Ensembles épais, fonctions multivoques, et théorie des possibilités

Thick Sets, Multiple-Valued Mappings, and Possibility Theory

Didier Dubois[1] Luc Jaulin[2] Henri Prade[1]
[1] IRIT – CNRS, 118, route de Narbonne, 31062 Toulouse Cedex 09, France
[2] Lab-STICC, ENSTA-Bretagne, 2 rue François Verny, 29200 Brest, France

Résumé :
Le transport d'une information incertaine par une fonction multivoque se rencontre dans différents cadres, allant du calcul de l'image d'un ensemble par une fonction inverse au transport Dempstérien d'un espace probabilisé par une fonction multivoque. On obtient alors des images supérieures et inférieures. Dans chaque cas on manipule des ensembles dits "épais" au sens de Jaulin, c'est-à-dire des ensembles mal connus bornés inférieurement et supérieurement. De tels ensembles mal connus se rencontrent sous diverses appellations dans la littérature, comme celles, par exemple, d'"ensemble-intervalle" chez Y. Y. Yao, de "twofold fuzzy sets" au sens de Dubois et Prade, ou d'"interval-valued fuzzy sets",... Différentes opérations peuvent être définies sur ces ensembles qui sont interprétés disjonctivement (incertitude épistémique), plutôt que conjonctivement. Le but de cette note est de proposer une vision unifiée de ces formalismes dans le cadre de la théorie des possibilités, ce qui permettra de donner des extensions graduelles à certains des calculs considérés.

Mots-clés :
Ensemble épais, calcul d'intervalles, théorie des possibilités, image inverse, incertitude.

Abstract:
Carrying uncertain information via a multivalued function can be found in different settings, ranging from the computation of the image of a set by an inverse function to the Dempsterian transfer of a probabilistic space by a multivalued function. We then get upper and lower images. In each case one handles so-called "thick sets' in the sense of Jaulin, i.e., lower and upper bounded ill-known sets. Such ill-known sets can be found under different names in the literature, e.g., "interval sets" after Y. Y. Yao, "twofold fuzzy sets" in the sense of Dubois and Prade, or "interval-valued fuzzy sets",... Various operations can then be defined on these sets, then understood in a disjunctive manner (epistemic uncertainty), rather than conjunctively. The intended purpose of this note is to propose a unified view of these formalisms in the setting of possibility theory, which should enable us to provide graded extensions to some of the considered calculi.

Keywords:
Thick set, interval analysis, possibility theory, inverse image, uncertainty

Thick set, interval analysis, possibility theory, inverse image, uncertainty.

1 Introduction

Les liens entre le calcul d'intervalles [16] et la théorie des possibilités [8] sont bien connus, de même que l'intérêt du calcul d'intervalles en commande robuste [15]. Les besoins d'approximation garantie ont conduit B. Desrochers et L. Jaulin à proposer un calcul original d'ensembles et d'intervalles "épais" [3, 4]. Cette petite note débute l'étude des liens de ce calcul avec d'autres travaux en matière d'incertitude développés dans le cadre des théories des possibilités et des ensembles flous.

2 Ensembles épais et autres

Un ensemble épais [4] $[[\mathbb{A}]]$ sur un référentiel U (en général $\mathbb{R}^n$) est un intervalle de 2^U défini par une paire $(\mathbb{A}_*, \mathbb{A}^*)$ telle que $\mathbb{A}_* \subset \mathbb{A}^*$, c'est-à-dire

$$[[\mathbb{A}]] = [\mathbb{A}_*, \mathbb{A}^*] = \{\mathbb{A} \in 2^U | \mathbb{A}_* \subset \mathbb{A} \subset \mathbb{A}^*\} \quad (1)$$

Il s'agit donc d'un ensemble mal connu compris en deux bornes. Formellement on peut le représenter par un ensemble flou dont la fonction d'appartenance μ est trivaluée : $\mu : U \to \{0, 1/2, 1\}$ comme par exemple dans les ensembles flous au sens de Gentilhomme [14] qui représentent des concepts avec une zone centrale $\mathbb{A}_*$ et une zone périphérique $\mathbb{A}^* \setminus \mathbb{A}_*$. Mais aussi les ensembles basés sur la logique de Kleene, pour lesquels, dans la zone périphérique, l'information pertinente pour conclure à l'appartenance ou non est incomplète (1/2 veut dire inconnu). Par exemple, c'est le cas pour les "rough sets" [17] où l'incertitude provient d'un manque d'attributs pour décrire parfaitement un ensemble d'objets, ou les twofold sets [7] où l'incertitude provient d'un manque d'information sur les valeurs des attributs des objets. Voir aussi les "interval sets" [18].

De tels ensembles ont aussi été introduits dans la littérature des ensembles flous à diverses époques. Mentionnons en particulier les ensembles flous intervallistes (Zadeh, Sambuc) qui sont des ensembles flous épais au sens de Jaulin, à savoir, des paires $(\mathbb{F}_*, \mathbb{F}^*)$ d'ensembles flous qui encadrent un ensemble flou mal connu $\mathbb{F}$: $\mu_{\mathbb{F}_*} \leq \mu_{\mathbb{F}} \leq \mu_{\mathbb{F}^*}$ (un cas particulier d'ensemble flou de type 2). On a aussi les "twofold fuzzy sets" [7], paires emboitées de deux ensembles flous (le support de l'un étant inclus dans le noyau de l'autre). Ils représentent un ensemble d'éléments qui appartiennent plus ou moins nécessairement (certainement) à $\mathbb{F}$, inclus dans un surensemble d'éléments qui appartiennent plus ou moins possiblement à $\mathbb{F}$, induit par le fait que l'information pertinente pour conclure à l'appartenance ou non est incomplète.

3 Ensembles épistémiques et ensembles ontiques

Pour donner un sens clair aux ensembles épais, il est important de comprendre ce que représentent les ensembles manipulés dans leur définition. Un ensemble, classique ou flou, peut représenter

— soit une entité constituée vue comme la *conjonction* de ses éléments - on parle alors d'*ensemble ontique*, puisque l'ensemble soit représente un objet réel, soit constitue l'entité qu'on cherche à identifier.

— soit un ensemble de valeurs possibles mutuellement exclusives pour une variable - on parle alors d'ensemble épistémique puisque reflétant une information imprécise sur la valeur de la variable.

Cette distinction est cruciale pour une manipulation correcte des ensembles dans les calculs. Ainsi un ensemble épais, comme ensemble d'ensembles, est épistémique et modélise un ensemble mal connu, lequel est considéré comme ontique ; par exemple, une zone dont on voudrait garantir la couverture [11], par exemple pour s'assurer qu'un robot puisse passer entre deux obstacles [12, 13] est un ensemble ontique.

4 La construction de Dempster

Un exemple d'intervalle épais est constitué par la paire d'ensembles de solutions inférieur A_* et supérieur A^* de l'équation ensembliste

$$f(S) = A \subseteq \Omega$$

où A est un ensemble et $f : U \to \Omega$ une fonction mal connue appartenant à un ensemble Γ de fonctions. Il s'agit d'un problème d'inversion ensembliste épais dont les solutions S satisfont toutes $S \in [A_*, A^*]$, avec

$$\begin{aligned} A_* &= \{u : \forall f \in \Gamma, \exists a \in A \mid a = f(u)\} \\ &= \{u : \Gamma(u) \subseteq A\} = \textstyle\bigcap_{f \in \Gamma} f^{-1}(A) ; \\ A^* &= \{u : \exists f \in \Gamma, \exists a \in A \mid a = f(u)\} \\ &= \{u : \Gamma(u) \cap A \neq \emptyset\} = \textstyle\bigcup_{f \in \Gamma} f^{-1}(A). \end{aligned}$$

Dempster [1] utilise ce modèle pour induire des probabilités inférieures et supérieures à partir d'un espace probabiliste (U, P) et d'une fonction multivoque $\Gamma : U \to \Omega$. Cette dernière représente la connaissance incomplète sur une variable aléatoire, soit une fonction f qui relie l'espace d'épreuves à l'espace d'observations Ω. La valeur $x = f(u)$ est une mesure d'une caractéristique de x. Si une expérience aléatoire donne un résultat u, l'observation résultante est une valeur mal connue $x = f(u) \in \Gamma(u)$ parce que l'instrument de mesure qui devrait fournir $x = f(u)$ est imparfait. On a affaire à une variable aléatoire mal observée.

On ne connaît donc pas la mesure de probabilité $P_f(A) = P(f^{-1}(A))$ de l'événement $f(u) \in A$ sur Ω, mais seulement des bornes supérieures : $P^*(A) = P(\{u : \Gamma(u) \cap A \neq \emptyset\})$ et inférieures : $P_*(A) = P(\{u : \Gamma(u) \subseteq A\})$. La même construction peut se faire à partir d'un espace possibiliste [6].

Dans ce modèle, on utilise donc la description de l'image inverse $f^{-1}(A)$ d'un événement $A \subset \Omega$ quand on connaît mal la fonction f. C'est le sous-ensemble épais $[A_*, A^*]$ de U. L'intervalle $[P_*(A), P^*(A)] = [P(A_*), P(A^*)]$ est la "probabilité" de ce sous-ensemble épais, et représente l'ensemble des valeurs possibles de la probabilité $P(f^{-1}(A))$ quand $f \in \Gamma$.

5 Cas de l'arithmétique d'intervalles

Une illustration de ce qui précède est constitué par la paire d'ensembles de solutions inférieur X_* et supérieur X^* de l'équation $x - u = v$ (et donc $x = u + v$) où $u \in M, v \in N$ M, N étant des intervalles.

On peut interpréter l'équation $x - u = v$ dans un contexte incertain de deux façons :

- Chercher l'ensemble
$X^* = \{x : \exists u \in M, v \in N, \text{tel que } x = u+v\}$.

- ou chercher l'ensemble
$X_* = \{x : \forall u \in M, \exists v \in N, \text{tel que } x = u+v\}$.

Ces ensembles maximaux et minimaux sont respectivement donnés par deux opérations d'addition ensemblistes :

$$\begin{aligned} X^* = M \oplus N =& \{x : (x \ominus M) \cap N \neq \emptyset\} \\ =& \{u + v : u \in M, v \in N\} \\ X_* = M \boxplus N =& \{x : (x \ominus M) \subseteq N\} \end{aligned}$$

avec $x \ominus M = \{x - u \mid u \in M\}$. $M \boxplus N$ est l'ensemble des x tel que $-M$ translaté de x soit inclus dans N.

Par exemple, si $M = [m, m']$ et $N = [n, n']$, on a $M \oplus N = [m + n, m' + n']$, et $M \boxplus N = [m + n', m' + n]$ si $m + n' \leq m' + n$ et $M \boxplus N = \emptyset$ sinon. On peut vérifier que $M \boxplus N \subseteq M \oplus N$, et que la longueur de $M \boxplus N$ est la longueur de M diminuée de celle de N.

L'opération $\oplus$ est dite optimiste, et l'opération $\boxplus$ est dite pessimiste. On vérifie :

$$\begin{aligned} M \boxplus N &= \bigcap_{f \in \Gamma} f^{-1}(N) = \bigcap_{u \in M} u \oplus N, \\ M \oplus N &= \bigcup_{f \in \Gamma} f^{-1}(N) = \bigcup_{u \in M} u \oplus N, \end{aligned}$$

où $u \oplus N = \{u + v \mid v \in N\}$ et $x \ominus M$ joue le rôle de $\Gamma(x)$. On a donc un cas particulier de la construction de Dempster où U est le domaine de X, et Ω est le domaine de v. Dans les notations de la section précédente, on devrait écrire $M \boxplus N = N_*$ et $M \oplus N = N^*$. Mais ici on voit que $M \boxplus N$ et $M \oplus N$ ne résolvent pas le même problème.

Si M et N sont des ensembles épistémiques représentant des valeurs mal connues, $M \oplus N$ décrit l'incertitude sur x induite par celle sur u et v. Si N représente un intervalle de tolérance à respecter, $M \boxplus N$ décrit les valeurs de x autorisées pour s'assurer que l'incertitude de $x - u$ reste bornée par N malgré ses fluctuations dues à la mauvaise connaissance sur u, décrite par M.

Mais, supposons que M et N soient ontiques, et représentent la position sur la ligne réelle de deux barres. Alors, la longueur de $M \oplus N$ est celle de la barre obtenue par concaténation de M et N. En revanche, $M \boxplus N$ est l'ensemble des points certainement couverts par la barre M si on la translate d'une longueur $v \in N$.

Ces deux opérations $\boxplus$ et $\oplus$ se généralisent quand M et N sont des intervalles flous [5].

Un exemple : le problème des deux chèvres : Soient deux chèvres, chacune attachée

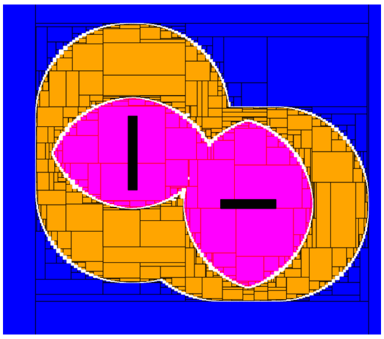

Figure 1 – La zone broutée contient l'ensemble $\mathbb{A}_*$ (gris) et est contenue dans $\mathbb{A}^*$ (gris + gris clair)

à un pieu par une corde de longueur 10m. La position des pieux, $\mathbf{m}_i$ pour la chèvre i, $i \in \{1,2\}$ est mal connue. On sait juste que $\mathbf{m}_1 \in [\mathbf{m}_1] = [0,1] \times [2,10]$ et $\mathbf{m}_2 \in [\mathbf{m}_2] = [10,16] \times [0,1]$. La zone broutée par la chèvre i est incertaine, du fait de l'incertitude sur la position des pieux. Elle est représentée par l'ensemble épais $[\![\mathbb{A}(i)]\!] = [\![\mathbb{A}_*(i), \mathbb{A}^*(i)]\!]$ avec $\mathbb{A}_*(i) = [\mathbf{m}_i] \boxplus \mathbb{D}$, $\mathbb{A}^*(i) = [\mathbf{m}_i] \oplus \mathbb{D}$, où $\mathbb{D}$ est le disque de centre 0 et de rayon 10. La zone broutée par au moins une des chèvres est un ensemble $\mathbb{A}$ qui appartient à l'ensemble épais $[\![\mathbb{A}]\!] = [\![\mathbb{A}(1)]\!] \cup [\![\mathbb{A}(2)]\!] = [\![\mathbb{A}_*(1) \cup \mathbb{A}_*(2), \mathbb{A}^*(1) \cup \mathbb{A}^*(2)]\!]$. L'ensemble $\mathbb{A}$ est un ensemble ontique, alors que les pavés $[\mathbf{m}_1]$,$[\mathbf{m}_2]$ sont épistémiques (en noir sur la figure). On est certain qu'aucune des chèvres n'atteindra la zone en gris foncé. Notons que les zones broutées possibles ne sont pas *tous* les sous-ensembles entre $\mathbb{A}_*(1) \cup \mathbb{A}_*(2)$ et $\mathbb{A}^*(1) \cup \mathbb{A}^*(2)$. L'ensemble épais est une approximation englobante des zones broutées effectivement possibles.

6 Conclusion

Cette note suggère des liens entre des travaux d'inspirations différentes. Les ensembles épais sont des paires emboitées de sous-ensembles classiques. Le cadre de la théorie des possibilités devrait permettre d'étendre leur calcul au cas d'ensembles épais flous, permettant ainsi d'introduire de la gradualité dans l'incertitude. Les travaux de Denœux et col. [2] peuvent être aussi vus comme une extension des ensembles épais aux fonctions de croyance.

Références

[1] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Statist.*, 38, 325-339,1967.

[2] T. Denœux, Z. Younes and F. Abdallah. Representing uncertainty on set-valued variables using belief functions. Artificial Intelligence, Vol. 174, Issues 7-8, pages 479-499, 2010.

[3] B. Desrochers, L. Jaulin, Computing a guaranteed approximation of the zone explored by a robot. *IEEE Trans. Automat. Contr.* 62(1) : 425-430, 2017.

[4] B. Desrochers, L. Jaulin, Thick set inversion. *Artificial Intelligence* 249 : 1-18, 2017.

[5] D. Dubois, H. Prade, Fuzzy set-theoretic differences and inclusions and their use in the analysis of fuzzy equations. *Control and Cybernetics* (Warsaw), 13, 129-146, 1984.

[6] D. Dubois, H. Prade. Evidence measures based on fuzzy information, *Automatica*, 21, 547-562, 1985.

[7] D. Dubois, H. Prade, Twofold fuzzy sets and rough sets - Some issues in knowledge representation. *Fuzzy Sets and Syst.*, 23 (1), 3-18, 1987.

[8] D. Dubois, H. Prade, *Théorie des Possibilités.* Masson, 1987.

[9] D. Dubois, H. Prade, On incomplete conjunctive information, *Computers and Mathematics with Applications*, 15(10), 797-810, 1988.

[10] D. Dubois, H. Prade, Gradualness, uncertainty and bipolarity : Making sense of fuzzy sets. *Fuzzy Sets and Systems*, 192, 3-24, 2012.

[11] L. Jaulin, Solving set-valued constraint satisfaction problems. *Computing*, 94 (2-4), 297-311, 2012.

[12] H. Farreny, H. Prade, Tackling uncertainty and imprecision in robotics. Proc. 3rd Inter. Symp. of Robotics Research, Gouvieux (Chantilly), (O. Faugeras, G. Giralt, eds.), M.I.T. Press, 85-91, Oct. 1985.

[13] H. Farreny, H. Prade, Uncertainty handling and fuzzy logic control in navigation problems. Proc. Inter. Conf. on Intelligent Autonomous Systems, Amsterdam, Dec. 8-11, 1986, 218-225.

[14] Y. Gentilhomme. Les ensembles flous en linguistique. *Cahiers de Linguistique Théorique et Appliquée* (Bucarest), 5, 47-63, 1968.

[15] L. Jaulin, M. Kieffer, O. Didrit, E. Walter, *Applied Interval Analysis : With Examples in Parameter and State Estimation, Robust Control and Robotics.* Springer, 2001.

[16] R. Moore, *Interval Analysis.* Prentice-Hall, 1966.

[17] Z. Pawlak. *Rough Sets. Theoretical Aspects of Reasoning about Data.* Kluwer Acad. Publ., Dordrecht, 1991.

[18] Y. Y. Yao, Interval sets and interval-set algebras. *Proc. of the 8th IEEE Int. Conf. on Cognitive Informatics* (ICCI'09), (G. Baciu et col., eds.), Hong Kong, June 15-17, 307-314, 2009.