



**HAL**  
open science

## Traitement évidentiel d'exemples non-conformes en discrimination

Benjamin Quost, Marie-Hélène Masson, Sébastien Destercke

► **To cite this version:**

Benjamin Quost, Marie-Hélène Masson, Sébastien Destercke. Traitement évidentiel d'exemples non-conformes en discrimination. 29èmes Rencontres Francophones sur la Logique Floue et ses Applications (LFA 2020), Oct 2020, Sète, France. hal-02944658

**HAL Id: hal-02944658**

**<https://hal.science/hal-02944658>**

Submitted on 21 Jun 2021

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

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

# Traitement évidentiel d'exemples atypiques en discrimination

B. Quost<sup>1</sup>      M.H. Masson<sup>2</sup>      S. Destercke<sup>1</sup>

<sup>1</sup> Université de Technologie de Compiègne

<sup>2</sup> Université de Picardie Jules Verne

UMR CNRS 7253 Heudiasyc, Sorbonne Universités, Université de Technologie de Compiègne

CS 60319 - 60203 Compiègne cedex, France

## Résumé :

Lorsque l'on classe un exemple sur la base d'un ensemble d'apprentissage, au moins trois situations dans lesquelles il est difficile de déterminer une classe unique peuvent se produire : soit il y a trop d'incertitude aléatoire car l'individu appartient à une zone de mélange, soit il y a trop d'incertitude épistémique car il y a peu de données dans la zone dans laquelle se situe l'exemple, soit l'exemple ne ressemble à aucun autre exemple d'apprentissage. Alors que les deux premières situations ont été explorées de manière assez intensive dans la littérature, la troisième situation a reçu en comparaison moins d'attention. Cet article présente des pistes de résolution ce problème dans le cadre de la théorie des fonctions de croyance.

## Mots-clés :

Fonctions de croyance, classification supervisée, incertitudes aléatoire et épistémique, gestion d'atypicité, détection de nouveauté

## Abstract:

When classifying an example on the basis of an observed population of (training) samples, at least three kinds of situations can arise where picking a single class may be difficult : high aleatory uncertainty due to the natural mixing of classes, high epistemic uncertainty due to the scarcity of training data, and non-conformity or atypicality of the example with respect to observations made so far. While the two first kinds of situations have been explored extensively, the last one still calls for a principled analysis. This paper is a first proposal to address this issue within the theory of belief function.

## Keywords:

Belief functions, Supervised classification, Epistemic and aleatoric uncertainty, Atypicality management, Novelty detection

## 1 Introduction

Dans un problème de discrimination, affecter une nouvelle instance sur la base d'un ensemble d'apprentissage peut être difficile en raison de sources d'incertitudes variées : l'incertitude aléatoire ou ambiguïté (les classes sont mélangées, aucune ne domine), l'incertitude épistémique (les données d'apprentissage sont pauvres), et la non conformité de l'exemple de test par rapport aux données d'apprentissage. Cette dernière source d'incertitude, bien

que reliée à l'incertitude épistémique, en diffère en ce qu'elle ne peut pas toujours être réduite grâce à la collecte de nouvelles données. C'est un point central en détection de nouveauté ou de points aberrants [2]. Dans cet article, nous étudions comment prendre en compte le caractère atypique de certaines instances dans le cadre de la théorie des fonctions de croyance. La théorie des fonctions de croyance introduite dans [3, 6], et développée par la suite par Smets [9], est un cadre adapté pour la représentation des incertitudes. Notons que la détection de données atypiques a déjà été étudiée dans d'autres cadres théoriques, comme avec la notion de rejet de distance [5] ou les prédictions conformelles [7], et a aussi été étudiée dans le cadre des fonctions de croyance (e.g., [1]) pour des cas spécifiques. Cependant, à notre connaissance, aucune approche de principe, générique, pour prendre en compte la non conformité, n'a été proposée dans le cadre des fonctions de croyance. Cet article peut être vu comme une première contribution à cette problématique. Nous donnons dans le premier paragraphe les notions de base nécessaires. Le paragraphe 3 discute certaines propriétés que l'on peut souhaiter lorsque la typicité des exemples est prise en compte, et enfin le paragraphe Section 4 propose différents modèles.

## 2 Notions de base

Nous commençons par rappeler dans ce paragraphe les notions de base des fonctions de croyance, un cadre riche et flexible pour la gestion des incertitudes que nous allons utiliser par la suite. Nous exposons ensuite le modèle imprécis de Dirichlet pour modéliser l'apparte-

nance aux différentes classes et, enfin, introduisons la notion de conformité d'un exemple.

## 2.1 Préliminaires sur les fonctions de croyance

On considère une variable  $\omega$  prenant ses valeurs dans un ensemble non ordonné  $\Omega = \{\omega_1, \dots, \omega_M\}$  appelé cadre de discernement. La connaissance partielle sur la valeur de  $\omega$  est représentée par une fonction de masse [6]  $m : 2^\Omega \rightarrow [0; 1]$  telle que

$$\sum_{A \subseteq \Omega} m(A) = 1. \quad (1)$$

Les ensembles  $A \subseteq \Omega$  tels que  $m(A) > 0$  sont appelés *éléments focaux* de  $m$ . Si  $m(A) = 1$  pour un  $A \subseteq \Omega$ ,  $m$  est dite *catégorique* et est noté  $m_A$  (si  $A = \Omega$ ,  $m_\Omega$  représente l'ignorance complète). On impose souvent que  $m(\emptyset) = 0$ ; sinon,  $m(\emptyset)$  peut prendre des interprétations variées, comme le degré de conflit après la fusion d'informations inconsistantes, ou le degré de croyance que  $\omega \notin \Omega$  (*hypothèse de monde ouvert*).

Toute fonction de masse peut être représentée de manière équivalente par une fonction de croyance  $bel$  et une fonction de plausibilité  $pl$ , définies, respectivement, pour tout  $A \subseteq \Omega$  par :

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

De nombreuses stratégies ont été proposées pour prendre des décisions sur la base de ces fonctions de croyance — voir, par exemple, [4]. Dans ce qui suit, nous noterons  $\delta$  tout opérateur de décision appliqué à une fonction de masse définie sur l'ensemble des classes. Par exemple, l'opérateur de *dominance d'intervalle*, qui peut conduire à une décision imprécise (en ce qu'il peut fournir un ensemble de classes possibles) est défini comme suit.

### Définition 1 (Dominance d'intervalle (DI))

Etant donnée une masse  $m$ , on dit que  $\omega_i$  domine  $\omega_j$ , noté  $\omega_i \succ \omega_j$ , si  $bel(\{\omega_i\}) > pl(\{\omega_j\})$ .

La règle de dominance d'intervalle consiste à fournir comme décision l'ensemble des classes

non dominées :

$$\delta_{ID}(m) = \{\omega_i : pl(\{\omega_i\}) \geq bel(\{\omega_j\}) \text{ pour tout } j \neq i\}. \quad (3)$$

## 2.2 Modélisation de l'appartenance aux classes

On suppose qu'une source fournit de l'information sur la classe d'un individu  $\mathbf{x}$  sous la forme d'une fonction de masse  $m$ . Cette masse est généralement obtenue à partir d'un ensemble de  $N$  individus d'apprentissage  $\mathbf{x}_i$  ( $i = 1, \dots, N$ ) observés dans la même région que  $\mathbf{x}$  et supposés être similaires à  $\mathbf{x}$ . Par exemple, si l'on utilise des arbres de décision,  $\mathbf{x}$  est classé grâce aux données appartenant à la même feuille que  $\mathbf{x}$ ; dans les  $k$  plus proches voisins, la décision se fonde sur les  $k$  individus d'apprentissage les plus proches de  $\mathbf{x}$ .

En classification prudente, la quantité d'information amenée par l'échantillon d'apprentissage est prise ne compte dans la phase de décision — permettant ainsi d'élaborer des stratégies prudentes, comme celle de retenir un ensemble de classes plausibles. Le modèle imprécis de Dirichlet (IDM) permet d'obtenir un tel modèle prudent des fréquences de classes sous la forme d'une fonction de masse  $m$ . En effet, si  $(n_1, n_2, \dots, n_M)$  denote les comptages associés aux différentes classes dans l'échantillon de taille  $N$ , avec  $\sum_k n_k = N$ , l'IDM produit la fonction de masse

$$m^{IDir}(\Omega) = \frac{s}{N+s}, \quad (4)$$

$$m^{IDir}(\{\omega_i\}) = \frac{n_i}{N+s} \quad \forall i = 1, \dots, M,$$

où le paramètre  $s$  peut être interprété comme le nombre de données additionnelles non observées lorsque l'on estime les probabilités des classes. Ce paramètre détermine le degré d'imprécision de l'estimation : plus il est élevé, plus l'inférence est prudente. De cette fonction de masse, on déduit les bornes

$$\left[ bel^{IDir}(\{\omega_i\}) = \frac{n_i}{N+s}; pl^{IDir}(\{\omega_i\}) = \frac{n_i+s}{N+s} \right],$$

qui rendent compte à la fois de l'incertitude aléatoire (qui intervient lorsque les décomptes

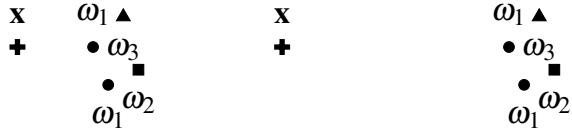


Figure 1 – Deux situations avec des comptages de classes identiques mais différents niveaux de typicité ; à gauche : conformité élevée ( $C = 0.8$ ) ; à droite : faible conformité ( $C = 0.2$ )

prennent des valeurs  $n_1, \dots, n_M$  similaires) et de l'incertitude épistémique (dans ce cas la largeur des intervalle croît avec la valeur de  $s/(N + s)$ ).

### 2.3 Conformité

Il faut noter que l'IDM ne tient pas compte de la typicité d'une instance de test, c'est à dire de sa ressemblance aux exemples d'apprentissage. On suppose ici que cette information de typicité est fournie de manière indépendante, sous la forme d'un score de conformité  $C \in [0; 1]$  : on a  $C = 0$  pour une instance de test complètement atypique, et  $C = 1$  pour une instance normale.

Sur la Figure 1 sont représentées deux situations dans lesquelles une instance  $\mathbf{x}$  doit être classée dans une des trois classes  $\{\omega_1, \omega_2, \omega_3\} = \Omega$ , sur la base de quatre exemples d'apprentissage. Le même IDM serait construit (les comptages de classes sont les mêmes) mais le niveau de typicité de  $\mathbf{x}$  est très différent d'un exemple à l'autre. Dans cet article, nous ne nous pencherons pas sur la manière dont ce niveau de typicité peut être quantifié (on pourrait par exemple exploiter la distance de l'exemple  $\mathbf{x}$  à son plus proche voisin).

Notre objectif est d'étudier comment une fonction de masse portant sur la classe d'une instance  $\mathbf{x}$  peut être modifiée (ou révisée) en fonction du niveau de typicité  $C$ . Pour cela, nous introduisons la notion d'*opérateur de conformité*  $Cf$ , qui transforme  $m$  en une nouvelle fonction de masse  $Cf[m, C]$ . On peut souhaiter que cet opérateur possède un certain nombre de propriétés, étudiées dans le paragraphe 3, sur la base desquelles différents opérateurs peuvent être proposés (voir le paragraphe 4).

## 3 Propriétés souhaitables des opérateurs de conformité

Dans la suite, par abus de notation,  $Cf[pl, \cdot]$  (respectivement,  $Cf[bel, \cdot]$ ) sera utilisé en lieu et place de la fonction de plausibilité (resp. la fonction de croyance) obtenue à partir de la fonction de masse modifiée  $Cf[m, \cdot]$ .

### Propriété 1 (Préservation des préférences)

Un opérateur de conformité  $Cf$  préserve l'information de préférence sur les classes si, pour tout  $C \in [0; 1]$ ,

$$\begin{aligned} pl(\{\omega_i\}) &\leq bel(\{\omega_j\}) \\ \Rightarrow Cf[pl, C](\{\omega_i\}) &\leq Cf[bel, C](\{\omega_j\}). \end{aligned} \quad (5)$$

Formulée autrement, cette propriété indique que la prise en compte de la conformité ne remet pas en cause la dominance d'intervalle entre les classes (voir l'équation (1)).

L'exemple sur la Figure 2 représente les frontières de décision d'un arbre appliqué aux données IRIS (seules deux variables ont été gardées à des fins illustratives). Les instances 1, 2 et 3 sont atypiques. La classe Setosa domine clairement les deux autres pour l'instance 1, une information qui peut raisonnablement être gardée pendant l'étape de révision. Cependant, cette stratégie est plus discutable pour l'instance 2, qui paraît être plus proche de la classe Versicolor que de la classe Setosa lorsque l'on tient compte de la dispersion des classes : dans ce cas, il semble légitime d'effacer l'information tirée du sous-ensemble de données associé à cette feuille de l'arbre.

De manière générale, garder l'information de préférence inférée grâce à la population de référence semble raisonnable si le modèle a peu de chance de confondre l'atypicité avec une autre source d'incertitude. Sur la Figure 2, l'instance 2 est à égale distance des classes Versicolor et Setosa, ce que l'arbre de décision est par nature incapable de détecter.

### Propriété 2 (Renforcement de la décision)

Un opérateur de conformité  $Cf$  renforce les décisions prises avec une stratégie  $\delta$  si

$$C \leq C' \Rightarrow \delta(Cf[m, C]) \subseteq \delta(Cf[m, C']). \quad (6)$$

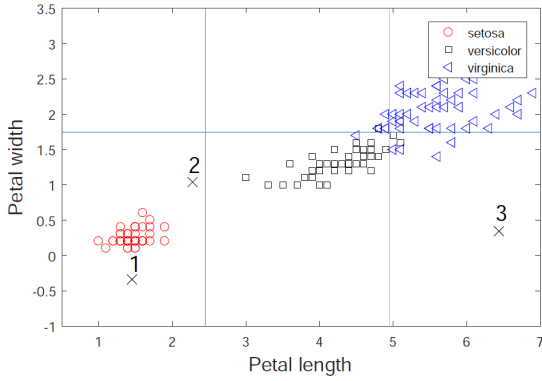


Figure 2 – Jeu de données Iris et plusieurs exemples atypiques

En d’autres termes, l’ensemble des classes plausibles pour une instance doit augmenter avec son niveau de conformité : à mesure qu’il devient atypique, les classes présumées plausibles doivent être retirées. Cela revient à poser l’hypothèse de monde ouvert, puisque les classes connues sont écartées de l’ensemble des classes possibles à mesure sur le niveau de conformité diminue, jusqu’à finir avec l’ensemble vide — comme avec les prédictions conformelles.

### Propriété 3 (Affaiblissement de la décision)

Un opérateur de conformité  $Cf$  affaiblit les décisions prises avec une stratégie  $\delta$  si

$$C \leq C' \Rightarrow \delta(Cf[m, C']) \subseteq \delta(Cf[m, C]). \quad (7)$$

Contrairement à la Propriété 2, la Propriété 3 fait l’hypothèse d’un monde clos, dans lequel  $\Omega$  est supposé contenir nécessairement toutes les classes possibles, mais l’information liée à une instance atypique est trop faible pour permettre une prédiction fiable.

Dans l’exemple sur la Figure 2, supposons que la décision pour l’instance 3 soit  $\delta(m_3) = \{\text{Versicolor}, \text{Virginica}\}$ . Alors, imposer la Propriété 2 reviendrait à écarter Versicolor, cette classe étant trop éloignée, alors que la Propriété 3 conduirait à l’ignorance complète, la classe Setosa étant intégrée à l’ensemble des classes plausibles.

D’autres propriétés peuvent être proposées, comme par exemple spécifier le comportement

de la règle de décision pour des instances non conformes extrêmes (pour  $C \rightarrow 0$ ). Ce sera fait dans le paragraphe suivant dans lequel nous proposons un certain nombre d’opérateurs ainsi que des règles de décision associées.

## 4 Quelques opérateurs de conformité et leur règle de décision

Ce paragraphe explore quelques opérateurs de conformité à la lumière des propriétés évoquées précédemment. Tous consistent à appliquer une transformation linéaire à la fonction de masses initiale en lien avec le niveau de conformité.

### 4.1 Affaiblissement classique en monde clos

Une première stratégie consiste à affaiblir<sup>1</sup>  $m$  en fonction du niveau  $1 - C$  d’atypicité :

$$Cf_1[m, C] = Cm + (1 - C)m_\Omega. \quad (8)$$

Dans le cas d’une fonction de masse générée par l’IDM, on a donc

$$\begin{cases} Cf_1[m, C](\{\omega_i\}) &= C \frac{n_i}{N + s}, \quad \forall i \\ Cf_1[m, C](\Omega) &= C \frac{s}{N + s} + (1 - C). \end{cases} \quad (9)$$

$Cf_1$  satisfait la Propriété 3 (affaiblissement de la décision) par rapport à  $\delta_{ID}$ , car le discounting rend les intervalles de croyance-plausibilité plus larges, aboutissant au cas extrême  $\delta_{ID}(Cf_1[m, 0]) = \Omega$ . Au contraire, et pour les mêmes raisons,  $Cf_1$  ne satisfait pas la Propriété 1 (préservation des préférences sur les classes). Une telle règle apparaît donc plus cohérente avec l’hypothèse d’un monde clos, dans laquelle les instances atypiques sont considérées comme étant difficiles à caractériser : une complète atypicité doit donc résulter en une ignorance totale.

### 4.2 Monde ouvert avec une classe “inconnue”

Notre deuxième opérateur  $Cf_2$  considère l’hypothèse de monde ouvert via une classe “inconnue”  $\omega_u$  : c’est à dire que  $Cf_2[m, C]$  est

1. L’affaiblissement  ${}^\varepsilon m$  de  $m$  par un facteur  $\varepsilon$  est défini par  ${}^\varepsilon m(A) = (1 - \varepsilon)m(A)$ , pour tout  $A \neq \Omega$ ; et  ${}^\varepsilon m(\Omega) = (1 - \varepsilon)m(\Omega) + \varepsilon$ .

maintenant une fonction de masse définie sur le référentiel  $\Theta = \Omega \cup \omega_u$  :

$$\text{Cf}_2 [m, C] = C m^{\uparrow\Theta} + (1 - C) m_{\omega_u}, \quad (10)$$

où l'extension vide  $m^{\uparrow\Theta}$  de  $m$  sur  $\Theta$  [8] est telle que  $m^{\uparrow\Theta}(A) = m(A)$  pour tout  $A \subseteq \Omega$  et  $m^{\uparrow\Theta}(A) = 0$  pour  $A \not\subseteq \Omega$  et  $m_{\omega_u}(\{\omega_u\}) = 1$ . Lorsqu'il est appliqué à la fonction de masse  $m^{IDir}$  générée par l'IDM, cet opérateur donne

$$\begin{cases} \text{Cf}_2 [m^{IDir}, C] (\{\omega_i\}) = C \frac{n_i}{N+s} & \forall i, \\ \text{Cf}_2 [m^{IDir}, C] (\{\omega_u\}) = 1 - C, \\ \text{Cf}_2 [m^{IDir}, C] (\Omega) = C \frac{s}{N+s}; \end{cases} \quad (11)$$

et, pour tout  $\omega_i \in \Omega$ , les valeurs de croyance et de plausibilités sont donc :

$$\begin{aligned} \text{Cf}_2 [bel^{IDir}, C] (\{\omega_i\}) &= C \frac{n_i}{N+s}, \\ \text{Cf}_2 [pl^{IDir}, C] (\{\omega_i\}) &= C \frac{n_i+s}{N+s}, \end{aligned} \quad (12)$$

et

$$\begin{aligned} \text{Cf}_2 [bel^{IDir}, C] (\{\omega_u\}) &= 1 - C, \\ \text{Cf}_2 [pl^{IDir}, C] (\{\omega_u\}) &= 1 - C. \end{aligned} \quad (13)$$

Appliquer  $\delta_{ID}$  à la masse mise à jour  $\text{Cf}_2 [m, C]$  (définie sur  $\Theta$ ) permet de satisfaire les Propriétés 1 et 2, avec le cas extrême  $\delta_{ID} (\text{Cf}_2 [m, 0]) = \{\omega_u\}$ . Notons aussi que

$$\begin{aligned} \delta_{ID} (\text{Cf}_2 [m, C]) \ni \omega_u \\ \Leftrightarrow \max_{\omega_j \in \Omega} \text{Cf}_2 [bel, C] (\{\omega_j\}) \leq 1 - C, \end{aligned}$$

$$\begin{aligned} \delta_{ID} (\text{Cf}_2 [m, C]) = \{\omega_u\} \\ \Leftrightarrow \max_{\omega_j \in \Omega} \text{Cf}_2 [pl, C] (\{\omega_j\}) < 1 - C. \end{aligned}$$

Cela a pour conséquence que l'ensemble des décisions inclura  $\{\omega_u\}$  seulement si le degré de support à chaque classe est faible. Cela suggère une autre stratégie dans laquelle  $\omega_u$  est laissée de côté lorsque l'on cherche les classes non dominées, et ajoutée a posteriori quand elle est aussi non dominée.

## Définition 2 (DI avec détection d'atypicité)

Etant donnée une masse  $m$  définie sur  $\Theta = \Omega \cup \{\omega_u\}$ , la règle de dominance d'intervalle avec détection d'atypicité est définie par

$$\delta_{ID:AT} (m) = \begin{cases} \delta_{ID} (m[\Omega]) \text{ si} \\ \min_{\omega_j \in \delta_{ID} (m[\Omega])} bel(\{\omega_j\}) > 1 - C, \\ \delta_{ID} (m[\Omega]) \cup \{\omega_u\} \text{ sinon,} \end{cases} \quad (14)$$

où le conditionnement  $m[\Omega]$  de  $m$  sur  $\Omega$  [8] est tel que  $m[\Omega](A) = \sum_{B \subseteq \Theta: B \cap \Omega = A} m(B)$ , pour tout  $A \subseteq \Omega$ .

En résumé, l'ensemble des classes non dominées est déterminé à partir des classes bien identifiées, et une alerte est envoyée si l'instance semble atypique. Cette stratégie satisfait la Propriété 1 et inclut la classe inconnue  $\omega_u$  lorsque  $C = 0$  : en particulier, on aura  $\delta_{ID:AT} (\text{Cf}_1 [m, 0]) = \Theta$ , et par ailleurs  $\delta_{ID:AT} (\text{Cf}_2 [m, 0]) = \omega_u$ .

## 4.3 Affaiblissement classique en monde ouvert

Nous proposons enfin un troisième opérateur pour lequel  $m$  est d'abord étendue sur  $\Theta$  puis affaiblie en fonction du niveau d'atypicité  $C$  :

$$\text{Cf}_3 [m, C] = C m^{\uparrow\Theta} + (1 - C) m_{\Theta}. \quad (15)$$

Dans le cas d'une masse  $m^{IDir}$  obtenue par l'IDM, on obtient donc :

$$\begin{cases} \text{Cf}_3 [m^{IDir}, C] (\{\omega_i\}) = C \frac{n_i}{N+s} & \forall i, \\ \text{Cf}_3 [m^{IDir}, C] (\Omega) = C \frac{s}{N+s}, \\ \text{Cf}_3 [m^{IDir}, C] (\Theta) = 1 - C; \end{cases} \quad (16)$$

et donc, pour tout  $\omega_i \in \Omega$ , on obtient les valeurs de croyance et de plausibilité suivantes :

$$\begin{aligned} \text{Cf}_3 [bel^{IDir}, C] (\{\omega_i\}) &= C \frac{n_i}{N+s}, \\ \text{Cf}_3 [pl^{IDir}, C] (\{\omega_i\}) &= C \frac{n_i+s}{N+s} + 1 - C, \end{aligned} \quad (17)$$

et

$$\begin{aligned} \text{Cf}_3 [bel^{IDir}, C] (\{\omega_u\}) &= 0, \\ \text{Cf}_3 [pl^{IDir}, C] (\{\omega_u\}) &= 1 - C. \end{aligned} \quad (18)$$

Notons qu'appliquer  $\delta_{ID}$  à  $Cf_3[m, \cdot]$  conduit à satisfaire la Propriété 3, puisque  $\delta_{ID}(Cf_3[m, \cdot]) = \Theta$ .

**Remarque 1 (Hypothèse de monde ouvert)**

*La classe inconnue  $\omega_u$  joue un rôle similaire à  $\emptyset$  dans l'hypothèse de monde classique (où  $m(\emptyset)$  quantifie la croyance qu'une instance appartient à une classe en dehors de  $\Omega$ ). Cependant, introduire  $\omega_u$  permet de distinguer la croyance que l'individu est atypique avec le degré de conflit possiblement issu de la combinaison de plusieurs fonctions de masses.*

**5 Conclusion et perspectives**

Le Tableau 1 résume les propriétés des opérateurs de conformité et les stratégies de décision associées, en rappelant la fonction de masse utilisée par chaque opérateur, l'ensemble des décisions obtenues pour  $C = 0$ , les propriétés satisfaites, et les hypothèses sur le cadre de discernement. Ces dernières dépendent de l'application visée : l'hypothèse de monde ouvert s'applique clairement en détection de nouveauté, mais est discutable en recherche d'anomalies ou en détection de données aberrantes. Ce travail ouvre plusieurs perspectives. Premièrement nous allons étudier si d'autres propriétés pourraient être requises. Ensuite, nous envisageons de comparer les différentes stratégies basées sur les opérateurs de conformité à d'autres approches traitant des données atypiques, comme les prédictions conformelles.

**Remerciements**

Ce travail a été accompli dans le cadre du LABEX MS2T(ANR-11-IDEX-0004-02) financé par le gouvernement français au travers du programme "investissements d'avenir" géré par l'Agence Nationale pour la Recherche.

**Références**

- [1] Aregui, A., Denoeux, T. : Novelty detection in the belief functions framework. In : Proceedings of IPMU. vol. 6, pp. 412–419 (2006)
- [2] Carreño, A., Inza, I., Lozano, J.A. : Analyzing rare event, anomaly, novelty and outlier detection terms under the supervised classification framework. Artificial Intelligence Review (2019)
- [3] Dempster, A. : Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics **38**, 325–339 (1967)
- [4] Denoeux, T. : Decision-making with belief functions : a review. International Journal of Approximate Reasoning **109**, 87–110 (2019)
- [5] Dubuisson, B., Masson, M. : A statistical decision rule with incomplete knowledge about classes. Pattern Recognition **26**(1), 155 – 165 (1993)
- [6] Shafer, G. : A mathematical Theory of Evidence. Princeton University Press, New Jersey (1976)
- [7] Shafer, G., Vovk, V. : A tutorial on conformal prediction. Journal on Machine Learning Research **9**, 371 – 421 (2008)
- [8] Smets, P. : Belief functions : the disjunctive rule of combination and the generalized bayesian theorem. I. J. of Approximate Reasoning **9**, 1–35 (1993)
- [9] Smets, P., Kennes, R. : The transferable belief model. Artificial Intelligence **66**, 191–234 (1994)

Opérateur de conformité	fonction de masse combinée a m	règle de décision associée	décisions pour C = 0	propriétés satisfaites			type de domaine	
				Prop. 1	Prop. 2	Prop. 3	ouvert	clos
Cf <sub>1</sub>	$m_{\Omega}$	$\delta_{ID}$	$\Omega$			×	×	
Cf <sub>2</sub>	$m_{\omega_u}$	$\delta_{ID}$	$\omega_u$	×	×			×
Cf <sub>2</sub>	$m_{\omega_u}$	$\delta_{ID:AT}$	$\delta_{ID}(m[\Omega]) \cup \{\omega_u\}$	×		×	×	×
Cf <sub>3</sub>	$m_{\Omega \cup \omega_u}$	$\delta_{ID}$	$\Omega \cup \omega_u$			×		×

Tableau 1 – Résumé des opérateurs et de leur règles de décision associées