
Mémoire scientifique

Pour l'obtention de

L'HABILITATION À DIRIGER DES RECHERCHES

UNIVERSITÉ D'ARTOIS

ÉCOLE DOCTORALE SCIENCES POUR L'INGÉNIEUR

LILLE-NORD DE FRANCE

DOMAINE DE RECHERCHE : Génie Informatique

Présentée par

David MERCIER

Corrections et fusion d'informations dans le cadre des fonctions de croyance. Applications.

Directeur de recherche : **M. LEFÈVRE Éric**

Soutenance le 4 décembre 2015
Devant la commission d'examen suivante

JURY

Mme. BLOCH Isabelle	Professeur, Télécom ParisTech, Paris	Rapporteur
M. DUBOIS Didier	Directeur de recherche, CNRS - IRIT, Toulouse	Rapporteur
Mme. ROMBAUT Michèle	Professeur, Université Joseph Fourier, Grenoble	Rapporteur
M. COLOT Olivier	Professeur, Université de Lille 1	Examineur
M. DENŒUX Thierry	Professeur, Université de technologie de Compiègne	Examineur
M. GONCALVES Gilles	Professeur, Université d'Artois	Examineur
M. JOLLY Daniel	Professeur émérite, Université d'Artois	Examineur
M. LEFÈVRE Éric	Professeur, Université d'Artois	Directeur

Remerciements

Je tiens tout d'abord à remercier très vivement Isabelle Bloch, professeur à Télécom ParisTech, Didier Dubois, directeur de recherche à l'IRIT et Michèle Rombaut, professeur à l'université Joseph Fourier de Grenoble, pour m'avoir fait l'honneur d'accepter de rapporter sur ce travail.

Je suis très heureux que Thierry Dencœux, professeur à l'université de technologie de Compiègne, ait pu participer à ce jury. Il a été mon directeur de thèse, il m'a beaucoup appris avec Marie-Hélène Masson, maître de conférences HDR à l'Université de Picardie Jules Verne, et ses travaux continuent d'inspirer nombre des miens.

Je voudrais aussi remercier Olivier Colot, professeur à l'Université de Lille 1, pour avoir accepté de faire partie de ce jury. Je trouve très appréciable le fait pour nos deux laboratoires de pouvoir collaborer, comme pour l'organisation de l'école sur les fonctions de croyance par exemple. J'adresse à John Klein, maître de conférences à l'Université de Lille 1, un merci également sur ce point et pour ses invitations à assister à des séminaires à Lille 1 sur ce thème.

Un grand merci également à Daniel Jolly, professeur émérite à l'université d'Artois et à Gilles Goncalves professeur à l'université d'Artois pour m'avoir accueilli il y a huit ans, en 2007, au sein du LGI2A. J'ai bénéficié d'un excellent environnement de travail, simple et convivial, où dès mon arrivée, j'ai eu les moyens de mener des travaux sur des projets attrayants.

Je souhaiterais aussi remercier tous mes co-auteurs, collègues et doctorants, sans qui les travaux présentés ici n'auraient pu voir le jour. J'ai des pensées particulières pour Éric Lefèvre, professeur à l'université d'Artois, avec qui j'ai commencé à travailler dès mon arrivée en 2007 au sein de l'axe décision et fusion d'informations du LGI2A, et qui m'a appris beaucoup de choses sur ce métier, pour Frédéric Pichon, maître de conférences à l'université d'Artois, pour toute son énergie et ses réflexions consacrées à développer les mécanismes de corrections présentés dans ce rapport (et à essayer de comprendre d'autres mystères autour des fonctions de croyance), et pour François Delmotte pour la direction des thèses de Mira Bou Farah et Samir Hachour.

Je remercie aussi les membres du LGI2A et les collègues de l'IUT, enseignants-chercheurs, secrétaires, ingénieurs, doctorants et post-doctorants, pour la bonne ambiance qui rend le travail bien plus agréable.

Enfin, je remercie Isabelle, qui partage ma vie, pour tout et pour bien des choses encore. J'ai une pensée également pour mes parents, mon frère, ma famille et mes amis.

Table des matières

Introduction	1
1 Concepts de base des fonctions de croyance	3
1.1 Introduction	3
1.2 Représentation des connaissances	3
1.2.1 Fonction de masse et interprétations	3
1.2.2 Définitions	4
1.2.3 Autres fonctions pour représenter les connaissances	5
1.2.4 Grossissement et raffinement	6
1.3 Manipulation des connaissances	7
1.3.1 Combinaison d'informations	7
1.3.2 Principe de minimum d'information	7
1.3.3 Marginalisation et extension vide	8
1.3.4 Conditionnement et déconditionnement	9
1.3.5 Mécanismes de correction de fonction de croyance	10
1.4 Prise de décision	13
1.5 Conclusion	14
2 Affaiblissement et renforcement contextuels	17
2.1 Introduction	17
2.2 Décompositions canoniques conjonctive et disjonctive	18
2.3 Affaiblissement contextuel	19
2.4 Renforcement contextuel	24
2.5 Apprentissage automatique des coefficients	26
2.5.1 Description du processus d'apprentissage	27
2.5.2 Cas de l'affaiblissement contextuel	28
2.5.3 Cas du renforcement contextuel	29
2.5.4 Discussions et exemples	29
2.5.5 Une expérimentation en classification	31
2.6 Conclusion et perspectives	33
3 Applications en fusion d'informations	35
3.1 Introduction	35
3.2 Application en communication inter-véhicule	35
3.2.1 Description de l'application et des objectifs	35
3.2.2 Solutions proposées pour la gestion des événements locaux	36
3.2.3 Simulations	39
3.2.4 Conclusion et perspectives	42
3.3 Association de données avec des fonctions de croyance.	44
3.3.1 Description de l'application et des objectifs	44

3.3.2	Un algorithme d'association de données utilisant des fonctions de croyance	45
3.3.3	Méthode de Dencœux et al.	49
3.3.4	Comparaison sur un scénario de suivi de cibles avec des données cinématiques	51
3.3.5	Conclusion et perspectives	53
3.4	Application au diagnostic du vieillissement du bobinage de moteurs électriques	54
3.4.1	Description de l'application et des objectifs	54
3.4.2	Solution développée	54
3.4.3	Résultats obtenus	57
3.4.4	Conclusion et perspectives	58
3.5	Conclusion sur ce chapitre	58
4	Projet de recherche	59
A	Preuve de l'équation (2.12)	63
B	Algorithme GNN (Global Nearest Neighbour)	65
C	Sélection de publications	67
	Bibliographie	165

Introduction

Ce document est une synthèse de mes principaux travaux de recherche, réalisés depuis la fin de ma thèse soutenue en décembre 2006, en vue de l'obtention du diplôme de l'habilitation à diriger des recherches.

Ces travaux ont pour cadre formel la **théorie des fonctions de croyances** (ou théorie de Dempster-Shafer) permettant de représenter et manipuler des données imparfaites (e.g. incertaines, imprécises, incomplètes). Celle-ci trouve ses origines dans les travaux de Dempster [18, 19] et Shafer [71]. Elle a aussi été particulièrement développée par Smets avec ses travaux concernant le modèle des croyances transférables (MCT) [78, 74].

Dans ce cadre, nous nous sommes intéressés au problème de la **correction d'une source d'information**. Nous entendons par cela le fait de détecter les biais d'une source afin de corriger/ajuster l'information qu'elle fournit. Pour cela, nous avons étendu, et étudié plus en profondeur, le mécanisme d'affaiblissement contextuel d'une fonction de croyance introduit durant ma thèse et en avons développé de nouveaux [56, 66]. Nous avons aussi montré qu'il était possible d'apprendre automatiquement les paramètres de ces mécanismes à partir de données étiquetées issues d'une source [61]. Ces coefficients permettent à la fois de révéler les biais de la source (les coefficients sont interprétables) et d'améliorer les performances de la source.

À côté de ce travail de développement de mécanismes de correction, nous avons été amenés à travailler sur plusieurs problèmes de fusions d'informations, c'est-à-dire des problèmes d'agrégation de connaissances issues de différentes sources dans le but d'améliorer la prise de décision [5]. Trois de ces problèmes de fusion d'informations sont développés dans la suite de ce rapport. Ces choix ont été principalement guidés par mon implication dans ces travaux. L'intérêt de ces applications est de montrer comment le cadre des fonctions de croyance peut simplement modéliser les informations disponibles, tout en assurant de bonnes performances, mais aussi, de montrer que toutes les modélisations de problèmes employant ce cadre n'aboutissent pas aux mêmes performances. La modélisation d'une solution dépend en effet d'une part, du choix des informations à modéliser, et d'autre part, de la manière dont ces informations ont été représentées par des fonctions de croyance.

La première application présentée se situe dans le domaine de la **communication véhicule à véhicule**, où nous avons développé, dans le cadre de la thèse (2010-2014) de Mira Bou Farah [6, 11], des méthodes utilisant des fonctions de croyance pour modéliser et fusionner des informations imparfaites au regard d'événements pouvant survenir sur la route, afin de fournir aux conducteurs une synthèse la plus proche possible de la réalité.

La seconde concerne le problème d'**association de données** où les informations relatives aux associations possibles entre objets sont données par des fonctions de croyance [59]. Ce problème consiste à mettre en correspondance chaque objet d'un premier ensemble avec zéro ou un objet d'un deuxième ensemble, de telle façon que

chaque objet ne peut être associé au maximum qu'à un autre objet. En d'autres termes, il s'agit de former des couples entre deux ensembles, sachant que certains objets peuvent rester seuls. Ce problème peut se rencontrer par exemple en suivi multi-objets lorsqu'il faut mettre en relation les observations d'un capteur avec les prédictions du système de suivi [3, 4].

Enfin, le troisième travail que nous avons sélectionné a été réalisé en collaboration avec des électrotechniciens du laboratoire systèmes électrotechniques et environnement (LSEE) de l'université d'Artois. Il concerne le **diagnostic du bobinage de moteur électrique** et l'estimation du vieillissement en fusionnant des informations issues de capteurs mesurant le champ magnétique extérieur de la machine [67, 60, 64, 63].

Cette partie est organisée en quatre chapitres.

Dans le chapitre 1, les notions de base sur les fonctions de croyance que nous avons employées ainsi que les notations utilisées sont abordées.

Dans le chapitre 2, nous présentons une synthèse de nos résultats [56, 66, 61] concernant les mécanismes d'affaiblissement et de renforcement contextuels. Nous mettons en particulier en avant l'intérêt de leurs apprentissages automatiques pour déterminer les biais potentiels d'une source et en améliorer les performances.

Dans le chapitre 3, nous exposons une contribution issue des travaux de thèse de Mira Bou Farah [6] portant sur la communication inter-véhicule. Nous résumons ensuite une méthode que nous avons développée autour de 2009 [57, 58, 59] pour résoudre le problème d'association avec des fonctions de croyance, et la comparons avec une méthode récemment introduite par Denœux et al. [24] sur une application de suivi de deux objets issue des travaux de thèse de Samir Hachour [34] portant sur le suivi et la classification multi-objets. Enfin, le système de fusion d'informations que nous avons développé pour le diagnostic du bobinage de moteur électrique est abordé.

Nos conclusions et perspectives pour chaque travail sont mentionnées en fin de chacun des chapitres précédents. Mon projet de recherche est abordé dans le dernier chapitre 4 de ce rapport.

Enfin, une sélection des publications les plus significatives de mon travail se trouve en annexe C.

Concepts de base de la théorie des fonctions de croyance employés.

1.1 Introduction

Les notions de base sur les fonctions de croyance employées dans nos travaux ainsi que les notations utilisées sont présentées dans ce chapitre. Il s'agit d'un résumé. Le lecteur désirant approfondir les notions abordées dans ce chapitre peut regarder les articles de Dempster [18, 19], le livre référence de Shafer [71], les articles de Smets dont [78, 74], ainsi que les cours et présentations en ligne de Denœux [20] notamment celles liées à l'école internationale sur les fonctions de croyance.

Le plan de ce chapitre est divisé en trois parties principales. Nous abordons d'abord la représentation des connaissances avec des fonctions de croyance en section 1.2, puis les manipulations possibles de ces informations en section 1.3, et enfin la prise de décision à partir d'une connaissance représentée par une fonction de croyance en section 1.4.

1.2 Représentation des connaissances

1.2.1 Fonction de masse et interprétations

Soit x une variable à valeurs dans un ensemble fini $\mathcal{X} = \{x_1, \dots, x_K\}$, appelé *cadre de discernement*, et soit x_0 la véritable valeur cherchée pour x .

Une *fonction de masse (FM)* $m^{\mathcal{X}}$ sur $\mathcal{X}$ est une fonction définie de l'ensemble des parties de $\mathcal{X}$ dans l'intervalle réel $[0, 1]$, telle que la somme des masses est égale à 1. Formellement :

$$\begin{aligned} m^{\mathcal{X}} : 2^{\mathcal{X}} &\rightarrow [0, 1] \\ A &\mapsto m^{\mathcal{X}}(A) , \end{aligned} \tag{1.1}$$

et

$$\sum_{A \subseteq \mathcal{X}} m^{\mathcal{X}}(A) = 1 . \tag{1.2}$$

Dans le MCT de Smets [78, 74] et dans le livre de Shafer [71], une FM représente une opinion pondérée d'un agent au regard de la vraie valeur x_0 prise par la variable x .

Une autre interprétation peut être donnée en considérant les travaux de Dempster [18, 19], et Kohlas et Monney [44]. Dans ces modèles, une FM $m^{\mathcal{X}}$ sur $\mathcal{X}$ est

induite par une source définie par un 4-uplet $(\Theta, P^\Theta, \Gamma, \mathcal{X})$ où :

- Θ est un ensemble fini ;
- P^Θ est une mesure de probabilité sur Θ ;
- Γ est une fonction définie de Θ dans $2^\mathcal{X}$.

La fonction Γ *interprète (transporte)* l'information donnée par P^Θ sur Θ dans $2^\mathcal{X}$. Pour tout $A \subseteq \mathcal{X}$, la masse $m^\mathcal{X}(A)$ peut alors être interprétée comme la probabilité que la vraie valeur cherchée x_0 de la variable d'intérêt x se trouve dans A et de ne rien savoir de plus. Formellement, pour tout $A \subseteq \mathcal{X}$, $m^\mathcal{X}(A) = P^\Theta(\{\theta \in \Theta, \Gamma(\theta) = A\})$.

1.2.2 Définitions

Chaque sous-ensemble A de $\mathcal{X}$ tel que $m(A) > 0$ est appelé *élément focal* de m .

Une FM telle que $m(A) = 1$ est dite *catégorique* ou *logique*. On sait juste que la valeur cherchée x_0 appartient à l'ensemble A . Elle est notée m_A . Un cas particulier est la FM $m_\mathcal{X}$ qui correspond à l'ignorance totale, elle est aussi appelée *fonction de masse vide*. Elle n'apporte aucune information (outre le fait que la variable d'intérêt x prend bien ses valeurs dans $\mathcal{X}$).

Une FM n'ayant pour éléments focaux que des éléments singletons est dite *bayésienne*. Elle correspond à une distribution de probabilité.

Une FM telle que $m(\emptyset) = 0$ est dite *normale*. Une FM non normale m peut être transformée en une FM normale notée m^* par l'*opération de normalisation* définie pour tout $A \subseteq \mathcal{X}$ par :

$$m^*(A) = \begin{cases} k \cdot m(A) & \text{si } A \neq \emptyset, \\ 0 & \text{sinon,} \end{cases} \quad (1.3)$$

avec $k = (1 - m(\emptyset))^{-1}$.

La *négation* [26] $\bar{m}$ d'une FM m est définie de la manière suivante : $\bar{m}(A) = m(\bar{A})$, $\forall A \subseteq \mathcal{X}$, où $\bar{A}$ est le complémentaire de A dans $\mathcal{X}$.

Une FM ayant pour éléments focaux $\mathcal{X}$ ainsi qu'un sous-ensemble $A \subset \mathcal{X}$, avec pour masses respectives w et $1 - w$, $w \in [0, 1]$, est appelée FM simple, et peut être simplement notée A^w [73, 23]. Une telle FM est illustrée par la Figure 1.1.

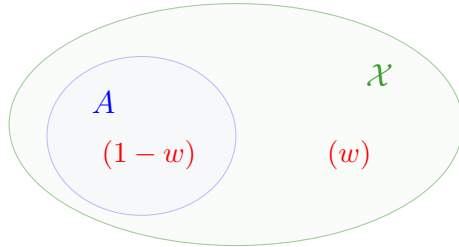


FIGURE 1.1 – Illustration des éléments focaux de la fonction de masse A^w .

La négation d'une fonction de masse simple peut être notée A_v avec v la masse sur l'ensemble vide. Elle possède deux éléments focaux : $\emptyset$ et $A \neq \emptyset$, avec pour masses respectives v et $1-v$, $v \in [0, 1]$. Une telle FM est appelée *FM simple négative* [73, 23].

1.2.3 Autres fonctions pour représenter les connaissances

Une FM m est en correspondance biunivoque avec une fonction de crédibilité bel , une fonction de plausibilité pl , une fonction d'implicabilité b et une fonction de communalité q définies, pour tout $A \subseteq \mathcal{X}$, respectivement par :

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B), \quad (1.4)$$

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \quad (1.5)$$

$$b(A) = bel(A) + m(\emptyset) = 1 - pl(\bar{A}), \quad (1.6)$$

et

$$q(A) = \sum_{B \supseteq A} m(B). \quad (1.7)$$

Les fonctions m , bel , pl , b et q représentent toute la même information sous une forme différente (voir par exemple [75] pour une synthèse de toutes les correspondances).

La fonction bel représente, pour tout sous-ensemble A de $\mathcal{X}$, le degré total de croyance supportant l'hypothèse $x \in A$, alors que pl représente la somme totale des croyances qui ne sont pas en contradiction avec A . Les fonctions bel et pl vérifient les propriétés suivantes :

- $bel(\emptyset) = pl(\emptyset) = 0$ et $bel(\mathcal{X}) \leq 1$ et $pl(\mathcal{X}) \leq 1$ (en considérant que la masse sur le vide peut être positive) ;
- $bel(A) \leq pl(A)$ et $pl(A) = 1 - bel(\bar{A})$ pour tout $A \subseteq \mathcal{X}$;
- Si m est bayésienne alors $bel = pl$ est une mesure de probabilité ;
- Si m est *consonante*, c'est-à-dire si m a ses éléments focaux emboîtés ($A_1 \subset A_2 \subset \dots \subset A_n$, avec A_i , $i \in \{1, 2, \dots, n\}$ les éléments focaux de m), alors pl est une mesure de possibilité et bel est la mesure duale de nécessité.

La fonction pl restreinte aux plausibilités sur les singletons, c'est-à-dire définie de $\mathcal{X}$ dans $[0, 1]$, est appelée *fonction contour*.

La fonction b représente le degré total de croyance supportant l'hypothèse $x \in A$ en incluant la masse sur l'ensemble vide, et nous avons les propriétés suivantes pour les fonctions b et q :

- $b(\emptyset) = m(\emptyset)$ et $b(\mathcal{X}) = 1$;
- $q(\emptyset) = 1$ et $q(\mathcal{X}) = m(\mathcal{X})$;
- $\bar{b}(A) = q(\bar{A})$ pour tout $A \subseteq \mathcal{X}$.

La fonction q est donc liée à la négation de la fonction b . La fonction q peut aussi être interprétée comme l'incertitude restant sur un ensemble après avoir conditionné sur cet ensemble, on a en effet $q(A) = mA$, pour tout $A \subseteq \mathcal{X}$, la définition du conditionnement $m[A]$ étant rappelé dans la section 1.3.4.1. Enfin, comme nous le rappellerons dans la section 1.3.1, les fonctions b et q ont respectivement un intérêt au niveau des calculs des combinaisons disjonctive et conjonctive.

1.2.4 Grossissement et raffinement

Le choix du cadre de discernement est toujours une histoire de convention comme le souligne Shafer [71, chapitre 6]. En effet, il est souvent possible de regrouper ou diviser des états du monde. Par exemple, on peut raisonner sur un cadre composé des individualités de personnes (Pierre, Paul, Marie, etc.) ou sur un cadre plus grossier constitué de deux éléments : hommes et femmes, ou encore sur un cadre composé des plus de cinquante ans, des trente-cinquante ans et des moins de trente ans. Aussi il est utile de considérer les concepts de *grossissement* et *raffinement* qui sont définies formellement de la façon suivante.

Étant donnés $\mathcal{Y}$ et $\mathcal{X}$ deux cadres de discernement, une fonction $\rho : 2^{\mathcal{Y}} \rightarrow 2^{\mathcal{X}}$ est appelée *raffinement* de $\mathcal{Y}$ si et seulement si :

1. l'ensemble $\{\rho(\{y\}), y \in \mathcal{Y}\} \subseteq 2^{\mathcal{X}}$ forme une partition de $\mathcal{X}$, et
2. pour tout $A \subseteq \mathcal{Y} : \rho(A) = \bigcup_{y \in A} \rho(\{y\})$.

Lorsqu'une telle application ρ existe, $\mathcal{Y}$ est appelé *grossissement* de $\mathcal{X}$, et $\mathcal{X}$ est appelé *raffinement* de $\mathcal{Y}$. La figure 1.2 illustre un exemple de grossissement et raffinement.

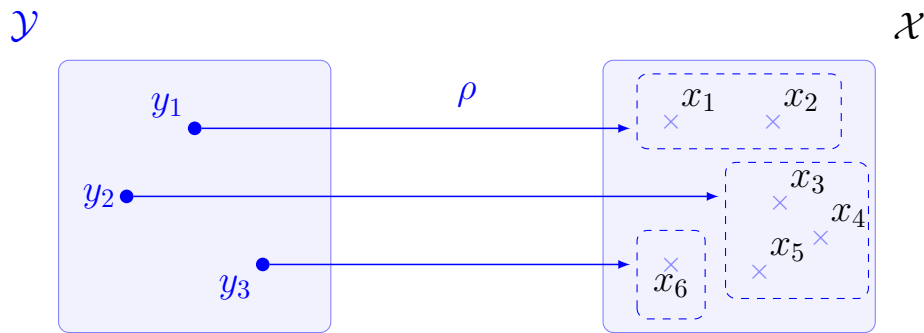


FIGURE 1.2 – Illustration d'un grossissement $\mathcal{Y}$ de $\mathcal{X}$ associé à un raffinement ρ de $\mathcal{Y}$.

1.3 Manipulation des connaissances

1.3.1 Combinaison d'informations

Deux fonctions de masse m_1 et m_2 issues de deux sources d'information fiables et distinctes, peuvent être combinées en utilisant la *Règle de Combinaison Conjonctive (RCC)* [78, 74, 77] ou *règle de Dempster non normalisée*, définie par :

$$(m_1 \odot m_2)(A) = m_1 \odot_2(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C), \quad \forall A \subseteq \mathcal{X}. \quad (1.8)$$

La règle de Dempster [71] correspond à la règle RCC suivie d'une étape de normalisation (1.3) :

$$(m_1 \oplus m_2)(A) = m_1 \oplus_2(A) = \begin{cases} \frac{m_1 \odot_2(A)}{1 - m_1 \odot_2(\emptyset)} & \text{si } \emptyset \neq A \subseteq \mathcal{X}, \\ 0 & \text{sinon.} \end{cases} \quad (1.9)$$

Si les sources sont distinctes mais seulement une des sources est fiable (et on ne sait pas laquelle), la *règle de combinaison disjonctive (RCD)* [26, 72] peut être appliquée. Elle est donnée, pour tout $A \subseteq \mathcal{X}$, par l'équation suivante :

$$(m_1 \ominus m_2)(A) = m_1 \ominus_2(A) = \sum_{B \cup C = A} m_1(B) \cdot m_2(C). \quad (1.10)$$

On peut mentionner les propriétés suivantes :

- RCC et RCD sont associatives, commutatives (ainsi peu importe l'ordre dans lequel sont combinées les informations) et ont pour élément neutre $m_{\mathcal{X}}$.
- RCC s'exprime simplement en termes de fonctions de communalité, on a en effet :

$$q_1 \odot_2 = q_1 \cdot q_2. \quad (1.11)$$

- RCD se calcule facilement en utilisant les fonctions d'implicabilité :

$$b_1 \ominus_2 = b_1 \cdot b_2. \quad (1.12)$$

1.3.2 Principe de minimum d'information

Le *principe de minimum d'information* issu du MCT [72] consiste à choisir, parmi un ensemble de fonctions de croyance compatibles avec les données d'une situation, la fonction de croyance la moins informative, c'est-à-dire celle qui n'ajoute pas d'informations arbitraires (non justifiées) aux données disponibles.

Une fois ce principe donné, il reste à définir la notion d'ordre entre fonctions de croyance qui permet de le mettre en place. Trois ordres partiels entre fonctions de croyance ont été donnés dans la littérature des années 80 par Yager [80], et Dubois et Prade [26, 27] :

- L'ordre suivant les plausibilités, appelé *ordre-pl* et noté $\sqsubseteq_{pl}$, tel que $m_1 \sqsubseteq_{pl} m_2$ (m_1 est plus informative/spécifique que m_2) si et seulement si $pl_1(A) \leq pl_2(A)$, pour tout $A \subseteq \mathcal{X}$.
- L'ordre suivant les communalités, appelé *ordre-q* et noté $\sqsubseteq_q$, tel que $m_1 \sqsubseteq_q m_2$ si et seulement si $q_1(A) \leq q_2(A)$, pour tout $A \subseteq \mathcal{X}$.
- L'ordre suivant la notion de *spécialisation* rappelée ci-après, appelé *ordre-s* et noté $\sqsubseteq_s$, tel que $m_1 \sqsubseteq_s m_2$ si et seulement si m_1 est une spécialisation de m_2 , c'est-à-dire qu'il existe une matrice $\mathbf{S}$ de terme générique $S(A, B)$, $A \subseteq \mathcal{X}$, $B \subseteq \mathcal{X}$, telle que $\sum_{B \subseteq \mathcal{X}} S(A, B) = 1$ pour tout $A \subseteq \mathcal{X}$, $S(A, B) > 0 \Rightarrow A \subseteq B$ pour tout $A \subseteq \mathcal{X}$ et $B \subseteq \mathcal{X}$, pour laquelle on a $m_1(A) = \sum_{B \subseteq \mathcal{X}} S(A, B)m_2(B)$ pour tout $A \subseteq \mathcal{X}$. La matrice $\mathbf{S}$ est appelée matrice de spécialisation [43, 75]. Chaque terme $S(A, B)$ représente la part de masse $m_2(B)$ qui descend vers $m_1(A)$ avec $A \subseteq B$.

Ces trois ordres ne sont pas équivalents, on a uniquement les relations suivantes [26] : $m_1 \sqsubseteq_s m_2 \Rightarrow m_1 \sqsubseteq_{pl} m_2$ et $m_1 \sqsubseteq_s m_2 \Rightarrow m_1 \sqsubseteq_q m_2$. Néanmoins tous généralisent la notion d'ordre entre ensembles : pour tout $A \subseteq \mathcal{X}$ et $B \subseteq \mathcal{X}$, $A \subseteq B \Leftrightarrow m_A \sqsubseteq m_B$ où $\sqsubseteq$ est n'importe quel ordre $\sqsubseteq_{pl}$, $\sqsubseteq_q$ ou $\sqsubseteq_s$.

On peut mentionner que d'autres relations d'ordre plus fortes existent comme celles basées sur la spécialisation Dempstérienne ($m_1 = m \odot m_2$) [43]) ou sur les fonctions de poids conjonctif w et disjonctif v [23] dont les définitions sont rappelées dans la section 2.2 du chapitre 2.

Le principe du minimum d'information permet en particulier d'introduire deux opérations dans le cadre des fonctions de croyance :

- l'*extension vide* : opération réalisant l'inverse de la projection ou marginalisation ;
- le *déconditionnement* : opération réalisant l'inverse du conditionnement.

Celles-ci sont rappelées dans les sections 1.3.3 et 1.3.4 suivantes.

1.3.3 Marginalisation et extension vide

1.3.3.1 Définitions

L'opération de *marginalisation* consiste à projeter une FM $m^{\mathcal{X}}$ dans un référentiel plus grossier $\mathcal{Y}$. Elle est notée avec une flèche vers le bas et est définie par :

$$m^{\mathcal{X} \downarrow \mathcal{Y}}(A) = \sum_{\{B \subseteq \mathcal{X} \mid \text{proj}(B \downarrow \mathcal{Y}) = A\}} m^{\mathcal{X}}(B), \forall A \subseteq \mathcal{Y}, \quad (1.13)$$

où $\text{proj}(B \downarrow \mathcal{Y})$ est le sous-ensemble de $\mathcal{Y}$ résultant de la projection de B sur $\mathcal{Y}$, c'est-à-dire l'image inverse de B par le raffinement ρ de $\mathcal{Y}$ vers $\mathcal{X}$ (cf section 1.2.4).

Généralement on ne peut pas retrouver une FM connaissant sa marginalisée. Néanmoins, il est possible d'appliquer le principe de minimum d'information et de trouver la FM la moins informative parmi les FM d'avant marginalisation compatibles avec la marginalisée. Cette opération est appelée extension vide. Notée avec

une flèche vers le haut, elle permet de transporter une FM $m^{\mathcal{Y}}$ dans un référentiel plus fin $\mathcal{X}$ et est définie par [71, 72] :

$$m^{\mathcal{Y} \uparrow \mathcal{X}}(\rho(A)) = m^{\mathcal{Y}}(A), \quad \forall A \subseteq \mathcal{Y}, \quad (1.14)$$

avec ρ le raffinement de $\mathcal{Y}$ vers $\mathcal{X}$.

1.3.3.2 Cas d'un espace produit

Dans le cas d'un espace produit, c'est-à-dire avec $\mathcal{Y} = \mathcal{X}_1$ et $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$, les définitions précédentes deviennent les suivantes.

Une FM $m^{\mathcal{X}_1 \times \mathcal{X}_2}$ définie sur un espace produit $\mathcal{X}_1 \times \mathcal{X}_2$ peut être marginalisée sur $\mathcal{X}_1$ par le transfert de chaque masse $m^{\mathcal{X}_1 \times \mathcal{X}_2}(B)$ vers la projection de B sur $\mathcal{X}_1$:

$$m^{\mathcal{X}_1 \times \mathcal{X}_2 \downarrow \mathcal{X}_1}(A) = \sum_{\{B \subseteq \mathcal{X}_1 \times \mathcal{X}_2 \mid \text{proj}(B \downarrow \mathcal{X}_1) = A\}} m^{\mathcal{X}_1 \times \mathcal{X}_2}(B), \quad \forall A \subseteq \mathcal{X}_1, \quad (1.15)$$

avec $\text{proj}(B \downarrow \mathcal{X}_1)$ le sous-ensemble de $\mathcal{X}_1$ résultat de la projection de B sur $\mathcal{X}_1$. Et l'extension vide d'une FM $m^{\mathcal{X}_1}$ sur $\mathcal{X}_1 \times \mathcal{X}_2$ est définie par :

$$m^{\mathcal{X}_1 \uparrow \mathcal{X}_1 \times \mathcal{X}_2}(B) = \begin{cases} m^{\mathcal{X}_1}(A) & \text{si } B = A \times \mathcal{X}_2, A \subseteq \mathcal{X}_1 \\ 0 & \text{sinon.} \end{cases} \quad (1.16)$$

Le transfert de masses intervenant lors d'une marginalisation (resp. extension vide) dans le cas d'un espace produit est illustré sur la figure 1.3(a) (resp. 1.3(b)). La figure 1.3(b) illustre bien le fait que dans le cas de l'extension vide, l'application du principe de minimum d'information consiste à chercher le plus grand sous-ensemble B de $\mathcal{X}_1 \times \mathcal{X}_2$ qui projeté sur $\mathcal{X}_1$ donne A .

1.3.4 Conditionnement et déconditionnement

1.3.4.1 Définitions

Le conditionnement d'une FM $m^{\mathcal{X}}$ sur un sous-ensemble B de $\mathcal{X}$ signifie la prise en compte d'une information selon laquelle il est certain que le sous-ensemble B de $\mathcal{X}$ contient la valeur d'intérêt cherchée. Le conditionnement $m^{\mathcal{X}}[B]$ d'une FM $m^{\mathcal{X}}$ sur $B \subseteq \mathcal{X}$ est donc simplement défini par la combinaison conjonctive de $m^{\mathcal{X}}$ avec la FM logique sur B :

$$m^{\mathcal{X}}[B] = m^{\mathcal{X}} \odot m_B. \quad (1.17)$$

Autrement dit, chaque masse $m^{\mathcal{X}}(A)$, $A \subseteq \mathcal{X}$, est transférée sur $A \cap B$.

Le déconditionnement est l'opération inverse du conditionnement. Étant donnée une FM conditionnelle $m^{\mathcal{X}}[B]$, il n'est généralement pas possible de retrouver la FM originelle d'avant conditionnement $m^{\mathcal{X}}$. Cependant, la FM la moins informative qui, conditionnée sur B , donnera $m[B]$, peut être calculée en transférant chaque masse $m[B](A)$, $A \subseteq B$, sur le plus grand ensemble de $\mathcal{X}$ qui lors du conditionnement sur

B verra sa masse transférée sur A . Il est facile de voir que ce plus grand ensemble est $A \cup \overline{B}$. Ainsi le déconditionnement de $m[B]$ sur $\mathcal{X}$, notée $m[B]^{\uparrow\mathcal{X}}$ (avec une double flèche vers le haut), est défini par :

$$m[B]^{\uparrow\mathcal{X}}(A \cup \overline{B}) = m[B](A), \forall A \subseteq B. \quad (1.18)$$

Les cas particuliers du conditionnement et déconditionnement dans un espace produit sont abordés dans le paragraphe suivant.

1.3.4.2 Cas d'un espace produit

Si $m^{\mathcal{X}_1 \times \mathcal{X}_2}$ est définie sur un espace produit $\mathcal{X}_1 \times \mathcal{X}_2$, et D est un sous-ensemble de $\mathcal{X}_2$, alors le conditionnement $m^{\mathcal{X}_1}[D]$ est définie par la combinaison de $m^{\mathcal{X}_1 \times \mathcal{X}_2}$ avec $m_D^{\mathcal{X}_2 \uparrow \mathcal{X}_1 \times \mathcal{X}_2}$ marginalisée sur $\mathcal{X}_1$:

$$m^{\mathcal{X}_1}[D] = \left(m^{\mathcal{X}_1 \times \mathcal{X}_2} \odot m_D^{\mathcal{X}_2 \uparrow \mathcal{X}_1 \times \mathcal{X}_2} \right)^{\downarrow \mathcal{X}_1}. \quad (1.19)$$

Quant au déconditionnement, il est définie par :

$$m^{\mathcal{X}_1}[D]^{\uparrow \mathcal{X}_1 \times \mathcal{X}_2}(A \times D \cup \mathcal{X}_1 \times \overline{D}) = m_1^{\mathcal{X}}[D](A), \quad \forall A \subseteq \mathcal{X}_1. \quad (1.20)$$

Le transfert de masses intervenant lors d'un tel conditionnement (resp. déconditionnement) est illustré sur la figure 1.3(c) (resp. 1.3(d)).

1.3.5 Mécanismes de correction de fonction de croyance

1.3.5.1 Affaiblissement

L'opération d'affaiblissement d'une FM $m^{\mathcal{X}}$ induite par une source S est définie dans [71, page 252] par :

$$\begin{cases} {}^{\alpha}m(A) &= (1 - \alpha)m(A), \quad \forall A \subset \mathcal{X}, \\ {}^{\alpha}m(\mathcal{X}) &= (1 - \alpha)m(\mathcal{X}) + \alpha, \end{cases} \quad (1.21)$$

où $\alpha \in [0, 1]$ est le taux d'affaiblissement.

Elle peut simplement s'écrire en utilisant la notation des FM logiques (cf section 1.2.2) de la façon suivante :

$${}^{\alpha}m = (1 - \alpha)m + \alpha m_{\mathcal{X}}. \quad (1.22)$$

De même, nous pouvons remarquer que l'affaiblissement résulte d'une combinaison disjonctive de m avec une FM simple négative :

$${}^{\alpha}m = m \odot \mathcal{X}_{1-\alpha}. \quad (1.23)$$

L'affaiblissement permet de prendre en compte la fiabilité de la source S représentée par le taux de fiabilité $\beta = 1 - \alpha$, $\beta \in [0, 1]$. Si la source est fiable, $\beta = 1$, alors l'information qu'elle a fourni est totalement gardée : ${}^\alpha m = m$. Si la source est non fiable, $\beta = 0$, alors l'information fournie est totalement écartée : ${}^\alpha m = m_{\mathcal{X}}$ (ignorance totale). Dans les cas intermédiaires, $\beta \in]0, 1[$, ${}^\alpha m$ est alors une somme pondérée de m et $m_{\mathcal{X}}$.

Comme l'a montré Smets dans [72, Section 5.7], l'affaiblissement d'une FM m_S issue d'une source S peut être obtenue à partir d'un modèle simple de fiabilité. Soit $\mathcal{R} = \{F, NF\}$ un cadre composé de deux éléments F et NF , tels que F indique que la source est fiable, et NF qu'elle est non fiable. On dispose alors de l'information suivante, $\alpha \in [0, 1]$:

$$\begin{cases} m^{\mathcal{R}}(\{F\}) &= 1 - \alpha \\ m^{\mathcal{R}}(\mathcal{R}) &= \alpha . \end{cases} \quad (1.24)$$

L'équation (1.24) indique que la source est fiable avec un degré $\beta = 1 - \alpha \in [0, 1]$.

La fiabilité d'une source est alors modélisée de la façon suivante :

- Si la source S est fiable, nous gardons l'information fournie par S : $m^{\mathcal{X}}[\{F\}] = m_S^{\mathcal{X}}$;
- Si la source S est non fiable, l'information fournie par S est écartée, et nous avons $m^{\mathcal{X}}[\{NF\}](\mathcal{X}) = 1$.

L'équation d'affaiblissement (1.22) résulte alors [72, Section 5.7][62] de l'opération suivante :

$${}^\alpha m = \left(m^{\mathcal{X}}[\{NF\}]^{\uparrow \mathcal{X} \times \mathcal{R}} \oplus m^{\mathcal{X}}[\{F\}]^{\uparrow \mathcal{X} \times \mathcal{R}} \oplus m^{\mathcal{R} \uparrow \mathcal{X} \times \mathcal{R}} \right)^{\downarrow \mathcal{X}} . \quad (1.25)$$

1.3.5.2 Affaiblissement contextuel basé sur un grossissement

L'opération d'affaiblissement (1.22) a été étendue durant mes travaux de thèse [51, 62, 53] pour pouvoir tenir compte d'informations plus fines sur la fiabilité de la source, en l'occurrence la source peut avoir une fiabilité différente en fonction de la véritable valeur x_0 de la variable d'intérêt. Plus généralement, le degré de fiabilité β de la source peut être différent en fonction du sous-ensemble A de $\mathcal{X}$, appelé *contexte*, dans lequel se trouve x_0 . Par exemple, le diagnostic d'un médecin peut être plus ou moins fiable en fonction de la maladie ou du type de maladie d'un patient. De même, un capteur de température peut avoir une fiabilité différente en fonction de la température réelle. L'opération d'affaiblissement développée est alors contrôlée par un vecteur de taux de fiabilité $\beta = (\beta_i, i \in \{1, \dots, N\})$ où chaque β_i , $i \in \{1, \dots, N\}$, représente la fiabilité de la source sachant que $x_0 \in A_i \subseteq \mathcal{X}$, l'ensemble $\mathcal{A} = \{A_i \subseteq \mathcal{X}, i \in \{1, \dots, N\}\}$ devant former une partition de $\mathcal{X}$.

Cette extension, appelée *affaiblissement contextuel basé sur un grossissement*, est alors obtenue à partir d'un ensemble $\mathcal{A}$ formant une partition de $\mathcal{X}$, sur lequel

la fiabilité de la source est exprimé, pour tout $A \in \mathcal{A}$, par :

$$\begin{cases} m^{\mathcal{R}}[A](\{F\}) &= 1 - \alpha_A = \beta_A , \\ m^{\mathcal{R}}[A](\mathcal{R}) &= \alpha_A , \end{cases} \quad (1.26)$$

où $\alpha_A \in [0, 1]$. La correction d'une FM m_S fournie par une source S connaissant cette information (1.26) est alors donnée par :

$$\left(m^{\mathcal{X}}[\{NF\}] \uparrow^{\mathcal{X} \times \mathcal{R}} \odot m^{\mathcal{X}}[\{F\}] \uparrow^{\mathcal{X} \times \mathcal{R}} \odot_{A \in \mathcal{A}} m^{\mathcal{R}}[A] \uparrow^{\mathcal{X} \times \mathcal{R}} \right)^{\downarrow \mathcal{X}} . \quad (1.27)$$

En notant simplement m la FM fournie par la source, le résultat de cette opération d'affaiblissement contextuel de m basé sur le grossissement $\mathcal{A}$ dépend uniquement de m et du vecteur de taux d'affaiblissement $\alpha = (\alpha_A, A \in \mathcal{A})$. Il est ainsi noté ${}^{\alpha}m$, et on a [62, Proposition 8][56] :

$${}^{\alpha}m = m \odot_{A \in \mathcal{A}} A_{\beta_A} . \quad (1.28)$$

Avec $\mathcal{A}$ composé d'un seul contexte $\mathcal{X}$, on retrouve l'équation (1.23) de l'affaiblissement classique (1.23).

Les résultats obtenus dans le cas général où $\mathcal{A}$ ne forme pas nécessairement une partition de $\mathcal{X}$ sont donnés dans la section 2.3 du prochain chapitre.

1.3.5.3 Autre mécanisme de correction

Nous présentons ici un autre mécanisme de correction généralisant l'affaiblissement (1.22) qui a aussi été développé durant mes travaux de thèse [51, 54].

Dans l'affaiblissement, nous avons considéré que la source pouvait se trouver dans deux états : fiable (F) et non fiable (NF), tels que $m^{\mathcal{X}}[\{F\}] = m_S^{\mathcal{X}}$ et $m^{\mathcal{X}}[\{NF\}] = m_{\mathcal{X}}$. Le mécanisme de correction introduit généralise ces hypothèses en considérant que la source peut se trouver dans N états R_i , $i \in \{1, \dots, N\}$ dont les significations sont donnés par N transformations m_i de m_S , tels que $m^{\mathcal{X}}[\{R_i\}] = m_i$, pour tout $i \in \{1, \dots, N\}$. En posant $\mathcal{R} = \cup_i \{R_i\}$ et en supposant que l'information dont nous disposons sur l'état dans lequel se trouve la source est donnée par $m^{\mathcal{R}}(\{R_i\}) = \nu_i$, $\forall i \in \{1, \dots, N\}$ avec $\sum_i \nu_i = 1$, l'information résultante de ce modèle peut alors être calculée par l'équation suivante :

$$\left(\odot_i m^{\mathcal{X}}[\{R_i\}] \uparrow^{\mathcal{X} \times \mathcal{R}} \odot m^{\mathcal{R}} \uparrow^{\mathcal{X} \times \mathcal{R}} \right)^{\downarrow \mathcal{X}} . \quad (1.29)$$

Comme démontré dans [51, Proposition 4.10] [54, Proposition 2], le résultat de l'équation (1.29) dépend uniquement du vecteur $\nu = (\nu_i, i \in \{1, \dots, N\})$. Il est noté ${}^{\nu}m$ et est donné par l'équation :

$${}^{\nu}m = \sum_i \nu_i m_i . \quad (1.30)$$

On retrouve une somme convexe comme dans l'équation (1.22) de l'affaiblissement qui est naturellement retrouvé en choisissant deux états R_1 et R_2 tels que $\mathcal{R} = \{R_1, R_2\}$ avec $m^{\mathcal{X}}[\{R_1\}] = m_S^{\mathcal{X}}$ et $m^{\mathcal{X}}[\{R_2\}] = m_{\mathcal{X}}$. On retrouve également [51, 54] d'autres corrections de la littérature comme le dé-affaiblissement introduit par Dencœux et Smets dans [25], et les corrections de Zhu et Basir introduites dans [81].

Nous avons employé un de ces mécanismes dans l'application de communication inter-véhicule sur laquelle nous avons travaillé [6, 11], et qui est présentée dans la section 3.2 de ce rapport. Nous avons choisi deux états R_1 et R_2 tels que si la source est dans l'état R_1 on garde l'information donnée par la source, on a totalement confiance dans ce qu'elle dit : $m^{\mathcal{X}}[\{R_1\}] = m_S^{\mathcal{X}}$, et si la source est dans l'état R_2 on sait que la vérité est $x_k \in \mathcal{X}$ quel que soit ce que la source nous dit : $m^{\mathcal{X}}[\{R_2\}] = m_{\{x_k\}}$. Nous avons donc appelé ce mécanisme *renforcement vers un élément singleton*, et il est défini par :

$${}^{\nu}m = (1 - \nu)m + \nu m_{\{x_k\}} , \quad (1.31)$$

avec $\nu \in [0, 1]$.

Remarquons que ce *renforcement vers un élément singleton* n'est pas un *renforcement* au sens où ${}^{\nu}m$ serait plus commise/informative que m (cf section 1.3.2). Ce n'est généralement pas le cas.

1.4 Prise de décision

Lorsqu'une décision doit être prise, des *axiomes de rationalité* [17] justifient la stratégie consistant à choisir la décision d parmi un ensemble de décisions possibles $\mathcal{D}$ minimisant le *risque espéré* défini par :

$$R(d) = \sum_{x_k \in \mathcal{X}} c(d, x_k) P^{\mathcal{X}}(\{x_k\}), \quad (1.32)$$

où $P^{\mathcal{X}} : 2^{\mathcal{X}} \rightarrow [0, 1]$ est une mesure de probabilité et $c : \mathcal{D} \times \mathcal{X} \rightarrow \mathbb{R}$ une fonction de coût, $c(d, x_k)$ représentant le coût de décider d sachant que la vérité est x .

En acceptant ces axiomes, au moment de prendre une décision, la fonction de croyance éventuellement obtenue après combinaison doit donc être transformée en une mesure de probabilité. Une solution [22, 74] consiste à utiliser la *transformation pignistique* en calculant la *probabilité pignistique* [76] définie par :

$$\text{Bet}P^{\mathcal{X}}(\{x_k\}) = \sum_{\{A \subseteq \mathcal{X}, x_k \in A\}} \frac{m(A)}{|A| (1 - m(\emptyset))}, \quad \forall x_k \in \mathcal{X} . \quad (1.33)$$

La décision choisie est alors celle qui minimise le risque pignistique défini par :

$$R_{\text{Bet}}(d) = \sum_{x_k \in \mathcal{X}} c(d, x_k) \text{Bet}P^{\mathcal{X}}(\{x_k\}) . \quad (1.34)$$

Dans le cas de coûts 0-1 avec $\mathcal{D} = \mathcal{X}$, c'est-à-dire $c(x_k, x_{k'}) = 1$ si $k = k'$, 0 sinon, la minimisation du risque pignistique (1.34) revient à choisir la décision de probabilité pignistique maximale.

Il est aussi possible d'ajouter à l'ensemble $\mathcal{D}$ précédent une décision de rejet d_0 . Celle-ci permet de ne pas prendre une autre des décisions lorsque le risque est trop important. Avec des coûts 0-1, $\mathcal{D} = \mathcal{X} \cup \{d_0\}$ et $c_0 = c(d_0, x_k) \forall k \in \{1, \dots, K\}$, la minimisation du risque pignistique (1.34) revient à choisir la décision :

$$\begin{aligned} d_0 \quad \text{si} \quad & \max_{k=1, \dots, K} \text{Bet}P(\{x_k\}) < 1 - c_0 \\ x_{k'} \quad \text{si} \quad & \text{Bet}P(\{x_{k'}\}) = \max_{k=1, \dots, K} \text{Bet}P(\{x_k\}) \geq 1 - c_0 . \end{aligned} \quad (1.35)$$

Il existe également une autre transformation [16, 2, 31] d'une fonction de croyance en mesure de probabilité. Elle est basée sur la fonction contour et est définie par :

$$plP^{\mathcal{X}}(\{x\}) = K^{-1} \cdot pl(\{x\}), \quad \forall x \in \mathcal{X} , \quad (1.36)$$

avec $K = \sum_{x \in \mathcal{X}} pl(\{x\})$ une constante de normalisation.

La transformation pignistique satisfait la propriété de linéarité suivante [76] :

$$\text{Bet}P(\alpha m_1 + (1 - \alpha)m_2) = \alpha \text{Bet}P(m_1) + (1 - \alpha) \text{Bet}P(m_2) , \quad (1.37)$$

tandis que la transformation basée sur la fonction contour [16] satisfait :

$$plP_{1 \odot 2} \propto plP_1 \cdot plP_2 . \quad (1.38)$$

Cette dernière propriété indique donc qu'il n'y a pas besoin d'effectuer la combinaison conjonctive pour obtenir la décision par maximum de plausibilité.

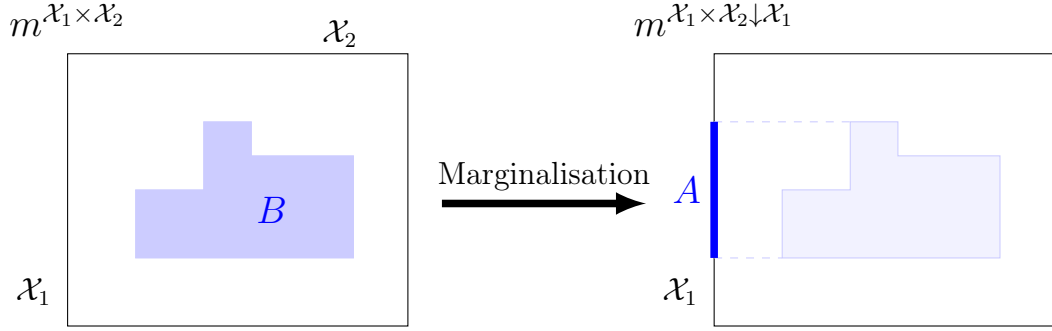
1.5 Conclusion

Les principaux concepts des fonctions de croyance que nous avons employés dans nos travaux ont été présentés dans ce chapitre. Il nous faut encore présenter les décompositions canoniques conjonctive et disjonctive en éléments simples qui sont vues au début du prochain chapitre au moment où nous les utilisons.

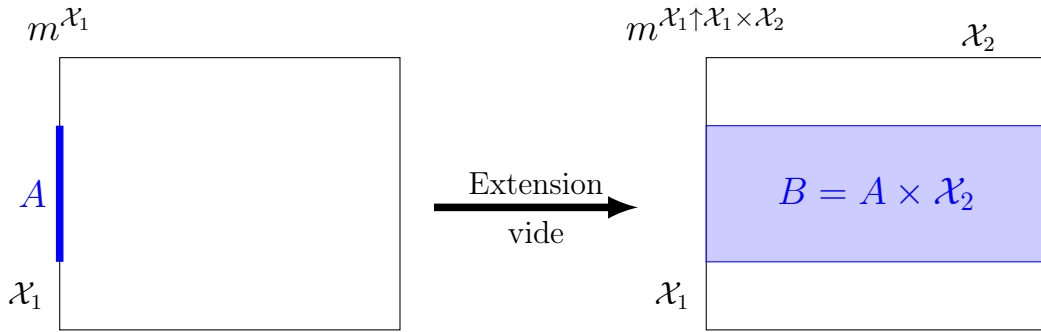
Comme nous contribuons à l'illustrer dans la suite de ce rapport, le cadre des fonctions de croyance permet de simplement modéliser des informations imprécises et incertaines, néanmoins toutes les méthodes employant le cadre des fonctions de croyance n'aboutissent pas aux mêmes performances. Ce qui compte pour un problème donné, ce sont les informations que l'on choisit de modéliser et la façon de les modéliser. La section 3.3 consacré au problème d'association de données en constitue un bon exemple.

Dans le prochain chapitre 2, mes nouveaux travaux effectués depuis ma thèse autour des développements des mécanismes d'affaiblissement et de renforcement

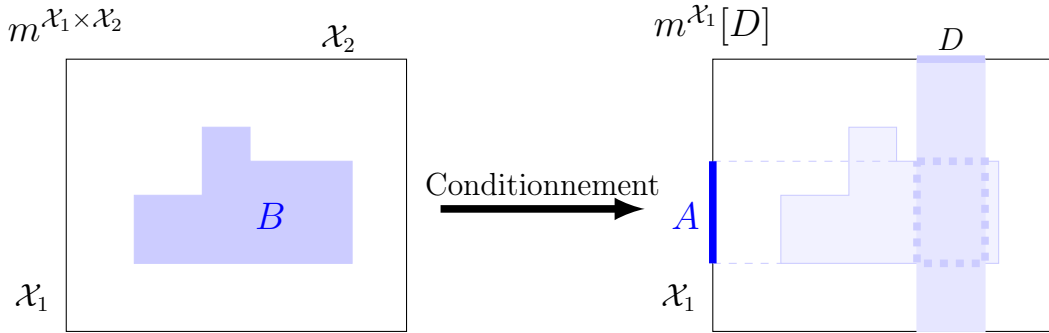
contextuels sont résumés.



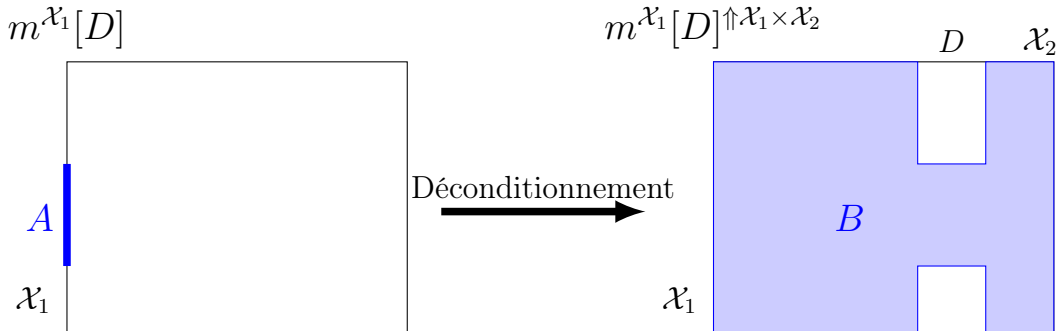
(a) Illustration (inspirée de [20]) du transfert de masse d'un ensemble $B \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ vers $A \subseteq \mathcal{X}_1$ lors de la **marginalisation** d'une FM $m^{\mathcal{X}_1 \times \mathcal{X}_2}$ sur $\mathcal{X}_1$.



(b) Illustration (inspirée de [20]) du transfert de masse d'un ensemble $A \subseteq \mathcal{X}_1$ sur $B \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ lors de l'opération d'**extension vide** d'une FM $m^{\mathcal{X}_1}$ sur $\mathcal{X}_1 \times \mathcal{X}_2$.



(c) Illustration (inspirée de [20]) du transfert de masse d'un ensemble $B \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ vers $A \subseteq \mathcal{X}_1$ lors du **conditionnement** d'une FM $m^{\mathcal{X}_1 \times \mathcal{X}_2}$ sur $D \subseteq \mathcal{X}_2$.



(d) Illustration (inspirée de [20]) du transfert de masse d'un ensemble $A \subseteq \mathcal{X}_1$ vers $B \subseteq \mathcal{X}_1 \times \mathcal{X}_2$ lors du **déconditionnement** d'une FM $m^{\mathcal{X}_1}[D]$ sur $\mathcal{X}_1 \subseteq \mathcal{X}_2$.

FIGURE 1.3 – Illustrations (inspirées de [20]) des transferts de masse lors des opérations de marginalisation, extension vide, conditionnement et déconditionnement dans le cas d'un espace produit.

Affaiblissement et renforcement contextuels

2.1 Introduction

Depuis les travaux sur l'affaiblissement contextuel basé sur un grossissement effectués durant ma thèse [51, 62] et résumés dans la section 1.3.5.2, nous avons cherché à étendre cet affaiblissement à un ensemble quelconque de contextes [56, 66], à explorer de nouveaux mécanismes de correction contextuelle [66] dont le renforcement contextuel [56, 66], et à mettre en évidence l'intérêt pratique de ces mécanismes [61].

Un résumé de nos contributions à ce jour est donné dans ce chapitre. Le terme *à ce jour* est à prendre en compte, car il s'agit d'un sujet sur lequel nous avons plusieurs travaux en cours.

Récemment, et contrairement à ce nous pensions dans [56], nous avons constaté que l'équation (1.27) n'aboutissait pas à l'équation (1.28) dans le cas général où $\mathcal{A}$ n'est plus une partition de $\mathcal{X}$. C'est en fait grâce au mécanisme de correction BBC (*Behaviour-based Correction*) issu des travaux de Pichon et al. [65], que nous avons trouvé un autre moyen [66] de dériver l'affaiblissement contextuel au sens de l'équation (1.28) pour un ensemble quelconque de contextes $\mathcal{A}$ de $\mathcal{X}$. Nous détaillons cela dans la section 2.3 de ce chapitre.

Le mécanisme dual de cet affaiblissement qui est le renforcement contextuel a été introduit dans [56] à partir de la négation de l'équation (1.28). Nous en avons trouvé une interprétation à partir du BBC dans [66].

Concernant l'intérêt pratique de ces mécanismes, nous avons montré qu'il était possible d'apprendre automatiquement les paramètres de ces mécanismes à partir de données étiquetées dans [61], ce qui permet par exemple d'entrevoir des perspectives en classification et en fusion d'informations.

Le plan de ce chapitre est le suivant. Nous commençons par rappeler en section 2.2 les concepts liés aux décompositions canoniques conjonctive et disjonctive. Nous abordons l'affaiblissement et le renforcement contextuels en sections 2.3 et 2.4, puis l'apprentissage de leurs coefficients respectifs avec une expérimentation en classification dans la section 2.5. Enfin, une conclusion et des perspectives à ce travail sont données dans la section 2.6.

2.2 Décompositions canoniques conjonctive et disjonctive

La décomposition canonique d'une FM consiste à décomposer cette fonction en une combinaison conjonctive $\odot$, RCC (1.8), ou disjonctive $\oplus$, RCD (1.8), de FM simples. Cette décomposition conjonctive ou disjonctive permet ainsi de représenter une information complexe sous forme d'*éléments* simples. Comme nous le détaillons dans cette section, ces décompositions permettent également d'obtenir deux nouvelles fonctions équivalentes aux fonctions m , pl , bel , q et b , qui sont la fonction de poids conjonctifs w et la fonction de poids disjonctif v . Enfin l'intérêt pour ces décompositions a aussi été accru avec l'introduction des combinaisons prudente $\oslash$ et hardie $\oslash$, introduites par Denœux [23], qui sont basées sur ces décompositions.

Étendant la notion de fonction séparable introduite par Shafer [71, chapter 4], Smets [73] a montré que n'importe quelle FM non dogmatique peut être décomposée de manière unique en *FM simples généralisées* notées $A^{w(A)}$, avec $A \subset \mathcal{X}$, et définies de $2^{\mathcal{X}}$ dans $\mathbb{R}$ par :

$$\begin{aligned} A^{w(A)} : \mathcal{X} &\mapsto w(A) \\ A &\mapsto 1 - w(A) \\ B &\mapsto 0, \forall B \in 2^{\mathcal{X}} \setminus \{A, \mathcal{X}\}, \end{aligned} \quad (2.1)$$

avec $w(A) \in [0, \infty[$.

La *décomposition canonique conjonctive* d'une FM non dogmatique m est alors donnée par :

$$m = \bigodot_{A \subset \mathcal{X}} A^{w(A)}. \quad (2.2)$$

La fonction $w : 2^{\mathcal{X}} \setminus \{\mathcal{X}\} \rightarrow]0, \infty[$ constitue une nouvelle représentation d'une FM non dogmatique. Elle est appelée *fonction de poids conjonctifs*. Remarquons que chaque poids conjonctif $w(A)$ représente l'incertitude dans la FM simple généralisée $A^{w(A)}$.

Dans [23], Denœux montre qu'une FM non normale m peut être décomposée de manière unique en FM simples négatives généralisées, notées $A_{v(A)}$, $A \neq \emptyset$, et définies de $2^{\mathcal{X}}$ dans $\mathbb{R}$ par :

$$\begin{aligned} A_{v(A)} : \emptyset &\mapsto v(A) \\ A &\mapsto 1 - v(A) \\ B &\mapsto 0, \forall B \in 2^{\mathcal{X}} \setminus \{\emptyset, A\}, \end{aligned} \quad (2.3)$$

avec $v(A) \in [0, \infty[$.

La *décomposition canonique disjonctive* d'une FM non normale m est alors définie ainsi :

$$m = \bigoplus_{A \subset \mathcal{X}, A \neq \emptyset} A_{v(A)}. \quad (2.4)$$

Comme la négation d'une FM non normale m est non dogmatique, on a :

$$\overline{m} = \bigcap_{A \subset \mathcal{X}} A^{\overline{w}(A)} = \overline{\bigcap_{A \subset \mathcal{X}} A^{w(A)}} = \bigcup_{A \subset \mathcal{X}} \overline{A^{w(A)}} = \bigcup_{A \supset \emptyset} A_{\overline{w}(\overline{A})} . \quad (2.5)$$

Ainsi, les fonctions v et w sont liées par la relation $v(A) = \overline{w}(\overline{A})$ pour tout $A \neq \emptyset$. La fonction $v : 2^\Omega \setminus \{\emptyset\} \rightarrow]0, \infty[$ est appelée *fonction de poids disjonctifs*, et se trouve être encore une autre représentation d'une fonction de croyance.

Finalement, les fonctions v et w vérifient les relations suivantes [23] :

– Pour tout w :

$$\prod_{A \subset \Omega} w(A) = m(\Omega) , \quad (2.6)$$

– Pour tout v :

$$\prod_{A \supset \emptyset} v(A) = m(\emptyset) , \quad (2.7)$$

– Pour tout $A \subset \Omega$: $A^{w_1(A)} \bigcap A^{w_2(A)} = A^{w_1(A)w_2(A)}$,

– Pour tout $A \supset \emptyset$: $A_{v_1(A)} \bigcup A_{v_2(A)} = A_{v_1(A)v_2(A)}$,

– La fonction w peut être facilement obtenue à partir de la fonction de communalité q ainsi :

$$\forall A \subset \Omega, \quad w(A) = \frac{\prod_{B \supseteq A, \text{par}(|A|) \neq \text{par}(|B|)} q(B)}{\prod_{B \supseteq A, \text{par}(|A|) = \text{par}(|B|)} q(B)} , \quad (2.8)$$

où $\text{par}(n)$ signifie la parité de l'entier n ($\text{par}(n) = 0$ si n est pair, 1 sinon),

– La fonction v peut être simplement obtenue de la fonction d'implicabilité b ainsi :

$$\forall A \supset \emptyset, \quad v(A) = \frac{\prod_{B \subseteq A, \text{par}(|A|) \neq \text{par}(|B|)} b(B)}{\prod_{B \subseteq A, \text{par}(|A|) = \text{par}(|B|)} b(B)} , \quad (2.9)$$

Nous résumons dans les deux sections suivantes nos contributions concernant l'affaiblissement et le renforcement contextuels.

2.3 Affaiblissement contextuel

Nous synthétisons ici nos résultats à ce jour [56, 66] concernant l'affaiblissement contextuel d'une fonction de croyance.

Nous commençons par poser la définition suivante.

Définition 2.1 (Affaiblissement contextuel) Soit $\mathcal{A}$ un ensemble de contextes de $\mathcal{X}$ (c'est-à-dire un ensemble de sous-ensembles de $\mathcal{X}$). Pour chaque $A \in \mathcal{A}$, on associe un paramètre réel $\beta_A \in [0, 1]$. L'affaiblissement contextuel $\alpha_\cup m$ d'une FM m est alors définie par :

$$\alpha_\cup m = m \bigcup_{A \in \mathcal{A}} A_{\beta_A} . \quad (2.10)$$

Une interprétation des paramètres β_A a donc été fournie dans la section 1.3.5.2 dans le cas où $\mathcal{A}$ forme une partition de $\mathcal{X}$, l'équation (1.27) revenant à l'équation (2.10). Chaque coefficient β_A , $A \in \mathcal{A}$, représente le degré de fiabilité de la source sachant que la vérité x_0 appartient à A , autrement dit, la fiabilité dans le contexte A . Plus précisément, une source fiable est caractérisée par $m^{\mathcal{X}}[\{F\}] = m_S$ avec m_S la FM fournie par la source, une source non fiable par $m^{\mathcal{X}}[\{NF\}](\mathcal{X}) = 1$, et le degré de fiabilité de la source β_A dépend du contexte $A \in \mathcal{A}$: $m^{\mathcal{R}}[A](\{F\}) = 1 - \alpha_A = \beta_A$ et $m^{\mathcal{R}}[A](\mathcal{R}) = \alpha_A$, pour tout $A \in \mathcal{A}$.

Dans [56], nous nous sommes intéressés à ce que devenait l'équation (1.27), qui peut être simplement écrite

$$\left(m^{\mathcal{X}}[\{F\}]^{\uparrow \mathcal{X} \times \mathcal{R}} \odot_{A \in \mathcal{A}} m^{\mathcal{R}}[A]^{\uparrow \mathcal{X} \times \mathcal{R}} \right)^{\downarrow \mathcal{X}} \quad (2.11)$$

puisque $m^{\mathcal{X}}[\{NF\}]$ correspond à une ignorance totale, lorsque l'ensemble $\mathcal{A}$ ne formait pas une partition. Nous avons alors affirmé dans [56][Théorème 1] que même dans cette situation, les équations (2.11) et (2.10) étaient équivalentes ce qui, nous l'avons découvert très récemment, n'est pas vrai. Nous avons commis une erreur et préparons actuellement un erratum sur ce point.

Nous avons depuis trouvé le résultat correct de l'équation (2.11) avec un ensemble $\mathcal{A}$ quelconque. Ce résultat est donné par l'équation suivante, avec m la FM donnée par la source :

$$m \odot \overline{\bigcup_{A \in \mathcal{A}} A_{\alpha_A}} = m \odot \left(\bigcap_{A \in \mathcal{A}} \overline{A}^{\alpha_A} \right). \quad (2.12)$$

La preuve de ce résultat est donné en annexe A. L'équation (2.12) constitue donc une sorte d'affaiblissement contextuel *bis*.

Exemple 2.1 Avec $\mathcal{X} = \{a, b, c\}$, $\mathcal{A} = \{\{a\}\}$, $\beta_{\{a\}}$ étant noté simplement par β_a , on a pour l'affaiblissement contextuel donnée par l'équation (2.10) :

$$m \odot \{a\}_{\beta_a} = m \odot \begin{cases} \emptyset & \mapsto \beta_a \\ \{a\} & \mapsto \alpha_a \end{cases}, \quad (2.13)$$

et pour l'opération définie par (2.12) :

$$m \odot \overline{\{a\}_{\alpha_a}} = m \odot \begin{cases} \mathcal{X} & \mapsto \alpha_a \\ \overline{\{a\}} = \{b, c\} & \mapsto \beta_a \end{cases}, \quad (2.14)$$

Ces deux résultats sont donc généralement différents.

Nous pouvons retrouver l'égalité cependant dans le cas par exemple où $\mathcal{X} = \{a\}$ et $\mathcal{A} = \{\{a\}\} = \{\mathcal{X}\}$, où l'on obtient :

$$m \odot \overline{\{a\}_{\alpha_a}} = m \odot \begin{cases} \mathcal{X} & \mapsto \alpha_a \\ \overline{\{a\}} = \emptyset & \mapsto \beta_a \end{cases} = m \odot \mathcal{X}_{\beta_a} = m \odot \{a\}_{\beta_a}, \quad (2.15)$$

ou bien dans le cas $\mathcal{X} = \{a, b\}$ et $\mathcal{A} = \{\{a\}, \{b\}\}$, où l'on a :

$$m \bigcirc \overline{\{a\}_{\alpha_a} \bigcirc \{b\}_{\alpha_b}} = m \bigcirc \left\{ \overline{\emptyset \mapsto \alpha_a} \bigcirc \left\{ \emptyset \mapsto \alpha_b \right. \right. \quad (2.16)$$

$$\left. \left. \left. \left. \left. \begin{array}{l} \overline{\{a\}} \mapsto \beta_a \alpha_b \\ \overline{\{b\}} \mapsto \alpha_a \beta_b \\ \overline{\{a, b\}} \mapsto \beta_a \beta_b \end{array} \right. \right. \right. \right. \right. \quad (2.17)$$

Or comme $\mathcal{X} = \{a, b\}$, on a finalement que (2.17) est équivalente à :

$$m \bigcirc \left\{ \begin{array}{l} \mathcal{X} \mapsto \alpha_a \alpha_b \\ \{b\} \mapsto \beta_a \alpha_b \\ \{a\} \mapsto \alpha_a \beta_b \\ \emptyset \mapsto \beta_a \beta_b \end{array} \right. = m \bigcirc \{a\}_{\beta_a} \bigcirc \{b\}_{\beta_b} \quad (2.18)$$

Formellement, on sait qu'on a l'égalité entre (2.10) et (2.12) dans les cas où $\mathcal{A}$ forme une partition de $\mathcal{X}$ [62]. Les conditions exactes pour cette égalité sont encore en cours de travaux.

Nous revenons maintenant au sens des paramètres β_A . D'un point de vue pratique, d'après l'équation (2.10), chaque coefficient β_A représente, lors de l'application de l'opération d'affaiblissement contextuel sur une FM m , la proportion de masse $m(B)$ restant sur B pour tout élément focal B de m , et $(1 - \beta_A)$ représente la part de $m(B)$ qui est transférée de B à $B \cup A$. La figure 2.1 illustre ce transfert de masses lors de l'affaiblissement contextuel d'une FM m , pour chaque contexte A , pour chaque élément focal B de m .

Dans [66], nous montrons que l'équation (2.10) peut aussi découler du mécanisme appelé *Behaviour-based correction (BBC)* (que l'on pourrait traduire en français par *correction basée sur le comportement*) introduit par Pichon et al. dans [65], et qui est défini à partir

- d'une FM $m^{\mathcal{Y}}$ fournie par une source,
- d'une FM $m^{\mathcal{H}}$ décrivant l'état (ou la configuration) h dans lequel se trouve la source,
- et pour tout $A \subseteq \mathcal{Y}$ d'une fonction Γ_A définie de $\mathcal{H}$ dans $2^{\mathcal{X}}$ indiquant comment interpréter l'information $y \in A \subseteq \mathcal{Y}$ pour chaque état $h \in \mathcal{H}$.

Le mécanisme BBC permet alors d'obtenir une information sur X en combinant ces

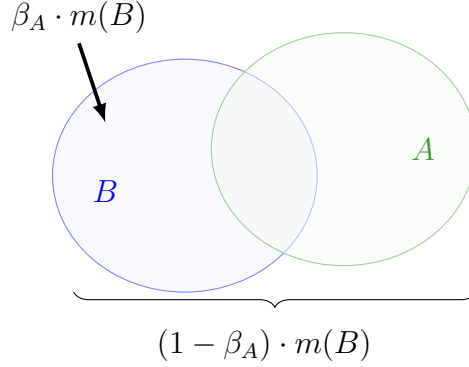


FIGURE 2.1 – Illustration des transferts de masses lors de l’affaiblissement contextuel d’une FM m pour un élément focal B de m et un contexte A de $\mathcal{A}$.

informations de la manière suivante :

$$\begin{aligned}
 & m^{\mathcal{X}}[m^{\mathcal{Y}}, m^{\mathcal{H}}, \{\Gamma_A, A \subseteq \mathcal{Y}\}] \\
 = & m^{\mathcal{X}}[m^{\mathcal{Y}}, m^{\mathcal{H}}, m_{\Gamma}^{\mathcal{H} \times \mathcal{X} \times \mathcal{Y}}] \quad \text{avec } m_{\Gamma}^{\mathcal{H} \times \mathcal{X} \times \mathcal{Y}}(\cup_{h \in \mathcal{H}, A \subseteq \mathcal{Y}} \{h\} \times A \times \Gamma_A(h)) = 1 \\
 = & \left(\left(m^{\mathcal{Y} \uparrow \mathcal{H} \times \mathcal{X} \times \mathcal{Y}} \oplus m_{\Gamma}^{\mathcal{H} \times \mathcal{X} \times \mathcal{Y}} \right)^{\downarrow \mathcal{H} \times \mathcal{X}} \oplus m^{\mathcal{H} \uparrow \mathcal{H} \times \mathcal{X}} \right)^{\downarrow \mathcal{X}}. \tag{2.19}
 \end{aligned}$$

Le résultat de cette opération [65] est alors donné, pour tout $C \subseteq \mathcal{X}$, par :

$$m^{\mathcal{X}}(C) = \sum_{H \subseteq \mathcal{H}} m^{\mathcal{H}}(H) \sum_{B: \Gamma_B(H)=C} m^{\mathcal{Y}}(B), \tag{2.20}$$

où, avec un abus usuel de notation, $\Gamma_B(H) = \cup_{h \in H} \Gamma_B(h)$.

Avec $\mathcal{Y} = \mathcal{X}$, et m la FM fournie par la source, on note simplement l’application du BBC par $f_{m^{\mathcal{H}}}(m)$.

Nous pouvons alors voir [66] que l’affaiblissement contextuel avec un seul contexte $A \subseteq \mathcal{X}$, *i.e.* $\mathcal{A} = \{A\}$, est équivalent au BBC avec $\mathcal{H}$ composé de deux états h_1 et h_2 tels que $m^{\mathcal{H}}(\{h_1\}) = \beta_A$, $m^{\mathcal{H}}(\{h_2\}) = 1 - \beta_A$, et l’interprétation des deux états étant donnée par $\Gamma_B(h_1) = B$ et $\Gamma_B(h_2) = B \cup A$, $\forall B \subseteq \mathcal{X}$. Avec les deux modèles, la même correction définie par

$$f_{m^{\mathcal{H}}}(m) = m \oplus A_{\beta_A}, \tag{2.21}$$

est obtenue.

L’équation de l’affaiblissement contextuel (2.10) peut alors être obtenue pour un ensemble quelconque de contextes $\mathcal{A}$ à partir d’une composition de $|\mathcal{A}|$ application(s) du BBC [66] :

$$\alpha_{\cup} m = m \oplus_{A \in \mathcal{A}} A_{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m_{A, \cup}^{\mathcal{H}}})(m), \tag{2.22}$$

où :

- $\mathcal{H} = \{n_A | A \subseteq \mathcal{X}\}$ où l'interprétation de chaque état est donné par $\Gamma_B(n_A) = B \cup A, \forall B \subseteq \mathcal{X}$;
- Pour tout $A \in \mathcal{A}$, les FM $m_{A,\cup}^{\mathcal{H}}$ sont définies par $m_{A,\cup}^{\mathcal{H}}(\{n_\emptyset\}) = \beta_A$ et $m_{A,\cup}^{\mathcal{H}}(\{n_A\}) = 1 - \beta_A$.

Dans [66], nous montrons également que nous pouvons donner un sens aux états $n_A, A \subseteq \mathcal{X}$, en terme de *biais* des déclarations de la source.

Si la source est dans l'état $n_A, A \subseteq \mathcal{X}$, nous disons que la source est *negativement biaisée dans A* et cela signifie que :

- la source n'est pas biaisée lorsqu'elle déclare que $x \in \overline{A}$ est une valeur possible pour x_0 ou $x \notin \overline{A}$ est une valeur possible pour x_0 . En d'autres termes, la source dit la vérité dans $\overline{A}$;
- la source n'est pas biaisée lorsqu'elle déclare $x \in A$ est une valeur possible pour x_0 , mais elle se trompe (ou ne dit pas la vérité) lorsqu'elle déclare $x \notin A$ est une valeur possible pour x_0 .

Au total, une source *negativement biaisée dans A* ne dit pas la vérité (ou se trompe) seulement quand elle dit $x \notin A$ est une valeur possible pour x_0 .

Ainsi, si la source nous indique $x \in B$ est une solution possible, $B \subseteq \mathcal{X}$ et qu'elle se trouve dans l'état (ou la configuration) n_A, A un contexte de $\mathcal{A}$, alors nous devons conclure que $x \in B \cup A$ est une solution possible. En effet, en considérant l'ensemble des valeurs possibles x dans $\mathcal{X}$ pour x_0 , on a, suite à la définition de l'état n_A donnée précédemment :

- Si la source indique que x est une valeur possible de x_0 (*i.e.* $x \in B$) :
 - Si x est dans le contexte A (*i.e.* $x \in A$), comme la source fait une déclaration positive alors la source dit la vérité, et x est donc bien une valeur possible pour x_0 ;
 - Si x n'est pas dans le contexte A ($x \notin A$), la source disant la vérité dans $\overline{A}$, x est bien une valeur possible pour x_0 .
- Si la source indique que x n'est pas une valeur possible de x_0 (*i.e.* $x \notin B$) :
 - Si x est dans le contexte A (*i.e.* $x \in A$), comme la source fait une déclaration négative, elle se trompe, donc x est une valeur possible pour x_0 ;
 - Si x n'est pas dans le contexte A ($x \notin A$), la source disant la vérité dans $\overline{A}$, x n'est en effet pas une valeur possible pour x_0 .

Au final, si une source déclare $x \in B, B \subseteq \mathcal{X}$, est une valeur possible pour x_0 étant dans l'état n_A , c'est qu'on a réellement $x \in (B \cap A) \cup (B \cap \overline{A}) \cup (\overline{B} \cap A) = A \cup B$ est une valeur possible pour x_0 . Les schémas illustrés sur les figures 2.2(a) et 2.2(b) illustrent cette correction de l'information fournie par la source si celle-ci se trouve dans l'état n_A .

Nous retrouvons alors l'interprétation donnée par la fonction Γ_B permettant de dériver l'affaiblissement contextuel (2.22).

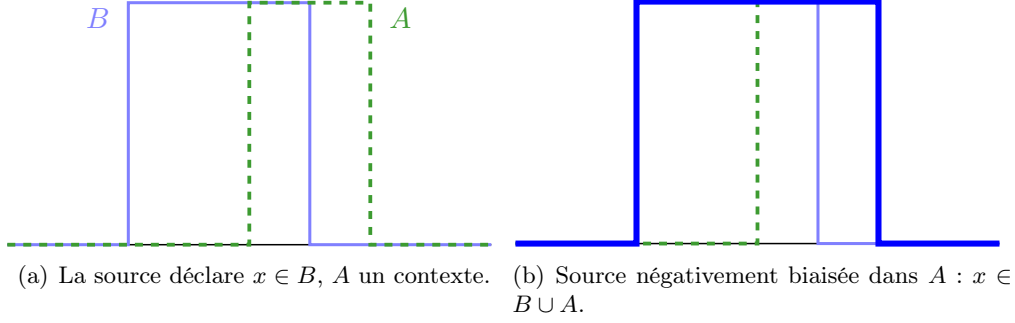


FIGURE 2.2 – Illustration d’une situation où la source déclare que $x \in B \subseteq \mathcal{X}$ est une valeur possible pour x_0 avec un contexte A . La figure (2.2(a)) représente les fonctions indicatrices des ensembles B et A . La figure (2.2(b)) représente l’information à conclure de la déclaration de la source qui est dans l’état n_A (si la source dit d’une valeur située dans A qu’elle n’est pas une valeur possible de x_0 , elle se trompe (ou elle ment), ce qu’elle dit des valeurs situées en dehors de A peut être gardé tel quel).

En comparant les deux approches permettant de dériver l’affaiblissement contextuel sur une partition $\mathcal{A}$, on peut remarquer que dans l’approche originelle [62], nous avons une connaissance contextuelle relative à la qualité de la source suivant la valeur prise par x_0 , alors que dans cette nouvelle dérivation [66], nous avons une connaissance contextuelle relative à ce que la source déclare sur une possible valeur pour x_0 . La notion de contexte est (un peu) différente entre les deux approches.

Le lien entre la dérivation originelle de l’affaiblissement contextuel donnée par l’équation (2.11) et la dérivation issu du BBC (2.19) reste encore à explorer, de même que la différence de résultat obtenu dans des cas où $\mathcal{A}$ ne forme pas une partition de $\mathcal{X}$ (les conditions exactes d’égalité entre (2.10) et (2.12) restant à déterminer).

2.4 Renforcement contextuel

Définition 2.2 (Renforcement contextuel) *Le renforcement contextuel d’une FM m est défini [56], avec $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$, par :*

$$\alpha_{\cap} m = m \bigcirc_{A \in \mathcal{A}} A^{\beta_A} . \quad (2.23)$$

D’après cette équation (2.23), chaque coefficient β_A représente la proportion de masse $m(B)$ restant sur B pour tout élément focal B de m , et $(1 - \beta_A)$ représente la part de $m(B)$ qui est transférée de B à $B \cap A$. La figure 2.3 illustre ce transfert de masses lors du renforcement contextuel d’une FM m , pour chaque contexte A , pour chaque élément focal B de m .

Nous n’avons pas trouvé d’équation équivalente à l’équation (2.11) pour dériver le renforcement contextuel. Nous l’avons introduit dans [56] comme l’opération associée à la négation de l’affaiblissement contextuel de la négation d’une masse non normale

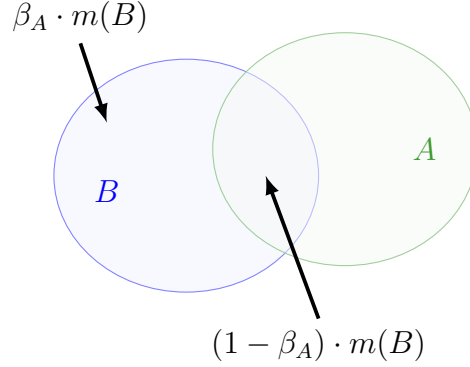


FIGURE 2.3 – Illustration des transferts de masses lors du renforcement contextuel d'une FM m pour un élément focal B de m et un contexte A de $\mathcal{A}$.

m , en remarquant que pour une telle masse nous avons :

$$\overline{\alpha \cup m} = \alpha \cap m \quad (2.24)$$

Cependant, de même que pour l'affaiblissement contextuel, nous avons montré dans [66] que le renforcement contextuel avec un ensemble $\mathcal{A}$ de contextes, peut être obtenu à partir d'une composition de $|\mathcal{A}|$ application(s) du BBC :

$$\alpha \cap m = m \bigodot_{A \in \mathcal{A}} A_{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m_{A, \cap}}^{\mathcal{H}})(m), \quad (2.25)$$

où :

- $\mathcal{H} = \{p_A | A \subseteq \mathcal{X}\}$ où l'interprétation de chaque état est donné par $\Gamma_B(p_A) = B \cap A, \forall B \subseteq \mathcal{X}$;
- Pour tout $A \in \mathcal{A}$, les FM $m_{A, \cap}^{\mathcal{H}}$ sont définies par $m_{A, \cup}^{\mathcal{H}}(\{p_{\emptyset}\}) = \beta_A$ et $m_{A, \cap}^{\mathcal{H}}(\{p_A\}) = 1 - \beta_A$.

Nous pouvons aussi donner un sens aux états $p_A, A \subseteq \mathcal{X}$, en terme de *biais* par rapport aux déclarations de la source [66]. Si la source est dans l'état $p_A, A \subseteq \mathcal{X}$, nous disons que la source est *positivement biaisée dans $\bar{A}$* . Cela signifie que :

- la source n'est pas biaisée lorsqu'elle déclare $x \in A$ ou $x \notin A$ est une valeur possible pour x_0 . En d'autres termes, la source dit la vérité dans A ;
- la source se trompe (ou ne dit pas la vérité) lorsqu'elle déclare $x \in \bar{A}$ est une valeur possible, et ne se trompe pas (dit la vérité) lorsqu'elle déclare $x \notin \bar{A}$ est une valeur possible pour x_0 .

En résumé, si la source est dans l'état p_A , elle se trompe (ou ne dit pas la vérité) uniquement quand elle indique que $x \in \bar{A}$ est une valeur possible pour x_0 .

Si la source nous indique $x \in B, B \subseteq \mathcal{X}$ est une valeur possible x_0 et qu'elle se trouve dans l'état p_A, A un contexte de $\mathcal{A}$, nous devons alors conclure que $x \in B \cap A$ est une valeur possible pour x_0 . Pour le vérifier, nous pouvons comme pour

l'affaiblissement contextuel, parcourir toutes les valeurs possibles x dans $\mathcal{X}$ pour x_0 . On obtient alors avec la définition de l'état p_A donnée précédemment, que :

- Si la source indique que x est une valeur possible de x_0 (*i.e.* $x \in B$) :
 - Si x est dans le contexte A (*i.e.* $x \in A$), la source disant la vérité dans A , x est bien une valeur possible pour x_0 .
 - Si x n'est pas dans le contexte A ($x \notin A$), comme la source fait une déclaration positive, elle ne dit pas la vérité et x n'est donc pas une valeur possible pour x_0 .
- Si la source indique que x n'est pas une valeur possible de x_0 (*i.e.* $x \notin B$) :
 - Si x est dans le contexte A (*i.e.* $x \in A$), la source disant la vérité dans A , x n'est pas une valeur possible pour x_0 .
 - Si x n'est pas dans le contexte A ($x \notin A$), comme la source fait une déclaration négative, elle dit la vérité et x n'est pas une valeur possible pour x_0 .

On trouve au final que si une source dans l'état p_A déclare $x_0 \in B$, $B \subseteq \mathcal{X}$, alors $x \in B \cap A$ est une valeur possible pour x_0 . Cette conclusion est illustrée par les schémas de la figure 2.4.

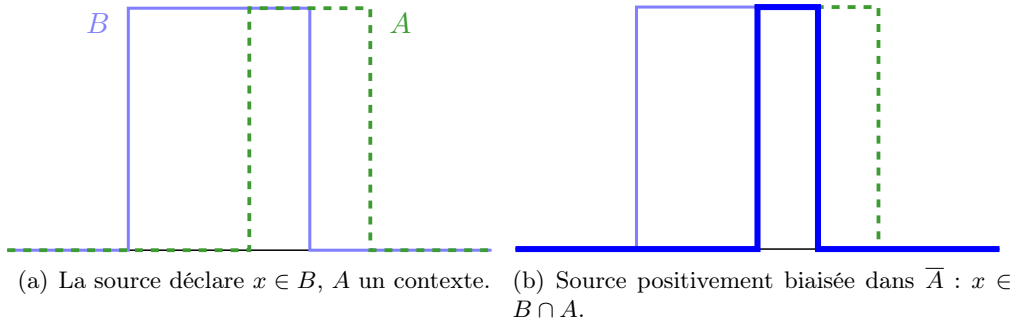


FIGURE 2.4 – Illustration d'une situation où la source déclare que $x \in B \subseteq \mathcal{X}$ est une valeur possible pour x_0 avec un contexte A . La figure (2.4(a)) représente les fonctions indicatrices des ensembles B et A . La figure (2.4(b)) représente l'information à conclure de la déclaration de la source qui est dans l'état p_A (si la source dit d'une valeur située en dehors A qu'elle est une valeur possible de x_0 , elle se trompe (ou elle ment), ce qu'elle dit des valeurs situées dans A peut être gardé tel quel).

2.5 Apprentissage automatique des coefficients

Dans cette section, nous abordons la question de l'apprentissage automatique des coefficients à partir de données étiquetées, pour d'un côté l'affaiblissement contextuel, et de l'autre le renforcement contextuel [61]. Nous commençons par décrire le processus d'apprentissage employé. Puis nous présentons le cas de l'affaiblissement contextuel suivi de celui du renforcement contextuel. Enfin nous discutons

Tableau 2.1 – Sorties de deux sources S_1 et S_2 concernant la classe de 4 objets dans $\mathcal{X} = \{a, h, r\}$. Exemple issu de [30, Table 1].

		$\{a\}$	$\{h\}$	$\{r\}$	$\{a, h\}$	$\{a, r\}$	$\{h, r\}$	$\mathcal{X}$	Vérité terrain
S_1	$m_{S_1}\{o_1\}$	0	0	0.5	0	0	0.3	0.2	a
	$m_{S_1}\{o_2\}$	0	0.5	0.2	0	0	0	0.3	h
	$m_{S_1}\{o_3\}$	0	0.4	0	0	0.6	0	0	a
	$m_{S_1}\{o_4\}$	0	0	0	0	0.6	0.4	0	r
S_2	$m_{S_2}\{o_1\}$	0	0	0	0.7	0	0	0.3	a
	$m_{S_2}\{o_2\}$	0.3	0	0	0.4	0	0	0.3	h
	$m_{S_2}\{o_3\}$	0.2	0	0	0	0	0.6	0.2	a
	$m_{S_2}\{o_4\}$	0	0	0	0	0	1	0	r

des différences de résultats obtenus sur des exemples et une expérimentation en classification.

2.5.1 Description du processus d'apprentissage

Une source S fournit en sortie une information, représentée sous la forme d'une FM m , concernant la classe dans $\mathcal{X} = \{x_1, \dots, x_K\}$ d'un objet donné en entrée. Nous disposons d'un ensemble d'apprentissage composé de n objets o_i , $i \in \{1, \dots, n\}$, dont la classe est connue, et au regard desquels la source a fourni des FM $m\{o_i\}$. Un exemple d'un tel ensemble d'apprentissage est donné dans le tableau 2.1.

Connaissant ces données, notre objectif est d'obtenir la meilleure correction pour cette source par affaiblissement contextuel d'un côté et la meilleure par renforcement contextuel de l'autre. Une fois ceci énoncé, il reste à définir la mesure de performance qui sera employée. Notre choix s'est porté sur la minimisation de la mesure d'erreur entre les informations corrigées et la vérité suivante :

$$E(\beta) = \sum_{i=1}^n \sum_{k=1}^K (\alpha pl\{o_i\}(\{x_k\}) - \delta_{i,k})^2, \quad (2.26)$$

où pour tout $i \in \{1, \dots, n\}$, $\alpha pl\{o_i\}$ est la fonction de plausibilité associée à la correction contextuelle de la sortie m de la source avec un vecteur de correction $\beta \in [0, 1]^{|A|}$, et $\delta_{i,k}$ est une variable binaire indiquant la classe de l'objet o_i : $\forall i \in \{1, \dots, n\}, \forall k \in \{1, \dots, K\}, \delta_{i,k} = 1$ si o_i a pour classe x_k , et $\delta_{i,k} = 0$ sinon.

Au lieu d'utiliser cette mesure (2.26) utilisant la fonction de plausibilité sur

les singletons (fonction contour), nous aurions pu utiliser par exemple une mesure basée sur la transformée pignistique (1.33) ou sur une distance [41]. L'emploi de cette mesure (2.26) a été motivée par trois raisons :

1. une raison pratique : comme nous le présentons aux paragraphes 2.5.2 et 2.5.3, l'optimisation de cette mesure pour l'affaiblissement contextuel comme pour le renforcement contextuel conduit à un problème simple avec une résolution rapide ;
2. une deuxième raison pratique : cette mesure est facilement interprétable et analysable ;
3. enfin, une raison plus théorique : cette mesure n'est pas moins justifiée que les autres (cf Section 1.4).

2.5.2 Cas de l'affaiblissement contextuel

Les plausibilités sur les singletons après avoir employé l'affaiblissement contextuel $\alpha_\cup m$ sur une FM m , avec un ensemble de contextes $\mathcal{A}$ et le vecteur de taux de fiabilité associé $\beta = (\beta_A, A \in \mathcal{A})$, sont données par [61, Proposition 1] :

$$\alpha_\cup pl(\{x\}) = 1 - (1 - pl(\{x\})) \prod_{A \in \mathcal{A}, x \in A} \beta_A . \quad (2.27)$$

Un moyen d'obtenir cette expression (2.27) consiste [61] à exprimer l'affaiblissement contextuel (2.10) en termes de fonctions d'implicabilité avec la propriété (1.12)), puis à utiliser la relation $pl(\{x\}) = 1 - b(\overline{\{x\}})$ pour tout $x \in \mathcal{X}$.

Nous montrons alors dans [61, Proposition 2] que, avec l'emploi d'un affaiblissement contextuel, la minimisation de E (2.26) est obtenue en utilisant le vecteur β composé des K coefficients $\beta_{\{x_k\}}$, c'est-à-dire un affaiblissement contextuel avec l'ensemble de contextes $\mathcal{A}$ composé des éléments singletons de $\mathcal{X}$.

Dans [62, Section 5], nous avons déjà étudié la minimisation de E (2.26) avec un tel affaiblissement contextuel basé sur la partition la plus fine de $\mathcal{X}$, i.e. $\mathcal{A} = \mathcal{X}$. Il s'agit d'un problème de moindres carrés, qui peut donc être résolu de manière efficace. On montre en effet [62, 61] qu'on a :

$$E(\beta) = \|\mathbf{Q}\beta - \mathbf{d}\|^2 \text{ avec } \mathbf{Q} = \begin{bmatrix} diag(\mathbf{pl}_1 - 1) \\ \vdots \\ diag(\mathbf{pl}_n - 1) \end{bmatrix} \text{ et } \mathbf{d} = \begin{bmatrix} \delta_1 - 1 \\ \vdots \\ \delta_n - 1 \end{bmatrix} , \quad (2.28)$$

avec $diag(\mathbf{v})$ la matrice diagonale composée des éléments du vecteur $\mathbf{v}$ sur sa diagonale, $\mathbf{pl}_i = (pl\{o_i\}(\{x_1\}), \dots, pl\{o_i\}(\{x_K\}))^T$, et $\delta_i = (\delta_{i,1}, \dots, \delta_{i,K})^T$ le vecteur colonne binaire indicateur de la classe des objets o_i .

2.5.3 Cas du renforcement contextuel

Dans le cas d'un renforcement contextuel $\alpha_{\cap} m$ d'une FM m , avec un ensemble de contextes $\mathcal{A}$ et un vecteur de correction associé $\beta = (\beta_A, A \in \mathcal{A})$, les plausibilités sur les singletons sont données par [61, Proposition 3] :

$$\alpha_{\cap} pl(\{x\}) = pl(\{x\}) \prod_{A \in \mathcal{A}, x \notin A} \beta_A . \quad (2.29)$$

Cette équation (2.29) peut être obtenue [61] en exprimant l'équation du renforcement contextuel (2.23) avec des fonctions de commonalité à partir de la propriété (1.11), puis à utiliser le fait que $pl(\{x\}) = q(\{x\})$ pour tout $x \in \mathcal{X}$.

Dans [61, Proposition 3], nous montrons alors que la minimisation de E (2.26) avec l'emploi d'un renforcement contextuel est obtenue en utilisant le vecteur $\beta = (\beta_{\overline{\{x_k\}}}, k \in \{1, \dots, K\})$, c'est-à-dire un renforcement contextuel avec l'ensemble de contextes $\mathcal{A}$ composé de tous les sous-ensembles de $\mathcal{X}$ de cardinalité $|\mathcal{X}| - 1$. Cette minimisation constitue aussi un problème de moindres carrés, puisqu'avec cette correction contextuelle, (2.26) s'écrit :

$$E(\beta) = \|\mathbf{P}\beta - \delta\|^2, \text{ with } \mathbf{P} = \begin{bmatrix} diag(\mathbf{pl}_1) \\ \vdots \\ diag(\mathbf{pl}_n) \end{bmatrix} \text{ and } \delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix}, \quad (2.30)$$

avec les mêmes notations que dans l'équation (2.28).

2.5.4 Discussions et exemples

Avant de montrer des exemples de résultats, nous pouvons commenter les différentes capacités de corrections de ces deux mécanismes.

Avec l'affaiblissement contextuel, on a $\alpha_{\cup} pl(\{x_k\}) = 1 - (1 - pl(\{x_k\}))\beta_{\{x_k\}}$ pour chaque $k \in \mathcal{X}$, $k \in \{1, \dots, K\}$ (2.27). Ainsi, avec $\beta_{\{x_k\}}$ variant dans l'intervalle $[0, 1]$, $\alpha_{\cup} pl(\{x_k\})$ peut prendre n'importe quelle valeur dans l'intervalle $[pl(\{x\}), 1]$. Cela signifie, qu'avec cette correction, chaque valeur $pl(\{x\})$ fournie par la source peut être amenée le proche possible de 1 (autrement dit affaiblie autant que désirée).

D'un autre côté, avec le renforcement contextuel (2.29), on a $\alpha_{\cap} pl(\{x_k\}) = pl(\{x_k\})\beta_{\overline{\{x_k\}}}$ pour tout $x_k \in \mathcal{X}$, $k \in \{1, \dots, K\}$. Par conséquent, $pl(\{x_k\})$ peut prendre n'importe quelle valeur dans $[0, pl(\{x_k\})]$ avec $\beta_{\overline{\{x_k\}}}$ variant de 0 à 1. En d'autres mots, le renforcement contextuel permet de rapprocher chaque valeur $pl(\{x\})$ fournie par la source vers 0 autant que souhaitée. Il renforce l'information donnée par la source en faisant décroître les plausibilités de certains singletons.

Afin d'améliorer la mesure d'erreur (2.26), l'affaiblissement contextuel peut donc rapprocher le plus proche possible de 1 la plausibilité sur le singleton étant la vérité terrain, alors que le renforcement contextuel peut diminuer toutes les plausibilités sur les singletons n'étant pas la vérité vers 0.

Tableau 2.2 – Résultats obtenus pour la minimisation de E à partir des données du tableau 2.1 pour l'affaiblissement contextuel d'un côté, et le renforcement contextuel de l'autre.

Correction	Source 1	Source 2
Affaib. contextuel	$\beta = (0.76, 1.00, 1.00)$ $E(\beta) = 3.39$	$\beta = (0.74, 1.00, 1.00)$ $E(\beta) = 4.81$
Renf. contextuel	$\beta = (0.94, 0.66, 0.38)$ $E(\beta) = 2.33$	$\beta = (0.65, 0.22, 0.55)$ $E(\beta) = 2.39$

L'exemple 2.2 suivant illustre ces différentes capacités d'ajustements.

Exemple 2.2 Supposons que $\mathcal{X} = \{a, b, c\}$ et que la vérité terrain est a .

Soit une FM m telle que $m(\{b, c\}) = 1$, i.e. $pl(\{a\}) = 0$ et $pl(\{b\}) = pl(\{c\}) = 1$.

Pour rapprocher, au sens de la mesure d'erreur E (2.26), pl de la réalité $pl(\{a\}) = 1$ et $pl(\{b\}) = pl(\{c\}) = 0$:

- par l'affaiblissement contextuel, on peut augmenter la plausibilité de la vérité $pl(\{a\})$ à 1,
- par le renforcement contextuel, on peut diminuer les plausibilités $pl(\{b\})$ et $pl(\{c\})$ à 0.

Sur cet exemple, la mesure d'erreur E minimale atteinte pour l'affaiblissement contextuel est $E = (1-1)^2 + (1-0)^2 + (1-0)^2 = 2$. Pour le renforcement contextuel, on arrive à obtenir $E = (0-1)^2 + (0-0)^2 + (0-0)^2 = 1$.

On peut donc remarquer que pour un objet, pour la mesure E , avec $\mathcal{X} = \{x_1, \dots, x_K\}$ et sans perdre de généralité avec x_1 la vérité, l'affaiblissement contextuel fait mieux que le renforcement contextuel si et seulement si $\sum_{k \in \{2, \dots, K\}} pl(\{x_k\})^2$ est inférieure à $(1 - pl(\{x_1\}))^2$. Sinon, le renforcement contextuel obtient une plus faible valeur pour la mesure d'erreur E .

Si on considère les performances sur un ensemble composé de plusieurs objets, il devient plus difficile de voir intuitivement quel mécanisme fera mieux. Nous montrons dans l'exemple 2.3 les résultats de l'application de ces méthodes sur les données du tableau 2.1.

Exemple 2.3 Les résultats de la minimisation de E obtenus à partir des données du tableau 2.1 pour l'affaiblissement contextuel (2.28) et pour le renforcement contextuel (2.30) sont indiqués dans le tableau 2.2 pour chaque source. Rappelons que $\beta = (\beta_{\{a\}}, \beta_{\{b\}}, \beta_{\{c\}})$ pour l'affaiblissement contextuel, et $\beta = (\beta_{\overline{\{a\}}}, \beta_{\overline{\{b\}}}, \beta_{\overline{\{c\}}})$ pour le renforcement contextuel, avec des significations différentes pour chaque coefficient.

Pour l'affaiblissement contextuel, $\beta_{\{h\}} = \beta_{\{r\}} = 1$ pour les deux sources, ce qui signifie que les deux sources sont fiables si la vérité cherchée $x_0 = h$ ou $x_0 = r$. Il

n'y a pas besoin de transférer une partie de la masse $m(B)$ fournie par une source sur $B \cup \{h\}$ ou $B \cup \{r\}$ avec $B \subseteq \mathcal{X}$ ($1 - \beta_{\{h\}} = 1 - \beta_{\{r\}} = 0$). Ce n'est pas le cas pour les objets de type a qui causent plus de problèmes, la source 1 étant légèrement plus fiable.

Concernant le renforcement contextuel, nous avons $\beta_{\overline{\{a\}}} > \beta_{\overline{\{h\}}}$ et $\beta_{\overline{\{a\}}} > \beta_{\overline{\{r\}}}$ pour les deux sources. Ainsi, au niveau des transferts de masses, on peut remarquer qu'il y a moins besoin de transférer des masses de $m(B)$ à $B \cap \overline{\{a\}} = B \cap \{h, r\}$ que sur $B \cap \overline{\{r\}} = B \cap \{a, h\}$ ou $B \cap \overline{\{h\}} = B \cap \{a, r\}$, ce qui confirme la meilleure fiabilité de la source sachant $x_0 = a$. Néanmoins, on peut observer des différences au niveau des valeurs des paramètres.

Enfin, nous pouvons observer également que nous réussissons à obtenir une plus faible valeur pour l'erreur E_{pl} avec le renforcement pour les deux sources, ce qui illustre l'intérêt d'avoir introduit ce mécanisme.

2.5.5 Une expérimentation en classification

Dans cette section, nous développons une expérimentation pour montrer l'intérêt des mécanismes d'affaiblissement et de renforcement contextuels pour améliorer les performances d'un classifieur évidentiel, c'est-à-dire fournissant des fonctions de croyance en sortie.

Pour ce faire, nous avons choisi le classifieur évidentiel des k plus proches voisins (*ev-knn*) introduit par Dencœux dans [21], et nous avons construit un problème de classification à 5 classes à partir des 5 distributions normales à 2 variables de moyennes respectives $\mu_{x_1} = (0, 0)$, $\mu_{x_2} = (2, 0)$, $\mu_{x_3} = (0, 2)$, $\mu_{x_4} = (2, 2)$, $\mu_{x_5} = (1, 1)$ avec pour matrice de covariance commune :

$$\Sigma = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}.$$

Nous avons généré 1000 instances pour chacune des 5 classes présentes dans $\mathcal{X} = \{x_1, x_2, x_3, x_4, x_5\}$, soit un total de 5000 instances illustrées sur la figure 2.5.

Le principe du classifieur *ev-knn* [21] est de prendre les k plus proches voisins étiquetés v_i , $i \in \{1, \dots, k\}$, d'une nouvelle instance a à classer. Chacune des classes des objets voisins v_i constitue alors une information en faveur de la classe de a qui est fonction de la distance de l'objet v_i à a . Avec $x^i \in \mathcal{X}$ la classe de l'objet voisin v_i , $i \in \{1, \dots, k\}$, chaque voisin v_i génère une FM m_i au regard de la classe de a qui est définie par :

$$m_i(A) = \begin{cases} \tau \cdot \exp^{-\gamma_{x^i} d(a, v_i)^2} & \text{si } A = \{x^i\}, \\ 1 - \tau \cdot \exp^{-\gamma_{x^i} d(a, v_i)^2} & \text{si } A = \mathcal{X}, \\ 0 & \text{sinon,} \end{cases} \quad (2.31)$$

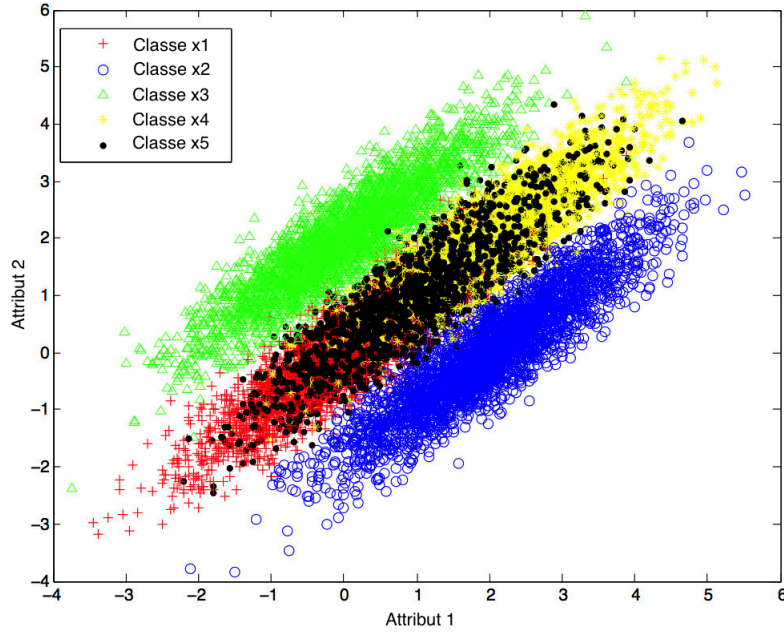


FIGURE 2.5 – Illustration des données générées pour un problème de classification à 5 classes avec 2 attributs.

avec τ une constante dans $[0, 1]$ choisie égale à 0.95 dans notre application, γ_x une constante positive dépendante de la classe $x \in \mathcal{X}$, et d une distance, la distance euclidienne ayant été choisie ici. Suivant une heuristique simple utilisée dans [21], γ_x a été choisie égale à l'inverse de la moyenne des distances entre chaque élément de la classe x (présent dans l'ensemble d'apprentissage du classifieur ev-knn).

La sortie du classifieur ev-knn est alors la combinaison par la règle de Dempster (1.9) des k fonctions de masse m_i .

Dans cette application, nous avons choisi de fixer k à 3. Ce n'est pas le paramètre permettant d'avoir les meilleures performances pour le classifieur ev-knn sur cette application, et le classifieur ev-knn n'est peut être pas non plus le meilleur des classifieurs crédibilistes pour cette application. L'idée de cette section est de montrer que les mécanismes contextuels peuvent apporter des améliorations à un classifieur existant dont on ne sait rien outre ses sorties. On trouve par exemple ce cas dans le domaine du tri du courrier postal (cf mon travail en thèse [51, 52]), où une entreprise peut acheter des lecteurs d'adresses postales d'une société concurrente, sans pour autant connaître leurs fonctionnements ni pouvoir les paramétrer, afin de les exploiter conjointement avec ses propres lecteurs d'adresses postales dans un système de fusion. Le rôle du classifieur *ev-knn* dans cette expérimentation est simplement de simuler une telle source.

Les données (les 5000 instances) sont divisées en trois parties :

1. Le premier tiers des données constitue l'ensemble d'apprentissage du classifieur ev-knn (apprentissage des paramètres γ_x , $x \in \mathcal{X}$, et constitution de l'ensemble

des voisins permettant de classer les nouvelles instances) ;

2. Le second tiers est utilisé pour apprendre les paramètres de l'affaiblissement et du renforcement contextuels à partir des sorties du classifieur ev-knn ;
3. Enfin le dernier tiers forme l'ensemble de test.

Les performances obtenues sur l'ensemble de test pour le classifieur ev-knn seul, pour le classifieur ev-knn avec une correction par affaiblissement contextuel et pour le classifieur ev-knn avec une correction par renforcement contextuel sont illustrées sur la figure 2.6 en utilisant des courbes ROC, c'est-à-dire la courbe du taux de vrais positifs en fonction du taux de faux positifs. Plus la courbe est proche de 1 et meilleure est la performance. À noter que les décisions ont été prises à partir des fonctions de plausibilité (1.36).

Comme nous pouvions nous y attendre, les instances des classes x_2 et x_3 , étant relativement disjointes des autres (cf figure 2.5) sont bien reconnues par le classifieur ev-knn (cf figures 2.6(b) et 2.6(c)). Les corrections contextuelles n'améliorent pas les résultats mais ne les détériorent pas.

Les classes x_1 , x_4 et x_5 se recoupent plus (cf figure 2.5) et le classifieur ev-knn a de moins bons résultats sur ces classe (cf figures 2.6(a), 2.6(d) et 2.6(e)). Nous pouvons alors observer une amélioration des performances dans ces situations grâce aux corrections contextuelles (l'affaiblissement faisant ici mieux que le renforcement).

2.6 Conclusion et perspectives

Nous avons résumé dans ce chapitre les résultats de nos travaux autour de l'affaiblissement et du renforcement contextuels, et montrer leurs intérêts au travers d'un exemple en classification où ils pourraient servir à améliorer les performances d'une source.

Les travaux de Pichon et al. [65] utilisant une fonction Γ interprétant des états h dans $\mathcal{X}$ ont permis de donner une justification à l'affaiblissement et au renforcement contextuels. Néanmoins, la découverte de cet affaiblissement contextuel bis (2.12) nous amène à vouloir comprendre la différence d'hypothèses faites entre l'équation (2.11) utilisant le déconditionnement [62, 56] et l'équation (2.19) utilisant une fonction d'interprétation Γ [65]. Ces travaux sont en cours.

Dans [66], nous sommes aussi allés plus loin en introduisant de nouveaux mécanismes de corrections contextuelles attachés aux interprétations suivantes :

1. $\mathcal{H} = \{\ell_A | A \subseteq \mathcal{X}\}$ chaque état étant interprété par $\Gamma_B(\ell_A) = B \oplus A, \forall B \subseteq \mathcal{X}$;
2. $\mathcal{H} = \{\ell'_A | A \subseteq \mathcal{X}\}$ chaque état étant interprété par $\Gamma_B(\ell'_A) = B \oplus A, \forall B \subseteq \mathcal{X}$;

Nous sommes actuellement en train d'étudier plus en profondeur ces nouveaux mécanismes.

Au niveau de l'apprentissage des coefficients des mécanismes, nous avons aussi travaillé à partir d'une information plus pauvre où nous aurions juste les matrices de confusion des sources [47, 55]. Une autre piste de recherche consiste à approfondir également ces méthodes pour voir comment nous pourrions obtenir d'autres méthodes d'apprentissage des coefficients.

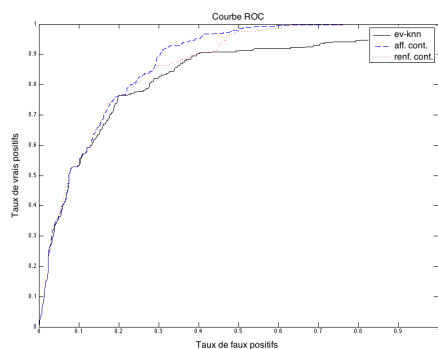
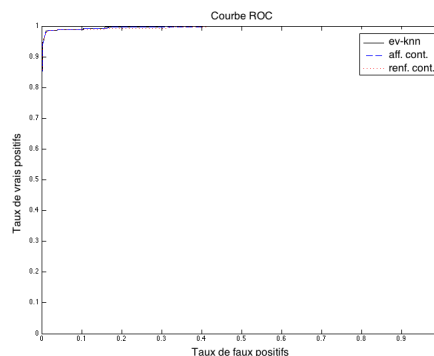
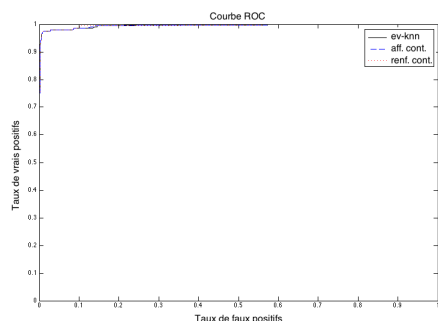
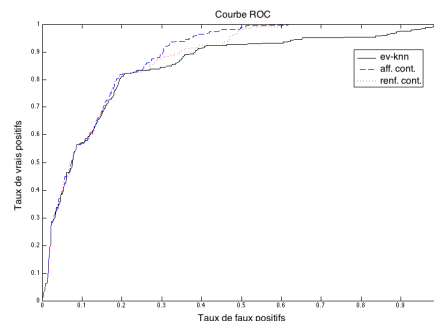
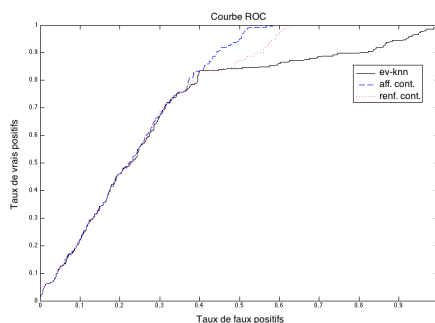
(a) Avec comme classe positive la classe x_1 .(b) Avec comme classe positive la classe x_2 .(c) Avec comme classe positive la classe x_3 .(d) Avec comme classe positive la classe x_4 .(e) Avec comme classe positive la classe x_5 .

FIGURE 2.6 – Performances sous forme de courbes ROC obtenues sur l'ensemble de test pour le classifieur ev-knn seul, pour le classifieur ev-knn avec une correction par affaiblissement contextuel et pour le classifieur ev-knn avec une correction par renforcement contextuel pour chacune des 5 classes.

La thèse de Gautier Levert a démarré en mars 2015 sur l'exploration de ces différentes perspectives de travaux.

Applications en fusion d'informations avec des fonctions de croyance

3.1 Introduction

Dans ce chapitre, trois applications en fusion d'informations sur lesquelles j'ai plus particulièrement travaillé, sont présentées.

La première, exposée en section 3.2, se situe dans le domaine de la communication véhicule à véhicule, et concerne une contribution issue des travaux de thèse de Mira Bou Farah [6, 11]. En plus d'apporter un nouvel exemple d'applications où le cadre des fonctions de croyance permet de modéliser les informations mises en jeu, elle montre l'intérêt d'un autre mécanisme de correction que l'affaiblissement classique (1.22) qui est le renforcement vers un élément singleton (1.31) abordé dans la section 1.3.5.3.

La seconde, présentée dans la section 3.3, met en avant un modèle que nous avons élaboré pour le problème d'association de données avec des fonctions de croyance. Cette application souligne les différences de performances possibles entre différentes modélisations crédales.

Enfin, la dernière application, abordée en section 3.4, porte sur le diagnostic du bobinage de moteur électrique et l'estimation de son vieillissement. Nous montrons comment le système de fusion d'informations a pu être simplement modélisé avec des fonctions de croyance, tout en assurant de bonnes performances.

Nos perspectives sur chacun de ces travaux sont mentionnées en fin de section.

3.2 Application en communication inter-véhicule

3.2.1 Description de l'application et des objectifs

Dans le cadre du projet *PLAiMOB* (*Plate-forme de simulation dédiée aux services de mobilité*) du contrat de projets état-région *CISIT* (*International Campus on Safety and Intermodality in Transportation*) [1] (2007 à 2013) et de la thèse de Mira Bou Farah [6, 11] (2010 à 2014), des méthodes utilisant des fonctions de croyance ont été développées pour modéliser et fusionner des informations imparfaites au regard d'événements pouvant survenir sur la route. Des exemples de tels échanges

d'informations entre véhicules concernant un accident, des travaux sur la route et une place de parking libre, sont illustrés sur la figure 3.1.

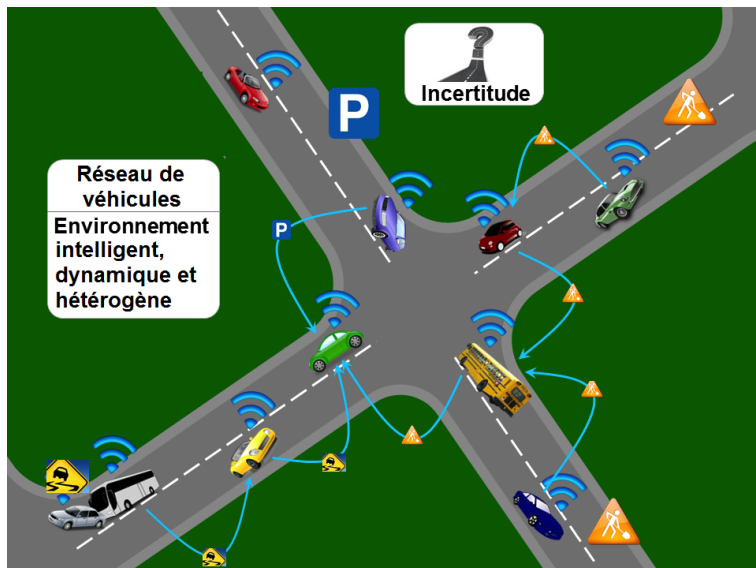


FIGURE 3.1 – Illustration d'échanges d'informations entre véhicules au regard d'événements pouvant survenir sur la route (accident, travaux, place de parking libre). Source [6, 11].

L'objectif est de fournir aux conducteurs une synthèse la plus proche possible de la réalité des événements sur la route (accidents, zones de travaux, embouteillages, nappes de brouillard, ...) à partir des informations reçues des autres véhicules. En particulier il n'y a pas de serveurs centraux, les seules bases de connaissances sont celles des véhicules.

Nous présentons ici une des contributions de ce travail [6, 11, 7, 10, 9, 8] qui est en lien avec la gestion des événements locaux, c'est-à-dire les événements situés localement dans l'espace comme un accident ou un véhicule en panne par exemple. Ces événements dits locaux ne s'étalent pas dans l'espace comme les embouteillages ou les nappes de brouillard qui sont traités autrement dans [6]. Ce choix de présenter cette contribution dans ce rapport est principalement dû à ce que l'emploi d'un autre mécanisme de correction que l'affaiblissement classique dans ce travail permet de mieux modéliser la réalité.

3.2.2 Solutions proposées pour la gestion des événements locaux

Un véhicule informe de la présence ou de l'absence d'un événement en créant un message $\mathbf{M}$ qui est un vecteur $(\mathbf{S}, \mathbf{t}, \mathbf{d}, \ell, \mathbf{m})$ composé des 5 éléments suivants :

1. la *source* $\mathbf{M.S}$ à l'origine de l'information, il s'agit du véhicule ayant perçu l'événement, ou ne l'ayant pas perçu alors qu'il était informé de son existence ;
2. le *type* $\mathbf{M.t}$ de l'événement (accident, travaux, ...) ;

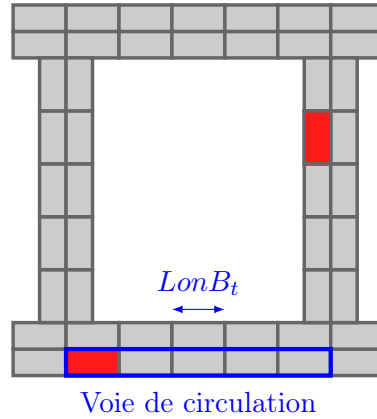


FIGURE 3.2 – Illustration d'un découpage en briques d'une carte. En rouge la présence d'un événement sur une brique. Source [6, 11].

3. la *date* $\mathbf{M.d}$ de perception de l'événement par la source S ;
4. la *localisation* $\mathbf{M.l}$ de l'événement qui correspond à une brique sur la carte qui a été discrétisée (cf figure 3.2). Ce découpage en brique est fonction du type de l'événement ;
5. la *fonction de masse* $\mathbf{M.m}$ représentant la confiance de la source S associée à l'événement. Cette fonction de masse s'exprime sur le cadre $\Omega = \{\exists, \nexists\}$, où $\exists$ indique que l'événement existe, et $\nexists$ qu'il n'est pas présent.

Les informations fournies par les conducteurs peuvent se faire par exemple via l'interface de saisie illustrée sur la figure 3.3. Cette interface a été déclinée dans une version pour téléphone Android™ dans [11]. Outre cette maquette, nous avons surtout travaillé avec des simulations Matlab™.

Plusieurs méthodes ont été construites en fonction de :

1. *La gestion des messages*. Deux stratégies ont été considérées. Une première stratégie consiste à garder tous les messages reçus et à les retransmettre à chaque véhicule croisé. Une deuxième stratégie consiste à fusionner les messages relatifs à chaque événement et à ne garder qu'un seul message issu de cette fusion pour chaque événement. Ainsi un seul message relatif à un événement et possiblement issu d'une fusion est présent dans la base d'un véhicule et peut être retransmis aux véhicules croisés. Le champ S d'un message fusionné contient l'union des sources ayant participé à l'information de ce message. Dans la première stratégie, les FM des messages concernant un même événement sont fusionnés conjonctivement pour donner la probabilité pignistique d'existence de l'événement au conducteur. Dans la deuxième stratégie, les FM des messages fusionnés sont combinés conjonctivement si les messages n'ont pas de sources en commun, sinon la combinaison prudente [23] introduite par T. Denœux est employée. Le conducteur reçoit aussi la probabilité pignistique de l'existence de l'événement.



FIGURE 3.3 – Interface de saisie possible pour créer les messages. Le conducteur indique le type de l'événement, s'il l'a perçu ou non (dans le cas où il aurait été informé de son existence), sa confiance par défaut à 0.6, et la localisation de l'événement (sur sa brique, sur la brique opposée, brique de droite ou brique de gauche). Source [6].

2. *La méthode de vieillissement.* Avant de combiner les FM des messages et de donner la probabilité pignistique résultante au conducteur, un mécanisme de vieillissement est employé sur les FM pour prendre en compte le temps écoulé depuis la perception de l'événement. Deux méthodes ont été considérées : l'affaiblissement classique (1.22) et le renforcement vers un élément singleton (1.31). L'affaiblissement de la FM $M.m$ d'un message M en fonction de la date de perception de l'événement $M.d$ consiste à considérer, qu'avec le temps, on ne peut rien dire sur la présence ou non de l'événement (méthode déjà employée dans cette application de communication inter-véhicule dans [15]). Au contraire le renforcement de $M.m$ vers $\#$ en fonction de $M.d$ met en avant le fait que l'événement avec le temps va disparaître.
3. La prise en compte de la *mise à jour du monde* [42, 49, 28], c'est-à-dire le fait de considérer que la vérité a changé à un instant donné du temps. Comme tous les véhicules sont sincères (la prise en compte de véhicules malveillants est une perspective de travail), lorsqu'un message reçu par un véhicule est en contradiction avec les informations qu'ils détenaient, tous les messages antérieurs à ce nouveau message sont supprimés.

Les méthodes testées sont alors celles qui sont résumées dans le tableau 3.1.

Dans la suite, nous avons choisi de présenter les résultats de deux scénarios de tests issus de [6]. Le premier met en avant les différences entre les deux mécanismes de vieillissement. Le second a pour but de comparer les méthodes introduites sur un scénario de plus grande échelle comprenant plus de véhicules et d'événements.

Tableau 3.1 – Résumé des méthodes testées pour la gestion des événements locaux.

Méthode	Combinaison
1 Garder les messages initiaux, affaiblir	Conjonctive
2 Garder les messages initiaux, renforcer	
3 Garder le résultat de la fusion, affaiblir	Conjonctive
4 Garder le résultat de la fusion, renforcer	ou prudente
5 Garder les messages initiaux, m à j du monde, affaiblir	Conjonctive
6 Garder les messages initiaux, m à j du monde, renforcer	

Ces scénarios ont été implémentés dans un simulateur sous Matlab™ (VMES Vehicle Message Exchange Simulator)¹ qui a été développé par Mira Bou Farah durant sa thèse.

3.2.3 Simulations

Dans un premier temps, nous présentons la mesure de performance employée pour comparer les méthodes. Puis, les deux scénarios de test sont présentés et leurs résultats analysés.

3.2.3.1 Mesure d'adéquation à la réalité employée

Les performances des méthodes sont mesurées en regardant l'adéquation entre les informations fournies aux conducteurs et la vérité terrain à chaque pas de temps. Formellement, la mesure de performance employée est définie pour chaque type d'événement t et véhicule v , à chaque instant τ , par :

$$Perf_t^{v,\tau} = 1 - \frac{\sum_{b \in B_t^{v,\tau}} \left(BetP_{t,b}^{v,\tau}(\{\exists\}) - R_{t,b}^\tau \right)^2}{|B_t^{v,\tau}|}, \quad (3.1)$$

où :

- $B_t^{v,\tau}$ englobe l'ensemble des briques sur lesquelles un événement est présent (réalité terrain) et l'ensemble des briques au regard desquelles des informations ont été donnés aux conducteurs ;
- $BetP_{t,b}^{v,\tau}$ est la probabilité pignistique dans la base du véhicule v au temps τ concernant la présence de l'événement (t, b) (sur une brique b il ne peut y avoir qu'un événement de type t) ;

1. Déposé sous licence CeCILL-B, ce simulateur est disponible sur le site du LGI2A à cette adresse : <http://www.lgi2a.univ-artois.fr/spip/spip.php?article182>

- $R_{t,b}^\tau$ est une variable binaire représentant la vérité terrain : $R_{t,b}^\tau = 1$ si l'événement (t, b) est présent dans la réalité terrain à l'instant τ , 0 sinon.

3.2.3.2 Tests sur deux scénarios

Dans ces expérimentations, le type d'événement *accident* est considéré, et la confiance associée aux FM des messages est égale à 0.6. Cela signifie que quand un véhicule perçoit un événement, il crée un message avec une FM : $\{\exists\}^{0.4}$, et quand un véhicule ne perçoit pas un événement présent dans sa base, il crée un message avec une FM : $\{\nexists\}^{0.4}$. La durée des accidents suit une loi normale D de moyenne μ et d'écart type σ . Nous choisissons un seuil d'obsolescence ou de suppression des messages *Suppr* tel que $P(D \leq \text{Suppr}) = .99\%$, soit $\text{Suppr} = \mu + u_{.99}$, où $u_{.99}$ est le 99^e centile de la loi normale centrée réduite (vallant 2.3263). Ce seuil permet de paramétrer les deux méthodes de vieillissement.

Pour chaque message M , au moment de fusionner les informations :

- l'affaiblissement de $M.m$ est donné par :

$${}^\alpha m = (1 - \alpha) m + \alpha m_\Omega , \quad (3.2)$$

avec le taux d'affaiblissement $\alpha = \frac{\text{date_actuelle} - M.d}{\text{Suppr}} \in [0, 1]$;

- et le renforcement vers $\nexists$ de $M.m$ par :

$${}^\nu m = (1 - \nu) m + \nu m_{\{\nexists\}} , \quad (3.3)$$

avec le taux de renforcement $\nu = \frac{\text{date_actuelle} - M.d}{\text{Suppr}} \in [0, 1]$.

Dès que la durée du seuil de suppression a été dépassée depuis la perception de l'événement, c'est-à-dire $\text{date_actuelle} - M.d \geq \text{Suppr}$, le message M est supprimé.

Scénario n°1 Dans ce premier scénario, illustré sur la figure 3.4, un accident est présent au début de la simulation et un véhicule v reçoit 2 messages l'informant de la présence de l'accident, puis 2 messages l'informant de l'absence de l'événement après sa disparition. La durée D de l'accident suit une loi normale de moyenne $\mu = 1800$ secondes (30mn) et d'écart type $\sigma = 300$.

En lançant une fois ce scénario nous obtenons pour la méthode 5 (garder les messages initiaux, màj du monde, affaiblir) la courbe de performances (3.1) en fonction du pas de temps illustrée sur la figure 3.5(a), et pour la méthode 6 (garder les messages initiaux, màj du monde, renforcer) celle donnée par la figure 3.5(b). L'objectif est d'avoir une performance la plus proche de 1 possible tout au long du scénario. Nous pouvons observer au cours du temps, l'accroissement vers 1 des performances des deux méthodes à la réception du premier message informant de l'événement. Puis avec le temps cette information est affaiblie pour la méthode 5, et renforcée vers $\nexists$ pour la méthode 6. Nous pouvons alors voir que le renforcement lorsque l'accident est toujours présent donne de moins bonnes performances que l'affaiblissement : les performances de la méthode 6 utilisant le renforcement

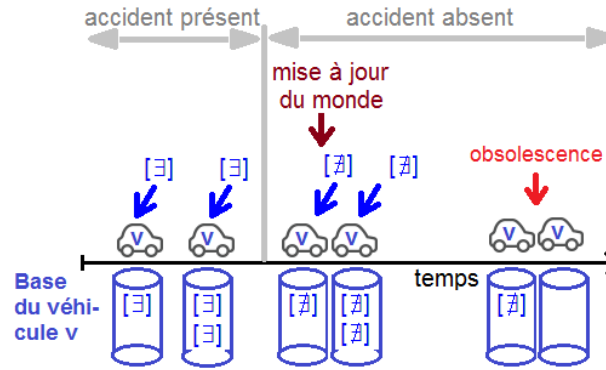


FIGURE 3.4 – Illustration du scénario n°1. Un véhicule v reçoit 2 messages l'informant de la présence de l'accident, puis 2 messages l'informant de l'absence de l'événement après sa disparition. Le symbole $[\exists]$ (resp. $[\nexists]$) indique un message en faveur de l'existence (resp. la non existence) de l'événement. Source [6].

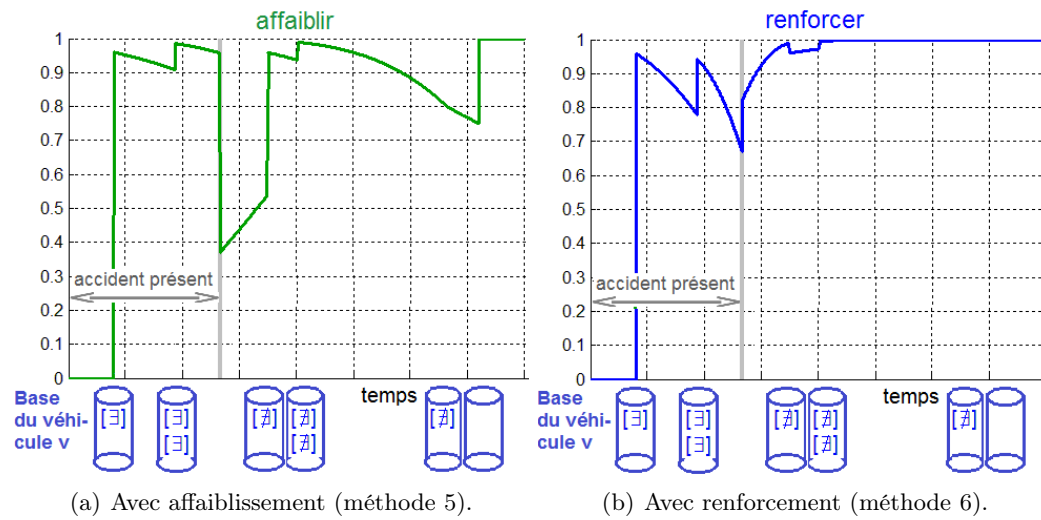


FIGURE 3.5 – Performances des méthodes 5 (garder les messages initiaux, mäj du monde, affaiblir) et 6 (garder les messages initiaux, mäj du monde, renforcer) concernant les informations données au conducteur du véhicule v au cours d'une instance du scénario n°1.

chutent plus vite que celles de la méthode 5 utilisant l'affaiblissement. Au moment de la réception d'un deuxième message informant de la présence de l'accident, les performances des deux méthodes croissent vers 1 avant de diminuer à nouveau sous l'effet des mécanismes de correction. Au moment de la disparition de l'accident la performance de la méthode 5 chute, alors que la méthode 6 a mieux anticipé cette disparition. Après la disparition de l'accident, avant d'avoir reçu le premier message informant de cette disparition, les performances des deux méthodes augmentent

mais dans cette situation c'est la courbe d'adéquation à la réalité de la méthode 6 qui progresse plus vite que celle de la méthode 5. Une fois les messages informant de la disparition reçus les performances des deux méthodes reviennent proche de 1. Néanmoins avec le temps, tant que les messages ne sont pas supprimés, l'affaiblissement présent dans la méthode 5 continue à être appliqué sur des FM en faveur de $\#$, ainsi les performances de la méthode 5 se dégradent, au contraire de celles de la méthode 6 utilisant le renforcement vers $\#$.

Afin d'obtenir des résultats de comparaison significatifs, nous avons pris un échantillon de taille 200 de durées d'accidents, et avons observé les valeurs moyennes des performances (3.1) obtenues par les méthode 5 et 6 sur ces 200 lancements. Les résultats sont les suivants :

- Méthode 5 (garder les messages initiaux, māj du monde, affaiblir) : 0.784 (écart type : 0.00601) ;
- Méthode 6 (garder les messages initiaux, māj du monde, renforcer) : 0.854 (écart type : 0.00439).

Sur ce scénario, la méthode 6 utilisant le renforcement vers $\#$ obtient donc de meilleures performances concernant son adéquation à la réalité que la méthode 5 employant l'affaiblissement.

Scénario n°2 Nous avons voulu tester les méthodes sur un scénario de plus grande échelle, et avons choisi ce scénario n°2 illustré sur la figure 3.6. Trois accidents vont survenir, chaque durée d'accident suivant la même loi $\mathcal{N}(600, 100^2)$, soit une durée moyenne de 10mn. Le premier accident (resp. le deuxième, le troisième) démarre au bout de 2mn (resp. 7mn, 12mn), la durée totale de la simulation étant de 1h. Deux sous-scénarios sont considérés, un premier avec un flot de 568 véhicules et un deuxième avec un flot de 184 véhicules, chaque véhicule ayant pour lieu de départ et lieu d'arrivée un des couples suivants illustrés la figure 3.6 : $(1, d)$, $(d, 1)$, $(2, c)$, $(c, 2)$, $(3, b)$, $(b, 3)$, $(4, a)$, $(a, 4)$.

Les résultats de performances moyennes obtenues sur l'ensemble des véhicules et des événements sur la durée de la simulation sont donnés dans le tableau 3.2. Ces tests n'ont pu être exécutés qu'une seule fois sur le simulateur développé du fait du temps de calcul qu'ils nécessitaient (de l'ordre de la semaine, les limites du simulateur étaient atteintes). Ces résultats montrent également la meilleure adéquation à la réalité des méthodes employant le renforcement par rapport à leurs duales employant l'affaiblissement.

3.2.4 Conclusion et perspectives

Nous avons présenté un résumé d'une partie du travail de thèse de Mira Bou Farah traitant des événements locaux. Nous avons choisi de montrer cet exemple d'application des fonctions de croyance en fusion d'informations pour mettre en avant un autre mécanisme de correction que l'affaiblissement qui est le renforcement vers un élément singleton. Ce mécanisme est mieux adapté au vieillissement des

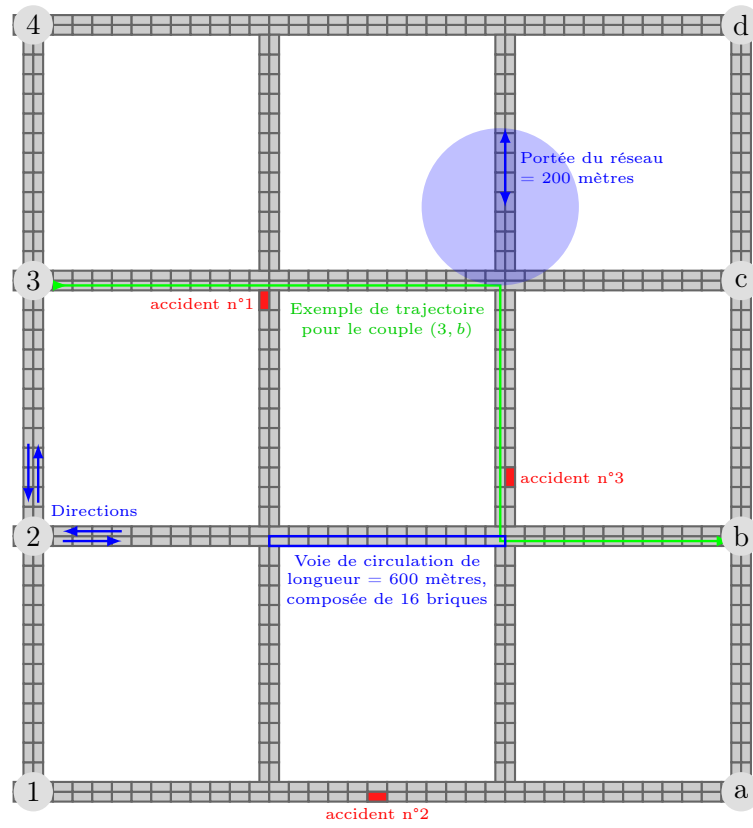


FIGURE 3.6 – Illustration du scénario n°2. Trois accidents surviendront. Chaque véhicule du scénario traverse la carte avec pour lieu de départ et lieu d'arrivée un des couples de positions $(1, d)$, $(d, 1)$, $(2, c)$, $(c, 2)$, $(3, b)$, $(b, 3)$, $(4, a)$ ou $(a, 4)$. Source [6].

informations concernant la présence sur la route d'un événement qui avec le temps disparaît.

Comme nous venons de le voir avec le second scénario, le simulateur introduit durant ces travaux arrive vite à des limites de fonctionnement. Une première perspective à ce travail est d'implémenter ces méthodes dans un simulateur du domaine, plus robuste et plus proche de la réalité comme par exemple SUMO (Simulation of Urban MObility) [45], et de vérifier les résultats obtenus sur les mêmes scénarios.

D'autres perspectives de travail se situent dans la gestion des événements spatiaux, c'est-à-dire des événements ayant une spatialité dont il est difficile de connaître les limites comme c'est le cas par exemple des embouteillages et des nappes de brouillard. Enfin, nous pouvons aussi évoquer la gestion des messages malveillants, sachant qu'un premier travail dans ce domaine avec des fonctions de croyance a déjà été réalisé [29].

Le problème d'association de données avec des fonctions de croyance est maintenant abordé dans la section suivante.

Tableau 3.2 – Performances moyennes obtenues sur l'ensemble des véhicules et des événements sur la durée de la simulation du scénario n°2 par chacune des méthodes (une seule simulation).

Moyennes sur toute la simulation (écart type = 0)	568 véhicules	184 véhicules
Garder les messages initiaux, affaiblir	(6) 0.7846	(5) 0.7047
Garder les messages initiaux, renforcer	(2) 0.8819	(1) 0.7913
Garder le résultat de la fusion, affaiblir	(5) 0.8051	(6) 0.6976
Garder le résultat de la fusion, renforcer	(4) 0.8448	(3) 0.7722
Garder les messages initiaux, mäj du monde, affaiblir	(3) 0.8475	(4) 0.7278
Garder les messages initiaux, mäj du monde, renforcer	(1) 0.8831	(2) 0.7906

3.3 Association de données avec des fonctions de croyance.

3.3.1 Description de l'application et des objectifs

Nous nous intéressons maintenant au problème d'association de données qui, étant donné deux ensembles finis dont les cardinalités peuvent être différentes, consiste à former des couples composés d'éléments de chacun des deux ensembles sachant que des éléments peuvent rester non associés.



(a) Véhicules détectés à l'avant d'un véhicule à un instant t (objets connus). (b) Véhicules détectés à l'avant du véhicule un instant plus tard $t' > t$ (objets perçus).

FIGURE 3.7 – Exemple de problème d'association de données appliqué aux véhicules intelligents. Comment les objets perçus X_i , $i \in \{1, 2, 3, 4\}$, et les objets connus Y_j , $j \in \{1, 2, 3, 4, 5\}$ sont-ils associés ?

Ce problème se rencontre notamment dans les applications de suivi multi-objets lorsqu'il y a besoin de mettre en correspondance les objets perçus (observés par un capteur) et une estimation de la position de ces objets donnée par un système de filtrage [3, 4]. Un exemple appliqué aux véhicules intelligents, qui est issu de notre laboratoire [59], est illustré sur la figure 3.7.

Ce problème d'affectation est très connu en recherche opérationnelle lorsqu'un coût numérique est donné à chaque association, il est résolu efficacement (en temps polynomial) grâce à l'algorithme hongrois (ou Kuhn-Munkres) [46, 12] qui calcule l'ensemble composé des associations dont la somme des coûts est minimale. Dans ces travaux, nous considérons la situation où les informations concernant les associations possibles entre chaque paire d'objets sont modélisées par des fonctions de croyance, ce qui permet de prendre en compte des informations plus fines au regard de chaque association.

Dans les années 2009 et 2010, les approches existantes utilisant des fonctions de croyance [70, 69, 32, 33, 50, 48] étaient basées sur des heuristiques. Nous avons alors proposé une solution [57, 58, 59] utilisant des outils classiques des fonctions de croyance pour résoudre ce problème, à savoir l'extension vide, la combinaison conjonctive et la prise de décision par maximum de probabilité pignistique. Cette méthode a en particulier mis en évidence les meilleures performances de la combinaison conjonctive par rapport à la méthode de fusion proposée par Rombaut et Cherfaoui [70, 69] sur un problème d'association de véhicules détectés devant un véhicule (cf figure 3.7) [59], ces données étant issues d'une caméra située au niveau du tableau de bord avant d'un véhicule.

Récemment, en 2014, Denœux et al. [24] ont introduit une nouvelle méthode répondant à deux problèmes posés par notre méthode, à savoir un problème de symétrie et une complexité calculatoire importante. Les auteurs ont aussi comparé les méthodes sur plusieurs tests où ils ont obtenu des résultats similaires de performances pour les deux méthodes.

Pour ce rapport, nous exposons en section 3.3.2 un résumé de la méthode que nous avons introduite pour résoudre le problème d'association avec des fonctions de croyance [57, 58, 59], puis la méthode de Denœux et al. dans la section 3.3.3, avant d'illustrer et de comparer ces méthodes, en section 3.3.4, au travers d'une application dans le cadre du suivi multi-cibles issue des travaux de thèse de Samir Hachour [34]. Enfin, une conclusion et des perspectives à ce travail sont exposées dans la section 3.3.5.

3.3.2 Un algorithme d'association de données utilisant des fonctions de croyance

Nous considérons d'abord les notations suivantes :

- X_i : désigne un objet i perçu, $i \in I = \{1, \dots, N\}$, N étant le nombre d'objets perçus,
- Y_j : désigne un objet j connu à l'instant précédent, $j \in J = \{1, \dots, M\}$, M étant le nombre d'objets connus,

- $*$: une proposition signifiant aucun objet. Si un objet perçu X_i est associé à la proposition $*$, cela signifie qu'il est apparu. Si un objet connu Y_j est associé à la proposition $*$, cela signifie qu'il a disparu, il n'est plus présent dans la scène en cours.

L'objectif du processus d'association consiste alors à trouver le meilleur ensemble de couples entre un ensemble d'objets perçus $\{X_1, X_2, \dots, X_N, *\}$ et un ensemble d'objets connus $\{Y_1, Y_2, \dots, Y_M, *\}$, sous les contraintes suivantes :

- chaque objet perçu X_i est associé avec au maximum un objet connu,
- chaque objet connu Y_j est associé avec au maximum un objet perçu,
- la proposition $*$ peut être associée à plusieurs objets.

Les cadres de discernement mis en jeu sont les suivants :

- Chaque cadre $\Omega_{i,j} = \{o_{i,j}, n_{i,j}\}$, $i \in I$, $j \in J$, correspond aux deux valeurs possibles (oui ou non) de la variable associée à la réponse à la question d'intérêt *Est-ce que l'objet X_i est associé à l'objet connu Y_j ?*
- Chaque cadre $\Omega_{X_i} = \{Y_1, Y_2, \dots, Y_M, *\} = \{1, 2, \dots, M, *\}$, $i \in I$, regroupe les réponses à la question d'intérêt *Quel objet est associé à l'objet perçu X_i ?*
- Chaque cadre $\Omega_{Y_j} = \{X_1, X_2, \dots, X_N, *\} = \{1, 2, \dots, N, *\}$, $j \in J$, regroupe les réponses à la question d'intérêt *Quel objet est associé à l'objet connu Y_j ?*

En entrée de la phase d'association nous avons donc $N \times M$ MF $m^{\Omega_{i,j}}$ s'exprimant sur l'association possible de chaque couple (i, j) .

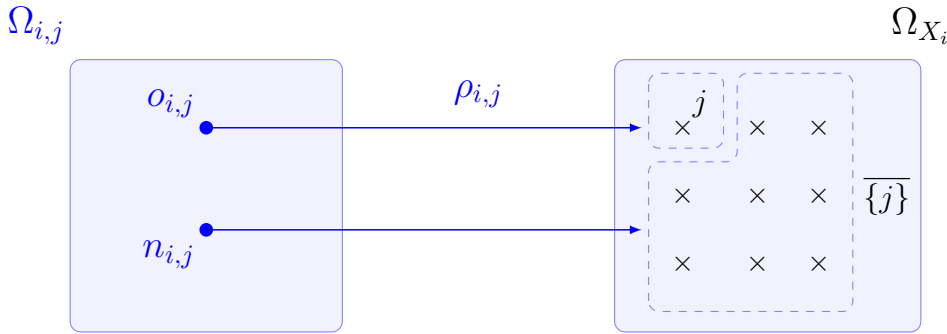


FIGURE 3.8 – Illustration du raffinement $\rho_{i,j}$ transportant l'information de $\Omega_{i,j}$ à Ω_{X_i} .

Une solution consiste alors à effectuer les opérations suivantes :

1. Exprimer chaque information $m^{\Omega_{i,j}}$ sur le cadre Ω_{X_i} par l'opération d'extension vide (1.14) associé au raffinement illustré sur la figure 3.8. Ce raffinement exprime simplement le fait que si i n'est pas associé à j alors i est associé à un autre objet que j (un objet dans $\overline{j}$). Le résultat $m^{\Omega_{i,j} \uparrow \Omega_{X_i}}$ est noté $m_j^{\Omega_{X_i}}$;
2. Combiner conjonctivement (1.8) ces FM $m_j^{\Omega_{X_i}}$, $j \in J$, le résultat étant noté $m^{\Omega_{X_i}}$;

3. Enfin, choisir la décision maximisant la probabilité $BetP^{\Omega_{X_1} \times \dots \times \Omega_{X_N}}$ (1.33) qui vérifient les contraintes d'associations.

Cette solution est illustrée par l'exemple suivant.

Exemple 3.1 On considère un objet perçu X_1 et deux objets connus Y_1 et Y_2 tels que :

$$\begin{cases} m^{\Omega_{1,1}}(\{y_{1,1}\}) = .2 \\ m^{\Omega_{1,1}}(\{n_{1,1}\}) = .45 \\ m^{\Omega_{1,1}}(\Omega_{1,1}) = .35 \end{cases} \quad \begin{cases} m^{\Omega_{1,2}}(\{y_{1,2}\}) = .45 \\ m^{\Omega_{1,2}}(\{n_{1,2}\}) = .15 \\ m^{\Omega_{1,2}}(\Omega_{1,2}) = .4 \end{cases}$$

En exprimant ces informations sur Ω_{X_1} , on a :

$$\begin{cases} m_1^{\Omega_{X_1}}(\{1\}) = .2 \\ m_1^{\Omega_{X_1}}(\overline{\{1\}}) = .45 \\ m_1^{\Omega_{X_1}}(\Omega_{X_1}) = .35 \end{cases} \quad \begin{cases} m_2^{\Omega_{X_1}}(\{2\}) = .45 \\ m_2^{\Omega_{X_1}}(\overline{\{2\}}) = .15 \\ m_2^{\Omega_{X_1}}(\Omega_{X_1}) = .4 \end{cases}$$

Puis, après combinaison :

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{\star\}$	$\{1, \star\}$	$\{2, \star\}$	$\{1, 2, \star\}$
$m^{\Omega_{X_1}}(A)$	.09	.11	.36	.07	.05	.18	.14
$BetP^{\Omega_{X_1}}(A)$		.20	.55	.25			

En conclusion nous avons alors :

- X_1 est associé à Y_2 .
- Y_1 a donc disparu.

Dans cette solution que nous venons de présenter, nous avons choisi d'étendre les informations initiales sur les cadres Ω_{X_i} , c'est-à-dire selon le point de vue des objets perçus (avec qui est associé chaque X_i ?). Nous aurions pu exprimer ces mêmes informations selon le point de vue des objets connus, c'est-à-dire étendre les informations exprimées sur les cadres $\Omega_{i,j}$ vers les cadres Ω_{Y_j} . L'exemple suivant développe cette solution en reprenant les données de l'exemple 3.1 précédent.

Exemple 3.2 En étendant les données initiales de l'exemple 3.1 sur les cadres Ω_{Y_1} et Ω_{Y_2} , nous obtenons :

$$\begin{cases} m_1^{\Omega_{Y_1}}(\{1\}) = .2 \\ m_1^{\Omega_{Y_1}}(\overline{\{1\}}) = .45 \\ m_1^{\Omega_{Y_1}}(\Omega_{Y_1}) = .35 \end{cases} \quad \begin{cases} m_1^{\Omega_{Y_2}}(\{1\}) = .45 \\ m_1^{\Omega_{Y_2}}(\overline{\{1\}}) = .15 \\ m_1^{\Omega_{Y_2}}(\Omega_{Y_2}) = .4 \end{cases}$$

Comme il n'y a qu'une seule information sur cet exemple pour chacun des objets connus Y_1 et Y_2 , il n'y a pas de combinaison à réaliser, et nous avons directement les résultats suivants en termes de probabilités pignistiques :

A	$\emptyset$	$\{1\}$	$\{\star\}$	$\{1, \star\}$
$m^{\Omega_{Y_1}}(A)$		.2	.45	.35
$BetP^{\Omega_{Y_1}}(A)$		.375	.625	1
$m^{\Omega_{Y_2}}(A)$		.45	.15	.4
$BetP^{\Omega_{Y_2}}(A)$		.65	.35	1

D'où :

- $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}(\{1, \star\}) = .375 \times .35 = .131$;
- $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}(\{\star, 1\}) = .625 \times .65 = .406$;
- $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}(\{\star, \star\}) = .625 \times .35 = .219$.

La probabilité jointe maximale est donc obtenue pour Y_1 associé à $\star$ et Y_2 associé à X_1 . On a le même résultat selon le point de vue des objets perçus X_i et selon le point de vue des objets connus Y_j . Malheureusement, ce n'est pas toujours le cas comme nous pouvons le voir avec l'exemple 3.3 suivant.

Exemple 3.3 Considérons un objet perçu X_1 , et deux objets connus Y_1 et Y_2 , tels que :

$$\begin{cases} m^{\Omega_{1,1}}(\{o_{1,1}\}) = 0.5 \\ m^{\Omega_{1,1}}(\{n_{1,1}\}) = 0 \\ m^{\Omega_{1,1}}(\Omega_{1,1}) = 0.5 \end{cases} \quad \begin{cases} m^{\Omega_{1,2}}(\{o_{1,2}\}) = 0.7 \\ m^{\Omega_{1,2}}(\{n_{1,2}\}) = 0.3 \\ m^{\Omega_{1,2}}(\Omega_{1,2}) = 0 \end{cases} \quad (3.4)$$

En exprimant les croyances du point de vue de l'objet perçu X_1 , on a :

$$\begin{cases} m_1^{\Omega_{X_1}}(\{1\}) = 0.5 \\ m_1^{\Omega_{X_1}}(\{\overline{1}\}) = 0 \\ m_1^{\Omega_{X_1}}(\Omega_{X_1}) = 0.5 \end{cases} \quad \begin{cases} m_2^{\Omega_{X_1}}(\{2\}) = 0.7 \\ m_2^{\Omega_{X_1}}(\{\overline{2}\}) = 0.3 \\ m_2^{\Omega_{X_1}}(\Omega_{X_1}) = 0 \end{cases}, \quad (3.5)$$

on obtient :

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{\star\}$	$\{1, \star\}$	$\{2, \star\}$	$\{1, 2, \star\}$
$m^{\Omega_{X_1}}(A)$	0.35	0.15	0.35	0	0.15	0	0
$BetP^{\Omega_{X_1}}(A)$		0.35	0.54	0.11			

(3.6)

Ainsi au regard de X_1 :

- X_1 est associé à Y_2 ,
- Y_1 a disparu.

Du côté de Y_1 et Y_2 , on a :

$$\begin{cases} m_1^{\Omega_{Y_1}}(\{1\}) = 0.5 \\ m_1^{\Omega_{Y_1}}(\overline{\{1\}}) = 0 \\ m_1^{\Omega_{Y_1}}(\Omega_{Y_1}) = 0.5 \end{cases} \quad \begin{cases} m_1^{\Omega_{Y_2}}(\{1\}) = 0.7 \\ m_1^{\Omega_{Y_2}}(\overline{\{1\}}) = 0.3 \\ m_1^{\Omega_{Y_2}}(\Omega_{Y_2}) = 0 \end{cases} \quad (3.7)$$

D'où :

A	$\{1\}$	$\{\star\}$
$BetP^{\Omega_{Y_1}}$	0.75	0.25
$BetP^{\Omega_{Y_2}}$	0.7	0.3

(3.8)

Comme $0.75 \times 0.3 > 0.7 \times 0.25$, $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}$ atteint son maximum valide en $\{1, \star\}$, ce qui implique une conclusion différente de la précédente selon X_1 :

- Y_1 associé à X_1 ,
- Y_2 a disparu.

3.3.3 Méthode de Denœux et al.

Récemment Denœux et al. [24] ont proposé une nouvelle solution à ce problème d'association. Comme les auteurs l'indiquent dans leurs tests [24, Section IV] il n'y a pas de différences notables de résultats entre la méthode qu'ils ont introduite et celle que nous venons d'exposer. Néanmoins, la solution de Denœux et al. possède deux importantes qualités par rapport à la nôtre :

1. Elle est symétrique ;
2. Et elle possède une complexité calculatoire bien moindre qui est la complexité de résolution de l'algorithme d'association classique avec des données numériques ($\mathcal{O}(n^3)$ par exemple pour l'algorithme hongrois [46]).

Pour réussir ces points, les auteurs de [24] sont parvenus à exprimer et manipuler les croyances sur le cadre $\mathcal{R}$ composé de toutes les relations possibles entre les objets X_i et Y_j , puis à prendre une décision à partir de ce cadre. Pour ce faire, ils ont utilisé la fonction contour et une décision basée sur le maximum de plausibilité.

Nous présentons cette méthode après avoir précisé les définitions et notations suivantes, pour tout $i \in I = \{1, \dots, N\}$ et $j \in J = \{1, \dots, M\}$:

- une variable binaire r_{ij} est définie telle que $r_{ij} = 1$ signifie que X_i est associé à Y_j , $r_{ij} = 0$ sinon ;
- l'ensemble $\mathcal{R}_{ij}$ désigne l'ensemble des relations où X_i est associé à Y_j , i.e. $\mathcal{R}_{ij} = \{r \in \mathcal{R}, r_{ij} = 1\}$;

- chaque FM initiale $m^{\Omega_{i,j}}$ est notée m_{ij} , et les réels $\alpha_{ij} < 1$ et $\beta_{ij} < 1$ sont introduits tels que $m_{ij}(\{o_{i,j}\}) = \alpha_{ij}$, $m_{ij}(\{n_{i,j}\}) = \beta_{ij}$ et $m_{ij}(\{o_{i,j}, n_{i,j}\}) = 1 - \alpha_{ij} - \beta_{ij}$.

La méthode de Denœux et al. [24] consiste alors à d'abord étendre par extension vide toutes les FM m_{ij} sur le cadre $\mathcal{R}$ en transférant α_{ij} sur $\mathcal{R}_{ij}$, β_{ij} sur $\overline{\mathcal{R}_{ij}}$, et $1 - \alpha_{ij} - \beta_{ij}$ sur $\mathcal{R}$.

La fonction contour $pl_{ij}^{\uparrow \mathcal{R}}$ associée à l'extension sur $\mathcal{R}$ de m_{ij} est alors donnée par :

$$pl_{ij}^{\uparrow \mathcal{R}}(r) = (1 - \beta_{ij})^{r_{ij}} (1 - \alpha_{ij})^{1-r_{ij}} \quad (3.9)$$

La fonction contour pl associée à la combinaison par la règle de Dempster (1.9) des $N \times M$ fonctions $pl_{ij}^{\uparrow \mathcal{R}}$ est alors donnée (d'après (1.11) et comme les fonctions de plausibilité et de communalité coïncident sur les singletons) par :

$$pl(r) \propto \prod_{i,j} (1 - \beta_{ij})^{r_{ij}} (1 - \alpha_{ij})^{1-r_{ij}}. \quad (3.10)$$

D'où

$$\ln pl(r) = \sum_{i,j} (r_{ij} \ln(1 - \beta_{ij}) + (1 - r_{ij}) \ln(1 - \alpha_{ij})) + C, \quad (3.11)$$

où C est une constante.

La relation maximisant la plausibilité pl est alors choisie, c'est-à-dire celle maximisant $\sum_{i,j} \omega_{ij} r_{ij}$ avec $\omega_{ij} = \ln \left(\frac{1-\beta_{ij}}{1-\alpha_{ij}} \right)$.

La méthode peut alors se résumer à la résolution du problème d'optimisation linéaire en nombres entiers suivant :

$$\max \left(\sum_{i \in I} \sum_{j \in J} \omega_{ij} r_{ij} \right) \quad (3.12)$$

avec les contraintes d'association suivantes :

$$\sum_{i \in I} r_{ij} \leq 1 \quad \forall j \in J \quad (3.13)$$

$$\sum_{j \in J} r_{ij} \leq 1 \quad \forall i \in I \quad (3.14)$$

$$r_{ij} \in \{0, 1\} \quad \forall i \in I, \quad \forall j \in J, \quad (3.15)$$

ce problème pouvant être résolu avec une complexité polynomiale par l'algorithme hongrois [46] par exemple.

3.3.4 Comparaison sur un scénario de suivi de cibles avec des données cinématiques

Dans ses travaux de thèse [34, 39, 36, 37, 35, 38] (2011 à 2015), Samir Hachour a développé et implémenté sous MatlabTM des méthodes pour le suivi et la classification multi-objets. La partie suivie est principalement basée sur des IMM (Interactive Multiple Model) [4] utilisant des filtres de Kalman. La partie classification étend à plusieurs objets des précédents de travaux de Smets et Ristic [79] portant sur le suivi et la classification avec des fonctions de croyance d'un objet à partir de données cinématiques. Comme indiqué en introduction, le suivi de plusieurs objets avec ce système de filtrage (IMM et filtres de Kalman) nécessite une phase d'association pour mettre en correspondance les mesures (observations) et les prédictions (pistes). C'est sur ce système de suivi à base d'IMM que nous comparons notre méthode et celle de Denœux et al.

Dans ce domaine, les objets perçus X_i , $i \in I = \{1, \dots, N\}$, sont les observations (mesures) z_i de la réalité, et chaque objet connu Y_j , $j \in J = \{1, \dots, M\}$, est une prédiction $\bar{z}_j$ fournie par un IMM en charge de suivre un objet. L'IMM fournit également en sortie la matrice de covariance S_j de $\bar{z}_j$ [4], ainsi la distance choisie entre les observations z_i et les prédictions $\bar{z}_j$ est la distance de Mahalanobis [4] qui est définie, z_i et $\bar{z}_j$ étant supposés avoir la même matrice de covariance S_j , par :

$$d_{ij} = \sqrt{(\bar{z}_j - z_i)' S_j (\bar{z}_j - z_i)}. \quad (3.16)$$

Pour chaque étape d'association, les FM initiales $m^{\Omega_{i,j}} = m_{ij}$ s'exprimant sur l'association entre chaque observation z_i et prédiction $\bar{z}_j$ sont construites à partir de cette distance d_{ij} (3.16). En s'inspirant de travaux de Denœux [21], elles sont définies par :

$$\begin{cases} m^{\Omega_{i,j}}(\{o_{i,j}\}) &= \beta \exp^{-\gamma d_{ij}} \\ m^{\Omega_{i,j}}(\{n_{i,j}\}) &= \beta (1 - \exp^{-\gamma d_{ij}}) \\ m^{\Omega_{i,j}}(\Omega_{i,j}) &= 1 - \beta \end{cases} \quad (3.17)$$

avec $\beta \in [0, 1]$ le degré de fiabilité de la source choisi égal à 0.9 dans cette application et γ un paramètre à fixer.

Nous comparons alors notre méthode et celle de Denœux et al. dans la situation de deux objets évoluant en parallèle l'un de l'autre, ce scénario étant illustré sur la figure 3.9.

En faisant varier le paramètre γ des FM m_{ij} (3.17), nous obtenons différents taux d'erreurs d'association présentés sur la figure 3.10. Nous avons également comparé les deux méthodes avec une méthode classique d'association dans ce domaine du suivi avec des IMM, qui est l'algorithme GNN (Global Nearest Neighbour) [4] résumé en annexe B. Ces résultats illustrés sur la figure 3.10 corroborent ceux donnés dans [24] : les deux méthodes ont aussi des performances similaires sur cette application. Nous pouvons aussi observer que sur ce scénario, le GNN possède également des performances voisines.

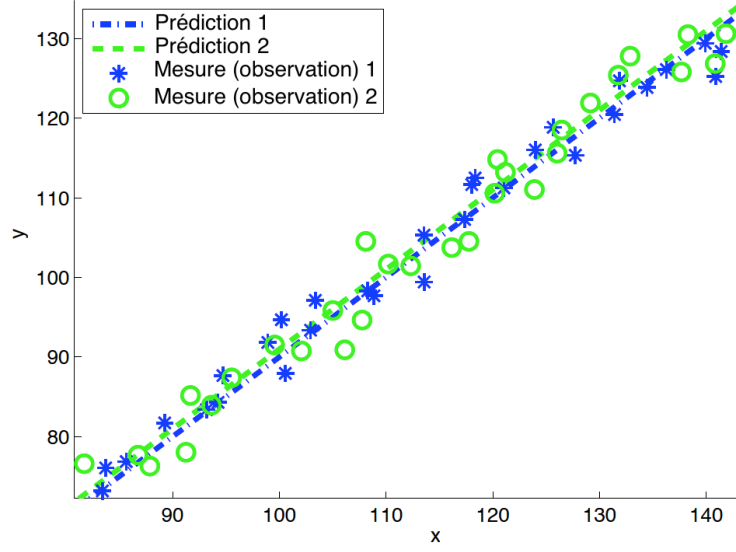


FIGURE 3.9 – Scénario sur lequel sont comparés les méthodes. Deux objets se déplacent en parallèle sur le plan. Source [34].

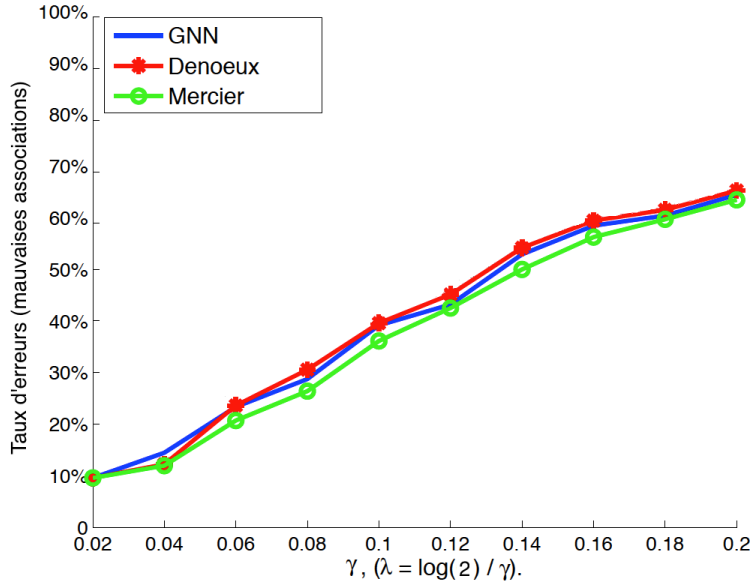


FIGURE 3.10 – Performances obtenues par chacune des méthodes sur le scénario illustré sur la figure 3.9 en fonction de la valeur du paramètre γ choisi dans l'équation des FM m_{ij} (3.17) pour la méthode que nous avons introduite et celle de Denœux et al. Le GNN possède un paramètre λ qui peut être déduit de γ (voir annexe B). Source [34].

Mentionnons pour conclure que ces mêmes observations peuvent être faites sur d'autres scénarios développés dans [34].

Enfin nous pouvons aussi voir le temps de calcul supérieur demandé par notre méthode sur la figure 3.11.

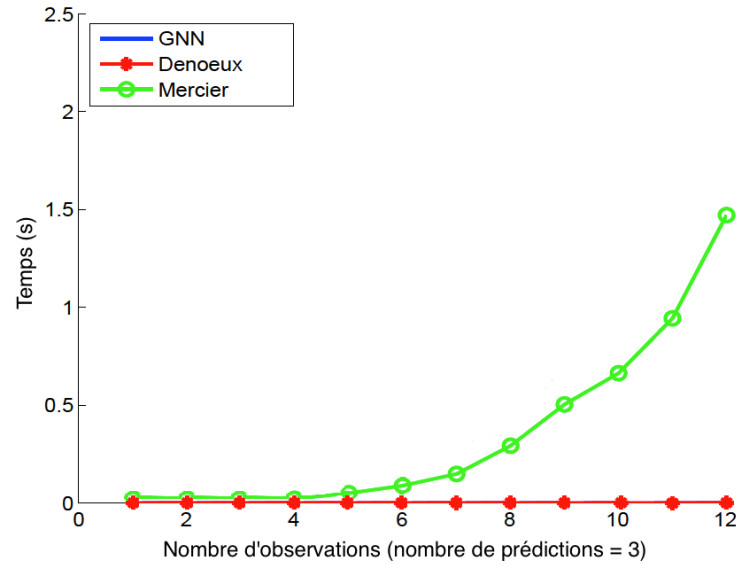


FIGURE 3.11 – Évolution du temps de calcul de chacune des méthodes en augmentant le nombre d'objets (augmentation du nombre d'observations pour 3 prédictions). Source [34].

3.3.5 Conclusion et perspectives

Nous venons de voir sur ce problème d'association que deux systèmes utilisant les fonctions de croyance et initialisés avec les mêmes informations, peuvent conduire à des différences de résultats (non symétrie, complexités significativement différentes). La bonne méthode consistait effectivement à exprimer toutes les informations sur l'ensemble $\mathcal{R}$ de toutes les relations possibles entre les objets, et à choisir la meilleure des relations. Mais autant exprimer toutes les croyances sur $\mathcal{R}$ était chose aisée, autant prendre la décision de la meilleure relation à partir de ce cadre était difficile pour nous au moment où nous pensions en termes de probabilités pignistiques. Contrairement à la probabilité pignistique qui nécessite la connaissance de toutes les cardinalités des sous-ensembles de $\mathcal{R}$ (ce qui nous semblait être un problème compliqué), la décision basée sur la plausibilité a uniquement besoin des plausibilités sur les singletons, ce qui permet aussi un calcul plus rapide.

Concernant les perspectives de ce travail, nous pouvons voir deux pistes. La première, déjà abordée dans [24], consiste à vouloir résoudre un problème de plus grande taille : le problème d'association avec un nombre quelconque d'ensemble d'objets. La seconde piste de travail est liée à l'application de suivi multi-objets

en faisant plusieurs hypothèses d'associations lorsque les incertitudes concernant les associations entre objets sont trop fortes. Les fonctions de croyance étant un modèle intéressant pour modéliser les informations pauvres (imprécises et incertaines) par rapport au modèle bayésien classique, des travaux en ce sens dans ce domaine pourraient donner des bons résultats.

Une dernière application sur laquelle j'ai travaillé est maintenant présentée. Même si, à ce jour, il ne s'agit principalement que d'un transfert de technologie, j'ai choisi de la mettre dans ce rapport car d'une part, je souhaite continuer à développer ce travail que je trouve prometteur (ce que je vais essayer de mettre en évidence dans la suite), et d'autre part, elle souligne une des difficultés de la modélisation d'un problème avec des fonctions de croyance, qui est la construction des fonctions de croyance.

3.4 Application au diagnostic du vieillissement du bobinage de moteurs électriques

3.4.1 Description de l'application et des objectifs

Grâce à des financements internes à l'université d'Artois (BQR - Bonus Qualité Recherche), nous avons pu travailler avec des électrotechniciens du laboratoire LSEE de l'université d'Artois sur le diagnostic du vieillissement du bobinage de machines électriques tournantes asynchrones ou synchrones, en utilisant le champ électromagnétique externe, ce qui a pour avantage de créer un diagnostic non invasif ne nécessitant pas de démontage de la machine [40, 64, 63, 68].

Nous présentons ici un résumé de l'approche développée dans [67, 60].

L'objectif de ce travail est donc de déterminer les défauts naissants des spires d'une machine électrique, afin de prévenir en particulier un arrêt brutal de celle-ci.

3.4.2 Solution développée

La solution développée se base sur des mesures du champ électromagnétique externe d'une machine tout en faisant augmenter la charge sur cette machine. Des travaux menés au sein du LSEE [13, 14, 68] ont en effet montré qu'en cas de défaut sur une spire, les mesures d'une certaine harmonique du champ électromagnétique issues de deux capteurs placés à 180 degrés l'un de l'autre évoluent en sens contraire avec l'augmentation de la charge, alors que ce n'est pas le cas s'il n'y a pas de défaut. La valeur de l'harmonique exploitée dépend du type de la machine à diagnostiquer [13, 14], par exemple pour la machine asynchrone triphasée utilisée dans nos travaux [67, 60] et illustrée sur la figure 3.12, la fréquence de l'harmonique est proche de 850 Hz.

Pour diagnostiquer un défaut, les capteurs doivent être placés en face de la phase en défaut. Il faut donc déplacer les capteurs dans plusieurs positions pour être sûr de diagnostiquer tous les défauts. Néanmoins, comme il n'est pas possible

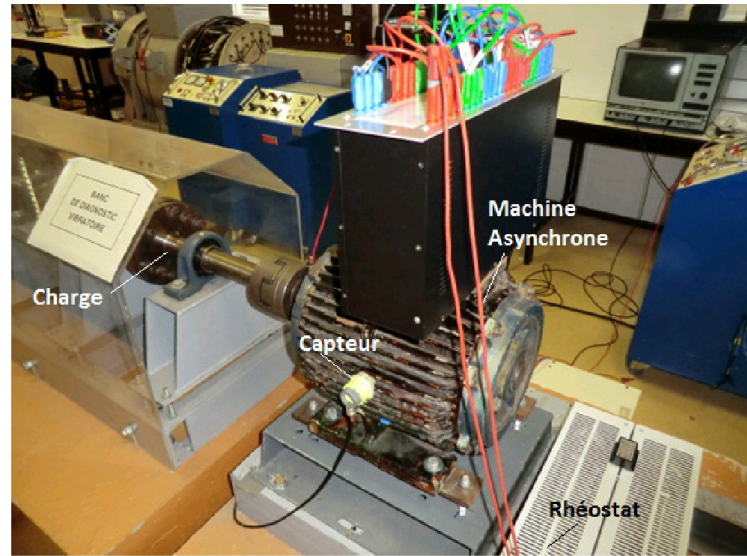


FIGURE 3.12 – Photo du banc d'essai et de la machine employés pour obtenir les résultats exposés dans ce chapitre dans la section 3.4.3. La machine peut fonctionner sous différentes conditions de charge. Les deux capteurs placés de chaque côté de la machine à l'opposé l'un de l'autre permettent de mesurer le champ électromagnétique externe de la machine. Le boîtier situé au-dessus de la machine permet de simuler des défauts (court-circuits) inter-spires. Source [60, 67].

de couvrir toute la périphérie de la machine, seules les quatre positions illustrées sur la figure 3.13 ont été utilisées.

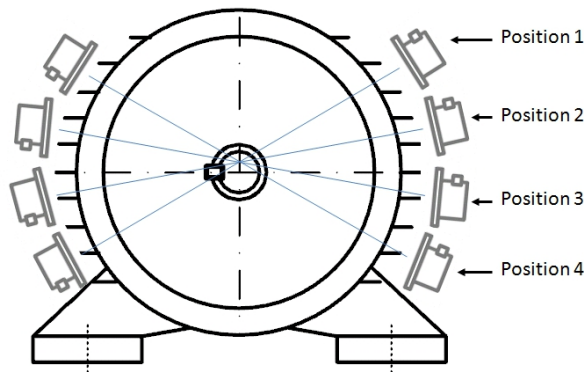


FIGURE 3.13 – Illustration des quatre positions de diagnostic possibles employés pour les capteurs. Source [13, 67, 60].

Pour chaque position i , $i \in \{1, 2, 3, 4\}$, des capteurs nous faisons quatre mesures issues de chacun des capteurs associées aux quatre choix de charges suivantes : 0W, 600W, 1000W et 1400W. Autrement dit, pour chaque position, nous partons

d'une charge nulle, puis nous augmentons la charge à 600W, puis 1000W, et enfin 1400W en relevant pour chaque charge la valeur de l'harmonique d'intérêt du champ électromagnétique. Si entre deux charges l'évolution de cette valeur n'est pas la même pour les deux capteurs, pour un capteur elle croît et pour l'autre elle décroît, alors il y a sûrement un défaut du bobinage. Un degré d'incertitude demeure car il est possible par exemple qu'il y ait eu des erreurs de mesures ou que les capteurs aient été mal alignés. Si pour aucune des montées en charge il n'y a eu de différences d'évolution, le fait qu'il y ait un défaut ou pas demeure incertain. Nous ne pouvons pas dire qu'il n'y a pas de défaut sachant qu'un défaut peut être sur une autre spire. Nous ne pourrions pas dire non plus qu'il n'y a pas de défaut sur telle spire car la localisation du défaut est très approximative [13, 14] (la localisation des défauts est une de nos perspectives de travail).

Nous avons fait le choix de modéliser cette technique de diagnostic dans le cadre des fonctions de croyance de la manière suivante. Nous considérons le cadre de discernement $\Omega = \{o, n\}$ composé de deux éléments o et n tels que o signifie que la machine a au moins un défaut, et n qu'elle n'a pas de défaut. Pour chaque position i , $i \in \{1, 2, 3, 4\}$, une FM $m_{var,i}^\Omega$ est créée ainsi :

- s'il y a une différence de variation durant une des montées en charge : $m_{var,i}(\{o\}) = 0.95$ et $m_{var,i}(\Omega) = 0.05$ (il y a sûrement un défaut) ;
- sinon (il n'y a eu aucune différence de variation durant les montées en charge) : $m_{var,i}(\{n\}) = 0.05$ et $m_{var,i}(\Omega) = 0.95$. Nous ne savons pas s'il y a un défaut, il y a une petite chance qu'il n'y en ait pas, il y a peut-être un défaut à une autre position.

Une autre information complémentaire peut-être exploitée. Elle concerne les valeurs numériques observées pour les mesures des capteurs. Pour une même charge, dans une même position, plus celles-ci sont éloignées plus on peut supposer la présence d'au moins un défaut. Pour chaque position i , $i \in \{1, 2, 3, 4\}$, nous avons donc défini des FM $m_{val,i}^\Omega$ à partir de trois seuils $S_j \in \mathbb{R}^+$, $j \in \{1, 2, 3\}$, suivant les valeurs absolues des différences de valeurs observées :

- si le maximum (en valeur absolue) des différences est inférieur à un seuil S_1 : $m_{val,i}(\{n\}) = 0.05$ et $m_{val,i}(\Omega) = 0.95$ (nous ne savons pas s'il y a un défaut mais peut-être que non) ;
- si cette valeur est comprise entre S_1 et un autre seuil S_2 , nous pouvons commencer à penser qu'il y a un défaut : $m_{val,i}(\{o\}) = 0.05$ et $m_{val,i}(\Omega) = 0.95$;
- avec une différence comprise entre S_2 et un autre seuil S_3 , la présence d'un défaut est fort probable : $m_{val,i}(\{o\}) = 0.5$ et $m_{val,i}(\Omega) = 0.5$;
- sinon (la différence est supérieure à S_3) un défaut est sûrement présent : $m_{val,i}(o) = 0.8$ et $m_{val,i}(\Omega) = 0.2$.

Exemple 3.4 *Supposons que pour une position i donnée nous ayons obtenu les résultats de mesures indiquées dans le tableau 3.3.*

Concernant l'information issue du sens de variation des mesures avec la montée en charge, il y a eu au moins une différence de variation entre les capteurs (par

Tableau 3.3 – Mesures obtenues de l’harmonique pour les différentes charges pour les deux capteurs (C_1 et C_2) dans une des positions $i \in \{1, 2, 3, 4\}$ illustrées sur la figure 3.13. Les variations des mesures entre chaque montée en charge sont indiquées pour chacun des capteurs. La dernière colonne donne la valeur absolue de la différence de mesures entre les deux capteurs pour la même charge.

Charge (W)	C_1 (μV)	Var. C_1 (μV)	C_2 (μV)	Var. C_2 (μV)	Diff. var. ?	Diff. mes. (μV)
0	314		436			122
600	103	−211	529	93	oui	426
1000	338	235	615	86	non	277
1400	503	165	596	−19	oui	93

exemple entre la charge nulle et celle à 600W : la mesure de C_1 a décru de 314 μV à 103 μV alors que celle de C_2 a augmenté de 436 μV à 529 μV). Par la méthode proposée on a alors $m_{var,i}(\{o\}) = 0.95$ et $m_{var,i}(\Omega) = 0.05$.

Concernant l’observation des différences de mesures effectuées par chaque capteur pour la même charge, on a un maximum de 426 μV obtenu pour la charge à 600 W. Avec des seuils S_1 , S_2 et S_3 fixés respectivement à 200 μV , 300 μV et 500 μV (ce sont les valeurs choisies expérimentalement dans notre test effectué dans la section 3.4.3), on a alors : $m_{val,i}(\{o\}) = 0.5$ et $m_{val,i}(\Omega) = 0.5$.

La solution proposée pour le diagnostic du bobinage d’une machine consiste alors à combiner conjonctivement (1.8) les huit FM obtenues pour les quatres positions et les deux méthodes, puis de prendre la décision maximisant la probabilité pignistique (1.33).

3.4.3 Résultats obtenus

Dans un premier temps, nous avons fait une série de mesures (au nombre de 13 incluant 1 cas sans défaut et 12 cas avec défaut répartis autour de la machine) pour choisir les seuils S_1 , S_2 et S_3 fixés respectivement à 200 μV , 300 μV et 500 μV . Puis nous avons pu faire des tests sur 39 mesures (3 cas sans défaut et 36 cas avec défaut répartis autour de la machine). La fusion basée uniquement sur le sens de variation a obtenu 34 bonnes décisions sur 39, soit un pourcentage d’environ 87.2% de réussite. En prenant en compte la valeur des mesures en plus du sens de variation, la fusion a obtenu 37 bonnes décisions sur 39, soit un pourcentage d’environ 94.9% de réussite.

Même si le nombre de mesures est limité, ces résultats montrent l’intérêt de notre solution employant des fonctions de croyance pour modéliser et fusionner les informations issues des capteurs.

3.4.4 Conclusion et perspectives

La méthode de diagnostic présentée ici possède l'avantage de ne pas nécessiter de connaissances sur l'état sain du système. De plus, elle est non invasive : elle s'effectue sans démontage de la machine, juste avec deux capteurs de champ magnétique externe. La modélisation dans le cadre des fonctions de croyance est très simple et fait que la méthode est facilement compréhensible comme ses résultats.

Le point faible de la méthode actuelle est que la construction des masses est très expérimentale (en particulier au niveau de l'affectation utilisant les seuils S_1 , S_2 et S_3). Une perspective de ce travail consiste à voir comment nous pourrions les construire d'une manière moins *ad hoc*, sachant qu'on ne souhaite pas une méthode nécessitant des connaissances sur l'état sain de la machine, donc sans apprentissage spécifique à une machine, afin de pouvoir l'appliquer directement sur n'importe quelle machine électrique en fonctionnement.

De même, nous pourrions développer des méthodes permettant de déterminer la localisation du défaut en plus de le diagnostiquer.

Enfin, la question du dépôt de brevet est revenue à nous récemment avec la signature en janvier 2014 d'une convention de partenariat entre l'université d'Artois et la SATT Nord (Société d'Accélération du Transfert Technologique) pour faciliter la maturation des projets, leurs protections, et le transfert technologique vers des industriels. Alors qu'au début de nos développements en 2007, ni l'université ni les laboratoires impliqués ne le souhaitaient, nous pourrions en fonction des nouveaux développements (nouvelles constructions des fonctions de masse, nouvelles informations prises en compte, localisation des défauts, etc.) décider d'aller vers un dépôt de brevet.

3.5 Conclusion sur ce chapitre

Dans ce chapitre, trois applications différentes de fusions d'informations utilisant des fonctions de croyance ont été présentées. Ce sont les trois applications principales sur lesquelles j'ai travaillé après ma thèse.

Ma conclusion personnelle et générale sur ces applications, est que plus les modèles sont simples tout en collant à la réalité, le cadre des fonctions de croyance permettant de telles modélisations, plus ils sont interprétables des personnes du métier, et plus il est possible pour ces dernières de maîtriser le système de fusion, de comprendre les résultats fournies et de détecter des anomalies.

Des perspectives pour chacune de ces applications ont été mentionnées. Certaines sont étendues dans mon projet de recherche exposé au prochain chapitre 4.

Projet de recherche

À court et moyen terme, j'ai l'intention de développer et approfondir des travaux autour des deux axes de ma recherche actuelle, à savoir :

1. Développer et améliorer des modèles formels de fusion d'informations, en particulier ceux employant les mécanismes de correction de fonctions de croyance.
2. Développer et améliorer des applications en fusion d'informations avec des fonctions de croyance.

Au niveau du premier axe, mon projet de travail concerne les perspectives détaillées en fin de chapitre 2. Je les complèterais en indiquant de possibles travaux en modélisation des incertitudes dans les chaînes logistiques, le LGI2A ayant pour applications privilégiées la logistique et le transport.

Je souhaite aussi éprouver les modèles formels dont je participerai à l'élaboration en contribuant aux développements d'applications, ce qui permettra également de susciter des idées d'amélioration. J'ai en tête l'idée d'aller plus loin dans les réalisations et les contributions, du fait des développements plus avancés des possibilités de maîtrise des corrections lors de la fusion (axe 1 de ma recherche), et de la meilleure connaissance du domaine pour des applications sur lesquelles j'aurai déjà travaillé.

Je fais maintenant un tour d'horizon des applications sur lesquelles je vais et pourrais travailler, dans un futur proche. Elles se situent côté *transport dit intelligent* dans les applications privilégiées du LGI2A, et dans un cadre plus large de fusion d'informations multi-sources.

La première application est déjà en cours de développement, puisqu'il s'agit de la thèse de Pauline Minary débutée en décembre 2014 et qui porte sur le floutage de visages dans des séquences vidéo issues de caméras de la SNCF. Cette thèse est réalisée au partenariat avec la SNCF (93) avec une convention industrielle de formation par la recherche (bourse cifre). Un premier travail a consisté à implémenter différents classifieurs permettant de reconnaître les visages sur des images. Nous allons voir comment réaliser une fusion de ces classifieurs pour améliorer les performances individuelles des classifieurs. Puis nous allons explorer la piste du suivi de visages pour améliorer la détection de ces derniers.

Je souhaite aussi poursuivre les applications autour du diagnostic non invasif de moteurs en utilisant le champ magnétique externe et de la communication inter-véhicule, qui ont été évoquées au chapitre 3. Je pense qu'il y a suffisamment de contenus à développer pour poursuivre ces travaux aux travers de thèses.

Pour l'application de diagnostic de moteurs, nous sommes actuellement en discussion avec des membres du LSEE pour développer à court terme un travail de

thèse à la fois sur les aspects théoriques de la signification des perturbations dans le champ magnétique externe en fonction des défauts et sur un système de fusion employant le cadre des fonctions de croyance. Il est en effet possible de modéliser par des équations les différents types de flux et tensions circulant au travers d'une machine électrique triphasée. Un court-circuit sur une spire a donc des conséquences théoriques mesurables par ces équations. Nous pourrions alors voir quelle nouvelle source d'information pourrait être exploitable et comment la modéliser et l'utiliser au mieux lors d'un processus de fusion avec ce que nous connaissons déjà.

Pour l'application sur les véhicules communicants, des applications grandeur nature étant en cours de test chez les constructeurs automobiles et les projets de recherche nombreux, nous pouvons jouer un rôle au niveau des systèmes haut niveaux de fusion des messages. Un premier travail est d'implémenter les modèles développés dans la thèse de Mira Fou Farah dans un simulateur plus réaliste, ce qui peut être fait dès l'année prochaine par un stage de master 1 informatique par exemple. Une fois qu'une plate-forme de simulation plus réaliste sera utilisable, nous pourrions essayer de mettre en évidence nos travaux en fusion d'informations et trouver un ou des partenaires académiques ou industriels.

J'ai également en tête l'application de fusion de décisions issues de lecteurs d'adresses postales que j'avais développée et implémentée durant ma thèse (2003-2006) effectuée au sein de l'entreprise Solystic (92) dans le domaine du tri du courrier postal. J'ai gardé des contacts réguliers avec les personnes avec lesquelles j'avais travaillé. À ce jour, ils continuent à accueillir des doctorants en cifre en privilégiant des travaux sur les méthodes de reconnaissance d'adresses postales (reconnaissance de caractères, de mots et de phrases). En fonction des résultats obtenus dans mon premier axe de recherche, en particulier une meilleure prise en compte des caractéristiques des classifieurs lors du processus de fusion, je pense pouvoir les convaincre qu'il est possible d'améliorer la combinaison de lecteurs d'adresses postales que j'avais développée, et qu'il y a là suffisamment de contenu, et de pourcentages de bonnes reconnaissances de courrier à gagner, pour débiter une nouvelle thèse sur ce sujet.

Une autre piste de travail en partenariat avec une entreprise concerne la société KRG Corporate, une startup de moins de 10 salariés dont les locaux sont situés à Montrouge (92), où est employé Gautier Levert qui démarre une thèse avec nous. Cette société maintient et développe une application nommée *seniorAdom* (<https://www.senioradom.com>) en charge de détecter une chute ou un malaise d'une personne âgée ou fragile, sans caméra (afin de garder une certaine intimité) ni bracelet porté par la personne (afin d'être la moins contraignante possible), seulement à l'aide de détecteurs de mouvements et d'ouverture de portes reliés à un boîtier d'assistance. La méthode actuelle analyse les anomalies par rapport au rythme de vie classique de la personne. Il s'agirait pour nous de voir comment contribuer à l'amélioration du système actuel en exploitant une fusion des détecteurs de mouvements et d'ouverture de portes.

Un autre projet qui pourrait constituer une sorte de synthèse de nos travaux de fusion multi-capteurs et de fusion de classifieurs en détection vidéo, concerne le

comptage de personnes à l'aide de capteurs au sol et d'images issues de caméras vidéos, par exemple dans des supermarchés, ou en ville lors de manifestations. Nous avons été sollicité sur ce point il y a 4 ou 5 ans par le centre d'innovation de Lille Métropole, parc de la Haute Borne à Villeneuve d'Ascq, qui cherchait un partenaire académique pour un industriel développant des solutions dans ce domaine. À l'époque nous n'étions pas encore prêts pour un tel travail, mais d'ici un ou deux ans, suivant les résultats obtenus avec la thèse de Pauline Minary, nous pourrions devenir un partenaire dans ce domaine.

Enfin, dans le cadre du projet Nisto et par l'intermédiaire de l'agence Boulogne développement, nous avons eu des contacts avec la RATP. Ils seraient intéressés par un système de communication point à point entre les bus et les panneaux d'annonces aux arrêts de bus. Ils sont actuellement obligés de déployer un réseau de communication coûteux pour faire parvenir les informations aux arrêts, alors qu'un système faisant transiter l'information via les bus qui apporteraient les informations aux arrêts coûterait moins cher. Ce projet comporte beaucoup de travail bas niveau à réaliser avec la construction du réseau point à point, des outils d'affichage, de stockage, de transferts et de calculs (pour ajouter une unité de fusion). Il faudrait dans un premier temps nous rapprocher d'une structure académique ou privée pour voir la faisabilité du projet et ensuite voir les gains possibles de l'ajout d'un système de fusion.

Concernant mon projet de recherche à long terme, je souhaiterais arriver à une meilleure compréhension des mécanismes de correction, à bien déterminer leurs places au regard des notions d'indépendance, de corrélation, de conditionnement. Je me pose aussi la question de l'exploitation des décompositions canoniques. L'évolution de mes connaissances sur ces points et les résultats obtenus dans les différentes applications guideront alors mes travaux dans ce futur éloigné.

Preuve de l'équation (2.12)

Nous donnons ici une preuve que le résultat de l'opération (2.11) est donné par l'équation (2.12), c'est-à-dire qu'on a :

$$\left(m^{\mathcal{X}}[\{F\}]^{\uparrow \mathcal{X} \times \mathcal{R}} \odot_{A \in \mathcal{A}} m^{\mathcal{R}}[A]^{\uparrow \mathcal{X} \times \mathcal{R}}\right)^{\downarrow \mathcal{X}} = m \odot \left(\bigodot_{A \in \mathcal{A}} \overline{A}^{\alpha_A}\right), \quad (\text{A.1})$$

avec $\mathcal{A}$ un ensemble quelconque de contextes de $\mathcal{X}$, β_A le degré de fiabilité dans chaque contexte $A \in \mathcal{A}$ et m la FM fournie par la source.

Ce résultat corrige le théorème 1 donné dans [56].

Notons A_i , $i \in I$, les contextes dans $\mathcal{A}$, et notons simplement β_i le degré de fiabilité β_{A_i} .

Pour tout $A_i \in \mathcal{A}$, le déconditionnement de $m^{\mathcal{R}}[A_i]$ sur $\mathcal{X} \times \mathcal{R}$ est donné par :

$$m^{\mathcal{R}}[A_i]^{\uparrow \mathcal{X} \times \mathcal{R}}(A_i \times \{F\} \cup \overline{A_i} \times \mathcal{R}) = \beta_i, \quad (\text{A.2})$$

$$m^{\mathcal{R}}[A_i]^{\uparrow \mathcal{X} \times \mathcal{R}}(\mathcal{X} \times \mathcal{R}) = \alpha_i. \quad (\text{A.3})$$

De plus, avec $A_i \in \mathcal{A}$, $A_j \in \mathcal{A}$ et $j \neq i$, on a :

$$(A_i \times \{F\} \cup \overline{A_i} \times \mathcal{R}) \cap (A_j \times \{F\} \cup \overline{A_j} \times \mathcal{R}) = (A_i \cup A_j) \times \{F\} \cup \overline{(A_i \cup A_j)} \times \mathcal{R}.$$

Ainsi on peut montrer par une récurrence rapide, que tous les éléments focaux de $\bigodot_{A \in \mathcal{A}} m^{\mathcal{R}}[A]^{\uparrow \mathcal{X} \times \mathcal{R}}$ sont de la forme $C \times \{F\} \cup \overline{C} \times \mathcal{R}$ avec $C = \bigcup_{i \in I' \subseteq I} A_i$, avec pour masse $\prod_{i \in I'} \beta_i \prod_{j \in I \setminus I'} \alpha_j$.

Autrement dit, pour tout $C = \bigcup_{i \in I' \subseteq I} A_i$ (c'est-à-dire, pour tout C composé d'une union d'éléments dans $\mathcal{A}$, I' étant l'ensemble des indices des A_i dans C) :

$$\bigodot_{A \in \mathcal{A}} m^{\mathcal{R}}[A]^{\uparrow \mathcal{X} \times \mathcal{R}}(C \times \{F\} \cup \overline{C} \times \mathcal{R}) = \prod_{i \in I'} \beta_i \prod_{j \in I \setminus I'} \alpha_j. \quad (\text{A.4})$$

D'un autre côté, pour tout $B \subseteq \mathcal{X}$, on a :

$$m^{\mathcal{X}}[\{F\}]^{\uparrow \mathcal{X} \times \mathcal{R}}(B \times \{F\} \cup \mathcal{X} \times \{NF\}) = m(B). \quad (\text{A.5})$$

Or, pour tout $B \subseteq \mathcal{X}$ et pour tout $C = \bigcup_{i \in I' \subseteq I} A_i$:

$$(C \times \{F\} \cup \overline{C} \times \mathcal{R}) \cap (B \times \{F\} \cup \mathcal{X} \times \{NF\}) = B \times \{F\} \cup \overline{C} \times \{NF\}. \quad (\text{A.6})$$

Après projection sur $\mathcal{X}$, l'opération $(m^{\mathcal{X}}[\{F\}]^{\uparrow \mathcal{X} \times \mathcal{R}} \odot_{A \in \mathcal{A}} m^{\mathcal{R}}[A]^{\uparrow \mathcal{X} \times \mathcal{R}})^{\downarrow \mathcal{X}}$ consiste

donc à transférer une portion $\prod_{i \in I'} \beta_i \prod_{j \in I \setminus I'} \alpha_j$ de chaque masse $m(B)$, $B \subseteq \mathcal{X}$, de B à $B \cup \overline{C}$, pour tout $C = \cup_{i \in I' \subseteq I} A_i$.

D'un autre côté, l'opération $m(\bigcup) (\bigodot_{A \in \mathcal{A}} \overline{A}^{\alpha_A})$ peut être réécrite de la manière suivante :

$$m(\bigcup) (\bigodot_{i \in I} \overline{A_i}^{\alpha_i}) = m(\bigcup) \left(\bigodot_{i \in I} \left\{ \begin{array}{ll} \mathcal{X} & \mapsto \alpha_i \\ \overline{A_i} & \mapsto \beta_i \end{array} \right. \right) \quad (\text{A.7})$$

Or pour tout $(i, j) \in I^2$, $i \neq j$, on a $\overline{A_i} \cap \overline{A_j} = \overline{A_i \cup A_j}$. Les éléments focaux de $\bigodot_{i \in I} \overline{A_i}^{\alpha_i}$ sont donc tous les éléments $\overline{C}$ avec $C = \cup_{i \in I' \subseteq I} A_i$ avec pour masse $\prod_{i \in I'} \beta_i \prod_{j \in I \setminus I'} \alpha_j$.

Par conséquent, l'opération $m(\bigcup) (\bigodot_{i \in I} \overline{A_i}^{\alpha_i})$ consiste aussi à transférer une portion $\prod_{i \in I'} \beta_i \prod_{j \in I \setminus I'} \alpha_j$ de chaque masse $m(B)$, $B \subseteq \mathcal{X}$, de B à $B \cup \overline{C}$, pour tout $C = \cup_{i \in I' \subseteq I} A_i$. Ce qui conclut à l'égalité (A.1).

□

Algorithme GNN (Global Nearest Neighbour)

Nous présentons ici l'algorithme d'association de données GNN (Global Nearest Neighbour) [4] utilisé dans le domaine du suivi multi-objets en particulier avec un filtrage par IMM.

Étant donnés N observations $z_i, i \in I = \{1, \dots, N\}$, et M prédictions $\bar{z}_j, j \in J = \{1, \dots, M\}$, l'algorithme GNN consiste à résoudre le problème d'optimisation linéaire en nombres entiers (donc résoluble en temps polynomiale par l'algorithme hongrois par exemple) suivant :

$$\min \sum_{i=1}^N \sum_{j=1}^M d_{ij} r_{ij}, \quad (\text{B.1})$$

sous les contraintes :

$$\sum_i^N r_{ij} = 1, \quad (\text{B.2})$$

$$\sum_j^{M+N} r_{ij} \leq 1, \quad (\text{B.3})$$

$$r_{ij} \in \{0, 1\}, \forall j \in \{1, \dots, N + M\}, \forall i \in \{1, \dots, N\}, \quad (\text{B.4})$$

où :

- r_{ij} est une variable binaire valant 1 si l'observation z_i est associée à la prédiction $\bar{z}_j$, 0 sinon ;
- N prédictions $\bar{z}_j, j \in \{M + 1, \dots, M + N\}$ ont été ajoutées (B.3) artificiellement aux M prédictions $\bar{z}_j, j \in \{1, \dots, M\}$, afin de gérer les apparitions et disparitions d'objets. Pour cela, les distances d_{ij} avec $j > M$ sont définies telles que $d_{ij} = \lambda$ si $j = M + i$, ∞ sinon. Un exemple est donné dans le tableau B.1 pour $N = 4$ et $M = 3$. La constante λ correspond à la distance maximale pour associer une observation à une prédiction. Une observation z_i associée au final à une prédiction $\bar{z}_j$ telle que $j > M$ constitue un nouvel objet (une apparition).

Afin de comparer dans nos tests (cf figure 3.10) le GNN avec les méthodes crédales utilisant les FM m_{ij} (3.17) dépendantes du paramètre γ , nous avons fait le lien suivant entre le paramètre γ et le paramètre λ . Ce dernier étant la distance

Tableau B.1 – Matrice des distances d'association pour le GNN avec 4 observations z_i , $i \in \{1, 2, 3, 4\}$, et 3 prédictions $\bar{z}_j$, $j \in \{1, 2, 3\}$.

	z_1	z_2	z_3	z_4
$\bar{z}_1$	d_{11}	d_{12}	d_{13}	d_{14}
$\bar{z}_2$	d_{21}	d_{22}	d_{23}	d_{24}
$\bar{z}_3$	d_{31}	d_{32}	d_{33}	d_{34}
$\bar{z}_4$	λ	∞	∞	∞
$\bar{z}_5$	∞	λ	∞	∞
$\bar{z}_6$	∞	∞	λ	∞
$\bar{z}_7$	∞	∞	∞	λ

à partir de laquelle une observation n'est pas associée à une prédiction connue, nous pouvons l'interpréter comme le moment où la valeur de la masse sur la non association de l'observation à la prédiction $m^{\Omega_{i,j}}(\{n_{i,j}\})$ dépasse la valeur de la masse en faveur de l'association de l'observation à la prédiction $m^{\Omega_{i,j}}(\{o_{i,j}\})$, c'est-à-dire en fixant $\exp^{-\gamma\lambda} = 1 - \exp^{-\gamma\lambda}$ soit $\lambda = \frac{-\log(.5)}{\gamma} = \frac{\log(2)}{\gamma}$.

Sélection de publications

J'ai choisi de mettre en annexe les trois articles de revues internationales suivants :

1. D. Mercier, É. Lefèvre, D. Jolly, **Object association with belief functions, an application with vehicles**, *Information Sciences*, Vol. 181, Issue 24, pp. 5485-5500, December 2011.
2. D. Mercier, É. Lefèvre, F. Delmotte, **Belief functions contextual discounting and canonical decompositions**, *International Journal of Approximate Reasoning*, Vol. 53, Issue 2, pp. 146-158, February 2012.
3. M. Bou Farah, D. Mercier, É. Lefèvre, F. Delmotte, **A high-level application using belief functions for exchanging and managing uncertain events on the road in vehicular ad-hoc networks**, *Annals of telecommunications : special issue on belief functions and uncertainty management in networks and telecommunication*, Vol. 69, Issue 3-4, pp. 185-199, April 2014.

et les deux articles de conférences internationales suivants :

1. F. Pichon, D. Mercier, F. Delmotte, É. Lefèvre, **Truthfulness in contextual information correction**, *3rd International Conference on Belief Functions, BELIEF 2014*, F. Cuzzolin (Ed.), pp. 11-20, Oxford, United Kingdom, September 26-28, 2014.
2. D. Mercier, F. Pichon, É. Lefèvre, F. Delmotte, **Learning contextual discounting and contextual reinforcement from labelled data**, *13th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU 2015, LNAI 9161*, S. Destercke and T. Denœux (Eds.), pp. 472-481, Compiègne, France, July 15-17, 2015.

Le premier article de revue résume nos travaux sur l'association de données avec une applications aux véhicules intelligents. Le second synthétise nos premiers travaux sur l'extension de l'affaiblissement contextuel, sous réserve de corriger l'équation (44) (résultat du théorème 1) par l'équation (2.12) de ce rapport (avec la preuve donnée en annexe A). Il a de ce fait surtout le mérite d'introduire le renforcement contextuel et de souligner les liens entre ces mécanismes et les décompositions canoniques. Le troisième est issu de la thèse de Mira Bou Farah, il montre l'intérêt de l'utilisation des fonctions de croyance en communication inter-véhicule décentralisée.

Les travaux des deux articles de conférences constituent les derniers publiés sur les mécanismes de correction contextuelle. Le premier avance de nouveaux mécanismes et de nouvelles perspectives, et le deuxième démontre un intérêt de ces

mécanismes avec des premières améliorations de performances de classifieurs. Ces deux articles sont aussi à l'origine d'une soumission d'un article de revue en cours de relecture.

Object Association with Belief Functions, an application with vehicles.

David Mercier, Éric Lefèvre, Daniel Jolly

Univ. Lille Nord de France, UArtois, EA 3926 LGI2A, France

Abstract

The problem tackled in this article consists in associating perceived objects detected a certain time with known objects previously detected, knowing uncertain and imprecise information regarding the association of each perceived objects with each known objects. For instance, this problem can occur during the association step of an association obstacle tracking process, especially in the context of vehicle driving aid. A contribution in the modelling of this association problem in the belief function framework is introduced. By interpreting belief functions as weighted opinions according to the Transferable Belief Model semantics, pieces of information regarding the association of known objects and perceived objects can be expressed in a common global space of association to be combined by the conjunctive rule of combination, and a decision making process using the pignistic transformation can be made. This approach is validated on real data.

Keywords: Object association, obstacle tracking, belief functions, Transferable Belief Model.

1. Introduction

In obstacle tracking, the association step consists in establishing a correlation between tracks (known objects) and targets (perceived objects) from information usually provided by different sensors or captors. Such a mapping can be even more complex depending on the number of targets and tracks, as well as the quality of the provided information. Introduced by Dempster [7] and Shafer [26], belief functions constitute a suitable framework for the representation and manipulation of imperfect information [4, 12, 13, 20, 30, 31]. Thus, next to architectures based on Bayesian probabilistic framework [2, 3], Rombaut [23, 24] develops a first modelling based on belief functions. In this

model, information regarding the association of couples (known objects, perceived objects) is represented by belief functions, which are combined using, for simplicity reasons, an adapted combination introduced by Rombaut. In [16] this latter model is developed by using a decision-making system based on belief matrices and the application of a coupling algorithm.

In this article, a modelling of this association step problem is introduced in the Smets' semantic approach of belief functions: the Transferable Belief Model (TBM) [29], a subjectivist and non-probabilistic interpretation of the Dempster-Shafer theory of belief function. In particular, it is shown that TBM classical tool like the conjunctive combination rule and the pignistic decision-making can be implemented and tested in a real time application, these experimental results demonstrating the effectiveness of this approach as compared to Rombaut's combination rule.

The works presented here extend a short version of this study presented in [19], as well as it reexpress and extend in the Transferable Belief Model a former model presented by some of the authors in [17]. At last, the association problem described here also presents some similarities with the works undertaken by Ristic and Smets in [22], which are discussed in this article.

This article is organized as follows. The TBM basic concepts we need are recalled in Section 2. An association algorithm based on belief functions is then introduced in Section 3 and discussed with previous approaches in Section 4. Then, experimental results on real data are presented in Section 5. Finally, Section 6 concludes this article.

2. Transferable Belief Model (TBM): basic concepts

The Transferable Belief Model (TBM) is a model of uncertain reasoning and decision-making based on two levels [14, 29]:

- the credal level, where available pieces of information are represented by belief functions, and manipulated;
- the pignistic or decision level, where belief functions are transformed into probability measures when a decision has to be made, and the expected utility is maximized.

2.1. Representing information with belief functions

2.1.1. Belief functions

The knowledge held by an agent is represented by the allocation of a finite mass of belief to subsets of the universe of discourse.

Let $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$, called the frame of discernment, be a finite set composed of all possible answers to a given question Q of interest. The beliefs held by a rational agent Ag regarding the answer to question Q can be quantified by a *mass function* or a *basic belief assignment (BBA)* $m_{Ag}^\Omega : 2^\Omega \rightarrow [0, 1]$ s.t.:

$$\sum_{A \subseteq \Omega} m_{Ag}^\Omega(A) = 1 . \quad (1)$$

The quantity $m_{Ag}^\Omega(A)$ represents the part of the unit mass allocated to the hypothesis that the answer to question Q is in the subset A of Ω . When there is no ambiguity, the notation m_{Ag}^Ω will be simplified as follows m^Ω or m .

- A subset A of Ω such that $m(A) > 0$ is called a *focal set* of m .
- A BBA m with only one focal set A is called a *categorical BBA* and is denoted m_A ; then $m_A(A) = 1$.
- Total ignorance is represented by the BBA m_Ω called the *vacuous BBA*.
- A *normal BBA* m satisfies the condition $m(\emptyset) = 0$.
- Let A be a subset of Ω , the cardinality of A , denoted $|A|$, is the number of elements of Ω in A ; if $|A| = 1$, A is said to be a *singleton*.

The belief and plausibility functions associated with a BBA m are defined, respectively, as:

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad \forall A \subseteq \Omega , \quad (2)$$

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \quad \forall A \subseteq \Omega . \quad (3)$$

Functions m , bel and pl are in one-to-one correspondence, and thus constitute different forms of the same information.

2.1.2. Refinements and Coarsenings

When applying the TBM to a real-world application, the determination of the frame of discernment Ω , which defines the set of states on which beliefs will be expressed, is a crucial step. As noticed by Shafer [26, chapter 6], the degree of granularity of Ω is always, to some extent, a matter of

convention, as any element of Ω representing a given state can always be split into several alternatives. Hence, it is fundamental to examine how a belief function defined on a frame may be expressed in a finer or, conversely, in a coarser frame. The concepts of refinement and coarsening can be defined as follows.

Let Θ and Ω denote two frames of discernment. A mapping $\rho : 2^\Theta \rightarrow 2^\Omega$ is called a *refining* of Θ (Figure 5) if it verifies the following properties:

1. The set $\{\rho(\{\theta\}), \theta \in \Theta\} \subseteq 2^\Omega$ is a partition of Ω , and
2. For all $A \subseteq \Theta$:

$$\rho(A) = \bigcup_{\theta \in A} \rho(\{\theta\}). \quad (4)$$

Θ is then called a coarsening of Ω , and Ω is called a refinement of Θ .

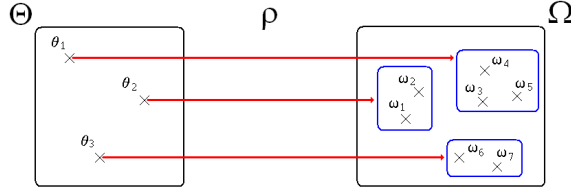


Figure 1: Illustration of a coarsening Θ of Ω associated with a refining ρ of Θ .

2.2. Manipulating information with belief functions

2.2.1. Vacuous extension

The *vacuous extension* operation allows one to convey a mass function m^Θ , expressing a state of belief on Θ , to a finer frame Ω , a refinement of Θ . Stemming from the *least committed principle* [27], this operation is denoted with an arrow pointing up, and is defined by:

$$m^{\Theta \uparrow \Omega}(\rho(A)) = m^\Theta(A), \quad \forall A \subseteq \Theta, \quad (5)$$

where ρ is the refining of Θ in Ω .

2.2.2. Combining beliefs

Two BBAs m_1 and m_2 , induced by distinct and reliable sources of information, can be combined using the *conjunctive rule of combination* (CRC), also called *unnormalized Dempster's rule of combination*, defined for all $A \subseteq \Omega$ by:

$$m_1 \odot m_2(A) = \sum_{B \cap C = A} m_1(B) m_2(C). \quad (6)$$

The normalization hypothesis ($m(\emptyset) = 0$) can be recovered with the following normalization step:

$$m_1 \oplus m_2(A) = \begin{cases} \frac{m_1 \odot_2(A)}{1 - m_1 \odot_2(\emptyset)} & \text{if } \emptyset \neq A \subseteq \Omega, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

This latter rule of combination is called *Dempster's rule of combination*.

2.3. Decision-making level

When a decision has to be made regarding the answer to question Q , some *rational principles* [6] justify the strategy consisting in choosing the decision d among a set of possible decisions $\mathcal{D}$, which minimizes the *expected risk* defined by:

$$R(d) = \sum_{\omega \in \Omega} c(d, \omega) P^\Omega(\{\omega\}), \quad (8)$$

where $P^\Omega : 2^\Omega \rightarrow [0, 1]$ is a probability measure and $c : \mathcal{D} \times \Omega \rightarrow \mathbb{R}$ a cost function, $c(d, \omega)$ representing the cost to decide d while the truth is ω .

At this level, the mass function m^Ω representing the available information regarding the answer to question Q belonging to Ω (resulting in practice from a fusion process) has then to be transformed in a probability measure. A solution [9] consists in computing the *pignistic probability* [28] defined by:

$$BetP^\Omega(\{\omega\}) = \sum_{\{A \subseteq \Omega, \omega \in A\}} \frac{m(A)}{|A| (1 - m(\emptyset))}, \quad \forall \omega \in \Omega. \quad (9)$$

The chosen decision is then the one that minimizes the pignistic risk defined by:

$$R_{Bet}(d) = \sum_{\omega \in \Omega} c(d, \omega) BetP^\Omega(\{\omega\}). \quad (10)$$

In the case of 0-1 costs with $\mathcal{D} = \Omega$, which means that $c(\omega_i, \omega_j) = 1$ if $i = j$, 0 otherwise, choosing the decision d which minimises the pignistic risk (10) is equivalent to choose the decision d which maximizes the pignistic probability (9).

An other case consists in choosing 0-1 costs with $\mathcal{D} = \Omega \cup \{d_0\}$, where d_0 , called *rejection decision* [9], consists in refusing to make a decision belonging to $\mathcal{D} \setminus \{d_0\}$ when the risk is judged too high. By denoting $c_0 = c(d_0, \omega_i) \forall i \in \{1, \dots, N\}$, minimizing the pignistic risk (10) is equivalent to choose the decision:

- d_0 if $\max_{i=1, \dots, N} BetP(\{\omega_i\}) < 1 - c_0$,

- ω_j if $BetP(\{\omega_j\}) = \max_{i=1,\dots,N} BetP(\{\omega_i\}) \geq 1 - c_0$.

The cost c_0 is called the *rejection cost*.

3. Object association algorithm

3.1. Representing information with belief functions

The first step when building belief functions is to define the universe of discourse.

Let us consider the following notations:

- X_i : represents a perceived object at time t , $i \in I = \{1, \dots, N\}$, N being the number of perceived objects at time t ;
- Y_j : represents a known object at previous time $t - 1$, $j \in J = \{1, \dots, M\}$, M being the number of known objects at time $t - 1$;
- $*$: is a proposition meaning “no object”.

The association process objective consists in finding the best possible association between a set of perceived objects $\{X_1, X_2, \dots, X_N, *\}$ and a set of known objects $\{Y_1, Y_2, \dots, Y_M, *\}$, under the following constraints:

- each perceived object X_i is associated with at most one known object;
- each known object Y_j is associated with at most one perceived object;
- proposition $*$ can be associated with any objects.

The frames of discernment involved in this application are then the followings:

- $\Omega_{i,j} = \{y_{i,j}, n_{i,j}\}$: contains the two possible answers (yes or no) to the question $Q_{i,j}$: “Is the perceived object X_i associated with the known object Y_j ?”;
- $\Omega_{X_i} = \{Y_1, Y_2, \dots, Y_M, *\}$: contains the set of possible answers to the question Q_{X_i} : “Who is associated with the perceived object X_i ?”, proposition $*$ meaning that X_i has appeared;
- $\Omega_{Y_j} = \{X_1, X_2, \dots, X_N, *\}$: contains the set of possible answers to the question Q_{Y_j} : “Who is associated with the known object Y_j ?”, proposition $*$ meaning that Y_j has disappeared or is hidden.

Let us remark that $\Omega_{Y_j} = \Omega_{Y_k}$, for all $j, k \in J$, and $\Omega_{X_i} = \Omega_{X_\ell}$, for all $i, \ell \in I$. Thus, Ω_{X_i} (respectively Ω_{Y_j}) can be denoted $\Omega_X \forall i$ (respectively $\Omega_Y \forall j$). At last, when there is no ambiguity, the frames elements will be simplified as follows:

- $\Omega_{X_i} = J \cup \{\star\} = \{1, \dots, M, \star\}$,
- $\Omega_{Y_j} = I \cup \{\star\} = \{1, \dots, N, \star\}$.

In the domain of intelligent vehicles, sensors or measures generally provide information regarding the association between each perceived object X_i and each known object Y_j [23, 24, 16, 15]. More precisely, initial information is represented by mass functions $m^{\Omega_{i,j}}$ on frames $\Omega_{i,j}$, $i \in I$, $j \in J$:

- the mass allocated to $\{y_{i,j}\}$ expresses information on the fact that X_i is associated with Y_j ;
- the mass allocated to $\{n_{i,j}\}$ expresses information on the fact that X_i is not associated with Y_j ;
- the mass allocated to $\Omega_{i,j} = \{y_{i,j}, n_{i,j}\}$ expresses the ignorance regarding the association of X_i and Y_j .

The association problem to solve can then be express in the following manner: $N \times M$ mass functions $m^{\Omega_{i,j}}$ being defined regarding the association of each perceived object X_i with each known object Y_j , how these pieces of information can be fused/merged to determine:

- Where do perceived objects X_i come from?
- What are known objects Y_j become?

3.2. Expressing pieces of information in a common frame

To answer these questions, the $N \times M$ mass functions can be combined when expressed on two possible common frames: Ω_X and Ω_Y .

Frames Ω_{X_i} and Ω_{Y_j} being refinements of $\Omega_{i,j}$, each information $m^{\Omega_{i,j}}$ can be expressed either on Ω_{X_i} or on Ω_{Y_j} by a vacuous extension (5):

$$m^{\Omega_{i,j} \uparrow \Omega_{X_i}}(\rho_{i,j}(A)) = m^{\Omega_{i,j}}(A), \quad \forall A \subseteq \Omega_{i,j}, \quad (11)$$

where $\rho_{i,j}$ is the refining of $\Omega_{i,j}$ on Ω_{X_i} illustrated in Figure 2, and defined by $\rho_{i,j}(\{y_{i,j}\}) = \{j\}$ and $\rho_{i,j}(\{n_{i,j}\}) = \overline{\{j\}}$.

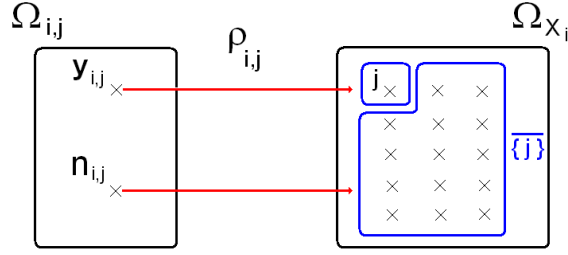


Figure 2: Refining $\rho_{i,j}$ allowing one to transport the information $m^{\Omega_{i,j}}$ on Ω_{X_i} .

Thus, for all $(i, j) \in I \times J$:

$$\begin{cases} m^{\Omega_{i,j} \uparrow \Omega_{X_i}}(\{j\}) &= m^{\Omega_{i,j}}(\{y_{i,j}\}) \\ m^{\Omega_{i,j} \uparrow \Omega_{X_i}}(\overline{\{j\}}) &= m^{\Omega_{i,j}}(\{n_{i,j}\}) \\ m^{\Omega_{i,j} \uparrow \Omega_{X_i}}(\Omega_{X_i}) &= m^{\Omega_{i,j}}(\Omega_{i,j}) \end{cases} \quad (12)$$

In the same manner, mass functions $m^{\Omega_{i,j}}$ can be vacuously extended on Ω_{Y_j} for all $(i, j) \in I \times J$:

$$\begin{cases} m^{\Omega_{i,j} \uparrow \Omega_{Y_j}}(\{i\}) &= m^{\Omega_{i,j}}(\{y_{i,j}\}) \\ m^{\Omega_{i,j} \uparrow \Omega_{Y_j}}(\overline{\{i\}}) &= m^{\Omega_{i,j}}(\{n_{i,j}\}) \\ m^{\Omega_{i,j} \uparrow \Omega_{Y_j}}(\Omega_{Y_j}) &= m^{\Omega_{i,j}}(\Omega_{i,j}) \end{cases} \quad (13)$$

In the following, mass functions $m^{\Omega_{i,j} \uparrow \Omega_{X_i}}$ (respectively $m^{\Omega_{i,j} \uparrow \Omega_{Y_j}}$) are denoted $m_j^{\Omega_{X_i}}$ (respectively $m_i^{\Omega_{Y_j}}$).

3.3. Combining mass functions

At this level:

- for each $i \in I = \{1, \dots, N\}$, M mass functions $m_j^{\Omega_{X_i}}$ have been created regarding the association of each object X_i toward the Y_j , the focal elements of each one being $\{j\}$, $\overline{\{j\}}$, and Ω_{X_i} .
- for each $j \in J = \{1, \dots, M\}$, N mass functions $m_i^{\Omega_{Y_j}}$ have been created regarding the association of each object Y_j toward the X_i , the focal elements of each one being $\{i\}$, $\overline{\{i\}}$, et Ω_{Y_j} .

The M mass functions $m_j^{\Omega_{X_i}}$, considered as distinct and reliable, are combined using the conjunctive rule of combination (6).

Let us denote $m^{\Omega_{X_i}}$ the resulting mass function:

$$m^{\Omega_{X_i}} = \odot_{j \in J} m_j^{\Omega_{X_i}} . \quad (14)$$

For all $k \in J$:

$$m^{\Omega_{X_i}}(\{k\}) = \sum_{\cap A_j = \{k\}} \prod_{j \in J} m_j^{\Omega_{X_i}}(A_j) , \quad (15)$$

where, for all $j \in J$, $A_j = \{j\}$, $\overline{\{j\}}$, or Ω_{X_i} .

But:

$$\begin{aligned} \cap_{j \in J} A_j = \{k\} &\Leftrightarrow A_k = \{k\} \text{ and } (A_j = \overline{\{j\}} \text{ or } \\ &\quad A_j = \Omega_{X_i}, \forall j \in J \setminus \{k\}), \\ &\Leftrightarrow A_k = \{k\} \text{ and } \\ &\quad A_j \neq \{j\}, \forall j \in J \setminus \{k\} . \end{aligned}$$

Thus, for all $k \in J$:

$$m^{\Omega_{X_i}}(\{k\}) = m_k^{\Omega_{X_i}}(\{k\}) \prod_{\substack{j=1 \\ j \neq k}}^M (1 - m_j^{\Omega_{X_i}}(\{j\})) . \quad (16)$$

Similarly, for all $K \subseteq J$:

$$\begin{aligned} m^{\Omega_{X_i}}(\overline{K}) &= \sum_{\cap A_j = \overline{K}} \prod_{j \in J} m_j^{\Omega_{X_i}}(A_j) , \\ &= \prod_{j \in K} m_j^{\Omega_{X_i}}(\overline{\{j\}}) \prod_{j \in \overline{K}} m_j^{\Omega_{X_i}}(\Omega_{X_i}) . \end{aligned}$$

In particular:

$$\begin{aligned} m^{\Omega_{X_i}}(\{\star\}) &= m^{\Omega_{X_i}}(\overline{J}) = \prod_{j \in J} m_j^{\Omega_{X_i}}(\overline{\{j\}}) , \\ m^{\Omega_{X_i}}(\Omega_{X_i}) &= m^{\Omega_{X_i}}(\overline{\emptyset}) = \prod_{j \in J} m_j^{\Omega_{X_i}}(\Omega_{X_i}) . \end{aligned}$$

At last:

$$m^{\Omega_{X_i}}(\emptyset) = \sum_{\cap A_j = \emptyset} \prod_{j \in J} m_j^{\Omega_{X_i}}(A_j) , \quad (17)$$

$$= \sum_{\substack{j, k \in J \\ j \neq k}} m_j^{\Omega_{X_i}}(\{j\}) m_k^{\Omega_{X_i}}(\{k\}) . \quad (18)$$

In the same manner, the N mass functions $m_i^{\Omega_{Y_j}}$ can also be conjunctively combined to result in a mass function $m^{\Omega_{Y_j}}$.

Example 1. *Let us consider one perceived object X_1 and two known objects Y_1 and Y_2 s.t.:*

$$\begin{cases} m^{\Omega_{Y_1}}(\{y_{1,1}\}) = .2 \\ m^{\Omega_{Y_1}}(\{n_{1,1}\}) = .45 \\ m^{\Omega_{Y_1}}(\Omega_{Y_1}) = .35 \end{cases} \quad \begin{cases} m^{\Omega_{Y_2}}(\{y_{1,2}\}) = .45 \\ m^{\Omega_{Y_2}}(\{n_{1,2}\}) = .15 \\ m^{\Omega_{Y_2}}(\Omega_{Y_2}) = .4 \end{cases} \quad (19)$$

By expressing this information on Ω_{X_1} (X_1 's point of view: with which known object, the perceived object X_1 is associated? In other words: Where does X_1 come from?), it is obtained:

$$\begin{cases} m_1^{\Omega_{X_1}}(\{1\}) = .2 \\ m_1^{\Omega_{X_1}}(\overline{\{1\}}) = .45 \\ m_1^{\Omega_{X_1}}(\Omega_{X_1}) = .35 \end{cases} \quad \begin{cases} m_2^{\Omega_{X_1}}(\{2\}) = .45 \\ m_2^{\Omega_{X_1}}(\overline{\{2\}}) = .15 \\ m_2^{\Omega_{X_1}}(\Omega_{X_1}) = .4 \end{cases} \quad (20)$$

The conjunctive combination of $m_1^{\Omega_{X_1}}$ and $m_2^{\Omega_{X_1}}$ provides the following result:

$$\begin{aligned} m^{\Omega_{X_1}}(\{1\}) &= .2 \times (1 - .45) = .2 \times .55 = .11 \\ m^{\Omega_{X_1}}(\{2\}) &= .45 \times (1 - .2) = .45 \times .8 = .36 \\ m^{\Omega_{X_1}}(\overline{\{1\}}) &= m^{\Omega_{X_1}}(\{2, \star\}) = .45 \times .4 = .18 \\ m^{\Omega_{X_1}}(\overline{\{2\}}) &= m^{\Omega_{X_1}}(\{1, \star\}) = .15 \times .35 = .05 \\ m^{\Omega_{X_1}}(\overline{\{1, 2\}}) &= m^{\Omega_{X_1}}(\{\star\}) = .45 \times .15 = .07 \\ m^{\Omega_{X_1}}(\Omega_{X_1}) &= m^{\Omega_{X_1}}(\{1, 2, \star\}) = .35 \times .4 = .14 \\ m^{\Omega_{X_1}}(\emptyset) &= .2 \times .45 = .09 \end{aligned} \quad (21)$$

3.4. Decision-making

The pignistic probability $BetP^{\Omega_{X_i}}$ (9) computed from $m^{\Omega_{X_i}}$ is defined for all $\omega \in \Omega_{X_i}$ by:

$$BetP^{\Omega_{X_i}}(\{\omega\}) = \sum_{\{A \subseteq \Omega_{X_i}, \omega \in A\}} \frac{m^{\Omega_{X_i}}(A)}{|A| (1 - m^{\Omega_{X_i}}(\emptyset))}. \quad (22)$$

Then, for all $k \in J$:

$$BetP^{\Omega_{X_i}}(\{k\}) = \mathcal{K}_1 \left[m_k^{\Omega_{X_i}}(\{k\}) \prod_{\substack{j=1 \\ j \neq k}}^M (1 - m_j^{\Omega_{X_i}}(\{j\})) \right. \\ \left. + \sum_{\substack{k \in \bar{K} \\ K \subseteq J}} \frac{1}{|\bar{K}|} \prod_{j \in K} m_j^{\Omega_{X_i}}(\overline{\{j\}}) \prod_{j \in \bar{K}} m_j^{\Omega_{X_i}}(\Omega_{X_i}) \right] , \quad (23)$$

where

$$\mathcal{K}_1 = \frac{1}{1 - m^{\Omega_{X_i}}(\emptyset)} = \frac{1}{1 - \sum_{\substack{j, k \in J \\ j \neq k}} m_j^{\Omega_{X_i}}(\{j\}) m_k^{\Omega_{X_i}}(\{k\})} . \quad (24)$$

And:

$$BetP^{\Omega_{X_i}}(\{\star\}) = \mathcal{K}_1 \sum_{K \subseteq J} \frac{1}{|\bar{K}|} \prod_{j \in K} m_j^{\Omega_{X_i}}(\overline{\{j\}}) \prod_{j \in \bar{K}} m_j^{\Omega_{X_i}}(\Omega_{X_i}) . \quad (25)$$

Once the pignistic probabilities $BetP^{\Omega_{X_i}}$ computed for each $i \in I$, the chosen decision is the one that maximizes the pignistic probability associated to the joint law $BetP^{\Omega_{X_1} \times \dots \times \Omega_{X_N}}$ which verifies the constraints expressed in Section 3.1.

This decision-making process is summed up by the following algorithm 1.

Similarly, an equivalently justified solution consists in computing the decision from the Y_j points of view, by maximizing the pignistic probability $BetP^{\Omega_{Y_1} \times \dots \times \Omega_{Y_M}}$.

Example 2 (Example 1 continued). *Let us consider again one perceived object X_1 and two known objects Y_1 and Y_2 with the information represented by the BBAs $m^{\Omega_{1,1}}$ and $m^{\Omega_{1,2}}$ defined by Equation 19.*

From X_1 's point of view, the conjunctive combination of $m_1^{\Omega_{X_1}}$ and $m_2^{\Omega_{X_1}}$ has been detailed in Example 1. The pignistic probability $BetP^{\Omega_{X_1}}$ regarding the association of X_1 is then given by:

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{\star\}$	$\{1, \star\}$	$\{2, \star\}$	$\{1, 2, \star\}$
$m^{\Omega_{X_1}}(A)$	.09	.11	.36	.07	.05	.18	.14
$BetP^{\Omega_{X_1}}(A)$		.20	.55	.25	.45	.80	1

Conclusion from X_1 's point of view:

Algorithm 1: Decision-making process regarding the perceived objects X_i

Data: $BetP^{\Omega_{X_i}}, i \in I = \{1, \dots, N\}$, pignistic probabilities of each perceived object X_i .

Result: The decision maximizing $BetP^{\Omega_{X_1} \times \dots \times \Omega_{X_N}}$ and verifying the association constraints.

begin

- Consider the $(M + 1)^N$ possible associations
- Delete from this set the associations which do not verify the association constraints expressed in Section 3.1
- Compute the products of the pignistic probabilities over these sets
- Select the association maximizing the products

end

1. The singleton maximizing $BetP^{\Omega_{X_1}}$ is $\{2\}$, so X_1 is associated with Y_2 ;
2. knowing that Y_1 is not associated, Y_1 has disappeared (or is hidden).

On the other hand, it is also possible to express the available information on Ω_{Y_1} and Ω_{Y_2} :

$$\begin{cases} m_1^{\Omega_{Y_1}}(\{1\}) = .2 \\ m_1^{\Omega_{Y_1}}(\overline{\{1\}}) = .45 \\ m_1^{\Omega_{Y_1}}(\Omega_{Y_1}) = .35 \end{cases} \quad \begin{cases} m_1^{\Omega_{Y_2}}(\{1\}) = .45 \\ m_1^{\Omega_{Y_2}}(\overline{\{1\}}) = .15 \\ m_1^{\Omega_{Y_2}}(\Omega_{Y_2}) = .4 \end{cases}$$

As there is only one perceived object X_1 , no combination is necessary:

A	$\emptyset$	$\{1\}$	$\{\star\}$	$\{1, \star\}$
$m^{\Omega_{Y_1}}(A)$		.2	.45	.35
$BetP^{\Omega_{Y_1}}(A)$		.375	.625	1
$m^{\Omega_{Y_2}}(A)$		.45	.15	.4
$BetP^{\Omega_{Y_2}}(A)$		.65	.35	1

From the association constraints (Section 3.1), the known objects (Y_1, Y_2) can be associated to $(1, \star)$, $(\star, 1)$, or $(\star, \star)$.

And, as:

- $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}(\{1, \star\}) = .375 \times .35 = .131;$
- $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}(\{\star, 1\}) = .625 \times .65 = .406;$
- $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}(\{\star, \star\}) = .625 \times .35 = .219,$

then $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}$ reaches its valid maximum at $\{\star, 1\}$, so (Y_1, Y_2) is associated with $(\star, 1)$; in other words, Y_1 has disappeared and Y_2 is associated with X_1 .

In the previous example, the decision coming from X_1 and the decision coming from the Y_j are the same.

Unfortunately, as illustrated by the following Example 3, the decision providing by the criteria of maximizing the joint pignistic probability can be different depending on which point of view (perceived objects X_i or known objects Y_j) it is computed.

Example 3. Let us considered one perceived object X_1 , and two known objects Y_1 and Y_2 , s.t.:

$$\begin{cases} m^{\Omega_{1,1}}(\{y_{1,1}\}) = .5 \\ m^{\Omega_{1,1}}(\{n_{1,1}\}) = 0 \\ m^{\Omega_{1,1}}(\Omega_{1,1}) = .5 \end{cases} \quad \begin{cases} m^{\Omega_{1,2}}(\{y_{1,2}\}) = .7 \\ m^{\Omega_{1,2}}(\{n_{1,2}\}) = .3 \\ m^{\Omega_{1,2}}(\Omega_{1,2}) = 0 \end{cases} .$$

By expressing the beliefs on the frames Ω_{X_i} :

$$\begin{cases} m_1^{\Omega_{X_1}}(\{1\}) = .5 \\ m_1^{\Omega_{X_1}}(\{\overline{1}\}) = 0 \\ m_1^{\Omega_{X_1}}(\Omega_{X_1}) = .5 \end{cases} \quad \begin{cases} m_2^{\Omega_{X_1}}(\{2\}) = .7 \\ m_2^{\Omega_{X_1}}(\{\overline{2}\}) = .3 \\ m_2^{\Omega_{X_1}}(\Omega_{X_1}) = 0 \end{cases} ,$$

the following results are obtained:

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{\star\}$	$\{1, \star\}$	$\{2, \star\}$	$\{1, 2, \star\}$
$m^{\Omega_{X_1}}(A)$	.35	.15	.35	0	.15	0	0
$BetP^{\Omega_{X_1}}(A)$		.35	.54	.11	.46	.65	1

Then, regarding object X_1 :

- X_1 is associated with Y_2 ,
- Y_1 has disappeared.

From Y_1 's and Y_2 's points of view:

$$\begin{cases} m_1^{\Omega_{Y_1}}(\{1\}) = .5 \\ m_1^{\Omega_{Y_1}}(\overline{\{1\}}) = 0 \\ m_1^{\Omega_{Y_1}}(\Omega_{Y_1}) = .5 \end{cases} \quad \begin{cases} m_1^{\Omega_{Y_2}}(\{1\}) = .7 \\ m_1^{\Omega_{Y_2}}(\overline{\{1\}}) = .3 \\ m_1^{\Omega_{Y_2}}(\Omega_{Y_2}) = 0 \end{cases} \quad (26)$$

So:

A	{1}	{★}
$BetP^{\Omega_{Y_1}}$	.75	.25
$BetP^{\Omega_{Y_2}}$	.70	.30

(27)

As $.75 \times .3 > .7 \times .25$, $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}$ reaches its valid maximum at $\{1, \star\}$, which implies that:

- Y_1 is associated with X_1 ,
- Y_2 has disappeared.

This decision is then different from the previous one.

Let us also remark that the introduction of a rejection decision, as presented in Section 2.3, can also imply a different decision according to the X_i or Y_j points of view. For instance, by choosing c_0 equal to 0.5 in the Example 2, from X_1 the same decision is made as $BetP^{\Omega_{X_1}}(\{2\}) \geq 1 - c_0$, however as $BetP^{\Omega_{Y_1} \times \Omega_{Y_2}}(\{\star, 1\}) < 1 - c_0$, the decision made according to the Y_j is d_0 (a rejection).

A practical solution consists in choosing a decision by favouring either the perceived objects or the known objects.

However, to relativize this problem, conflicting decisions happen in very few cases in the particular application described in Section 5, less than 1% of the cases in this example.

4. Discussion

4.1. What's new in comparison to Rombaut and Gruyer's approaches?

The approach presented in this article differs mainly from Rombaut and Gruyer's approaches [23, 16] by regarding two points:

1. the combination of BBAs $m_j^{\Omega_{X_i}} = m^{\Omega_{i,j} \uparrow \Omega_{X_i}}$ and $m_i^{\Omega_{Y_j}} = m^{\Omega_{i,j} \uparrow \Omega_{Y_j}}$;
2. the decision-making process.

In both Rombaut's approach [23] and Gruyer's approach [16], BBAs $m_j^{\Omega_{X_i}}$ and $m_i^{\Omega_{Y_j}}$ are not classically conjunctively combined with (14). To simplify the combination and to make it computationally efficient, it is proposed to allocate masses only on singletons and the universe. Thus the following mergers are proposed, $\forall i \in I$:

$$\begin{aligned}
m_{Rombaut}^{\Omega_{X_i}}(\{\emptyset\}) &= m^{\Omega_{X_i}}(\{\emptyset\}) \\
m_{Rombaut}^{\Omega_{X_i}}(\{k\}) &= m^{\Omega_{X_i}}(\{k\}), \quad \forall k \in J, \\
m_{Rombaut}^{\Omega_{X_i}}(\{\star\}) &= m^{\Omega_{X_i}}(\{\star\}) \\
m_{Rombaut}^{\Omega_{X_i}}(\Omega_{X_i}) &= 1 - m^{\Omega_{X_i}}(\{\emptyset\}) - m^{\Omega_{X_i}}(\{\star\}) \\
&\quad - \sum_{k \in J} m^{\Omega_{X_i}}(\{k\}).
\end{aligned} \tag{28}$$

In [16], the authors suggest a decision-making system based on BBAs $m^{\Omega_{X_i}}$ and $m^{\Omega_{Y_j}}$ whose focal elements, thanks to Rombaut's combination, are either a singleton or the universe. In outline:

- An association matrix $N \times M$ is built such that each of its elements (i, j) is equal to the product $m^{\Omega_{X_i}}(\{j\}) \times m^{\Omega_{Y_j}}(\{i\})$. Each row i is then associated with a perceived object X_i , and each column j is associated with a known object Y_j .
- If necessary, fictive objects are added to make the latter matrix squared.
- A coupling algorithm, the Hungarian algorithm, is then applied to this matrix, this latter algorithm providing an optimal decision regarding the sum of the beliefs.
- A final treatment deals with the objects appearance.

Applied to the examples presented in [23] and [16], the model presented in this article and Gruyer's approach lead to the same results.

Example 4 (Rombaut [23]). *Let us consider two perceived objects X_i and two known objects Y_j , s.t.:*

$$\begin{aligned}
\begin{cases} m^{\Omega_{1,1}}(\{y_{1,1}\}) &= 0.8 \\ m^{\Omega_{1,1}}(\{n_{1,1}\}) &= 0.1 \\ m^{\Omega_{1,1}}(\Omega_{1,1}) &= 0.1 \end{cases} & \quad \begin{cases} m^{\Omega_{1,2}}(\{y_{1,2}\}) &= 0.7 \\ m^{\Omega_{1,2}}(\{n_{1,2}\}) &= 0.2 \\ m^{\Omega_{1,2}}(\Omega_{1,2}) &= 0.1 \end{cases} \\
\begin{cases} m^{\Omega_{2,1}}(\{y_{2,1}\}) &= 0.8 \\ m^{\Omega_{2,1}}(\{n_{2,1}\}) &= 0.1 \\ m^{\Omega_{2,1}}(\Omega_{2,1}) &= 0.1 \end{cases} & \quad \begin{cases} m^{\Omega_{2,2}}(\{y_{2,2}\}) &= 0.6 \\ m^{\Omega_{2,2}}(\{n_{2,2}\}) &= 0.3 \\ m^{\Omega_{2,2}}(\Omega_{2,2}) &= 0.1 \end{cases}
\end{aligned} \tag{29}$$

By expressing these beliefs over Ω_{X_1} and Ω_{X_2} , it is obtained:

$$\begin{cases} m_1^{\Omega_{X_1}}(\{1\}) = 0.8 \\ m_1^{\Omega_{X_1}}(\overline{\{1\}}) = 0.1 \\ m_1^{\Omega_{X_1}}(\Omega_{X_1}) = 0.1 \end{cases} \begin{cases} m_2^{\Omega_{X_1}}(\{2\}) = 0.7 \\ m_2^{\Omega_{X_1}}(\overline{\{2\}}) = 0.2 \\ m_2^{\Omega_{X_1}}(\Omega_{X_1}) = 0.1 \end{cases} \begin{cases} m_1^{\Omega_{X_2}}(\{1\}) = 0.8 \\ m_1^{\Omega_{X_2}}(\overline{\{1\}}) = 0.1 \\ m_1^{\Omega_{X_2}}(\Omega_{X_2}) = 0.1 \end{cases} \begin{cases} m_2^{\Omega_{X_2}}(\{2\}) = 0.6 \\ m_2^{\Omega_{X_2}}(\overline{\{2\}}) = 0.3 \\ m_2^{\Omega_{X_2}}(\Omega_{X_2}) = 0.1 \end{cases}$$

Pignistic probabilities are then defined by:

A	{1}	{2}	{*}
$BetP^{\Omega_{X_1}}$	0.57	0.34	0.09
$BetP^{\Omega_{X_2}}$	0.65	0.25	0.10

(30)

Thus $BetP^{\Omega_{X_1} \times \Omega_{X_2}}$ reaches its valid maximum at $\{2, 1\}$, and regarding the X_i :

1. X_1 is associated with Y_2 ;
2. X_2 is associated with Y_1 ;

By expressing the beliefs regarding objects Y_j , we have:

$$\begin{cases} m_1^{\Omega_{Y_1}}(\{1\}) = 0.8 \\ m_1^{\Omega_{Y_1}}(\overline{\{1\}}) = 0.1 \\ m_1^{\Omega_{Y_1}}(\Omega_{Y_1}) = 0.1 \end{cases} \begin{cases} m_2^{\Omega_{Y_1}}(\{2\}) = 0.8 \\ m_2^{\Omega_{Y_1}}(\overline{\{2\}}) = 0.1 \\ m_2^{\Omega_{Y_1}}(\Omega_{Y_1}) = 0.1 \end{cases} \begin{cases} m_1^{\Omega_{Y_2}}(\{1\}) = 0.7 \\ m_1^{\Omega_{Y_2}}(\overline{\{1\}}) = 0.2 \\ m_1^{\Omega_{Y_2}}(\Omega_{Y_2}) = 0.1 \end{cases} \begin{cases} m_2^{\Omega_{Y_2}}(\{2\}) = 0.6 \\ m_2^{\Omega_{Y_2}}(\overline{\{2\}}) = 0.3 \\ m_2^{\Omega_{Y_2}}(\Omega_{Y_2}) = 0.1 \end{cases}$$

Then:

A	{1}	{2}	{*}
$BetP^{\Omega_{Y_1}}$	0.47	0.47	0.06
$BetP^{\Omega_{Y_2}}$	0.52	0.33	0.15

(31)

Conclusion from Y_1 's and Y_2 's point of view:

1. $BetP^{\Omega_{X_1} \times \Omega_{X_2}}$'s maximum for a valid association is obtained in $\{2, 1\}$, then Y_1 is associated with X_2 , and Y_2 is associated with X_1 .

Example 5 (Gruyer [16]). In this example, three perceived objects (X_i)

and four objects (Y_j) are considered such that:

$$\begin{aligned} &\begin{cases} m^{\Omega_{1,1}}(\{y_{1,1}\}) = 0.80 \\ m^{\Omega_{1,1}}(\{n_{1,1}\}) = 0 \\ m^{\Omega_{1,1}}(\Omega_{1,1}) = 0.20 \end{cases} \begin{cases} m^{\Omega_{1,2}}(\{y_{1,2}\}) = 0 \\ m^{\Omega_{1,2}}(\{n_{1,2}\}) = 0.99 \\ m^{\Omega_{1,2}}(\Omega_{1,2}) = 0.01 \end{cases} \begin{cases} m^{\Omega_{1,3}}(\{y_{1,3}\}) = 0 \\ m^{\Omega_{1,3}}(\{n_{1,3}\}) = 0.97 \\ m^{\Omega_{1,3}}(\Omega_{1,3}) = 0.03 \end{cases} \begin{cases} m^{\Omega_{1,4}}(\{y_{1,4}\}) = 0 \\ m^{\Omega_{1,4}}(\{n_{1,4}\}) = 0.99 \\ m^{\Omega_{1,4}}(\Omega_{1,4}) = 0.01 \end{cases} \\ &\begin{cases} m^{\Omega_{2,1}}(\{y_{2,1}\}) = 0.57 \\ m^{\Omega_{2,1}}(\{n_{2,1}\}) = 0 \\ m^{\Omega_{2,1}}(\Omega_{2,1}) = 0.43 \end{cases} \begin{cases} m^{\Omega_{2,2}}(\{y_{2,2}\}) = 0.57 \\ m^{\Omega_{2,2}}(\{n_{2,2}\}) = 0 \\ m^{\Omega_{2,2}}(\Omega_{2,2}) = 0.43 \end{cases} \begin{cases} m^{\Omega_{2,3}}(\{y_{2,3}\}) = 0 \\ m^{\Omega_{2,3}}(\{n_{2,3}\}) = 0.52 \\ m^{\Omega_{2,3}}(\Omega_{2,3}) = 0.48 \end{cases} \begin{cases} m^{\Omega_{2,4}}(\{y_{2,4}\}) = 0 \\ m^{\Omega_{2,4}}(\{n_{2,4}\}) = 0.99 \\ m^{\Omega_{2,4}}(\Omega_{2,4}) = 0.01 \end{cases} \\ &\begin{cases} m^{\Omega_{3,1}}(\{y_{3,1}\}) = 0 \\ m^{\Omega_{3,1}}(\{n_{3,1}\}) = 0.99 \\ m^{\Omega_{3,1}}(\Omega_{3,1}) = 0.01 \end{cases} \begin{cases} m^{\Omega_{3,2}}(\{y_{3,2}\}) = 0.61 \\ m^{\Omega_{3,2}}(\{n_{3,2}\}) = 0 \\ m^{\Omega_{3,2}}(\Omega_{3,2}) = 0.39 \end{cases} \begin{cases} m^{\Omega_{3,3}}(\{y_{3,3}\}) = 0 \\ m^{\Omega_{3,3}}(\{n_{3,3}\}) = 0.52 \\ m^{\Omega_{3,3}}(\Omega_{3,3}) = 0.48 \end{cases} \begin{cases} m^{\Omega_{3,4}}(\{y_{3,4}\}) = 0 \\ m^{\Omega_{3,4}}(\{n_{3,4}\}) = 0.99 \\ m^{\Omega_{3,4}}(\Omega_{3,4}) = 0.01 \end{cases} \end{aligned}$$

Pignistic probabilities regarding objects X_i are given by:

A	{1}	{2}	{3}	{4}	{★}
$BetP^{\Omega_{X_1}}$	0.90	0.00	0.00	0.00	0.10
$BetP^{\Omega_{X_2}}$	0.44	0.44	0.03	0.00	0.08
$BetP^{\Omega_{X_3}}$	0.00	0.77	0.06	0.00	0.16

(32)

The association result can be obtained from Algorithm 1. However, when numerous probabilities are equal to zero, the following Algorithm 2 can be employed.

The use of Algorithm 2 on this example is illustrated in Figure 3.

$BetP^{\Omega_{X_1} \times \Omega_{X_2} \times \Omega_{X_3}}$ reaches its maximum $.9 \times .08 \times .77$ at $\{1, \star, 2\}$. Thus:

1. X_1 is associated with Y_1 ;
2. X_2 has appeared;
3. X_3 is associated with Y_2 ;
4. then Y_3 and Y_4 have disappeared.

On the other side, by considering Y_j 's' points of view:

A	{1}	{2}	{3}	{★}
$BetP^{\Omega_{Y_1}}$	0.69	0.26	0.00	0.05
$BetP^{\Omega_{Y_2}}$	0.00	0.42	0.49	0.09
$BetP^{\Omega_{Y_3}}$	0.01	0.20	0.20	0.59
$BetP^{\Omega_{Y_4}}$	0.05	0.05	0.05	0.985

(33)

$BetP^{\Omega_{Y_1} \times \Omega_{Y_2} \times \Omega_{Y_3} \times \Omega_{Y_4}}$ reaches its maximum $.69 \times .49 \times .59 \times .985$ at $\{1, 3, \star, \star\}$.

Then:

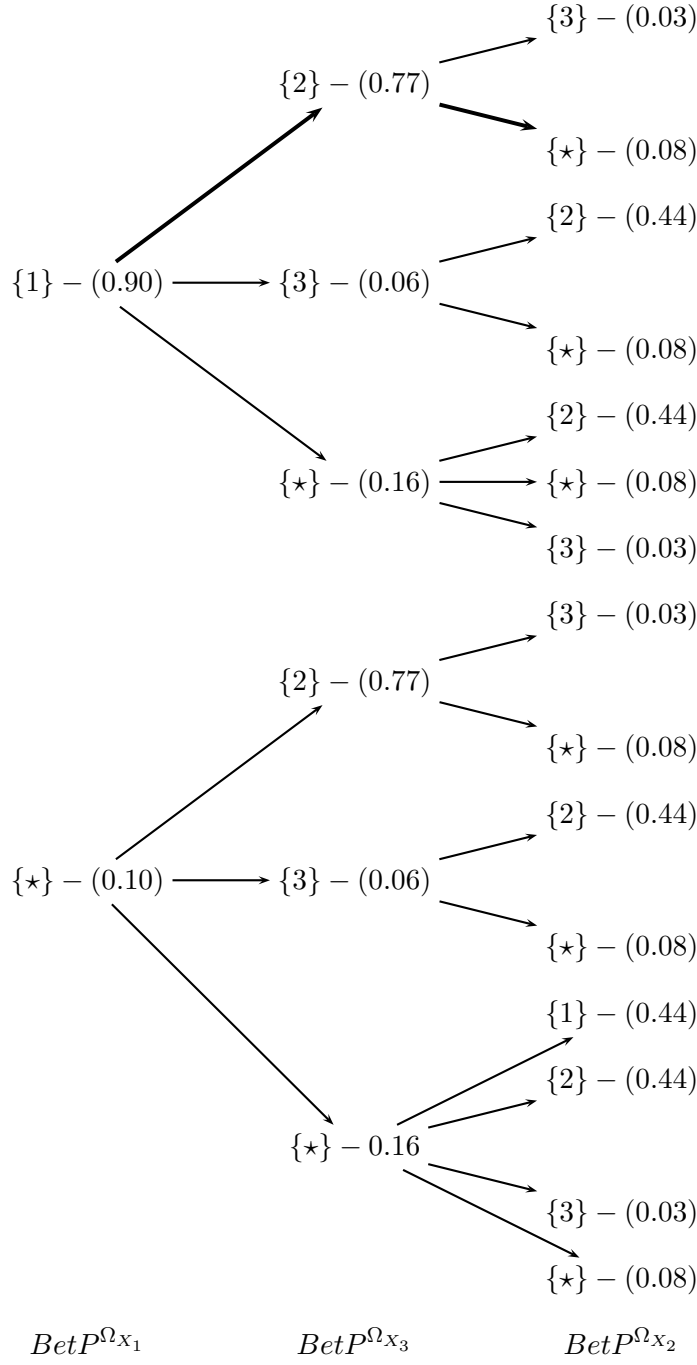


Figure 3: Illustration of the use of Algorithm 2 on Example 5

Algorithm 2: Decision regarding the perceived objects X_i

Data: $BetP^{\Omega_{X_i}}, i \in I = \{1, \dots, N\}$, pignistic probabilities of each perceived object X_i .

Result: The decision maximizing $BetP^{\Omega_{X_1} \times \dots \times \Omega_{X_N}}$ and verifying the association constraints

begin

- Rank in decreasing order the pignistic probabilities $BetP^{\Omega_{X_i}}$ according to the number of singletons which probabilities are equal to zero.
- Delete each branch which do not correspond to a valid association
- Compute the products of the pignistic probabilities on each branch
- Select the association maximizing the products

end

- Y_1 is associated with X_1 ;
- Y_2 is associated with X_3 ;
- Y_3 has disappeared;
- Y_4 has disappeared;
- and, X_2 has appeared.

In this example, the solution proposed by Gruyer in [16] leads to the same conclusion.

4.2. About Ristic and Smets' approach

The problem tackled by Ristic and Smets in [22] is somewhat different from the association problem described in this article. Ristic and Smets consider a given volume of interest containing an unknown number of objects. While sensors we consider give information regarding the associations of each object detected at a time step t , with previous objects detected at a previous time step $t - 1$, Ristic and Smets's sensors provide information regarding the class of each object they have detected in the scene, for instance helicopter, airplane, ... The "association problem" they try to solve consists then in determining the number of objects as well as the class of

each one. Besides, the appearance and disappearance of objects do not take directly part of their problem. The application of Ristic and Smets' works to our problem is consequently not straightforward.

However, some technical points of this model should be taken into account and investigated.

Following [10], the authors remark that the mass given to the empty set, after conjunctively combining two BBAs expressing themselves on the class of two different objects is equal to the belief that these two objects do not belong to the same class, an idea already present in [1] (multi-sensor fusion for submarine detection) and in [25] (intelligence clustering).

At last, the criteria the authors maximize is based on the plausibility of each possible associations. As justified in [28], the pignistic transformation has been chosen to make the decision in this article. A first investigation in the direction of the plausibility consists in using the plausibility-probability transformation [5].

5. Results on real data

In this section, the approach presented in this article (Section 3) is compared to the approach of Rombaut and Gruyer on real data coming from a DV camera placed behind the windshield of a car. This DV camera has a CCD sensor, a 720×576 pixels resolution, an angle ranging from -0.5 to $+0.5$ radians (i.e. approximately $\pm 30^\circ$), and works at 25 images per second ($\Delta_t = .04s$), a filmed image example being presented in Figure 4.

The video sequence allowing one to compare the two approaches includes about 3250 images corresponding to a 130-second playing time. Each images contains 1 to 6 objects. The distribution of the number of objects in each image is illustrated in Figure 5. It can be observed that the sequence to analyse is mainly composed of images with 1 object, 2 objects, 3 objects or more.

During the sequence, 75 distinct objects were manually identified as illustrated in Figure 4 (no automatic treatment of the images), the number of associations to realize being equal to 6800. The ground truth is then



Figure 4: Four vehicles in a selected image in the sequence.

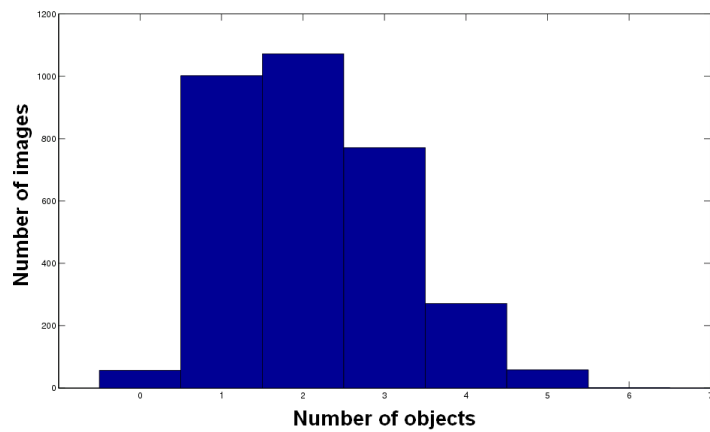


Figure 5: Number of objects in each image during the sequence.

known. It allows us to compute the good recognition rate (GRR) of each approach during this sequence. Formally, this rate is defined as follows:

$$GRR = \frac{\text{Number of correct associations}}{\text{Number of association to realize}}$$

In the same manner, the rejection rate (RR) and the error rate (ER) are defined by:

$$RR = \frac{\text{Number of rejections}}{\text{Number of association to realize}}$$

$$ER = \frac{\text{Number of incorrect associations}}{\text{Number of association to realize}}.$$

The sum of these rates is equal to 1.

Distance and angle criteria allow the creation of two mass functions $m_{distance}^{\Omega_{i,j}}$ and $m_{angle}^{\Omega_{i,j}}$, regarding the association between each perceived object X_i and known object Y_j .

The distance was estimated as a function of the height and the width in pixels of the object observed in the scene thanks to an interpolation method illustrated in Figure 6.

On the other hand, the angle between two objects is computed from the gravity centre of the perceived object in the image (Figure 4).

The measurements provided are very noisy. For instance, there can be a variation of $20m$ for the same object from an image to the next one. Likewise, angle variations can be as high as 100%, from $0.01rd$ to $0.02rd$ for two consecutive measurements of the same object.

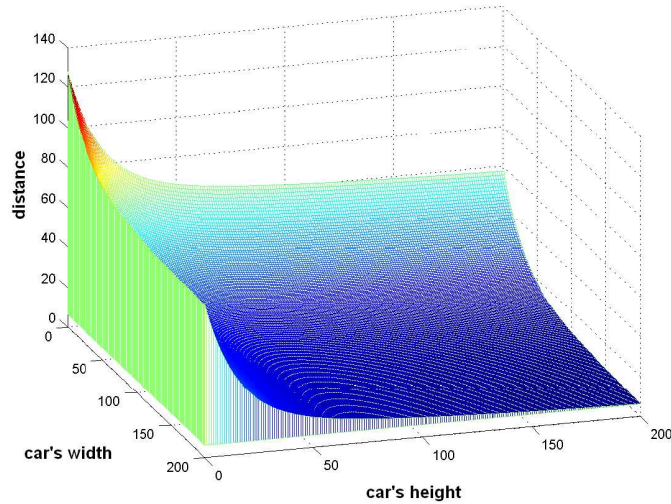


Figure 6: Interpolation function giving the distance in meter depending on the height and the width in pixels of the object in the scene.

In this application, masses are fixed in the following way:

$$\begin{cases} m^{\Omega_{i,j}}(\{y_{i,j}\}) &= \beta \phi_{i,j}(e_{i,j}) \\ m^{\Omega_{i,j}}(\{n_{i,j}\}) &= \beta (1 - \phi_{i,j}(e_{i,j})) \\ m^{\Omega_{i,j}}(\Omega_{i,j}) &= 1 - \beta \end{cases} \quad (34)$$

where:

- $0 < \beta < 1$ is a constant representing the degree of reliability of the source of information (cf the discounting operation [26, page 252], see [18, 21] for other correction mechanisms).
- $\phi_{i,j}(\cdot)$ is a monotone decreasing function s.t. $\phi_i(0) = 1$ and $\lim_{e \rightarrow \infty} \phi_i(e) = 0$;
- $e_{i,j}$ is the dissimilarity measure between the perceived object X_i and the known object Y_j , which means the difference of distance and the difference of angle in this application.

The function $\phi_{i,j}$ is chosen as follows [8]:

$$\phi_{i,j}(e_{i,j}) = \exp^{-(e_{i,j})^2}. \quad (35)$$

Constant β being fixed at 0.9, these two mass functions are combined thanks to the Dempster's rule of combination to obtain a mass function $m^{\Omega_{i,j}}$:

$$m^{\Omega_{i,j}} = m_{distance}^{\Omega_{i,j}} \oplus m_{angle}^{\Omega_{i,j}} \quad \forall i \in I, \forall j \in J. \quad (36)$$

The association model presented in Section 3 only need one BBA expressing the information regarding the association between object X_i and object Y_j . In this application, we are lucky enough to have two information sources. Thus these two pieces of information are firstly combined using a well justified rule for the combination of two distinct sources. The choice to combine theses sources at this step, and the choice of the rule have been left for further study.

In Figure 7, the good recognition rate of the two approaches presented in this article obtained in this video sequence is represented as a function of the rejection cost (Section 2.3). It can be observed that as soon as the rejection cost becomes greater than 0, the good recognition rates obtained with the conjunctive combination are greater than those obtained with Rombaut's combination, which is recalled to be also used in Gruyer's approach.

Let us note that the decisions have been computed on the basis of the perceived objects. As mentioned in Section 3.4, these decisions are not necessary identical with those computed from the known objects point of view. However, as illustrated in Figure 8, this conflicting decision rate remains very low in this application (from 0% to less than 1% depending on the rejection cost). Let us also recall that, as illustrated at the end of Example 2, the introduction of a rejection cost enhances the appearance of conflicting decisions.

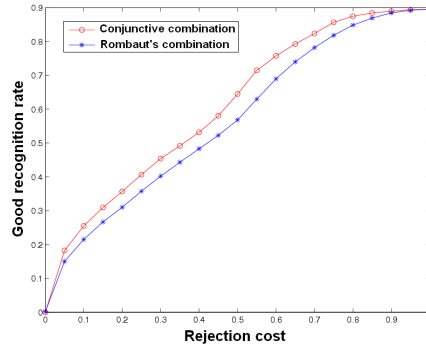


Figure 7: Good recognition rate in function of the rejection cost.

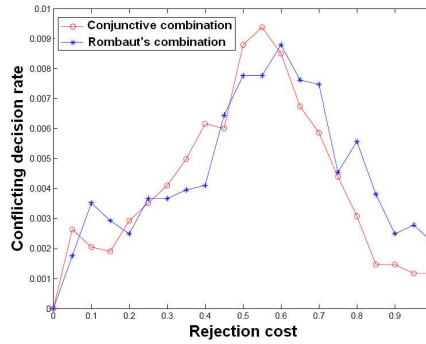


Figure 8: Conflicting decision rate in function of the rejection cost.

6. Conclusion and prospects

In this article, a modelling of the association problem in the belief function framework has been presented. It has been shown in particular, how tools from the theory of belief functions such as the vacuous extension, the

conjunctive combination rule and the pignistic transformation can be applied. Validated on real data, this approach can perform better good recognition rates than Rombaut's initial approach as soon as a rejection cost is introduced.

Concerning the prospects, even if it concerns a reduce number of cases, the resolution of the possible conflicting decisions between perceived and known objects have to be deeper investigated, even if it seems an integral part of this problem.

The decomposition of the BBAs [11] expressing the beliefs regarding the associations between known objects and perceived objects could also be studied in order to use a more adapted rule.

Subsequently, this approach should be enhanced by introducing information coming from the tracking of vehicles at time steps preceding the current analysis.

Acknowledgements

These works have been financed by the French region Nord-Pas de Calais under the project CISIT (Campus International pour la Sécurité et l'Intermodalité des Transports).

References

- [1] A. Ayoun and Ph. Smets. Data association in multi-target detection using the transferable belief model. *International Journal of Intelligent Systems*, 16(10):1167–1182, 2001.
- [2] Y. Bar-Shalom. *Multitarget/Multisensor Tracking: Applications and Advances vol III*, Artech House, 2000.
- [3] S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.
- [4] I. Bloch. Fusion of Information under Imprecision and Uncertainty, Numerical Methods, and Image Information Fusion. *Multisensor Data Fusion*, A.K. Hyder et al (Eds.), pp. 267–294, 2002.
- [5] B.R. Cobb, P.P. Shenoy. On the plausibility transformation method for translating belief function models to probability models. *International Journal of Approximate Reasoning*, 41(3):314–330, 2006.

- [6] M.H. DeGroot. *Optimal Statistical Decisions*. McGraw-Hill, New York, 1970.
- [7] A. Dempster. A generalization of Bayesian inference. *Journal of Royal Statistical Society*, B 30:205-247, 1968.
- [8] T. Denœux. A k-nearest neighbour classification rule based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(5):804-813, 1995.
- [9] T. Denœux. Analysis of evidence-theoretic decision rules for pattern classification. *Pattern Recognition*, 30(7):1095–1107, 1997.
- [10] T. Denœux and M. Masson. EVCLUS: Evidential Clustering of Proximity Data. *IEEE Transactions on Systems, Man and Cybernetics B*, 34(1):95–109, 2004.
- [11] T. Denœux. Conjunctive and Disjunctive Combination of Belief Functions Induced by Non Distinct Bodies of Evidence. *Artificial Intelligence*, 172:234-264, 2008.
- [12] T. Denœux. Extending stochastic ordering to belief functions on the real line. *Information Sciences*, 179(9):1362–1376, 2009.
- [13] D. Dubois and R.R. Yager. Fuzzy set connectives as combinations of belief structures. *Information Sciences*, 66(3):245–276, 1992.
- [14] D. Dubois, H. Prade, and Ph. Smets. Representing partial ignorance. *IEEE Transactions on Systems, Man and Cybernetics*, 26(3):361–377, 1996.
- [15] M. El Najjar and P. Bonnifait. A road-matching method for precise vehicle localization using belief theory and Kalman filtering. *Autonomous Robots*, 19(2):173-191, 2005.
- [16] D. Gruyer, C. Royère, R. Labayrade and D. Aubert. Credibilistic multi sensor fusion for real time application, application to obstacle detection and tracking. *IEEE Int. Conf. on Advanced Robotics, ICAR'2003*, paper P366, 2003.
- [17] Y. Lemeret, E. Lefevre and D. Jolly. Improvement of an association algorithm for obstacle tracking. *Information Fusion*, 9(2):234-245, 2008.

- [18] D. Mercier, B. Quost and T. Denœux. Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 9(2):246–258, 2008.
- [19] D. Mercier, É. Lefèvre and D. Jolly. Object association in the TBM framework, application to vehicle driving aid. In the *6th International Symposium on Imprecise Probability: Theories and Applications, ISIPTA 2009*, Durham, United Kingdom, July 14-18, 2009.
- [20] D. Mercier, G. Cron, T. Denœux and M.-H. Masson. Decision fusion for postal address recognition using belief functions. *Expert Systems with Applications*, 36(3):5643–5653, 2009.
- [21] D. Mercier, T. Denœux and M.-H. Masson. Belief function correction mechanisms. *Studies in Fuzziness and Soft Computing*, B. Bouchon-Meunier et al. (Eds.), Vol. 249, pp. 203–222, January 2010.
- [22] B. Ristic and Ph. Smets. Global cost of assignment in the TBM framework for association of uncertain ID reports. *Aerospace Science and Technology*, 11(4):303–309, 2007.
- [23] M. Rombaut. Decision in multi-obstacle matching process using the theory of belief. *Advances in Vehicle Control and Safety, AVCS98*, pp. 63–68, 1998.
- [24] M. Rombaut and V. Cherfaoui. Decision making in data fusion using Dempster-Shafer’s theory. *Symposium on Intelligent Components and Instrumentation for Control Applications*, 1997.
- [25] J. Schubert. Managing inconsistent intelligence. In *Proceedings of the 3rd International Conference on Information Fusion, FUSION’2000*, pp. TuB4/10–16, Paris, France, 2000.
- [26] G. Shafer. *A Mathematical Theory Of Evidence*. Princeton University Press, Princeton, N.J., 1976.
- [27] Ph. Smets. Belief functions: the disjunctive rule of combination and the generalized bayesian theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.
- [28] Ph. Smets. Decision making in the TBM: the necessity of the pignistic transformation. *International Journal of Approximate Reasoning*, 38(2):133–147, 2005.

- [29] Ph. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–243, 1994.
- [30] Y.M. Wang and J.B. Yang and D.L. Xu and K.S. Chin. On the combination and normalization of interval-valued belief structures. *Information Sciences*, 177(5):1230-1247, 2007.
- [31] W.Z. Wu. Attribute reduction based on evidence theory in incomplete decision systems. *Information Sciences*, 178(5):1355-1371, 2008.

Belief Functions Contextual Discounting and Canonical Decompositions

David Mercier, Éric Lefèvre, François Delmotte

*Univ. Lille Nord de France, F-59000, Lille, France;
UArtois, LGI2A, F-62400, Béthune, France*

Abstract

In this article, the contextual discounting of a belief function, a classical discounting generalization, is extended and its particular link with the canonical disjunctive decomposition is highlighted. A general family of correction mechanisms allowing one to weaken the information provided by a source is then introduced, as well as the dual of this family allowing one to strengthen a belief function.

Keywords: Belief Functions, Discounting, Reinforcement, Canonical Decompositions.

1. Introduction

In the Dempster-Shafer theory of belief functions [3, 23], the reliability of a source of information is classically taken into account by the discounting operation [23, page 252], which transforms a belief function into a weaker, less informative one. This operation is usually important in uncertain information management [1, 2, 6, 7, 8, 10, 13, 17, 16, 18, 30].

Introduced in [12], the contextual discounting is a refinement of the discounting operation. It takes into account the fact that the reliability of a source of information can be expected to depend on the true answer of the question of interest.

For instance, in medical diagnosis, depending on his/her speciality, experience or training, a physician may be more or less competent to diagnose some types of diseases. Likewise, in target recognition, a sensor may be

Email addresses: david.mercier@univ-artois.fr (David Mercier),
eric.lefevre@univ-artois.fr (Éric Lefèvre), francois.delmotte@univ-artois.fr
(François Delmotte)

more capable of recognizing some types of targets while being less effective for other types.

In this contextual model, the agent in charge of the fusion process or the decision making can hold some knowledge regarding the reliability of a source of information in different contexts, which forms a partition of the universe of discourse. For example, a sensor in charge of recognizing targets can be more or less reliable depending on the fact that the target is a helicopter (h), an airplane (a) or a rocket(r). Here, subsets $\{h\}$, $\{r\}$ and $\{a\}$ form the finest partition of the universe $\Omega = \{a, h, r\}$. However, this model cannot handle the reliability knowing that the target is a helicopter or a rocket ($\{h, r\}$) as well as the reliability knowing that the target is an airplane or a helicopter ($\{a, h\}$), because the sets $\{a, h\}$ and $\{h, r\}$ do not form a partition of Ω .

This last step is reached in this article. Contextual discounting exposed previously in [12] is then shown to be a particular case of a more general correction process [14, 11] allowing the discounting of a belief function in a finer way. In particular, a simple expression of this mechanism is given in the form of disjunctive combinations. At last, the dual version of the contextual discounting, allowing one to reinforce a belief function, is also introduced. This article extends deeply a first version of this work presented in [15].

To develop the justifications of these mechanisms, belief functions are interpreted as expressing weighted opinions, irrespective of any underlying probability distributions, and the Transferable Belief Model [26, 27, 29] is adopted.

This article is organized as follows. Background material needed on belief functions is recalled in Section 2. Contextual discounting is extended in Section 3. A dual reinforcement process is introduced in Section 4. Different aspects of these mechanisms are then discussed in Section 5, and finally, Section 6 concludes this article.

2. Belief functions: basic concepts

2.1. Representing information

Let us consider an agent Ag in charge of making a decision regarding the answer to a given question Q of interest.

Let $\Omega = \{\omega_1, \dots, \omega_K\}$, called the *frame of discernment*, be the finite set containing the possible answers to question Q .

The information held by agent Ag regarding the answer to question Q can be quantified by a *basic belief assignment (BBA)* or a *mass function*

m_{Ag}^Ω , defined as a function from 2^Ω to $[0, 1]$, verifying:

$$\sum_{A \subseteq \Omega} m_{Ag}^\Omega(A) = 1 . \quad (1)$$

Function m_{Ag}^Ω describes the state of knowledge of agent Ag regarding the answer to question Q belonging to Ω . By extension, it also represents an item of evidence that induces such a state of knowledge. The quantity $m_{Ag}^\Omega(A)$ is interpreted as the part of the unit mass allocated to the hypothesis: “the answer to question Q is in the subset A of Ω ”.

When there is no ambiguity, the full notation m_{Ag}^Ω will be simplified to m^Ω , or even m .

The following definitions and notations are considered.

Definition 1 (Focal element). *A subset A of Ω such that $m(A) > 0$ is called a focal element of m .*

Definition 2 (Categorical BBA). *A BBA m with only one focal element A is said to be categorical and is denoted m_A ; we thus have $m_A(A) = 1$.*

Definition 3 (Vacuous mass function). *Total ignorance is represented by the BBA m_Ω , called the vacuous mass function.*

Definition 4. *A BBA m is said to be:*

- dogmatic if $m(\Omega) = 0$;
- non-dogmatic if $m(\Omega) > 0$;
- normal if $m(\emptyset) = 0$;
- subnormal if $m(\emptyset) > 0$;
- simple if m has no more than two focal sets, Ω being included.

Definition 5 (Negation of a BBA). *Function $\overline{m}$ denotes the negation of m [5], defined by $\overline{m}(A) = m(\overline{A})$, for all $A \subseteq \Omega$ such that $\overline{A}$ is the complement of A in Ω .*

Definition 6. *The belief, plausibility, implicability and commonality functions associated with a mass function m are respectively defined by:*

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B), \quad (2)$$

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B), \quad (3)$$

$$b(A) = bel(A) + m(\emptyset) = 1 - pl(\bar{A}), \quad (4)$$

and

$$q(A) = \sum_{B \supseteq A} m(B), \quad (5)$$

for all $A \subseteq \Omega$.

Functions bel , pl , b and q are all in one-to-one correspondence [31], and represent then the same information.

2.2. Combining pieces of information

Two BBAs m_1 and m_2 induced by distinct and reliable sources of information can be combined using the *conjunctive rule of combination (CRC)*, also referred to as the *unnormalized Dempster's rule of combination*, defined for all $A \subseteq \Omega$ by:

$$m_1 \odot m_2(A) = \sum_{B \cap C = A} m_1(B) m_2(C). \quad (6)$$

Alternatively, if we only know that at least one of the sources is reliable, BBAs m_1 and m_2 can be combined using the *disjunctive rule of combination (DRC)*, defined for all $A \subseteq \Omega$ by:

$$m_1 \oplus m_2(A) = \sum_{B \cup C = A} m_1(B) m_2(C). \quad (7)$$

2.3. Marginalization and vacuous extension on a product space

A mass function defined on a product space $\Omega \times \Theta$ may be *marginalized* on Ω by transferring each mass $m^{\Omega \times \Theta}(B)$ for $B \subseteq \Omega \times \Theta$ to its projection on Ω :

$$m^{\Omega \times \Theta \downarrow \Omega}(A) = \sum_{\substack{B \subseteq \Omega \times \Theta, \\ \text{Proj}(B \downarrow \Omega) = A}} m^{\Omega \times \Theta}(B), \quad (8)$$

for all $A \subseteq \Omega$ where $\text{Proj}(B \downarrow \Omega)$ denotes the projection of B onto Ω .

Conversely, it is usually not possible to retrieve the original BBA $m^{\Omega \times \Theta}$ from its marginal $m^{\Omega \times \Theta \downarrow \Omega}$ on Ω . However, the *least committed*, or *least informative BBA* [25] such that its projection on Ω is $m^{\Omega \times \Theta \downarrow \Omega}$ may be computed. This defines the *vacuous extension* of m^Ω in the product space $\Omega \times \Theta$, noted $m^{\Omega \uparrow \Omega \times \Theta}$, and given by:

$$m^{\Omega \uparrow \Omega \times \Theta}(B) = \begin{cases} m^\Omega(A) & \text{if } B = A \times \Theta, A \subseteq \Omega, \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

2.4. Conditioning and ballooning extension on a product space

Conditional beliefs represent knowledge that is valid provided that an hypothesis is satisfied. Let m be a mass function and $B \subseteq \Omega$ an hypothesis; the *conditional belief function* $m[B]$ is given by:

$$m[B] = m \odot m_B. \quad (10)$$

If $m^{\Omega \times \Theta}$ is defined on the product space $\Omega \times \Theta$, and θ is a subset of Θ , the conditional BBA $m^\Omega[\theta]$ is defined by combining $m^{\Omega \times \Theta}$ with $m_\theta^{\Theta \uparrow \Omega \times \Theta}$, and marginalizing the result on Ω :

$$m^\Omega[\theta] = \left(m^{\Omega \times \Theta} \odot m_\theta^{\Theta \uparrow \Omega \times \Theta} \right) \downarrow^\Omega. \quad (11)$$

Assume now that $m^\Omega[\theta]$ represents the agent's beliefs on Ω conditionally on θ , i.e., in a context where θ holds. There are usually many BBAs on $\Omega \times \Theta$, whose conditioning on θ yields $m^\Omega[\theta]$. Among these, the least committed one is defined for all $A \subseteq \Omega$ by:

$$m^\Omega[\theta] \uparrow^{\Omega \times \Theta} (A \times \theta \cup \Omega \times \bar{\theta}) = m^\Omega[\theta](A). \quad (12)$$

This operation is referred to as the *deconditioning* or *ballooning extension* [25] of $m^\Omega[\theta]$ on $\Omega \times \Theta$.

2.5. Discounting

When receiving a piece of information represented by a mass function m , agent Ag may have some doubts regarding the reliability of the source that provided this information. Such metaknowledge can be taken into account using the discounting operation introduced by Shafer [23, page 252], and defined by:

$${}^\alpha m = (1 - \alpha)m + \alpha m_\Omega, \quad (13)$$

where $\alpha \in [0, 1]$.

A discount rate α equal to 1, means that the source is not reliable and the piece of information it provides cannot be taken into account, so Ag 's knowledge remains vacuous: $m_{Ag}^\Omega = {}^1 m = m_\Omega$. On the contrary, a null discount rate indicates that the source is fully reliable and the piece of information is entirely accepted: $m_{Ag}^\Omega = {}^0 m = m$. In practice, however, agent Ag usually does not know for sure whether the source is reliable or not, but has some degree of belief expressed by:

$$\begin{cases} m_{Ag}^\mathcal{R}(\{R\}) &= 1 - \alpha \\ m_{Ag}^\mathcal{R}(\mathcal{R}) &= \alpha, \end{cases} \quad (14)$$

where $\mathcal{R} = \{R, NR\}$, R and NR standing, respectively, for “the source is reliable” and “the source is not reliable”. This formalization yields expression (13), as demonstrated by Smets in [25, Section 5.7].

The discounting operation (13) of a BBA m is also equivalent to the disjunctive combination (7) of m with the mass function m_0^Ω defined by:

$$m_0^\Omega(A) = \begin{cases} \beta & \text{if } A = \emptyset \\ \alpha & \text{if } A = \Omega \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

with $\alpha \in [0, 1]$ and $\beta = 1 - \alpha$.

Indeed:

$$m \oplus m_0^\Omega(A) = m(A)m_0^\Omega(\emptyset) = \beta m(A) = {}^\alpha m(A), \quad \forall A \subset \Omega, \quad (16)$$

and

$$\begin{aligned} m \oplus m_0^\Omega(\Omega) &= m(\Omega)m_0^\Omega(\emptyset) + m_0^\Omega(\Omega) \sum_{A \subseteq \Omega} m(A) \\ &= \beta m(\Omega) + \alpha = {}^\alpha m(\Omega). \end{aligned} \quad (17)$$

2.6. Contextual Discounting based on a coarsening

Let $\Theta = \{\theta_1, \dots, \theta_L\}$ be a coarsening of Ω , which means that $\theta_1, \dots, \theta_L$ form a partition of Ω [23, chapter 6].

Unlike (14), in the contextual model, agent Ag is assumed to hold beliefs on the reliability of the source of information conditionally on each θ_ℓ , $\ell \in \{1, \dots, L\}$:

$$\begin{cases} m_{Ag}^{\mathcal{R}}[\theta_\ell](\{R\}) &= 1 - \alpha_\ell = \beta_\ell \\ m_{Ag}^{\mathcal{R}}[\theta_\ell](\mathcal{R}) &= \alpha_\ell. \end{cases} \quad (18)$$

For all $\ell \in \{1, \dots, L\}$, $\beta_\ell + \alpha_\ell = 1$, and β_ℓ represents the degree of belief that the source is reliable knowing that the true answer of the question of interest belongs to θ_ℓ .

In the same way as in the discounting operation (13), agent Ag considers that the source can be in two states: reliable or not reliable [25, 12]:

- If the source is reliable (state R), the information m_S^Ω it provides becomes Ag ’s knowledge. Formally, $m_{Ag}^\Omega[\{R\}] = m_S^\Omega$.
- If the source is not reliable (state NR), the information m_S^Ω it provides is discarded, and Ag remains in a state of ignorance: $m_{Ag}^\Omega[\{NR\}] = m_\Omega$.

The knowledge held by agent Ag , based on the information m_S^Ω from a source S as well as metaknowledge $m_{Ag}^\mathcal{R}$ concerning the reliability of the source can then be computed by:

- Deconditioning the L BBAs $m_{Ag}^\mathcal{R}[\theta_\ell]$ on the product space $\Omega \times \mathcal{R}$ using (12);
- Deconditioning $m_{Ag}^\Omega[\{R\}]$ on the same product space $\Omega \times \mathcal{R}$ using (12) as well;
- Combining them using the CRC (6);
- Marginalizing the result on Ω using (8).

Formally:

$$m_{Ag}^\Omega[m_S^\Omega, m_{Ag}^\mathcal{R}] = \left(\odot_{\ell=1}^L m_{Ag}^\mathcal{R}[\theta_\ell]^{\uparrow\Omega \times \mathcal{R}} \odot m_{Ag}^\Omega[\{R\}]^{\uparrow\Omega \times \mathcal{R}} \right)^{\downarrow\Omega}. \quad (19)$$

As shown in [12], the resulting BBA m_{Ag}^Ω , only depends on m_S and on the vector $\alpha = (\alpha_1, \dots, \alpha_L)$ of discount rates. It is then denoted by $\mathfrak{G}m$.

Proposition 1 ([12, Proposition 8]). *Contextual discounting $\mathfrak{G}m$ of a BBA m is equal to the disjunctive combination of m with a BBA m_0^Ω such that:*

$$m_0^\Omega = m_1^\Omega \odot m_2^\Omega \odot \dots \odot m_L^\Omega, \quad (20)$$

where each m_ℓ^Ω , $\ell \in \{1, \dots, L\}$, is defined by:

$$m_\ell^\Omega(A) = \begin{cases} \beta_\ell & \text{if } A = \emptyset \\ \alpha_\ell & \text{if } A = \theta_\ell \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

Remark 1. *Two special cases of this discounting operation can be considered.*

- If $\Theta = \{\Omega\}$ denotes the trivial partition of Ω in one class, combining m with m_0 defined by (15) is equivalent to combining m with m_0 defined by (20), so this contextual discounting operation is identical to the classical discounting operation.
- If $\Theta = \Omega$, the finest partition of Ω , this discounting is simply called contextual discounting and denoted αm . It is defined by the disjunctive combination of m with the BBA $m_1^\Omega \odot m_2^\Omega \odot \dots \odot m_K^\Omega$, where each m_k^Ω , $k \in \{1, \dots, K\}$ is defined by $m_k^\Omega(\emptyset) = \beta_k$ and $m_k^\Omega(\{\omega_k\}) = \alpha_k$.

2.7. Canonical conjunctive and disjunctive decompositions

In [28], extending the notion of separable BBA introduced by Shafer [23, chapter 4], Smets shows that each non-dogmatic BBA m can be uniquely decomposed into a conjunctive combination of *generalized simple BBAs* (*GS-BBAs*), denoted $A^{w(A)}$ with $A \subset \Omega$, and defined from 2^Ω to $\mathbb{R}$ by:

$$\begin{aligned} A^{w(A)} : \quad \Omega &\mapsto w(A) \\ A &\mapsto 1 - w(A) \\ B &\mapsto 0, \forall B \in 2^\Omega \setminus \{A, \Omega\}, \end{aligned} \quad (22)$$

with $w(A) \in (0, \infty)$.

The function $w: 2^\Omega \setminus \{\Omega\} \rightarrow (0, \infty)$ is yet another representation of a non-dogmatic mass function and is called the *conjunctive weight function*.

Let us note that the higher is the weight $w(A)$, the higher is the incertitude on A .

The canonical conjunctive decomposition of a non-dogmatic BBA m is then given by:

$$m = \bigodot_{A \subset \Omega} A^{w(A)}. \quad (23)$$

In [4], Denceux introduces another decomposition: the canonical disjunctive decomposition of a subnormal BBA into *negative GSBBAs* (*NGSBBAs*), denoted $A_{v(A)}$ with $A \supset \emptyset$, and defined from 2^Ω to $\mathbb{R}$ by:

$$\begin{aligned} A_{v(A)} : \quad \emptyset &\mapsto v(A) \\ A &\mapsto 1 - v(A) \\ B &\mapsto 0, \forall B \in 2^\Omega \setminus \{\emptyset, A\}, \end{aligned} \quad (24)$$

with $v(A) \in (0, \infty)$.

Every subnormal BBA m can be canonically decomposed into a disjunctive combination of NGSBBAs:

$$m = \biguplus_{A \supset \emptyset} A_{v(A)}. \quad (25)$$

Indeed, as remarked in [4], the negation of a BBA m can also be conjunctively decomposed as soon as m is subnormal:

$$\begin{aligned} \bar{m} &= \bigodot_{A \subset \Omega} A^{\bar{w}(A)} \text{ (as } \bar{m} \text{ is non-dogmatic)} \\ \Rightarrow m &= \bigodot_{A \subset \Omega} A^{\bar{w}(A)} = \biguplus_{A \subset \Omega} A^{\bar{w}(A)} = \biguplus_{A \supset \emptyset} A_{\bar{w}(\bar{A})}. \end{aligned} \quad (26)$$

The relation between functions v and w is then $v(A) = \bar{w}(\bar{A})$ for all $A \neq \emptyset$, and function $v: 2^\Omega \setminus \{\emptyset\} \rightarrow (0, \infty)$, called the *disjunctive weight function*, is another representation of a subnormal mass function.

Let us remark that behind the Equation (26) we retrieve the fact that the disjunctive and conjunctive rules of combination are linked by De Morgan's laws [5]:

$$\begin{aligned}\overline{m_1 \odot m_2} &= \overline{m_1} \odot \overline{m_2}, \\ \overline{m_1 \odot m_2} &= \overline{m_1} \odot \overline{m_2},\end{aligned}\tag{27}$$

for all BBAs m_1 and m_2 .

Practically, functions w and v have the following properties [4]:

- for all w :

$$\prod_{A \subset \Omega} w(A) = m(\Omega), \tag{28}$$

- for all v :

$$\prod_{A \supset \emptyset} v(A) = m(\emptyset), \tag{29}$$

- for all subset $A \subset \Omega$: $A^{w_1(A)} \odot A^{w_2(A)} = A^{w_1(A)w_2(A)}$,
- for all subset $A \supset \emptyset$: $A_{v_1(A)} \odot A_{v_2(A)} = A_{v_1(A)v_2(A)}$,
- function w can be conveniently obtained from the commonality function q as follows:

$$\forall A \subset \Omega, \quad w(A) = \frac{\prod_{B \supseteq A, \text{Parity}(|A|) \neq \text{Parity}(|B|)} q(B)}{\prod_{B \supseteq A, \text{Parity}(|A|) = \text{Parity}(|B|)} q(B)}, \tag{30}$$

where $\text{Parity}(n)$ means the parity of an integer n ($\text{Parity}(n) = 0$ if n is even, 1 otherwise),

- likewise, function v can be computed from the implicability function b as follows:

$$\forall A \supset \emptyset, \quad v(A) = \frac{\prod_{B \subseteq A, \text{Parity}(|A|) \neq \text{Parity}(|B|)} b(B)}{\prod_{B \subseteq A, \text{Parity}(|A|) = \text{Parity}(|B|)} b(B)}, \tag{31}$$

For “quasi-Bayesian” BBAs, another convenient way to compute w is given by the following property.

Proposition 2 ([4, Proposition 1]). *Let m be a BBA which focal sets are $\Omega, A_1, A_2, \dots, A_n$, and possibly $\emptyset$, such that the n subsets A_k verifies*

$A_i \cap A_j = \emptyset$ for all $i, j \in \{1, \dots, n\}$. The conjunctive weight function w associated with m is then defined by:

$$w(A) = \begin{cases} \frac{m(\Omega)}{m(A_k) + m(\Omega)} & \text{if } A = A_k, \\ m(\Omega) \prod_{k=1}^n (1 + \frac{m(A_k)}{m(\Omega)}) & \text{if } A = \emptyset, \\ 1 & \text{otherwise.} \end{cases} \quad (32)$$

Remark 2. If m has only one focal element in addition to $\emptyset$ and Ω , Proposition 2 holds as well.

The dual version of this property to compute disjunctive weights can be obtained as follows.

Proposition 3. Let m be a BBA which focal sets are $\emptyset, A_1, A_2, \dots, A_n$, and possibly Ω , such that the n subsets A_k verifies $A_i \cup A_j = \Omega$ for all $i, j \in \{1, \dots, n\}$. The disjunctive weight function v associated with m is then defined by:

$$v(A) = \begin{cases} \frac{m(\emptyset)}{m(A_k) + m(\emptyset)} & \text{if } A = A_k, \\ m(\emptyset) \prod_{k=1}^n (1 + \frac{m(A_k)}{m(\emptyset)}) & \text{if } A = \Omega, \\ 1 & \text{otherwise.} \end{cases} \quad (33)$$

Proof 1. Focal sets of $\bar{m}$ are $\Omega, \overline{A_1}, \overline{A_2}, \dots, \overline{A_n}$, and possibly $\emptyset$, such that $\overline{A_i} \cap \overline{A_j} = \emptyset$ for all $i, j \in \{1, \dots, n\}$. Consequently, from Proposition 2, the conjunctive weight function $\bar{w}$ associated with $\bar{m}$ is given by:

$$\bar{w}(A) = \begin{cases} \frac{\bar{m}(\Omega)}{\bar{m}(\overline{A_k}) + \bar{m}(\Omega)} & \text{if } A = \overline{A_k}, \\ \bar{m}(\Omega) \prod_{k=1}^n (1 + \frac{\bar{m}(\overline{A_k})}{\bar{m}(\Omega)}) & \text{if } A = \emptyset, \\ 1 & \text{otherwise.} \end{cases} \quad (34)$$

Then:

$$v(A) = \bar{w}(\bar{A}) = \begin{cases} \frac{\bar{m}(\Omega)}{\bar{m}(\overline{A_k}) + \bar{m}(\Omega)} & \text{if } \bar{A} = \overline{A_k}, \\ \bar{m}(\Omega) \prod_{k=1}^n (1 + \frac{\bar{m}(\overline{A_k})}{\bar{m}(\Omega)}) & \text{if } \bar{A} = \emptyset, \\ 1 & \text{otherwise,} \end{cases} \quad (35)$$

and, as $\bar{m}(A) = m(\bar{A}) \forall A$, Equation (33) is obtained.

□

Remark 3. Proof 1 also implies, from Remark 2, that if m has only one focal element in addition to $\emptyset$ and Ω , Equation (33) is still valid.

3. Extending the contextual discounting

In this section, the contextual discounting operation on a coarsening is extended to any subsets of the frame of discernment, and a general formulation linked with the canonical disjunctive decomposition of a BBA is developed.

According to the previous definitions (22) and (24), BBAs m_ℓ , $\ell \in \{1, \dots, L\}$, defined in (21) by $m_\ell(\emptyset) = \beta_\ell$ and $m_\ell(\theta_\ell) = \alpha_\ell$, can be denoted $\theta_{\ell\beta_\ell}$ or θ_{β_ℓ} in a simple way.

From (20) and (25), the contextual discounting on a coarsening $\Theta = \{\theta_1, \dots, \theta_L\}$ of Ω of a subnormal BBA m is thus defined by:

$$\begin{aligned}\alpha m &= m \circledcirc \theta_{\beta_1} \circledcirc \dots \circledcirc \theta_{\beta_L} \\ &= \bigoplus_{A \supset \emptyset} A_{v(A)} \circledcirc \theta_{\beta_1} \circledcirc \dots \circledcirc \theta_{\beta_L} .\end{aligned}$$

In particular:

- The classical discounting (15) of a subnormal BBA $m = \bigoplus_{A \supset \emptyset} A_{v(A)}$ is defined by:

$$\alpha m = \Omega_{\beta v(\Omega)} \bigoplus_{\Omega \supset A \supset \emptyset} A_{v(A)} ; \quad (36)$$

- The contextual discounting (Remark 1) of a subnormal BBA $m = \bigoplus_{A \supset \emptyset} A_{v(A)}$ is defined by:

$$\alpha m = \bigoplus_{\omega_k \in \Omega} \{\omega_k\}_{\beta_k v(\{\omega_k\})} \bigoplus_{A \subset \Omega, |A| > 1} A_{v(A)} . \quad (37)$$

These discounting operations can then be viewed as particular cases of a more general correction mechanism defined by:

$$\alpha^\cup m = \bigoplus_{A \supset \emptyset} A_{\beta_A v(A)}, \quad (38)$$

where $\beta_A \in [0, 1]$ for all $A \neq \emptyset$ and α is the vector $\{\alpha_A\}_{A \neq \emptyset}$.

In [12], the interpretation of each β_A has been given only in the case where the union of the subsets A forms a partition of Ω , β_A being interpreted as the degree of belief held by the agent regarding the fact that the source is reliable, knowing that the value searched belongs to A .

Instead of considering (18), let us now suppose that agent Ag holds beliefs regarding the reliability of the source, conditionally on each subset A of Ω :

$$\begin{cases} m_{Ag}^{\mathcal{R}}[A](\{R\}) &= 1 - \alpha_A = \beta_A \\ m_{Ag}^{\mathcal{R}}[A](\mathcal{R}) &= \alpha_A , \end{cases} \quad (39)$$

where $\alpha_A \in [0, 1]$.

In the same way as in Section 2.6, the knowledge held by agent Ag , based on the information m_S^Ω from a source and on metaknowledge $m_{Ag}^\mathcal{R}$ (39) regarding the reliability of this source, can be computed as follows:

$$m_{Ag}^\Omega[m_S^\Omega, m_{Ag}^\mathcal{R}] = \left(\odot_{A \subseteq \Omega} m_{Ag}^\mathcal{R}[A]^{\uparrow \Omega \times \mathcal{R}} \odot m_{Ag}^\Omega[\{R\}]^{\uparrow \Omega \times \mathcal{R}} \right)^{\downarrow \Omega}. \quad (40)$$

Proposition 4. *The BBA m_{Ag}^Ω resulting from (40) only depends on m_S^Ω and the vector $\alpha = \{\alpha_A\}_{A \subseteq \Omega}$. It is equal to the disjunctive combination of m_S^Ω with a BBA m_0^Ω defined by:*

$$m_0^\Omega(C) = \prod_{\cup A=C} \alpha_A \prod_{\cup B=\overline{C}} \beta_B, \quad \forall C \subseteq \Omega. \quad (41)$$

Proof 2. *See Appendix A.1. $\square$*

Like in the case of contextual discounting operations considered in Section 2.6, BBA m_0^Ω defined in Proposition 4 admits a simple decomposition described in the following proposition.

Proposition 5. *The BBA m_0^Ω defined in Proposition 4 can be rewritten as:*

$$m_0^\Omega = \odot_{A \supset \emptyset} A_{\beta_A}. \quad (42)$$

Proof 3. *Directly from (41) and the definition (7) of the DRC. $\square$*

From (42), the contextual discounting resulting from (40) of a subnormal BBA $m = \odot_{A \supset \emptyset} A_{v(A)}$ is then defined by:

$$\odot_{A \supset \emptyset} A_{v(A)} \odot_{A \supset \emptyset} A_{\beta_A} = \odot_{A \supset \emptyset} A_{\beta_A v(A)} = \alpha^\cup m. \quad (43)$$

Correction mechanism $\alpha^\cup m$ is then the general formulation for a contextual discounting on any subsets of Ω .

The following theorem sums up the contextual discounting operation in its general formulation.

Theorem 1. *On the one hand, agent Ag receives an information m from a source S .*

On the other hand, agent Ag knows that this source is reliable with a degree β_A in different contexts A of Ω , which means that the source is reliable with a degree β_A knowing that the true answer to the question Q of interest belongs to A . Let us note $\mathcal{A}$ the set containing these contexts.

Then, agent Ag 's mass function is given by the contextual discounting αm of m defined by:

$$\alpha m = m \bigoplus_{A \in \mathcal{A}} A_{\beta_A} . \quad (44)$$

Moreover, if m is subnormal then:

$$\begin{aligned} \alpha m &= \bigoplus_{A \supset \emptyset} A_{v(A)} \bigoplus_{A \in \mathcal{A}} A_{\beta_A} , \\ &= \bigoplus_{A \notin \mathcal{A}} A_{v(A)} \bigoplus_{A \in \mathcal{A}} A_{\beta_A v(A)} . \end{aligned} \quad (45)$$

Example 1. Let us consider a two-lane road section, the question of interest concerning the number of lanes where the traffic is flowing freely. Frame of discernment Ω is equal to $\{\omega_0, \omega_1, \omega_2\}$ with:

- ω_0 standing for “0 lane is free”: both are blocked,
- ω_1 meaning “1 lane is free”: one is blocked,
- ω_2 signifying “2 lanes are free”: traffic is flowing freely on both lanes.

A source S provides information regarding the traffic on this road section. For instance:

$$\begin{cases} m(\emptyset) &= .1 , \\ m(\{\omega_0\}) &= .8 , \\ m(\Omega) &= .1 . \end{cases} \quad (46)$$

On the other side, you know that the source is very reliable in case of heavy traffic (situation $\{\omega_0, \omega_1\}$), and less reliable when the traffic is rather light (situation $\{\omega_1, \omega_2\}$). Formally, let us suppose that $\beta_{\{\omega_0, \omega_1\}} = .8$ and $\beta_{\{\omega_1, \omega_2\}} = .6$.

Contextual discounting of m is then given by:

$$\alpha m = m \odot \{\omega_0, \omega_1\}.8 \odot \{\omega_1, \omega_2\}.6 . \quad (47)$$

From Remark 3, the disjunctive weight function v associated with m can be computed in the following manner:

$$\begin{cases} v(\{\omega_0\}) &= \frac{m(\emptyset)}{m(\emptyset) + m(\{\omega_0\})} = \frac{1}{9} , \\ v(\Omega) &= m(\emptyset) \left(1 + \frac{m(\{\omega_0\})}{m(\emptyset)}\right) = .1 \left(1 + \frac{.8}{.1}\right) = .9 . \end{cases} \quad (48)$$

Contextual discounting of m knowing $\beta_{\{\omega_0, \omega_1\}} = .8$ and $\beta_{\{\omega_1, \omega_2\}} = .6$ is then given by:

$$\begin{aligned} \alpha m &= m \odot \{\omega_0, \omega_1\}.8 \odot \{\omega_1, \omega_2\}.6 \\ &= \{\omega_0\}_{\frac{1}{9}} \odot \Omega_{.9} \odot \{\omega_0, \omega_1\}.8 \odot \{\omega_1, \omega_2\}.6 , \end{aligned} \quad (49)$$

which can be also written by definition from (24) in the following manner:

$$\begin{aligned}\alpha_m &= \left\{ \begin{array}{ll} \emptyset & \mapsto 1/9 \\ \{\omega_0\} & \mapsto 8/9 \end{array} \right\} \oplus \left\{ \begin{array}{ll} \emptyset & \mapsto .9 \\ \Omega & \mapsto .1 \end{array} \right\} \oplus \left\{ \begin{array}{ll} \emptyset & \mapsto .8 \\ \{\omega_0, \omega_1\} & \mapsto .2 \end{array} \right\} \oplus \left\{ \begin{array}{ll} \emptyset & \mapsto .6 \\ \{\omega_1, \omega_2\} & \mapsto .4 \end{array} \right\} \\ &= \left\{ \begin{array}{ll} \emptyset & \mapsto .05 \\ \{\omega_0\} & \mapsto .38 \\ \{\omega_0, \omega_1\} & \mapsto .11 \\ \{\omega_1, \omega_2\} & \mapsto .03 \\ \Omega & \mapsto .43 \end{array} \right. \quad (50)\end{aligned}$$

4. Dual version: a new reinforcement process

In a similar way, a correction mechanism for a non-dogmatic BBA m can be defined from the conjunctive decomposition of m as follows:

$$\alpha_{\cap} m = \bigodot_{A \subset \Omega} A^{\beta_A w(A)} ; \quad (51)$$

where $\forall A \subset \Omega, \beta_A \in [0, 1]$, and α is the vector $\{\alpha_A\}_{A \subset \Omega}$.

The smaller is the conjunctive weight, the higher is the mass on A . This process allows then the reinforcement of a BBA m .

Correction mechanisms $\alpha_{\cap} m$ (38) and $\alpha_{\cup} m$ (51) are related in the following way.

Let us consider a subnormal BBA m , $\bar{m}$ is then non-dogmatic:

$$\alpha_{\cap} \bar{m} = \bigodot_{A \subset \Omega} A^{\beta_A \bar{w}(A)} . \quad (52)$$

Then:

$$\begin{aligned}\overline{\alpha_{\cap} \bar{m}} &= \overline{\bigodot_{A \subset \Omega} A^{\beta_A \bar{w}(A)}} \\ &= \bigoplus_{A \subset \Omega} A^{\beta_A \bar{w}(A)} \\ &= \bigoplus_{A \supset \emptyset} A_{\beta_A \bar{w}(\bar{A})} \\ &= \bigoplus_{A \supset \emptyset} A_{\beta_A v(A)} \\ &= \alpha_{\cup} m\end{aligned} \quad (53)$$

These two correction mechanisms can thus be seen as belonging to a general family of correction mechanisms.

In a nutshell, if m is subnormal, a contextual discounting of m is given by:

$$\alpha m = \bigoplus_{A \supset \emptyset} A_{\beta_A v(A)} , \quad (54)$$

and, the negation of a contextual discounting of $\bar{m}$ defines a *dual reinforcement process*:

$$\overline{\alpha \bar{m}} = \overline{\bigoplus_{A \supset \emptyset} A_{\beta_A \bar{v}(A)}} = \bigodot_{A \subset \Omega} A^{\beta_A w(A)} . \quad (55)$$

The application of this reinforcement process as well as its comparison with other correction mechanisms [20, 14] has been left for future researches.

5. Discussions

5.1. About the mass transfer during a contextual discounting

As it can be observed in particular in Example 1, the contextual discounting allows the transfer of masses on intermediate subsets between focal sets and the frame of discernment Ω . More precisely (cf Equation (44)), contextual discounting consists in transferring each mass to its union with subsets not precisely known by the source.

The following example illustrates a case where a sensor is totally reliable if an object is of a certain type, and not reliable for another type.

Example 2. *Let us consider a sensor in charge of the recognition of two types of objects: $\Omega = \{a, b\}$.*

The sensor knows very well how to recognize objects of type a, whereas it is not the case for objects of type b. Which means that:

- *if an object is of type a, the sensor will recognize it;*
- *if an object is of type b, the sensor will hesitate and make mistakes.*

For instance, a possible confusion matrix for such a sensor is represented in Table 1.

Table 1: Confusion matrix associated with source S.

Decision	a	b
Truth		
a	10	0
b	5	5

A contextual discounting knowing that $\beta_{\{a\}} = 1$ and $\beta_{\{b\}} = 0$ (S totally reliable in context $\{a\}$, not reliable in context $\{b\}$) applied on a mass function m provided by S is given by:

$$\alpha m = m \odot \{a\}_1 \odot \{b\}_0 = m \odot \{b\}_0 . \quad (56)$$

In particular:

- *if $m(\{a\}) = 1$ then $\alpha m(\{a, b\}) = 1$,*
- *if $m(\{b\}) = 1$ then $\alpha m(\{b\}) = 1$.*

In other words, if the source says it is an object of type a : it is an object of type a or an object the source does not recognize. And, in the particular case where the source says it is an object of type b , and the source knows very well the other types of object: it remains that the object is of type b . If it had been an object of type a , the source would have said it, because objects of type a are very well recognized by the source.

5.2. On the notion of reliability

In the simple Example 2, we have a situation where when the source decides b : it is indeed b , and when the source decides a : the truth is a or b . However, the source is totally reliable for a and not for b . The notion of reliability present in this article has then to be clearly distinguished from a different notion of reliability which would be linked with the reliability of the decision made by a source.

In the contextual discounting, the definition of the reliability (39) is given conditionally on Ω by:

$$m_{Ag}^{\mathcal{R}}[A](\{R\}) = \beta_A , \quad (57)$$

and not conditionally on results of a decision-making process:

$$m_{Ag}^{\mathcal{R}}[\text{"The source decides A"}](\{R\}) = \beta_A . \quad (58)$$

Let us remark that this second definition will imply to define the notion of conditioning on processes (even belief functions?) which is not known at present by the authors.

5.3. Discounting vs combination of discountings

Let us consider a separable BBA m , which means that $m = \bigodot_{A \subset \Omega} A^{w(A)}$ with $w(A) \in [0, 1]$ for all $A \subset \Omega$.

As recently exposed in [9], a classical discounting on each simple BBA $A^{w(A)}$ can be undertaken with a discount rate $\alpha_A \in [0, 1]$, the result being $A^{\beta_A w(A) + \alpha_A}$.

The discounted simple BBAs can then be conjunctively combined which yields to the following discounting operation:

$$\bigodot_{A \subset \Omega} A^{\beta_A w(A) + \alpha_A} . \quad (59)$$

This operation, restricted to separable BBAs, is different from a contextual operation. The conjunctive combination of discounted BBAs is not in general a discounting of the BBAs combination.

However, we can remark that the dual of the contextual discounting operation (51), which is nevertheless a reinforcement, has a close formulation.

5.4. To be subnormal and non-dogmatic

In order to exploit simple expressions obtained with both conjunctive and disjunctive canonical decompositions, should each mass function be subnormal and non-dogmatic?

As already mentioned in [4], it may be argued that most (if not all) pieces of information provided in real life applications are imperfect, and then the mass on the frame of discernment should be always strictly positive.

For instance, let us consider a coin tossing and a universe equal to $\{heads, tails\}$. As remarked by Denceux, it is absolutely not certain in practice that the coin is perfectly balanced. An appropriate BBA may then be $m(\{heads\}) = 0.5(1 - \epsilon)$, $m(\{tails\}) = 0.5(1 - \epsilon)$ and $m(\Omega) = \epsilon$ for some small $\epsilon > 0$.

However, we can also add that when we are trying to model a real life problem in all its complexity, some approximations and some unforeseeable situations may occur, and then some doubts on the model are possible, so the mass on the conflict should also be strictly positive: the coin can fall against a book and lands on edge, someone may also catch the coin in the air, it can likewise break when falling on the ground, ...

Consequently, a more appropriate BBA may be:

$$\begin{cases} m(\emptyset) &= \epsilon_2 , \\ m(\{heads\}) &= 0.5(1 - \epsilon_1)(1 - \epsilon_2) , \\ m(\{tails\}) &= 0.5(1 - \epsilon_1)(1 - \epsilon_2) , \\ m(\Omega) &= \epsilon_1(1 - \epsilon_2) , \end{cases} \quad (60)$$

for some small $\epsilon_1, \epsilon_2 > 0$.

Nevertheless, accepting $m(\emptyset) > 0$ from the creation of m seems not as easy as accepting $m(\Omega) > 0$.

The discussion on accepting a strictly positive mass on the empty set has been conducted on several occasions in Smets' articles on the TBM, in particular in [32]. The mass on the empty set plays an alarm role, it means "there is problem somewhere". Several main causes not necessarily exclusive have been mentioned:

1. the non-exclusivity of the frame of discernment: as previously exposed, the coin can fall neither heads nor tails;
2. one assumption for the application of a conjunctive combination is not respected: one of the source is not reliable, sources are not distinct, or sources do not refer to the same objects.

However in the present article, a strictly positive mass on the empty set is accepted from the creation of a BBA, which seems harder to accept as it remains only the non-exclusivity argument to tolerate $m(\emptyset) > 0$.

So what else if this cause is excluded: the coin is tossed in a particular room where the coin can not fall on edge, no one can catch it, the coin is unbreakable, ... In fact, it can be argued that it is impossible to predict all the situations and to cover the whole complexity of a real situation. By definition, something unpredictable may always occur, and it always exists some alternatives not considered [32].

Besides, as discussed in [32, Section 3.5], trying to close the world [21, 22, 33] by adding an extra element covering all the possible alternatives not initially considered in Ω , seems not a solution as it involves among other things problems of mass allocation, and questions on the cardinality of this new element in particular when making a decision.

More generally, mass $m(\emptyset)$ reflects the appropriateness, the accuracy of the model with the reality. So putting a mass on the empty set strictly equal to zero (not a thousandth nor a millionth, really zero) seems very delicate as by definition real situations are associated with unforeseeable problems.

At last, just as Dencœux proposed a slight discounting to obtain a non-dogmatic BBA from a dogmatic one, a subnormal BBA can result from a normal one by a slight reinforcement. Small enough to be considered as negligible, the additions of masses on Ω and $\emptyset$ will not change anything on the modeling but will add a huge flexibility by the use of conjunctive and disjunctive decompositions.

6. Conclusion and future work

In this article, the contextual discounting operation of a belief function has been extended to any subsets, and a simple and practical expression, based on disjunctive combinations, to compute it has been given. This expression has highlighted the close relationship between contextual discounting and canonical disjunctive decomposition. The dual expression of this discounting, allowing one to strengthen a belief function, has also been exposed.

Future work will aim at testing it on real data. Likewise, it would also be interesting to automatically learn the coefficients of these correction mechanisms from data, as done for the classical and the contextual discounting operations [6, 12].

Acknowledgements

The authors are very grateful to the anonymous reviewers for their valuable comments which helped to improve the clarity (and the accuracy) of this article, and which have allowed the authors to develop their arguments.

This work has been financed by the French région Nord-Pas de Calais under the project CISIT (Campus International pour la Sécurité et l'Intermodalité des Transports).

Appendix A. Proofs

Appendix A.1. Proofs of Proposition 4

For each $A \subseteq \Omega$, the deconditioning of $m_{Ag}^{\mathcal{R}}[A]$ on $\Omega \times \mathcal{R}$ is given by:

$$m_{Ag}^{\mathcal{R}}[A]^{\uparrow\Omega \times \mathcal{R}}(A \times \{R\} \cup \bar{A} \times \mathcal{R}) = \beta_A, \quad (\text{A.1})$$

$$m_{Ag}^{\mathcal{R}}[A]^{\uparrow\Omega \times \mathcal{R}}(\Omega \times \mathcal{R}) = \alpha_A. \quad (\text{A.2})$$

With $A \neq B$:

$$\begin{aligned} (A \times \{R\} \cup \bar{A} \times \mathcal{R}) \cap (B \times \{R\} \cup \bar{B} \times \mathcal{R}) \\ = (A \cup B) \times \{R\} \cup \overline{(A \cup B)} \times \mathcal{R}. \end{aligned}$$

Then:

$$\begin{aligned} \odot_{A \subseteq \Omega} m_{Ag}^{\mathcal{R}}[A]^{\uparrow\Omega \times \mathcal{R}}(C \times \{R\} \cup \bar{C} \times \mathcal{R}) \\ = \prod_{\cup D = \bar{C}} \alpha_D \prod_{\cup E = C} \beta_E, \quad \forall C \subseteq \Omega, \end{aligned}$$

or, by exchanging the roles of C and $\bar{C}$:

$$\begin{aligned} \odot_{A \subseteq \Omega} m_{Ag}^{\mathcal{R}}[A]^{\uparrow\Omega \times \mathcal{R}}(\bar{C} \times \{R\} \cup C \times \mathcal{R}) \\ = \prod_{\cup D = C} \alpha_D \prod_{\cup E = \bar{C}} \beta_E, \quad \forall C \subseteq \Omega. \end{aligned}$$

It remains to combine conjunctively $m_{Ag}^{\Omega}[\{R\}]^{\uparrow\Omega \times \mathcal{R}}$ and $\odot_{A \subseteq \Omega} m_{Ag}^{\mathcal{R}}[A]^{\uparrow\Omega \times \mathcal{R}}$ which have focal sets of the form $B \times \{R\} \cup \Omega \times \{NR\}$ and $\bar{C} \times \{R\} \cup C \times \mathcal{R}$, respectively, with $B, C \subseteq \Omega$. The intersection of two such focal sets is:

$$\begin{aligned} (\bar{C} \times \{R\} \cup C \times \mathcal{R}) \cap (B \times \{R\} \cup \Omega \times \{NR\}) \\ = B \times \{R\} \cup C \times \{NR\}, \end{aligned}$$

and it can be obtained only for a particular choice of B and C . Then:

$$\begin{aligned} \odot_{A \subseteq \Omega} m_{Ag}^{\mathcal{R}}[A]^{\uparrow \Omega \times \mathcal{R}} \odot m_{Ag}^{\Omega}[\{R\}]^{\uparrow \Omega \times \mathcal{R}}(B \times \{R\} \cup C \times \{NR\}) \\ = \left[\prod_{\cup D=C} \alpha_D \prod_{\cup E=\overline{C}} \beta_E \right] m_S^{\Omega}(B) . \end{aligned} \quad (\text{A.3})$$

Finally, the marginalization of this BBA on Ω is given for all subsets A of Ω , by:

$$\alpha_m(A) = \sum_{B \cup C=A} \left[\prod_{\cup D=C} \alpha_D \prod_{\cup E=\overline{C}} \beta_E \right] m_S^{\Omega}(B) . \quad (\text{A.4})$$

Let us note that the above proof has many similarities with proofs presented in [12, Sections A.1 and A.3].

References

- [1] I. Bloch. Defining belief functions using mathematical morphology - Applications to image fusion under imprecision. *International Journal of Approximate Reasoning*, volume 48, pages 437–465, 2008.
- [2] F. Delmotte and G. Gacquer. Detection of defective sources with belief functions. In *Proceedings of the 12th Information Processing and Management of Uncertainty in Knowledge-Based Systems International Conference (IPMU'08)*, L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (Ed.), pages 337–344, Torremolinos (Malaga), Spain, June 22–27, 2008.
- [3] A. Dempster. Upper and Lower Probabilities Induced by Multivalued Mapping. *Annals of Mathematical Statistics*, volume AMS-38, pages 325–339, 1967.
- [4] T. Denœux. Conjunctive and Disjunctive Combination of Belief Functions Induced by Non Distinct Bodies of Evidence. *Artificial Intelligence*, volume 172, pages 234–264, 2008.
- [5] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, volume 12, pages 193–226, 1986.
- [6] Z. Elouedi, K. Mellouli and Ph. Smets. Assessing sensor reliability for multisensor data fusion with the transferable belief model. *IEEE Transactions on Systems, Man and Cybernetics B*, volume 34, pages 782–787, 2004.

- [7] S. Fabre, A. Appriou and X. Briottet. Presentation and description of two classification methods using data fusion based on sensor management. *Information Fusion*, volume 2, pages 49–71, 2001.
- [8] M. Ha-Duong, Hierarchical fusion of expert opinions in the Transferable Belief Model, application to climate sensitivity, *International Journal of Approximate Reasoning*, volume 49, issue 3, pages 555-574, November 2008.
- [9] A. Kallel and S. Le Hégarat-Masclé, Combination of partially non-distinct beliefs: the cautious adaptive rule, *International Journal of Approximate Reasoning*, volume 50, issue 7, pages 1000–1021, 2009.
- [10] J. Klein, C. Lecomte and P. Mich, Hierarchical and conditional combination of belief functions induced by visual tracking, *International Journal of Approximate Reasoning*, volume 51, issue 4, pages 410-428, March 2010.
- [11] D. Mercier, T. Denœux and M.-H. Masson. A parameterized family of belief functions correction mechanisms. In *Proceedings of the 12th Information Processing and Management of Uncertainty in Knowledge-Based Systems International Conference, IPMU'08*, L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (Ed.), pages 306-313, Torremolinos (Malaga), Spain, June 22-27, 2008.
- [12] D. Mercier, B. Quost and T. Denœux. Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, volume 9, pages 246–258, 2008.
- [13] D. Mercier, G. Cron, T. Denœux and M.-H. Masson. Decision fusion for postal address recognition using belief functions. *Expert Systems with Applications*, volume 36, issue 3, part 1, pages 5643–5653, 2009.
- [14] D. Mercier, T. Denœux and M.-H. Masson. Belief function correction mechanisms. *Studies in Fuzziness and Soft Computing*, B. Bouchon-Meunier et al. (Eds.), volume 249, pages 203–222, January 2010.
- [15] D. Mercier. Extending the contextual discounting of a belief function thanks to its canonical disjunctive decomposition. *Proceedings of the 1st Workshop on Belief Functions*, paper 61, Brest, France, April 1-2, 2010.

- [16] N. Milisavljević, I. Bloch, S. van den Broek and M. Acheroy. Improving mine recognition through processing and Dempster-Shafer fusion of ground-penetrating radar data, *Pattern Recognition*, volume 36, pages 1233–1250, 2003.
- [17] R. Muñoz-Salinas, R. Medina-Carnicer, F.J. Madrid-Cuevas and A. Carmona-Poyato, Multi-camera people tracking using evidential filters, *International Journal of Approximate Reasoning*, volume 50, issue 5, pages 732–749, May 2009.
- [18] F. Périsset, D. Mercier, É. Lefèvre and D. Roger. Robust Diagnostics of Stator Insulation Based on High Frequency Resonances Measurements, *IEEE Transactions on Dielectrics and Electrical Insulation*, volume 16, issue 5, pages 1496–1502, 2009.
- [19] F. Pichon. Belief functions: canonical decompositions and combination rules, PhD Thesis, Université de Technologie de Compiègne, March, 2009.
- [20] F. Pichon, D. Dubois and T. Denœux. Relevance and Truthfulness in Information Fusion Correction and Fusion. *International Journal of Approximate Reasoning*, this issue, 2011.
- [21] M. Rombaut. Decision in multi-obstacle matching process using the theory of belief. In *Proceedings of the 1st International Conference on Advances in Vehicle Control and Safety, AVCS'98*, pages 63–68, Amiens, France, 1998.
- [22] C. Royere, D. Gruyer and V. Cherfaoui. Data association within belief theory. In *Proceedings of the 3rd International Conference on Information Fusion (FUSION'00)*, paper TuD29, Paris, France, July 10-13, 2000.
- [23] G. Shafer. A mathematical theory of evidence. Princeton University Press, Princeton, N.J., 1976.
- [24] P. Smets. Belief functions. *Non Standard Logics for Automated Reasoning*, P. Smets, E. H. Mamdani, D. Dubois, and H. Prade (Ed.), Academic Press, London, pages 253–286, 1988.
- [25] P. Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, volume 9, pages 1–35, 1993.

- [26] P. Smets and R. Kennes. The Transferable Belief Model. *Artificial Intelligence*, volume 66, pages 191–243, 1994.
- [27] P. Smets. What is Dempster-Shafer’s model? In *Advances in the Dempster-Shafer theory of evidence*, R. R. Yager, M. Fedrizzi and J. Kacprzyk (Ed.), Wiley, New-York, pages 5–34, 1994.
- [28] P. Smets. The canonical decomposition of a weighted belief. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI’95)*, Morgan Kaufman (Ed.), San Mateo, California, USA, pages 1896–1901, 1995.
- [29] P. Smets. The Transferable Belief Model for quantified belief representation. In *Handbook of Defeasible reasoning and uncertainty management systems*, D. M. Gabbay and Ph. Smets (Ed.), Kluwer Academic Publishers, Dordrecht, The Netherlands, volume 1, pages 267–301, 1998.
- [30] P. Smets. Data Fusion in the Transferable Belief Model. *Proceedings of the 3rd International Conference on Information Fusion, FUSION’2000*, Paris, France, July 10-13, pages PS21–PS33, 2000.
- [31] P. Smets. The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, volume 31, issue 12, pages 1-30, 2002.
- [32] P. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion*, volume 8, issue 4, pages 387–412, 2007.
- [33] R.R. Yager. Hedging in the combination of evidence. *Journal of Information and Optimization Science*, volume 4, pages 73-81, 1983.

A high-level application using belief functions for exchanging and managing uncertain events on the road in vehicular ad-hoc networks

Mira Bou Farah¹ · David Mercier¹ ·
Éric Lefèvre¹ · François Delmotte¹

Received: date / Accepted: date

Abstract This article introduces a high-level system using belief functions for exchanging and managing imperfect information about events on the road in vehicular ad-hoc networks. The main purpose of this application is to provide the most reliable information for the driver from multiple messages received informing the driver about events on the roads. This system and some variants are tested using a Matlab™ simulator. An implementation with Android smartphones using a Bluetooth technology to exchange the messages is also introduced.

Keywords Vehicular Ad-hoc Networks (VANET) · Events on the road · Imperfect information exchange · Belief functions · Information fusion

1 Introduction

Cars are nowadays by far the most common means of transport. But this popularity leads to safety and environmental problems. Research is conducted in parallel in Europe [11,16,18], in the USA [26] and in Japan [12] to develop Intelligent Transportation Systems (ITS) and to improve safety and driver comfort by in particular looking for the reduction of the number of undesirable events on the roads such as accidents and traffic jams.

Ad-hoc networks are wireless networks capable of organizing themselves without any infrastructure. Instead of communicating through a centralized access point, they are composed of wireless nodes communicating with each other to exchange information. In a mobility context, they are called MANET (Mobile Ad hoc Networks) [5,13,20,1].

¹ Université Lille Nord de France, F-59000 Lille, France
UArtois, LIGI2A, F-62400, Béthune, France
Corresponding Author: D. Mercier
Tel.: +33 (0)3 21 63 23 49
E-mail: david.mercier@univ-artois.fr

Vehicular Ad-hoc Networks (VANET) are MANET networks applied to Inter-Vehicle Communication (IVC) where nodes are highly dynamic. Two main modes of communication are known in IVC: Vehicle to Infrastructure (V2I) and Vehicle to Vehicle (V2V). A V2V communication example is given by the VESPA system (Vehicular Event Sharing with a mobile P2P Architecture) [2,6,8] which allows vehicles to share different types of information such as accidents or emergency brakings. It guides in particular drivers to find a parking place or to get traffic information. Another example is given by TrafficView [17] a road data dissemination system using V2V communication, allowing continuous exchange of position and speed information between vehicles to inform drivers about road situation (traffic, weather, quickest way, ...).

One important component of V2V communication concerns the data transmission protocol, and currently no universal standard has emerged [7]. Communications can be realized using DSRC (Dedicated Short-Range Communications), Wi-Fi (ad-hoc), Bluetooth, 3G or 4G cellular networks, SDARS (Satellite Digital Audio Radio Service), etc.

The present work concerns V2V communications, where information is decentralized and vehicles build their own information. It is focused on the high-level aspect of merging and organizing the messages received by a vehicle. To demonstrate the applicability of the proposed approach to real systems, a simple Bluetooth system is also developed and tested on-line. Extensions of this approach with other technologies, such as with Wi-Fi devices for instance, is beyond the scope of this paper.

In vehicular networks, the environment is very dynamic. Events on the road appears and disappears over time. Vehicles receive a large amount of information often uncertain, redundant or outdated. Then to manage the exchanges of such imperfect information between vehicles, a method is proposed in this article using the belief functions theory [21,22], belief functions constituting a rich and flexible framework to handle imprecise and uncertain information [9, 19].

Early work integrating belief functions has already been developed by Cherfaoui et al. in [4]. It has been extended in [3] where the authors propose a first model able to manage different events on the same segment of road, and a simple strategy for sending messages is exposed. Each vehicle can either send new messages created by the driver regarding events on the road or transfer messages received from other vehicles, each vehicle building its own map of the current situation without delivering to other vehicles its results of combination or deductions.

In this article, this previous model is deeply extended and tested using a MatlabTM simulator. A new mechanism to model the aging of a message is considered and compared with the first approach which is based on a discounting process considering that the information contained in a message decreases in reliability over time. The new one interprets the aging of a message as a reinforcement in the fact that the event has disappeared as do all events over time. Two strategies for sending messages are also studied: a first one where

all messages are kept in vehicles data bases and all messages are exchanged, and a second approach aiming at saving bandwidth and spacing where only one merged message by event is retained in each vehicle. At last, an implementation of this model is realized using Android smartphones and Bluetooth connections to illustrate a simple concrete case of use and to demonstrate that smartphones can manage belief functions calculations.

This article is organized as follows. Basic concepts on belief functions are recalled in Section 2. A credal model for representing and managing imperfect information about events on the road is exposed in Section 3. It is simulated and tested in Section 4. Then, an example of implementation with Android smartphone application using Bluetooth transmissions is presented in Section 5. Finally, Section 6 concludes this paper and gives prospects.

2 Belief Functions: Basic Concepts

In this article, belief functions are handled with Smets' semantic approach: the Transferable Belief Model (TBM) [22,24], where belief functions are interpreted as weighted opinions with no underlying probabilistic measure.

Two levels are distinguished in the TBM:

- the credal level, where available pieces of information are represented by belief functions, and manipulated;
- the pignistic or decision level, where belief functions are transformed into probability measures when a decision has to be made, and the expected utility is maximized.

2.1 Representing Information

Let us consider $\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$, called *frame of discernment*, a finite set of possible answers to a given question Q of interest.

A knowledge regarding the answer to question Q can be quantified by a *basic belief assignment (BBA)*, also called a *mass function*, denoted by m and defined from 2^Ω (the set of all possible subsets of Ω) to $[0, 1]$ such that the sum of all the masses is equal to 1:

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

Each mass $m(A)$, with $A \subseteq \Omega$, represents the part of belief allocated to the fact that the true answer to question Q belongs to A . In particular $m(\Omega)$ represents the degree of ignorance of the source which has provided the information m .

Each subset A of Ω such that $m(A) > 0$ is called a *focal element* of m .

A BBA m is said to be *categorical* if it has only one focal element A , which means that $m(A) = 1$. It is then denoted by m_A . In particular m_Ω represents

the categorical BBA such that $m(\Omega) = 1$ which corresponds to a state of total ignorance for the agent. BBA m_Ω is also called the *vacuous belief function*.

A mass function m having less than two focal elements including Ω is called a *simple mass function* and verifies:

$$\begin{cases} m(A) = 1 - \omega \\ m(\Omega) = \omega \end{cases} \quad (2)$$

with $A \subset \Omega$ and $\omega \in [0, 1]$. It can be conveniently noted A^ω .

2.2 Manipulating Information

2.2.1 Discounting operation

When receiving a piece of information m , an agent may have some doubts regarding the reliability of the source which has provided this BBA. Such a metaknowledge can be taken into account by using the discounting operation [21, page 252] defined by:

$$\begin{cases} {}^\alpha m(A) = (1 - \alpha) m(A), & \forall A \subset \Omega, \\ {}^\alpha m(\Omega) = (1 - \alpha) m(\Omega) + \alpha, \end{cases} \quad (3)$$

or more simply:

$${}^\alpha m = (1 - \alpha) m(A) + \alpha m_\Omega, \quad (4)$$

where $\alpha \in [0, 1]$ is called the discount rate.

Coefficient $\beta = (1 - \alpha)$ represents the degree of reliability regarding the information which have been provided [21, 15].

2.2.2 Reinforcement operation

On the contrary, an agent may want to reinforce a BBA which seems to be too cautious in the sense that it is not specific enough. This operation can be realized using the following correction mechanism [14]:

$${}^\nu m = (1 - \nu) m(A) + \nu m^{tr}, \quad (5)$$

where $\nu \in [0, 1]$ is called the reinforcement rate, and m^{tr} is the mass function m totally reinforced.

Transformation m^{tr} of m corresponds to the mass function expected by the agent when m has to be fully reinforced. Different possible choices are given in [14].

In next Section 3.5 a reinforcement process is introduced with BBA m^{tr} chosen equal to a particular m_{ω_i} with $\omega_i \in \Omega$, which means that according to the reinforcement rate ν the agent has some reason to believe that ω_i is the true value he/she searches.

2.2.3 Conjunctive rule of combination

Two mass functions m_1 and m_2 obtained from distinct and reliable sources, can be combined using the conjunctive rule of combination denoted by $\odot$ and defined by:

$$m_1 \odot m_2(A) = m_1 \odot_2(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C), \quad \forall A \subseteq \Omega. \quad (6)$$

With this combination, masses are transferred to focal elements intersections.

2.2.4 Cautious rule of combination

If mass functions m_1 and m_2 are obtained from non-distinct reliable sources, they can be combined by using the cautious rule of combination [10], denoted by $\oslash$ and defined as follows:

$$m_{1 \oslash 2}^\Omega = \bigodot_{A \subseteq \Omega} A^{w_1^\Omega(A) \wedge w_2^\Omega(A)}, \quad (7)$$

where $\wedge$ denotes the minimum operator, and w is the conjunctive weight function [23] defined by:

$$w(A) = \prod_{A \subseteq B} q(B)^{(-1)^{|B|-|A|+1}}, \quad \forall A \subset \Omega. \quad (8)$$

The cautious rule of combination is commutative, associative and idempotent (which means that: $m \oslash m = m$).

2.3 Making a decision

At this level, the mass function m defined on Ω and representing the available information regarding the answer to question Q , and generally resulting in practice from a fusion process, is transformed into a probability measure with the *pignistic transformation* [25] defined by:

$$BetP^\Omega(\{\omega\}) = \sum_{\{A \subseteq \Omega, \omega \in A\}} \frac{m(A)}{|A| (1 - m(\emptyset))}, \quad \forall \omega \in \Omega. \quad (9)$$

3 A credal model for representing and managing imperfect information about events

In this section, messages format and the different parts composing the proposed credal model allowing drivers to have a more reliable view of situations on the road based on exchanged messages are exposed.

Vehicles are assumed equipped with a Global Positioning System (GPS). Locations and dates of events are assumed to be automatically generated as soon as drivers need to send messages regarding events.

3.1 Exchanged messages format

Vehicles exchange messages about events on the road, each message giving information regarding one event.

A message M is represented as a 5-tuple (S, t, d, ℓ, m) described in Table 1.

Table 1 Messages format

Attribute	Description
S	Source which has perceived the event
t	Type of the event
d	Date and Time when S has detected the event
ℓ	Location where S has detected the event
m	Mass function representing the confidence of S regarding the fact that the event is present

Each attribute $x \in \{S, t, d, \ell, m\}$ of a message M will be denoted by $M.x$:

- $M.S$ designates the source of information which have perceived the event. It is not necessarily the source which have transferred the message M .
- $M.t$ indicates the type t of the event reported by message M . Different types of events can be considered such as accident, parking place, traffic-jam, fog blanket, animals on the road, etc.
- $M.d$ points out the date when $M.S$ has created the message to inform about its presence. It is generally not the date at which message M has been received.
- $M.\ell$ indicates the location ℓ of the event reported by message M .
- At last, $M.m$ denotes the mass function m , held by vehicle source S , which is defined on the frame of discernment $\Omega = \{\exists, \bar{A}\}$ where:
 - $\exists$ stands for "the event, which is of type t , is present at time d at location ℓ ."
 - and $\bar{A}$ means "the event, which is of type t , is not present at time d at location ℓ ."

3.2 Sending process

Each vehicle manages an inner database composed of the messages it has either received or created.

All the messages M_i^j concerning an event j are regrouped into a table denoted by M^j .

Each time a vehicle can exchange messages with another vehicle, all the messages present in its database are exchanged.

3.3 Reception process

The reception of a message by a vehicle is realized sequentially by the procedure illustrated in Figure 1.

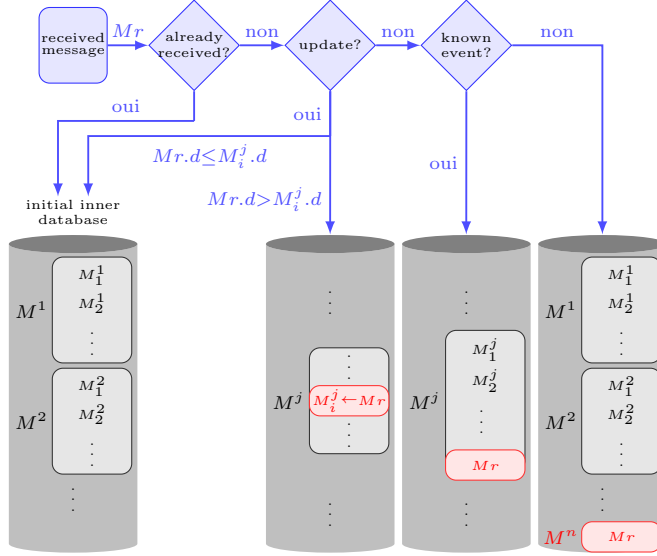


Fig. 1 Messages reception process.

Case of a message already received If the received message has been already received, it is ignored. Otherwise the message goes to the next test.

Case of an update The same source may send a message and then send an update of this message, which means that it sends a new message corresponding to the same event to correct the information previously sent.

Such an update example is illustrated in Figure 2, the date d_2 of the updated message is more recent than d_1 , locations ℓ_1 and ℓ_2 of the messages are neighbouring, mass functions m_1 and m_2 may be different or not.

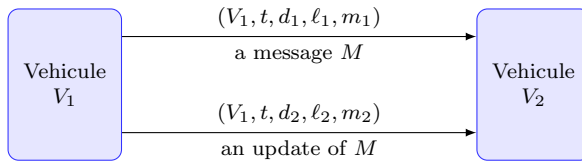


Fig. 2 Illustration of a message update example.

To determine an update, a threshold, denoted λ_t and depending on the type t of the event, is used. It corresponds to the maximum value for a difference of location to consider that a message can be an update of another.

More precisely, the system will consider that a message M_1 is an update of a message M_2 if:

- $M_1.S = M_2.S$;
- and $\Delta(M_1.\ell, M_2.\ell) < \lambda_t$ with Δ a distance;
- and $M_1.d > M_2.d$.

If a message present in the database is an update of the received message, the latter is ignored. If the received message is an update of a message present in the database, then it replaces it. Otherwise the message goes to the next test.

Grouping messages corresponding to the same event The received message has not been already received and is not an update of a message present in the database, it remains to look for the table M^j in which to put it, in other words to look for which events j it corresponds to.

Let $\mathcal{M}$ be the set of all the messages M_i^j corresponding to an event j of the same type of the event associated with the received message Mr and which is close in location, which means that $\mathcal{M}$ is composed of all the messages M_i^j verifying: $Mr.t = M_i^j.t$ and $\Delta(Mr.\ell, M_i^j.\ell) < \lambda_t$ with Δ a distance measure.

If $\mathcal{M}$ contains at least one message, the received message Mr is inserted into the same table of the message M_i^j present in $\mathcal{M}$ which is the nearest in location.

If $\mathcal{M}$ is empty, a new event j is created, its table M^j containing only one message which is the received message Mr .

3.4 Messages obsolescence

Over time, all messages stored in tables are not necessarily useful: some of them become obsolete. Then messages having a perception date too old are deleted.

The deletion is realized using a threshold, denoted Del_t depending on the type t of the event.

So, in parallel with the other processes, each message M_i^j such that:

$$\Delta(now, M_i^j.d) > Del_t, \quad (10)$$

with Δ a distance measure, is suppressed.

3.5 Data Fusion: giving an overview of the situation to the driver

From the database composed of the tables M^j , themselves composed of messages M_i^j corresponding to a same event j , an overview of the existing events with their probabilities can be given to the driver, in the following manner:

- First, each mass function $M_i^j.m$ is corrected with a correction rate equal to $\frac{\Delta(now, M_i^j.d)}{Del_t}$ to take into account the aging of the message. The older the message, the more corrected it is.

Two correction mechanisms can be used:

- the discounting operation (cf Equation (4)): $\alpha_i^j M_i^j.m = (1 - \alpha_i^j) M_i^j.m + \alpha_i^j m_\Omega$, where the correction rate $\alpha_i^j = \frac{\Delta(now, M_i^j.d)}{Del_t}$ corresponds to a discounting rate. With this operation, over time $\alpha_i^j M_i^j.m$ tends to the total ignorance m_Ω .
- the reinforcement process (cf Equation (5)): $\nu_i^j M_i^j.m = (1 - \nu_i^j) M_i^j.m + \nu_i^j m_{\bar{J}}$, where the correction rate $\nu_i^j = \frac{\Delta(now, M_i^j.d)}{Del_t}$ is a reinforcement rate. With this operation, over time $\nu_i^j M_i^j.m$ tends to $m_{\bar{J}}$ meaning that event j has disappeared.

Let us recall that in both cases, messages are discarded as soon as $\Delta(now, M_i^j.d) > Del_t$ (cf Equation (10)).

- Then, for each event j , corrected mass functions ($\alpha_i^j M_i^j.m$ or $\nu_i^j M_i^j.m$) are conjunctively combined using Equation (6).
- At last, pignistic probabilities regarding each event presence are obtained using Equation (9).

These two methods are tested and compared in Section 4.3.1.

An overview of the whole proposed system is given in Figure 3.

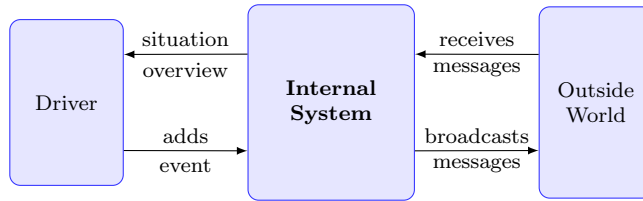


Fig. 3 General flowchart of the model. The internal system manages new messages, repeats messages which have still an interest, deletes useless one, and offers an overview at driver's request.

3.6 Variant: exchanging and keeping only the fusion results in databases

In this section, a variant of the first method is presented to save bandwidth and limit the number of messages stored in databases. In this second model, only the fusion results are kept in databases and exchanged between vehicles.

Only the changes between these two models are exposed in the sequel.

3.6.1 Messages and databases

Each vehicle database contains only one message associated with each event.

A message M^j associated with event j is a 5-tuple (S, t, d, ℓ, m) where:

- $M^j.S$ is the set of sources which have contributed to messages creation concerning the event j ;
- $M^j.t$ is the type of the event j ;
- $M^j.d$ is the date and time of the most recent message concerning the event j ;
- $M^j.\ell$ is the set of locations of the event j ;
- $M^j.m$ is the mass function representing the confidence of sources $M^j.S$ regarding the fact that the event is present.

When possible, vehicles exchange their messages, each one being viewed as a synthesis of the information related to an event on the road.

3.6.2 Reception and fusion process

A received message Mr concerning an event j already identified in the vehicle database (they are of the same type and close in location) is fused with message M^j as follows:

- The new set of sources is equal to $M^j.S \cup Mr.S$;
- The date of M^j becomes the most recent date among $M^j.d$ and $Mr.d$;
- The new mass function of M^j is obtained as follows:
 - The mass function of the message having the oldest date among M^j and Mr is corrected to take into consideration its aging (correction rate being equal to $\frac{|\Delta(Mr.d, M^j.d)|}{\Delta t}$).
 - If $Mr.S \cap M^j.S = \emptyset$ (before adding $Mr.S$ to $M^j.S$), the new mass function of M^j is obtained from the conjunctive rule of combination Equation (6) of the corrected mass functions of $M^j.m$ and $Mr.m$.
 - Otherwise, the cautious rule, Equation (7), is used to combine these two corrected mass functions.

If the event j is not already identified in the vehicle database, message M^j is created with the attributes of Mr : $M^j.S = \{Mr.S\}$, $M^j.t = Mr.t$, $M^j.d = Mr.d$, $M^j.\ell = Mr.\ell$ and $M^j.m = Mr.m$.

A comparison of the model keeping all the messages and the second model keeping only the fused messages is realized in Section 4.3.2 with both discounting and reinforcement correction operations.

4 Experimental tests using MatlabTM simulations

4.1 Simulator

In order to test and compare the proposed methods, a simulator has been developed in MatlabTM such that different scenarios can be tested.

A scenario is built with the following components:

- A scenario time axis τ with start and end dates, and a sampling period $\Delta\tau$.
- Type, location, start and end dates of each real event on the map.
- Attributes related to vehicles given by:
 - V : the number of vehicles.
 - $NetRange$: the network range in meter allowing V2V communication.
 - $Speed$: the average vehicle speed in km/h when no event is present on the road.
 - $SDRate_t$: the slow-down rate, depending on event type, used to compute new vehicle speed when an event exists at its location at τ moment. In this case, new vehicle speed is equal to $(1 - SDRate_t) \cdot Vit$ in km/h from τ to $\tau + 1$.
- The following attributes to obtain the path of each vehicle:
 - τ_d : the moment τ of vehicle departure;
 - ℓ_d : the departure vehicle location;
 - ℓ_a : the arrival vehicle location.

At each τ moment, the following actions are realized:

- each vehicle creates a message when an event is perceived;
- each vehicle denies an event present in vehicle database if the driver does not perceive it;
- each vehicle communicates its messages to neighboring vehicles in the network range $NetRange$.

4.2 Performance measure: Adequacy to the reality

The common purpose of the proposed models is to provide the most reliable information for the drivers. Thus performance rates of models are measured for each type t of event and for each vehicle v by the adequacy to the reality of the information given to the driver.

Formally, at each time step τ , the set equal to the union of the events j present in the vehicle database and the existing events in the reality is considered and denoted by $J_t^{v,\tau}$, and performance rates are computed for each type t of event and for each vehicle v at each time τ by:

$$Perf_t^{v,\tau} = 1 - \frac{\sum_{j \in J_t^{v,\tau}} (BetP_j^{v,\tau}(\{\exists\}) - R_j^\tau)^2}{|J_t^{v,\tau}|}, \quad (11)$$

where:

- $R_j^\tau = 1$ if event j is present at time τ , 0 otherwise.
- $|J_t^{v,\tau}|$: number of events of type t present in vehicle v database or existing in the reality at time τ .
- $BetP_j^{v,\tau}$: pignistic probability of vehicle v at time τ concerning the presence of the event j . If no message corresponding to event j is present in vehicle v database then $Betp_j^{v,\tau}(\{\exists\}) = 0$.

4.3 Experimentations

The methods exposed in Section 3 are tested and compared through different scenarios.

In all the scenarios exposed here, an accident starts at the beginning of the simulation and lasts 60 minutes, which means that it is present from $\tau = 0$ to $\tau = 900$ ($\Delta\tau = 4s$).

The confidence of all created messages is equal to 0.6: the mass function of the messages created is either $m(\{\exists\}) = 0.6, m(\Omega) = 0.4$ when vehicle confirms an event or $m(\{\neg\}) = 0.6, m(\Omega) = 0.4$ when vehicle denies an event.

4.3.1 Comparing messages aging mechanisms: discounting and reinforcement

Two scenarios are presented to compare aging mechanisms using the method where all messages are kept in vehicles databases.

Scenario n°1 In the first scenario, a vehicle v receives a message indicating the presence of the accident few minutes after its beginning. After a delay of Del_t equal to 100 minutes, the message is suppressed from the vehicle database being considered as obsolete.

In the following, in each figure illustrating the adequacy to the reality of a vehicle, the messages added to the vehicle database or suppressed from the vehicle database are illustrated by the following symbols just above the time τ where they occur:

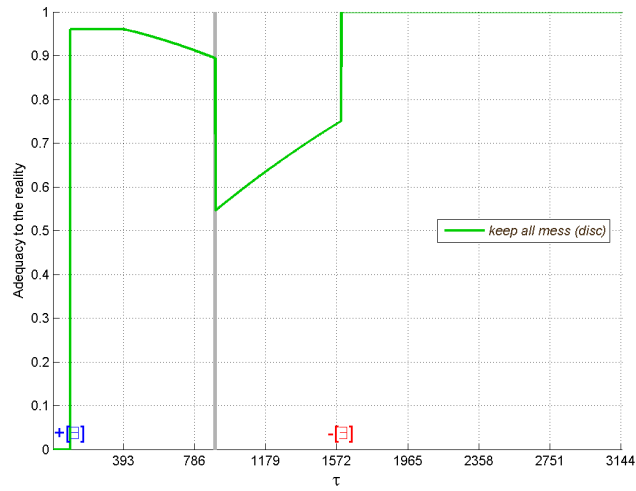
- $+n[\exists]$ which means that n new messages confirming the event are added to vehicle database;
- $+n[\neg\exists]$ which means that n new messages denying the event are added to vehicle database;
- $-n[\exists]$ which means that n messages confirming the event are deleted from vehicle database;
- $-n[\neg\exists]$ which means that n messages denying the event are deleted from vehicle database.

A grey vertical line indicates the disappearance of the event.

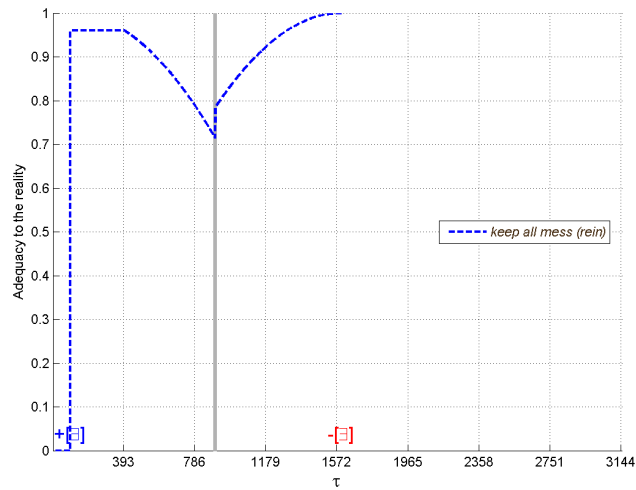
Figure 4 illustrates the adequacy to the reality of vehicle v during this first scenario.

The discounting mechanism tends to the total ignorance while the reinforcement process tends to the negation of the event. Therefore, the discounting mechanism gives a better result (a better adequacy to the reality) while the accident has not disappeared (the accident disappearing after the grey vertical line). On the contrary, the reinforcement process which anticipates the disappearance provides better results after the disappearance.

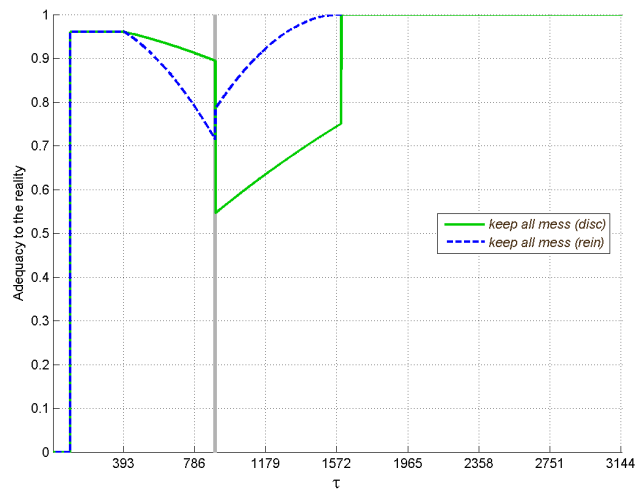
Scenario n°2 In this scenario, a vehicle v receives two messages confirming the accident, and two other messages denying the accident 10 minutes and 30



(a) with a discounting correction



(b) with a reinforcement correction



(c) Discounting and reinforcement corrections on the same figure

Fig. 4 Adequacy to the reality of the information given to the driver with the first model (keeping all the messages) during scenario n°1. Symbol $+[\emptyset]$ indicates the date of the arriving of a message which mentions the accident existence. Symbol $-[\emptyset]$ indicates the deletion of the message mentioning the existence of the accident. The grey vertical line corresponds to the disappearance of the accident. Accident starts at $\tau = 0$.

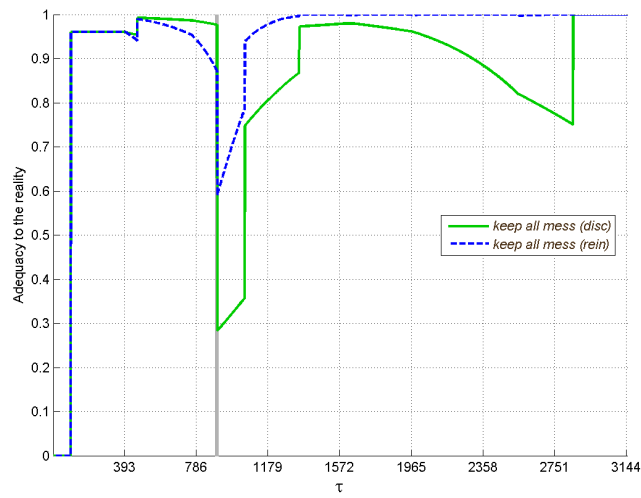
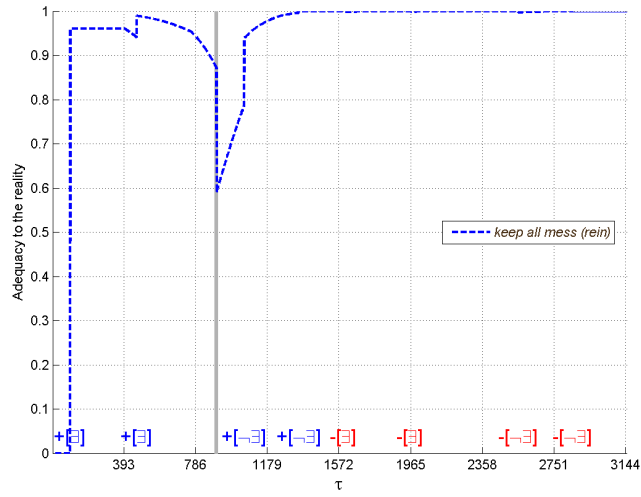
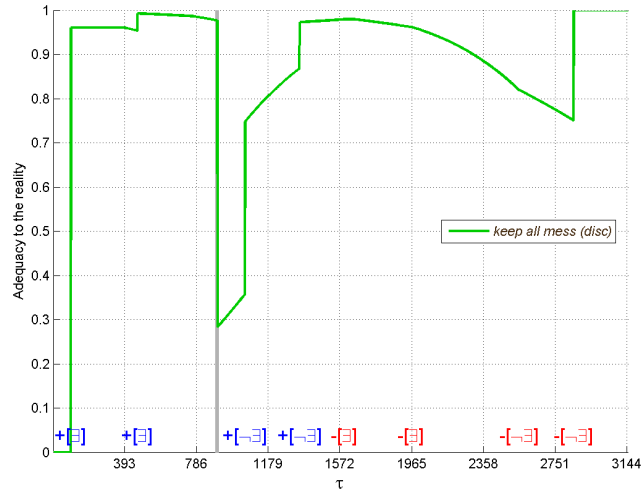


Fig. 5 Adequacy to the reality of the information given to the driver with the first model (keeping all the messages) during scenario n°2.

minutes after its disappearance. The adequacy to the reality of vehicle v is illustrated in Figure 5.

This scenario shows that after receiving messages denying the accident (from $\tau > 1050$, 10 minutes and 30 minutes after the accident disappearance), adequacy tends quickly to 100% with the reinforcement process while it is not the case with the discounting process. Indeed the discounting process tends to transfer the masses to the ignorance giving a 50% bet for the existence and 50% to the non-existence of the event.

4.3.2 Comparing methods: keep all messages, keep fusion results or keep yes/no last message

Two scenarios are presented to compare the method keeping all the messages (first method) and the second method described in Section 3.6 which conserves and exchanges only the fused messages. Each method are tested using the discounting or the reinforcement process for messages agings.

Scenario n°3 Three vehicles $v1$, $v2$ and $v3$ are considered in addition to a vehicle v .

This scenario proceeds as follows:

- a vehicle $v1$ creates a message $M1$ confirming the accident;
- vehicle $v1$ shares its message with vehicles $v2$ and $v3$;
- vehicle $v2$ creates a message $M2$ confirming the accident;
- vehicle $v2$ shares its database with v ;
- vehicle $v3$ creates a message $M3$ confirming the accident;
- vehicle $v3$ shares its database with v .

The database of the vehicle v depends on the method used:

- with the method where all messages are kept in vehicle database: vehicle v database contains $M1$ and $M2$, then $M1$, $M2$ and $M3$;
- with the method where only a synthesis message is kept in vehicle database: vehicle v database contains the fusion result message received from $v2$ (fusion of $M1$ and $M2$), then at the reception of the fusion result contained in $v3$ (fusion of $M1$ and $M3$), v applies an aging mechanism to both fusion results and combine them using the cautious rule of combination and keep in its database only the final fusion result.

The adequacy to the reality of the vehicle v is illustrated in Figure 6.

The method where vehicle keeps all messages in its database gives better results: before accident disappearance, the discounting mechanism gives the best result; then, the reinforcement mechanism gives the best result. This is due to the following reasons:

- the result of the conjunctive combination is higher than (or equal to) the result of the cautious conjunctive combination;
- discounting or reinforcing messages before combining them gives a higher result than combining messages before applying an aging mechanism.

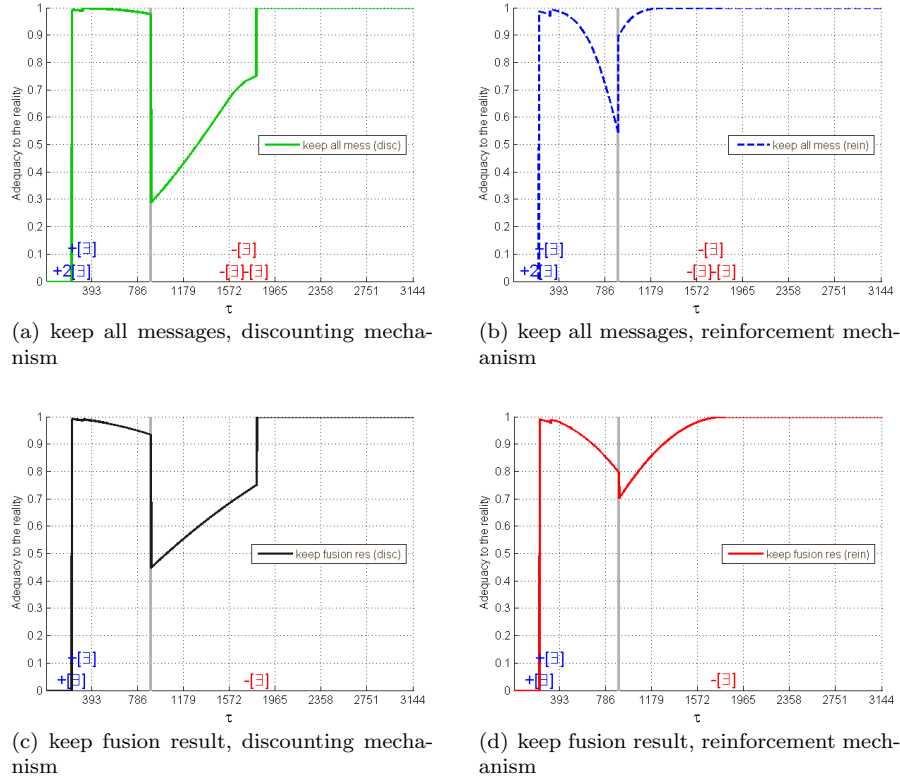


Fig. 6 Adequacy to the reality of the information contained in vehicle v for each method during scenario $n^{\circ}3$

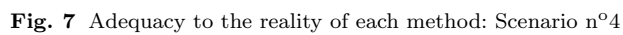
Scenario $n^{\circ}4$ In this scenario, previous methods are also compared with a very simple method called *y/n last mess* for short, where vehicles communicate simple messages telling *yes* or *no* an event exists; for each event, they keep in their databases only the last message. As all the drivers tell the truth (all the messages carry a true information), the last message is always the most recent true information.

In this scenario, a vehicle receives from distinct sources two messages confirming the accident, and two messages denying it after its disappearance.

Adequacy to the reality of the vehicle for each method is illustrated in Figure 7.

This scenario shows that after receiving messages denying the accident, the method where all messages are kept in vehicle database is still the best method when using reinforcement mechanism.

The result of the trivial method *y/n last mess* shows that this method does not allow predicting the accident disappearance, and gives a bad result after accident disappearance until the reception of a first message denying the



event. It means that even in an ideal not-very-real situation where drivers tell always the truth and messages are always received, a simple solution is not so evident.

5 An example of implementation with Android smartphones using Bluetooth transmissions

The purpose of this section is to illustrate how the first model presented in Section 3 can be implemented in Android smartphones with Bluetooth connections to illustrate a simple concrete case of use and to demonstrate that smartphones can manage belief functions calculations.

The choice of the Bluetooth technology is purely a matter of programming convenience. Data appearing in the following figures have been obtained on foot. Indeed no real tests were possible by car (or only at very slow speed) due to the fact that, with the Bluetooth version used by the authors, connection times were too long to allow smartphones to exchange data with cars passing on opposite sides on the road. As pointed out in the introduction of the article, the choice of the transmission technology is out of the scope of this article, the work presented here only concerns the application layer.

The following paragraphs detail both functional and technical aspects in relation with the implementation of the application and its possible services.

5.1 Application User Interfaces

As illustrated in Figure 3, the driver represents the unique actor of the system. He/She might play both sender and receiver roles, messages being periodically propagated "broadcasted" to the drivers within neighborhood coverage.

Application User Interfaces are based on four main screens:

- *Dashboard* (Figure 8): summarizes the description of the current vehicle environment. It lists a set of events sorted by their distance to the vehicle (Figure 8(a)). In addition to the event type, each list element shows the distance which is the average of the distances of the messages composing this event, the orientation of the event which is also the average of messages orientations and its pignistic probability which is the result of the aggregation of all the messages describing the event (cf Section 3.5). The time displayed near an event is the age of the most recent message in the list of messages constituting this event. For example, from Figure 8(a), there is a traffic jam in front of the vehicle at about 113 meters with a probability .98, and the most recent message concerning this event was received 1mn and 41s ago.

The dashboard event list is periodically updated, taking into account time, new location, and received messages. Besides, the user have the possibility to launch the update manually. A detailed event view is also available. For each event, the set of belonging messages is made accessible by a simple selection of an event from dashboard. For example, a pressure on the first event "Traffic_jam" on the dashboard illustrated in Figure 8(a), gives the result illustrated in Figure 8(b) where messages corresponding to this event can be observed. For each message, the distance between the vehicle and the event, the age of the message and the value of the mass function are

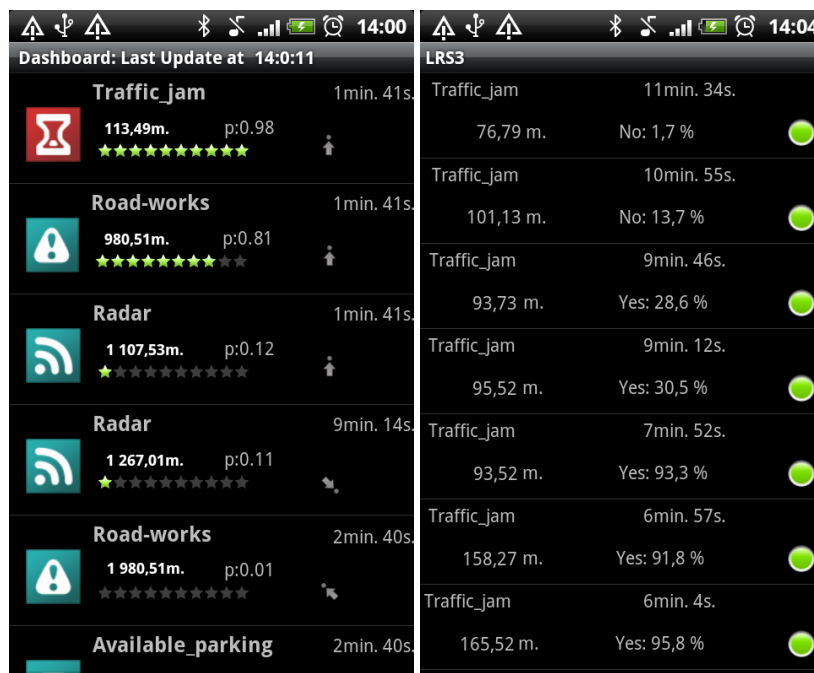


Fig. 8 Application dashboard.

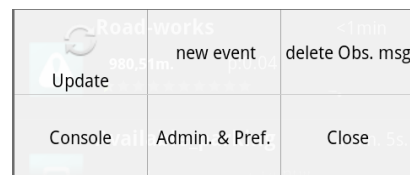


Fig. 9 Application menu

given. Each driver gives a simple mass function, then only masses on "yes, the event is present on the road at this time at my position" or "no it does not exist at this time at my position" are given (the complementary to one being given to Ω).

- *New event*: accessible via the menu interface (Figure 9), the new event interface (Figure 10(a)) affords to drivers the possibility to point out and report a new message describing an event. Drivers choose first the type of the event, then the purpose of the message: "yes, the event is present on the road at this time at my position" or "no it does not exist at this time at my position", and finally their degree of belief (in yes or no, the com-

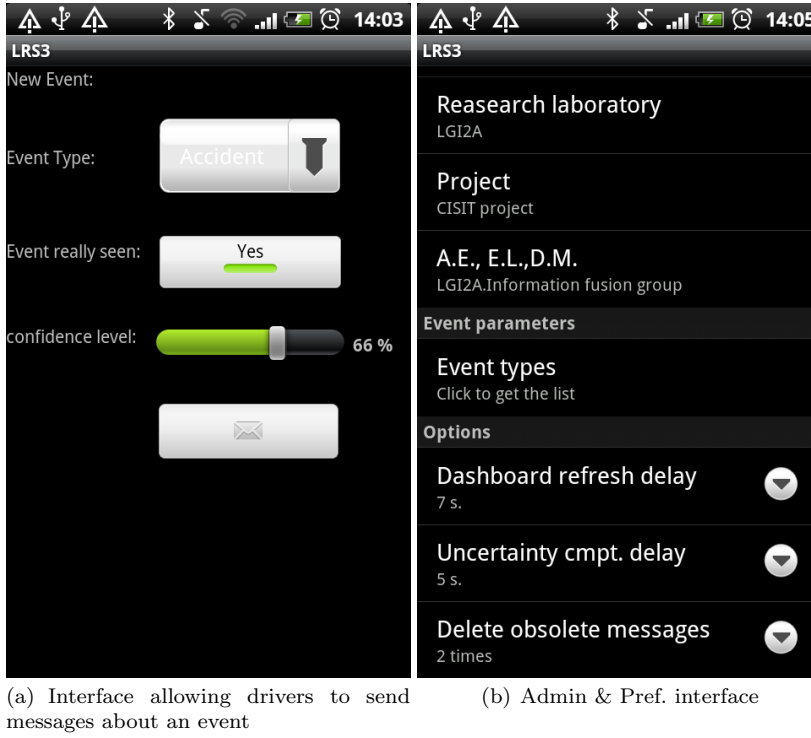


Fig. 10 Other user interfaces

- plementary being given to Ω). Once the message created, it will be added to the database and periodically propagated to the neighboring vehicles.
- *Console*: is the user interface which traces application evolution and provides more detailed information about application components life cycle. Since there are several threads started by the application, the monitoring of the application is required. The console could be seen as an interactive log file. Such an interface is not of a great interest of final users but remains a considerable issue for debugging or/and experimentation.
 - *Admin & Pref.* (Figure 10(b)): allows the user to manage event types as well as the configuration and the update of application parameters.

5.1.1 Software architecture: main application entities

Figure 11 illustrates the information system entities managed by the application. Message is the main entity. Each message belongs to a unique event. Messages classification is based on the event type and proximity indicators like location and time. The accuracy of information stream exchanged throughout messages are assessed, aggregated and merged according to their respective mass. For example, if an event is described by two messages, the respective

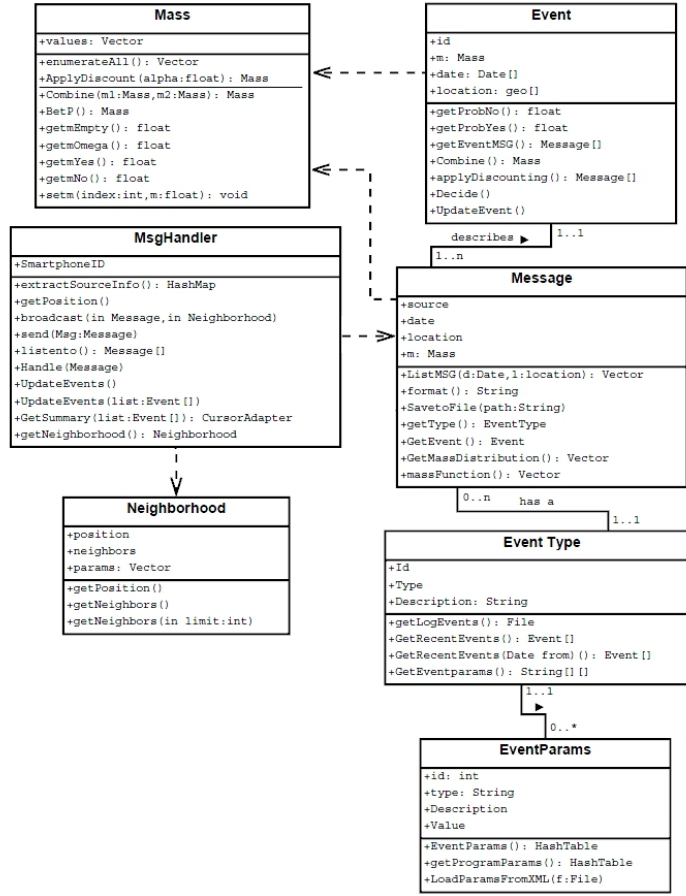


Fig. 11 Main application entities

masses are, first discounted, then combined and finally event consistency is assigned as a probability. Message handler class refers to the set of handler, periodic handler and procedures applied to both message and event databases. Finally, neighborhood entity denotes the set of devices discovers within neighborhood. We should note that the diagram illustrated by Figure 11, only covers representative functional entities. The class diagram takes into account more detailed and technical aspects of the application.

The application evolves through three states:

- Listening: the associated thread listen and waits for incoming connection. Once the connection is established, the current state changes to *receive* state.
- Receive: after state switching, a socket is created and the input stream is redirected. Then, binary data is read until the end of the stream. Binary data is converted to object messages. Next they are added to database.

- Send: Send state is used to propagate messages to neighboring smart phones. All non obsolete messages are sent to the connected devices.

6 Conclusion and future work

In this article, two models based on belief functions with two different message aging systems have been presented to manage uncertain messages informing drivers about events on the road in vehicular networks. A MatlabTM simulator has been developed to test them, and an implementation using Android smartphones have also been exposed to illustrate the feasibility of a concrete realization of these models.

One way to improve the method consists in trying to determine the age of an event from the series of messages related to it. It has indeed been observed that ideally messages created at the end of an event or after its disappearance should be reinforced towards the non existence of the event while messages confirming an event and created before its end should be kept. The knowledge of the beginning and the end of an event may significantly improve the results.

Another improvement concerns the spatial aspect of an event. In the models presented in this article an event takes place at a location ℓ which is sufficient for an accident but maybe not for a traffic jam which may spread to a large area. Likewise an accident can cause a traffic jam so the relations between types of events may also been considered.

Acknowledgements This work has been financed by the French region Nord-Pas de Calais under the project CISIT (Campus International pour la Sécurité et l'Intermodalité des Transports).

The authors are very grateful to the VESPA team, in particular Thierry Delot and Sylvain Lecomte from LAMIH Laboratory, Univ. of Valenciennes, for having helped them in their developments.

The authors would like to thank the anonymous reviewers for their valuable comments which have helped them to improve the clarity and the quality of this article.

References

1. A. Boukerche, Performance comparison and analysis of ad hoc routing algorithms, IEEE International Conference on Performance, Computing and Communications, pp. 171–178, (2001)
2. N. Cenerario, T. Delot, S. Ilarri, Dissemination of information in inter-vehicle ad hoc networks, IEEE Intelligent Vehicles Symposium (IV'2008), IEEE Computer Society, Eindhoven (The Netherlands), pp. 763–768, June (2008)
3. M. Bou Farah, D. Mercier, É. Lefevre, F. Delmotte, Towards a robust exchange of imperfect information in inter-vehicle ad-hoc networks using belief functions, IEEE Intelligent Vehicles Symposium, IV'2011, pp. 436–441, Baden-Baden, Germany, June 5–9 (2011)
4. V. Cherfaoui, T. Denoeux, Z.L. Cherfi, Distributed data fusion: application to confidence management in vehicular networks, International Conference on Information Fusion (FUSION'2008), pp. 846–853, Germany (2008)
5. S. Corson, J. Macker, Mobile ad hoc networking (MANET): routing protocol performance issues and evaluation considerations, RFC 2501 (1999)

6. B. Defude, T. Delot, S. Ilarri, J-L Zechinelli-Martini, N. Cenerario, Data aggregation in VANETs: the VESPA approach, *Mobiquitous'08 - First International Workshop on Computational Transportation Science (IWCTS)*, ACM digital library, Dublin, Ireland (2008)
7. L. Delgrossi, T. Zhang, *Vehicle Safety Communications: Protocols, Security, and Privacy*, John Wiley & Sons, Inc., Hoboken, NJ, USA (2012).
8. T. Delot, N. Cenerario, S. Ilarri, Estimating the relevance of information in inter-vehicle ad hoc networks, *International Workshop on Sensor Network Technologies for Information Explosion Era (SeNTIE'08)*, IEEE Computer Society, pp. 151–158, Beijing, China (2008)
9. F. Delmotte, P. Smets, Target identification based on the transferable belief model interpretation of Dempster-Shafer model, *IEEE Transactions on Systems, Man and Cybernetics A*, vol. 34, pp. 457–471 (2004)
10. T. Denœux, Conjunctive and disjunctive combination of belief functions induced by non distinct bodies of evidence, *Artificial Intelligence*, vol. 172, issue 2-3, pp. 234-264, February (2008).
11. European CAR 2 CAR Communication Consortium, <http://www.car-to-car.org>
12. ITS in Japan, <http://www.mlit.go.jp/road/ITS/>
13. X. Hong, K. Xu, M. Gerla, Scalable routing protocols for mobile ad hoc networks xi-aoyan, *IEEE Network Magazine*, vol. 16, pp. 11-21 (2002)
14. D. Mercier, T. Denœux, M.-H. Masson, Belief function correction mechanisms, *Studies in Fuzziness and Soft Computing*, B. Bouchon-Meunier et al. (Eds.), vol. 249, pp. 203–222, January (2010).
15. D. Mercier, É. Lefèvre, F. Delmotte, Belief functions contextual discounting and canonical decompositions, *International Journal of Approximate Reasoning*, vol.53, issue 2, pp. 146-158 (2012)
16. P. Morsink, R. Hallouzi, I. Dagli, C. Cseh, L. Schafers, M. Nelisse, D.D. Bruin, CarTalk 2000: Development of a cooperative ADAS based on vehicle-to-vehicle communication, *World Congress and Exhibition on Intelligent Transport Systems and Services*, Madrid, Spain, November (2003)
17. T. Nadeem, S. Dashtinezhad, C. Liao, L. Iftode, TrafficView: Traffic data dissemination using car-to-car communication, *ACM Mobile Computing and Communications Review (MC2R)*, Special Issue on Mobile Data Management, vol. 8, issue 3, pp. 6–19, July (2004)
18. D. Reichardt, M. Miglietta, L. Moretti, P. Morsink, W. Schulz, CarTALK 2000: safe and comfortable driving based upon inter-vehicle-communication, *IEEE Intelligent Vehicle Symposium*, vol. 2, pp. 545–550 (2002)
19. B. Ristic, P. Smets, Belief function theory on the continuous space with an application to model based classification, *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU'04*, pp. 4–9 (2004)
20. E.M. Royer, C.-K. Toh, A review of current routing protocols for ad hoc mobile wireless networks, *IEEE Personal Communications*, vol 6, issue 2, pp. 46-55 (1999)
21. G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press (1976)
22. P. Smets and R. Kennes, The Transferable Belief Model, *Artificial Intelligence*, vol. 66, pp. 191–243 (1994)
23. P. Smets, The Canonical Decomposition of a Weighted Belief, *International Joint Conferences on Artificial Intelligence*, pp. 1896–1901 (1995).
24. P. Smets, The Transferable Belief Model for quantified belief representation, *handbook of defeasible reasoning and uncertainty management systems*, D. M. Gabbay and Ph. Smets (Ed.), Kluwer Academic Publishers, Dordrecht, The Netherlands, vol. 1, pp. 267–301 (1998)
25. P. Smets, Decision making in the TBM: the necessity of the pignistic transformation, *International Journal of Approximate Reasoning*, vol. 38, issue 2, pp. 133-147 (2005)
26. US Department of Transportation RITA ITS program, <http://www.rita.dot.gov/>

Truthfulness in contextual information correction

Frédéric Pichon, David Mercier, François Delmotte, and Éric Lefèvre

Univ. Lille Nord de France, F-59000 Lille, France

UArtois, LGI2A, F-62400, Béthune, France

`{frederic.pichon,david.mercier,francois.delmotte,eric.lefevre}@univ-artois.fr`

Abstract. Recently, a dual reinforcement process to contextual discounting was introduced. However, it lacked a clear interpretation. In this paper, we propose a new perspective on contextual discounting: it can be seen as successive corrections corresponding to simple contextual lies. Most interestingly, a similar interpretation is provided for the reinforcement process. Two new contextual correction mechanisms, which are similar yet complementary to the two existing ones, are also introduced.

Keywords: Dempster-Shafer theory, Belief functions, Information correction, Discounting.

1 Introduction

Information correction has received quite a lot of attention in recent years in belief function theory (see, *e.g.*, [9, 11]). It is an important question that deals with how an agent should interpret a piece of information received from a source about a parameter $\mathbf{x}$ defined on a finite domain $\mathcal{X} = \{x_1, \dots, x_K\}$. Classically, the agent has some knowledge regarding the reliability of the source and, using the discounting operation [12], he is able to take into account that knowledge and to modify, or *correct*, the initial piece of information accordingly.

Since its inception, the discounting operation has been extended in different ways. Notably, Mercier *et al.* [10, 9] consider the case where one has some knowledge about the reliability of the source, conditionally on different subsets (contexts) A of $\mathcal{X}$, leading to the so-called contextual discounting operation. One may also refine the discounting operation in order to take into account knowledge about the source truthfulness [11]. Of particular interest for the present work is the dual reinforcement operation to contextual discounting introduced in [9]. Mercier *et al.* [9] show that this correction mechanism amounts to the negation [6] of the contextual discounting of the negation of the initial information, but unfortunately they do not go further in providing a clear interpretation for this interesting operation.

In this paper, we study further contextual correction mechanisms. We present (Section 3) a new framework for handling detailed meta-knowledge about source

truthfulness. Using this framework, we then derive the contextual discounting operation (Section 4.1) and its dual (Section 4.2), leading to a new perspective on the former and an interpretation for the latter. We proceed (Section 4.3) with the introduction of two new contextual correction mechanisms, whose interpretations are similar yet complementary to the two existing ones. Background material on belief function theory is first recalled in Section 2.

2 Belief function theory: necessary notions

In this section, we first recall basic concepts of belief function theory. Then, we present existing correction mechanisms that are of interest for this paper.

2.1 Basic concepts

In this paper, we adopt Smets' Transferable Belief Model (TBM) [14], where the beliefs held by an agent Ag regarding the actual value taken by $\mathbf{x}$ are modeled using a belief function [12] and represented using an associated mass function. A mass function (MF) on $\mathcal{X}$ is defined as a mapping $m : 2^{\mathcal{X}} \rightarrow [0, 1]$ verifying $\sum_{A \subseteq \mathcal{X}} m(A) = 1$. Subsets A of $\mathcal{X}$ such that $m(A) > 0$ are called *focal sets* of m . A MF having focal sets $\mathcal{X}$ and $A \subset \mathcal{X}$, with respective masses w and $1 - w$, $w \in [0, 1]$, may be denoted by A^w . A MF having focal sets $\emptyset$ and $A \neq \emptyset$, with respective masses v and $1 - v$, $v \in [0, 1]$, may be denoted by A_v . The negation $\bar{m}$ of a MF m is defined as $\bar{m}(A) = m(\bar{A})$, $\forall A \subseteq \mathcal{X}$, where $\bar{A}$ denotes the complement of A [6].

Beliefs can be aggregated using so-called combination rules. In particular, the conjunctive rule, which is the unnormalized version of Dempster's rule [5], is defined as follows. Let m_1 and m_2 be two MFs, and let $m_1 \odot_2$ be the MF resulting from their combination by the conjunctive rule denoted by $\odot$. We have:

$$m_1 \odot_2 (A) = \sum_{B \cap C = A} m_1(B) m_2(C), \quad \forall A \subseteq \mathcal{X}. \quad (1)$$

Other combination rules of interest for this paper are the disjunctive rule $\oslash$ [6], the exclusive disjunctive rule $\oplus$ and the equivalence rule $\odot$ [13]. Their definitions are similar to that of the conjunctive rule: one merely needs to replace $\cap$ in (1) by, respectively, $\cup$, $\sqcup$ and $\sqcap$, where $\sqcup$ (exclusive OR) and $\sqcap$ (logical equality) are defined respectively by $B \sqcup C = (B \cap \bar{C}) \cup (\bar{B} \cap C)$ and $B \sqcap C = (B \cap C) \cup (\bar{B} \cap \bar{C})$ for all $B, C \subseteq \mathcal{X}$. The interpretations of these four rules are discussed in detail in [11].

2.2 Correction mechanisms

Knowledge about a source reliability is classically taken into account in the TBM through the *discounting* operation. Suppose a source S providing a piece

of information represented by a MF m_S . Let β , with $\beta \in [0, 1]$, be Ag 's degree of belief that the source is reliable. Ag 's belief m on $\mathcal{X}$ is then defined by [12]:

$$m(\mathcal{X}) = \beta m_S(\mathcal{X}) + (1 - \beta), \quad m(A) = \beta m_S(A), \forall A \subseteq \mathcal{X}. \quad (2)$$

Mercier *et al.* [9] consider the case where Ag has some knowledge about the source reliability, conditionally on different subsets A of $\mathcal{X}$. Precisely, let β_A , with $\beta_A \in [0, 1]$, be Ag 's degree of belief that the source is reliable in context $A \subseteq \mathcal{X}$ and let $\mathcal{A}$ be the set of contexts for which Ag possesses such contextual meta-knowledge. Ag 's belief m on $\mathcal{X}$ is then defined by the following equation known as *contextual discounting* that subsumes discounting (recovered for $\mathcal{A} = \{\mathcal{X}\}$):

$$m = m_S \bigcup_{A \in \mathcal{A}} A_{\beta_A}. \quad (3)$$

In addition, a dual reinforcement process to contextual discounting, called contextual reinforcement hereafter, is introduced in [9]. Let m_S be a MF provided by a source S . The contextual reinforcement of m_S is the MF m defined by:

$$m = m_S \bigcap_{A \in \mathcal{A}} A^{\beta_A}, \quad (4)$$

with $\beta_A \in [0, 1]$, $A \in \mathcal{A}$. Mercier *et al.* [9] show that this correction amounts to the negation of the contextual discounting of the negation of m_S . However, they do not go further in providing a clear explanation as to what meta-knowledge on the source this correction of m_S corresponds. One of the main results of this paper is to provide such an interpretation.

3 A refined model of source truthfulness

In the correction schemes recalled in Section 2.2, the reliability of a source is assimilated to its relevance as explained in [11]. In [11], Pichon *et al.* assume that the reliability of a source involves in addition another dimension: its truthfulness. Pichon *et al.* [11] note that there exists various forms of lack of truthfulness for a source. For instance, for a sensor, it may take the form of a systematic bias. However, Pichon *et al.* [11] study only the crudest description of the lack of truthfulness, where a non truthful source is a source that declares the contrary of what it knows. According to this definition, from a piece of information of the form $\mathbf{x} \in B$ for some $B \subseteq \mathcal{X}$ provided by a relevant source S , one must conclude that $\mathbf{x} \in B$ or $\mathbf{x} \in \overline{B}$, depending on whether the source S is assumed to be truthful or not.

In this section, we propose a new and refined model of source truthfulness that allows the integration of more detailed meta-knowledge about the lack of truthfulness of an information source.

3.1 Elementary truthfulness

Assume that a relevant source provides a piece of information on the value taken by $\mathbf{x}$ of the form $\mathbf{x} \in B$, for some $B \subseteq \mathcal{X}$. Let us now consider a particular value

$x \in \mathcal{X}$. Either $x \in B$ or $x \notin B$, that is, the source may tell that x is possibly the actual value of $\mathbf{x}$ or it may tell that x is not a possibility for the actual value of $\mathbf{x}$. Furthermore, for each of those two possible declarations by the source about the value x , one may have some knowledge on whether the source is truthful or not. For instance, one may believe that the source is truthful when it tells that x is a possibility – in which case one must conclude that x is possibly the actual value of $\mathbf{x}$ if the source does tell that x is a possibility for $\mathbf{x}$ – and that it lies when it tells that x is not a possibility – in which case one must conclude that x is possibly the actual value of $\mathbf{x}$ if the source does tell that x is not a possibility for $\mathbf{x}$.

To account for such detailed knowledge about the behavior of the source, let us introduce two binary variables $\mathbf{p}_x$ and $\mathbf{n}_x$, with respective frames $\mathcal{P}_x = \{p_x, \neg p_x\}$ and $\mathcal{N}_x = \{n_x, \neg n_x\}$: p_x (resp. $\neg p_x$) corresponds to the state where the source is truthful (resp. not truthful) when it tells that x is possibly the actual value for $\mathbf{x}$; n_x (resp. $\neg n_x$) corresponds to the state where the source is truthful (resp. not truthful) when it tells that x is not a possibility for the actual value of $\mathbf{x}$.

Now, we can define a variable $\mathbf{t}_x$ with associated frame $\mathcal{T}_x = \mathcal{P}_x \times \mathcal{N}_x$, which contains four states $t_x = (p_x, n_x)$, $\neg t_x^n = (p_x, \neg n_x)$, $\neg t_x^p = (\neg p_x, n_x)$ and $\neg t_x = (\neg p_x, \neg n_x)$ allowing us to model the global truthfulness of the source with respect to the value x : t_x corresponds to the case where the source tells the truth whatever it says about the value x , in short the source is said to be truthful for x ; $\neg t_x^n$ corresponds to the case of a source that lies only when it tells that x is not a possibility for $\mathbf{x}$, which will be called a negative liar for x ; $\neg t_x^p$ corresponds to the case of a source that lies only when it says that x is a possibility for $\mathbf{x}$, which will be called a positive liar for x ; $\neg t_x$ corresponds to the case where the source lies whatever it says about the value x , in short the source is said to be non truthful for x .

There are thus four possible cases:

1. Suppose the source tells x is possibly the actual value of $\mathbf{x}$, *i.e.*, the information $\mathbf{x} \in B$ provided by the source is such that $x \in B$.
 - (a) If the source is assumed to be truthful (t_x) or a negative liar ($\neg t_x^n$), then one must conclude that x is possibly the actual value of $\mathbf{x}$;
 - (b) If the source is assumed to be a positive liar ($\neg t_x^p$) or non truthful ($\neg t_x$), then one must conclude that x is not a possibility for the actual value of $\mathbf{x}$;
2. Suppose the source tells x is not a possibility for the actual value of $\mathbf{x}$, *i.e.*, $x \notin B$.
 - (a) If the source is assumed to be in state t_x or in state $\neg t_x^p$, then one must conclude that x is not a possibility for the actual value of $\mathbf{x}$;
 - (b) If the source is assumed to be in state $\neg t_x^n$ or in state $\neg t_x$, then one must conclude that x is possibly the actual value of $\mathbf{x}$;

3.2 Contextual truthfulness

Let $\mathcal{T}$ denote the possible states of S with respect to its truthfulness for all $x \in X$. By definition, $\mathcal{T} = \times_{x \in \mathcal{X}} \mathcal{T}_x$. $\mathcal{T}$ is clearly a big space, however we will be interested in this paper only by a smaller subspace of $\mathcal{T}$, which we define below.

Let $h_A^{t_1, t_2} \in \mathcal{T}$, $A \subseteq \mathcal{X}$, $t_1, t_2 \in \mathcal{T}_x$, denote the state where the source is in state t_1 for all $x \in A$, and in state t_2 for all $x \notin A$. For instance, let $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$, $A = \{x_3, x_4\}$, $t_1 = \neg t_x^p$ and $t_2 = t_x$, then $h_A^{t_1, t_2} = h_{\{x_3, x_4\}}^{\neg t_x^p, t_x} = (t_{x_1}, t_{x_2}, \neg t_{x_3}^p, \neg t_{x_4}^p)$, *i.e.*, the source is a positive liar for x_3 and x_4 , and is truthful for x_1 and x_2 .

Consider now the following question: what must one conclude about $\mathbf{x}$ when the source tells $\mathbf{x} \in B$ and is assumed to be in some state $h_A^{t_1, t_2}$? To answer this question, one merely needs to look in turn at each $x \in \mathcal{X}$ and to consider 4 cases for each of those $x \in \mathcal{X}$: 1) $x \notin B$ and $x \notin A$; 2) $x \notin B$ and $x \in A$; 3) $x \in B$ and $x \notin A$; 4) $x \in B$ and $x \in A$. Table 1 lists, for each of the 4 cases and for all states $h_A^{t_1, t_2}$, $t_1, t_2 \in \mathcal{T}_x$, whether one should deduce that a given value $x \in \mathcal{X}$ is possibly the actual value of $\mathbf{x}$ or not – the former is indicated by a 1 and the latter by a 0 in columns $h_A^{t_1, t_2}$, $t_1, t_2 \in \mathcal{T}_x$.

Table 1. Interpretations of the source testimony according to its contextual truthfulness.

$x \in B$	$x \in A$	$\neg t_x^p, \neg t_x^p$	$\neg t_x^p, t_x$	$t_x, \neg t_x^p$	t_x, t_x	$\neg t_x^n, \neg t_x^n$	$\neg t_x^n, t_x$	$t_x, \neg t_x^n$	t_x, t_x	$\neg t_x^p, \neg t_x^p$	$\neg t_x^p, t_x$	$t_x, \neg t_x^p$	t_x, t_x	$\neg t_x^n, \neg t_x^n$	$\neg t_x^n, t_x$	$t_x, \neg t_x^n$	t_x, t_x
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

According to Table 1, when the source is assumed to be in, *e.g.*, state $h_A^{t_x, \neg t_x^p}$, *i.e.*, the source is truthful for all $x \in A$ and a positive liar for all $x \in \bar{A}$, then one should deduce that $x \in \mathcal{X}$ is a possible value for $\mathbf{x}$ iff $x \in B$ and $x \in A$, and therefore, since this holds for all $x \in \mathcal{X}$, one should deduce that $\mathbf{x} \in B \cap A$. For instance, consider state $h_{\{x_3, x_4\}}^{t_x, \neg t_x^p}$ and testimony $\mathbf{x} \in \{x_1, x_3\}$, then one should deduce $\{x_1, x_3\} \cap \{x_3, x_4\} = \{x_3\}$.

Another interesting state is $h_A^{\neg t_x^n, t_x}$, *i.e.*, the source is a negative liar for all $x \in A$ and truthful for all $x \in \bar{A}$, in which case $x \in \mathcal{X}$ is a possible value for $\mathbf{x}$ iff $x \in B$ or $x \in A$, and thus one should conclude that $\mathbf{x} \in B \cup A$. More generally, as can be seen from Table 1, the couples $(t_1, t_2) \in \mathcal{T}_x^2$ yields all possible binary Boolean connectives.

Of particular interest in this paper are the states $h_A^{t_x, \neg t_x^p}$ and $h_A^{\neg t_x^n, t_x}$, which have already been discussed, and the states $h_A^{t_x, \neg t_x}$ (the source is truthful for all $x \in A$ and non truthful for all $x \in \bar{A}$) and $h_A^{\neg t_x, t_x}$ (the source is non truthful for all $x \in A$ and truthful for all $x \in \bar{A}$), which yield respectively $\mathbf{x} \in B \cap A$ and

$\mathbf{x} \in B \sqcup A$. Accordingly, we will consider in the sequel only the following subspace $\mathcal{H} \subseteq \mathcal{T}$: $\mathcal{H} = \{h_A^{t_1, t_2} | A \subseteq \mathcal{X}, (t_1, t_2) \in \{(t_x, \neg t_x^p), (\neg t_x^n, t_x), (t_x, \neg t_x), (\neg t_x, t_x)\}\}$.

Following [11], we can encode the above reasoning by a multivalued mapping $\Gamma_B : \mathcal{H} \rightarrow \mathcal{X}$ indicating how to interpret the information $\mathbf{x} \in B$ in each state $h \in \mathcal{H}$; we have for all $A \subseteq \mathcal{X}$:

$$\Gamma_B(h_A^{t_x, \neg t_x^p}) = B \cap A, \Gamma_B(h_A^{\neg t_x^n, t_x}) = B \cup A, \Gamma_B(h_A^{t_x, \neg t_x}) = B \sqcap A, \Gamma_B(h_A^{\neg t_x, t_x}) = B \sqcup A.$$

If the knowledge about the source state is imprecise and given by $H \subseteq \mathcal{H}$, then one should deduce the image $\Gamma_B(H) := \bigcup_{h \in H} \Gamma_B(h)$ of H by Γ_B .

3.3 Uncertain testimony and meta-knowledge

More generally, both the testimony provided by the source and the knowledge of Ag about the source truthfulness may be uncertain. Let m_S be the uncertain testimony and $m^{\mathcal{H}}$ the uncertain meta-knowledge. In such case, the *Behavior-Based Correction* (BBC) procedure introduced by Pichon *et al.* [11], can be used to derive Ag knowledge on $\mathcal{X}$. It is represented by the MF m defined by [11]:

$$m(C) = \sum_{H \subseteq \mathcal{H}} m^{\mathcal{H}}(H) \sum_{B: \Gamma_B(H)=C} m_S(B), \quad \forall C \subseteq \mathcal{X}. \quad (5)$$

For convenience, we may denote by $f_{m^{\mathcal{H}}}(m_S)$ the BBC of m_S according to meta-knowledge $m^{\mathcal{H}}$, *i.e.*, we have $m = f_{m^{\mathcal{H}}}(m_S)$ with m defined by (5).

4 Interpretation of contextual corrections

In this section, we propose a new perspective on contextual discounting by recovering it using the framework introduced in Section 3. Then, using a similar reasoning, we provide an interpretation for contextual reinforcement. Finally, we introduce two new contextual correction schemes that are complementary to the two existing ones.

4.1 Contextual discounting in terms of BBCs

Let us consider a particular contextual lie among those introduced in Section 3.2: the states $h_A^{\neg t_x^n, t_x}$, $A \subseteq \mathcal{X}$, corresponding to the assumptions that the source is a negative liar for all $x \in A$ and truthful for all $x \in \overline{A}$. Among these states, $h_{\emptyset}^{\neg t_x^n, t_x}$ admits a simpler interpretation: it corresponds to assuming that the source is truthful $\forall x \in \mathcal{X}$.

Theorem 1. *Let m_S be a MF. We have, $\forall A$ and with $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$:*

$$m_S \bigoplus_{A \in \mathcal{A}} A_{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m_{A, \cup}^{\mathcal{H}}})(m_S), \quad (6)$$

where $\circ$ denotes function composition (*i.e.*, $(g \circ f)(x) = g(f(x))$) and where $m_{A, \cup}^{\mathcal{H}}$ is defined by $m_{A, \cup}^{\mathcal{H}}(\{h_{\emptyset}^{\neg t_x^n, t_x}\}) = \beta_A$, $m_{A, \cup}^{\mathcal{H}}(\{h_A^{\neg t_x^n, t_x}\}) = 1 - \beta_A$, $\forall A \in \mathcal{A}$.

Proof. This theorem can be shown by applying for each $A \in \mathcal{A}$, $\mathcal{A}$ being finite, the following property:

$$f_{m_{A,\cup}^{\mathcal{H}}}(m_S) = m_S \odot A_{\beta_A}, \quad \forall A \in \mathcal{A}, \quad (7)$$

which is shown as follows.

From (5) and the definition of $m_{A,\cup}^{\mathcal{H}}$, $\forall C \subseteq \mathcal{X}$:

$$f_{m_{A,\cup}^{\mathcal{H}}}(m_S)(C) = \beta_A \sum_{B:B=C} m_S(B) + (1 - \beta_A) \sum_{B:B \cup A=C} m_S(B). \quad (8)$$

Which means:

$$f_{m_{A,\cup}^{\mathcal{H}}}(m_S) = \beta_A m_S + (1 - \beta_A)(m_S \odot m_A), \quad (9)$$

with m_A a MF defined by $m_A(A) = 1$.

On the other hand, $\forall A \in \mathcal{A}$:

$$m_S \odot A_{\beta_A} = m_S \odot \begin{cases} A \mapsto 1 - \beta_A \\ \emptyset \mapsto \beta_A \end{cases} = \beta_A m_S + (1 - \beta_A)(m_S \odot m_A). \quad (10)$$

□

In other words, contextual discounting, which appears on the left side of (6), corresponds to successive behavior-based corrections – one for each context $A \in \mathcal{A}$ – where for each context A , we have the following meta-knowledge: with mass β_A the source is truthful for all $x \in \mathcal{X}$, and with mass $1 - \beta_A$ the source is a negative liar for all $x \in A$ and truthful for all $x \in \bar{A}$.

Successive corrections of an initial piece of information is a process that may be encountered when considering a chain of sources, where the information provided by an initial source may be iteratively corrected by the sources down the chain according to the knowledge each source has on the behavior of the preceding source. The chain of sources problem is an important and complex one, which has received different treatments in logic [4], possibility theory [1] and belief function theory [2, 3]: in particular a solution involving successive corrections, precisely successive discountings, was proposed in [1]. The fact that contextual discounting may be relevant for this problem had not been remarked yet.

4.2 Contextual reinforcement in terms of BBCs

Let us consider another kind of contextual lie: the states $h_A^{t_x, \neg t_x^p}$, $A \subseteq \mathcal{X}$, corresponding to the assumptions that the source is truthful for all $x \in A$ and a positive liar for all $x \in \bar{A}$. Among these states, $h_{\mathcal{X}}^{t_x, \neg t_x^p}$ has the same simple interpretation as $h_{\emptyset}^{\neg t_x^p, t_x}$.

Theorem 2. Let m_S be a MF. We have, $\forall A$ and with $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$:

$$m_S \bigcirc_{A \in \mathcal{A}} A^{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m_{A, \cap}^{\mathcal{H}}})(m_S), \quad (11)$$

where $m_{A, \cap}^{\mathcal{H}}$ is defined by $m_{A, \cap}^{\mathcal{H}}(\{h_{\mathcal{X}}^{t_x, \neg t_x^p}\}) = \beta_A$, $m_{A, \cap}^{\mathcal{H}}(\{h_A^{t_x, \neg t_x^p}\}) = 1 - \beta_A$, $\forall A \in \mathcal{A}$.

Proof. The proof is similar to that of Theorem 1. $\square$

Theorem 2 is important as it constitutes the first known interpretation for contextual reinforcement. It shows that, similarly to contextual discounting, contextual reinforcement (left side of (11)) corresponds to successive behavior-based corrections – one for each context. The only difference between the two correction mechanisms is what is assumed with mass $1 - \beta_A$: with the former that the source is a negative liar for all $x \in A$ and truthful for all $x \in \bar{A}$, whereas with the latter that the source is truthful for all $x \in A$ and a positive liar for all $x \in \bar{A}$.

Example 1. Let us consider a series of three agents: agent 1 reports to agent 2, who reports in turn to agent 3. Let m_i denote the beliefs of agent i on $\mathcal{X} = \{x_1, x_2, x_3\}$ and let $m_i^{\mathcal{H}}$, $i > 1$, denote the meta-knowledge of agent i about agent $i - 1$. Furthermore, assume that $m_2^{\mathcal{H}}(\{h_{\mathcal{X}}^{t_x, \neg t_x^p}\}) = 0.6$ and $m_2^{\mathcal{H}}(\{h_{\{x_1, x_2\}}^{t_x, \neg t_x^p}\}) = 0.4$, that is, agent 2 believes with mass 0.6 that agent 1 is truthful for all $x \in \mathcal{X}$, and with mass 0.4 that agent 1 is truthful for x_1 and x_2 and a positive liar for x_3 . Suppose further that $m_3^{\mathcal{H}}(\{h_{\mathcal{X}}^{t_x, \neg t_x}\}) = 0.8$ and $m_3^{\mathcal{H}}(\{h_{\{x_2, x_3\}}^{t_x, \neg t_x}\}) = 0.2$. From Theorem 2, we have

$$\begin{aligned} m_2 &= m_1 \bigcirc \{x_1, x_2\}^{0.6}, \\ m_3 &= m_2 \bigcirc \{x_2, x_3\}^{0.8}, \\ m_3 &= m_1 \bigcirc \{x_1, x_2\}^{0.6} \bigcirc \{x_2, x_3\}^{0.8}. \end{aligned}$$

4.3 Two new contextual correction mechanisms

Contextual discounting and contextual reinforcement are based on corrections induced by simple pieces of meta-knowledge $m_{A, \cup}^{\mathcal{H}}$ and $m_{A, \cap}^{\mathcal{H}}$ respectively. In practice, those pieces of meta-knowledge transform a testimony $\mathbf{x} \in B$ as follows: they both allocate mass β_A to B , and mass $1 - \beta_A$ to $B \cup A$ and to $B \cap A$, respectively.

Now, as we have seen in Section 3.2, there exist states $h_A^{t_1, t_2} \in \mathcal{T}$ that lead to other binary Boolean connectives than the disjunction and the conjunction. This suggests a way to extend contextual discounting and contextual reinforcement. Of particular interest are states $h_A^{t_x, \neg t_x}$ (the source is truthful for all $x \in A$ and non truthful for all $x \in \bar{A}$) and $h_A^{\neg t_x, t_x}$ (the source is non truthful for all $x \in A$ and truthful for all $x \in \bar{A}$), which yield respectively $\mathbf{x} \in B \sqcap A$ and $\mathbf{x} \in B \sqcup A$. Indeed, the properties satisfied by connectives $\sqcap$ and $\sqcup$ allow us to obtain similar relations as those obtained for contextual discounting and contextual reinforcement:

Theorem 3. Let m_S be a MF. We have, $\forall \mathcal{A}$ and with $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$:

$$m_S \bigcirc_{A \in \mathcal{A}} A^{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m_{A, \sqcap}}^{\mathcal{H}})(m_S), \quad (12)$$

$$m_S \bigcup_{A \in \mathcal{A}} A^{\beta_A} = (\circ_{A \in \mathcal{A}} f_{m_{A, \sqcup}}^{\mathcal{H}})(m_S), \quad (13)$$

where $m_{A, \sqcap}^{\mathcal{H}}$ and $m_{A, \sqcup}^{\mathcal{H}}$ are defined by $m_{A, \sqcap}^{\mathcal{H}}(\{h_{\mathcal{X}}^{t_x, \neg t_x}\}) = m_{A, \sqcup}^{\mathcal{H}}(\{h_{\emptyset}^{\neg t_x, t_x}\}) = \beta_A$, and $m_{A, \sqcap}^{\mathcal{H}}(\{h_A^{t_x, \neg t_x}\}) = m_{A, \sqcup}^{\mathcal{H}}(\{h_A^{\neg t_x, t_x}\}) = 1 - \beta_A$, $\forall A \in \mathcal{A}$.

Proof. The proof is similar to that of Theorem 1. $\square$

Eqs. (12) and (13) are the $\sqcap$ and $\sqcup$ counterparts to Eqs. (6) and (11), which are based on connectives $\cup$ and $\cap$. Hence, if contextual discounting and contextual reinforcement are renamed as $\sqcup$ -contextual correction and $\sqcap$ -contextual correction, then Eqs. (12) and (13) may be called $\sqcap$ -contextual correction and $\sqcup$ -contextual correction. Let us also stress that although the $\sqcap$ and $\sqcup$ -contextual correction mechanisms are based on less classical combination rules than contextual discounting and contextual reinforcement, these two new contextual correction schemes seem to be as reasonable from the point of view of the meta-knowledge that they correspond to. Actually, their interpretations are even simpler since they rely on the classical assumptions of truthfulness and non truthfulness, whereas contextual discounting and contextual reinforcement involve negative and positive lies, which are less conventional. Finally, we note that the computational complexity of the $\sqcap$ and $\sqcup$ -contextual correction mechanisms is similar to that of $\cup$ and $\cap$ -contextual correction mechanisms: it merely corresponds to the complexity of applying $|\mathcal{A}|$ combinations by the rules $\bigcirc$ and $\bigcup$, respectively, where $|\mathcal{A}|$ denotes the cardinality of $\mathcal{A}$.

5 Conclusion

Using a new framework for handling detailed meta-knowledge about source truthfulness, a new view on contextual discounting and an interpretation for contextual reinforcement were proposed. In addition, two similar yet complementary contextual correction mechanisms were introduced.

Future work will be dedicated to the application of contextual correction mechanisms. Similarly as contextual discounting [7, 10], their parameters β_A , $A \in \mathcal{A}$, could be obtained from a confusion matrix or learnt from training data, and then they could be used in classification problems. Other potential applications include those involving chain of sources communicating pieces of information between themselves, as is the case in vehicular ad-hoc networks [8].

References

1. T. Baerecke, T. Delavallade, M.-J. Lesot, F. Pichon, H. Akdag, B. Bouchon-Meunier, P. Capet, and L. Cholvy. Un mode de cotation pour la veille informationnelle en source ouverte. In *6me colloque Veille Stratgique Scientifique et Technologique, Toulouse, France*, 2010.

2. L. Cholvy. Evaluation of information reported: A model in the theory of evidence. In E. Hüllermeier, R. Kruse, and F. Hoffmann, editors, *Proc. of the Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Dortmund, Germany*, volume 80 of *CCIS*, pages 258–267. Springer-Verlag, 2010.
3. L. Cholvy. Collecting information reported by imperfect information sources. In S. Greco, B. Bouchon-Meunier, G. Coletti, M. Fedrizzi, B. Matarazzo, and R. R. Yager, editors, *Proc. of the Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Catania, Italy*, volume 299 of *CCIS*, pages 501–510. Springer, 2012.
4. L. Cholvy. Un modèle de cotation pour la veille informationnelle en source ouverte. In *European Workshop on Multi-agent Systems, Toulouse, France*, 2013.
5. A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
6. D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12(3):193–226, 1986.
7. Z. Elouedi, E. Lefèvre, and D. Mercier. Discountings of a belief function using a confusion matrix. In *Int. Conf. on Tools with Artificial Intelligence*, volume 1, pages 287–294, 2010.
8. M. Bou Farah, D. Mercier, E. Lefèvre, and F. Delmotte. A high-level application using belief functions for exchanging and managing uncertain events on the road in vehicular ad hoc networks. *Annals of telecommunications*, 69(3-4):185–199, 2014.
9. D. Mercier, E. Lefèvre, and F. Delmotte. Belief functions contextual discounting and canonical decompositions. *International Journal of Approximate Reasoning*, 53(2):146 – 158, 2012.
10. D. Mercier, B. Quost, and T. Dencœur. Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 9(2):246 – 258, 2008.
11. F. Pichon, D. Dubois, and T. Dencœur. Relevance and truthfulness in information correction and fusion. *International Journal of Approximate Reasoning*, 53(2):159–175, 2012.
12. G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.
13. Ph. Smets. The α -junctions: combination operators applicable to belief functions. In D. M. Gabbay, R. Kruse, A. Nonnengart, and H. J. Ohlbach, editors, *1st Int. J. Conf. on Qual. and Quant. Pract. Reas., Bad Honnef, Germany*, volume 1244 of *Lecture Notes in Computer Science*, pages 131–153. Springer, 1997.
14. Ph. Smets and R. Kennes. The Transferable Belief Model. *Artificial Intelligence*, 66:191–243, 1994.

Learning contextual discounting and contextual reinforcement from labelled data

David Mercier, Frédéric Pichon, Éric Lefèvre, and François Delmotte

Univ. Lille Nord de France, F-59000 Lille, France
UArtois, LGI2A, F-62400, Béthune, France
`{firstname.lastname@univ-artois.fr}`

Abstract. This paper addresses the problems of learning from labelled data contextual discounting and contextual reinforcement, two correction schemes recently introduced in belief function theory. It shows that given a particular error criterion based on the plausibility function, for each of these two contextual correction schemes, there exists an optimal set of contexts that ensures the minimization of the criterion and that finding this minimum amounts to solving a constrained least-squares problem with as many unknowns as the domain size of the variable of interest.

Keywords: Belief functions, Learning, Contextual Discounting, Contextual Reinforcement, Source Biases, Expert Tuning.

1 Introduction

Classically, in belief function theory, the correction of the information provided by a source concerning a variable of interest $\mathbf{x}$ defined on a finite domain $\mathcal{X}$, is achieved using the discounting operation [8–10]. This operation admits one parameter, a real β belonging to $[0, 1]$, reflecting the degree of *reliability* of the source of information [9, Section 5.7] [5, Section 2.5].

Discounting operation has been extended by Mercier *et al.* in [5], where it is considered that one may have some knowledge about the reliability of a source, conditionally on different subsets (*contexts*) of $\mathcal{X}$, the set of contexts forming a partition of $\mathcal{X}$. This operation, called *contextual discounting based on a coarsening*, is controlled by a vector of parameters β_A , each β_A belonging to $[0, 1]$ and reflecting the degree of reliability of the source given context $A \subseteq \mathcal{X}$ (in other words, knowing that the true value of $\mathbf{x}$ lies in A). In this same article [5, Section 5], following preceding work from Elouedi *et al.* for the classical discounting [3], a computationally efficient method to automatically learn from labelled data the parameters β_A of a contextual discounting based on a coarsening, once a partition (a set of contexts) has been fixed, is also introduced; the idea is to find the parameters values which minimize a measure of discrepancy between the ground truth and the outputs of the source corrected according to the parameters values. This method is potentially useful to improve a source performance in, *e.g.*, a classification application, as well as to discover its contextual reliability. However, the problem of finding the optimal partition of $\mathcal{X}$ for a given source was left open.

In [6], Mercier *et al.* have extended this contextual discounting based on a coarsening to be applicable to any set of contexts (*i.e.*, the set of contexts no longer needs to form a partition of $\mathcal{X}$). This mechanism is therefore simply called *contextual discounting (CD)*. The *contextual reinforcement (CR)* of a source is also introduced as the dual of CD. This new correction operation is also controlled by a vector of parameters β_A in $[0, 1]$ associated with a set of subsets (contexts) A of $\mathcal{X}$. However, the interpretation of CR was not clear (it was only known that CR amounts to the negation [2] of the CD of the negation of the information provided by the source) and the problem of learning CD and CR from labelled data was not tackled. Recently, in [7], Pichon *et al.* gave an interpretation to CR: it amounts to assuming, for each context A , that the source is truthful with mass β_A , and that with mass $(1 - \beta_A)$ it lies only when it tells that the true value of $\mathbf{x}$ is in $\overline{A}$; but the question of learning CD and CR from data remained open.

In this paper, we address the problems of learning from labelled data contextual discounting and contextual reinforcement, the former problem being only partially addressed so far – its solution is restricted to the case where a set of contexts has been fixed beforehand and where this set must also form a partition of $\mathcal{X}$ – and the latter problem not being addressed at all. Especially, we show that given the discrepancy measure used in [5], there exists an optimal set of contexts for the most general form of CD (the CD proposed in [6] and that does not require the set of contexts to form a partition of $\mathcal{X}$) that ensures the minimization of the measure, and that finding this minimum amounts to a computationally simple optimization problem (a constrained least-squares problem with K unknowns, K being the size of $\mathcal{X}$). Furthermore, we show that a similar result holds for CR. In addition, an illustrative example of the proposed learning of CD and CR is given. This example is also useful to make insightful additional remarks on CD and CR, and in particular the potential superiority of the recently introduced CR mechanism over CD, to improve a source performance.

This paper is organized as follows. Required basic concepts on belief functions and contextual correction mechanisms are exposed in Section 2. Learning of CD and of CR is formally studied in Section 3 and illustrated in Section 4, where a comparison of CD and CR correction capacities is also presented. Finally, Section 5 concludes the paper.

2 Belief functions and contextual correction mechanisms: basic concepts and notations

2.1 Representation and combination of beliefs

The necessary background material on the representation and combination of beliefs is given here.

Representation of beliefs A *mass function (MF)* represents an agent’s opinion regarding a variable of interest $\mathbf{x}$ taking values in a finite domain $\mathcal{X}$ [8, 10].

It is defined as a mapping $m : 2^{\mathcal{X}} \rightarrow [0, 1]$ verifying $\sum_{A \subseteq \mathcal{X}} m(A) = 1$. The negation $\bar{m}$ of a mass function m is defined as $\bar{m}(A) = m(\bar{A})$, $\forall A \subseteq \mathcal{X}$.

A MF m defined by $m(\mathcal{X}) = w$ and $m(A) = 1 - w$, with $w \in [0, 1]$ and $A \subset \mathcal{X}$, can be conveniently noted A^w . Likewise a MF m such that $m(\emptyset) = v$ and $m(A) = 1 - v$, with $v \in [0, 1]$, $A \subseteq \mathcal{X}$, $A \neq \emptyset$, can be conveniently noted A_v .

A MF m is in one-to-one correspondence with a plausibility function pl , a commonality function q and an implicability function b , which are respectively defined by: $pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$, $q(A) = \sum_{A \subseteq B} m(B)$ and $b(A) = \sum_{B \subseteq A} m(B)$, for all $A \subseteq \mathcal{X}$.

Combination Two MF m_1 and m_2 can be combined using the *conjunctive rule of combination* [10] denoted by $\odot$ and defined by:

$$(m_1 \odot m_2)(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C), \quad \forall A \subseteq \mathcal{X}. \quad (1)$$

Numerous combination rules exist [11] to merge mass functions. The other combination of particular interest in this paper is the *disjunctive rule of combination* $\oslash$ [2, 9] defined by replacing the symbol $\cap$ in (1) by $\cup$.

Let us also recall that if $m = m_1 \odot m_2$, the corresponding commonality functions verify $q = q_1 \cdot q_2$, and if $m = m_1 \oslash m_2$, the corresponding implicability functions verify $b = b_1 \cdot b_2$.

2.2 Contextual discounting and reinforcement of a belief function

Throughout this paper, m_S is a MF defined on $\mathcal{X}$, provided by a source S , and $\mathcal{A}$ is a set of subsets (contexts) of $\mathcal{X}$.

Contextual discounting (CD) The contextual discounting [5–7] of m_S is the MF m defined, with $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$, by:

$$m = m_S \oslash_{A \in \mathcal{A}} A_{\beta_A}. \quad (2)$$

The classical discounting [8, 9] is retrieved when $\mathcal{A}$ is composed of just one element which is the whole domain $\mathcal{X}$:

$$m = m_S \oslash \mathcal{X}_{\beta} = \beta m_S + (1 - \beta)m_{\mathcal{X}}, \quad (3)$$

with $m_{\mathcal{X}}$ defined by $m_{\mathcal{X}}(\mathcal{X}) = 1$.

In practice, β_A represents the proportion of $m_S(B)$ which remains on B , and $(1 - \beta_A)$ represents the part of $m_S(B)$ transferred to $B \cup A$, $\forall A \in \mathcal{A}$ and $\forall B \subseteq \mathcal{X}$.

Contextual reinforcement (CR) The contextual reinforcement [6, 7] of m_S is the MF m defined, with $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$, by:

$$m = m_S \odot_{A \in \mathcal{A}} A^{\beta_A}. \quad (4)$$

In practice, β_A represents the fraction of $m_S(B)$ remaining on B , and $(1 - \beta_A)$ the part of $m_S(B)$ transferred to $B \cap A$, $\forall A \in \mathcal{A}$ and $\forall B \subseteq \mathcal{X}$.

3 Learning CD and CR from labelled data

3.1 Description of the learning process

In this section, we study how to automatically learn CD and CR from:

1. A training set describing the outputs of a source (expressed in the form of a MF) regarding the classes in $\mathcal{X} = \{x_1, \dots, x_K\}$ of n objects o_i , $i \in \{1, \dots, n\}$ (A small illustrative example is given in Section 4 in Table 2);
2. And a measure of discrepancy to be minimized between the corrections of the mass functions provided by the source and the reality.

In this paper, the following measure of discrepancy between the corrected information and the ground truth has been chosen:

$$E_{pl}(\beta) = \sum_{i=1}^n \sum_{k=1}^K (pl\{o_i\}(\{x_k\}) - \delta_{i,k})^2, \quad (5)$$

where $\forall i \in \{1, \dots, n\}$, $pl\{o_i\}$ is the plausibility function obtained from a contextual correction of the output m_S of the source with a vector of coefficients $\beta \in [0, 1]^{|\mathcal{A}|}$. The binary variable $\delta_{i,k}$ indicates the class of o_i as follows: $\forall i \in \{1, \dots, n\}$, $\forall k \in \{1, \dots, K\}$, $\delta_{i,k} = 1$ if object o_i belongs to the class x_k , and $\delta_{i,k} = 0$ otherwise.

Our choice to use measure E_{pl} (5) is mostly based on the fact that, as it will be seen in Propositions 2 and 4, its minimization has the advantage to yield constrained least-squares problems, which can be solved efficiently. Moreover, it was the one used in the approach proposed in [5], which we are clearly extending with this current work. At last, using the plausibility on singletons is in accordance with the Shafer [8] and Smets [10] *singular* [1] interpretation of belief functions, which is adopted in this paper. However, we may note that other measures of discrepancy could be used, *e.g.*, a measure based on the pignistic probability [10] or on a distance measure [4], but then it is not guaranteed that their minimization can be performed efficiently.

3.2 Learning CD

Plausibilities on the singletons after having applied CD on a MF m_S provided by a source are given by next proposition.

Proposition 1. *Let $m = m_S \bigoplus_{A \in \mathcal{A}} A_{\beta_A}$, $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$, be the CD of a MF m_S . The plausibility function associated with m is defined for all $x \in \mathcal{X}$ by:*

$$pl(\{x\}) = 1 - (1 - pl_S(\{x\})) \prod_{A \in \mathcal{A}, x \in A} \beta_A. \quad (6)$$

Proof. See Appendix A.1

Next proposition indicates that the minimization of E_{pl} when CD has been applied, is obtained using the vector β composed of the K parameters $\beta_{\{x_k\}}$, which means the parameters associated with the singletons of $\mathcal{X}$. Moreover the minimization of E_{pl} using this vector constitutes a constrained least-squares problem which can then be solved efficiently using standard algorithms.

Proposition 2. *The minimization of E_{pl} with CD is obtained using the vector $\beta = (\beta_{\{x_k\}}, k \in \{1, \dots, K\})$ and constitutes a constrained least-squares problem as (5) can then be rewritten as:*

$$E_{pl}(\beta) = \|\mathbf{Q}\beta - \mathbf{d}\|^2 \text{ with } \mathbf{Q} = \begin{bmatrix} \text{diag}(\mathbf{pl}_1 - 1) \\ \vdots \\ \text{diag}(\mathbf{pl}_n - 1) \end{bmatrix} \text{ and } \mathbf{d} = \begin{bmatrix} \delta_1 - 1 \\ \vdots \\ \delta_n - 1 \end{bmatrix}, \quad (7)$$

with $\text{diag}(\mathbf{v})$ a square diagonal matrix with the elements of vector $\mathbf{v}$ on the main diagonal, and with $\mathbf{pl}_i = (pl_S\{o_i\}(\{x_1\}), \dots, pl_S\{o_i\}(\{x_K\}))^T$, and $\delta_i = (\delta_{i,1}, \dots, \delta_{i,K})^T$ the column vector of 0-1 class indicator variables for object o_i .

Proof. See Appendix A.2

This answers a prospect given in [5] concerning the study of the set of contexts which yields the best possible value for the measure of discrepancy E_{pl} . The answer given here is that there will be no smaller value reachable for E_{pl} than the one obtained with the set of the singletons of $\mathcal{X}$ with associated coefficients $\beta = (\beta_{\{x_k\}}, k \in \{1, \dots, K\})$.

3.3 Learning CR

Plausibilities on the singletons after having applied CR are given in next proposition.

Proposition 3. *Let $m = m_S \bigoplus_{A \in \mathcal{A}} A^{\beta_A}$, $\beta_A \in [0, 1]$, $\forall A \in \mathcal{A}$, be the CR of a MF m_S . The plausibility function associated with m is defined for all $x \in \mathcal{X}$ by:*

$$pl(\{x\}) = pl_S(\{x\}) \prod_{A \in \mathcal{A}, x \notin A} \beta_A. \quad (8)$$

Proof. See Appendix A.3

Proposition 4. *The minimization of E_{pl} with CR is obtained using the vector $\beta = (\beta_{\{x_k\}}, k \in \{1, \dots, K\})$ and constitutes a constrained least-squares problem as (5) can then be written as:*

$$E_{pl}(\beta) = \|\mathbf{P}\beta - \delta\|^2, \text{ with } \mathbf{P} = \begin{bmatrix} \text{diag}(\mathbf{pl}_1) \\ \vdots \\ \text{diag}(\mathbf{pl}_n) \end{bmatrix} \text{ and } \delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix}, \quad (9)$$

with the same notations as in Proposition 2.

Proof. See Appendix A.4

4 CD and CR learnings: comments and illustration

4.1 CD and CR respective correction capacities

The differences between CD and CR concerning their respective plausibilities ranges on singletons after having been applied are briefly discussed here.

With CD, as $pl(\{x_k\}) = 1 - (1 - pl_S(\{x_k\}))\beta_{\{x_k\}}$ for each $k \in \mathcal{X}$, $k \in \{1, \dots, K\}$, with $\beta_{\{x_k\}}$ varying in $[0, 1]$, $pl(\{x_k\})$ can take any values in the interval $[pl_S(\{x_k\}), 1]$. It means that with CD the value on each singleton $pl_S(\{x\})$ can be shifted as close to 1 as required, in other words weakened as required.

In contrast, with CR, as $pl(\{x_k\}) = pl_S(\{x_k\})\beta_{\overline{\{x_k\}}}$ for each $x_k \in \mathcal{X}$, $k \in \{1, \dots, K\}$, $pl(\{x_k\})$ can take any values in $[0, pl_S(\{x_k\})]$ with $\beta_{\overline{\{x_k\}}}$ varying from 0 to 1. With CR, the value of the plausibility on each singleton can then be carried as close to 0 as necessary. In other words, CR strengthens the information provided by the source by decreasing the plausibilities on certain singletons.

The following example illustrates these different capacities of adjustment to the reality on simple scenarios for CD and CR.

Example 1. Let us suppose that $\mathcal{X} = \{a, b, c\}$ and, without lack of generality, that the ground truth is a .

Let us suppose that a source n°1 outputs a mass $m_S(\{b, c\}) = 1$ which means that $pl_S(\{a\}) = 0$ and $pl_S(\{b\}) = pl_S(\{c\}) = 1$. To bring closer source n°1 output and the reality: CD can increase $pl_S(\{a\})$ to 1; CR can decrease $pl_S(\{b\})$ to 0 and $pl_S(\{c\})$ to 0.

This example is taken again in Table 1, and two more situations are considered: a source n°2 giving $m_S(\{c\}) = 1$, that is $pl_S(\{a\}) = pl_S(\{b\}) = 0$ and $pl_S(\{c\}) = 1$ and a source n°3 giving $m_S(\{a, b\}) = 1$, that means $pl_S(\{a\}) = pl_S(\{b\}) = 1$ and $pl_S(\{c\}) = 0$.

Table 1. Attainable plausibilities with CD and CR for three sources outputs.

	Ground truth	Source n°1	CD	CR	Source n°2	CD	CR	Source n°3	CD	CR
$pl(\{a\})$	1	0	1	0	0	1	0	1	1	1
$pl(\{b\})$	0	1	1	0	0	0	0	1	1	0
$pl(\{c\})$	0	1	1	0	1	1	0	0	0	0
				CD: $E_{pl} = 2$					CD: $E_{pl} = 1$	
				CR: $E_{pl} = 1$					CR: $E_{pl} = 1$	
									CD: $E_{pl} = 1$	
									CR: $E_{pl} = 0$	

As it can be observed in Table 1, CD can improve only one value of plausibility: the plausibility on the ground truth by increasing it as close as possible to 1, whereas CR can improve all the other plausibility values (all except the one associated with the ground truth) by decreasing them as near as possible to 0. CR has then more degrees of flexibility to improve the plausibility output of the source. Situations where CD can be of more help than CR, are in particular

those where all the plausibilities on singletons which are not the ground truth are equal to zero, for example: $pl_S(\{b\}) = pl_S(\{c\}) = 0$ and $pl_S(\{a\}) = 0.5$, which means $m_S(\{a\}) = 0.5$ and $m_S(\emptyset) = 0.5$.

4.2 An illustrative example

Inspired from [3] and [5, Section 5], we consider the following small example of target recognition illustrated in Table 2.

Example 2. Two sensors are in charge of recognizing flying objects which can be airplanes (a), helicopters (h) or rockets (r). Data are composed of 4 known objects on which two sensors have expressed their outputs as MF on $\mathcal{X} = \{a, h, r\}$.

Table 2. Outputs of two sensors regarding the classes of 4 objects which can be airplanes (a), helicopters (h) or rockets (r). Data come from [3, Table 1].

		$\{a\}$	$\{h\}$	$\{r\}$	$\{a, h\}$	$\{a, r\}$	$\{h, r\}$	$\mathcal{X}$	Ground truth
Sensor 1	$m_{S_1}\{o_1\}$	0	0	0.5	0	0	0.3	0.2	a
	$m_{S_1}\{o_2\}$	0	0.5	0.2	0	0	0	0.3	h
	$m_{S_1}\{o_3\}$	0	0.4	0	0	0.6	0	0	a
	$m_{S_1}\{o_4\}$	0	0	0	0	0.6	0.4	0	r
Sensor 2	$m_{S_2}\{o_1\}$	0	0	0	0.7	0	0	0.3	a
	$m_{S_2}\{o_2\}$	0.3	0	0	0.4	0	0	0.3	h
	$m_{S_2}\{o_3\}$	0.2	0	0	0	0	0.6	0.2	a
	$m_{S_2}\{o_4\}$	0	0	0	0	0	1	0	r

Results of the minimization of E_{pl} for CD and CR are summarized in Table 3 for both sensors 1 and 2. Let us recall that $\beta = (\beta_{\{a\}}, \beta_{\{h\}}, \beta_{\{r\}})$ for CD, and $\beta = (\beta_{\overline{\{a\}}}, \beta_{\overline{\{b\}}}, \beta_{\overline{\{c\}}})$ for CR with different meanings for each correction mechanism.

Table 3. Results for the minimization of E_{pl} with the data in Table 2 for each contextual correction mechanism for both sensors 1 and 2.

Contextual correction	Sensor 1	Sensor 2
CD	$\beta = (0.76, 1.00, 1.00)$	$\beta = (0.74, 1.00, 1.00)$
	$E_{pl}(\beta) = 3.39$	$E_{pl}(\beta) = 4.81$
CR	$\beta = (0.94, 0.66, 0.38)$	$\beta = (0.65, 0.22, 0.55)$
	$E_{pl}(\beta) = 2.33$	$E_{pl}(\beta) = 2.39$

For CD it can be observed that $\beta_{\{h\}} = \beta_{\{r\}} = 1$ for both sensors, which means that both sensors are reliable to detect objects h and r . There is no need to transfer a portion of mass $m_S(B)$ to $B \cup \{h\}$ or $B \cup \{r\}$ with $B \subseteq \mathcal{X}$

($1 - \beta_{\{h\}} = 1 - \beta_{\{r\}} = 0$). It is not the case for objects of type a which cause problems for both sensors, sensor 1 being slightly more reliable.

Minimizing E_{pl} with CR confirms that both sensors are more truthful to recognize objects h and r as $\beta_{\{a\}} = 0.94 > \beta_{\{h\}} > \beta_{\{r\}}$ for sensor 1 and $\beta_{\{a\}} = 0.65$ for sensor 2. In terms of mass transfers, there is less need to transfer a portion of mass $m_S(B)$ to $B \cap \{a\} = B \cap \{h, r\}$ ($1 - 0.94 = 0.06$ for sensor 1, and $1 - 0.65 = 0.35$ for sensor 2) than to $B \cap \{h\} = B \cap \{a, r\}$ (0.34 for sensor 1, and 0.78 for sensor 2) or $B \cap \{r\} = B \cap \{a, h\}$ (0.62 for sensor 1, and 0.45 for sensor 2). With these data, CR also permits to obtain lower values for E_{pl} than those reached with CD, which confirms the advantages of CR over CD exposed in Section 4.1 and in Example 1 concerning the minimization of E_{pl} .

4.3 On the absence of link between learning CR and CD

Even if CR and CD are related (CR amounts to the negation of the CD of the negation of the information provided by the source [6]), CR and CD parameters minimizing E_{pl} (5) cannot be deduced analytically from each other.

Let us consider next example which is a slight modification of example 2.

Example 3. By modifying in Table 2, MF $m_{S_1}\{o_1\}$ by $m_{S_1}\{o_1\}(\{r\}) = 0.5282$, $m_{S_1}\{o_1\}(\{h, r\}) = 0.3$ and $m_{S_1}\{o_1\}(\mathcal{X}) = 0.1718$ (information coming from Sensor 1 is slightly deteriorated, the truth being a), learning of CD parameters for sensors 1 and 2 yields the same vector $\beta = (0.74, 1.00, 1.00)$, while the learning of CR parameters yields $\beta = (0.92, 0.68, 0.38)$ for sensor 1 and $\beta = (0.65, 0.22, 0.55)$ for sensor 2.

Example 3 shows that knowing the vector β minimizing E_{pl} for CD does not imply knowing the vector β minimizing E_{pl} for CR.

5 Conclusion

In this paper, we have studied the learning of CD and of CR from labelled data, given a measure of discrepancy based on the plausibility function. We have shown that for each of these two contextual correction schemes, there exists an optimal set of contexts that ensures the minimization of the measure and that finding this minimum amounts to solving a constrained least-squares problem with K unknowns. These results can find applications in at least two domains: learning the biases of a source of information (what are the characteristics of a source?) and in source tuning (how to tune a source to obtain the best performances?). Future work will consist in exploiting these mechanisms in more complex applications and investigating the tuning of the combination of several sources.

References

1. Destercke, S., Burger, T.: Toward an axiomatic definition of conflict between belief functions. *IEEE Trans. Cybern.* 43(2), 585–596 (2012)

2. Dubois, D., Prade, H.: A Set-Theoretic View of Belief Functions: logical operations and approximations by fuzzy sets. *Int. J. Gen. Syst.* 12, 193–226 (1986)
3. Elouedi, Z., Mellouli, K., Smets, P.: Assessing Sensor Reliability for Multisensor Data Fusion Within the Transferable Belief Model. *IEEE Trans. Syst. Man Cybern. Part B.* 34(1), 782–787 (2004)
4. Jousselme, A.-L., Maupin, P.: Distances in evidence theory: Comprehensive survey and generalizations, *Int. J. Approx. Reason.* 53(2), 118–145 (2012)
5. Mercier, D., Quost, B., Dencœux, T.: Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Inf. Fusion.* 9(2), 246–258 (2008)
6. Mercier, D., Lefèvre, É., Delmotte, F.: Belief functions contextual discounting and canonical decompositions. *Int. J. Approx. Reason.* 53(2), 146–158 (2012)
7. Pichon, F., Mercier, D., Lefèvre, É., Delmotte, F.: Truthfulness in contextual information correction. In: Cuzzolin, F. (Ed.) *Belief Functions: Theory and Applications. Third International Conference, BELIEF 2014, Oxford, UK, September 26–28, 2014. Proceedings. LNCS, vol. 8764, pp. 11–20. Springer, Heidelberg (2014)*
8. Shafer, G.: *A mathematical theory of evidence.* Princeton University Press, Princeton, N.J., (1976)
9. Smets, P.: Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem. *Int. J. Approx. Reason.* 9, 1–35 (1993)
10. Smets, P., Kennes, R.: The Transferable Belief Model. *Artif. Intell.* 66, 191–243 (1994)
11. Smets, P.: Analyzing the combination of conflicting belief functions. *Inf. Fusion* 8, 387–412 (2007)

A Appendices

A.1 Proof of Proposition 1

As $m = m_S \bigoplus_{A \in \mathcal{A}} A_{\beta_A}$, CD is given in terms of implicability functions by: $b = b_S \prod_{A \in \mathcal{A}} b_{\beta_A}$ with $b_{\beta_A}(B) = 1$ if $A \subseteq B$, β_A otherwise, for all $B \subseteq \mathcal{X}$. Thus, for all $B \subseteq \mathcal{X}$: $b(B) = b_S(B) \prod_{A \in \mathcal{A}, A \not\subseteq B} \beta_A$, and consequently, for all $x \in \mathcal{X}$:

$$\begin{aligned}
 pl(\{x\}) &= 1 - b(\overline{\{x\}}) = 1 - b_S(\overline{\{x\}}) \prod_{A \in \mathcal{A}, A \not\subseteq \overline{\{x\}}} \beta_A \\
 &= 1 - b_S(\overline{\{x\}}) \prod_{A \in \mathcal{A}, x \in A} \beta_A = 1 - (1 - pl_S(\{x\})) \prod_{A \in \mathcal{A}, x \in A} \beta_A.
 \end{aligned}$$

A.2 Proof of Proposition 2

From Proposition 1, after having applied CD on m_S , the discrepancy measure E_{pl} (5) can be written: $E_{pl}(\beta) = \sum_{k=1}^K E_{pl}(\beta, x_k)$, with for all $k \in \{1, \dots, K\}$:

$$E_{pl}(\beta, x_k) := \sum_{i=1}^n \left(\left(1 - (1 - pl_S\{o_i\}(\{x_k\})) \prod_{A \in \mathcal{A}, x_k \in A} \beta_A \right) - \delta_{i,k} \right)^2. \quad (10)$$

As $E_{pl}(\beta, x_k) \geq 0$ for all $k \in \{1, \dots, K\}$, the minimum value of $E_{pl}(\beta)$ is obtained when each $E_{pl}(\beta, x_k)$ reaches its minimum.

Besides, as all coefficients β_A belong to $[0, 1]$, for each $x_k, k \in \{1, \dots, K\}$, the product $\prod_{A \in \mathcal{A}, x_k \in A} \beta_A$ of coefficients β_A in $E_{pl}(\beta, x_k)$ (10) also belongs to $[0, 1]$ and can be denoted by a variable $\beta_k \in [0, 1]$. Hence, for each $k \in \{1, \dots, K\}$, the minimum of $E_{pl}(\beta, x_k)$ is reached for a particular value of β_k .

Now, we can remark that each coefficient $\beta_{\{x_k\}} \in [0, 1], k \in \{1, \dots, K\}$, only appears in the expression of $E_{pl}(\beta, x_k)$ (10), $k \in \{1, \dots, K\}$. Hence, choosing $\beta_k = \beta_{\{x_k\}}$ for all k (which means choosing $\mathcal{A}$ composed of the set of singletons of $\mathcal{X}$) constitutes then a solution, *i.e.*, a set of contexts for which the minimum value of $E_{pl}(\beta)$ is reached.

Each value of E_{pl} is then reachable using the vector β of coefficients $\beta_k := \beta_{\{x_k\}}, k \in \{1, \dots, K\}$, and as already mentioned in [5, Section 5.1], the computation of the coefficient β with CD based on the singletons is a constrained least-squares problem. Indeed, for all $k \in \{1, \dots, K\}$, and for all $i \in \{1, \dots, n\}$:

$$pl\{o_i\}(\{x_k\}) - \delta_{i,k} = 1 - (1 - pl_S\{o_i\}(\{x_k\}))\beta_k - \delta_{i,k} \quad (11)$$

$$= (pl_S\{o_i\}(\{x_k\}) - 1)\beta_k - (\delta_{i,k} - 1). \quad (12)$$

Then (5) can be rewritten as (7).

A.3 Proof of Proposition 3

As $m = m_S \bigcirc_{A \in \mathcal{A}} A^{\beta_A}$, the CR is determined in terms of commonality functions by $q = q_S \prod_{A \in \mathcal{A}} q^{\beta_A}$ with $q^{\beta_A}(B) = 1$ if $B \subseteq A$, β_A otherwise, for all $B \subseteq \mathcal{X}$. Then, for all $B \subseteq \mathcal{X}$: $q(B) = q_S(B) \prod_{A \in \mathcal{A}, B \not\subseteq A} \beta_A$, which means that after having applied CR, plausibilities on singletons are defined, for all $x \in \mathcal{X}$, by:

$$pl(\{x\}) = q(\{x\}) = q_S(\{x\}) \prod_{A \in \mathcal{A}, x \notin A} \beta_A = pl_S(\{x\}) \prod_{A \in \mathcal{A}, x \notin A} \beta_A$$

A.4 Proof of Proposition 4

From Proposition 3, for each $k \in \{1, \dots, K\}$, coefficient $\beta_{\overline{\{x_k\}}}$ takes its values in $[0, 1]$ and only appears in $pl(x_k)$ when a CR has been applied. Then, with the same reasoning as for the CD case, the minimum value of E_{pl} with CR can be reached using the set of contexts $\{\overline{x_k} = \mathcal{X} \setminus \{x_k\}, k \in \{1, \dots, K\}\}$.

The minimization of E_{pl} with CR based on the vector $\beta = (\beta_k := \beta_{\overline{\{x_k\}}}, k \in \{1, \dots, K\})$ is also a constrained least-squares problem as (5) can be written as (9) (as $\forall k \in \{1, \dots, K\}$ and $\forall i \in \{1, \dots, n\}$, $pl\{o_i\}(\{x_k\}) - \delta_{i,k} = pl_S\{o_i\}(\{x_k\})\beta_k - \delta_{i,k}$).

Bibliographie

- [1] International Campus on Safety and Intermodality in Transportation CISIT project (2007-2013) : <http://www.cisit.org>.
- [2] A. APPRIOU : Probabilités et incertitude en fusion de données multi-senseurs. *Revue Scientifique et Technique de la Défense*, 11:27–40, 1991.
- [3] Y. BAR-SHALOM et T. FORTMANN : *Tracking and data association*. Academic Press, Boston, 1988.
- [4] S. BLACKMAN et R. POPOLI : *Design and Analysis of Modern Tracking*. Systems, Artech House, 1999.
- [5] I. BLOCH : *Fusion d'informations en traitement du signal et des images*. Hermès, Paris, France, 2003.
- [6] M. BOU FARAH : *Méthodes utilisant des fonctions de croyance pour la gestion des informations imparfaites dans les réseaux de véhicules*. Thèse de doctorat, Université d'Artois, 2014.
- [7] M. BOU FARAH, D. MERCIER, F. DELMOTTE, É. LEFÈVRE et S. LAGRUE : Methods Handling Accident and Traffic Jam Information with Belief Functions in VANETs. In F. CUZZOLIN, éditeur : *Proceedings of the 3rd International Conference on Belief Functions, BELIEF 2014*, volume 8764 de *Lecture Notes in Computer Science*, pages 124–133, Oxford, United Kingdom, septembre 2014. Springer International Publishing.
- [8] M. BOU FARAH, D. MERCIER, É. LEFÈVRE et F. DELMOTTE : Towards a robust exchange of imperfect information in inter-vehicle ad-hoc networks using belief functions. In *Proceedings of the 4th Intelligent Vehicles Symposium, IV 2011*, pages 436–441, juin 2011.
- [9] M. BOU FARAH, D. MERCIER, É. LEFÈVRE et F. DELMOTTE : Un processus V2V d'échanges et de gestion d'informations imparfaites basé sur des fonctions de croyance. In *21e Rencontres Francophones sur la Logique Floue et ses Applications, LFA 2012*, pages 71–78, novembre 2012.
- [10] M. BOU FARAH, D. MERCIER, É. LEFÈVRE et F. DELMOTTE : Exchanging dynamic and imprecise information in V2V networks with belief functions. In *Proceedings of the 16th International IEEE Conference on Intelligent Transport Systems, ITSC 2013, The Hague, The Netherlands*, pages 967–972, octobre 2013.
- [11] M. BOU FARAH, D. MERCIER, É. LEFÈVRE et F. DELMOTTE : A high-level application using belief functions for exchanging and managing uncertain events on the road in vehicular ad-hoc networks. *Annals of telecommunications, special issue on belief functions and uncertainty management in networks and telecommunication*, 69:185–199, avril 2014.

- [12] F. BOURGEOIS et J-C. LASSALLE : An Extension of the Munkres Algorithm for the Assignment Problem to Rectangular Matrices. *Communication. ACM*, 14(12):802–804, décembre 1971.
- [13] A. CEBAN : *Méthode globale de diagnostic des machines électriques*. Thèse de doctorat, Université d'Artois, 2012.
- [14] A. CEBAN, R. PUSCA et R. ROMARY : Study of Rotor Faults in Induction Motors Using External Magnetic Field Analysis. *IEEE Transactions on Industrial Electronics*, 59(5):2082–2093, mai 2012.
- [15] V. CHERFAOUI, T. DENŒUX et Z.L. CHERFI : Distributed data fusion : application to confidence management in vehicular networks. *In Proceedings of the 11th International Conference on Information Fusion, FUSION 2008*, pages 846–853, Cologne, Germany, 2008.
- [16] B.R. COBB et P.P. SHENOY : On the plausibility transformation method for translating belief function models to probability models. *International Journal of Approximate Reasoning*, 41(3):314–330, April 2006.
- [17] M.H. DEGROOT : *Optimal Statistical decisions*. McGraw-Hill, New York, 1970.
- [18] A.P. DEMPSTER : Upper and lower probabilities induced by a multiple valued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
- [19] A.P. DEMPSTER : A generalization of bayesian inference. *Journal of the Royal Statistical Society*, 30:205–247, 1968.
- [20] T. DENŒUX : Teaching and Talks on belief functions : <https://www.hds.utc.fr/~tdenoeux/>.
- [21] T. DENŒUX : A k-nearest neighbor classification rule based on dempster-shafer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(5):804–813, 1995.
- [22] T. DENŒUX : Analysis of evidence-theoretic decision rules for pattern classification. *Pattern Recognition*, 30(7):1095–1107, 1997.
- [23] T. DENŒUX : Conjunctive and disjunctive combination of belief functions induced by non distinct bodies of evidence. *Artificial Intelligence*, 172(2–3):234–264, 2008.
- [24] T. DENŒUX, N. EL ZOGHBY, V. CHERFAOUI et A. JOUGLET : Optimal Object Association in the Dempster-Shafer Framework. *IEEE Transactions on Cybernetics*, 44(12):2521–2531, décembre 2014.
- [25] T. DENŒUX et P. SMETS : Classification using belief functions : the relationship between the case-based and model-based approaches. *IEEE transactions on Systems, Man and Cybernetics B*, 36(6):1395–1406, 2006.
- [26] D. DUBOIS et H. PRADE : A set-theoretic view of belief functions : logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12:193–226, 1986.

- [27] D. DUBOIS et H. PRADE : The principle of minimum specificity as a basis for evidential reasoning. In B. BOUCHON et R.R. YAGER, éditeurs : *Uncertainty in Knowledge-Based Systems*, pages 75–84. Springer Verlag, 1987.
- [28] D. DUBOIS et H. PRADE : Focusing vs. belief revision : A fundamental distinction when dealing with generic knowledge. In D.M. GABBAY, R. KRUSE, A. NONNENGART et H. OHLBACH, éditeurs : *Proceedings of the 4th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU 1997*, volume 1244 de *Lecture Notes in Computer Science*, pages 96–107, Bad Honnef, Germany, juin 1997. Springer Berlin Heidelberg.
- [29] N. EL ZOGHBY, V. CHERFAOUI, B. DUCOURTHIAL et T. DENÈUX : Distributed Data fusion for detecting Sybil attacks in VANETs Belief Functions : Theory and Applications. In T. DENÈUX et M.H. MASSON, éditeurs : *Proceedings of the 2nd International Conference on Belief Functions, BELIEF 2012*, volume 164 de *Advances in Intelligent and Soft Computing*, pages 351–358, Compiègne, France, mai 2012. Springer Berlin Heidelberg.
- [30] Z. ELOUEDI, K. MELLOULI et P. SMETS : Assessing Sensor Reliability for Multisensor Data Fusion Within the Transferable Belief Model. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, 34(1):782–787, 2004.
- [31] S. FABRE, A. APPRIOU et X. BRIOTTET : Presentation and description of two classification methods using data fusion based on sensor management. *Information Fusion*, 2:49–71, 2001.
- [32] D. GRUYER et V. CHERFAOUI : Matching and decision for vehicle tracking in road situation. In *Proceedings of International Conference on Intelligent Robots and Systems, IROS 1999*, volume 1, pages 29–34 vol.1, Kyongju, Korea, octobre 1999.
- [33] D. GRUYER, C. ROYÈRE, R. LABAYRADE et D. AUBERT : Credibilistic multisensor fusion for real time application, application to obstacle detection and tracking. In *Proceedings of IEEE International Conference on Advanced Robotics, ICAR 2003*, pages 1462—1467, Coimbra, Portugal, juin 2003.
- [34] S. HACHOUR : *Suivi et classification d’objets multiples : contributions avec la théorie des fonctions de croyance*. Thèse de doctorat, Université d’Artois, 2015.
- [35] S. HACHOUR, F. DELMOTTE et D. MERCIER : A distributed solution for multi-object tracking and classification. In *Proceedings of the 17th International Conference on Information Fusion, FUSION 2014, paper 323*, Salamanca, Spain, juillet 2014.
- [36] S. HACHOUR, F. DELMOTTE et D. MERCIER : A new parameterless credal method to track-to-track assignment problem. In F. CUZZOLIN, éditeur : *Proceedings of the 3rd International Conference on Belief Functions, BELIEF 2014*, volume 8764 de *Lecture Notes in Computer Science*, pages 403–411, Oxford, United Kingdom, septembre 2014. Springer International Publishing.
- [37] S. HACHOUR, F. DELMOTTE et D. MERCIER : Comparison of Credal Assignment Algorithms in Kinematic Data Tracking Context. In A. LAURENT,

- O. STRAUSS, B. BOUCHON-MEUNIER et R.R. YAGER, éditeurs : *Proceedings of the 15th Information Processing and Management of Uncertainty in Knowledge-Based Systems International Conference, IPMU 2014*, volume 444 de *Communications in Computer and Information Science*, pages 200–211, Montpellier, France, juillet 2014. Springer International Publishing.
- [38] S. HACHOUR, F. DELMOTTE, D. MERCIER et É. LEFÈVRE : Multi-sensor multi-target tracking with robust kinematic data based credal classification. *In Proceedings of the 8th Workshop, Sensor Data Fusion : Trends, Solutions, Application, SDF 2013, paper Thu1130*, Bonn, Germany, octobre 2013.
- [39] S. HACHOUR, F. DELMOTTE, D. MERCIER et É. LEFÈVRE : Object tracking and credal classification with kinematic data in a multi-target context. *Information Fusion*, 20:174–188, novembre 2014.
- [40] H. HENAO, C. DEMIAN et G.A. CAPOLINO : A frequency-domain detection of stator winding faults in induction machines using an external flux sensor. *IEEE Transactions on Industry Applications*, 39:1272–1279, 2003.
- [41] A.-L. JOUSSELME et P. MAUPIN : Distances in evidence theory : Comprehensive survey and generalizations. *International Journal of Approximate Reasoning*, 53(2):118–145, 2012.
- [42] H. KATSUNO et A. MENDELZON : On the difference between updating a knowledge base and revising it. *In J.F. ALLEN, R. FIKES et E. SANDEWALL, éditeurs : Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning, KR 1991*, pages 387–394, Cambridge, MA, USA, avril 1991. Morgan Kaufmann.
- [43] F. KLAWONN et P. SMETS : The dynamic of belief in the transferable belief model and specialization-generalization matrices. *In B. D'Ambrosio D. DUBOIS, M. P. Wellman et Ph. SMETS, éditeurs : Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, pages 130–137, San Mateo, CA, 1992.
- [44] J. KOHLAS et P.-A. MONNEY : *A Mathematical Theory of Hints. An Approach to the Dempster-Shafer Theory of Evidence*, volume 425. Springer-Verlag, Berlin, 1995.
- [45] D. KRAJZEWICZ, J. ERDMANN, M. BEHRISCH et L. BIEKER : Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3-4):128–138, décembre 2012.
- [46] H. KUHN : The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955.
- [47] É. LEFÈVRE, F. PICHON, D. MERCIER, Z. ELOUEDI et B. QUOST : Estimation de méta-connaissances à partir de matrices de confusion pour la correction de fonctions de croyance. *In 23e Rencontres Francophones sur la Logique Floue et ses Applications, LFA 2014*, pages 115–122, octobre 2014.
- [48] Y. LEMERET, É. LEFÈVRE et D. JOLLY : Improvement of an association algorithm for obstacle tracking. *Information Fusion*, 9(2):234–245, 2008.

- [49] I. LEVI : *The enterprise of knowledge*. MIT Press, Kambridge, MA, USA, 1980.
- [50] N. MEGHERBI, S. AMBELLOUIS, O. COLOT et F. CABESTAING : Multimodal data association based on the use of belief functions for multiple target tracking. *In Proceedings on the 8th International Conference on Information Fusion, FUSION 2005*, pages 900—906, Philadelphia, USA, juillet 2005.
- [51] D. MERCIER : *Fusion d'informations pour la reconnaissance automatique d'adresses postales dans le cadre de la théorie des fonctions de croyance*. Thèse de doctorat, Université de Technologie de Compiègne, 2006.
- [52] D. MERCIER, G. CRON et B. BENYOUB : Method of merging postal OCR using credibility functions, European Patent Application EP1835444, 2007.
- [53] D. MERCIER, G. CRON, T. DENÈUX et M.-H. MASSON : Decision fusion for postal address recognition using belief functions. *Expert Systems with Applications*, 36:5643–5653, 2009.
- [54] D. MERCIER, T. DENÈUX et M.-H. MASSON : Belief function correction mechanisms. *Studies in Fuzziness and Soft Computing*, B. Bouchon-Meunier et al. (Eds.), 249:203–222, janvier 2010.
- [55] D. MERCIER, Z. ELOUEDI et É. LEFÈVRE : Sur l'affaiblissement d'une fonction de croyance par une matrice de confusion. *In 19e Rencontres Francophones sur la Logique Floue et ses Applications, LFA 2010*, pages 277–283, novembre 2010.
- [56] D. MERCIER, É. LEFÈVRE et F. DELMOTTE : Belief functions contextual discounting and canonical decompositions. *International Journal of Approximate Reasoning*, 53(2):146–158, 2012.
- [57] D. MERCIER, É. LEFÈVRE et D. JOLLY : Object association in the TBM framework, application to vehicle driving aid. *In Proceedings of the 6th International Symposium on Imprecise Probability : Theories and Applications, ISIPTA 2009*, pages 317–326, Durham, United Kingdom, juillet 2009.
- [58] D. MERCIER, É. LEFÈVRE et D. JOLLY : Association pour le suivi d'objets dans le cadre des fonctions de croyance, appliquée aux véhicules intelligents. *Revue des Nouvelles Technologies de l'Information*, E-18:25–50, février 2010.
- [59] D. MERCIER, É. LEFÈVRE et D. JOLLY : Object association with belief functions, an application with vehicles. *Information Sciences*, 181(24):5485–5500, décembre 2011.
- [60] D. MERCIER, É. LEFÈVRE, R. PUSCA, C. DEMIAN et R. ROMARY : Étude préliminaire de l'application de la fusion d'informations pour le diagnostic de défauts de bobinages de machines à courant alternatif. *In 21e Rencontres Francophones sur la Logique Floue et ses Applications, LFA 2012*, pages 87–94, novembre 2012.
- [61] D. MERCIER, F. PICHON, É. LEFÈVRE et F. DELMOTTE : Learning Contextual Discounting and Contextual Reinforcement from Labelled Data. *In S. DETERCKE et T. DENÈUX, éditeurs : Proceedings of the 13th European Conference*

- on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECS-QARU 2015*, volume 9161 de *Lecture Notes in Computer Science*, pages 472–481, Compiègne, France, juillet 2015. Springer International Publishing.
- [62] D. MERCIER, B. QUOST et T. DENÈUX : Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 9(2):246–258, avril 2008.
 - [63] F. PÉRISSE, É. LEFÈVRE, D. MERCIER, D. ROGER et E. MATÉO : Diagnostic based on high frequency resonances and information fusion. In *Proceedings of the 17th IEEE International Symposium on Electrical Insulation, ISEI 2008*, pages 628–631, Vancouver, Canada, juin 2008.
 - [64] F. PÉRISSE, D. MERCIER, É. LEFÈVRE et D. ROGER : Robust diagnostics of stator insulation based on high frequency resonances measurements. *IEEE Transactions on Dielectrics and Electrical Insulation*, 16(5):1496–1502, octobre 2009.
 - [65] F. PICHON, D. DUBOIS et T. DENÈUX : Relevance and truthfulness in information correction and fusion. *International Journal of Approximate Reasoning*, 53(2):159–175, 2012.
 - [66] F. PICHON, D. MERCIER, F. DELMOTTE et É. LEFÈVRE : Truthfulness in contextual information correction. In F. CUZZOLIN, éditeur : *Proceedings of the 3rd International Conference on Belief Functions, BELIEF 2014*, volume 8764 de *Lecture Notes in Computer Science*, pages 11–20, Oxford, United Kingdom, septembre 2014. Springer International Publishing.
 - [67] R. PUSCA, C. DEMIAN, D. MERCIER, É. LEFÈVRE et R. ROMARY : An improvement of a diagnosis procedure for AC machines using two external flux sensors based on a fusion process with belief functions. In *38th Annual Conference on IEEE Industrial Electronics Society, IECON 2012*, pages 5096–5101. IEEE, octobre 2012.
 - [68] R. PUSCA, R. ROMARY, A. CEBAN et J-F. BRUDNY : An Online Universal Diagnosis Procedure Using Two External Flux Sensors Applied to the AC Electrical Rotating Machines. *Sensors*, 10(11):10448–10466, novembre 2010.
 - [69] M. ROMBAUT : Decision in multi-obstacle matching process using the theory of belief. In *Advances in Vehicle Control and Safety, AVCS 98*, pages 63–68, 1998.
 - [70] M. ROMBAUT et V. CHERFAOUI : Decision making in data fusion using Dempster-Shafer’s theory. In *Proceedings of the 3th IFAC Symposium on Intelligent Components and Instrumentation for Control Applications*, pages 375–379, Annecy, France, juin 1997.
 - [71] G. SHAFFER : *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.
 - [72] P. SMETS : Belief functions : the disjunctive rule of combination and the generalized bayesian theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.

- [73] P. SMETS : The canonical decomposition of a weighted belief. *In Proceedings of the 14th International Joint Conference on Artificial Intelligence, (IJCAI'95)*, pages 1896–1901, San Mateo, CA, 1995. Morgan Kaufmann Publishers.
- [74] P. SMETS : The Transferable Belief Model for quantified belief representation. *In D. M. GABBAY et P. SMETS, éditeurs : Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 1, pages 267–301. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
- [75] P. SMETS : The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, 31(1–2):1–30, 2002.
- [76] P. SMETS : Decision making in the TBM : the necessity of the pignistic transformation. *International Journal of Approximate Reasoning*, 38(2):133–147, 2005.
- [77] P. SMETS : Analyzing the combination of conflicting belief functions. *Information Fusion*, 8(4):387–412, 2007.
- [78] P. SMETS et R. KENNES : The Transferable Belief Model. *Artificial Intelligence*, 66:191–243, 1994.
- [79] P. SMETS et B. RISTIC : Kalman filter and joint tracking and classification based on belief functions in the TBM framework. *Information Fusion*, 8(1):16–27, janvier 2007.
- [80] R.R. YAGER : The entailment principle for Dempster-Shafer granules. *International Journal of Intelligent Systems*, 1:247–262, 1986.
- [81] H. ZHU et O. BASIR : Extended discounting scheme for evidential reasoning as applied to MS lesion detection. *In P. SVENSSON et J. SCHUBERT, éditeurs : Proceedings of the 7th International Conference on Information Fusion, FUSION 2004*, pages 280–287, Stockholm, Sweden, juin 2004.

Corrections et fusion d'informations dans le cadre des fonctions de croyance. Applications.

Résumé : Ce document est une synthèse de mes travaux de recherche en vue de l'obtention du diplôme d'habilitation à diriger des recherches. Ces travaux ont pour cadre formel la théorie des fonctions de croyance (ou théorie de Dempster-Shafer) qui est un modèle de représentation des incertitudes généralisant le cadre des probabilités comme cas particulier. Dans ce modèle, un résumé de mes travaux concernant des mécanismes de correction de l'information fournie par une source est donné, de même qu'une application en classification. Nos contributions concernant trois applications en fusion d'informations avec des fonctions de croyance sont également exposées. Ces applications concernent la communication véhicule à véhicule, l'association de données avec des fonctions de croyance, et le diagnostic du vieillissement du bobinage d'un moteur électrique.

Mots clés : Fonctions de croyance, Théorie de Dempster-Shafer, Fusion d'informations, corrections contextuelles, Association de données, Aide à la décision, Applications.

Corrections and informaton fusion Belief functions and information fusion. Developments and applications.

Abstract : This document is a summary of research I have worked on to obtain the *habilitation à diriger des recherches* diploma. This work has been realized with the belief function theory (or Dempster-Shafer theory), a model for uncertainty representation which generalizes in particular the probability theory. In this model, a summary of our work related with correction mechanisms of the information provided by a source is given, as well as an application in classification. Our contributions in three applications in information fusion with belief functions are also exposed. These applications concern car to car communication, data association with belief functions, the diagnosis of the winding ageing of an electric machine.

Keywords : Belief functions, Dempster-Shafer theory, Information fusion, Contextual corrections, Data association, Decision support systems, Applications.