Un argument pour le choix entre décision pignistique et maximum de plausibilité en théorie de l'évidence

Alain NIFLE et Roger REYNAUD

Institut d'Electronique Fondamentale, université d'Orsay et
AEROSPATIALE ESPACE & DEFENSE, établissement des Mureaux

RÉSUMÉ

Nous abordons ici un aspect du problème de décision en théorie de l'évidence, à savoir le choix du critère associé. Lorsque la décision doit porter sur les singletons du cadre de discernement, deux approches sont couramment rencontrées : par maximum de probabilité pignistique ou par maximum de plausibilité. Nous proposons un argument fondé sur l'interprétation de la modélisation par masses permettant de choisir entre ces deux approches. En effet, dans le cas d'ensembles aléatoires, la fusion de Dempster pourrait être vue, selon le type d'entrée modélisée par masses, ou bien comme une extension de la fusion bayésienne classique (a priori et vraisemblance) conduisant au maximum de probabilité pignistique, ou bien comme une extension de la fusion de vraisemblances, tout aussi classique en théorie des probabilités et conduisant à un critère de maximum de plausibilité.

ABSTRACT

We point out a side of the decision problem in evidence theory, namely, the criterion choice. When the decision must be done on the singletons of the frame of discernment, two approaches are commonly chosen : by maximum of pignistic probability or maximum of plausibility. We propose an argument based on the modelisation interpretation which allows to choose between these two approaches. As a matter of fact, in the case of random set, the Dempster fusion could be seen, according to the type of the inputs, as an extension of the classic Bayes fusion (a priori and likelihood) leading to maximum pignistic probability or as an extension of the likelihood fusion, also classic in probability theory, and leading to a maximum of plausibility criterion.

1 Introduction

La pertinence du raisonnement entièrement probabiliste provient du sens précis que l'on peut conférer aux coefficients de probabilités. Ainsi, l'interprétation fréquentiste conduit directement aux différentes formules bien connues. La fusion bayésienne peut, notamment de ce point de vue, être démontrée bien qu'elle puisse aussi être obtenue en imposant que les degrés de probabilités satisfassent un certain nombre de contraintes vues comme des axiomes [BLOCH].

La difficulté du raisonnement en théorie de l'évidence [SHAFER] résulte de l'existence de multiples interprétations possibles et non consensuelles du sens des masses affectées aux sous ensembles du cadre de discernement, conduisant naturellement à de multiples méthodes. La combinaison proposée par Dempster est le plus souvent utilisée mais on trouve typiquement deux approches en ce qui concerne l'étape de décision sur les hypothèses singletons : par critère de maximum de plausibilité [APPRIOU] ou bien de maximum de probabilité pignistique [SMETS].

L'objet de ce papier est de proposer un argument permettant de choisir entre ces deux approches décisionnelles dans certains cas particuliers, en fonction du type (et donc du sens) d'entrée disponible et en supposant des coûts d'erreur de décision équirépartis. Nous établissons un parallèle entre ces deux approches de décision et les deux types de représentations, probabilité et vraisemblance, que l'on trouve à la base des systèmes de décision en théorie des probabilités.

Dans un premier temps, nous rappelons le mécanisme d'extension de répartition permettant, d'un point de vue strictement probabiliste, de passer des informations sur des sous-ensembles non singletons, à des informations sur les hypothèses singletons du cadre de discernement. Nous verrons, notamment, que ce mécanisme diffère selon que les incertitudes sont représentées par des probabilités ou des vraisemblances. Dans un deuxième temps, nous montrons que les transformations pignistiques ou d'élaboration des plausibilités peuvent être vues comme un élargissement, au cadre de la théorie de l'évidence, du concept d'extension de répartition. Ceci nous permet alors de nous positionner quant à l'approche décisionnelle.
(En raison de l'associativité et la commutativité de la combinaison de Dempster, nous limitons notre étude, sans perte de généralité, à deux sources d'information).

2 Rappel des approches probabilistes

2.1 Source d'informations sur singletons

On considère un problème de classification. On note E le cadre de discernement composé de l'ensemble des hypothèses possibles, supposées exclusives et exhaustives : $E=\{H_1,..,H_n\}$.

Modélisation. On suppose un capteur apte à fournir une mesure m d'une grandeur x, capteur dont l'erreur est caractérisée par la distribution probabiliste : $P(m/x) \quad \forall (m,x) \in \Re^2$.

On suppose, de plus, que l'on a accès aux répartitions de probabilités de la grandeur x conditionnellement aux hypothèses du cadre de discernement, à savoir : $P(x/H_i) \quad \forall H_i \in E$ et $\forall x \in \Re$.

Quand une mesure m_0 est disponible, il est possible d'évaluer la vraisemblance de chaque hypothèse H_i au vu de cette mesure. Ainsi, un capteur fournit une répartition de vraisemblances avec :

$$V_{m_0}(H_i) = P(m_O / H_i) = \int_{-\infty}^{+\infty} P(m_0 / x)\, P(x / H_i)\, dx \qquad (1)$$

Fusion et décision. S'il est nécessaire de décider une hypothèse, alors deux cas sont envisageables. Ou bien les probabilités a priori $P(H_i) \quad \forall H_i \in E$ sont disponibles et on peut envisager le critère de maximum de probabilité a posteriori après fusion Bayésienne classique (fusion vraisemblance et a priori que l'on notera $Vrais \oplus A\ Priori$). Ou bien elles ne le sont pas et dans ce cas, on peut envisager le critère de maximum de vraisemblance. Il est à noter que le choix de ce critère revient exactement à considérer une répartition a priori uniforme, correspondant au choix du maximum d'entropie [GUIASU], à savoir un minimum d'information introduite (au sens de Shannon) et à prendre le critère de maximum de probabilité a posteriori.

Si un deuxième observateur fournit une mesure indépendante de la première, alors il est possible de fusionner ces deux informations (fusion de vraisemblances que l'on notera $Vrais_1 \oplus Vrais_2$) pour obtenir la vraisemblance fusionnée normalisée de chaque hypothèse :

$$V_{m_1,m_2}(H_i) = \frac{V_{m_1}(H_i) V_{m_2}(H_i)}{\sum_{H_j} V_{m_1}(H_j) V_{m_2}(H_j)} \qquad (2)$$

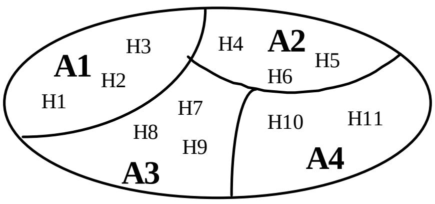
Il existe donc deux types de fusion probabiliste : la fusion $Vrais \oplus A\ Priori$ conduisant à une probabilité a posteriori et la fusion $Vrais_1 \oplus Vrais_2$ conduisant à une vraisemblance.

2.2 Sources d'informations sur des sous-ensembles composés

Modélisation. Des informations de type vraisemblance (issues d'un capteur) ou de type probabilité a priori peuvent être fournies sur une partition de E.

$$E = \{ H_1, \ldots, H_i, \ldots, H_n \} = \{ A_1, \ldots, A_p, \ldots, A_m \} \qquad (3)$$

Exemple de partition de E :



Ainsi, un capteur peut fournir des informations représentées par des vraisemblances de sous-ensembles non singletons de E. Par exemple, un capteur peut être sensible à l'âge d'un homme et un autre à sa taille.

Mécanismes d'extension aux hypothèses singletons. Suivant le type d'incertitude (vraisemblance ou probabilité), on expose ici deux mécanismes d'extension de leur répartition définie sur une partition vers une répartition sur les hypothèses singletons. En plus d'être nécessaires pour la décision, ces extensions de répartitions permettent les fusions probabilistes $Vrais \oplus A\ Priori$ et $Vrais_1 \oplus Vrais_2$.

Pour A_p, un élément de la partition de E, nous nous plaçons dans le cas où on dispose de $P(A_p)$. Les probabilités étant additives, en répartissant de manière équilibrée les probabilités sur les singletons, on prendra, de façon cohérente avec le principe de maximum d'entropie [GUIASU] :

$$\boxed{\forall H_i \in E, \quad P(H_i) = \frac{P(A_p)}{card(A_p)}\ avec\ A_p \supset H_i} \qquad (4)$$

Dans le cas où nous disposons de la vraisemblance de A_p :

$$V_{m_0}(A_p) = P(m_0 / A_p) = P(m_0 / (\bigcup_{H_i \in A_p} H_i))$$

$$= \frac{\sum_{H_i \in A_p} P(m_0 / H_i)\, P(H_i)}{\sum_{H_i \in A_p} P(H_i)}$$

donc : $$V_{m_0}(A_p) = \frac{\sum_{H_i \in A_p} V(H_i) P(H_i)}{\sum_{H_i \in A_p} P(H_i)} \neq \sum_{H_i \in A_p} V(H_i) \qquad (5)$$

La vraisemblance de A_p est donc le barycentre des vraisemblances des hypothèses singletons affectées des coefficients de probabilités a priori. Par conséquent, en répartissant de façon équilibrée les probabilités a priori et les vraisemblances, on obtient le mécanisme d'extension de vraisemblances suivant :

$$\boxed{\begin{array}{l} \forall H_i \in E,\ V(H_i) = V(A_p) \text{ avec } A_p \supset H_i \\ \text{et } V_{normalisée}(H_i) = \dfrac{V(H_i)}{\sum_{H_i \in E} V(H_j)} \end{array}} \qquad (6)$$

Notons qu'en faisant ce choix, la relation (5) est vérifiée indépendamment des probabilités a priori.

Fusion et décision. Bien que nous restions, dans ce chapitre, dans un cadre strictement probabiliste (les ensembles A_p étant exclusifs et exhaustifs), nous utiliserons, par soucis de clarté, les notations et le formalisme de la théorie de l'évidence. On note donc $m_v(X)=V(X)$ (masse au sens de vraisemblance) et $m_p(X)=P(X)$ (masse au sens de probabilité).

En notant m_{be} une répartition de masses uniforme bayésienne sur l'ensemble des singletons de E, m_A la répartition de masses dont les éléments focaux sont les parties A_p de E et m_H la répartition de masses m_A étendue aux singletons, alors le mécanisme d'extension de répartition de vraisemblances de la relation (6) peut être exprimé par la combinaison au sens de Dempster suivante :

$$m_H = m_A \oplus m_{be} \qquad (7)$$

Le mécanisme d'extension de probabilités exprimé par (4) est, quant à lui, difficilement représentable par une combinaison de Dempster. On définit donc un opérateur, noté $Pig\,(m_A)$, correspondant à cette extension en référence à la transformation dite "pignistique" proposée par [SMETS], définie dans le cadre plus général de la théorie de l'évidence et dont la restriction au cas d'une répartition sur une partition donne (4).

Toujours en restant dans un cadre strictement probabiliste et où les informations disponibles le sont sur une partition de E, les fusions $Vrais \oplus A\ Priori$ et $Vrais_1 \oplus Vrais_2$ restent possibles en appliquant le principe suivant : on étend puis on fusionne classiquement (voir § 2.1).

Ainsi, dans le cas $Vrais \oplus A\ Priori$, la fusion entre m_V et m_P donne la masse m suivante :

$$m = (\,m_v \oplus\ m_{be}\,) \oplus\ pig\ (\,m_p\,) =\ m_v \oplus\ pig\ (\,m_p\,) \qquad (8)$$

Exemple 1:

$\backslash m_V$ / $m_P\backslash$	H_1H_2 2/3	H_3 1/3
H_1 1/3		
H_2H_3 2/3		

$\Rightarrow$

$\backslash m_V$ / $m_P\backslash$	H_1 2/5	H_2 2/5	H_3 1/5
H_1 1/3	2/5 H_1		
H_2 1/3		2/5 H_2	
H_3 1/3			1/5 H_3

et dans le cas $Vrais_1 \oplus Vrais_2$, la fusion est la suivante :

$$m = (\,m_{v1} \oplus\ m_{be}\,)\ \oplus\ (\,m_{v2} \oplus\ m_{be}\,) \qquad (9)$$

Exemple 2 :

$\backslash m_{V1}$ / $m_{V2}\backslash$	H_1H_2 2/3	H_3 1/3
H_1 1/3	2/8 H_1	
H_2H_3 2/3	4/8 H_2	2/8 H_3

$\Leftrightarrow$

$\backslash m_{V1}$ / $m_{V2}\backslash$	H_1 2/5	H_2 2/5	H_3 1/5
H_1 1/3	2/8 H_1		
H_2 2/5		4/8 H_2	
H_3 2/5			2/8 H_3

Finalement, deux types d'extensions coexistent, suivant le type d'incertitudes mises en jeu. Ces deux mécanismes, constituant deux approches, permettent une fusion d'informations données sur deux partitions (qui peuvent être différentes) du cadre de discernement E. Ces extensions de répartition puis la fusion peuvent d'ailleurs être appliquées de chacune des deux partitions de base P_1 et P_2 vers la partition $P = P_1 \cap P_2$.

3 Elargissement au cadre de la théorie de l'évidence

3.1 Interprétation des masses et fusion

On suppose maintenant que les sources d'information considérées, que constituent une répartition de probabilités a priori ou une répartition de vraisemblances, sont mal connues. On modélise alors cette imperfection par une répartition de masses sur tous les sous ensembles du cadre de discernement E. De même, pour être cohérent avec l'approche strictement probabiliste, on considère que les répartitions de masses ont un sens correspondant ou bien à un élargissement du concept de probabilité, ou bien à un élargissement du concept de vraisemblances. Si les sous-ensembles focaux forment une partition de E alors, dans le premier cas on étendra les masses aux singletons suivant (4), dans le second cas, suivant (6). La masse issue de la fusion aura donc un sens dépendant du type d'incertitude que représentent les masses à combiner. Ainsi :

$$\text{masse}_{Probabilité} \oplus \text{masse}_{Vraisemblance} = \text{masse}_{Probabilité} \qquad \text{(type 1)}$$

$$\text{masse}_{Vraisemblance} \oplus \text{masse}_{Vraisemblance} = \text{masse}_{Vraisemblance} \qquad \text{(type 2)}$$

On montre ici que cet éclairage sémantique et la contrainte de cohérence avec le cas d'informations probabilistes sur des partitions de E impliquent que la fusion de type 1 n'est envisageable que dans certains cas particuliers. Prenons un exemple. On suppose $E = \{\,H_1\,, H_2\,, H_3\,\}$ et une première source d'information issue d'un capteur aveugle, fournissant donc la répartition de vraisemblances suivante :

$$V\,(\,H_1\,) = V\,(\,H_2\,) = V\,(\,H_3\,) = 1/3$$

modélisé par : $m_v\,(\,H_1\,) = m_v\,(\,H_2\,) = m_v\,(\,H_3\,) = 1/3$ (10)

On suppose de plus que l'on dispose de la répartition de probabilités a priori suivante :

$$P\,(\,H_1\,) = 1/3 \ \text{ et } \ P\,(\,H_2 \cup H_3\,) = 2/3$$

modélisé par : $m_p\,(\,H_1\,) = 1/3$ et $m_p\,(\,H_2\,, H_3\,) = 2/3$ (11)

La fusion au sens de Dempster donne alors : $m = m_v \oplus m_p$

$\backslash m_V$ / $m_P\backslash$	H_1 1/3	H_2 1/3	H_3 1/3
H_1 1/3	1/5 H_1		
H_2H_3 2/3		2/5 H_2	2/5 H_3

$$\Rightarrow\ m\,(\,H_1\,) = 1/5\ ,\ \ m\,(\,H_2\,) = m\,(\,H_3\,) = 2/5 \qquad (12)$$

En toute rigueur, cette répartition de masses doit être interprétée comme une probabilité et non une vraisemblance. Ce résultat implique donc $P\,(\,H_2 \cup H_3) = 4/5$. Il y a là une incohérence. En effet, le capteur n'apportant en réalité aucune information, le résultat devrait être cohérent avec la distribution de probabilités a priori.

De plus, si on reprend le principe de fusion établi en (8), on a :

$\backslash m_V$ / $m_P\backslash$	H_1 1/3	H_2 1/3	H_3 1/3
H_1 1/3			
H_2H_3 2/3			

$\Rightarrow$

$\backslash m_V$ / $m_P\backslash$	H_1 1/3	H_2 1/3	H_3 1/3
H_1 1/3	1/3 H_1		
H_2 1/3		1/3 H_2	
H_3 1/3			1/3 H_3

Donc $m\,(\,H_1\,) = m\,(\,H_2\,) = m\,(\,H_3\,) = 1/3$. Cette masse résultante, différente, a pour sens une probabilité et est cohérente avec les données d'entrée ($P(H_1 \cup H_2) = 2/3$). En fait, on peut considérer que la fusion de type 1 n'a de sens que si la condition suivante est vérifiée :

$$pig\ (m_p \oplus m_v) = pig\ (m_p) \oplus m_v \qquad (13)$$

Il suffit pour satisfaire (13) qu'en notant A_i et B_j les sous-ensembles focaux respectivement de m_p (au sens de probabilités) et m_v (au sens de vraisemblances) :

$$\forall A_i, \forall B_j,\ \ A_i \subset B_j \ \text{ ou } \ A_i \cap B_j = \varnothing \qquad (14)$$

La masse résultante a alors un sens de probabilités. Pour l'exemple précédant, la condition (13) n'est pas vérifiée et une fusion de ce type n'a pas de sens. Par contre, cette relation est toujours vérifiée quand la masse a priori est bayésienne et donc la fusion a un sens.

Maintenant, si on reprend l'exemple précédant, en considérant cette fois-ci que les deux sources sont des vraisemblances, alors la fusion de type 2 au sens de Dempster reste cohérente. En effet, la masse résultante (12) a cette fois-ci un sens de vraisemblance et est totalement cohérente avec les entrées (d'après (6)), à savoir un capteur aveugle et un autre capteur fournissant les vraisemblances $V(H_1) = 1/3$ et $V(H_2 \cup H_3) = 2/3$. De plus, ce résultat est identique à celui issu du mécanisme de fusion applicable ici, et formulé en (9).

Enfin, il est important de noter que :

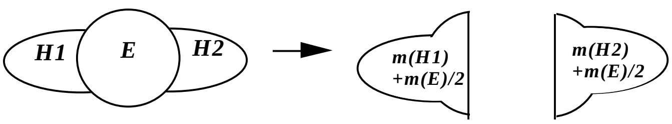
$$(m_{v1} \oplus m_{v2}) \oplus m_{be} = (m_{v1} \oplus m_{be}) \oplus (m_{v2} \oplus m_{be}) \qquad (15)$$

Ceci montre qu'indépendamment des masses (de vraisemblances) à fusionner, la combinaison de Dempster est toujours cohérente avec le mécanisme de fusion exposé en (9).

3.2 Décision

On se place ici dans la configuration où la décision d'une des hypothèses est demandée. Il est alors nécessaire de choisir une mesure de confiance affectée aux seuls singletons de E. Quand on considère la théorie de l'évidence comme un élargissement au cas de sources mal connues de la théorie des probabilités, alors nous avons vu que l'interprétation des masses est essentielle. Par conséquent, la transformation permettant la décision, à savoir l'élaboration d'une répartition de masses bayésienne de décision notée m_D, doit tenir compte du sens donné aux masses. Si le sens est une probabilité, alors la transformation pignistique [SMETS] semble naturelle, à savoir :

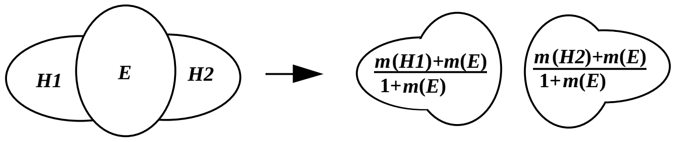
$$\forall H_i \in E, \quad m_D(H_i) = \sum_{A_j \supset H_i} \frac{m_p(A_j)}{card(A_j)} \qquad (16)$$



La décision se fait donc sur critère de maximum de probabilité pignistique.

Par contre, si les masses ont un sens de vraisemblance alors la transformation doit être cohérente avec l'extension d'une répartition de vraisemblances définie sur une partition du cadre de discernement. La transformation suivante, dont la restriction à une partition donne (6), semble mieux indiquée :

$$\forall H_i \in E, m'(H_i) = \sum_{A_j \supset H_i} m_v(A_j) \quad , m_D(H_i) = \frac{m'(H_i)}{\sum_{H_j \in E} m'(H_j)} \qquad (17)$$



En remarquant que $m'(H_j) = Pl(H_j)$, on aboutit finalement à une décision fondée sur le critère de maximum de plausibilité [APPRIOU], équivalente à une décision sur critère de maximum de probabilité a posteriori après fusion de type 1 avec une répartition de masses bayésienne uniforme. En raison de (17) et (6), ce critère de décision peut être vu comme un élargissement au cadre de la théorie de l'évidence du critère de maximum de vraisemblance.

3 Conclusion

En se plaçant dans un cadre d'intelligence artificielle et en ne considérant aucunement l'aspect aléatoire des sous ensembles de E, alors les masses représentent des confiances transférables [SMETS], issues d'expertises ou de traitements symboliques, sur des ensembles non nécessairement disjoints. Dans ce cas, différentes approches décisionnelles telles que le maximum de crédibilité, de probabilité pignistique ou de plausibilité sont difficilement comparables a priori.

Nous nous sommes placés dans un cadre concernant des ensembles aléatoires permettant de nous raccrocher aux raisonnements probabilistes. De ce fait, il est possible de conserver l'interprétation fréquentiste, et par la contrainte d'une continuité du mécanisme de décision dans les cas de répartition de masses sur une partition E, de se positionner en matière de décision. Dans ce cadre, nous avons pu montrer deux situations qui, en fonction de l'interprétation initiale des masses, sont opérationnelles avec l'un des modes de décision et pas avec l'autre.

Références

[APPRIOU] A. APPRIOU. "*Probabilités et Incertitude en fusion de données Multisenseurs*", Revue scientifique et technique de la défense 1992.

[BLOCH] I. BLOCH, H. MAITRE (1994), "*Fusion de données en traitement d'images : modèles d'information et décision*". Traitement du Signal. Signal, Image et Parole. Vol 11, N°6. Numéro spécial 1994. Fusion de données. Revue GRETSI.

[GUIASU] S. GUIASU, R. THEODORESCU, "*Incertitude et information*", Les presses de l'université LAVAL (1971).

[SHAFER] G. SHAFER, "*A mathematical theory of evidence*", Princeton University Press 1975.

[SMETS] P. SMETS et R. KENNES, "*The transferable belief model*", Artificial intelligence, Vol 66 N°2 pp191-234, 1994.