



**HAL**  
open science

# Un aperçu de la théorie des ensembles flous et quelques applications (III)

Didier Dubois

► **To cite this version:**

Didier Dubois. Un aperçu de la théorie des ensembles flous et quelques applications (III). Bulletin pour les sous-ensembles flous et leurs applications, 1980, 5, pp.69–75. hal-04564609

**HAL Id: hal-04564609**

**<https://hal.science/hal-04564609>**

Submitted on 30 Apr 2024

**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.

# UN APERÇU DE LA THÉORIE DES SOUS-ENSEMBLES FLOUS ET QUELQUES APPLICATIONS (III)

*Didier DUBOIS*

CERT / DERA

## QUELQUES APPLICATIONS

La théorie des sous-ensembles flous peut s'avérer très utile pour la modélisation des systèmes à composantes humaines et des processus trop complexes pour qu'une représentation détaillée soit viable ou même ait un sens. Comme l'a signalé Zadeh, à mesure que la complexité d'un système s'accroît, les informations précises relatives à ce système ont de moins en moins de sens. On cite ici quelques applications spécifiques, à la classification, la modélisation du langage naturel, la commande de processus, les modèles de diagnostic, les modèles décisionnels, la recherche opérationnelle.

### 1 - CLASSIFICATION

Un algorithme classique de Taxonomie est ISODATA (Ball and Hall 1967) qui cherche à partitionner un ensemble de données  $X = \{x_1 \dots x_n\}$  en  $m$  sous-ensembles  $\{F_i \mid i = 1, m\}$  et à déterminer les centres  $v_i$  de chaque sous-ensemble, de manière à ce qu'on puisse approximer  $X$  par  $\{v_i \mid i=1, m\}$ .  $X$  est supposé inclus dans  $\mathbb{R}^n$ , muni de la distance euclidienne  $d$ . On veut que les groupes  $F_i$  soient compacts et bien séparés, ce qui signifie intuitivement que deux éléments d'un groupe sont toujours plus proches que deux éléments de groupes différents. En pratique, on cherche les  $F_i$  et les  $v_i$  qui minimisent

$$J(F,v) = \sum_{i=1}^m \sum_{x \in F_i} d(x, v_i)^2 \quad (F = \{F_i\}_i, v = \{v_i\}_i)$$

$J(F,v)$  est l'erreur quadratique que l'on commet en assimilant les  $F_i$  aux  $v_i$ .  
L'algorithme ISODATA est itératif et peut se schématiser ainsi

- 1 choisir une partition  $F^0 = \{F_1, \dots, F_m\}$
- 2 calculer les centres des  $F_i$ , soit  $v_i$
- 3 construire une nouvelle partition  $F^1$  en affectant chaque  $x_j$  au groupe  $F_i$  tel que  $d(x_j, v_i)$  est minimum
- 4 si on obtient la même partition qu'auparavant arrêt. Sinon retour en 2.

On obtient ainsi toujours des  $F_i$  et des  $v_i$ , même quand des groupes compacts et bien séparés n'existent pas. De plus l'algorithme ISODATA ne calcule qu'un minimum local de  $J$ . Ce qu'on sait, c'est que si les groupes bien séparés existent alors ils sont un point fixe de l'algorithme.

Pour pallier à cet inconvénient dû au peu de souplesse de la représentation, on peut remplacer la notion de partition par celle de partition floue :

$F_1, \dots, F_m$  est une partition floue de  $X$  si et seulement si  $\sum_{i=1}^m \mu_{F_i}(x) = 1 \quad \forall x \in X$

où les  $F_i$  sont des sous-ensembles flous. Un exemple simple de partition floue est formé d'un sous-ensemble flou  $A$  et son complémentaire  $\bar{A}$ .

On cherche donc une partition floue et des centres qui minimisent

$$J(F,v) = \sum_{i=1}^m \sum_{x \in X} [\mu_{F_i}(x)]^w d(x, v_i)^2$$

L'algorithme ISODATA a été adopté par Bezdek (1974) qui a étudié la fiabilité de la solution. Il a notamment montré que pour  $w = 2$ , le résultat de l'algorithme reflète généralement le flou de la meilleure partition possible, c'est-à-dire indique si, en définitive, il est licite ou non de partitionner le jeu initial de données en  $m$  groupes. Pour plus de détails, voir Dunn (1974), Bezdek (1974).

D'autres approches floues en taxonomie existent, notamment

Ruspini (1969), (1973), Gitman et Levine (1970), etc... Une bibliographie plus complète se trouve dans Dubois-Prade /6/.

N.B. Le terme "classification" renvoie parfois aussi aux problèmes de documentation automatisée. L'application du flou à la définition d'un Thesaurus est étudiée par Reisinger (1974), Radecki (1976). L'interrogation d'une base de données documentaires est abordée par Tahani (1977) et Yager (1980). Mais le premier à utiliser le flou en ce domaine reste Negoita (1973).

## 2 - MODELISATION DU LANGAGE NATUREL

La plupart des tentatives pour représenter le langage naturel sont structurelles, c'est-à-dire qu'elles tiennent compte de la syntaxe des phrases. Zadeh (1971) a développé une approche sémantique, c'est-à-dire qui cherche à représenter le sens des mots. Le sens d'un mot est vu par Zadeh comme l'ensemble (flou) des objets auquel il s'applique. Ainsi "grand" est un sous-ensemble flou de tailles.

Le sens d'une phrase telle que "Jean est grand" est une procédure de calcul qui fournit l'information contenue dans la phrase sous forme d'une distribution de possibilité. "Jean est grand" se traduit en une distribution de possibilité sur la variable "taille de Jean". A chaque valeur de taille on associe un nombre entre 0 et 1 qui traduit la possibilité pour que la taille de Jean ait effectivement cette valeur. Zadeh a entrepris de fonder une théorie du raisonnement approché où on puisse tirer des conclusions à partir de prémisses vagues et avec des règles d'inférence approximatives. (Zadeh 1978). On distingue 4 types de phrases correspondant à quatre règles de traductions :

- les règles de modification s'appliquent à des phrases telles que

"Jean est très grand"

- les règles de composition s'appliquent à des phrases telles que

" Si le volume est très petit, la pression sera très forte".

- les règles de quantification, à des phrases telles que

"La plupart des américains sont grands".

- les règles de qualification

- . en vérité : "il est vrai que la température soit très fraîche"
- . en probabilité : "il est probable qu'il fera froid demain"
- . en possibilité : "il est possible qu'il gèle fort cette nuit"

La théorie du raisonnement approché permet de calculer la distribution de possibilité associée à chacune de ces phrases, connaissant les sous-ensembles flous qu'elles contiennent

\* Par exemple "Jean est très grand" est formellement du type  $M$  est  $mA$  où  $A$  est un sous-ensemble flou sur un référentiel  $X$  (ici les tailles).  $m$  est un adverbe de modification,  $M$  est un objet donné.  $mA$  est un sous-ensemble flou construit à partir de  $A$  par une formule du type

$$\mu_{mA}(x) = f(\mu_A(x)) \quad (\text{pour } m = \text{"très"}, \mu_{\text{très}A}(x) = \mu_A(x)^2)$$

"Jean est très grand" se traduit par

$$\pi_{\text{Taille}(\text{nom=Jean})} = \text{"très grand"} = \mu_A^2$$

\* Dans "la plupart des américains sont grands", il y a deux sous-ensembles flous :  $A = \text{"grand"}$  sur l'univers des tailles, et "La plupart" qui est une proportion floue, sous-ensemble flou  $\hat{p}$  de  $0,1$ . La proportion des grands américains dépend du cardinal du sous-ensemble flou "grands américains" c'est-à-dire :

$$\text{Prop (Grand)} = \frac{\sum_{u \in U} \mu_A(\text{taille}(u))}{|U|}$$

cù  $U$  est la population américaine. La phrase se traduit donc par

$$\pi_{\text{prop(Grand)}} = \text{la plupart} = \mu_{\hat{p}}$$

Zadeh a formulé cette approche en donnant la spécification d'un langage de représentation du langage naturel : PRUF. On conçoit tout l'intérêt d'un tel langage dans une optique d'intelligence artificielle (Zadeh 1977)

### 3 - COMMANDE LINGUISTIQUE DE PROCESSUS

La théorie du raisonnement approché permet de quantifier des stratégies de commande de processus exprimées en langage naturel. Lorsque le processus est trop complexe pour être modélisé, mais qu'un opérateur humain parvient malgré tout, grâce à son expérience, à le commander, il est possible d'implémenter un organe flou de commande qui soit la traduction en règles logiques floues de la stratégie s'exprime souvent sous la forme de propositions conditionnelles du genre:

" Si la température est élevée, et que sa variation est positive forte, alors diminuer beaucoup la pression. "

C'est-à-dire formellement.

Si X est A et Y est B alors Z est C

où X, Y sont des variables observées et Z la variable de commande.

La traduction usuelle d'une telle règle est  $\pi(x,y,z) = \mu_{A \wedge B \rightarrow C}(x,y,z)$  soit  $\pi(x,y,z) = \min(\mu_A(x), \mu_B(y), \mu_C(z))$

Si plusieurs règles sont disponibles on convient de prendre l'union de ces règles, c'est-à-dire l'union des relations floues. Soit R la relation floue résultante ; si on a des observations floues  $\tilde{x}$  et  $\tilde{y}$ , la composition sup-min permet d'inférer une commande floue  $\tilde{z}$  :

$$\mu_{\tilde{z}}(z) = \sup_{x,y} \min(\mu_{\tilde{x}}(x), \mu_{\tilde{y}}(y), \mu_R(x,y,z))$$

En fait, il arrive souvent que les observations soient des valeurs non floues  $\bar{x}, \bar{y}$  et on a

$$\mu_{\tilde{z}}(z) = \mu_R(\bar{x}, \bar{y}, z)$$

Pour choisir une commande non floue qui soit compatible avec la sortie du contrôleur on convient de choisir  $\bar{z}$  tel que  $\mu_{\tilde{z}}(\bar{z}) = \sup \mu_{\tilde{z}}(z)$ . Si plusieurs valeurs sont possibles, on peut choisir la moyenne arithmétique de ces valeurs.

Cette approche de la commande des processus complexes a fait l'objet de nombreuses expérimentations sur des cas d'école (machine à vapeur, Mamdani et Assilian 1975) mais aussi des cas réels (régulation de température, Kickert et Van Nanta Lanke 1976 ; système d'échange de chaleur, Østergaard 1977).

Pour une bibliographie exhaustive, voir Tong (1977) ou Dubois et Prade /6/.

De nouvelles techniques destinées à améliorer la règle sup-min décrite ci-dessus voient actuellement le jour (e.g., Baldwin et Guild 1980)

(à suivre)

### Bibliographie choisie

#### Ouvrages généraux et recueils d'articles

1. KAUFMANN A. Introduction à la théorie des sous-ensembles flous. Tomes I à IV, Masson, Paris 1973-78
2. NEGOITA C.V., RALESCU D. : Applications of fuzzy sets to systems analysis, ISR II, Birkhäuser Verlag 1975
3. ZADEH L.A., FU K.S., TANAKA K., SHIMURA M. (eds) : Fuzzy sets and their applications to cognitive and decision processes, Academic Press 1975.
4. GUPTA M.M., SARIDIS G.N., GAINES B.R., (eds) : Fuzzy automata and decision processes, North Holland 1977
5. GUPTA M.M., RAGADE R.K., YAGER R.R. : Advances in fuzzy set theory and applications, North Holland 1979.
6. DUBOIS D., PRADE H. : Fuzzy sets and systems : theory and applications Academic Press 1980
7. WANG P.P., CHANG S.K. : Fuzzy sets theory and application to Policy analysis and Information Systems, Plenum Press 1980.

#### Classification

- BALL G., HALL D. : A clustering technique for summarizing multivariate data Behavioral Sci. 12, 153-155, 1967
- DUNN J.S. : Well separated clusters and optimal fuzzy partitions - Jr. Cybernetics. Vol. 3 n°3 pp 32-57 - 1974
- BEZDEK J.C. : Clustervalidity with fuzzy sets - J. Cybernetics Vol.3 n°3 pp 58-73, 1974
- RUSPINI E.H. : A new approach to clustering - Information and Control - Vol. 15 - pp 22-32 , 1969
- RUSPINI E.H. : New experimental Results in fuzzy clustering - Inf. Sci. Vol. 6 pp. 273-284, 1973
- GITMAN I., LEVINE M.D. : A algorithm for detecting unimodal fuzzy sets and its application as a clustering technique - IEEE Trans. Computers 19, 583-593 - 1970
- NEGOITA C.V. : On the application of the fuzzy sets separation theorem for automatic classification in information retrieval systems. Inf. Sci. Vol.5 pp.279-286 - 1973
- NEGOITA C.V., FLONDOR P. : On fuzziness in Information Retrieval - Int. J. on Man-Machine Studies - Vol.8 - pp. 711-716 - 1976.

- REISINGER L. : On fuzzy thesauri - Compstat 1974 - Proc. Symp. Comput. Stat. 119-127 Physica-Verlag - Vienna 1974
- RADECKI T. : Mathematical model of information retrieval system based on the concept of fuzzy the thesaurus - Inf.Process. & Manag. 12, 313-318 - 1976.
- TAHANI V. : A conceptual frame work for fuzzy query processing - A step towards very intelligent data systems - Inf. Process. et Manag. 13, 289-303 - 1977
- YAGER R.R. : A logical on-line bibliographic searcher : An application of fuzzy sets IEEE Trans. Systems Man et Cybernetics 10 n°1, 51-53 - 1980

### Langage naturel

- Zadeh L.A. : Quantitative fuzzy semantics - Inf. Science 3, 159-176 - 1971
- ZADEH L.A. : A theory of approximate reasoning - in Machine Intelligence 19 D. Michie Cd. American Elsevier NY 1978
- ZADEH L.A. : Pruf, A meaning representation language for natural language Memo. UCB/ERLM77-61 University of California Berkeley 1977  
Int. J. Man-Machine Stud. 10, 395-460, 1978

### Commande de processus

- MAMDANI E.H., ASSILIAN S. : An experiment in linguistic synthesis with a fuzzy logic controller - Int. J. on Man-Machine Studies 7, 1-13- 1975
- KICKERT W.J.M., VAN NAUTA LEMKE H.R. : Application of a fuzzy controller in a warm water plant - Automatica 12, 301-308, 1976
- ØSTERGAARD J.J. : Fuzzy logic control of a heat exchanger process, in /4/ pp 285-320
- TONG R.M. : A control engineering review of fuzzy systems - Automatica 13, 559-569 - 1977
- BALDWIN J.F., GUILD N.C.F. : Modelling controllers using fuzzy relations, Kybernetes 9, 223-229 - 1980.