



HAL
open science

Modèle des Croyances Transférables : Représentation des connaissances, Fusion d'informations, Décision

Emmanuel Ramasso, Michèle Rombaut, Denis Pellerin

► **To cite this version:**

Emmanuel Ramasso, Michèle Rombaut, Denis Pellerin. Modèle des Croyances Transférables : Représentation des connaissances, Fusion d'informations, Décision. 2007. <hal-00283144>

HAL Id: hal-00283144

<https://hal.science/hal-00283144v1>

Submitted on 29 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

RAPPORT INTERNE - TECHNICAL REPORT



Modèle des Croyances Transférables : Représentation des connaissances, Fusion d'informations, Décision

Transferable Belief Model : Knowledge representation, Information fusion, Decision

Emmanuel Ramasso, Michèle Rombaut, Denis Pellerin

Emails : e.ramasso@yahoo.fr, rombaut@gipsa-lab.inpg.fr, pellerin@gipsa-lab.inpg.fr

GIPSA-lab, Département Images et Signal

46, avenue Félix Viallet

38031 GRENOBLE, France

Ce document propose une introduction au Modèle des Croyances Transférable (TBM : Transferable Belief Model) particulièrement adapté à la modélisation des connaissances sur l'état d'un système, à la fusion des informations concernant ce système et à la prise de décision sur son état. Il est extrait de la thèse de E. Ramasso intitulée *Reconnaissance de séquences d'états par le Modèle des Croyances Transférables - Application à l'analyse de vidéos d'athlétisme*¹, plus précisément il correspond au chapitre 2 du manuscrit.

Les principes et mécanismes de base décrits dans ce document sont illustrés d'exemples simples facilitant la compréhension.

¹Thèse soutenue le 5 décembre 2007 à l'Université Joseph Fourier de Grenoble

Table des matières

1 Outils de base	4
1.1 Présentation des concepts sur un cas d'école : le lancé de dé	4
1.2 Formalisme des fonctions de croyance dans le TBM	6
1.3 Le principe du minimum d'information	10
1.4 Règles de combinaisons de base	11
1.5 Prise de décision à partir de fonctions de croyance	13
1.6 Gestion du conflit	14
2 Opérations avancées	15
2.1 Le Théorème de Bayes Généralisé	16
2.2 Opérations sur les cadres de discernement	23
2.3 Déconditionnement	27
2.4 Le PMI pour transformer une distribution de probabilités en BBA	31
2.5 Combinaison de sources non distinctes	32
2.6 Affaiblissement généralisé dit <i>contextuel</i>	35

Introduction

Le *Modèle des Croyances Transférables* (TBM : *Transferable Belief Model*) est un cadre formel générique développé par Ph. Smets [SK94] pour la représentation et la combinaison des connaissances. Le TBM est basé sur la définition de fonctions de croyance fournies par des sources d'information pouvant être complémentaires, redondantes et éventuellement non-indépendantes. Il propose un ensemble d'opérateurs permettant de combiner ces fonctions. Il est donc naturellement employé dans le cadre de la fusion d'informations pour améliorer l'analyse et l'interprétation de données issues de sources d'informations multiples.

L'un des points fondamentaux qui caractérisent le TBM est la différenciation des niveaux *Représentation des connaissances* et *Décision*. Cette différenciation est beaucoup moins prépondérante pour d'autres approches, particulièrement pour le modèle probabiliste pour lequel la *décision* est souvent le seul objectif visé. Les mécanismes de raisonnement du TBM sont donc regroupés en deux niveaux comme illustré sur la figure 1 :

- Le niveau *credal* (du latin « credo » signifiant « je crois ») siège de la représentation des connaissances (partie statique), des combinaisons et du raisonnement sur ces connaissances (partie dynamique),
- Le niveau *pignistique* (du latin « pignus » signifiant « un pari ») impliquant une prise de décision en prenant éventuellement en compte le risque et/ou le gain associés à cette décision.

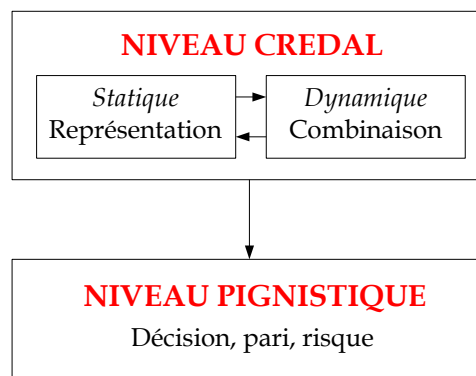


FIG. 1 : Représentation abstraite des mécanismes dans le *Modèle des Croyances Transférables*.

Ce document est divisé en deux chapitres : le premier est consacré à la description des outils de base du TBM et le second est dédié à des opérateurs permettant de régler des problèmes de fusion complexes.

1 Outils de base

Ce chapitre répond aux questions suivantes (Fig. 2) :

- Comment *représenter* la connaissance d’une source d’information sous forme de fonctions de croyance et quels avantages peut-on tirer d’une telle représentation ?
- Comment *combiner* plusieurs sources de fonctions de croyance afin de résumer l’information et d’améliorer la prise de décision ?
- Comment *prendre une décision* à partir de fonctions de croyance ?

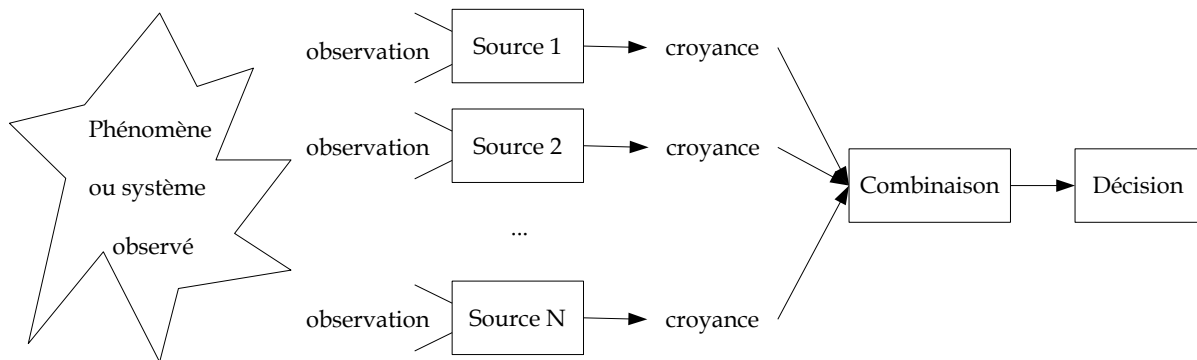


FIG. 2 : Fusion d’information : des sources de croyance émettent une opinion pondérée sur l’état du système observé. Les différentes sources sont combinées afin de synthétiser la compréhension du système et améliorer la décision sur l’état du système après combinaison.

Tout d’abord, nous présentons les *concepts* importants à partir d’un exemple simple : le lancé de dé. Puis, nous décrivons le *formalisme mathématique* des fonctions de croyance ainsi que les *outils de base* permettant de *combiner* ces fonctions et de prendre des décisions en rapport avec leurs valeurs.

1.1 Présentation des concepts sur un cas d’école : le lancé de dé

Nous considérons un dé à six faces (numérotées 1, 2 . . . 6) lancé par un joueur. Les chiffres possibles pouvant être tirés par le dé sont appelés *hypothèses* et l’ensemble des hypothèses forment le *cadre de discernement* généralement noté Ω , par exemple $\Omega = \{F_1, F_2, F_3, F_4, F_5, F_6\}$ où F_i est la face montrant le chiffre i . Les hypothèses sont supposées *exclusives* signifiant que deux hypothèses ne peuvent être vraies simultanément. L’exclusivité se traduit par une intersection nulle entre les hypothèses, par exemple $\{F_1\} \cap \{F_3\} = \emptyset$.

La face tirée à l’aide du dé est en partie cachée et seul le point central est visible. Le joueur émet alors une opinion pondérée sur le chiffre du dé ce qui fait de lui une *source* d’information. Au vu de la connaissance disponible après le lancé du dé, le joueur affirme que le chiffre est soit F_1 , soit F_3 ou soit F_5 . Comment modéliser cette connaissance ?

Dans le cadre de la *théorie des probabilités*, la réponse est donnée par le principe d’*équiprobabilité* : chaque face se voit affectée une même probabilité de $\frac{1}{6}$ telle que la somme des probabilités soit égale

à 1. La *distribution* de probabilités pour ce jet de dé est alors égale à la distribution uniforme :

$$P^\Omega(F_1) = P^\Omega(F_2) = P^\Omega(F_3) = \frac{1}{3}$$

La probabilité est affectée seulement à des hypothèses singletons.

Dans le cadre de la *théorie des fonctions de croyance*, la connaissance est modélisée par une *distribution de masses de croyance*. Dans l'exemple du dé, la totalité de la masse de croyance, c'est à dire une unité comme en probabilité, est affectée à la *proposition* $\{F_1, F_3, F_5\}$ formée d'une *union d'hypothèses* et la distribution se réduit alors à un élément :

$$m^\Omega(\{F_1, F_3, F_5\}) = 1$$

La masse affectée à la proposition $\{F_1, F_3, F_5\}$ *ne donne aucune information* concernant la croyance sur chacune des hypothèses F_1 , F_3 ou F_5 composant cette proposition. On dit qu'une masse de croyance modélise *explicitement le doute* (ou l'ignorance) sur cet ensemble d'hypothèses. En comparaison, dans le cadre de la théorie des probabilités, la valeur de la probabilité sur une union d'hypothèses est déduite *implicitement* de la probabilité sur les hypothèses singletons par une somme des probabilités à laquelle on retire la partie commune, c'est à dire la probabilité de l'intersection ou encore du couple. On dit qu'une mesure de probabilité est *additive*. Par exemple :

$$P^\Omega(\{F_3, F_5\}) = P^\Omega(F_3) + P^\Omega(F_5) - P^\Omega(\{(F_3, F_5)\})$$

où la probabilité sur l'intersection ($P^\Omega(\{(F_3, F_5)\})$) représente la partie commune des événements « la face est F_3 » et « la face est F_5 » qu'il est nécessaire de retirer pour éviter de compter les événements deux fois. Si les hypothèses sont exclusives, comme c'est le cas tout au long de ce manuscrit, alors :

$$\{F_3\} \cap \{F_5\} = \emptyset \Rightarrow P^\Omega(\{(F_3, F_5)\}) = 0 \Rightarrow P^\Omega(\{F_3, F_5\}) = P^\Omega(F_3) + P^\Omega(F_5)$$

car $P^\Omega(\emptyset) = 0$ par définition de la mesure de probabilité.

Une fonction de masses *n'est pas additive* ce qui signifie que la masse sur une union d'hypothèses *n'est pas égale* à la somme des masses des hypothèses composant l'union [KW99] et cette propriété permet de représenter explicitement le doute entre des hypothèses.

Outre la représentation du doute, un des avantages des fonctions de croyance est leur généralité. En effet, elles représentent mathématiquement un sur-ensemble des fonctions de probabilités : une distribution de masses où seules les hypothèses singletons ont une masse non nulle peut être interprétée comme une distribution de probabilités. Nous verrons aussi que les fonctions de croyance généralisent mathématiquement les fonctions de possibilités [KW99]. Chacune de ces fonctions (probabilités, possibilités et croyances) possède des caractéristiques [KW99] et permettent de modéliser la connaissance différemment.

La section suivante est dédiée à la présentation du formalisme mathématique des fonctions de croyance dans le cadre du Modèle des Croyances Transférables (TBM). Notons que plusieurs

modèles d'interprétation des fonctions de croyance ont été proposés : le Modèle de Dempster [Dem67, Dem68], le Modèle de Shafer [Sha76], le Modèle des Croyances Transférables [SK94] et le Modèle des Hints [KM95] entre autres, chacun disposant de spécificités propres comme des opérateurs de combinaison ou de décision. Nous nous plaçons dans le cadre axiomatisé et formalisé du Modèle des Croyances Transférables développé par Ph. Smets.

1.2 Formalisme des fonctions de croyance dans le TBM

La figure 2 représente un synopsis de fusion d'information. L'objectif du système est de déterminer l'état *réel* ω_0 d'un système étudié à partir d'observations.

L'état du système est supposé prendre des valeurs *discrètes* appelées *hypothèses* ω_k et l'ensemble des K hypothèses possibles pouvant expliquer l'état du système est appelé *cadre de discernement* (ou *univers de discours*) généralement noté :

$$\Omega = \{\omega_1 \dots \omega_k \dots \omega_K\} = \bigcup_{k=1}^K \{\omega_k\}$$

Le nombre d'hypothèses $|\Omega| = K$ composant Ω est appelé *cardinal* et les hypothèses sont considérées *exclusives*.

L'observation du système est assurée par des *agents*² qui fournissent, entre autres choses, une opinion pondérée représentée par une fonction de croyance (Fig. 2) concernant l'état réel $\omega_0 \subseteq \Omega$ du système. Les agents jouent le rôle de *sources* de croyance.

1.2.1 Définition de la fonction de masse

Les fonctions de croyance peuvent être représentées de différentes manières. La représentation la plus communément utilisée est celle de la *fonction de masses de croyance* notée m^Ω définie par :

$$\begin{aligned} m^\Omega &: 2^\Omega \rightarrow [0, 1] \\ &B \mapsto m^\Omega(B) \end{aligned} \quad (1)$$

où l'espace puissance $2^\Omega = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_3\}, \dots, \{\omega_1, \dots, \omega_K\}\}$ rassemble tous les sous-ensembles possibles formées des hypothèses et unions d'hypothèses de Ω . Notons que la source de croyance S est souvent inscrite en indice de la fonction de masses, par exemple m_S^Ω .

Une proposition $B \subseteq \Omega$, par exemple $B = \{\omega_1, \omega_3\}$, représente explicitement le doute entre les hypothèses composant B et la masse de croyance $m^\Omega(B)$ allouée à B ne donne aucune information sur les hypothèses et sous-ensembles composant B [KW99]. Cette caractéristique provient de la nature non-additive des fonctions de masses et représente, comme nous l'avons vu dans l'exemple du *lancé de dé*, une différence fondamentale avec la théorie des probabilités où la mesure utilisée est additive.

²Terme générique issu de la communauté de l'*Intelligence Artificielle* représentant des capteurs, des humains ...

Une *distribution de masses* de croyance (BBA : *basic belief assignment*) fournie par une source S est un ensemble de masses de croyance concernant des propositions $B \subseteq \Omega$ (ou de manière équivalente $B \in 2^\Omega$) vérifiant :

$$\sum_{B \subseteq \Omega} m_S^\Omega(B) = 1 \quad (2)$$

La masse $m_S^\Omega(B)$ représente la part de croyance de la source S en la proposition « $\omega_0 \in B$ » où ω_0 est l'état réel du système observé. Un élément B dont la masse est non nulle ($m_S^\Omega(B) > 0$) est appelé *élément focal* de la BBA. Un exemple de BBA est donné colonne « quelconque » (colonne 6) du tableau 1. Quelques exemples particuliers de BBA sont à noter :

BBA normale : la masse sur l'ensemble vide est nulle, i.e. $m^\Omega(\emptyset) = 0$. Dans le cas contraire, la BBA est dite *sous-normale*. Le passage d'une BBA *sous-normale* m^Ω à une BBA *normale* m_*^Ω peut être effectué par le processus de *normalisation* défini par :

$$\begin{aligned} m_*^\Omega(B) &= \frac{m^\Omega(B)}{1 - m^\Omega(\emptyset)}, \quad \forall B \subseteq \Omega, B \neq \emptyset \\ m_*^\Omega(\emptyset) &= 0 \end{aligned} \quad (3)$$

BBA catégorique : un seul élément de l'ensemble des propositions est un élément focal, i.e. $m^\Omega(B) = 1, B \subseteq \Omega$ et $m^\Omega(C) = 0, C \subseteq \Omega$ et $C \neq B$.

BBA vide : l'unique élément focal est Ω , i.e. $m^\Omega(\Omega) = 1$ et indique que la source n'a aucune information sur l'état ω_0 (dernière colonne intitulée « vide » du tableau 1).

BBA dogmatique : la proposition Ω n'est pas un élément focal, i.e. $m^\Omega(\Omega) = 0$ (colonne « bayésienne » du tableau 1).

BBA simple : il existe au plus deux éléments focaux dont Ω (la BBA est donc non-dogmatique), i.e. pour $\eta \in [0, 1]$, $m^\Omega(B) = 1 - \eta$ pour $B \subset \Omega$ et $m^\Omega(\Omega) = \eta$. Ces BBA sont généralement notées B^η .

BBA Bayésienne : les éléments focaux sont des singletons et donc la BBA m^Ω est une distribution de probabilités avec $m^\Omega(\{\omega_k\}) > 0, \omega_k \in \Omega$. Un exemple d'une telle BBA est donné dans la 4^e colonne intitulée « Bayésienne » du tableau 1. Dans la littérature, des BBA *quasi-Bayésiennes* sont fréquemment utilisées dont les éléments focaux sont les singletons ainsi que Ω .

BBA consonante : les éléments focaux sont emboîtés, i.e. $B \subseteq C \subseteq D \dots$. Un exemple de BBA consonante est donné dans la 5^e colonne intitulée « consonante » du tableau 1. Ces BBA ont des caractéristiques communes avec les possibilités [Sme00b, Sme00c, DGPS01].

BBA sans conflit interne : elles sont caractérisées par des éléments focaux dont les intersections ne sont pas vides, i.e. $\cap\{\omega \subseteq \Omega / m^\Omega(\omega) > 0\} \neq \emptyset$. Par conséquent ces BBA sont normales et les BBA consonantes sont sans conflit interne.

Lorsque le cadre de discernement est *exhaustif*, cela signifie qu'il contient forcément l'hypothèse permettant de décrire le système observé (sinon le cadre de discernement n'est pas adapté au problème traité). Dans ce cas, la masse sur l'ensemble vide est nulle, i.e. $m^\Omega(\emptyset) = 0$, et la masse est *normale*.

On parle alors de *monde fermé* (le « monde » étant le cadre de discernement), comme c'est le cas dans le modèle de Shafer [Sha76]. Dans le cadre du TBM, le cadre de discernement peut être *non exhaustif* et donc la masse sur l'ensemble vide peut être non nulle (et la masse *sous-normale*). On parle alors de *monde ouvert*. Le monde ouvert permet par exemple de redéfinir le cadre de discernement en ajoutant une hypothèse [DS04] en fonction de la valeur de la masse sur l'ensemble vide. Cette nouvelle hypothèse peut correspondre à une *hypothèse de rejet*. Une masse positive sur l'ensemble vide peut être interprétée de différentes manières et peut poser quelques problèmes comme nous le verrons dans le paragraphe 1.6. Le passage du monde ouvert au monde fermé consiste à *redistribuer* la masse de l'ensemble vide sur les autres sous-ensembles de l'espace puissance 2^Ω . Plusieurs *règles de redistribution* existent [LCV02] dont la *normalisation* présentée précédemment.

$B \subseteq \Omega$	Précise	Imprécise	Bayésienne	Consonante	Quelconque	Vide
\emptyset	0	0	0	0	0.1	0
$\{\omega_1\}$	0	0	0.3	0.35	0.07	0
$\{\omega_2\}$	0	0	0.6	0	0.12	0
$\{\omega_1, \omega_2\}$	0	0	0	0.2	0.22	0
$\{\omega_3\}$	1	0	0.1	0	0.31	0
$\{\omega_1, \omega_3\}$	0	1	0	0	0.08	0
$\{\omega_2, \omega_3\}$	0	0	0	0	0.01	0
$\{\omega_1, \omega_2, \omega_3\}$	0	0	0	0.45	0.09	1

TAB. 1 : Exemple de BBA m^Ω exprimée sur un cadre de discernement $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Le cas crédibiliste (fonction de masses de croyance quelconque) permet de couvrir tous les cas (6^e colonne), i.e. de la connaissance précise et certaine (colonne 2) à l'ignorance totale (colonne 7). Inspiré de [VS05].

1.2.2 Transformations de la fonction de masses

A partir de la fonction de masses m^Ω , d'autres fonctions en *relation biunivoque* peuvent être déduites : elles représentent la *même information* mais sous une forme différente. Ces fonctions ont diverses interprétations mais en ce qui nous concerne, nous les utiliserons essentiellement pour simplifier les calculs de combinaisons. Elles sont toutes définies de 2^Ω vers $[0, 1]$ mais la somme des éléments focaux n'est pas forcément unitaire. Les fonctions élémentaires que nous utiliserons sont les suivantes :

Plausibilité pl : la plausibilité d'un élément B est la part maximale de croyance qui pourrait soutenir B :

$$pl^\Omega(B) = \sum_{C \cap B \neq \emptyset} m^\Omega(C), \quad \forall B \subseteq \Omega \quad (4a)$$

$$m^\Omega(B) = \sum_{C \subseteq B} (-1)^{|B|-|C|+1} pl^\Omega(\overline{C}), \quad \forall B \subseteq \Omega \quad (4b)$$

Crédibilité (ou croyance) bel : la crédibilité d'un élément B est la part de croyance spécifiquement (sans le conflit) allouée à B :

$$bel^\Omega(B) = \sum_{\emptyset \neq C \subseteq B} m^\Omega(C), \quad \forall B \subseteq \Omega \quad (5a)$$

$$m^\Omega(B) = \sum_{C \subseteq B} (-1)^{|B|-|C|} bel^\Omega(C), \quad \forall B \subseteq \Omega \quad (5b)$$

Implicabilité b : l'implicabilité d'un élément B est la somme des masses allouées aux sous-ensembles de B tels que leur véracité implique la véracité de B :

$$b^\Omega(B) = \sum_{C \subseteq B} m^\Omega(C) = bel^\Omega(B) + m^\Omega(\emptyset), \quad \forall B \subseteq \Omega \quad (6a)$$

$$m^\Omega(B) = \sum_{C \subseteq B} (-1)^{|B|-|C|} b^\Omega(C), \quad \forall B \subseteq \Omega \quad (6b)$$

Communalité q : la communalité d'un élément B représente la somme des masses allouées aux sur-ensembles de B et donc qui ont B en commun :

$$q^\Omega(B) = \sum_{C \supseteq B} m^\Omega(C), \quad \forall B \subseteq \Omega \quad (7a)$$

$$m^\Omega(B) = \sum_{B \subseteq C} (-1)^{|C|-|B|} q^\Omega(C), \quad \forall B \subseteq \Omega \quad (7b)$$

Le tableau 2 donne une indication sur l'utilisation de ces fonctions dans le cadre de cette thèse. Un résumé de ces transformées est donné dans [Sme02a] et l'outil graphique TBMLab³ contient quelques procédures Matlab[®] pour les conversions.

Sigle	Nom	Utilité
m	Masse	Représentation de la connaissance
pl	Plausibilité (Eq. 4a)	Théorème de Bayes Généralisé, vraisemblances
bel	Crédibilité (Eq. 5a)	Représentation de la connaissance
b	Implicabilité (Eq. 6a)	Règle de combinaison disjonctive (DRC)
q	Communalité (Eq. 7a)	Règle de combinaison conjonctive (CRC)

TAB. 2 : Utilisation des fonctions de croyance dans le cadre de nos travaux.

Exemple 1 Le tableau 3 présente le résultat de la transformation de la BBA donnée de la colonne « quelconque » du tableau 1 en plausibilité (pl), crédibilité (bel), implicabilité (b) et communalité (q). Il est important de noter que le calcul de la crédibilité bel nécessite la normalisation préalable (Eq. 3) de la distribution de masses pour éliminer la masse sur l'ensemble vide.

³Disponible à l'adresse <http://iridia.ulb.ac.be/~psmets/>.

	m	pl	bel	b	q
\emptyset	0.1	0	0	0.1	1
$\{\omega_1\}$	0.07	0.46	0.08	0.17	0.46
$\{\omega_2\}$	0.12	0.44	0.13	0.22	0.44
$\{\omega_1, \omega_2\}$	0.22	0.59	0.45	0.51	0.31
$\{\omega_3\}$	0.31	0.49	0.34	0.41	0.49
$\{\omega_1, \omega_3\}$	0.08	0.78	0.51	0.56	0.17
$\{\omega_2, \omega_3\}$	0.01	0.83	0.48	0.54	0.1
$\{\omega_1, \omega_2, \omega_3\}$	0.09	0.9	1	1	0.09

TAB. 3 : Résultat de la transformation de la BBA donnée dans la 6^e colonne du tableau 1 en plausibilité (pl), crédibilité (bel), implicabilité (b) et communalité (q).

1.3 Le principe du minimum d'information

Le principe du minimum d'information (PMI) est au cœur d'un grand nombre de mécanismes développés dans le TBM. L'idée est la suivante : lorsque l'on doit choisir une BBA parmi un ensemble de BBAs possibles, le PMI impose de choisir celle qui est la *moins engagée* ou, autrement dit, la *moins informative*.

Une approche *quantitative* consiste à calculer le *taux d'incertitude* des BBA et de choisir celle qui maximise l'incertitude minimisant ainsi le parti pris. Un grand nombre de méthodes ont été proposées pour mesurer le taux d'incertitude à partir d'informations imprécises [KW99, Har99] (voir aussi la mesure proposée dans [DM04] dans le cadre du *clustering*). La plus répandue des mesures est la *non-spécificité* [DP86b] basée sur la cardinalité des éléments focaux pondérée par la valeur des masses :

$$N(m) = \sum_{\emptyset \neq A \subseteq \Omega} m(A) \cdot \log_2(|A|) \quad (8)$$

Cette mesure est bornée sur $[0, \log_2(|\Omega|)]$.

Plus la mesure de non-spécificité est grande moins la BBA est informative. Par exemple, la BBA vide est totalement non-spécifique tandis qu'une BBA précise (telle qu'une BBA Bayésienne) est totalement spécifique. Le PMI basé sur la non-spécificité maximise donc la valeur de ce critère jouant un rôle similaire au maximum d'entropie en théorie de l'information. Il est par ailleurs possible d'utiliser des relations d'ordre entre des fonctions de croyance (approche *qualitative* du PMI) [DP86a, Yag87, Sme00c, Sme00b].

Exemple 2 Les BBAs données dans les colonnes intitulées « Bayésienne », « consonante » et « quelconque » (4^e, 5^e et 6^e colonnes) du tableau 1 ont pour non-spécificité $N(m_{4e}) = 0$, $N(m_{5e}) \approx 0.91$ et $N(m_{6e}) \approx 0.45$ respectivement. La BBA consonante est ici la moins informative. La non-spécificité d'une BBA Bayésienne est nulle car les éléments focaux sont des singletons ce qui traduit un parti pris total.

1.4 Règles de combinaisons de base

La figure 2 illustre un cas de *fusion d'information* qui consiste à combiner des informations hétérogènes issues de plusieurs sources afin d'améliorer la prise de décision [Blo05]. Plusieurs opérateurs de fusion de BBA ont été développés dans le cadre du TBM. Les règles de base présentées dans ce paragraphe sont basées sur l'hypothèse que les sources de croyance sont *distinctes* [Sme93, YSM01, YSM02a, YSM02b], concept proche de l'*indépendance* en statistique.

1.4.1 La CRC pour combiner des sources distinctes et fiables

Soient deux BBA m_1^Ω et m_2^Ω distinctes et définies sur le même cadre de discernement Ω . Les sources sont supposées *fiables* [Sha76]. La *règle de combinaison conjonctive* (CRC : *conjunctive rule of combination*), notée \odot , des BBAs m_1^Ω et m_2^Ω , ou de leur communalité associée q_1^Ω et q_2^Ω , est définie $\forall B, C, D \subseteq \Omega$ par les équations suivantes :

$$\begin{aligned} m_{1 \odot 2}^\Omega(D) &= (m_1^\Omega \odot m_2^\Omega)(D) = \sum_{B \cap C = D} m_1^\Omega(B) \cdot m_2^\Omega(C) \\ q_{1 \odot 2}^\Omega(D) &= (q_1^\Omega \odot q_2^\Omega)(D) = q_1^\Omega(D) \cdot q_2^\Omega(D) \end{aligned} \quad (9)$$

L'intersection intervenant dans la CRC joue un effet de *spécialisation* (voir les références et définitions de [Sme02a]) en transférant la masse sur des sous-ensembles de cardinalité plus faible. Ainsi, lorsque l'une des BBA est Bayésienne alors le résultat est une BBA Bayésienne. Notons que l'utilisation des communalités réduit la CRC à un produit membre à membre simplifiant ainsi les calculs.

Les propriétés notables de la règle CRC sont les suivantes :

La CRC n'est pas idempotente : la CRC est associative et commutative mais elle n'est pas idempotente ce qui signifie que $(m_1^\Omega \odot m_1^\Omega) \neq m_1^\Omega$ et par conséquent il n'est pas possible de combiner des sources dépendantes ou en interaction. Dans la section 2, nous présenterons un opérateur récent proposé par Denœux [Den07] qui est associatif, commutatif et idempotent. Notons que l'élément neutre est la BBA vide et que $m^\Omega(\emptyset) = 1$ est l'élément absorbant.

La CRC souligne le conflit : lorsque l'intersection des éléments B et C est vide ($B \cap C = \emptyset$), les sources sont conflictuelles. Il est nécessaire d'identifier les raisons de ce conflit car l'ensemble vide est absorbant par l'intersection. Des règles de redistribution ont été proposées [LCV02] pour annuler son effet mais il est généralement préférable d'analyser ses origines (sujet de § 1.6).

La BBA issue d'une CRC peut être sous-normale : la masse sur l'ensemble vide peut être non nulle [Sme92] si les sources sont conflictuelles. Si la normalisation est effectuée alors la CRC revient à la *règle orthogonale* de Dempster [Dem67, Dem68] (notée \oplus) utilisée dans le modèle de Shafer [Sha76, Sme94a]. La CRC normalisée est définie par :

$$m_{1 \oplus 2}^\Omega(D) = (m_1^\Omega \oplus m_2^\Omega)(D) = \begin{cases} \frac{m_{1 \odot 2}^\Omega(D)}{1 - m_{1 \odot 2}^\Omega(\emptyset)}, & \forall D \subseteq \Omega \text{ si } D \neq \emptyset \\ 0 & \text{si } D = \emptyset \end{cases} \quad (10)$$

La CRC est une opération linéaire : elle peut donc être décrite sous forme matricielle [Sme02a].

Cette notation permet de mettre en évidence l'idée de *transfert des masses* sous-jacent au Modèle des Croyances *Transférables* (voir [Mer06] pour une application concernant la reconnaissance d'adresses postales).

1.4.2 Le conditionnement pour combiner des informations certaines

La CRC permet de définir l'opération de *conditionnement*. Conditionner une BBA m^Ω par un élément $B \subseteq \Omega$ consiste à restreindre le cadre des propositions possibles 2^Ω à celles ayant une intersection non nulle avec B . Ainsi, lors d'un conditionnement, les masses affectées à $C \subseteq \Omega$ sont transférées sur $C \cap B$. La BBA *conditionnelle* qui en résulte est notée $m^\Omega[B]$ (les crochets traduisent le conditionnement) et ces éléments sont donnés par la règle de *conditionnement de Dempster non normalisée* :

$$m^\Omega[B](C) = \sum_{D \subseteq \overline{B}} m^\Omega(C \cup D), \quad \forall C \subseteq B \quad (11)$$

L'opération de conditionnement est particulièrement adaptée dans les cas où l'on dispose d'une information *certaine* (mais qui peut être imprécise, par exemple la colonne 3 du tableau 1). Le conditionnement revient à effectuer une combinaison conjonctive avec une BBA *catégorique sur B* (i.e. $m^\Omega(B) = 1$) et donc la BBA issue d'un conditionnement est souvent sous-normale (la gestion du conflit est décrite dans § 1.6).

1.4.3 La DRC pour combiner des sources distinctes et non-fiables

La *règle de combinaison disjonctive* (DRC : *disjunctive rule of combination*), notée \oplus , a été proposée [DP86b, Sme93] pour combiner des BBAs dont l'une au moins est fiable, sans pouvoir quantifier la fiabilité ni savoir quelle source est fiable. Elle est définie $\forall D \subseteq \Omega$ par les équations suivantes :

$$\begin{aligned} m_{1 \oplus 2}^\Omega(D) &= (m_1^\Omega \oplus m_2^\Omega)(D) = \sum_{B \cup C = D} m_1^\Omega(B) \cdot m_2^\Omega(C) \\ b_{1 \oplus 2}^\Omega(D) &= (b_1^\Omega \oplus b_2^\Omega)(D) = b_1^\Omega(D) \cdot b_2^\Omega(D) \end{aligned} \quad (12)$$

La DRC est une règle associative, commutative mais non idempotente (nous décrirons un opérateur disjonctif idempotent dans la section 2) dont l'élément neutre est la BBA $m^\Omega(\emptyset) = 1$ et l'élément absorbant la BBA vide. Cette règle est conservatrice car elle transfère les masses sur des sur-ensembles des éléments focaux. Par conséquent, les éléments focaux résultants peuvent être de cardinal élevé. On dit que la BBA résultante est moins spécialisée que les BBA d'origine ou encore que la DRC est un processus de *généralisation* [Sme02a] (le contraire de la spécialisation). Notons que l'utilisation des implicabilités (b) permet de réduire la DRC à un produit membre à membre et que la DRC est une opération linéaire au même titre que la CRC pouvant être décrite sous forme matricielle [Sme02a].

Remarque 1.1 La dualité des règles \oplus et \ominus est mise en valeur dans les règles de De Morgan s'appliquant aux fonctions de croyance [DP86b].

1.4.4 Règle d'affaiblissement *simple* pour intégrer la fiabilité des sources

L'application de la DRC lorsque les sources de BBA ne sont pas fiables peut s'avérer trop conservatrice pouvant mener à une BBA complètement non-informative (BBA vide). Dans certaines applications, il est cependant possible de *quantifier la fiabilité* ce qui permet d'appliquer une règle de combinaison moins conservatrice dans le cas de sources non fiables.

La prise en compte de la fiabilité dans le cadre des fonctions de croyance porte le nom d'*affaiblissement* car le processus consiste à pondérer la masse des éléments d'une BBA. Les premiers travaux sur l'affaiblissement dans le cadre des fonctions de croyance ont été développés par Shafer [Sha76], axiomatisés par Smets [Sme93] et généralisés par Mercier et Denœux [MQD07, Mer06]. L'affaiblissement généralisé a été appelé *affaiblissement contextuel*.

L'affaiblissement « simple » (de Shafer) d'une BBA m^Ω est défini comme la pondération de chaque masse $m^\Omega(B)$ de la distribution par un coefficient $(1 - \alpha)$ appelé *fiabilité* et où $\alpha \in [0, 1]$ est le *taux d'affaiblissement*. Formellement :

$$\begin{aligned} {}^\alpha m^\Omega(B) &= (1 - \alpha) \cdot m^\Omega(B), \quad \forall B \subset \Omega \\ {}^\alpha m^\Omega(\Omega) &= (1 - \alpha) \cdot m^\Omega(\Omega) + \alpha \end{aligned} \quad (13)$$

Par le PMI, le reste de la masse (après pondération) est transféré sur l'élément d'ignorance Ω . Plus la source est fiable (et donc plus le taux d'affaiblissement α est faible), moins la BBA est modifiée alors que, plus la fiabilité diminue, plus la BBA issue de l'affaiblissement tend vers la BBA vide.

1.5 Prise de décision à partir de fonctions de croyance

Les systèmes de fusion d'information, qu'ils soient basés sur la théorie des probabilités, des possibilités ou des fonctions de croyance, ont pour finalité la prise de décision. Prendre une décision consiste à choisir une hypothèse sur un *cadre de pari*, généralement le cadre de discernement Ω . La prise de décision peut être réalisée de façon automatique ou laissée à la responsabilité de l'utilisateur final (par exemple dans le cas de *l'aide au diagnostic* dans le domaine de la Médecine).

1.5.1 Prise de décision à partir des probabilités pignistiques

Dans le cadre du TBM, la phase de décision s'appuie sur la *distribution de probabilités pignistiques* [Sme05b] notée $\text{BetP}\{m^\Omega\}$ obtenue à partir de la BBA m^Ω . La transformée pignistique consiste à répartir de manière équiprobable la masse d'une proposition B sur les hypothèses contenues dans B . Formellement :

$$\begin{aligned} \text{BetP}\{m^\Omega\} &: \Omega \rightarrow [0, 1] \\ \omega_k &\mapsto \text{BetP}\{m^\Omega\}(\omega_k) \end{aligned} \quad (14)$$

où $\text{BetP}\{m^\Omega\}(\omega_k), \forall \omega_k \in \Omega$ est donnée par :

$$\text{BetP}\{m^\Omega\}(\omega_k) = \frac{1}{(1 - m^\Omega(\emptyset))} \sum_{B \subseteq \Omega, \omega_k \in B} \frac{m^\Omega(B)}{|B|} \quad (15)$$

Les différentes propriétés de cette transformation sont analysées dans [Sme05b].

La décision est généralement prise en choisissant l'élément ω_k possédant la plus grande probabilité pignistique :

$$\omega_0 = \operatorname{argmax}_{\omega_k \in \Omega} \mathbf{BetP}\{m^\Omega\}(\omega_k) \quad (16)$$

Exemple 3 La transformée pignistique de la BBA donnée dans la 6^e colonne du tableau 1 fournit la distribution pignistique suivante : $\mathbf{BetP}\{m^\Omega\}(\omega_1) \approx 0.28$, $\mathbf{BetP}\{m^\Omega\}(\omega_2) \approx 0.29$ et $\mathbf{BetP}\{m^\Omega\}(\omega_3) \approx 0.43$. La décision sur l'état du système ω_0 pourrait être $\omega_0 = \omega_3$ signifiant que l'état réel du système est l'état ω_3 .

1.5.2 Utilisation des plausibilités pour la prise de décision

L'agent qui se base sur la transformée pignistique lors de la phase de prise de décision a un comportement rationnel en maximisant l'*utilité espérée*. Alternativement, les *plausibilités* et les *crédibilités* peuvent être utilisées lorsque l'agent a une attitude plutôt *optimiste* ou *pessimiste*.

Dans le cas des plausibilités, on dit que l'agent a une attitude optimiste dans le sens où la plausibilité d'un élément représente la valeur maximale de la masse de croyance sur cet élément qui pourrait être atteinte si des informations supplémentaires parvenaient au système de fusion (ce dernier utilisant la combinaison conjonctive pour intégrer cette nouvelle information). La distribution de probabilités sur les singletons à partir de laquelle une décision peut être prise est alors donnée $\forall \omega_k \in \Omega$ par :

$$\mathbf{PIP}\{m^\Omega\}(\omega_k) = \frac{pl^\Omega(\omega_k)}{\sum_{\omega_k \in \Omega} pl^\Omega(\omega_k)} \quad (17)$$

Notons que la plausibilité d'un singleton $pl^\Omega(\omega_k)$ est égale à sa communalité $q^\Omega(\omega_k)$. De la même manière, un critère adapté à un agent « pessimiste » peut être obtenue en utilisant les crédibilités (la crédibilité d'un singleton $bel^\Omega(\omega_k)$ est simplement égal à sa masse après normalisation de la BBA m^Ω). On peut remarquer que pour tout élément $A \in 2^\Omega$ on a :

$$bel^\Omega(A) \leq \mathbf{BetP}\{m^\Omega\}(A) \leq pl^\Omega(A) \quad (18)$$

et en particulier lorsque les plausibilités et les crédibilités sont calculées sur les singletons.

Remarque 1.2 Le cadre de pari utilisé lors de la décision peut être différent du cadre de discernement Ω de la BBA [Sme94b, Sme02b, Sme00a, Mer06] et le coût d'une décision peut être intégré dans la prise de décision (voir [Mer06] pour un exemple d'application dans le cadre de la reconnaissance d'adresses postales).

1.6 Gestion du conflit

Lors de la combinaison par la CRC de deux BBAs m_1^Ω et m_2^Ω , l'intersection entre les éléments focaux peut être vide et dans ce cas la BBA résultante a une masse non nulle sur l'ensemble vide, i.e. $m_{1 \odot 2}^\Omega(\emptyset) > 0$. Cette valeur est appelée *conflit* et quantifie la discordance entre les sources de croyance. Il est nécessaire d'analyser ce conflit pour deux raisons majeures :

- L'ensemble vide étant absorbant par l'opérateur CRC, la combinaison avec une troisième BBA m_3^Ω ne fera qu'augmenter la valeur du conflit, i.e. $m_{1 \odot 2 \odot 3}^\Omega(\emptyset) \geq m_{1 \odot 2}^\Omega(\emptyset)$.
- Lors de la phase de décision, la fonction de croyance est transformée en distribution de probabilités pignistiques sur les singletons. Or, cette distribution est normalisée par l'opposé du conflit (Eq. 16) mais si le conflit est élevé, quel est le sens de la distribution de probabilités obtenue ?

Il est nécessaire d'analyser les origines du conflit qui peuvent être multiples [Sme05a]. Une des origines est la non exhaustivité du cadre de discernement. Ainsi, la quantité de masse de conflit peut être interprétée comme la croyance sur une nouvelle hypothèse [DS04]. On parle alors de « TBM en monde ouvert ».

Une autre raison pouvant expliquer le conflit est le non respect des hypothèses sous jacentes à l'application de la règle de combinaison conjonctive employée. Par exemple, la CRC (Eq. 9) suppose que les BBA sont distinctes [Sme93, YSM01, YSM02a, YSM02b] or si ce n'est pas le cas il est préférable d'utiliser la règle prudente proposée récemment par Denœux [Den07].

Une troisième origine possible du conflit provient de la nature de la combinaison conjonctive employée et de celle des BBA. Par exemple, la CRC n'est pas idempotente et par conséquent la combinaison d'une BBA avec elle-même modifie la BBA initiale et peut induire du conflit. Ce conflit est alors appelé *conflit interne*. Pour éviter ce phénomène, deux solutions existent. La première consiste à construire des distributions de masses consonantes en s'inspirant des sous-ensembles flous comme dans [HCR05, GCBR05, RPRP07]. La seconde consiste à changer la règle de combinaison conjonctive et à utiliser par exemple une règle minimisant le conflit [Cat03].

Le conflit peut être aussi interprété comme une véritable source d'information par exemple pour regrouper des fonctions de croyance [Sch04, DM04], associer des objets [AS01], détecter de nouveaux objets [DS04], suivre des cibles [RS06] ou détecter des contours dans les images [CCFM04]. Nous l'avons exploité pour détecter des changements d'états et filtrer des croyances évoluant au cours du temps [RRP07c] ainsi que pour reconnaître des séquences d'états [RPR06, RRP07a, RRP07b].

Dans certaines applications, il peut être intéressant de traiter le conflit automatiquement et pour cela il existe des règles dites de *redistribution* [LCV02]. Celle de Yager par exemple transfère le conflit sur le doute total (Ω) tandis que la règle de Dubois et Prade, aussi appelée règle *adaptive*, transfère le conflit sur l'union des propositions impliquées dans le conflit.

2 Opérations avancées

Dans cette section, nous présentons des outils du TBM permettant de traiter des problèmes de représentation des connaissances et de combinaison complexes. Plus précisément, des éléments de réponse aux questions suivantes sont apportés :

- Comment traiter des problèmes de *diagnostic* ou de *prédiction* ?
- Comment combiner des fonctions de croyance dont les *cadres de discernement* sont *différents* ?
- Comment combiner des distributions de *probabilités* avec des fonctions de *croyances* ?
- Comment combiner des fonctions de croyance dont les sources ne sont *pas indépendantes* ?

2.1 Le Théorème de Bayes Généralisé

Nous décrivons tout d’abord le rôle et le principe du Théorème de Bayes sur un exemple concernant la prédiction météorologique (voir aussi [Sme93] pour un exemple concernant le diagnostic médical et [DS04] pour l’identification de cibles). Nous décrivons ensuite le Théorème de Bayes Généralisé (GBT : *Generalized Bayesian Theorem*) et donnons quelques exemples d’application.

2.1.1 Le principe sur un exemple : problème de prédiction du temps

Soit un problème de *prédiction météorologique*. On note $\Omega_j = \{E_j, P_j, N_j\}$ l’espace de discernement concernant le temps au jour j signifiant que le temps est soit ensoleillé (E_j) soit pluvieux (P_j) ou soit nuageux (N_j).

Une *base de connaissance* liant le temps du jour précédent $j - 1$ au temps du jour courant j a été établie suite à des campagnes de mesures. La base est décrite dans le tableau 4 : chaque ligne décrit la croyance sur le temps au jour j conditionnellement au temps du jour précédent $j - 1$. Ces croyances sont définies sur Ω_j conditionnellement aux singletons du cadre de discernement Ω_{j-1} et sont notées $m^{\Omega_j}[E_{j-1}]$, $m^{\Omega_j}[P_{j-1}]$ et $m^{\Omega_j}[N_{j-1}]$ (une ligne par BBA conditionnelle).

	\emptyset_j	$\{E_j\}$	$\{P_j\}$	$\{E_j, P_j\}$	$\{N_j\}$	$\{E_j, N_j\}$	$\{P_j, N_j\}$	Ω_j
$[E_{j-1}]$	0	0.59	0.01	0.03	0.07	0.21	0.01	0.08
$[P_{j-1}]$	0	0.06	0.31	0.02	0.45	0.11	0.02	0.03
$[N_{j-1}]$	0	0.14	0.12	0.04	0.63	0.01	0.05	0.01

TAB. 4 : Base de connaissance liant l’état du temps au jour j connaissant l’état du temps au jour $j - 1$.

A partir de cette base deux problèmes seront traités :

1. *Raisonnement par inférence ascendante* ou *diagnostic* : Disposant d’une BBA m^{Ω_j} traduisant le temps au jour j , comment déterminer le temps au jour précédent $j - 1$, i.e. comment calculer la BBA $m^{\Omega_{j-1}}$? On s’intéresse aux *causes* d’une observation.
2. *Raisonnement par inférence descendante* ou *prédiction* : Disposant d’une BBA m^{Ω_j} traduisant le temps au jour j , comment *prédire* la BBA sur le temps du lendemain, i.e. comment calculer la BBA $m^{\Omega_{j+1}}$? On s’intéresse aux *effets* d’une hypothèse sur une base de connaissance.

Les solutions sont apportées par des mécanismes d’*inférence* dont le plus connu est le théorème de Bayes dans le cadre de la théorie des probabilités. Ce théorème a été généralisé aux fonctions de croyance (GBT) et on parle alors d’*inférence évidentielle*⁴.

⁴Le terme « évidentiel » étant utilisé pour qualifier des mécanismes basés sur les fonctions de croyance.

2.1.2 La solution probabiliste : le théorème de Bayes

On appelle $\Theta = \{\theta_1, \dots, \theta_i, \dots\}$ l'espace des hypothèses et $\Omega = \{\omega_1, \dots, \omega_k, \dots\}$ l'espace des observations⁵. En théorie des probabilités, la relation entre les observations $\omega_k \in \Omega$ et les hypothèses $\theta_i \in \Theta$ est représentée par un jeu de probabilités conditionnelles donnant la distribution de probabilités sur les observations $\omega_k \in \Omega$ sachant chaque hypothèse $\theta_i \in \Theta$, i.e. $P^\Omega(\cdot|\theta_i)$. Les systèmes d'inférence probabilistes décrits dans ce paragraphe sont à la base de nombreux systèmes actuels, par exemple pour le traitement de signaux audio [Rab89] et de vidéo [Mur02].

Raisonnement par inférence ascendante ou diagnostic Suite à l'observation d'un système, une source d'information fournit son *opinion sur les valeurs des hypothèses* θ_i dans Θ (concernant l'état du système par exemple) sous la forme d'une vraisemblance définie sur Ω conditionnellement à l'hypothèse $\theta_i \in \Theta$ et que l'on note $P(\omega_k|\theta_i)$. Il est alors possible de remonter aux causes de cette observation en estimant les distributions de probabilités a posteriori $P^\Theta(\cdot|\omega_k)$ par le *théorème de Bayes* :

$$P(\theta_i|\omega_k) = \frac{P(\theta_i) \cdot P(\omega_k|\theta_i)}{\sum_{\theta_i \in \Theta} P(\theta_i) \cdot P(\omega_k|\theta_i)}, \forall \theta_i \in \Theta \quad (19)$$

ou de façon moins formelle :

$$\text{a posteriori sur } \Theta \propto \begin{array}{c} \text{a priori} \\ \text{sur les hypothèses} \end{array} \times \begin{array}{c} \text{vraisemblance des observations} \\ \text{conditionnelle aux hypothèses} \end{array} \quad (20)$$

Notons que le symbole \propto signifie « proportionnel à », sous-entendu à une constante près assurant que la somme des éléments est unitaire. Ce symbole sera utilisé à plusieurs reprises dans ce manuscrit.

Exemple 4 Problème de diagnostic météorologique : *on cherche à connaître les causes du temps au jour courant j disposant d'une base de connaissance liant le temps au jour j et $j - 1$. Autrement dit, on souhaite déterminer le temps du jour précédent $j - 1$ pouvant avoir causé le temps au jour courant j . Pour cela, on cherche à estimer les distributions conditionnelles $P^{\Omega_{j-1}}(\cdot|\omega_j)$ en utilisant le théorème de Bayes. Notons que $\Omega_j (\equiv \Omega)$ est l'espace des observations et que $\Omega_{j-1} (\equiv \Theta)$ est celui des hypothèses.*

Comme on ne dispose initialement que d'une base de connaissance sous la forme de fonctions de croyance (tableau 4) et que l'on emploie des outils probabilistes, il est nécessaire de transformer les BBA conditionnelles $m^{\Omega_j}[\omega_{j-1}]$ (Tab. 4) en probabilités conditionnelles $P^{\Omega_j}(\cdot|\omega_{j-1})$ en utilisant la transformation pignistique (Eq. 16). Les résultats sont donnés dans le tableau 5. Ne disposant d'aucune connaissance a priori sur le temps au jour précédent $j - 1$, la distribution de probabilités a priori est donc uniforme, i.e. $P(\omega_{j-1}) = 1/3, \forall \omega_{j-1} \in \Omega_{j-1}$. On dispose ainsi des éléments pour appliquer le théorème de Bayes.

Les distributions a posteriori obtenues par le théorème de Bayes figurent dans le tableau 7a. Elles ont logiquement les mêmes valeurs que les vraisemblances du tableau 5 (en inversant les conditionnements) : n'ayant aucune connaissance a priori sur le temps à $j - 1$, la base de connaissance n'est donc

⁵Nous préférons utiliser Θ et Ω au lieu de Ω_{j-1} et Ω_j pour définir les outils dans un cadre général.

	$\{E_j\}$	$\{P_j\}$	$\{N_j\}$
$[E_{j-1}]$	0.74	0.05	0.21
$[P_{j-1}]$	0.13	0.34	0.53
$[N_{j-1}]$	0.17	0.17	0.66

TAB. 5 : Base de connaissance probabiliste $P^{\Omega_j}(\cdot|\omega_{j-1})$ issue de la transformation pignistique des distributions du tableau 4.

pas modifiée a posteriori. Notons que les valeurs en gras représentent les maximums de probabilités pour chaque conditionnement.

Raisonnement par inférence descendante ou prédiction Disposant d'une base de connaissances $P^{\Omega}(\cdot|\theta_i)$ donnant la probabilité d'apparition d'une observation conditionnellement aux hypothèses, il est possible d'estimer la distribution sur les observations $P(\omega_k)$ suite à une hypothèse émise $P(\theta_i)$ à l'aide de l'équation de conditionnement :

$$P(\omega_k) = \sum_{\theta_i \in \Theta} P(\theta_i) \cdot P(\omega_k|\theta_i) \quad (21)$$

Cette équation permet d'estimer les effets d'une hypothèse sur les observations et elle est classiquement utilisée pour les problèmes de prédiction [Mur02].

Exemple 5 Problème de prédiction météorologique : on souhaite déterminer le temps au jour $j + 1$, i.e. la distribution de probabilités $P^{\Omega_{j+1}}$. On suppose dans un premier temps n'avoir aucune information a priori sur j et on fait donc appel au principe d'équiprobabilité permettant de traduire cette méconnaissance par une distribution uniforme : $P^{\Omega_j}(E_j) = P^{\Omega_j}(P_j) = P^{\Omega_j}(N_j) = 1/3$. On utilise la base de connaissance probabiliste du tableau 5 donnant la relation entre le temps du jour $j - 1$ et celui du lendemain j . On dispose donc des éléments pour appliquer l'équation 21 générant une prédiction probabiliste dont le résultat est donné à la première ligne de l'équation 22 :

$$P^{\Omega_{j+1}}(E_{j+1}) = 0.35 \quad P^{\Omega_{j+1}}(P_{j+1}) = 0.19 \quad P^{\Omega_{j+1}}(N_{j+1}) = 0.46 \quad (22)$$

par exemple $P^{\Omega_{j+1}}(E_{j+1}) = \sum_{\omega_j \in \Omega_j} P(\omega_j) \cdot P(\omega_{j+1}|\omega_j) = 1/3 \cdot 0.74 + 1/3 \cdot 0.13 + 1/3 \cdot 0.17 \approx 0.35$. Cette distribution montre que le temps prédit est probablement nuageux.

On suppose à présent qu'un a priori sur le temps à j existe : on sait que le temps n'est pas ensoleillé. En terme de probabilités, on dispose d'une distribution a priori de la forme $P^{\Omega_j}(\{P_j\}) = P^{\Omega_j}(\{N_j\}) = 1/2$ et on obtient :

$$P^{\Omega_{j+1}}(\{E_{j+1}\}) = 0.15 \quad P^{\Omega_{j+1}}(\{P_{j+1}\}) = 0.25 \quad P^{\Omega_{j+1}}(\{N_{j+1}\}) = 0.60 \quad (23)$$

2.1.3 La solution TBM : le Théorème de Bayes Généralisé

Smets a proposé une généralisation du théorème de Bayes dans le cadre du TBM qu'il a nommé le Théorème de Bayes Généralisé (GBT) [Sme93] (voir aussi les travaux d'Appriou [App91]). Par rapport à sa version probabiliste, le GBT est capable de manipuler des fonctions de croyance.

Outre le fait que les fonctions de croyance permettent une modélisation plus fine de la connaissance et qu'elles généralisent les probabilités et les possibilités, elles permettent de modéliser la *méconnaissance* sur l'a priori nécessaire au calcul de la distribution a posteriori.

Le Théorème de Bayes Généralisé A la différence du théorème de Bayes qui génère une distribution de probabilités a posteriori, le GBT génère une distribution de croyances pouvant prendre la forme de plausibilités, de communalités, de masses. . .

Les expressions analytiques des distributions générées par le GBT ont été dérivées par Smets [Sme93]. Parmi les plus utilisées, trois sont basées sur les plausibilités $pl^\Omega[\theta_i]$ définies conditionnellement aux *singletons* $\theta_i \in \Theta$ réduisant ainsi le coût de stockage en mémoire lors de l'implémentation⁶. Les expressions analytiques sont basées sur l'hypothèse que l'a priori sur Θ est vide. Le GBT génère alors les plausibilités, communalités et masses a posteriori suivantes :

$$pl^\Theta[\omega](\theta) = 1 - \prod_{\theta_i \in \theta} (1 - pl^\Omega[\theta_i](\omega)) \quad (24a)$$

$$q^\Theta[\omega](\theta) = \prod_{\theta_i \in \theta} pl^\Omega[\theta_i](\omega) \quad (24b)$$

$$m^\Theta[\omega](\theta) = \prod_{\theta_i \in \theta} pl^\Omega[\theta_i](\omega) \cdot \prod_{\theta_i \in \bar{\theta}} (1 - pl^\Omega[\theta_i](\omega)) \quad (24c)$$

Exemple 6 Problème de diagnostic météorologique : *Nous appliquons le GBT pour déterminer la BBA sur le temps à $j - 1$ conditionnellement au temps à j sans a priori sur Ω_{j-1} . Les résultats du GBT sont obtenus à partir des données du tableau 4 en trois étapes :*

1. Transformation des BBA conditionnelles $m^{\Omega_j}[\omega_{j-1}]$ du tableau 4 en plausibilités conditionnelles $pl^{\Omega_j}[\omega_{j-1}]$ en utilisant l'équation 4a,
2. Application du GBT (Eq. 24c) :

$$m^{\Omega_{j-1}}[\omega_j](\omega) = \prod_{\omega_{j-1} \in \omega} pl^{\Omega_j}[\omega_{j-1}](\omega_j) \cdot \prod_{\omega_{j-1} \in \bar{\omega}} (1 - pl^{\Omega_j}[\omega_{j-1}](\omega_j))$$

Les plausibilités $pl^{\Omega_{j-1}}[B], \forall B \subseteq \Omega_j$ sont explicitées dans le tableau 6.

3. Application de la transformation pignistique après le passage en BBA (Eq. 4b). La distribution de probabilités pignistiques associée à la BBA $m^{\Omega_{j-1}}[E_j]$ est donnée à la première ligne du tableau 7b et montre que le temps au jour $j - 1$ était probablement ensoleillé avec une pignistique de 0.77.

⁶Le coût mémoire d'une base de connaissance probabiliste $P^\Omega(\cdot|\theta_i)$ est de $|\Omega| \times |\Theta|$ tandis que pour une base crédibiliste, elle est de $2^{|\Omega|} \times |\Theta|$ pour les plausibilités et $2^{|\Omega|} \times 2^{|\Theta|}$ pour les masses.

	\emptyset_{j-1}	$\{E_{j-1}\}$	$\{P_{j-1}\}$	$\{E_{j-1}, P_{j-1}\}$	$\{N_{j-1}\}$	$\{E_{j-1}, N_{j-1}\}$	$\{P_{j-1}, N_{j-1}\}$	Ω_{j-1}
$[\emptyset_j]$	0	0	0	0	0	0	0	0
$[E_j]$	0	0.910	0.220	0.930	0.200	0.928	0.376	0.944
$[P_j]$	0	0.130	0.380	0.461	0.220	0.321	0.516	0.579
$[\{E_j, P_j\}]$	0	0.930	0.550	0.969	0.370	0.956	0.717	0.980
$[N_j]$	0	0.370	0.610	0.754	0.700	0.811	0.883	0.926
$[\{E_j, N_j\}]$	0	0.990	0.690	0.997	0.880	0.999	0.963	1
$[\{P_j, N_j\}]$	0	0.410	0.940	0.965	0.860	0.917	0.992	0.995
$[\Omega_j]$	0	1	1	1	1	1	1	1

TAB. 6 : Les plausibilités a posteriori $pl^{\Omega_{j-1}}[\omega_j]$ définies sur Ω_{j-1} conditionnellement aux hypothèses et sous-ensembles d'hypothèses de Ω_j obtenues par le GBT.

La distribution de probabilités pignistiques conditionnelles obtenues par le GBT (Tab. 7b) est différente de la distribution de probabilités conditionnelles obtenue par le théorème de Bayes (Tab. 7a). En particulier, les probabilités maximales pour chaque conditionnement sont différentes et surtout les valeurs maximales ne concernent pas les mêmes singletons⁷.

	$\{E_{j-1}\}$	$\{P_{j-1}\}$	$\{N_{j-1}\}$
$[E_j]$	0.74	0.05	0.21
$[P_j]$	0.13	0.34	0.53
$[N_j]$	0.17	0.17	0.66

a) Probabilité a posteriori après application du théorème de Bayes avec a priori uniforme.

	$\{E_{j-1}\}$	$\{P_{j-1}\}$	$\{N_{j-1}\}$
$[E_j]$	0.77	0.12	0.11
$[P_j]$	0.16	0.55	0.29
$[N_j]$	0.19	0.36	0.44

b) Probabilités pignistiques a posteriori après application du GBT avec a priori vide.

TAB. 7 : Distributions de probabilités $P^{\Omega_{j-1}}(\cdot|\omega_j)$ au jour $j - 1$ conditionnellement aux éléments de Ω_j .

La différence vient du fait que la base de connaissance a été transformée en probabilités et par conséquent, une connaissance a priori forte a été ajoutée.

Dualité GBT – DRC La relation $pl^{\Omega}[\theta_1](\theta_2) = pl^{\Omega}(\theta_1 \cap \theta_2), \forall \theta_1 \subseteq \Theta, \theta_2 \subseteq \Theta$ (avec $pl^{\Omega}(\theta_1) = 1$) [Sha76] permet d'écrire $pl^{\Omega}[\theta](\omega) = pl^{\Omega \times \Theta}(\omega \times \theta)$. Par conséquent, l'égalité suivante est vérifiée :

$$pl^{\Omega}[\theta](\omega) = pl^{\Theta}[\omega](\theta) \quad (25)$$

Le couplage de l'égalité 25 avec l'équation 24a permet d'obtenir une expression de la DRC basée sur les plausibilités [Sme93]. Celle-ci est fréquemment utilisée car elle minimise le stockage mémoire et les coûts de calcul par rapport à la DRC classique (Eq. 12) :

$$pl^{\Omega}[\theta](\omega) = 1 - \prod_{\theta_i \in \Theta} (1 - pl^{\Omega}[\theta_i](\omega)) \quad (26)$$

⁷Notons que, si au lieu de transformer les BBA conditionnelles en pignistique on choisit plutôt de les transformer en plausibilités normalisées **PIP** (Eq. 17), les résultats numériques sont différents mais les maximums concernent les mêmes singletons que dans le cas pignistique.

Prise en compte d'a priori sur les hypothèses et prédiction Lorsqu'un a priori m_o^Θ existe sur Θ (les hypothèses), pouvant être vide ou prendre toute autre forme (cf. exemples de BBA dans le tableau 1), il est pris en compte en utilisant la forme conjonctive suivante [Sme93] :

$$pl^\Omega(\omega) = \sum_{\theta \subseteq \Theta} m_o^\Theta(\theta) \cdot pl^\Omega[\theta](\omega), \forall \omega \subseteq \Omega \quad (27)$$

où $pl^\Omega[\theta](\omega)$ est donnée par l'équation 26. L'équation 27, qui est analogue à l'équation 21 dans le cadre des probabilités, est généralisable aux fonctions de croyance quelconques (fonctions de masses m , communalités $q \dots$) et que l'on note $f^\Omega \in \{m^\Omega, pl^\Omega, bel^\Omega, b^\Omega, q^\Omega\}$. L'expression généralisant l'équation 27 est donc [Sme93] :

$$f^\Omega(\omega) = \sum_{\theta \subseteq \Theta} m_o^\Theta(\theta) \cdot f^\Omega[\theta](\omega), \forall \omega \subseteq \Omega \quad (28)$$

Comme le montre l'équation 28, il est nécessaire de disposer des croyances définies sur Θ conditionnellement aux sous-ensembles $\omega = \bigcup_k \omega_k \subseteq \Omega$. Comme cela a été précisé par Smets [Sme93], on ne dispose généralement que des croyances définies conditionnellement aux singletons car elles sont plus simples à déterminer. Il s'agit alors de calculer les croyances définies conditionnellement aux sous-ensembles d'hypothèses à l'aide d'une règle disjonctive. La DRC basée sur les plausibilités a été décrite précédemment mais il est aussi possible d'utiliser une DRC sur les BBA conditionnelles :

$$m^\Omega[\theta] = \bigoplus_{\theta_i \in \theta} m^\Omega[\theta_i], \forall \theta \subseteq \Theta \quad (29)$$

Par exemple la BBA $m^{\Omega_j}[\{E_{j-1}, P_{j-1}\}]$ est obtenue par l'application de la DRC sur $m^{\Omega_j}[E_{j-1}]$ et $m^{\Omega_j}[P_{j-1}]$ en supposant ces BBA distinctes. Le résultat est donné à la quatrième ligne du tableau 8.

Exemple 7 Problème de prédiction météorologique : *On souhaite prédire le temps au jour $j + 1$ à partir de la base de connaissances et des données a priori. Contrairement à la prédiction probabiliste de l'exemple 5, nous disposons des BBA conditionnelles nécessaires à l'application du GBT (Tab. 4).*

On applique ensuite l'équation 28 (avec $f \equiv m$) pour prédire la BBA recherchée. Pour appliquer cette équation, nous devons déterminer les BBA $m^{\Omega_j}[\omega_{j-1}]$ définies sur Ω_j conditionnellement aux hypothèses et sous-ensembles d'hypothèses de Ω_{j-1} . Pour cela, nous avons deux solutions :

- *Appliquer la DRC avec les plausibilités : ceci requière de transformer les BBA conditionnelles du tableau 4 en plausibilités puis d'appliquer l'équation 26 et enfin de transformer les plausibilités conditionnelles obtenues en BBA (Eq. 4b),*
- *Ou appliquer la DRC définie à l'équation 29 directement sur les BBA conditionnelles du tableau 4.*

Les deux solutions mènent aux BBA du tableau 8.

Dans un premier temps, on suppose que l'a priori sur Ω_j est vide (on ne connaît rien du temps actuel). L'équation 28 est donc appliquée avec $m_o^{\Omega_j}(\Omega_j) = 1$ et en utilisant les BBA conditionnelles

	\emptyset_j	$\{E_j\}$	$\{P_j\}$	$\{E_j, P_j\}$	$\{N_j\}$	$\{E_j, N_j\}$	$\{P_j, N_j\}$	Ω_j
$[\emptyset_{j-1}]$	1	0	0	0	0	0	0	0
$[E_{j-1}]$	0	0.590	0.010	0.030	0.070	0.210	0.010	0.080
$[P_{j-1}]$	0	0.060	0.310	0.020	0.450	0.110	0.020	0.030
$[\{E_{j-1}, P_{j-1}\}]$	0	0.035	0.003	0.207	0.032	0.472	0.036	0.215
$[N_{j-1}]$	0	0.140	0.120	0.040	0.630	0.010	0.050	0.010
$[\{E_{j-1}, N_{j-1}\}]$	0	0.083	0.001	0.105	0.044	0.552	0.027	0.188
$[\{P_{j-1}, N_{j-1}\}]$	0	0.008	0.037	0.071	0.283	0.192	0.303	0.105
$[\Omega_{j-1}]$	0	0.005	0	0.068	0.020	0.396	0.036	0.475

TAB. 8 : Les masses définies sur Ω_j conditionnellement aux hypothèses et sous-ensembles d'hypothèses de Ω_{j-1} obtenues par la DRC.

obtenues précédemment (Tab. 8). La BBA prédite est alors la suivante :

$$\begin{aligned}
m^{\Omega_{j+1}}(\emptyset_{j+1}) &= 0 & m^{\Omega_{j+1}}(\{E_{j+1}\}) &= 0.005 \\
m^{\Omega_{j+1}}(\{P_{j+1}\}) &= 0 & m^{\Omega_{j+1}}(\{E_{j+1}, P_{j+1}\}) &= 0.068 \\
m^{\Omega_{j+1}}(\{N_{j+1}\}) &= 0.02 & m^{\Omega_{j+1}}(\{E_{j+1}, N_{j+1}\}) &= 0.396 \\
m^{\Omega_{j+1}}(\{P_{j+1}, N_{j+1}\}) &= 0.036 & m^{\Omega_{j+1}}(\Omega_j) &= 0.475
\end{aligned} \tag{30}$$

dont la distribution de probabilité pignistique est donnée par :

$$\text{BetP}\{m^{\Omega_{j+1}}\}(E_{j+1}) \approx 0.40 \quad \text{BetP}\{m^{\Omega_{j+1}}\}(P_{j+1}) \approx 0.21 \quad \text{BetP}\{m^{\Omega_{j+1}}\}(N_{j+1}) \approx 0.39 \tag{31}$$

Cette distribution indique que le temps prédit est soit ensoleillé soit nuageux (les valeurs des probabilités sont trop proches pour départager les deux hypothèses). Les valeurs obtenues sont sensiblement différentes du cas probabiliste (Ex. 22) du fait de la modification de la base de connaissance pour appliquer le Théorème de Bayes.

Dans un deuxième temps, on modifie l'a priori : on sait que le temps actuel à j n'est pas ensoleillé. L'équation 28 est appliquée avec $m_o^{\Omega_j}(\{P_j, N_j\}) = 1$ et en utilisant les BBA conditionnelles du tableau 8. La BBA prédite est alors la suivante :

$$\begin{aligned}
m^{\Omega_{j+1}}(\emptyset_{j+1}) &= 0 & m^{\Omega_{j+1}}(\{E_{j+1}\}) &= 0.008 \\
m^{\Omega_{j+1}}(\{P_{j+1}\}) &= 0.037 & m^{\Omega_{j+1}}(\{E_{j+1}, P_{j+1}\}) &= 0.071 \\
m^{\Omega_{j+1}}(\{N_{j+1}\}) &= 0.284 & m^{\Omega_{j+1}}(\{E_{j+1}, N_{j+1}\}) &= 0.192 \\
m^{\Omega_{j+1}}(\{P_{j+1}, N_{j+1}\}) &= 0.303 & m^{\Omega_{j+1}}(\Omega_j) &= 0.105
\end{aligned} \tag{32}$$

dont la distribution de probabilité pignistique est donnée par :

$$\text{BetP}\{m^{\Omega_{j+1}}\}(E_{j+1}) \approx 0.17 \quad \text{BetP}\{m^{\Omega_{j+1}}\}(P_{j+1}) \approx 0.26 \quad \text{BetP}\{m^{\Omega_{j+1}}\}(N_{j+1}) \approx 0.57 \tag{33}$$

Cette distribution indique que le temps du lendemain est probablement nuageux. On peut remarquer que les valeurs sont assez proches du cas probabiliste 23 car on a ajouté un a priori.

GBT et DRC pour la propagation de croyances dans les réseaux évidentiels Le GBT et la DRC sont au cœur des réseaux évidentiels [Sme93, XS94, YSM03] généralisant les réseaux Bayésiens [Mur02]. Un réseau est un modèle graphique dont l'objectif est de représenter les connaissances sous forme de graphes orientés mettant en évidence la causalité entre des variables d'un système étudié.

L'équation 28 est représentée sur la figure 3 sous la forme d'un modèle graphique évidentiel. Cette représentation montre un graphe orienté à deux variables prenant leurs valeurs discrètes sur Θ et Ω . L'arc orienté partant du nœud Θ et allant vers le nœud de Ω signifie que « $\theta \subseteq \Theta$ cause $\omega \subseteq \Omega$ » (sous l'hypothèse θ , on observe ω) et donc il existe un lien de causalité (dépendance) entre les deux variables.

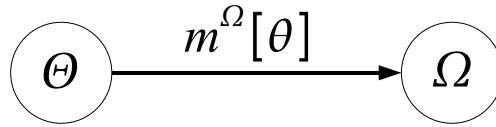


FIG. 3 : Modèle graphique représentant le lien de causalité entre deux variables à valeurs discrètes dans Θ et Ω . Le lien de causalité est traduit par un conditionnement ($m^\Omega[\theta], \forall \theta \subseteq \Theta$).

Si un ou plusieurs nœuds d'un réseau reçoit une BBA (instanciation, par exemple Θ reçoit m^Θ), on peut calculer son influence sur l'ensemble du réseau grâce à des mécanismes de *propagation de croyance*. Pour illustrer ces mécanismes, on considère le réseau très simple de la figure 3 avec deux variables où l'arc orienté, qui représente les liens de causalité, est caractérisé par un jeu de plausibilités conditionnelles $pl^\Omega[\theta_i], \forall \theta_i \in \Theta$:

Propagation forward : si le nœud Θ reçoit une BBA m^Θ , cette dernière est *propagée* vers Ω (en suivant la flèche d'où le qualificatif *forward*) par l'équation 27 qui calcule les plausibilités $pl^\Omega[\theta]$ conditionnelles aux sous-ensembles $\theta \subseteq \Theta$ par la DRC et applique une règle conjonctive pour calculer l'impact sur Ω .

Propagation backward : si le nœud Ω reçoit une BBA m^Ω , cette dernière est *propagée* vers Θ (dans le sens contraire à celui indiqué par la flèche d'où le qualificatif *backward*) par une règle conjonctive similaire à 27 :

$$pl^\Theta(\theta) = \sum_{\omega \subseteq \Omega} m^\Omega(\omega) \cdot \left(1 - \prod_{\omega_i \in \omega} (1 - pl^\Theta[\omega_i](\theta)) \right), \forall \theta \subseteq \Theta \quad (34)$$

où les plausibilités $pl^\Theta[\omega_i]$ sont obtenues par le GBT (Eq. 24a).

Propagation dans les deux directions : si les deux nœuds reçoivent une BBA, il suffit d'appliquer la propagation *forward* avec m^Θ , puis la propagation *backward* avec m^Ω et enfin de combiner les deux résultats avec la CRC.

Dans l'exemple de la prédiction météorologique, un modèle graphique peut être obtenu, similaire à celui de la figure 3, où le nœud Ω_j (le temps actuel) est représenté par Θ et le nœud Ω_{j+1} (le temps du jour suivant) par Ω explicitant le fait que la connaissance du temps au jour j *influe* sur la prédiction du temps à $j + 1$.

2.2 Opérations sur les cadres de discernement

Dans le paragraphe 1.4, nous avons présenté des mécanismes de base de combinaison de fonctions de croyance dont le rôle est de prendre en compte toutes les informations disponibles afin d'améliorer la prise de décision. Ces mécanismes sont basés sur l'hypothèse que les fonctions de croyance sont définies sur le même espace de discernement mais cette contrainte limite les applications pratiques.

Le TBM propose des outils performants et adaptés lorsque les cadres de discernement sont différents. Afin d'illustrer l'utilisation de ces outils, nous considérons le cas de la combinaison de deux fonctions de croyance : la première est m_1^Ω définie sur un premier cadre Ω , la seconde est m_2^Θ définie sur Θ .

Le principe général de ces mécanismes consiste en trois étapes :

1. Construire le cadre de discernement commun,
2. Exprimer les croyances sur le cadre de discernement commun obtenu à l'étape 1. Les mécanismes sont basés sur le Principe du Minimum d'Information (PMI),
3. Combiner les fonctions de croyance obtenues à l'étape 2 par les règles de combinaison classiques.

Les deux mécanismes de base qui permettent le changement d'espace sont le *raffinement*, où l'espace devient plus *détaillé*, et le *grossissement*, où l'espace devient plus *grossier*.

2.2.1 Raffinement et extension vide

Raffinement Un raffinement associe à chaque hypothèse $\omega_k \in \Omega$ une ou plusieurs hypothèses $\theta_i \in \Theta$. L'application $\rho : 2^\Omega \rightarrow 2^\Theta$ est un raffinement de Ω si et seulement si ρ vérifie les contraintes suivantes :

- l'ensemble $\{\rho(\{\omega_k\}), \omega_k \in \Omega\} \subseteq 2^\Theta$ est une partition de Θ ,
- $\forall B \subseteq \Omega, \rho(B) = \bigcup_{\omega_k \in B} \rho(\{\omega_k\})$

L'image d'une union d'hypothèses est l'union des images par ρ de chacune des hypothèses.

La figure 4 présente un exemple de raffinement où le cadre Θ est la version *raffinée* de Ω . Les flèches partant de l'ensemble Ω et allant vers Θ décrivent le raffinement.

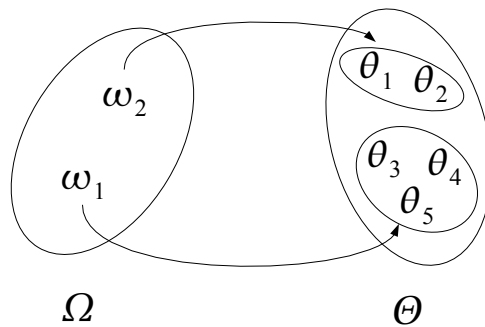


FIG. 4 : Raffinement de cadre de discernement : Θ est un raffinement de Ω .

Extension vide Après le raffinement d'un cadre de discernement Ω en Θ , il s'agit de redéfinir sur Ω la distribution de masses de croyance initialement définie sur Ω . Pour cela, on utilise l'*extension vide* [Sha76, Sme93] qui est un processus conservatif de réallocation des masses de croyance respectant le PMI. L'opérateur associé est noté « \uparrow » et il est défini par :

$$\begin{aligned} 2^\Omega &\rightarrow 2^\Theta \\ m^\Omega(B) &\mapsto m^\Theta(\rho(B)) = m^{\Omega\uparrow\Theta}(\rho(B)) \end{aligned} \quad (35)$$

Exemple 8 Dans l'exemple de la figure 4, la masse $m^\Omega(\omega_2)$ est transférée sur $\{\theta_1, \theta_2\}$ après raffinement. Ce transfert ne privilégie ni θ_1 ni θ_2 justifiant ainsi le terme d'*extension vide* (sous-entendu, l'*a priori* sur $\{\theta_1, \theta_2\}$ est vide). La masse $m^\Omega(\{\omega_1, \omega_2\})$ est quant à elle transférée sur $\rho(\{\omega_1\}) \cup \rho(\{\omega_2\}) = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$. Après raffinement, la BBA m^Θ est définie par :

$$\begin{aligned} m^\Theta(\{\theta_1, \theta_2\}) &= m^\Omega(\{\omega_2\}) & m^\Theta(\{\theta_3, \theta_4, \theta_5\}) &= m^\Omega(\{\omega_1\}) \\ m^\Theta(\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}) &= m^\Omega(\{\omega_1, \omega_2\}) \end{aligned} \quad (36)$$

Première utilisation pratique : définition de BBAs sur un espace produit L'opération d'*extension vide* permet de combiner deux BBAs m_1^Ω et m_2^Θ définies sur des cadres de discernement disjoints. Pour cela, les BBAs doivent être en premier lieu étendues sur un espace commun défini par l'*espace produit* $\Omega \times \Theta = \{(\omega_k, \theta_i), \forall k \in \{1 \dots |\Omega|\}, \forall i \in \{1 \dots |\Theta|\}\}$. Par exemple, m_1^Ω est redéfinie sur $m_1^{\Omega \times \Theta}$ de la manière suivante :

$$m_1^{\Omega\uparrow\Omega \times \Theta}(C) = \begin{cases} m_1^\Omega(B) & \text{si } C = B \times \Theta \\ 0 & \text{sinon} \end{cases} \quad (37)$$

La masse initialement allouée à la proposition $B \subseteq \Omega$ est réallouée à $B \times \Theta$ après *extension vide* minimisant ainsi tout a priori sur Θ (aucune hypothèse de Θ n'est favorisée car la source de croyance fournissant m_1^Ω ne connaît rien de Θ). Notons que la transformation $B \subseteq \Omega \mapsto B \times \Theta$ est appelée *extension cylindrique* de B sur $\Omega \times \Theta$.

La combinaison des deux BBAs m_1^Ω et m_2^Θ est alors donnée par :

$$\begin{aligned} m_{1 \odot 2}^{\Omega \times \Theta}(D) &= (m_1^{\Omega\uparrow\Omega \times \Theta} \odot m_2^{\Theta\uparrow\Omega \times \Theta})(D) \\ &= \begin{cases} m_1^\Omega(B) \cdot m_2^\Theta(C) & \text{si } D = B \times C, B \subseteq \Omega, C \subseteq \Theta \\ 0 & \text{sinon} \end{cases} \end{aligned} \quad (38)$$

Exemple 9 Nous considérons un exemple de reconnaissance de formes où l'on cherche à reconnaître le type d'un panneau de signalisation. Pour cela, on dispose de deux sources de croyance :

- une première source de croyance est capable de distinguer deux types de forme : ronde (R_f) et triangulaire (T_f). Le cadre de discernement est $\Omega = \{R_f, T_f\}$.
- une deuxième source discrimine les couleurs : bleu (B_c), vert (V_c) et magenta (M_c). Le cadre de discernement est $\Theta = \{B_c, V_c, M_c\}$.

On cherche à combiner conjonctivement les deux premières BBA définies par :

$$\begin{aligned}
m_1^\Omega(\{R_f\}) &= 0.3 & m_2^\Theta(\{M_c\}) &= 0.7 \\
m_1^\Omega(\{T_f\}) &= 0.5 & m_2^\Theta(\{B_c, V_c\}) &= 0.1 \\
m_1^\Omega(\{R_f, T_f\}) &= 0.2 & m_2^\Theta(\{B_c, V_c, M_c\}) &= 0.2
\end{aligned} \tag{39}$$

Pour cela, on utilise l'équation 38 : on calcule l'extension vide des BBA sur l'espace produit $\Omega \times \Theta = \{(R_f, B_c), (R_f, V_c), (R_f, M_c), (T_f, B_c), (T_f, V_c), (T_f, M_c)\}$ suivie d'une combinaison CRC. La BBA résultante $m_{1 \otimes 2}^{\Omega \times \Theta}$ comporte 9 éléments focaux :

$$\begin{aligned}
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(T_f, M_c)\}) &= 0.35 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(R_f, M_c)\}) &= 0.21 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(R_f, M_c), (T_f, M_c)\}) &= 0.14 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(T_f, B_c), (T_f, V_c), (T_f, M_c)\}) &= 0.1 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(R_f, B_c), (R_f, V_c), (R_f, M_c)\}) &= 0.06 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(T_f, B_c), (T_f, V_c)\}) &= 0.05 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(R_f, B_c), (R_f, V_c), (R_f, M_c), (T_f, B_c), (T_f, V_c), (T_f, M_c)\}) &= 0.04 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(R_f, B_c), (R_f, V_c)\}) &= 0.03 \\
m_{1 \otimes 2}^{\Omega \times \Theta}(\{(R_f, B_c), (R_f, V_c), (T_f, B_c), (T_f, V_c)\}) &= 0.02
\end{aligned} \tag{40}$$

Par exemple, $m_{1 \otimes 2}^{\Omega \times \Theta}(\{(T_f, M_c)\}) = m_1^\Omega(\{T_f\}) \cdot m_2^\Theta(\{M_c\}) = 0.7 \cdot 0.5 = 0.35$.

Deuxième utilisation pratique : fusion de BBAs avec cadres de discernement hiérarchisés Le concept de *hiérarchie* (ou de catégorie) est utilisé dans de nombreuses applications et permet de représenter les connaissances sous la forme d'arbres de décision. Cette représentation est facilement *interprétable* et possède une signification réelle pour un utilisateur ou un expert. Un exemple concernant les automobiles est proposé par Delmotte et Smets [DS04], l'objet de leur article étant l'identification de cibles aériennes. Dans [Mer06], un système hiérarchique de reconnaissance d'adresses postales est proposé où l'aspect hiérarchique permet, d'une part, de reconnaître soit complètement soit partiellement les adresses et, d'autre part, d'estimer les distributions de masses à partir de matrices de confusion.

2.2.2 Grossissement et marginalisation

Si Θ est un raffinement de Ω on dit que Ω est un *grossissement* de Θ (Fig. 4). Le grossissement est donc l'opération duale du raffinement et permet de réduire le cadre de discernement en le rendant plus grossier. L'application ρ définie dans le cadre du raffinement n'est pas toujours surjective et par conséquent il existe des éléments $\theta \subseteq \Theta$ qui ne sont pas l'image d'éléments de Ω . Par exemple, sur la figure 4, l'élément $\{\theta_1, \theta_3\}$ n'est pas une image par ρ des éléments de Ω . Il existe deux méthodes pour palier ce problème appelées *réductions intérieures* et *extérieures* [Sha76]. Nous encourageons le lecteur intéressé par ces mécanismes à consulter les références [DY02, Quo06] comprenant des exemples détaillés.

L'opération de *marginalisation* est un cas particulier de grossissement lorsque le cadre de discernement avant marginalisation est défini par un espace produit $\Theta \times \Omega$ et que le sous-ensemble d'hypothèses considéré pour le grossissement est un des cadres de l'espace produit. Par exemple, la marginalisation sur Ω d'une BBA initialement définie sur l'espace produit $\Theta \times \Omega$ est donnée par :

$$m^{\Theta \times \Omega \downarrow \Omega}(B) = \sum_{\substack{C \subseteq \Theta \times \Omega \\ \text{Proj}(C \downarrow \Omega) = B}} m^{\Theta \times \Omega}(C), \quad \forall B \subseteq \Omega \quad (41)$$

où $\text{Proj}(C \downarrow \Omega)$ est la projection de C sur Ω . La masse d'un élément $m^\Omega(B)$ est obtenue par la somme des masses des éléments de l'espace produit contribuant à B .

Il n'est généralement pas possible de retrouver la BBA initiale après marginalisation. Si l'on est intéressé par recouvrir une BBA, le Principe du Minimum d'Information (PMI) impose de choisir la BBA la moins informative de toutes les BBA possibles et cette BBA est donnée par l'extension vide (Eq. 38).

Exemple 10 La marginalisation de la BBA $m_{1 \oplus 2}^{\Omega \times \Theta}$ obtenue dans l'exemple 9 sur Ω redonne m_1^Ω et la marginalisation sur Θ redonne m_2^Θ car les deux BBA ont été étendues avec l'opérateur d'extension vide sur $\Omega \times \Theta$. Par contre, si l'on combine cette BBA avec une BBA définie par $m_3^{\Omega \times \Theta}(\{T_f\} \times \Theta) = 1$ alors la BBA obtenue est notée $m_{123}^{\Omega \times \Theta}$ (par souci de simplicité, on ne note pas les combinaisons CRC en indices lorsque plus que 2 BBA sont combinées, c'est à dire que l'on assimile $m_{1 \oplus 2 \oplus 3}^{\Omega \times \Theta}$ à $m_{123}^{\Omega \times \Theta}$). La BBA $m_{123}^{\Omega \times \Theta}$ est donnée par :

$$\begin{aligned} m_{123}^{\Omega \times \Theta}(\{(T_f, M_c)\}) &= 0.49 & m_{123}^{\Omega \times \Theta}(\emptyset) &= 0.3 \\ m_{123}^{\Omega \times \Theta}(\{(T_f, B_c), (T_f, V_c), (T_f, M_c)\}) &= 0.14 & m_{123}^{\Omega \times \Theta}(\{(T_f, B_c), (T_f, V_c)\}) &= 0.07 \end{aligned} \quad (42)$$

La marginale sur Θ est :

$$\begin{aligned} m_{123}^{\Omega \times \Theta \downarrow \Theta}(\{M_c\}) &= 0.49 & m_{123}^{\Omega \times \Theta \downarrow \Theta}(\emptyset) &= 0.3 \\ m_{123}^{\Omega \times \Theta \downarrow \Theta}(\{B_c, V_c, M_c\}) &= 0.14 & m_{123}^{\Omega \times \Theta \downarrow \Theta}(\{B_c, V_c\}) &= 0.07 \end{aligned}$$

L'extension vide de ces marginales sur $\Omega \times \Theta$ ne redonne pas $m_{123}^{\Omega \times \Theta}$.

2.3 Déconditionnement

Disposant d'une BBA conditionnée $m^\Omega[B], B \subseteq \Omega$, il n'est généralement pas possible de retrouver la BBA d'origine m^Ω avant conditionnement. Une solution consiste à choisir la BBA la moins informative parmi toutes les BBA possibles et elle est obtenue par l'opérateur de *déconditionnement* [Sme93], noté \uparrow , menant à :

$$m^\Omega[B] \uparrow^\Omega(C \cup \bar{B}) = m^\Omega[B](C), \quad \forall C \subseteq B \quad (43)$$

où \bar{B} est le complémentaire de B . La masse initialement affectée à un sous-ensemble C avant déconditionnement (i.e. $m^\Omega[B](C)$) est transférée sur les sous-ensembles composés de C et ne contenant pas B (i.e. $C \cup \bar{B}$). Cette transformation assure que le (re)conditionnement sur B de la BBA obtenue donne de nouveau $m^\Omega[B]$, propriété obtenue grâce au PMI.

Il existe plusieurs applications du déconditionnement. Les plus intéressantes que nous décrivons dans les paragraphes suivants sont :

- *Traduire des règles implicatives en BBA* : comment traduire des implications logiques pondérées en fonctions de croyance ?
- *Fusionner des BBA conditionnelles sur un espace produit* : disposant de plusieurs BBAs $m^\Omega[\theta_i]$ définies sur un cadre Ω conditionnellement à des hypothèses θ_i d'un second cadre de discernement Θ , comment combiner ces BBAs ?

Cet opérateur permet par ailleurs d'ajouter des hypothèses dans un cadre de discernement [DS04] : disposant d'une BBA définie $\Omega_0 = \{\omega_1, \omega_2\}$, le déconditionnement redéfinit cette BBA sur $\Omega_1 = \Omega_0 \cup \{\omega_3\} = \{\omega_1, \omega_2, \omega_3\}$.

2.3.1 Le déconditionnement pour la traduction de règles implicatives

Le déconditionnement a été exploité par Ristic et Smets [RS05] pour la traduction de règles d'implication dans le cadre de l'identification de cibles. Ce type de règle est très bien adapté pour ajouter des connaissances a priori fournies par des experts.

Une règle d'implication R mettant en jeu deux cadres de discernement Ω et Θ est du type :

$$\ll R : \text{si } \theta \subseteq \Theta \text{ alors } \omega \subseteq \Omega \text{ avec une croyance } \alpha \in [0, 1] \gg$$

que l'on peut par ailleurs noter par $\theta \Rightarrow \omega$.

A partir de l'expression logique de la règle R , il est possible de construire une BBA que l'on pourra ensuite combiner avec d'autres BBA si l'application le nécessite. L'expression de la BBA équivalente à la règle est une BBA conditionnelle :

$$m_R^\Omega[\theta](\omega) = \alpha \quad (44a)$$

$$m_R^\Omega[\theta](\Omega) = 1 - \alpha \quad (44b)$$

Dans les systèmes réels, les BBA peuvent être exprimées sur des espaces produits. Si l'on souhaite combiner les BBA conditionnelles de règles implicatives avec d'autres BBA, il peut s'avérer utile d'exprimer les BBA des règles sur l'espace produit $\Omega \times \Theta$. La transformation d'une BBA conditionnelle en une BBA sur un espace produit est réalisée par l'opérateur de déconditionnement [RS05] qui, appliqué à la règle R (Eq. 44), fournit une BBA exprimée sur $\Omega \times \Theta$ et définie par :

$$m_R^\Omega[\theta]^{\uparrow \Omega \times \Theta}(D) = \begin{cases} \alpha & \text{si } D = (\omega \times \theta) \cup (\bar{\theta} \times \Omega) \\ 1 - \alpha & \text{si } D = (\Omega \times \Theta) \end{cases} \quad (45)$$

Pour déterminer et interpréter l'élément focal $D = (\omega \times \theta) \cup (\bar{\theta} \times \Omega)$, nous pouvons écrire la règle R comme l'union de deux règles :

- une première indiquant que si θ apparaît alors ω apparaît avec une croyance α : « $R_1 : \text{si } \theta \subseteq \Theta \text{ alors } \omega \subseteq \Omega \text{ avec une croyance } \alpha_1 \in [0, 1] \gg$,
- une seconde indiquant que si autre chose que θ apparaît, i.e. $\bar{\theta}$, alors on ne sait rien de la conclusion sur Θ : « $R_2 : \text{si } \bar{\theta} \subseteq \Theta \text{ alors } \Omega \text{ avec une croyance } \alpha_2 = 1 \gg$.

Après déconditionnement, la masse α est donc transférée sur l'ensemble $D = (\omega \times \theta) \cup (\bar{\theta} \times \Omega)$ composé de l'ensemble $\omega \times \theta$ pour lequel on dispose d'une connaissance (d'après l'équation 44) et de l'ensemble $\bar{\theta} \times \Omega$ représentant l'extension cylindrique de $\bar{\theta}$ sur $\Omega \times \Theta$ et modélisant l'ignorance sur Ω lorsque $\bar{\theta}$ est vrai. Notons que la BBA issue du déconditionnement respecte le Principe du Minimum d'Information et son conditionnement sur θ (avec marginalisation sur Θ) redonne la BBA initiale conditionnée [Sme93].

Exemple 11 Nous reconsidérons l'exemple 9 sur la reconnaissance de panneaux de signalisation. En plus des informations sur la forme et la couleur, on dispose d'une troisième source sous la forme d'une règle implicite spécifiée par un expert. Cette règle indique que si le panneau est rond alors il est bleu avec une croyance de 0.8, les 0.2 manquant représentent l'ignorance de l'expert concernant la couleur du panneau sous l'hypothèse que ce dernier soit rond. La règle est donc de la forme :

$$R : R_f \Rightarrow B_c, \alpha = 0.8$$

On souhaite combiner la règle avec la BBA $m_{1\ominus 2}^{\Omega \times \Theta}$ donnée à l'équation 40 (exemple 9). Pour cela, on procède en trois étapes :

1. L'expression de la règle sous la forme d'une BBA conditionnelle :

$$m_R^\ominus[R_f](\{B_c\}) = 0.8 \quad m_R^\ominus[R_f](\Theta) = 0.2 \quad (46)$$

2. La transformation de cette BBA sur l'espace produit $\Omega \times \Theta$ afin de la combiner avec $m_{1\ominus 2}^{\Omega \times \Theta}$. Pour cela, on utilise le déconditionnement de la règle :

$$m_R^{\Omega \times \Theta} = m_R^\ominus[R_f]^{\uparrow \Omega \times \Theta} \quad (47)$$

Le résultat du déconditionnement est la BBA suivante :

$$m_R^{\Omega \times \Theta}(\{(R_f, B_c), (T_f, B_c), (T_f, V_c), (T_f, M_c)\}) = 0.8 \quad m_R^{\Omega \times \Theta}(\Omega \times \Theta) = 0.2 \quad (48)$$

où $\Omega \times \Theta = \{(T_f, B_c), (T_f, V_c), (T_f, M_c), (R_f, B_c), (R_f, V_c), (R_f, M_c)\}$ est un espace produit. Cette BBA est équivalente à la forme conditionnelle car son conditionnement par R_f suivie d'une marginalisation sur Θ redonne la BBA conditionnelle.

3. La combinaison par la CRC de la BBA $m_{1\ominus 2}^{\Omega \times \Theta}$ avec la règle est exprimée de la façon suivante :

$$m_{12R}^{\Omega \times \Theta} = m_R^\ominus[R_f]^{\uparrow \Omega \times \Theta} \oplus m_{1\ominus 2}^{\Omega \times \Theta} \quad (49)$$

La BBA résultante comporte 13 éléments focaux dont les 5 plus élevés sont les suivants :

$$\begin{aligned} m_{12R}^{\Omega \times \Theta}(\{(T_f, M_c)\}) &= 0.462 \\ m_{12R}^{\Omega \times \Theta}(\emptyset) &= 0.168 \\ m_{12R}^{\Omega \times \Theta}(\{(T_f, B_c), (T_f, V_c), (T_f, M_c)\}) &= 0.1 \\ m_{12R}^{\Omega \times \Theta}(\{(R_f, B_c)\}) &= 0.072 \\ m_{12R}^{\Omega \times \Theta}(\{(T_f, B_c), (T_f, V_c)\}) &= 0.05 \\ &\dots \end{aligned} \quad (50)$$

Par exemple, $m_{12R}^{\Omega \times \Theta}(\{(T_f, M_c)\}) = m_{1 \odot 2}^{\Omega \times \Theta}(\{(T_f, M_c)\}) \cdot (m_R^{\Omega \times \Theta}(\{(R_f, B_c), (T_f, B_c), (T_f, V_c), (T_f, M_c)\}) + m_R^{\Omega \times \Theta}(\Omega \times \Theta)) + m_{1 \odot 2}^{\Omega \times \Theta}(\{(R_f, M_c), (T_f, M_c)\}) \cdot m_R^{\Omega \times \Theta}(\{(R_f, B_c), (T_f, B_c), (T_f, V_c), (T_f, M_c)\}) = 0.35 + 0.14 \cdot 0.8 = 0.462$.

Les marginales $m_{12R}^{\Omega \times \Theta \downarrow \Omega}$ et $m_{12R}^{\Omega \times \Theta \downarrow \Theta}$ sur Ω et Θ sont données par les BBA suivantes :

$$\begin{aligned} m_{12R}^{\Omega}(\{T_f\}) &= 0.612 & m_{12R}^{\Theta}(\{M_c\}) &= 0.532 \\ m_{12R}^{\Omega}(\emptyset) &= 0.168 & m_{12R}^{\Theta}(\emptyset) &= 0.168 \\ m_{12R}^{\Omega}(\{R_f\}) &= 0.132 & m_{12R}^{\Theta}(\{B_c, V_c, M_c\}) &= 0.152 \\ m_{12R}^{\Omega}(\{R_f, T_f\}) &= 0.088 & m_{12R}^{\Theta}(\{B_c, V_c\}) &= 0.076 \\ & & m_{12R}^{\Theta}(\{B_c\}) &= 0.072 \end{aligned} \quad (51)$$

2.3.2 Le déconditionnement pour la fusion de BBA conditionnelles

On suppose disposer d'un jeu de BBAs $m^{\Omega}[\theta_i], \forall \theta_i \in \Theta$ définies sur un cadre de discernement $\Omega = \{\omega_1, \omega_2 \dots\}$ conditionnellement à chacun des *singletons* θ_i d'un deuxième cadre de discernement $\Theta = \{\theta_1, \theta_2 \dots\}$.

Les BBA conditionnelles peuvent être combinées sur l'espace produit $\Omega \times \Theta$ obtenue par la combinaison conjonctive des BBA déconditionnées :

$$m^{\Omega \times \Theta} = \bigodot_i m^{\Omega}[\theta_i]^{\uparrow \Omega \times \Theta} \quad (52)$$

Chaque sous-ensemble S de l'espace produit $\Omega \times \Theta$ se voit allouer la masse :

$$m^{\Omega \times \Theta}(S) = \prod_{\theta_i \in \Theta} m^{\Omega}[\theta_i](v), \quad \forall S \subseteq \Omega \times \Theta \quad (53)$$

où $v = ((\theta_i \times \Omega) \cap S)^{\downarrow \Omega}$. Cette expression analytique des éléments de la BBA résultante a été dérivée par Smets [Sme78, Sme93]. De la même manière, l'expression analytique de la communalité des sous-ensemble $S \subseteq \Omega \times \Theta$ est donnée par :

$$q^{\Omega \times \Theta}(S) = \prod_{\theta_i \in \Theta} q^{\Omega}[\theta_i](v), \quad \forall S \subseteq \Omega \times \Theta \quad (54)$$

L'opération de déconditionnement utilisée de cette manière a permis de dériver le Théorème de Bayes Généralisé (§ 2.1) [Sme93]. Notons que les croyances conditionnelles sont supposées distinctes. Si ce n'est pas le cas, il faut utiliser la *combinaison conjonctive prudente* (Eq. 64).

Exemple 12 Nous reprenons l'exemple de la prédiction du temps décrit dans le paragraphe 2.1.1 et notamment les croyances conditionnelles $m^{\Omega_j}[E_{j-1}]$, $m^{\Omega_j}[P_{j-1}]$ et $m^{\Omega_j}[N_{j-1}]$ données dans le tableau 4. Il est possible de déterminer la BBA équivalente sur $\Omega_{j-1} \times \Omega_j$ à l'ensemble de ces BBA conditionnelles. Pour cela, deux étapes sont nécessaires : déconditionner chaque BBA puis les combiner. On a donc :

$$m^{\Omega_{j-1} \times \Omega_j} = m^{\Omega_j}[E_{j-1}]^{\uparrow \Omega_{j-1} \times \Omega_j} \odot m^{\Omega_j}[P_{j-1}]^{\uparrow \Omega_{j-1} \times \Omega_j} \odot m^{\Omega_j}[N_{j-1}]^{\uparrow \Omega_{j-1} \times \Omega_j} \quad (55)$$

Le résultat de la combinaison comporte environ 343 éléments focaux (potentiellement, il existe 2^9 éléments) et nous présentons seulement les 6 plus grandes valeurs approchées ci-dessous :

$$\begin{aligned}
m^{\Omega_{j-1} \times \Omega_j}(\{(N_{j-1}, N_j), (P_{j-1}, N_j), (E_{j-1}, E_j)\}) &= 0.168 \\
m^{\Omega_{j-1} \times \Omega_j}(\{(N_{j-1}, N_j), (P_{j-1}, P_j), (E_{j-1}, E_j)\}) &= 0.116 \\
m^{\Omega_{j-1} \times \Omega_j}(\{(N_{j-1}, N_j), (P_{j-1}, N_j), (E_{j-1}, N_j), (E_{j-1}, E_j)\}) &= 0.060 \\
m^{\Omega_{j-1} \times \Omega_j}(\{(N_{j-1}, N_j), (P_{j-1}, P_j), (E_{j-1}, N_j), (E_{j-1}, E_j)\}) &= 0.041 \\
m^{\Omega_{j-1} \times \Omega_j}(\{(N_{j-1}, N_j), (P_{j-1}, N_j), (P_{j-1}, E_j), (E_{j-1}, E_j)\}) &= 0.041 \\
m^{\Omega_{j-1} \times \Omega_j}(\{(N_{j-1}, E_j), (P_{j-1}, N_j), (E_{j-1}, E_j)\}) &= 0.037
\end{aligned} \tag{56}$$

2.4 Le PMI pour transformer une distribution de probabilités en BBA

L'estimation des distributions de masses (BBA) est un problème difficile. Une solution consiste à exploiter les outils performants d'estimation de distributions de probabilités puis de transformer ces distributions en BBA. La transformation « probabilités \rightarrow BBA » est nécessaire car les combinaisons CRC et DRC perdent leur intérêt lorsque les BBA combinées sont Bayésiennes.

Disposant d'une distribution de probabilités pignistiques \mathbf{BetP}^Ω , le problème est de déterminer une BBA m^Ω dont la transformation pignistique redonne \mathbf{BetP}^Ω . Il existe généralement plusieurs BBA qui vérifient cette propriété et, dans cette *famille*, le Principe du Minimum d'Information (PMI) impose de choisir la BBA qui est la moins *engagée*, c'est à dire la moins informative.

L'expression de la BBA est définie formellement de la façon suivante (théorème 7.1 de [Sme00c, Sme00b]) : disposant d'une distribution de probabilités pignistiques \mathbf{BetP}^Ω sur les singletons $\omega_1, \omega_2 \dots \omega_K$ de Ω (avec $K = |\Omega|$) dont les éléments de la distribution sont ordonnés :

$$\mathbf{BetP}^\Omega(\omega_1) \geq \mathbf{BetP}^\Omega(\omega_2) \geq \dots \geq \mathbf{BetP}^\Omega(\omega_k) \geq \dots \geq \mathbf{BetP}^\Omega(\omega_K)$$

Le PMI permet de construire la BBA la moins engagée \hat{m}^Ω associée à \mathbf{BetP}^Ω composée d'éléments focaux emboîtés $A = \{\omega_1, \omega_2 \dots \omega_k\}$ tels que :

$$\hat{m}^\Omega(A) = |A| \cdot (\mathbf{BetP}^\Omega(\omega_k) - \mathbf{BetP}^\Omega(\omega_{k+1})) \tag{57}$$

où $\mathbf{BetP}^\Omega(\omega_{K+1}) = 0$ par définition. Smets [Sme00b, Sme00c] a montré que cette BBA est *consonante* et la distribution de plausibilités associée est une distribution de possibilités. Notons qu'une méthode a été proposée pour passer d'une distribution de possibilités à une distribution de masses dont l'expression est assez similaire [DP82] à l'expression précédente.

Exemple 13 On pose $\Omega = \{\omega_1, \omega_2, \omega_3\}$ et on définit la distribution de probabilités pignistiques ainsi que les deux BBA suivantes :

$$\begin{array}{lll}
\mathbf{BetP}^\Omega(\{\omega_1\}) = 0.6 & m_1^\Omega(\{\omega_1\}) = 0.3 & m_2^\Omega(\{\omega_1\}) = 0.45 \\
\mathbf{BetP}^\Omega(\{\omega_2\}) = 0.2 & m_1^\Omega(\{\omega_1, \omega_2\}) = 0.4 & m_2^\Omega(\{\omega_1, \omega_2\}) = 0.1 \\
\mathbf{BetP}^\Omega(\{\omega_3\}) = 0.2 & m_1^\Omega(\{\omega_1, \omega_3\}) = 0.2 & m_2^\Omega(\{\omega_2, \omega_3\}) = 0.2 \\
& m_1^\Omega(\{\omega_3\}) = 0.1 & m_2^\Omega(\{\omega_1, \omega_3\}) = 0.1 \\
& & m_2^\Omega(\Omega) = 0.15
\end{array}$$

Les BBA m_1^Ω et m_2^Ω ont la même distribution de probabilités pignistiques définie par BetP^Ω .

Afin de faire un choix entre les deux BBA, on calcule la non-spécificité (Eq. 8). La non-spécificité de m_1^Ω est

$$N(m_1) = (0.3 + 0.1) \cdot \log_2(1) + (0.4 + 0.2) \cdot \log_2(2) = 0.6$$

et celle de m_2^Ω est

$$N(m_2) = 0.45 \cdot \log_2(1) + (0.1 + 0.2 + 0.1) \cdot \log_2(2) + 0.15 \cdot \log_2(3) = 0.6377$$

Par conséquent, la BBA la moins engagée est m_2 (puisque'elle est la moins spécifique) et correspond à celle que l'on devrait choisir si on avait seulement ces deux BBA dans la famille des BBA possibles. En reprenant l'expression 57, on détermine la BBA la moins informative parmi toutes les BBA possibles qui est donnée par la BBA consonante définie par :

$$m_3^\Omega(\{\omega_1\}) = 0.4 \quad m_3^\Omega(\Omega) = 0.6 \quad (58)$$

dont la non-spécificité est

$$N(m_3) = 0.951 > N(m_2) > N(m_1)$$

La masse sur $\{\omega_1\}$ est par exemple obtenue de la manière suivante : $m_3^\Omega(\{\omega_1\}) = 1 \cdot (0.6 - 0.2) = 0.4$ et celle sur Ω par $m_3^\Omega(\Omega) = 3 \cdot (0.2 - 0) = 0.6$.

2.5 Combinaison de sources non distinctes

Un grand nombre de mécanismes décrits dans les paragraphes précédents étaient basés sur des règles de combinaison conjonctive ou disjonctive sous l'hypothèse que les sources de croyance étaient *distinctes*, concept proche de l'*indépendance* en statistique [Sme93, YSM01, YSM02a, YSM02b].

Une règle de combinaison conjonctive dite *prudente* (CCRC : *Cautious Conjunctive Rule of Combination*), notée $\textcircled{\wedge}$, et une règle de combinaison disjonctive *hardie* (BDRC : *Bold Disjunctive Rule of Combination*), notée $\textcircled{\vee}$, ont été proposées par Denœux [Den07] (à partir de travaux initiés par Smets [Sme95]) pour combiner des sources *non-distinctes*. Ces règles sont associatives, commutatives et surtout *idempotentes* à la différence des règles usuelles CRC et DRC. Les quatre règles et leur utilisation sont représentées dans le tableau 9.

	sources fiables	sources non fiables
sources distinctes	$\textcircled{\cap}$	$\textcircled{\cup}$
sources non distinctes	$\textcircled{\wedge}$	$\textcircled{\vee}$

TAB. 9 : Les règles de combinaison et leur contexte d'utilisation [Den07].

Les règles CCRC et BDRC sont basées sur la décomposition canonique des BBA et sur des opérateurs flous que nous décrivons dans les paragraphes suivants.

2.5.1 Décomposition canonique d'une BBA

L'idée de la décomposition canonique [Sha76] est d'exprimer une BBA non-dogmatique ($m^\Omega(\Omega) > 0$) sous la forme d'une combinaison conjonctive ou disjonctive de BBA simples B^w . On rappelle que B^w est une écriture simplifiée signifiant que la BBA est formée de deux éléments focaux : $m^\Omega(B) = 1 - w$ et $m^\Omega(\Omega) = w$.

Pour définir la décomposition canonique de BBA quelconques, Smets [Sme95] a proposé le concept de fonction simple généralisée (GSBBA : *generalized simple BBA*) définie par une fonction $\mu : 2^\Omega \rightarrow \mathfrak{R}$ et vérifiant les contraintes suivantes :

$$\mu(B) = 1 - w \quad (59a)$$

$$\mu(\Omega) = w \quad (59b)$$

$$\mu(C) = 0, \quad \forall C \in 2^\Omega \setminus \{B, \Omega\} \quad (59c)$$

avec $B \neq \Omega$ et $w \in [0, +\infty[$. Lorsque $w \leq 1$, la BBA obtenue est une BBA *simple*.

La *décomposition canonique conjonctive* d'une BBA non-dogmatique, qui existe et qui est unique, est définie comme la combinaison conjonctive de GSBBA :

$$m^\Omega = \bigcap_{B \subset \Omega} B^{w(B)} \quad (60)$$

où les *poids* de la décomposition canonique conjonctive $w(B) \in [0, +\infty)$, $\forall B \subset \Omega$ sont définis par l'équation suivante :

$$w^\Omega(B) = \prod_{B \subseteq C} q^\Omega(C)^{(-1)^{|C|-|B|+1}}, \quad \forall B \subseteq \Omega \quad (61)$$

De la même manière, la *décomposition canonique disjonctive* [Den07] d'une BBA *sous normale* ($m^\Omega(\emptyset) > 0$) existe, est unique et elle est définie par :

$$m^\Omega = \bigcup_{B \neq \emptyset} B_{v(B)} \quad (62)$$

où la notation $B_{v(B)}$ représente une distribution de masses avec deux éléments focaux : $m^\Omega(\emptyset) = v(B)$ et $m^\Omega(B) = 1 - v(B)$, et les poids $v(B)$ sont donnés par :

$$v(B) = \overline{w}(\overline{B}) \quad (63)$$

Pour calculer les poids $v(B)$ il suffit de calculer la BBA \overline{m}^Ω , puis les poids associés $\overline{w}(B)$ par l'équation 61 et enfin de prendre le complémentaire de chaque poids (voir [Den07] pour des exemples, cf. exemple 14 ci-après).

2.5.2 Définitions des règles prudente (CCRC) et hardie (BDRC)

La CCRC et la BDRC sont basés sur des opérateurs flous de type *t-norme* (opérateur conjonctif flou) et *t-conorme* (opérateur disjonctif flou). Dencœux [Den07] a ainsi montré qu'une infinité d'opérateurs de combinaison conjonctive « prudente » peut être déterminée conférant au TBM un large ensemble de méthodes de combinaison de fonctions de croyance. Dans ce paragraphe, nous

nous limitons à un cas particulier : celui de la t-norme « minimum » utilisée à la fois dans la CCRC et la BDRC.

La CCRC \otimes est définie de la manière suivante :

$$m_1^\Omega \otimes m_2^\Omega = \bigcap_{B \subset \Omega} B^{w_1^\Omega(B) \wedge w_2^\Omega(B)} \quad (64)$$

où \wedge est l'opérateur « minimum ». Notons que la BBA vide n'est pas l'élément neutre de la CCRC [Den07].

La BDRC \oplus est donnée par :

$$m_1^\Omega \oplus m_2^\Omega = \bigcup_{B \neq \emptyset} B_{v_1(B) \vee v_2(B)} \quad (65)$$

où les poids $v_i(B)$ sont calculés à partir des poids $w_i(B)$ par l'équation 63.

Exemple 14 Nous reprenons l'exemple de la prédiction du temps. On dispose de deux sources d'information $m_1^{\Omega_j}$ et $m_2^{\Omega_j}$ sur le temps de la journée définies sur $\Omega_j = \{E_j, P_j, N_j\}$ et que l'on souhaite combiner pour améliorer la prise de décision sur le temps au jour j .

Le tableau ci-dessous donne les détails sur les différentes fonctions utilisées dans la CCRC et la BDRC et notamment les BBA m_i , leur négation \bar{m}_i , et les poids des décompositions (w_i et v_i). On remarque que certains poids (w_i et v_i) sont supérieurs à l'unité ce qui signifie que la BBA associée est une GSBBA.

	m_1^Ω	m_2^Ω	\bar{m}_1^Ω	\bar{m}_2^Ω	w_1^Ω	w_2^Ω	\bar{w}_1^Ω	\bar{w}_2^Ω	v_1^Ω	v_2^Ω
\emptyset_j	0.07	0.01	0.15	0.46	0.93	1	2.61	6.66	1	1
$\{E_j\}$	0	0	0.56	0	1	1.134	0.111	1	1	1
$\{P_j\}$	0	0	0	0.36	1.883	1	1	0.027	1	1
$\{E_j, P_j\}$	0.22	0.17	0	0	0.405	0.73	1	1	0.241	0.056
$\{N_j\}$	0	0	0.22	0.17	1	1	0.241	0.056	1	1
$\{E_j, N_j\}$	0	0.36	0	0	1	0.561	1	1	1	0.027
$\{P_j, N_j\}$	0.56	0	0	0	0.211	1	1	1	0.111	1
Ω_j	0.15	0.46	0.07	0.01	1	1	1	1	2.61	6.66

Les résultats de la combinaison conjonctive par la règle conjonctive usuelle (CRC) et la règle prudente (CCRC) sont donnés dans le tableau de gauche ci-dessous. Nous donnons aussi les résultats de la combinaison disjonctive usuelle (DRC) et hardie (BDRC). Pour le cas conjonctif, nous présentons les pignistiques dans le tableau de droite.

	$m_{1 \otimes 2}^\Omega$	$m_{1 \oplus 2}^\Omega$	$m_{1 \otimes 2}^\Omega$	$m_{1 \oplus 2}^\Omega$
\emptyset	0.079	0.261	0.001	0
$\{E_j\}$	0.079	0.051	0	0
$\{P_j\}$	0.095	0.245	0	0
$\{E_j, P_j\}$	0.164	0.066	0.051	0.007
$\{N_j\}$	0.202	0.131	0	0
$\{E_j, N_j\}$	0.054	0.035	0.025	0.016
$\{P_j, N_j\}$	0.258	0.167	0.006	0.003
Ω_j	0.069	0.045	0.917	0.973

	BetP ₁	BetP ₂	BetP _{1 \otimes 2}	BetP _{1 \oplus 2}
$\{E_j\}$	0.17	0.42	0.23	0.16
$\{P_j\}$	0.47	0.24	0.36	0.51
$\{N_j\}$	0.36	0.34	0.41	0.33

Les transformations pignistiques des BBA combinées (m_1^Ω et m_2^Ω) et des BBA obtenues par les combinaisons conjonctives ($m_{1\otimes 2}^\Omega$, $m_{1\oplus 2}^\Omega$) mènent à une décision différente (voir les valeurs en gras, tableau en bas à droite) : pluvieux si la décision est prise sur m_1 , ensoleillé si la décision est prise sur m_2 , nuageux si la décision est prise sur le résultat de la CRC et pluvieux si la décision est prise sur le résultat de la CCRC. La décision P_j à partir des pignistiques sur le résultat de la CCRC est en accord avec la décision sur m_1 cependant, la valeur du conflit en sortie de la CCRC reste élevée (0.261). La CRC favorise quant à elle un élément (N_j) sur lequel on aurait pas forcément parié avant fusion (relativement aux pignistiques sur les masses initiales). Si on analyse les deux BBA initiales, on s'aperçoit qu'elles sont plutôt conflictuelles concernant E_j et P_j et plutôt concensuelles concernant N_j . Ceci explique que, par la CRC, c'est N_j qui l'emporte (consensus par l'intersection).

Ces opérateurs s'appliquant aux sources non distinctes sont particulièrement intéressants dans le cadre des extensions de méthodes probabilistes où des hypothèses d'indépendance sont souvent émises pour simplifier les calculs (et pour appliquer l'opérateur « produit »). C'est le cas, par exemple, dans le cadre des réseaux Bayésiens [Mur02] pour lesquels il est presque systématiquement admis que les distributions observées sont indépendantes conditionnellement aux états. Dans le cadre des réseaux évidentiels [XS94, Sme93] (qui étendent les réseaux Bayésiens aux fonctions de croyance), les distributions de croyance non indépendantes peuvent être combinées avec la BDRC ou la CCRC (en fonction des cas). La CCRC et la BDRC représentent alors des alternatives intéressantes de gestion de l'indépendance avec un panel d'opérateurs possibles.

2.6 Affaiblissement généralisé dit *contextuel*

L'affaiblissement *simple* décrit dans le paragraphe 1.4.4 agit de la même manière (avec le même coefficient) quelle que soit l'hypothèse. La généralisation de l'affaiblissement est basée sur l'idée que la fiabilité d'une source peut agir différemment sur les différentes hypothèses du cadre de discernement [MQD07, Mer06].

La généralisation passe par deux étapes : la définition d'une partition notée $\Theta = \{\theta_1, \theta_2, \dots, \theta_G\}$ du cadre de discernement $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$ et la quantification des coefficients d'affaiblissement α_g , $g \in \{1 \dots G\}$ sur chaque élément de la partition Θ . Chaque élément $\theta_g \in \Theta$ de la partition se voit donc affecter un taux d'affaiblissement α_g .

Disposant de G éléments θ_g dans la partition et donc de G taux d'affaiblissement rassemblés dans le vecteur $\alpha = [\alpha_1 \dots \alpha_g \dots \alpha_G]$, l'affaiblissement d'une BBA m^Ω génère une BBA notée αm^Ω définie par :

$$\alpha m^\Omega = m^\Omega \odot m_1^\Omega \odot m_2^\Omega \dots \odot m_g^\Omega \dots \odot m_G^\Omega \quad (66)$$

où les BBA m_g^Ω , $g \in \{1 \dots G\}$ sont définies par :

$$\begin{aligned} m_g^\Omega(\emptyset) &= 1 - \alpha_g \\ m_g^\Omega(\theta_g) &= \alpha_g \end{aligned} \quad (67)$$

Lorsque la partition est égale à Ω , i.e. $\Theta = \{\Omega\}$ (il n'existe donc qu'un seul coefficient de fiabilité), l'affaiblissement généralisé devient un affaiblissement simple (§ 1.4.4).

Exemple 15 On considère la BBA m_2^Ω définie dans l'exemple 14 où $\Omega = \{E_j, P_j, N_j\}$. On considère que la source de croyance fournissant m_2^Ω n'est fiable qu'à 60% en ce qui concerne l'hypothèse E_j et qu'à 85% pour les hypothèses P_j et N_j . On crée alors une partition $\Theta = \{\theta_1, \theta_2\}$ où $\theta_1 = \{E_j\}$ et $\theta_2 = \{P_j, N_j\}$. Les coefficients d'affaiblissement sont $\alpha = [0.4, 0.15]$ et la BBA affaiblie αm_2^Ω a les éléments focaux suivants :

$$\begin{aligned} \alpha m_2^\Omega(\emptyset) &= 0.0051 & \alpha m_2^\Omega(\{E_j\}) &= 0.0034 \\ \alpha m_2^\Omega(\{E_j, P_j\}) &= 0.1870 & \alpha m_2^\Omega(\{P_j, N_j\}) &= 0.3729 \\ \alpha m_2^\Omega(\Omega) &= 0.4316 \end{aligned} \quad (68)$$

La masse sur E_j était nulle avant affaiblissement et devient positive (faiblement) après le processus d'affaiblissement contextuel car ce dernier est une généralisation qui a pour effet de transférer des masses vers des sur-ensembles (la non-spécificité augmente). Ainsi, la masse 0.01 allouée à l'ensemble vide avant affaiblissement est en partie transférée vers E_j qui se voit allouer une masse de $0.01 \cdot (0.4 \cdot 0.85) = 0.0034$ par le processus d'affaiblissement (Eq. 66 suivie de Eq. 67).

Conclusion

Le Modèle des Croyances Transférables (TBM) permet de représenter et combiner la connaissance à l'aide de fonctions de croyance. Le TBM offre des outils pour combiner des BBA définies sur des espaces différents, pouvant être plus ou moins précis ou grossiers. Des opérateurs de combinaison conjonctif et disjonctif permettent par ailleurs de fusionner des BBA provenant de sources pouvant être plus ou moins fiables et non-distinctes. Ces mécanismes de base ont été utilisés pour de nombreux problèmes (voir l'état de l'art dressé par Vannoorenberghe [Van03] en 2003) et notamment en reconnaissance des formes et classification [Den04, Quo06, DS06].

Références

- [App91] A. Appriou. Probabilités et incertitudes en fusion de données multi-senseurs. *Revue Scientifique et Technique de la Défense*, 11 :27,40, 1991.
- [AS01] A. Ayoun and Ph. Smets. Data association in multi-target detection using the Transferable Belief Model. *Int. Jour. of Intelligent Systems*, 16 :1167–1182, 2001.
- [Blo05] I. Bloch. Fusion d'informations numériques : panorama méthodologique. In *Journées Nationales de la Recherche en Robotique*, Guidel, Morbihan, Octobre 2005. <http://jnrr05.irisa.fr/>.
- [Cat03] M. Cattaneo. Combining belief functions issued from dependent sources. In *Proc. of the Third Int. Symposium on Imprecise Probabilities and Their Applications*, pages 133–147, Lugano, Switzerland, 2003. Carleton Scientific.
- [CCFM04] A.S. Capelle, O. Colot, and C. Fernandez-Maloigne. Evidential segmentation scheme of multi-echo mr images for the detection of brain tumors using neighborhood information. *Information Fusion*, 5 :203–216, 2004.

- [Dem67] A.P. Dempster. Upper and lower probabilities induced by multiple valued mappings. *Annals of Mathematical Statistics*, 38 :325–339, 1967.
- [Dem68] A.P. Dempster. A generalization of Bayesian inference. *Journal of the Royal Statistical Society*, 30 :205–247, 1968.
- [Den04] T. Denoeux. Théorie des fonctions de croyance et classification. Séminaire, thème AS-TRID, UTC, Juin 2004.
- [Den07] T. Denoeux. Conjunctive and disjunctive combination of belief functions induced by non distinct bodies of evidence. *Artificial Intelligence*, 2007. To appear.
- [DGPS01] D. Dubois, M. Grabisch, H. Prade, and Ph. Smets. Using the transferable belief model and a qualitative possibility theory approach on an illustrative example : the assessment of the value of a candidate. *Int. Jour. of Intelligent Systems*, 16 :1245–1272, 2001.
- [DM04] T. Denoeux and M.H. Masson. EVCLUS : evidential clustering of proximity data. *IEEE Trans. Systems, Man and Cybernetics B*, 34(1), February 2004.
- [DP82] D. Dubois and H. Prade. On several representations of an uncertainty body of evidence. In M. M. Gupta and E. Sanchez, editors, *Fuzzy Information and Decision Processes*, pages 167–181, North-Holland, New-York, 1982.
- [DP86a] D. Dubois and H. Prade. The principle of minimum specificity as a basis for evidential reasoning. In *Uncertainty in knowledge-based systems*, volume 12, pages 75–84. Springer-Verlag, 1986.
- [DP86b] D. Dubois and H. Prade. A set-theoretic view of belief functions : Logical operations and approximations by fuzzy sets. *Int. Jour. of General Systems*, 12 :193–226, 1986.
- [DS04] F. Delmotte and Ph. Smets. Target identification based on the Transferable Belief Model interpretation of Dempster-Shafer model. *IEEE Trans. on Systems, Man and Cybernetics*, 34(4), July 2004.
- [DS06] T. Denoeux and Ph. Smets. Classification using belief functions : The relationship between the case-based and model-based approaches. *IEEE Trans. on Systems, Man and Cybernetics*, 2006. To appear.
- [DY02] T. Denoeux and A.B. Yaghlane. Approximating the combination of belief functions using the fast Moebius transform in a coarsened frame. *Int. Jour. of Approximate Reasoning*, 37 :77–101, 2002.
- [GCBR05] V. Girondel, A. Caplier, L. Bonnaud, and M. Rombaut. Belief theory-based classifiers comparison for static human body postures recognition in video. *Int. Jour. of Signal Processing*, 2(1) :29–33, 2005.
- [Har99] D. Harmanec. Measures of uncertainty and information. 1999.
- [HCR05] Z. Hammal, A. Caplier, and M. Rombaut. Belief theory applied to facial expressions classification. In *Int. Conf. on Advances in Pattern Recognition*, Bath, United Kingdom, 2005.

- [KM95] J. Kholas and P.A. Monney. A mathematical theory of hints : An approach to the Dempster-Shafer theory of evidence. In *Lecture Notes in Economics and Mathematical Systems*, volume 425. Springer-Verlag, 1995.
- [KW99] G.J. Klir and M.J. Wierman. *Uncertainty-based information. Elements of generalized information theory, 2nd edition*. Studies in fuzzyness and soft computing. Physica-Verlag, 1999.
- [LCV02] E. Lefevre, O. Colot, and P. Vannoorenberghe. Belief function combination and conflict management. *Information Fusion*, 3 :149–162, 2002.
- [Mer06] D. Mercier. *Fusion d'informations pour la reconnaissance automatique d'adresses postales dans le cadre de la théorie des fonctions de croyance*. PhD thesis, Université de Technologie de Compiègne, December 2006.
- [MQD07] D. Mercier, B. Quost, and T. Denoeux. Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 2007. To appear.
- [Mur02] K. P. Murphy. *Dynamic Bayesian Networks : Representation, inference and learning*. PhD thesis, UC Berkeley (CSD), 2002.
- [Quo06] Benjamin Quost. *Combinaison de classifieurs binaires dans le cadre de la théorie des fonctions de croyance*. PhD thesis, Université de Technologie de Compiègne, November 2006.
- [Rab89] L.R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of the IEEE*, 77 :257–285, 1989.
- [RPR06] E. Ramasso, D. Pellerin, and M. Rombaut. Belief Scheduling for the recognition of human action sequence. In *Int. Conf. on Information Fusion*, Florence, Italia, 2006.
- [RPRP07] E. Ramasso, C. Panagiotakis, M. Rombaut, and D. Pellerin. Human action recognition in videos based on the transferable belief model - application to athletics jumps. *Pattern Analysis and Applications Journal*, 2007. Accepted, doi :10.1007/s10044-007-0073-y.
- [RRP07a] E. Ramasso, M. Rombaut, and D. Pellerin. Algorithmes forward, backward et viterbi dans le cadre du modèles des croyances transférables - application à la classification de vidéos. In *21th colloque GRETSI : Traitement du signal et des Images*, Troyes, France, 11-14 Septembre 2007.
- [RRP07b] E. Ramasso, M. Rombaut, and D. Pellerin. Forward-backward-viterbi procedures in the transferable belief model for state sequence analysis using belief functions. In *European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume Lecture Notes in Computer Sciences, Hammamet, Tunisia, Oct. 31-Nov. 2 2007. Springer-Verlag. Accepted.
- [RRP07c] E. Ramasso, M. Rombaut, and D. Pellerin. State filtering and change detection using TBM conflict - application to human action recognition in athletics videos. *IEEE Trans. on Circuits and Systems for Video Technology*, 17(7) :944–949, 2007. Accepted for publication.

- [RS05] B. Ristic and Ph. Smets. Target identification using belief functions and implication rules. *IEEE Trans. Aerospace and Electronic Systems*, 41(3) :1097–1102, 2005.
- [RS06] B. Ristic and Ph. Smets. The TBM global distance measure for the association of uncertain combat ID declarations. *Information Fusion*, 7 :276–284, 2006.
- [Sch04] J. Schubert. Clustering belief functions based on attracting and conflicting metalevel evidence using Potts spin mean field theory. *Information Fusion*, 5(4) :309–318, 2004.
- [Sha76] G. Shafer. *A mathematical theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
- [SK94] Ph. Smets and R. Kennes. The Transferable Belief Model. *Artificial Intelligence*, 66(2) :191–234, 1994.
- [Sme78] Ph. Smets. *Un modèle mathématico-statistique simulant le processus du diagnostic médical*. PhD thesis, Université Libre de Bruxelles, Brussels, Belgium, 1978. (in French).
- [Sme92] Ph. Smets. The nature of the unnormalized beliefs encountered in the Transferable Belief Model. In *Proceedings of the 8th Annual Conference on Uncertainty in Artificial Intelligence*, pages 292–297, 1992.
- [Sme93] Ph. Smets. Beliefs functions : The Disjunctive Rule of Combination and the Generalized Bayesian Theorem. *Int. Jour. of Approximate Reasoning*, 9 :1–35, 1993.
- [Sme94a] Ph. Smets. *Advances in the Dempster-Shafer Theory of Evidence - What is Dempster-Shafer's model ?*, pages 5–34. Wiley, 1994.
- [Sme94b] Ph. Smets. Belief induced by the knowledge of some probabilities. In *Uncertainty in artificial intelligence*, pages 523–530. Morgan Kaufmann, 1994.
- [Sme95] Ph. Smets. The canonical decomposition of a weighted belief. In *In Int. Joint Conf. on Artificial Intelligence*, pages 1896–1901, San Mateo, Ca, 1995. Morgan Kaufman.
- [Sme00a] Ph. Smets. Belief functions and the Transferable Belief Model. Available on www.sipta.org/documentation/belief/belief.ps, 2000.
- [Sme00b] Ph. Smets. Quantified possibility theory seen as an hyper cautious transferable belief model. In Cepadues-Editions, editor, *Rencontres francophones sur les logiques floues et ses applications*, pages 343–353, La Rochelle, France, 2000.
- [Sme00c] Ph. Smets. Quantitative epistemic Possibility Theory seen as an hyper-cautious Transferable Belief Model. Technical Report 18, IRIDIA, 2000.
- [Sme02a] Ph. Smets. The application of the matrix calculus to belief functions. *Int. Jour. of Approximate Reasoning*, 31(1-2) :1–30, 2002.
- [Sme02b] Ph. Smets. Decision making in a context where uncertainty is represented by belief functions. In *Belief Functions in Business Decisions*, pages 17–61. Physica-Verlag, 2002.
- [Sme05a] Ph. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion*, 8(4) :387–412, October 2005.
- [Sme05b] Ph. Smets. Decision making in the TBM : The necessity of the pignistic transformation. *Int. Jour. of Approximate Reasoning*, 38 :133–147, 2005.

- [Van03] P. Vannoorenberghe. Un état de l'art sur les fonctions de croyance appliquées au traitement de l'information. *Revue I3*, 3(1) :9–45, 2003.
- [VS05] P. Vannoorenberghe and Ph. Smets. Partially supervised learning by a Credal EM approach. In *Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 3571, pages 956–967, 2005.
- [XS94] H. Xu and Ph. Smets. Evidential reasoning with conditional belief functions. *Uncertainty in Artificial Intelligence*, 94 :598–606, 1994.
- [Yag87] R.R. Yager. On the Dempster-Shafer framework and new combination rules. *Information Sciences*, 41 :93–137, 1987.
- [YSM01] B. Yaghlane, Ph. Smets, and K. Mellouli. Belief function independence : I. The marginal case. *Int. Jour. of Approximate Reasoning*, 29 :47–70, 2001.
- [YSM02a] B. Yaghlane, Ph. Smets, and K. Mellouli. Belief function independence : II. The conditionnal case. *Int. Jour. of Approximate Reasoning*, 31(1) :31–75, October 2002.
- [YSM02b] B. Yaghlane, Ph. Smets, and K. Mellouli. Independence concept for belief functions. In *Technologies for constructing intelligent systems : Tools*. Physica-Verlag, 2002.
- [YSM03] B. Ben Yaghlane, Ph. Smets, and K. Mellouli. Directed evidential networks with conditionnal belief functions. In *Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 291–305, 2003.