

PROBABILITE D'EVENEMENTS FLOUS
APPLICATION AU PROBLEME DU FILTRAGE

ANDRES V.

Equipe I.A. et Robotique
Laboratoire "Langages et Systèmes Informatiques"
Université Paul Sabatier
118, route de Narbonne
31062 TOULOUSE Cedex - FRANCE

On a déjà vu (CAYROL 82) que pour estimer le degré d'appariement entre un objectif F et une donnée D, on dispose des 2 indices de filtrage :

$$\Pi(F;D) = \sup \mu_{D \cap F}(x) = \sup \{\min(\mu_D(x), \mu_F(x))\} \quad (1)$$

$$\text{et } N(F;D) = \inf \mu_{D \cap F}(x) = \inf \mu_{\neg D \cup F}(x) = \inf \{\max(1 - \mu_D(x), \mu_F(x))\} \quad (2)$$

qui sont respectivement la possibilité et la nécessité que la donnée D satisfasse l'objectif F.

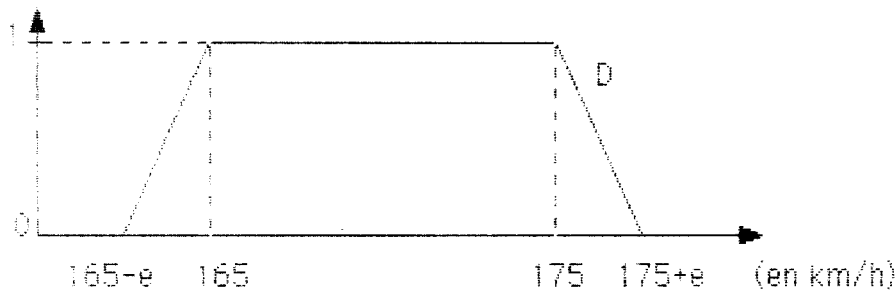
Pour calculer ces 2 indices on dispose :

- de l'événement flou F, qui spécifie l'objectif (ex: "la voiture doit être rapide")
- de la distribution de possibilité $\mu_D(x)$ de la variable donnée, (ex: vitesse maximale d'une CX)

Dans le problème qui nous préoccupe (ANDRES 86a, ANDRES 86b), les distributions de possibilité des données proviennent de la "fuzzification" d'un catalogue de données (l'auto-journal) qui pour un critère, fournit l'intervalle dans lequel se trouve la donnée relative à un type de voiture.

ex :	voiture	vitesse maximale	
	CX	[165 km/h;175 km/h]	
	R4	[120 km/h;130 km/h]	(tableau 1)
	GS	[130 km/h;145 km/h]	

En pratique, les intervalles fournis par ce tableau peuvent s'interpréter de façon floue sous la forme de distributions de possibilités. Par exemple, dans le cas de la vitesse maximale des CX :

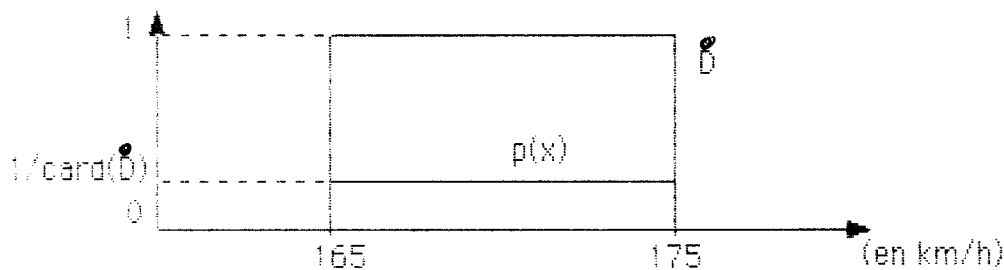


ce est par exemple l'erreur commise par l'instrument de mesure, ou la dispersion des vitesses réelles des voitures).

A partir des données du tableau 1, on peut aussi déduire une distribution de probabilité pour la variable "vitesse maximale des CX".

On propose de prendre une distribution de probabilité équirépartie, et allouée seulement sur l'intervalle spécifié par les données.

$$\text{c.à.d. : } p(x) = \begin{cases} 1/\text{card}(\hat{D}) & \text{si } x \in \hat{D} \\ 0 & \text{sinon} \end{cases} \quad (3)$$



rappel : Dans le cas d'une variable continue, le cardinal scalaire d'un ensemble flou F est remplacé par : $\text{card}(F) = \int_A \mu_F(x).dx$ (4)

Le cardinal d'un intervalle strict est donc égal à sa longueur. \square

Il y a bien entendu d'autres manières de définir une telle distribution de probabilité, mais celle-ci est intéressante pour plusieurs raisons :

1) La probabilité engendrée par cette distribution reste calculable pour des produits ou sommes cartésiennes d'ensembles flous, car elle est "semi-décomposable".

2) la probabilité engendrée par cette distribution appartient à la classe de probabilités vérifiant la condition de cohérence :

$$\Pi(F;D) \geq P(F;D) \geq N(F;D) \quad \forall F, \forall D \quad (5)$$

En effet,

proposition : Si la masse totale d'une distribution de probabilité est allouée seulement sur le noyau de la distribution de possibilité correspondante, alors la probabilité engendrée par cette distribution vérifie la condition (5).

$$\begin{aligned} \text{dém :} \quad & P(F;D) = \int_{\mathcal{A}} \mu_F(x).p(x).dx = \int_{\mathcal{B}} \mu_F(x).p(x).dx \\ \text{et} \quad & \Pi(F;D) = \sup \mu_{D \cap F}(x) \geq \sup \mu_{D \cap \bar{F}}(x) = \sup \mu_{\bar{F}}(x) \\ \Rightarrow \quad & P(F;D) \leq \int_{\mathcal{B}} \Pi(F;D).p(x).dx = \Pi(F;D) \end{aligned}$$

par suite $N(F;D) = 1 - \Pi(\bar{F};D) \leq 1 - P(\bar{F};D) = P(F;D)$ □

Cette probabilité vérifie en fait une condition plus forte que (5), à savoir :

$$P(A;D) > 0 \Rightarrow \Pi(A;D) = 1 \quad \text{pour } A \text{ non-flou} \quad (6)$$

i.e. pour qu'un événement non-flou soit un tant soit peu probable, il faut qu'il soit d'abord totalement possible. ((6) n'est plus vrai pour A flou, où on a seulement (5).)

3) Cette distribution de probabilité équirépartie sur le noyau, traduit "bien" (du point de vue probabiliste) notre incertitude totale quant à la valeur réelle de la variable à l'intérieur du noyau.

Avec la distribution de probabilité ainsi définie, on obtient donc un indice d'appariement supplémentaire pour l'objectif F : la probabilité de l'événement flou F, définie par Zadeh (ZADEH 68) comme :

$$P(F) = \int_{\mathcal{A}} \mu_F(x).p(x).dx \quad (7)$$

N.B.1 : On emploie indifféremment la notation $P(F)$ ou $P(F;D)$, cette dernière servant uniquement à spécifier la distribution à laquelle la probabilité fait référence.

N.B.2 : Il existe d'autres définitions de la probabilité d'un événement flou, notamment la définition d'une probabilité floue (on peut en effet s'étonner qu'un événement flou puisse engendrer une probabilité précise). On consultera

(DUBOIS 80, YAGER 79) pour plus de précision. □

C'est bien une probabilité, puisque cette définition vérifie encore :

$$- P(\Omega) = 1 \quad (8)$$

$$- P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (9)$$

avec les définitions classiques de l'union et de l'intersection ensembliste floue :

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$$

$$\text{et } \mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$$

remarque : notons que si l'on considère l'ensemble flou D, dont la fonction d'appartenance est donnée par la distribution de possibilité $\mu_D(x)$, on a

$$\text{alors } P(F/D) = P(F) \quad (10)$$

en effet,

$$\begin{aligned} P(F/D) &= P(F \cap D) && (\text{car } P(D) = \int_A \mu_D(x).p(x).dx = \int_B p(x).dx = 1) \\ &= \int_A \mu_{F \cap D}(x).p(x).dx \\ &= \int_A \min(\mu_F(x), \mu_D(x)).p(x).dx \\ &= \int_A \min(\mu_F(x).p(x), p(x)).dx \\ &= \int_A \mu_F(x).p(x).dx \end{aligned}$$

$$\Rightarrow P(F/D) = P(F)$$

$$(\text{à rapprocher de : } \Pi(D \cap F; D) = \Pi(F; D) \quad (11)$$

$$N(D \cap F; D) = N(F; D) \quad (12))$$

Une formule simple pour P(F)

Si on prend la distribution de probabilité définie par (3), on a alors :

$$P(F) = \int_A \mu_F(x).p(x).dx = \frac{1}{\text{card}(\hat{D})} \cdot \int \mu_F(x).dx = \text{card}(F \cap \hat{D}) / \text{card}(\hat{D}) \quad (13)$$

remarque : La formule (13) fait immédiatement penser à la formule voisine, $\text{card}(F \cap D) / \text{card}(D)$, mais cette expression ne s'interprète de façon cohérente qu'en tant que probabilité conditionnelle, avec une distribution de probabilité constante et répartie au moins sur le support de D.

Par exemple, si la distribution de probabilité est constante sur le

support de D , et vaut donc $1/\text{card}(S(D))$, on a :

$$P(F/D) = P(F \cap D) / P(D) = \frac{\text{card}(F \cap D \cap S(D)) / \text{card}(S(D))}{\text{card}(D \cap S(D)) / \text{card}(S(D))} = \frac{\text{card}(F \cap D)}{\text{card}(D)} \quad (14)$$

Indépendamment de la forme de la distribution de probabilité, l'interprétation de $|F \cap D|/|D|$ en tant que probabilité conditionnelle ne présente en fait que peu d'intérêt, en effet : $P(F/D) = P(d \in F / d \in D)$ où d est la variable qui suit précisément la distribution de probabilité relative à D . Le fait de conditionner $P(F)$ par l'information $d \in D$ n'apporte donc rien ! Il s'agit en fait plutôt d'une normalisation que d'un réel conditionnement.

Dans le cas où D est non-flou on a d'ailleurs $P(F/D) = P(F)$, puisque $P(D) = 1$.

Des formules équivalentes à (13) s'obtiennent en prenant par exemple des distributions de probabilités équiréparties sur l'espérance de D , $\text{Esp}(D)$ ou sur le support de D , $S(D)$. Elles conduisent à :

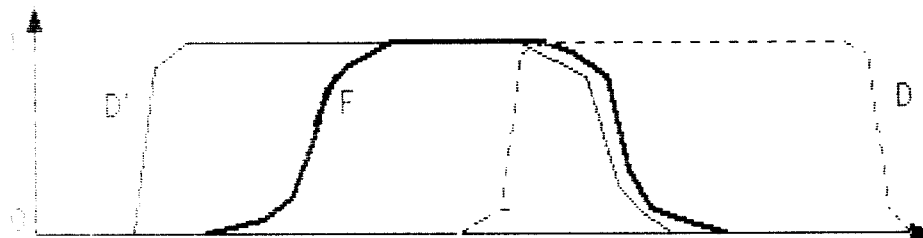
$$P(F;D) = \text{card}(F \cap \text{Esp}(D)) / \text{card}(\text{Esp}(D)) \quad (15)$$

$$\text{et } P(F;D) = \text{card}(F \cap S(D)) / \text{card}(S(D)) \quad (16)$$

mais ces probabilités ne vérifient pas l'inégalité (5), sauf dans le cas où on a affaire à des ensembles flous trapézoïdaux.

Complémentarité des indices d'appariement $P(F;D)$, $\Pi(F;D)$ et $N(F;D)$

A l'origine, on a cherché un indice d'appariement supplémentaire, car on s'est aperçu que derrière des couples d'appariement (Π, N) identiques, pouvaient se cacher des données en fait très différentes (cf schéma).



$\Pi(F, D') = \Pi(F, D) = 1$ et $N(F, D') = N(F, D) = 0$, mais visiblement D' a "plus de chances" de matcher avec F que D .

Pour les discriminer on a alors naturellement pensé à l'indice de

recouvrement $\text{card}(F \cap D) / \text{card}(D)$, mais cet indice ne s'étend pas au cas où l'objectif global est composé d'un produit ou d'une somme cartésienne d'objectifs élémentaires, car on n'a pas de formule décomposable pour le calculer; et de plus il ne vérifie pas la condition de cohérence (5).

L'indice $P(F;D)$ défini en (13) est par contre encore calculable pour des produits ou sommes cartésiennes d'objectifs élémentaires, et est de plus extensible au cas où les objectifs sont pondérés.

L'indice $P(F;D)$ appliqué au produit cartésien d'événements flous

C'est le cas où l'objectif global est une conjonction d'objectifs élémentaires (ex : rapide et grande). (La donnée est par nature toujours conjonctive : $D=D_1 \times D_2$).

Pour des ensembles F_1 et F_2 non-flous, si on considère les distributions de probabilité engendrées par D_1 et D_2 comme étant indépendantes, on obtient naturellement une formule décomposable pour $P(F_1 \times F_2)$:

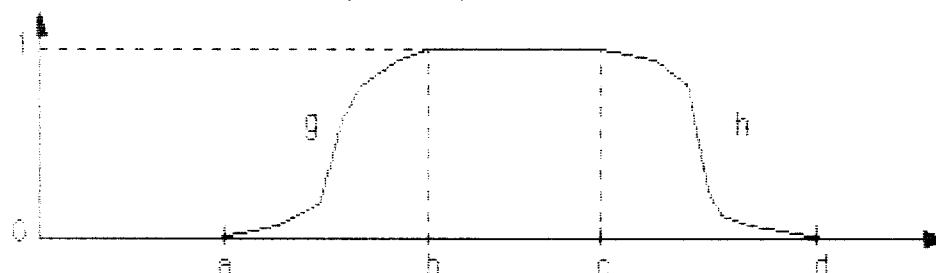
$$P(F_1 \times F_2) = P(F_1).P(F_2) \quad (17)$$

Par contre, dans le cas où F_1 et F_2 sont flous, même si on considère les distributions comme étant indépendantes, on n'obtient plus une formule décomposable, puisque le produit cartésien de 2 ensembles flous ne vérifie plus $\text{card}(F_1 \times F_2) = \text{card}(F_1).\text{card}(F_2)$. (DUBOIS 85a).

On dispose néanmoins d'une formule "semi-décomposable", puisque :

proposition :
$$P(F) = \int_0^1 P(F_y) dy \quad (18)$$

gém : Pour F un ensemble flou quelconque



$$P(F) = \int_a^d \mu_F(x) dP(x) = \int_a^d \mu_F(x) dP(x) = \int_a^d P(x) d\mu_F(x) \quad (\text{int. par parties})$$

$$\begin{aligned}
\Rightarrow P(F) &= \int_a^c P(x).dh(x) + \int_b^a P(x).dg(x) \\
&= \int_a^1 P(h^{-1}(y)).dy - \int_0^1 P(g^{-1}(y)).dy \quad (\text{chgt. de variable}) \\
&= \int_a^1 \left(\int_{g^{-1}(y)}^{h^{-1}(y)} dP(u) \right).dy = \int_a^1 P(F_y).dy \quad \square
\end{aligned}$$

rg : On notera (DUBOIS 85a) que la probabilité telle que définie en (18), est égale au cardinal de la probabilité floue $P^*(F)$ définie par Yager comme étant :

$$\mu_{P^*(F)}(P(F_y)) = y \quad (19) \square$$

Calcul pratique de $P(F)$

Avec la probabilité définie en (3), (18) peut se réécrire plus simplement. En effet, dans ce cas :

$$\begin{aligned}
P(F) &= \int_a^1 P(F_y).dy \\
&= \int_a^{N(F;\hat{D})} P(F_y).dy + N(F;\hat{D}) \int_{N(F;\hat{D})}^{\Pi(F;\hat{D})} P(F_y).dy + \Pi(F;\hat{D}) \int^1 P(F_y).dy
\end{aligned}$$

$$\text{et } \Pi(F;\hat{D})=1 \Rightarrow \int_{\Pi(F;\hat{D})}^1 P(F_y).dy = 0$$

$$\Pi(F;\hat{D}) < 1 \Rightarrow \forall y \geq \Pi(F;\hat{D}), P(F_y) = |F_y \cap \hat{D}| / |\hat{D}| = |\emptyset| / |\hat{D}| = 0$$

$$\Rightarrow \int_{\Pi(F;\hat{D})}^1 P(F_y).dy = 0$$

$$\text{donc } \forall \Pi(F;\hat{D}), \int_{\Pi(F;\hat{D})}^1 P(F_y).dy = 0.$$

duallement

$$N(F;\hat{D})=0 \Rightarrow \int_a^{N(F;\hat{D})} P(F_y).dy = 0 = N(F;\hat{D})$$

$$N(F;\hat{D}) > 0 \Rightarrow \forall y \leq N(F;\hat{D}), P(F_y) = |F_y \cap \hat{D}| / |\hat{D}| = |\hat{D}| / |\hat{D}| = 1$$

$$\Rightarrow \int_a^{N(F;\hat{D})} P(F_y).dy = N(F;\hat{D})$$

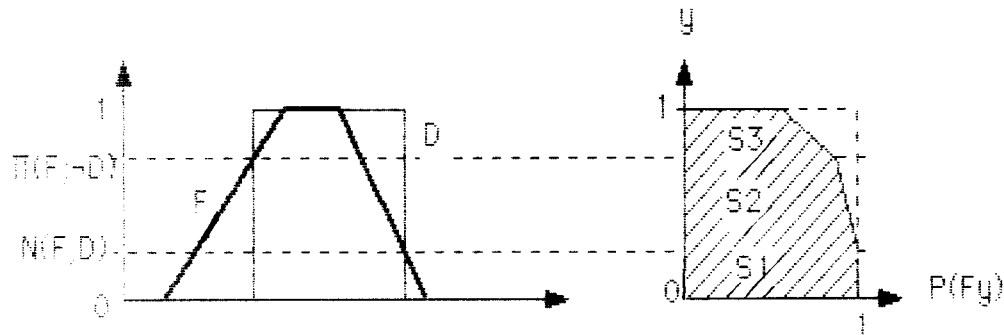
$$\text{donc } \forall N(F;\hat{D}), \int_a^{N(F;\hat{D})} P(F_y).dy = N(F;\hat{D})$$

$$\Rightarrow P(F) = N(F;\hat{D}) + N(F;\hat{D}) \int_{N(F;\hat{D})}^{\Pi(F;\hat{D})} P(F_y).dy \quad (20)$$

1) Cas où F est un ensemble flou unidimensionnel

Dans ce cas, $P(F_y) = |F_y \cap \hat{D}| / |\hat{D}|$ est linéaire par morceaux en fonction de y , et l'on doit alors distinguer 2 cas, suivant le nombre de morceaux :

1.1) $\Pi(F; -\hat{D}) < 1 \Rightarrow \Pi(F; \hat{D}) = 1 \Rightarrow$ on a 2 morceaux :



$$\begin{aligned}
 P(F) &= N(F; \hat{D}) + \int_{N(F; \hat{D})}^{\Pi(F; -\hat{D})} P(F_y) dy + \int_{\Pi(F; -\hat{D})}^1 P(F_y) dy \\
 &= N(F; \hat{D}) + (1/2) \cdot \{ (\Pi(F; -\hat{D}) - N(F; \hat{D})) \cdot (P(F_{\Pi(F; -\hat{D})}) + P(F_{N(F; \hat{D})})) \\
 &\quad + (1 - \Pi(F; -\hat{D})) \cdot (P(F_1) + P(F_{\Pi(F; -\hat{D})})) \} \\
 &= N(F; \hat{D}) + (1/2) \cdot \{ (\Pi(F; -\hat{D}) - N(F; \hat{D})) \cdot (P(F_{\Pi(F; -\hat{D})}) + P(S(F))) \\
 &\quad + (1 - \Pi(F; -\hat{D})) \cdot (P(F) + P(F_{\Pi(F; -\hat{D})})) \} \\
 &\text{(car : } P(F_{N(F; \hat{D})}) = P(S(F)), \quad \forall N(F; \hat{D}) \quad (21))
 \end{aligned}$$

1.2) $\Pi(F; -\hat{D}) = 1 \Leftrightarrow \Pi(F; \hat{D}) \leq 1 \Rightarrow$ on a un seul morceau :

$$\begin{aligned}
 P(F) &= N(F; \hat{D}) + \int_{N(F; \hat{D})}^{\Pi(F; \hat{D})} P(F_y) dy \\
 &= N(F; \hat{D}) + (1/2) \cdot (\Pi(F; \hat{D}) - N(F; \hat{D})) \cdot (P(F_{\Pi(F; \hat{D})}) + P(F_{N(F; \hat{D})})) \\
 &= N(F; \hat{D}) + (1/2) \cdot (\Pi(F; \hat{D}) - N(F; \hat{D})) \cdot (P(F) + P(S(F))) \\
 &\text{(car : } P(F_{\Pi(F; \hat{D})}) = P(F), \quad \forall \Pi(F; \hat{D}) \quad (22))
 \end{aligned}$$

Malgré la complexité apparente de ces formules, l'algorithme de calcul de $P(F; D)$ reste tout de même relativement rapide, puisque dans tous les sous-cas (en fonction de $\Pi(F; \hat{D})$ et $N(F; \hat{D})$), sauf un, ces formules se simplifient.

proposition : si $D=[a;b]$ et F est tel que $\mu_F(x)$ est continue en a ou b

$$\text{alors} \quad \Pi(F;D) \geq N(F;D) \quad (23)$$

preuve : $\mu_F(x)$ continue en a

$$\Rightarrow \Pi(F;D) = \sup \mu_F(x) \geq \mu_F(a) \geq \inf \mu_F(x) = N(F;D)$$

idem pour $\mu_F(x)$ continue en b . □

2) Cas où F est un produit cartésien d'ensembles flous

Si F est un produit cartésien d'ensembles flous, $F = \prod_{i=1}^n (F_i)$. La fonction $P(F;D) = \int_a^b \prod_{i=1}^n P((F_i);y)$ sera polynomiale (de degré maximal n) par morceaux, avec au maximum $2.n$ morceaux.

L'ensemble des bornes des morceaux est :

$$Y = 0 \cup \{N(F_i;D_i)\} \cup \{\Pi(F_i;-D_i)\} \cup 1$$

Plutôt que de calculer toutes ces bornes, et faire une intégration linéaire qui ne sera de toute façon qu'une approximation supérieure, on propose de calculer (20) par une somme de Riemann, ce qui est du point de vue algorithmique plus rapide.

L'indice $P(F;D)$ appliqué à la somme cartésienne d'événements flous

C'est le cas où l'objectif global est une disjonction d'objectifs élémentaires (ex : rapide ou grande).

Le produit et la somme cartésienne étant définis par min et max, comme pour l'intersection et l'union, on a :

$$P(F_1+F_2; D_1 \times D_2) = P(F_1;D_1 \times D_2) + P(F_2;D_1 \times D_2) - P(F_1 \times F_2; D_1 \times D_2)$$

$$\Rightarrow P(F_1+F_2; D_1 \times D_2) = P(F_1;D_1) + P(F_2;D_2) - P(F_1 \times F_2; D_1 \times D_2) \quad (24)$$

$$\begin{aligned} \text{car } P(F_i;D_1 \times D_j) &= \int_a^b \mu_{F_i}(x).p_{ij}(x,y).dx.dy \\ &= \int_a^b \mu_{F_i}(x).p_i(x).dx \cdot \int_a^b p_j(y).dy \\ &= \int_a^b \mu_{F_i}(x).p_i(x).dx = P(F_i;D_i) \end{aligned} \quad (25)$$

remarque . on a toujours $P(\text{OU}_{\alpha}^n(F_i)) \in [0;1]$, en effet .

$$\begin{aligned} P(F_1+F_2) &= P(F_1) + P(F_2) - P(F_1 \times F_2) \\ &= \int_0^1 [P((F_1)_y) + P((F_2)_y) - P((F_1)_y)P((F_2)_y)] dy \quad \in [0;1] \end{aligned}$$

d'où par récurrence $P(\text{OU}_{\alpha}^n(F_i)) \in [0;1]$.

L'indice $P(F;D)$ appliqué à l'agrégation d'événements flous pondérés.

On est confronté au même problème qu'avec les indices Π et N (évidemment puisque $\Pi \geq P \geq N$), en effet, lors du calcul de la probabilité d'un produit cartésien d'événements flous, si un seul des événements, par exemple F_j vérifie $P(F_j)=0$, alors la probabilité globale vaut 0. En effet :

$$\begin{aligned} P(F_j)=0 &\Rightarrow P((F_j)_y)=0 \quad \forall y \in [0;1] \\ &\Rightarrow P(\text{ET}_{\alpha}^n(F_i)) = \int_0^1 \prod_{i=1}^n P((F_i)_y) dy = 0 \end{aligned}$$

Dualement, lors du calcul de la probabilité d'une somme cartésienne d'événements flous, si un seul des événements, par exemple F_j vérifie $P(F_j)=1$, alors la probabilité globale vaut 1. En effet :

$$\begin{aligned} P(F_j)=1 &\Rightarrow P((F_j)_y)=1 \quad \forall y \in [0;1] \\ &\Rightarrow P(F_g \times F_j) = \int_0^1 P((F_g)_y) P((F_j)_y) dy = P(F_g) \quad \forall F_g \end{aligned}$$

et donc $P(\text{OU}_{\alpha}^n(F_i)) = P(F_j) + P(\text{OU}_{\alpha}^{n-1}(F_i)) - P(\text{OU}_{\alpha}^{n-1}(F_i) \times F_j)$
 $= 1 + P(\text{OU}_{\alpha}^{n-1}(F_i)) - P(\text{OU}_{\alpha}^{n-1}(F_i)) = 1$

Dans les 2 cas, si la réalisation de l'événement en question est peu importante, ce résultat est navrant. Il faut donc chercher un moyen de pondérer qui atténue l'effet minimisant d'une agrégation ET et l'effet maximisant d'une agrégation OU.

Lorsqu'on veut prendre en compte le fait que les différents objectifs peuvent ne pas avoir la même importance, deux choix sont possibles :

- soit, laisser les objectifs tels quels, et prendre en compte la pondération au niveau de l'opération d'agrégation.
- soit, prendre en compte la pondération au niveau des objectifs, et faire une agrégation normale.

Pour les indices d'appariement Π et N , avec la pondération adoptée, on a vu qu'il existe une équivalence entre ces 2 traitements. En effet, si les π_i sont les importances des objectifs F_i :

$$\Pi(ET(F_i, \pi_i)) = \min\{ \max(\Pi(F_i, 1-\pi_i)) \} = \Pi(ET(F_i^*)) \quad (26)$$

$$N(ET(F_i, \pi_i)) = \min\{ \max(N(F_i, 1-\pi_i)) \} = N(ET(F_i^*)) \quad (27)$$

$$\Pi(OU(F_i, \pi_i)) = \max\{ \min(\Pi(F_i, \pi_i)) \} = \Pi(OU(F_i^*)) \quad (28)$$

$$N(OU(F_i, \pi_i)) = \max\{ \min(N(F_i, \pi_i)) \} = N(OU(F_i^*)) \quad (29)$$

$$\text{ou} \quad \mu_{F^*}(x) = \max(\mu_F(x), 1-\pi) \quad (30)$$

$$\mu_{F^*}(x) = \min(\mu_F(x), \pi) \quad (31)$$

avec cette pondération, on atténue l'effet minimisant d'une agrégation ET, et l'effet maximisant d'une agrégation OU. (ANDRES 86b, DUBOIS 85b).

Les coefficients π_i s'interprétant de façon possibiliste, il ne paraît pas naturel de vouloir les exploiter dans une agrégation probabiliste (du moins pas directement, tout au plus après leur avoir fait subir une transformation possibilité \rightarrow probabilité, idoine).

On propose donc plutôt de les utiliser en adoptant le deuxième choix, c'est à dire en posant :

$$P(ET(F_i, \pi_i)) = P(ET(F_i^*)) \quad (32)$$

$$\text{et} \quad P(OU(F_i, \pi_i)) = P(OU(F_i^*)) \quad (33)$$

où F^* et F_* sont définis par (30) et (31).

Il y a bien entendu une quantité d'autres façons de définir F^* et F_* pour pondérer F avec une importance π , et il n'est pas sûr que la façon de pondérer adoptée soit la meilleure. Néanmoins, on peut tenter de justifier ce choix intuitivement, en notant que puisque ce choix est bon pour les 2 indices Π et N , par continuité, il doit aussi l'être pour P , puisque $\Pi \geq P \geq N$. Cette question est de toute façon encore à l'étude.

Conclusion

Pour que la probabilité d'un produit cartésien soit décomposable, il est nécessaire de faire l'hypothèse d'indépendance des différentes distributions de probabilité, ce n'est pas toujours justifié, ex : nerveuse et sobre, cas où l'on serait même tenté d'écrire : sobre \Rightarrow \neg nerveuse, ce qui est bien loin de l'indépendance !

On peut néanmoins lever cette critique en notant qu'elle ne s'applique qu'à une conjonction d'objectifs incohérents, cas où pour la plupart des voitures, la probabilité est donc voisine de 0 (à moins d'avoir une voiture "incohérente" dans la base de données).

L'un des reproches qu'on entend parfois à l'égard des probabilités, est que pour un produit cartésien d'événements indépendants, dès que le produit cartésien contient un peu trop d'événements, la probabilité globale (résultant d'un produit) tend vers 0, quelles que soient les probabilités partielles, ce qui rend le résultat peu significatif.

Avec les ensembles flous, la probabilité du produit cartésien échappe à cette critique, puisqu'elle vérifie :

$$P(F_1 \times F_2) \geq P(F_1) \cdot P(F_2) \quad \text{si } F_1 \text{ et } F_2 \text{ sont flous.} \quad (34)$$

Cette probabilité ne va en fait plus varier tout à fait comme un produit, mais comme une combinaison d'un produit et d'un minimum, ce qui l'empêchera de décroître trop vite.

Les questions qui se posent maintenant sont, possédant les 3 indices d'appariement Π , P et N , dont la signification est clairement définie :

- ces 3 indices sont-ils suffisants pour permettre un classement des données valable ?, ou faut-il encore chercher d'autres indices ? Par exemple, des "indices horizontaux" comme l'entropie d'Higashi et Klir ou la distance de Hausdorff étendue aux ensembles flous.
- comment utiliser ces différents indices pour classer les données (en dehors de l'ordre de Pareto !).

Références

- /ANDRES 86a/ ANDRES V., DUBOIS D., PRADE H., TESTEMALE C. (1986)
Notion d'importance relative dans des requêtes à une base de données imprécises : Expérimentation sur micro-ordinateur. Journées d'étude, Systèmes de gestion de bases de données sur micro-ordinateurs. La Rochelle. Eyrolles, Paris.
- /ANDRES 86b/ ANDRES V. Une application de la théorie des possibilités : Le Filtrage Sémantique Souple. Rapport de DEA 86.
- /CAYROL82/ CAYROL M., FARRENY H., PRADE H., (1982) Fuzzy Pattern Matching. Kybernetes 11, 103-116.
- /DUBOIS 80/ DUBOIS D., PRADE H., (1980) Fuzzy Sets and Systems : Theory and applications, Vol. 144 in Mathematics in Sciences and Engineering series. Academic Press, New-York, 36-41.
- /DUBOIS 85a/ DUBOIS D., PRADE H. Fuzzy cardinality and the modelling of imprecise quantification. Fuzzy sets and systems 16 (1985) 199-230. North-Holland Publishing Company.
- /DUBOIS 85b/ DUBOIS D., PRADE H. (1985). Théorie des possibilités. Application à la représentation des connaissances en informatique. MASSON. PARIS.
- /YAGER 79/ YAGER R., A note on probability of fuzzy events. Information sciences 18, 113-129 (1979).
- /ZADEH 68/ ZADEH L., Probability measures of fuzzy events. Journal of mathematical analysis and applications 23, 421-427 (1968).