Approche générique de la gestion de l'incertain dans les processus de fusion multisenseur

Generic approach of the uncertainty management in multisensor fusion processes

Alain Appriou

ONERA, BP 72, 92322 Chatillon cedex
Tél. : 0146734919, Fax : 0146734163
Alain.Appriou@onera.fr

Manuscrit reçu le 30 avril 2004

Résumé et mots clés

Le problème de la gestion de l'incertain au sein des systèmes multisenseurs est abordé dans le cadre fédérateur de la théorie de l'évidence. Un opérateur de base appelé extension est élaboré, dont on montre qu'il permet de générer, sous une forme générale, les différents opérateurs utiles à la constitution d'une chaîne complète et cohérente de traitement d'observations incertaines multiples, depuis leur modélisation jusqu'à la prise de décision exigée. Cette formulation permet, en particulier, de raisonner sur des croyances exprimées individuellement dans des référentiels hétérogènes, sans avoir à les ramener dans un cadre de discernement commun. Les conditions particulières qui permettent de retrouver les opérateurs traditionnels sont précisées, et la mise en œuvre pratique de l'opérateur générique proposé est décrite. Quelques exemples illustrent le bon usage qui peut ainsi être fait du traitement de l'incertain.

Incertain, théorie de l'évidence, fusion de données, systèmes multisenseurs, aide à la décision.

Abstract and key words

The uncertainty management in multisensor systems is considered in the federative framework of the theory of evidence. A basic operator, named extension, is elaborated. It provides a general formulation of the different operators that constitute a complete and coherent processing of multiple uncertain observations, from their modeling up to the required decision making. This formulation allows to reason with believes individually expressed on heterogeneous frames of discernment, without using a common one. The particular conditions that lead to the traditional operators are specified, and the implementation of the new generic operator is described. A few examples illustrate a suitable management of uncertainty processing.

Uncertainty, theory of evidence, data fusion, multisensor systems, decision-making.

1. Les systèmes multisenseurs

L'association de capteurs multiples est prioritairement motivée par l'acquisition d'observations suffisamment complémentaires pour satisfaire des exigences croissantes en matière de robustesse des systèmes face au contexte et à l'environnement, d'acuité et de richesse de l'information exploitée, et de réactivité opérationnelle.

L'élaboration d'un processus de fusion des données adapté à ce besoin exige cependant une parfaite maîtrise des difficultés liées à deux grandes contraintes incontournables : d'une part la disparité des informations traitées, que ce soit en termes d'incertitude, d'imprécision, d'incomplétude, de fiabilité, de subjectivité, ou de pertinence, et d'autre part la complexité des tâches d'analyse et la sûreté des conclusions induites. Un certain nombre de fonctionnalités propres aux traitements de fusion requièrent en particulier des développements spécifiques :

- la modélisation et l'intégration dans un même formalisme d'informations de nature différente, pouvant chacune relever d'approches théoriques variées : observations, connaissances *a priori*, apprentissages, fiabilités des données ;

- l'association de données ambiguës, dans l'espace ou dans le temps ;

- la combinaison de sources disparates : prise en compte de référentiels distincts, intégration d'informations contextuelles, gestion de conflits, exploitation des dépendances, traitement dynamique, asynchronisme ;

- l'élaboration des architectures de traitement adaptées aux besoins, que ce soit en termes de niveau de fusion, de centralisation ou de distribution, de hiérarchisation, ou d'organisation fonctionnelle ;

- les principes, les critères, et les niveaux de décision, le processus devant en outre pouvoir gérer la valeur relative des décisions les plus pertinentes, compte tenu de l'information disponible.

La mise en œuvre conjointe d'un certain nombre de cadres théoriques appropriés est certainement le moyen le plus judicieux pour surmonter ces difficultés de façon circonstanciée. Parmi les approches envisageables, les théories de l'incertain permettent de façon privilégiée de repousser les limites des méthodes classiques en se conformant plus strictement au traitement de la seule information explicitement disponible. Nous allons nous attacher à montrer dans la suite que la théorie de l'évidence peut en particulier être considérée comme un cadre fédérateur dans lequel un formalisme général peut être élaboré pour répondre de façon coordonnée aux différentes préoccupations évoquées précédemment.

Ces développements bénéficient de réflexions antérieures qui avaient déjà permis d'établir la cohérence des opérateurs propres à la réalisation d'une chaîne complète de traitement pour la fusion multisenseur [APP 02] [APP 04]. Il est ici proposé l'élaboration d'un opérateur générique, appelé extension, propre notamment à manipuler toutes les relations plus ou moins bien définies entre des référentiels différents, afin d'exprimer dans un référentiel souhaité une information donnée sous une forme quelconque. Il conduit en particulier à une formulation générale d'un ensemble cohérent d'opérateurs qui permettent de raisonner à partir de propositions exprimées individuellement sur des référentiels hétérogènes, sans avoir à les ramener sur un référentiel commun, et ceci jusqu'à une prise de décision adaptable au contexte opérationnel. Quelques illustrations didactiques particulières de ces éléments mettent en évidence tout l'avantage qui peut être tiré des théories de l'incertain.

2. Notions de base

La théorie de l'évidence [SHA 76] repose sur la définition d'un cadre de discernement E, ensemble de I hypothèses exclusives et exhaustives $H_i (i \in [1,I])$. 2^E désigne alors l'ensemble des 2^I sous-ensembles de E, incluant l'ensemble vide.

2.1. Fonctions élémentaires

Trois fonctions, définies de 2^E sur l'intervalle [0,1], permettent de caractériser la vraisemblance de chacun des sous-ensembles de E :

- la fonction de masse $m(.)$, qui représente la vraisemblance attachée à l'un des singletons H_i appartenant au sous-ensemble de E en argument, sans qu'il soit néanmoins possible de distinguer entre ces singletons ; elle est telle que :

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

$$m(\varnothing) = 0 \tag{2}$$

- la fonction de croyance $Cr(.)$, qui peut être interprétée comme la vraisemblance minimale du sous-ensemble de E en argument, et qui est définie de façon unique à partir d'une fonction de masse donnée :

$$Cr(B) = \sum_{A \subseteq B} m(A) \tag{3}$$

- la fonction de plausibilité $Pl(.)$, qui peut être interprétée comme la vraisemblance maximale du sous-ensemble de E en argument, et qui est elle aussi reliée de façon biunivoque aux fonctions de masse et de croyance, respectivement par les relations :

$$Pl(B) = \sum_{A \cap B \neq \varnothing} m(A) \tag{4}$$

$$Pl(B) = 1 - Cr(\neg B) \tag{5}$$

où $\neg B$ désigne le complémentaire de B dans E.

Les éléments focaux d'une fonction de masse $m(.)$ sont par ailleurs les éléments A de 2^E tels que $m(A)$ n'est pas nul.

La théorie de l'évidence a été introduite au paragraphe 1 comme un cadre fédérateur des théories de l'incertain, car les autres théories, et notamment les principales que sont les probabilités et la théorie des possibilités, manipulent des notions qui ne sont que des cas particuliers des notions de la théorie de l'évidence que nous venons de rappeler. En effet, lorsque les éléments focaux d'une fonction de masse constituent une partition de E, ou un sous-ensemble de celle-ci, alors cette fonction de masse est une fonction de masse Bayésienne, qui est strictement identique aux fonctions de croyance et de plausibilité correspondantes, toutes ces fonctions étant alors identiques à la notion classique de fonction de probabilité. Lorsqu'au contraire les éléments focaux sont tous inclus les uns dans les autres, alors les fonctions de croyance et de plausibilité sont respectivement réduites aux notions de nécessité et de possibilité utilisées dans la théorie des possibilités. Il convient de noter qu'en conséquence, hormis un cas trivial d'intérêt très limité, les notions de probabilité et de possibilité sont en revanche parfaitement incompatibles.

Un intérêt majeur de la théorie de l'évidence est donc de permettre de traiter conjointement, dans un même cadre théorique, des informations probabilistes et des informations possibilistes, pourtant *a priori* incompatibles. Corrélativement, elle permet une interprétation plus large des données par rapport à chacune des deux théories probabiliste et possibiliste.

2.2. Conditionnement

Le conditionnement permet de prendre en compte une information selon laquelle il est certain que le sous-ensemble A de E est vérifié. Il transforme en conséquence une fonction de masse quelconque $m(.)$ sur E en une fonction de masse $m(./A)$ dont tous les éléments focaux sont inclus dans A, ceci en transférant la masse de chaque élément focal de $m(.)$ sur sa partie incluse dans A, et en renormalisant par rapport à la masse qui a ainsi pu être affectée :

$$m(B/A) = \frac{\sum\limits_{C \cap A = B} m(C)}{\sum\limits_{C \cap A \neq \varnothing} m(C)} \tag{6}$$

Cette expression peut également s'écrire à l'aide des fonctions de plausibilité, pour tout sous-ensemble B du cadre de discernement E :

$$Pl(B/A) = \frac{Pl(B \cap A)}{Pl(A)} \tag{7}$$

Le déconditionnement est l'opération inverse du conditionnement. Connaissant une fonction de masse $m(./A)$ sur un sous-ensemble A de E, le but est de reconstituer une fonction de masse complète $m(.)$ sur E. L'expression (6) met clairement en évidence l'indétermination du problème si l'on ne connaît que $m(./A)$, le nombre d'inconnues étant alors nécessairement supérieur au nombre d'équations. Il convient donc d'utiliser toute connaissance complémentaire disponible, et au-delà de rechercher la fonction $m(.)$ de spécificité minimale, c'est-à-dire qui affecte au maximum la masse non définie aux éléments focaux les plus grands possibles.

De façon plus pragmatique, l'inversion de (7) fournit directement la plausibilité de tous les sous-ensembles de A :

$$\forall B \subseteq A \quad Pl(B) = Pl(B/A)Pl(A) \tag{8}$$

Outre $Pl(./A)$, la reconstruction complète de $m(.)$ exige donc d'estimer $Pl(A)$, ainsi que la plausibilité de tous les sous-ensembles de E non inclus dans A. Dans le cas général où ces grandeurs sont totalement inconnues, le minimum de spécificité déjà évoqué nous conduit à les prendre égales à 1. Ceci définit entièrement la fonction de masse $m(.)$ ainsi obtenue par l'opérateur de déconditionnement classique :

$$\forall B \subseteq A \quad m(B \cup \neg A) = m(B/A) \tag{9}$$

2.3. Raffinement, grossissement

Un raffinement R associe à chaque hypothèse H_i^1 d'un cadre de discernement $E^1 = \{H_1^1,...,H_{I1}^1\}$ un sous-ensemble $R\left(H_i^1\right)$ d'un autre cadre de discernement $E^2 = \{H_1^2,...,H_{I2}^2\}$ tel que $\{R\left(H_1^1\right),...,R\left(H_{I1}^1\right)\}$ constitue une partition de E^2. L'opération de raffinement revient donc à considérer que chaque singleton H_i^1 de E^1 est lui-même représentatif d'un ensemble d'hypothèses plus détaillées $R\left(H_i^1\right)$. Une fonction de masse $m^1(.)$ définie sur E^1 fournit donc naturellement, par une opération dite d'extension minimale, une fonction de masse $m^2(.)$ sur E^2 :

$$\forall A \subseteq E^1 \quad m^2\left(R(A)\right) = m^1(A) \tag{10}$$

Le grossissement est l'opération inverse R^{-1} du raffinement R. Il consiste donc à regrouper les singletons du cadre de discernement E^2 en sous-ensembles exclusifs $R\left(H_i^1\right)$, qui sont alors associés aux singletons H_i^1 du cadre de discernement E^1. Une fonction de masse $m^2(.)$ définie sur E^2 conduit dans ces conditions à une fonction de masse $m^1(.)$ sur E^1 par la transformation :

$$m^1(A) = \sum_{\substack{B \subseteq E^2 \\ A = \left\{H_i^1 / R\left(H_i^1\right) \cap B \neq \varnothing\right\}}} m^2(B) \tag{11}$$

3. Élaboration d'un opérateur générique : l'extension

Sur les seules bases qui précèdent, nous allons ici nous attacher à établir un opérateur qui permet de déterminer sur un espace E_s la fonction de masse $m_s(.)$ résultant de la connaissance d'une fonction de masse $m_e(.)$ sur un espace E_e et des relations plus ou moins incertaines qui lient les éléments de E_s à ceux de E_e. Cette situation est par exemple celle où une observation nous fournirait une vraisemblance $m_e(.)$ sur un espace d'attributs E_e, et où l'on chercherait la vraisemblance $m_s(.)$, sur un espace de classes d'objets E_s, induite par la connaissance d'une description imprécise ou incertaine des objets selon les attributs observés. Nous allons voir que l'opérateur cherché est strictement équivalent aux opérateurs de conditionnement-déconditionnement et de raffinement-grossissement introduits au paragraphe précédent, dans le sens où il est établi à partir de ces seuls opérateurs, et que ceux-ci ne sont réciproquement que des mises en œuvres particulières de l'opérateur développé ici. Nous verrons par ailleurs dans la suite de l'article que l'opérateur obtenu permet de générer les différents opérateurs utiles en fusion de données, sous la forme la plus générale qu'il peut s'avérer souhaitable d'exprimer. Nous utiliserons dans la suite le terme d'extension pour désigner cet opérateur, par analogie avec le principe du même nom utilisé dans le cadre des sous-ensembles fous. Supposons plus précisément que nous disposons des données suivantes :

- $Pl_e(.)$, une fonction de plausibilité sur E_e, éventuellement incomplète ;

- $Pl_s\,(./B \subseteq E_e)$, une fonction de plausibilité sur E_s, éventuellement incomplète, valide lorsque le sous-ensemble B de E_e est avéré de façon certaine ; cette fonction de plausibilité peut être disponible pour un certain nombre de sous-ensembles B de E_e , mais pas nécessairement pour tous ; elle permet de désigner de façon incertaine les éléments de E_s qui sont vraisemblables dès lors que l'un ou l'autre des éléments de B $\subseteq E_e$ est vérifié ;

- $E_r \subseteq E_e$, l'ensemble des éléments de E_e qui sont susceptibles d'être associés à un élément quelconque de E_s ; il peut en effet être nécessaire de formaliser des incompatibilités entre certains éléments de E_e et certains éléments de E_s (par exemple certaines altitudes ou certaines vitesses pour des classes particulières de véhicules aériens) ; les éléments de E_s qui ne sont compatibles avec aucun élément de E_e peuvent être gérés par la mise à zéro des termes correspondants des $Pl_s(./B \subseteq E_e)$; en revanche les éléments de E_e qui ne sont compatibles avec aucun élément de E_s ne peuvent être pris en compte de cette façon, car une fonction de plausibilité ne peut pas être nulle partout, par définition ; ils doivent donc être identifiés dans $E_e - E_r$, et traités de façon appropriée.

3.1. Formulation de l'opérateur d'extension

L'opérateur envisagé peut être établi en commençant par chercher à évaluer sur le produit cartésien $E_s \times E_r$ la fonction de masse $m_{sr}(.)$ susceptible d'intégrer toute la connaissance disponible. Pour tout sous-ensemble A de E_s et tout sous-ensemble B de E_e, une fonction de plausibilité $Pl_{se}(.)$ sur $E_s \times E_e$ peut dans un premier temps être évaluée à partir de la décomposition évidente suivante :

$$Pl_{se}(A \times B) = Pl_{se}\left((A \times E_e) \cap (E_s \times B)\right) \tag{12}$$

Elle peut ainsi être conditionnée sur $E_s \times B$ par application de (7) :

$$Pl_{se}\,(A \times E_e/E_s \times B) = \frac{Pl_{se}(A \times B)}{Pl_{se}(E_s \times B)} \tag{13}$$

Or $Pl_{se}(E_s \times B)$ peut être obtenue à partir de la plausibilité $Pl_e(B)$ de B évaluée sur E_e, par raffinement de E_e sur $E_s \times E_e$. L'application de (10) aux fonctions de masse associées, pour $R(B) = E_s \times B$, procure ainsi :

$$Pl_{se}(E_s \times B) = Pl_e(B) \tag{14}$$

De même $Pl_{se}(A \times E_e/E_s \times B)$ peut être obtenue à partir de la plausibilité $Pl_s(A/E_s \times B)$ de A évaluée sur E_s sachant que $E_s \times B$ est réalisé de façon certaine, par raffinement de E_s sur $E_s \times E_e$. E_s étant par définition toujours réalisé, le conditionnement par $E_s \times B$ se résume en fait à un conditionnement par B dans E_e, et l'application de (10) aux fonctions de masse associées, pour $R(A) = A \times E_e$, donne :

$$Pl_{se}\,(A \times E_e/E_s \times B) = Pl_s\,(A/B \subseteq E_e) \tag{15}$$

L'expression de $Pl_{se}(A \times B)$ devient donc, à la lumière de (13), (14), et (15) :

$$Pl_{se}\,(A \times B) = Pl_s\,(A/B \subseteq E_e)\,Pl_e(B) \tag{16}$$

La fonction de plausibilité $Pl_{sr}(.)$ associée à la fonction de masse cherchée $m_{sr}(.)$ sur $E_s \times E_r$ est finalement obtenue en conditionnant de $Pl_{se}(.)$ de $E_s \times E_e$ sur $E_s \times E_r$ à l'aide de (7), la prise en compte de (16) amenant finalement :

$$Pl_{sr}(A \times B) = \frac{Pl_s(A/B \subseteq E_r)Pl_e(B)}{Pl_{se}(E_s \times E_r)}, \tag{17}$$

Ceci fournit, par application de (14) à $B = E_r$, le résultat attendu :

$$Pl_{sr}(A \times B) = \frac{Pl_s(A/B \subseteq E_r)Pl_e(B)}{Pl_e(E_r)}, \tag{18}$$

Il n'y a plus alors qu'à déterminer la fonction de masse $m_{sr}(.)$ associée à $Pl_{sr}(.)$. Cette dernière n'étant en général pas complète, $m_{sr}(.)$ est la fonction de masse de minimum de spécificité (maximum de masse sur les éléments focaux les plus grands) qui respecte les contraintes (18) disponibles. Cette détermination peut être réalisée à l'aide de l'algorithme présenté au paragraphe 3.2.

La connaissance ainsi élaborée peut enfin être ramenée sur E_s par grossissement de $E_s \times E_r$ sur E_s. Compte tenu de la relation particulière entre les ensembles concernés, la fonction de masse $m_s(.)$ qui en résulte sur E_s est simplement donnée par :

$$m_s(A) = \sum_{B \subseteq E_r} m_{sr}(A \times B) \tag{19}$$

En résumé, l'opérateur d'extension établi consiste, à partir des données listées en introduction du paragraphe 3, à réaliser les 3 opérations suivantes :

- Détermination des $Pl_{sr}(A \times B)$ à l'aide de (18), pour toutes les données disponibles ;

- Détermination de la fonction de masse $m_{sr}(.)$ de minimum de spécificité satisfaisant les valeurs de $Pl_{sr}(A \times B)$ obtenues à l'étape précédente, selon l'algorithme décrit au paragraphe 3.2 ;

- Détermination de la fonction de masse $m_s(.)$ sur E_s à partir de $m_{sr}(.)$, à l'aide de (19).

Il convient de noter, comme annoncé en introduction du paragraphe 3, qu'il est aisé de retrouver à partir de cet opérateur les opérateurs de conditionnement, de déconditionnement, de raffinement, et de grossissement, grâce à chaque fois à un choix approprié de E_s, E_e, E_r, et des $Pl_s(./B \subseteq E_e)$.

3.2. Algorithme de détermination de $m_{sr}(.)$ à partir des $Pl_{sr}(A \times B)$

Cet algorithme est en fait une méthode qui permet de résoudre le problème très général de détermination d'une fonction de masse $m(.)$ associée à une fonction de plausibilité $Pl(.)$ connue de façon incomplète sur un cadre de discernement E. La fonction $m(.)$ étant de ce fait indéterminée, on recherche parmi les fonctions possibles celle dont la spécificité $Sp(m)$ est minimale :

$$Sp(m) = \sum_{A \subseteq E} \frac{m(A)}{|A|} \tag{20}$$

Le principe de l'algorithme consiste à considérer au départ l'incertitude totale (toute la masse sur E), puis à satisfaire une à une les plausibilités données en redistribuant à chaque fois en conséquence tout ou partie de la masse attribuée à un élément focal A sur un sous-ensemble de cet élément A, A et ce sous-ensemble étant choisis de façon à ne provoquer qu'un accroissement minimal de $Sp(m)$ donné par (20).

Plus précisément, les différentes étapes de l'algorithme sont les suivantes :

- Initialiser $m(.)$ par $m(E) = 1$;

- Considérer successivement chaque $B_j \subseteq E$ de plausibilité connue, par ordre de cardinal décroissant ; à cardinal égal, peu importe leur ordre de prise en compte ;

 – Pour chaque B_j, calculer :

$$\Delta_j = \left(\sum_{\substack{A_i \cap B_j \neq \varnothing \\ A_i \subseteq E}} m(A_i) \right) - Pl(B_j) \tag{21}$$

 • Si $\Delta_j = 0$, passer au B_j suivant, s'il en reste ;

 • Si $\Delta_j > 0$, considérer l'élément focal A_i de $m(.)$ en l'état, tel que :

$$\begin{aligned} &A_i \cap B_j \neq \varnothing, \quad A_i - B_j \neq \varnothing, \\ &\left(\left|A_i - B_j\right|^{-1} - |A_i|^{-1} \right) \ \text{minimal} \end{aligned} \tag{22}$$

  • Si $\Delta_j > m(A_i)$, transférer la masse $m(A_i)$ sur $A_i - B_j$, recalculer Δ_j selon (21), puis traiter successivement de la même façon les A_i satisfaisant la condition (22) tant que $\Delta_j > m(A_i)$;

  • Si $\Delta_j = m(A_i)$, transférer la masse $m(A_i)$ sur $A_i - B_j$, puis passer au B_j suivant, s'il en reste ;

  • Si $\Delta_j < m(A_i)$, transférer une masse égale à Δ_j de A_i sur $A_i - B_j$, conserver $(m(A_i) - \Delta_j)$ sur A_i, puis passer au B_j suivant, s'il en reste ;

- Lorsque tous les B_j de plausibilité connue ont été pris en compte, $m(.)$ obtenue est la fonction de masse cherchée.

On notera que, selon le conditionnement du problème, la solution peut n'être pas unique.

À titre illustratif, les différentes étapes de cette méthode sont données pour un exemple numérique particulier. Soit la fonction de plausibilité incomplète $Pl(.)$ sur $E = \{H_1, H_2, H_3\}$, définie par :

$Pl(H_1) = 0{,}8;\ Pl(H_2) = 0{,}3;\ Pl(H_1 \cup H_2) = 0{,}9;$
$Pl(H_2 \cup H3) = 0{,}6$;

Les différentes évolutions de $m(.)$ sont les suivantes :

- Initialisation
 $m(E) = 1$;
- Prise en compte de $Pl(H_2 \cup H_3)$
 $m(E) = 0{,}6$; $m(H_1) = 0{,}4$;
- Prise en compte de $P_l(H_1 \cup H_2)$
 $m(E) = 0{,}5$; $m(H_1) = 0{,}4$; $m(H_3) = 0{,}1$;
- Prise en compte de $Pl(H_2)$
 $m(E) = 0{,}3$; $m(H_1 \cup H_3) = 0{,}2$; $m(H_1) = 0{,}4$; $m(H_3) = 0{,}1$;
- Prise en compte de $Pl(H_1)$
 $m(E) = 0{,}2$; $m(H_1 \cup H_3) = 0{,}2$;
 $m(H_2 \cup H_3) = 0{,}1$; $m(H_1) = 0{,}4$; $m(H_3) = 0{,}1$;

La solution obtenue est ici unique. Il est aisé de vérifier *a posteriori* qu'elle satisfait bien les quatre plausibilités imposées, et

que toute modification de $m(.)$ sous ces quatre contraintes conduit bien à une augmentation de $Sp(m)$ donnée par (20).
Pour la mise en œuvre de l'opérateur d'extension qui nous intéresse, il est clair qu'il suffit d'appliquer cet algorithme à $m(.) = m_{sr}(.)$ et $Pl(.) = Pl_{sr}(.)$.

3.3. Mises en œuvre immédiates de l'opérateur d'extension

En premier lieu, l'opérateur d'extension établi précédemment permet bien sûr de gérer la dépendance entre une grandeur X exprimée sur $E_e = E_x = \{X_1,...,X_{Nx}\}$ et une grandeur Y exprimée sur $E_s = E_y = \{Y_1,...,Y_{Ny}\}$. En particulier une information de vraisemblance sur la première permet d'élaborer une vraisemblance de l'autre, connaissant certaines relations pouvant les unir. Dans un problème de classification par exemple, X est typiquement un attribut (taille, forme, couleur,...) susceptible de caractériser des objets, et dont l'observation sur un objet présenté fournit une fonction de plausibilité $Pl_x(.)$ sur E_x. Corrélativement Y est alors une classe d'objets répertoriée dans E_y, et qu'une connaissance préalable a permis de plus ou moins caractériser en termes d'attribut X, sous la forme d'une plausibilité conditionnelle $Pl_y(./B \subseteq E_x)$ sur E_y (Par exemple, un petit objet dans l'espace aérien peut être un drone ou un missile, jamais un avion de ligne). L'opérateur proposé permet de déduire de tout cela la vraisemblance des différentes classes auxquelles l'objet observé peut appartenir.
Dans un problème d'évaluation, X et Y peuvent être toutes les deux des attributs (par exemple taille et vitesse), une information sur l'un permettant alors d'accéder à une connaissance sur l'autre grâce aux relations physiques qui les lient (par exemple : les véhicules les plus rapides sont en général les plus petits).
Un autre usage direct et important de l'opérateur d'extension concerne la mise à jour des connaissances dans le temps, typiquement pour évaluer une situation à partir d'une observation antérieure et d'une connaissance sur l'évolution possible des choses depuis cette observation. E_e et E_s représentent alors le même espace physique, mais considéré à deux instants différents. L'exemple didactique présenté au paragraphe 6 illustre cette utilisation de l'opérateur, et met en évidence tout l'intérêt des théories de l'incertain qui en résulte, par rapport à des approches plus classiques.
Enfin un usage un peu particulier de cet opérateur consiste à l'utiliser « à l'envers » pour identifier les dépendances entre 2 grandeurs. L'idée est de considérer d'une part une observation qui fournit la plausibilité conjointe $Pl_{se}(A \times B)$ sur E_{se}, et d'autre part une observation qui fournit la plausibilité $Pl_e(B)$ sur E_e seul. Grâce à (16) il est alors possible d'identifier $Pl_s(A/B \subseteq E_e)$, par exemple à des fins de classification ou d'évaluation ultérieures, telles que présentées plus haut.

4. Gestion de la fiabilité

Étant donnée une fonction de masse $m(.)$ définie sur un cadre de discernement E, l'opérateur précédent permet d'établir la prise en compte d'une confiance q, à valeur dans [0,1], que l'on est en droit d'accorder à $m(.)$. Il faut pour cela considérer le cadre de discernement $E_F = \{F, \neg F\}$, où F désigne le fait que $m(.)$ est fiable. La valeur $(1-q)$ représente en pratique la perte de confiance avérée, sans que toutefois aucune certitude ne soit jamais établie quant à la fiabilité effective de la source considérée ($q = 0$ signifie que des facteurs de nature à invalider l'évaluation $m(.)$ ont été constatés, mais $q = 1$ signifie simplement qu'aucun élément perturbateur n'a été observé, sans pour autant garantir avec certitude que $m(.)$ est parfaitement représentatif de la réalité). La confiance q doit donc être interprétée sur E_F sous la forme :

$$Pl_F(F) = q \tag{23}$$

Notons que la donnée de cette plausibilité comme seule contrainte du problème conduit à modéliser l'information de fiabilité par la fonction de masse de minimum de spécificité sur E_F qui la satisfait :

$$m_F(F) = 0\ ; \quad m_F(\neg F) = 1 - q\ ; \quad m_F(E_F) = q \tag{24}$$

Nous pouvons dans ce contexte appliquer directement l'opérateur d'extension du paragraphe 3.1 en considérant que :
- $E_e = E_r = E_F$;
- $E_s = E$;
- $Pl_s(./B \subseteq E_e)$ est connue pour $B = F, Pl(./F \in E_F)$ étant la fonction de plausibilité associée à $m(.)$ donnée sur $E_s = E$, puisque celle-ci ne doit bien sûr être considérée qu'à la condition qu'elle soit fiable ;
- $Pl_e(B)$ est connue pour $B = F$, et donnée par (23) sur $E_e = E_F$.

Nous obtenons ainsi immédiatement l'opérateur d'affaiblissement classique de la théorie de l'évidence, $m^q(.)$ désignant la fonction de masse finalement obtenue sur E :

$$\begin{aligned} &\forall A \subset E, \quad A \neq E, \quad m^q(A) = q\ m(A) \\ &m^q(E) = 1 - q(1 - m(E)) \end{aligned} \tag{25}$$

Une conclusion intéressante de ceci est que l'affaiblissement q mis en œuvre dans (25) peut être déterminé à partir de la connaissance d'une fonction de masse $m_F(.)$ sur E_F, de type (24).

5. Combinaison d'observations

5.1. Formulation générale de la combinaison

L'opérateur d'extension établi précédemment permet de proposer une formulation générale de la combinaison de sources distinctes. Supposons que deux grandeurs X et Y soient respectivement évaluées sur les cadres de discernement $E_x = \{X_1,...,X_{Nx}\}$ et $E_y = \{Y_1,...,Y_{Ny}\}$, et que l'on cherche à en déduire l'évaluation d'une grandeur Z appartenant à un cadre de discernement $E_z = \{Z_1,...,Z_{Nz}\}$, moyennant la connaissance préalable de relations entre la grandeur Z et les grandeurs X et Y. Typiquement X et Y sont des grandeurs discriminantes, et Z est une classe de situations pour laquelle les grandeurs X et Y possibles sont connues de façon plus ou moins sûre.

Nous pouvons dans ces conditions appliquer l'opérateur d'extension sur les données suivantes :

- $E_e = E_x \times E_y$;

- $E_r =$ ensemble E_{xy} des seuls couples (X_i, Y_h) admissibles de $E_x \times E_y$, selon le problème traité ;

- $E_s = E_z$;

- $Pl_e(B) = Pl_{xy}(B)$, résultat de l'évaluation obtenue pour $B \subseteq E_x \times E_y$ à partir de l'observation conjointe des grandeurs X et Y ; dans le cas le plus fréquent où les grandeurs X et Y sont indépendantes, elles peuvent être observées et évaluées séparément, fournissant respectivement $m_x(B_i)$ sur E_x et $m_y(B_h)$ sur E_y, et $Pl_{xy}(B)$ est définie par :

$$m_{xy}(B) = m_x(B_i)\, m_y(B_h) \tag{26}$$

pour $B = B_i \times B_h$

La connaissance des dépendances entre les grandeurs peut aussi permettre d'écrire directement, à partir d'évaluations distinctes sur E_x et E_y :

$$m_{xy}(B) = m_x(B_i / B_h \subseteq E_y)\, m_y(B_h) \tag{27}$$

pour $B = B_i \times B_h$

Ce peut être le cas, par exemple, lorsque le processus d'évaluation de X dépend de la réalisation de Y.

- $Pl_s(A/B \subseteq E_e) = Pl_z(A/B \subseteq E_x \times E_y)$, plausibilité sur E_z qui formalise la connaissance que l'on a concernant les relations susceptibles d'unir la grandeur Z (typiquement l'ensemble A de classes) et les grandeurs X et Y (typiquement l'ensemble B des couples d'attributs observés).

Le résultat fournit directement l'opérateur général de combinaison, qui consiste à effectuer les opérations suivantes :

- Détermination sur $E_z \times E_x \times E_y$ de :

$$Pl_{zxy}(A \times B) = \frac{Pl_z(A/B \subseteq E_{xy}) Pl_{xy}(B)}{Pl_{xy}(E_{xy})} \tag{28}$$

où $Pl_{xy}(B)$ est donnée par (26) si X et Y sont indépendantes, par (27) si leurs dépendances sont connues, ou déterminée par observation conjointe de X et Y ;

- Détermination sur $E_z \times E_x \times E_y$ de la fonction de masse $m_{zxy}(.)$ de minimum de spécificité satisfaisant les valeurs de $Pl_{zxy}(A \times B)$ obtenues à l'étape précédente, selon l'algorithme décrit au paragraphe 3.2 ;

- Détermination de la fonction de masse cherchée :

$$m_z(A) = \sum_{B \subseteq E_{xy}} m_{zxy}(A \times B) \tag{29}$$

Bien évidemment, l'opérateur général de combinaison présenté ici peut être généralisé à la combinaison d'un nombre quelconque de sources, supérieur à 2. Il est par ailleurs commutatif, les grandeurs X et Y jouant des rôles symétriques. Il n'est en revanche pas associatif sous sa forme générale, mais l'associativité peut être retrouvée pour certains des cas particuliers classiques mis en évidence au paragraphe 5.2 (par exemple pour la somme orthogonale, la règle disjonctive, ou la technique de *hedging*).

5.2. Cas particuliers

Ce formalisme général de combinaison peut être particularisé pour retrouver toutes les règles de combinaison de sources connues à ce jour. En particulier, pour les plus usitées :

- Si $E_x = E_y = E_z, Pl_z(A/B \subseteq E_x \times E_y) = 1$ si $\exists (X_i, Y_h) \in B$ et $Z_k \in A$, t.q. $X_i = Y_h = Z_k$, $Pl_z(A/B \subseteq E_x \times E_y) = 0$ si $\nexists (X_i, Y_k) \in B$ et $Z_k \in A$, t.q. $X_i = Y_h = Z_k$, et E_{xy} est l'ensemble des couples (X_i, Y_h) tels que $X_i = Y_h$, alors l'opérateur de combinaison correspond à la somme orthogonale :

$$m_z(A) = \frac{\sum_{B_i \cap B_h = A \neq \emptyset} m_x(B_i) m_y(B_h)}{1 - \sum_{B_i \cap B_h = \emptyset} m_x(B_i) m_y(B_h)} \tag{30}$$

- Si $E_x = E_y = E_z$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 1$ pour $A \cap (B_i \cup B_h) \neq \emptyset$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 0$ pour $A \cap (B_i \cup B_h) = \emptyset$, et E_{xy} est directement l'ensemble $E_x \times E_y$, alors l'opérateur de combinaison correspond à la règle disjonctive :

$$m_z(A) = \sum_{B_i \cup B_h = A} m_x(B_i) m_y(B_h) \tag{31}$$

- Si $E_x = E_y = E_z, Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 1$ pour $A \cap B_i \cap B_h \neq \emptyset$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 1$ pour $B_i \cap B_h = \emptyset$ et $A \cap (B_i \cup B_h) \neq \emptyset$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 0$ pour $B_i \cap B_h \neq \emptyset$ et $A \cap B_i \cap B_h = \emptyset$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 0$ pour $B_i \cap B_h = \emptyset$ et $A \cap (B_i \cup B_h) = \emptyset$, et E_{xy} est directement l'ensemble $E_x \times E_y$, alors l'opérateur de combinaison correspond à la règle mixte de Dubois et Prade :

$$m_z(A) = \sum_{B_i \cap B_h = A} m_x(B_i) m_y(B_h) + \sum_{\substack{B_i \cup B_h = A \\ B_i \cap B_h = \emptyset}} m_x(B_i) m_y(B_h) \tag{32}$$

- Si $E_x = E_y = E_z$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 1$ pour $A \cap B_i \cap B_h \neq \emptyset$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 1$ pour $B_i \cap B_h = \emptyset$ et A quelconque, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 0$ pour $B_i \cup B_h \neq \emptyset$ et $A \cap B_i \cap B_h = \emptyset$, et E_{xy} est directement l'ensemble $E_x \times E_y$, alors l'opérateur de combinaison correspond à la règle de Yager :

$$m_z(A) = \sum_{B_i \cap B_h = A \neq \varnothing} m_x(B_i) m_y(B_h)$$
$$m_z(E_z) = m_z(E_z) + \sum_{B_i \cap B_h = \varnothing} m_x(B_i) m_y(B_h) \tag{33}$$

- Si $E_x = E_y$, $E_z = E_x + \{e\}$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 1$ pour $A \cap ((B_i \cap B_h) \cup \{e\}) \neq \varnothing$, $Pl_z(A/B_i \times B_h \subseteq E_x \times E_y) = 0$ pour $A \cap ((B_i \cap B_h) \cup \{e\}) = \varnothing$, et E_{xy} est directement l'ensemble $E_x \times E_y$, alors l'opérateur de combinaison correspond à la technique de *hedging* :

$$m_z(A \cup e) = \sum_{B_i \cap B_h = A \neq \varnothing} m_x(B_i) m_y(B_h)$$
$$m_z(e) = \sum_{B_i \cap B_h = \varnothing} m_x(B_i) m_y(B_h) \tag{34}$$

Il convient par ailleurs de remarquer que, sous les hypothèses du monde ouvert proposé par Ph. Smets ($m(\varnothing) \neq 0$), il suffit d'autoriser E_{xy} à être directement l'ensemble $E_x \times E_y$, dans les conditions propres à la somme orthogonale, pour retrouver la règle non normalisée correspondant à ce contexte.

5.3. Le conflit

Le conflit, c'est-à-dire la contradiction qui peut apparaître entre deux sources, peut avoir des origines diverses, bien précises. Il doit donc être diagnostiqué et traité de façon appropriée. Globalement, les causes de conflit sont liées, soit à un problème de fiabilité des sources, soit à l'utilisation d'un cadre de discernement mal adapté. Les différentes règles de combinaison connues à ce jour, éventuellement associées à un affaiblissement circonstancié, fournissent les compromis souhaitables pour toutes les causes de conflit identifiables. Elles permettent en particulier de gérer l'information conflictuelle tout en conservant un résultat de combinaison qui reste le plus informatif compte tenu des conditions rencontrées. Une synthèse complète sur ce sujet est discutée dans [APP 02]. Au-delà et en support de ce savoir-faire, il convient de noter que l'opérateur général de combinaison présenté ici offre un cadre privilégié pour mettre en œuvre une telle gestion du conflit. D'une part toute contradiction peut être analysée dans les meilleures conditions sur $Pl_{xy}(B)$, et $Pl_{xy}(E_{xy})$ en fournit une mesure immédiate. D'autre part les paramètres de la combinaison peuvent être adaptés par ajustement au mieux de E_{xy}, E_z, et des $Pl_z(A/B \subseteq E_x \times E_y)$, sans oublier la gestion de la fiabilité des sources. Ceci peut en particulier permettre de retrouver automatiquement la plus appropriée des règles classiques de combinaison, mais surtout de dégager une combinaison hybride plus satisfaisante. Le développement de telles procédures dépasse néanmoins le cadre du présent article.

6. Exemple d'application à la mise à jour des connaissances

Nous allons illustrer, à l'aide de l'exemple qui suit, la mise en œuvre de l'opérateur d'extension et de certains de ses opérateurs associés, et ainsi mettre en évidence le bénéfice apporté par les théories de l'incertain, tel qu'il permet de l'exploiter.

Imaginons que l'on s'intéresse à la classification de pixels d'images du sol entre 4 labels : sable, eau, champs, forêt. Nous disposons de deux images d'un même lieu relevées à des dates différentes, et nous nous intéressons à un pixel en bordure d'une rivière, qui correspond à du sable lors de la première prise d'image, mais qui correspond à de l'eau lors de la deuxième prise d'image, suite à une crue de la rivière intervenue entre temps. Corrélativement, l'interprétation de la première image conduit pour notre pixel aux probabilités de classification suivantes :

P(sable) $= 0{,}50$; P(eau) $= 0{,}05$; P(champs) $= 0{,}40$; et P(forêt) $= 0{,}05$;

où le sable est bien le plus probable, mais où un risque de confusion avec les champs apparaît. L'interprétation de la deuxième image pour le même pixel conduit naturellement à des probabilités de classification différentes :

P(sable) $= 0$; P(eau) $= 0{,}50$; P(champs) $= 0{,}45$; et P(forêt) $= 0{,}05$;

L'eau est cette fois bien la plus probable, mais toujours avec un risque de confusion avec les champs, détrempés à cette époque. L'objectif est donc de fusionner les deux images, sachant que le sable a pu devenir de l'eau entre les deux prises d'images, pour conforter le fait que notre pixel est bien de l'eau à l'époque de la deuxième prise d'image, en diminuant le risque de confusion avec les champs. Ce problème s'inscrit en fait dans un processus itératif où l'on souhaite à chaque étape prédire l'état d'un pixel compte tenu des observations antérieures, et mettre à jour la connaissance ainsi obtenue à l'aide de l'observation disponible à cette étape. L'objectif est donc de réaliser une sorte de filtrage de Kalman qui porterait sur des descriptions autres que l'état de la dynamique d'une cible.

À titre de référence, nous allons dans un premier temps considérer l'approche probabiliste adaptée au problème. Cette approche procède donc en deux étapes :

- Prédiction : l'interprétation de la première image est ramenée à la date de la deuxième image en considérant que ce qui était du sable peut maintenant indifféremment être du sable ou de l'eau ; ceci conduit à répartir la probabilité 0,5 du sable en 0,25 sur l'eau et 0,25 sur le sable, pour fournir la prédiction de l'interprétation de la première image à la date de la seconde :

P(sable) $= 0{,}25$; P(eau) $= 0{,}30$; P(champs) $= 0{,}40$; et P(forêt) $= 0{,}05$;

- Mise à jour par fusion de la prédiction issue de la première image avec l'interprétation de la seconde, qui procure le résultat final :

P(sable) $= 0$; P(eau) $= 0{,}45$; P(champs) $= 0{,}54$; et P(forêt) $= 0{,}01$;

Nous voyons que cette procédure conduit à une erreur de classification, puisque la classe « champs » apparaît comme la plus probable. La fusion n'a donc ici fait que renforcer la confusion. Nous allons maintenant considérer la théorie de l'évidence, et la mise en œuvre de l'opérateur d'extension introduit précédemment, les probabilités données étant simplement considérées comme des fonctions de masse Bayésiennes (éléments focaux disjoints, *cf.* paragraphe 2.1). Cette approche suit les deux étapes identifiées pour les probabilités, seuls les opérateurs étant différents :

- Prédiction : l'interprétation de la première image est ramenée à la date de la deuxième image en considérant que ce qui était du sable peut maintenant indifféremment être du sable ou de l'eau ; l'application de l'opérateur générique conduit à transférer la masse 0,5 du sable sur l'élément focal (sable ∪ eau), pour fournir la prédiction de l'interprétation de la première image à la date de la seconde :

m(sable ∪ eau) $= 0{,}5$; m(eau) $= 0{,}05$; m(champs) $= 0{,}40$; et m(forêt) $= 0{,}05$;

- Mise à jour par combinaison conjonctive de la prédiction issue de la première image avec l'interprétation de la seconde, qui procure le résultat final :

m(sable) $= 0$; m(eau) $= 0{,}60$; m(champs) $= 0{,}39$; et m(forêt) $= 0{,}01$;

Cette procédure conduit cette fois non seulement à une bonne reconnaissance de la classe « eau », mais de surcroît à la baisse escomptée du risque de confusion avec les champs.
Cet exemple montre en particulier l'intérêt de pouvoir conserver toute la vraisemblance disponible sur un ensemble d'hypothèses indiscernables, sans être obligé de la répartir entre ces hypothèses, et donc de diminuer les chances de chacune. Il montre également l'aptitude de l'opérateur d'extension à exploiter cet avantage de la théorie de l'évidence.

7. Décision : recherche des hypothèses les plus vraisemblables

L'étape ultime de la fusion des informations multisenseurs consiste à reconnaître, à partir des croyances élaborées, la nature de la situation observée (nombre de cibles présentes, identité, caractérisation, ou localisation d'une cible particulière,...).

7.1. Procédure de décision

Plutôt que de chercher de façon classique, mais souvent abusive, l'hypothèse unique la plus vraisemblable parmi l'ensemble des hypothèses *a priori* envisageables, l'objectif est ici de sélectionner le sous-ensemble A, dans le cadre de discernement E sur lequel est finalement exprimée l'évaluation qui résulte de la mise en œuvre des opérateurs précédents, qui a le plus de chances de contenir la bonne hypothèse. Afin de ne pas déclarer systématiquement le cadre de discernement total, il convient néanmoins que le sous-ensemble A choisi réalise également un compromis avec un critère de cardinal $|A|$ minimal.
En conséquence, nous devons considérer comme entrées du processus de décision, d'une part la fonction de masse $m(.)$ établie sur $E = \{H_i\}$ à partir des observations, et d'autre part une fonction de masse $m_d(.)$ définie sur $E_d = \{d_A\}$, où d_A représente le choix respectivement de chaque sous-ensemble A de E. Cette fonction de masse $m_d(.)$ doit rassembler tous les souhaits *a priori* concernant la décision escomptée. C'est typiquement une fonction de masse Bayésienne exprimée sur E_d par :

$$m_d(d_A) = K_d \lambda_A g(|A|) \tag{35}$$

La fonction $g(.)$ est une fonction monotone décroissante quelconque du cardinal du sous-ensemble A évalué, qui permet d'éviter de déclarer l'ensemble E dans sa totalité en ajustant un compromis entre la certitude et la spécificité de l'ensemble des hypothèses retenues. Elle peut préférentiellement revêtir la forme :

$$g(|A|) = \left(\frac{1}{|A|}\right)^r \tag{36}$$

où r est un paramètre appartenant à [0,1] qui permet de choisir une attitude parmi un *continuum* de principes de décision allant du choix d'un singleton $(r = 1)$ à l'indécision totale $(r = 0)$.
Le coefficient λ_A intègre le manque de connaissance sur l'une quelconque des hypothèses appartenant à A, afin d'éviter de choisir systématiquement les hypothèses sur lesquelles peu d'information est disponible. La constante K_d est un facteur de normalisation qui assure la compatibilité avec la notion de fonction de masse.
Les fonctions de masse $m(.)$ et $m_d(.)$ peuvent être combinées à l'aide de l'opérateur de combinaison du paragraphe 5.1, en y imposant simplement que le sous-ensemble A déclaré contienne l'hypothèse H_i représentative de la réalité observée :

- $E_x = E$; $E_y = E_d$; $E_z = E_d$; $E_{xy} = \{(H_i, d_A) \in E \times E_d / H_i \in A\}$;

- $Pl_{xy}(B)$ est déterminée à partir de (26) pour $m_x(\cdot) = m(\cdot)$ et $m_y(\cdot) = m_d(\cdot)$;

- $Pl_z(A/B \subseteq E_x \times E_y) = Pl_d(D/V \times C \subseteq E \times E_d) = 1$ si $\exists\, d_A \in D \cap C$ et $H_i \in V$, tels que $H_i \in A$, et $Pl_z(A/B \subseteq E_x \times E_y) = Pl_d(D/V \times C \subseteq E \times E_d) = 0$ ailleurs.

Ceci fournit une fonction de masse Bayésienne sur E_d, dont il suffit de rechercher le maximum pour déterminer la décision optimale d_A^* qui s'exprime en conséquence :

$$d_a^* = \arg\left(\max_{A\subset E}(m_d(d_A)Pl(A))\right) \tag{37}$$

où $m_d(.)$ est bien sûr donnée par (35).
En fait, cette procédure peut conduire à des ambiguïtés de décision qui doivent prioritairement faire l'objet d'une recherche paramétrique de la solution de cardinal supérieur le plus proche qui ne soit plus ambiguë (action sur r de (36)). Si cette démarche ne s'avère pas satisfaisante, les solutions équivalentes obtenues peuvent être départagées en cherchant parmi elles le sous-ensemble A dont le complémentaire est de surcroît le moins vraisemblable, au sens du développement qui a conduit à (37) :

$$d_a^* = \arg\left(\min_{A\subset E}(m_d(d_{\neg}A)Pl(\neg A))\right) \tag{38}$$

Tout ceci laisse la liberté d'ajuster un compromis entre vraisemblance et précision pour l'ensemble A des hypothèses finalement déclarées. Il convient par ailleurs de noter que cette méthode dégage bien l'ensemble des singletons qui a le plus de chances de contenir la bonne solution, et non l'ensemble des singletons qui ont individuellement le plus de chances d'être la bonne solution, seul accessible par les approches classiques. Corrélativement, un ensemble de singletons dégagés ici peut être plus petit que l'union de solutions ambiguës de cardinal moindre, voire contenir des singletons différents. Enfin cette approche met à l'abri de déclarations abusives en regard de l'information disponible, en ne levant que les ambiguïtés qui peuvent l'être légitimement. Ces différents points sont illustrés dans les exemples qui suivent.

7.2. Quelques exemples pratiques

Les trois exemples qui suivent mettent en évidence les comportements annoncés de la procédure de décision qui vient d'être exposée. Les différentes solutions qu'elle est en mesure de fournir dans chaque cas sont comparées entre elles, puis comparées à une technique de référence souvent utilisée dans les réalisations pratiques, à savoir la recherche du singleton H_i qui maximise la probabilité pignistique :

$$P(H_i) = \sum_{\substack{A\subseteq E\\ H_i\in A}} \frac{m(A)}{|A|} \tag{39}$$

7.2.1. Premier exemple

Considérons sur $E = \{H_1, H_2, H_3\}$ la fonction de masse définie par :

$m(H_1) = 0{,}2$; $m(H_2) = 0{,}1$; $m(H_3) = 0.3$;

$m(H_1 \cup H_2) = 0{,}2$; $m(H_2 \cup H_3) = 0{,}1$;

$m(E) = 0{,}1$;

Lorsque tous les λ_A sont égaux à 1, si la décision est contrainte par $g(.)$ à déclarer un singleton, alors il en résulte une ambiguïté totale entre les trois singletons de E, ceux-ci ayant tous les trois la même plausibilité 0,5. En revanche, la solution de cardinal 2 la plus plausible est unique, $H_1 \cup H_3$, et de plausibilité bien supérieure, 0,9. De surcroît, si cela est nécessaire, le critère secondaire (38) permet de lever l'ambiguïté entre les singletons en déclarant un singleton unique, $H_3(Pl(H_1 \cup H_2)$ minimale). Cette dernière solution est conforme à la solution imposée par le critère pignistique (39).

7.2.2. Deuxième exemple

Sur le même cadre de discernement, avec toujours tous les λ_A égaux à 1, analysons cette fois la fonction de masse définie par :

$m(H_1) = 0{,}1$; $m(H_3) = 0{,}3$;

$m(H_1 \cup H_2) = 0{,}5$; $m(H_2 \cup H_3) = 0{,}1$;

La recherche du singleton unique le plus vraisemblable nous conduit ici à une ambiguïté entre H_1 et H_2, tous les deux de plausibilité 0,6. En revanche, la solution de cardinal 2 est unique, mais il s'agit de $H_1 \cup H_3$, de plausibilité 1, et non de $H_1 \cup H_2$ comme l'analyse des seuls singletons l'aurait laissé penser. De façon tout à fait cohérente, la levée d'ambiguïté entre les singletons à l'aide du critère secondaire (38) procure H_1 comme solution unique. Dans cette situation le critère pignistique conserve une ambiguïté entre H_1 et H_3 qui ne peut être levée, contrairement à ce que permet la méthode proposée, mais reste néanmoins cohérent avec les résultats de cette dernière.

7.2.3. Troisième exemple

Nous pouvons enfin, sur un dernier exemple à la fois très simple à interpréter, mais aussi tout à fait caractéristique de l'esprit de la théorie de l'évidence, nous convaincre de la légitimité de la méthode proposée ici. Toujours sur le même cadre de discernement, avec toujours tous les λ_A égaux à 1, considérons maintenant la fonction de masse suivante :

$m(H_1) = 0{,}4$;

$m(H_2 \cup H_3) = 0{,}6$;

Si l'on se rappelle la signification de la notion de masse introduite au paragraphe 2.1, cette fonction de masse exprime simplement qu'il est plus vraisemblable qu'il s'agisse de H_2 ou H_3 que de H_1, sans toutefois qu'il soit possible de distinguer entre H_2 et H_3.
Il est immédiat de vérifier que la solution de cardinal 1 fournit bien une ambiguïté entre H_2 et H_3, sans que, ni le passage à une solution de cardinal supérieur, ni l'utilisation du critère secondaire (38), ne permette de modifier d'une façon quelconque cette conclusion. La procédure de décision adoptée traduit donc bien strictement le contenu de la fonction de masse testée. En revanche le critère pignistique conduit dans les mêmes conditions à déclarer de façon erronée le singleton H_1.

8. Modélisation des données

Les différents opérateurs présentés précédemment exigent, pour pouvoir être appliqués aux observations fournies par des senseurs, que ces dernières soient exprimées sous la forme de fonctions de masse. Supposons que l'information à modéliser revête globalement la forme très générale d'un ensemble de vraisemblances $V_h \in [0,1]$, $h \in [1,H]$, chacune respectivement accordée à une proposition P_i, $i \in [1,I]$, compte tenu du signal s_j, $j \in [1,J]$, relevé par un des senseurs disponibles. Il est clair qu'un senseur peut permettre d'évaluer une ou plusieurs propositions, sans forcément toutes les évaluer, mais qu'il peut également n'en traiter aucune, par exemple s'il est indisponible. Plusieurs senseurs peuvent par ailleurs bien sûr traiter la même proposition. Les développements qui suivent n'ayant néanmoins de sens que si au moins une vraisemblance est disponible, nous avons en pratique $1 \leq H \leq IJ$.

Concrètement une proposition P_i peut prendre les formes les plus diverses imposées par chaque application. Elle peut par exemple être relative à l'identité d'une cible (« la cible analysée est un hélicoptère », ou « la cible détectée est une plate-forme amie »), au nombre de cibles présentes dans l'espace aérien (« 5 cibles sont présentes », ou « les cibles présentes sont au moins au nombre de 3 »),... Il est très important de noter que, contrairement aux fondements des travaux antérieurs [APP 02] [APP 04], l'ensemble des propositions P_i n'a ici aucune raison de constituer une description exclusive et exhaustive du monde traité, c'est-à-dire qu'il ne constitue pas un ensemble d'hypothèses exclusives et exhaustives tel qu'habituellement exigé par l'emploi de la théorie de l'évidence.

Un facteur de fiabilité $f_h \in [0,1]$ est par ailleurs associé à chaque vraisemblance V_h. Il a pour but d'exprimer l'aptitude de cette vraisemblance V_h à caractériser la proposition P_i sur laquelle elle porte, compte tenu de la qualité de la connaissance disponible. Il peut en particulier, par exemple, figurer la plus ou moins bonne représentativité de la connaissance *a priori* utilisée pour élaborer V_h, en regard du contexte opérationnel rencontré, ou l'observabilité du signal utilisé. Il est typiquement élaboré à partir de l'observation z_k d'une variable contextuelle discriminante quant à la proposition F_h : « V_h est fiable », voire d'un vecteur de variables contextuelles.

8.1. Prise en compte de la vraisemblance V_h et de sa fiabilité f_h

En pratique V_h a généralement une valeur de réfutation, que nous retiendrons dans la suite, c'est-à-dire que $(1 - V_h)$ représente la certitude attachée au fait que la proposition P_i n'est pas vérifiée. Prenons par exemple le problème classique de reconnaissance de cible où la proposition P_i serait « La cible observée est un F16 », et où la vraisemblance V_h serait établie en comparant des attributs mesurés sur la cible observée à ceux d'un F16, tels qu'ils sont connus. Il est clair que plus les attributs de la cible observée diffèrent de ceux d'un F16 (V_h décroissante), plus il est certain qu'il ne s'agit pas d'un F16. En revanche si la ressemblance est grande (V_h élevée), il n'est pas pour autant certain qu'il s'agisse d'un F16 : il peut y avoir confusion avec une autre cible ayant fortuitement les mêmes attributs. Il apparaît donc clairement que dans ces conditions V_h doive être interprétée comme la plausibilité de P_i sur le cadre de discernement $E_i = \{P_i, \neg P_i\}$. La connaissance apportée par V_h se résume donc à :

$$Pl_h(P_i) = V_h \;\text{ sur }\; E_i = \{P_i, \neg P_i\} \tag{40}$$

Ceci conduit, sur le même cadre de discernement, à la fonction de masse de minimum de spécificité :

$$m_h(P_i) = 0\,; \quad m_h(\neg P_i) = 1 - V_h\,; \quad m_h(E_i) = V_h \tag{41}$$

L'interprétation du facteur de fiabilité qui est donnée au paragraphe 4 revient à le modéliser, de façon parfaitement duale, par la fonction de masse $m_{Fh}(.)$ sur $E_h = \{F_h, \neg F_h\}$:

$$m_{Fh}(F_h) = 0\,; \quad m_{Fh}(\neg F_h) = 1 - f_h\,; \quad m_{F_h}(E_h) = f_h \tag{42}$$

Corrélativement elle permet de modifier la fonction de masse $m_h(.)$ donnée par (41) à l'aide de l'opérateur d'affaiblissement (25), pour donner la fonction de masse $m_h(.)$ sur E_i qui doit finalement être prise en compte :

$$\begin{aligned} &m'_h(P_i) = 0\,; \quad m_h(\neg P_i) = f_h(1 - V_h)\,; \\ &m'_h(E_i) = 1 - f_h(1 - V_h) \end{aligned} \tag{43}$$

Il convient de remarquer que cette fonction de masse est la fonction de masse de minimum de spécificité qui correspond à la seule donnée de :

$$Pl'_h(P_i) = 1 - f_h(1 - V_h) \;\text{ sur }\; E_i = \{P_i, \neg P_i\} \tag{44}$$

Il suffit donc de considérer cette seule donnée comme entrée des opérateurs développés précédemment. Il convient également de constater que le modèle (43) est en parfait accord avec les modèles définis antérieurement dans le cas particulier d'un ensemble d'hypothèses P_i exclusives et exhaustives [APP 02] [APP 04].

8.2. Expression de V_h et f_h en fonction des observations

En accord avec le début du paragraphe 8, la vraisemblance V_h est classiquement établie à partir de l'observation d'un signal s_j et de la caractérisation préalable de la proposition P_i sur l'espace E_j des signaux possibles *a priori* pour s_j. La formulation de cette caractérisation permet d'élaborer l'expression appropriée de V_h. Les plus courantes de ces formulations conduisent en pratique à des expressions très simples :

- S'il s'agit d'une description incertaine à l'aide d'une densité de probabilité $p(./P_i)$ sur l'espace E_j des signaux possibles, une modélisation de minimum de spécificité, en accord entre (41) et les modélisations établies antérieurement dans le cas particulier d'un ensemble de propositions P_i exclusives et exhaustives [APP 02], fournit l'expression :

$$V_h = R_j \; p(s_j/P_i) \quad \text{avec} \quad R_j = \max_{i \in I'} \left(\sup_{x \in E_j} (p(x/P_i)) \right)^{-1} \qquad (45)$$

où $I' \subseteq I$ désigne les propositions P_i qui font l'objet d'une caractérisation préalable sur E_j.

- S'il s'agit d'une description imprécise à l'aide d'un sous-ensemble flou défini par la fonction d'appartenance $\mu_i(.)$ sur l'espace E_j, les mêmes raisons procurent cette fois :

$$V_h = \mu_i(s_j) \qquad (46)$$

- Si P_i est caractérisée par une valeur s_{ij} précise et certaine sur E_j, nous retrouvons comme cas particulier, à la fois de (45) et (46) :

$$\begin{aligned} V_h &= 1 \quad \text{si} \quad s_j = s_{ij} \\ V_h &= 1 \quad \text{si} \quad s_j \neq s_{ij} \end{aligned} \qquad (47)$$

Par ailleurs, la parfaite dualité obtenue entre la modélisation de la vraisemblance V_h (41) et la modélisation de sa fiabilité f_h (42) permet de retenir pour l'expression de cette dernière, en fonction de l'observation de la variable contextuelle z_k et de la caractérisation préalable de la proposition F_h sur l'espace E_k des valeurs possibles *a priori* pour z_k, des expressions strictement similaires à celles qui viennent d'être données pour V_h :

- S'agissant d'une description incertaine à l'aide d'une densité de probabilité $p(./F_h)$ sur l'espace E_k des valeurs possibles de la variable contextuelle z_k, nous avons l'expression :

$$f_h = R_k \; p(z_k/F_h) \quad \text{avec} \quad R_k = \max_{h \in H'} \left(\sup_{x \in E_k} (p(x/F_h)) \right)^{-1} \qquad (48)$$

où $H' \subseteq H$ désigne les fiabilités F_h qui sont caractérisées par la variable contextuelle définie sur E_k .

- S'agissant d'une description imprécise à l'aide d'un sous-ensemble flou défini par la fonction d'appartenance $\mu_h(.)$ sur l'espace E_k , nous avons cette fois :

$$f_h = \mu_h(z_k) \qquad (49)$$

- Enfin, si F_h est caractérisée par une valeur z_{hk} précise et certaine sur E_k , nous retrouvons comme cas particulier, à la fois de (48) et (49) :

$$\begin{aligned} f_h &= 1 \quad \text{si} \quad z_k = z_{hk} \\ f_h &= 0 \quad \text{si} \quad z_k \neq z_{hk} \end{aligned} \qquad (50)$$

9. Conclusion

L'opérateur d'extension développé au paragraphe 3 dans le cadre de la théorie de l'évidence, à partir des seules notions de conditionnement et de raffinement, a permis de dégager un ensemble d'opérateurs susceptibles de constituer une chaîne complète et cohérente de traitement, depuis une modélisation appropriée des données disponibles, jusqu'à une prise de décision adaptable au besoin opérationnel. Ces opérateurs, qui recouvrent comme cas particuliers les opérateurs classiques de la théorie de l'évidence, permettent de raisonner sur un ensemble de propositions les plus disparates, exprimées individuellement dans des référentiels hétérogènes, sans avoir à les ramener sur un cadre de discernement commun comme l'exige classiquement la théorie de l'évidence. De plus, le processus synthétisé s'accommode naturellement de la donnée des seules plausibilités identifiables, différentes de 1. L'outil obtenu est *a priori* de nature à satisfaire les exigences exprimées au paragraphe 1 en matière de fusion multisenseur, et permet en particulier de valoriser l'apport des théories de l'incertain pour le traitement d'observations incertaines multiples dans ce contexte.

Références

[APP 98] A. APPRIOU, « Uncertain data aggregation in classification and tracking processes », *Aggregation and Fusion of Imperfect Information*, B. Bouchon-Meunier, ed., Physica Verlag, 1998.

[APP 01] A. APPRIOU, « Situation Assessment Based on Spatially Ambiguous Multisensor Measurements », *International Journal of Intelligent Systems*, Volume 16, Issue 10 (Special Issue : Data and Knowledge Fusion), John Wiley & Sons, Inc., October, 2001.

[APP 02] A. APPRIOU, « Discrimination Multisignal par la Théorie de l'Évidence », *Reconnaissance des Formes et Décision en Signal, Traité IC2, Information, Commande, Communication*, Éditions Hermès, 2002.

[APP 04] A. APPRIOU, « Processus d'agrégation pour la décision », *Revue des Nouvelles Technologies de l'Information*, Numéro Spécial «Systèmes d'Information pour l'Aide à la Décision en Ingénierie système», Cépaduès Editions, RNTI-1, 2004.

[DEN 97] T. DENŒUX, « Analysis of evidence-theoristic decision rules for pattern classification », *Pattern Recognition*, Vol. 30, N° 7, 1997.

[DUB 88] D. DUBOIS, H. PRADE, « Representation and combination of uncertainty with belief functions and possibility measures », *Computer Intelligence*, 4, 1988.

[DUB 92] D. DUBOIS, H. PRADE, « On the combination of evidence in various mathematical frameworks », *Reliability Data Collection and Analysis*, J. FLAMM & T. LUISI, eds., 1992.

[SHA 76] SHAFER G., A *Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey, 1976.

[SME 86] P. SMETS, « Combining non-distinct evidences », *Proceedings*, Conference of North America Fuzzy Information Processing Society (NAFIPS' 86), New Orleans, 1986.

[SME 88] P. SMETS, « Belief functions », *Non-standard logics for automated reasoning* , P. SMETS, E. MAMDANI, D. DUBOIS, H. PRADE, eds, Academic Press, London, 1988.

[SME 90] P. SMETS, « The combination of evidence in the transferable belief model », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, n° 5, may, 1990.

[YAG 83] R.R. YAGER, « Entropy and specificity in a mathematical theory of evidence », *International Journal General Systems*, vol. 9, 1983.

[YAG 86] R.R. YAGER, « A general approach to decision making with evidential knowledge », *Uncertainty in Artificial Intelligence*, L.N. KANAL & J.F. LEMMER éd., Elsevier Science Publishers, BV North-Holland, 1986.

[YAG 87] R.R. YAGER, « Optimal alternative selection in the face of evidential knowledge », *Optimization models using fuzzy sets and possibility theory*, J. KACPRZYK & S.A. ORLOVSKI, eds, D. Reidel Publishing Company, 1987.

Alain **Appriou**

Né le 26 juin 1952, il est Ingénieur de l'École Centrale de Nantes (1975), Ingénieur Supélec (1977), Maître ès Sciences en Mécanique (1975), et habilité à diriger des recherches en Sciences (1995).

Directeur de recherche à l'ONERA, il a été responsable de recherches concernant le traitement des signaux, des images, et de l'information, le développement de systèmes complexes, les méthodes numériques, les techniques radar, la navigation, et le guidage. Il s'est plus personnellement investi dans la fusion de données et le traitement du signal pour le développement de systèmes de senseurs, avec un intérêt particulier pour les théories de l'incertain, et notamment pour la théorie de l'évidence. Il coordonne actuellement l'ensemble des actions de l'Office en matière de systèmes.

Parallèlement, il enseigne dans différentes écoles d'ingénieurs, et dans différents séminaires. Membre de Comités de Publication de revues internationales, il est également impliqué dans l'organisation de divers congrès internationaux. Il est enfin Président du Club Traitement du Signal de la SEE.