'information disponible en sortie.