

Revue Internationale de

ISSN 0980-1472

systemique

Vol. 11, N° 1, 1997

afcet

DUNOD

AFSCET

Revue Internationale de
systemique

Revue
Internationale
de Sytémique

volume 11, numéro 1, pages 109 - 127, 1997

Acquisition de connaissances imparfaites :
mise en évidence d'une fonction d'appartenance

Nathalie Aladenise et Bernadette Bouchon-Meunier

Numérisation Afscet, mars 2016.



Creative Commons

**ACQUISITION DE CONNAISSANCES IMPARFAITES :
MISE EN ÉVIDENCE D'UNE FONCTION D'APPARTENANCE**

Nathalie ALADENISE et Bernadette BOUCHON-MEUNIER ¹

Résumé

L'acquisition et la modélisation des connaissances imparfaites sont des étapes importantes dans l'élaboration d'un système. La théorie des sous-ensembles flous offre une bonne représentation de ces connaissances. Cependant, la construction des fonctions d'appartenance demeure un problème mal résolu. Cet article propose quatre grandes catégories de méthodes de construction de ces fonctions : les méthodes automatiques qui permettent de se passer d'expert, les méthodes probabilistes et les méthodes statistiques où l'avis du plus grand nombre fait foi, les méthodes psychométriques pour interroger l'expert, et une méthode de construction d'une fonction par interpolation à partir de quelques points.

Abstract

The acquisition and the representation of imperfect knowledge are important steps in the construction of a system. Fuzzy set theory provides interesting tools for this purpose. Nevertheless, it is generally a difficult problem to determine the necessary membership functions. This paper presents four categories of methods to obtain these functions: automatic methods which can be used in the case where no expert is available, probabilistic and statistical methods based on the opinion of a large number of individuals, psychometrical methods to elicit the functions from experts, and an interpolative method to construct a function from a small number of data.

1. LAFORIA, Université Pierre-et-Marie-Curie, case 169, 4, place Jussieu, 75252 Paris Cedex 05. e-mail: aladenis@laforia.ibp.fr. e-mail: bouchon@laforia.ibp.fr.

I. INTRODUCTION

L'efficacité d'un système dépend en grande partie de la qualité des connaissances qui le composent. C'est pourquoi la phase d'acquisition de ces connaissances est primordiale lors de sa conception.

En particulier, les connaissances vagues, les concepts qualitatifs ou quantitatifs non précisément limités, sont particulièrement difficiles à manipuler. Une méthode de modélisation possible pour ces connaissances imprécises consiste à utiliser la théorie des sous-ensembles flous, introduite par Zadeh en 1965 ([Za 65], [Bo 94], [Bo 95], [Co 95]).

Or il n'existe pas de méthodologie standard pour fabriquer les fonctions d'appartenance, brique de base de cette théorie. La difficulté se trouve dans la représentation des termes linguistiques par un modèle numérique. Il faut, pour y parvenir, employer des méthodes adéquates pour interroger les experts détenteurs de ces connaissances. Mais s'il est facile de comprendre « très grand », il est difficile de numériser « très » ou « grand ».

Cet article présente différentes méthodes de mise en évidence de ces fonctions. Le paragraphe 2 effectue un bref rappel sur la théorie des sous-ensembles flous. Le paragraphe 3 présente les techniques automatiques de détermination de fonctions d'appartenance les plus utilisées. Les paragraphes 4 et 5 exposent les différentes méthodes statistiques et probabilistes, ainsi que les méthodes psychométriques. Enfin, le paragraphe 6 décrit une méthode de construction analytique de fonction d'appartenance, par interpolation entre plusieurs points connus.

A l'issue de cet exposé, seront soulevés les problèmes de choix de la forme de la fonction d'appartenance, d'agrégation de connaissances et de dépendance des concepts vagues par rapport au contexte.

II. UTILISATION DE LA THÉORIE DES SOUS-ENSEMBLES FLOUS POUR MODÉLISER LES TERMES VAGUES

L'une des fonctions de la théorie des sous-ensembles flous est de modéliser les termes linguistiques vagues ou imprécis, en particulier les qualificatifs exprimant l'appartenance graduelle à un ensemble, comme par exemple les adjectifs « petit », « grand » ([Za 65], [Bo 94], [Bo 95], [Za 75], [Co 95]). Chacun de ces termes est défini sur un ensemble de référence X : l'ensemble des tailles pour « petit » et « grand », l'ensemble des prix pour « cher ».

Si A est une catégorie imprécise de l'ensemble de référence X , il existe des éléments de X dont l'appartenance à A est partielle. Une façon de modéliser une telle catégorie est de la décrire comme une classe avec un continuum de degrés d'appartenance, c'est-à-dire comme un sous-ensemble flou de X dont la fonction d'appartenance serait $\mu_A(x)$, $x \in X$. A chaque terme linguistique, on fait ainsi correspondre une fonction d'appartenance μ_{petit} , μ_{grand} , μ_{cher} ...

Pour tout sous-ensemble flou A de X , défini par sa fonction d'appartenance $\mu_A : X \rightarrow [0, 1]$, $\mu_A(x)$ représente la valeur de vérité du fait « $x \in A$ ».

- Si « $x \in A$ » est totalement faux, $\mu_A(x) = 0$.
- Si « $x \in A$ » est vrai sans aucun doute possible, alors $\mu_A(x) = 1$.
- Pour tous les éléments x tels que « $x \in A$ » soit *plus ou moins vrai*, $0 < \mu_A(x) < 1$.

On définit : le *noyau* de A : $\text{noy}(A) = \{x | \mu_A(x) = 1\}$, et le *support* de A : $\text{supp}(A) = \{x | \mu_A(x) \neq 0\}$.

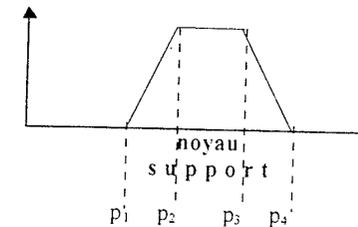


Figure 1. Exemple de fonction d'appartenance.

Lorsqu'on utilise la théorie des sous-ensembles flous pour formaliser un problème, on manipule autant de fonctions d'appartenance qu'il y a de caractérisations de concepts. Mais, ces fonctions d'appartenance, d'où proviennent-elles ? Comment les détermine-t-on ?

Dans la littérature, on se réfère souvent à l'utilisation de fonctions d'appartenance sans plus d'explications quant à leur obtention. On se contente là de contourner la question de la représentation des connaissances de l'expert. On peut supposer qu'il s'agit d'une explicitation des connaissances totalement empirique, dépendante de l'individu et du contexte qui entoure cet individu.

Pour répondre plus formellement à cette question, nous allons, dans les paragraphes suivants, présenter différentes méthodes de construction d'une fonction d'appartenance. On peut les classer en quatre catégories :

les méthodes automatiques, les méthodes statistiques et probabilistes, les méthodes psychométriques et les méthodes géométriques de construction.

III. APERÇU DES MÉTHODES AUTOMATIQUES

Les méthodes les plus usitées, à l'heure actuelle, pour déterminer les fonctions d'appartenance sont celles à base de réseaux de neurones ([Ma 92], [Me 93]). D'autres méthodes automatiques, utilisant des algorithmes génétiques, commencent aussi à donner des résultats intéressants ([Bä 94], [He 94], [Go 89]). Comme la logique floue, ces méthodes manipulent des informations numériques, ce qui facilite leur synergie. Elles ont, de plus, l'avantage de pouvoir se passer d'expert; la phase interactive de recueil de données n'est pas nécessaire ici : un ensemble de données d'apprentissage composé d'exemples significatifs est suffisant. Or, cet ensemble peut être, par exemple, le résultat d'une expérience ou d'un sondage.

L'obtention des fonctions d'appartenance se déroule en deux phases : création d'une fonction d'appartenance primaire mal ajustée (voire, dans certains cas, aléatoire), et ajustement de cette fonction de base par optimisation. Cette étape s'appellera *apprentissage supervisé avec rétro-propagation* dans le cas des réseaux de neurones, et application d'*opérations génétiques* jusqu'à *émergence* d'une fonction optimale, dans le cas des algorithmes génétiques.

Les réseaux neuronaux

Le connexionnisme, en raison de ses capacités à manipuler des données numériques, permet bien des mariages fructueux avec la théorie des sous-ensembles flous. Entre autres, la capacité d'apprentissage des réseaux de neurones apporte une aide appropriée pour la détermination des fonctions d'appartenance. L'association des techniques neuronales et floues est baptisée le « neuro-flou ».

Le principe des réseaux neuronaux est simple : on possède au départ, une base d'exemples composée de couples (données d'entrée, résultat désiré) qui va servir à l'apprentissage du réseau. Cet apprentissage consiste à ajuster les paramètres du réseau au fur et à mesure qu'il prend connaissance des données initiales. La particularité de l'*apprentissage supervisé* est que, de plus, on mesure l'erreur du résultat produit par le réseau, par rapport au résultat

attendu, ce qui permet un ajustement supplémentaire des *poids* (paramètres) de manière à diminuer cette erreur.

On utilise donc les propriétés d'apprentissage et de généralisation des réseaux de neurones pour la détermination des fonctions d'appartenance. Les données d'apprentissage sont expérimentales et ponctuelles; c'est le réseau qui interpole les résultats pour fabriquer la fonction complète.

La notion d'approximation due aux réseaux neuronaux n'est pas gênante dans le contexte de la logique floue, puisque cette dernière manipule des données vagues. En mettant en évidence une fonction d'appartenance, on cherche à se rapprocher le plus possible de la représentation mentale du terme linguistique. De plus, si deux experts différents fournissent des données d'apprentissage contradictoires, ceci ne pose pas de problème, car leurs deux avis seront pris en compte à égale importance, pour l'ajustement des poids du réseau, lors de l'apprentissage. La recherche de consensus est automatique.

Les algorithmes génétiques

En génétique, une population initiale d'individus, décrits par des chromosomes, est soumise à des transformations telles que des mutations ou des croisements, les individus résultants étant les plus adaptés à la survie. Par analogie, dans la méthode des algorithmes génétiques, on se donne une population de vecteurs ou d'éléments d'information appelés *chromosomes* qui correspondent à une solution au problème posé (ici, la recherche d'une fonction d'appartenance). Chaque chromosome représente donc, sous forme codée, les paramètres permettant de définir une fonction d'appartenance. Cette population de départ est souvent aléatoire. On définit ensuite des transformations (*sélection* par une mesure d'adéquation (« *fitness* »), *mutation*, et *recombinaison*) qui s'appliquent à la population de vecteurs. Les applications successives des trois opérations génétiques vont discriminer ou promouvoir certains chromosomes. La *mutation* de gènes favorise l'apparition de nouvelles propriétés; l'élimination d'éléments les moins pertinents est effectuée par la *sélection*; la prédominance de certaines caractéristiques par rapport à d'autres est mise en valeur par *recombinaisons* successives. Le résultat de l'application de ces opérations est la mise en évidence de la meilleure solution possible au problème posé, donc de la fonction d'appartenance la plus pertinente.

De la même famille que les algorithmes génétiques, les stratégies d'évolution ([Mi 92], [Bä 94]) ont un fonctionnement analogue, à la différence près que le chromosome des algorithmes génétiques contient ses

informations sous forme binaire, alors que le vecteur d'information des stratégies d'évolution contient des réels ([Bä 94], [He 94]).

Le principal défaut de ces deux méthodes est que les fonctions de calcul (calcul des poids dans les réseaux de neurones, ou calcul des opérations génétiques) ne sont ni intuitives ni explicites vis-à-vis de l'utilisateur. Par contre, elles ont l'avantage d'être automatiques, c'est-à-dire que ce n'est pas l'expert qui fabrique la fonction d'appartenance, mais un algorithme, et également l'avantage de fonctionner à partir d'une base d'exemples; ces exemples ne sont pas nécessairement fournis par des experts. Elles ont aussi l'intérêt de manipuler directement des nombres, alors que tout le problème de la fabrication des fonctions d'appartenance est la numérisation de concepts linguistiques. Ces facilités expliquent l'essor de ces méthodes pour la construction de systèmes à base de commande floue ces dernières années.

La classification

La classification n'est pas, à proprement parler, une méthode automatique au même sens que les deux précédentes, car elle ne donne pas directement des fonctions d'appartenance. Cependant, les fonctions d'appartenance sont souvent induites par un processus de classification. La fonction d'appartenance représente ici le taux d'adéquation de chaque élément à la classe à laquelle il a été attaché.

Par exemple, on pourra associer à chaque taille de X la proportion d'individus, mesurant exactement cette taille, classés dans l'ensemble « grand ».

La fonction d'appartenance d'un élément peut encore être fonction de la distance qui sépare chaque élément du prototype de sa classe; le prototype aura alors un degré d'appartenance égal à 1.

IV. MÉTHODES STATISTIQUES ET PROBABILISTES

Les deux premières méthodes présentées ci-dessous ont un fondement statistique et partent du principe que, plus une représentation est approuvée par un grand nombre d'individus, plus elle est satisfaisante. La troisième utilise des observations de situations et non plus des interrogations d'individus, avec la même approche fréquentielle. Les deux dernières sont basées sur l'utilisation de densités de probabilité.

Méthode statistique « Oui-Non »

Un principe simple pour obtenir le degré d'appartenance d'un élément d'un ensemble de référence à un terme linguistique (modélisé par un ensemble flou) est la méthode appelée « Oui-Non », proposée initialement par Black [BI 37] et utilisée, entre autres, par Hisdal et Gaines ([Hi 88], [Ga 78], [Gi 88], [Zh 93]). Elle consiste à interroger chaque membre d'une population avec une question qui n'autorise qu'une réponse binaire : oui ou non. Par exemple, lorsque l'ensemble de référence est celui des tailles, pour savoir si un homme de 1,72 m est « grand », la question pourra être formulée ainsi : « Est-il vrai qu'un homme d'1,72 m soit grand ? ». Le degré d'appartenance de « 1,72 m » à l'ensemble flou « grand » sera égal à la proportion de réponses « Oui » dans l'ensemble des réponses.

Méthode statistique d'estimation d'ensemble

Dans cette méthode, proposée par Wang [Wa 83], on demande aux sujets de décrire leur acception de A par un sous-ensemble de X . Par définition, la fonction d'appartenance de l'élément x à A , $\mu_A(x)$, prendra comme valeur la fréquence avec laquelle la description de A donnée par les sujets p_i interrogés (notée $A(p_i)$) contenait x .

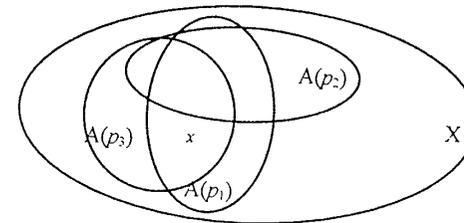


Figure 2. Méthode statistique d'estimation d'ensemble.

Sur cet exemple, $\mu_A(x) = 2/3$. Deux des trois individus interrogés pensent que x possède la caractéristique A .

Histogramme de fréquences normalisé

Cette méthode, présentée par Chauvin [Ch 95], utilise la fréquence d'apparition comme critère déterminant l'appartenance à une classe. Plus un élément est rencontré souvent dans une classe, plus il est représentatif de la classe et donc plus sa fonction d'appartenance à cette classe est grande.

En pratique, les éléments représentatifs de la caractéristique A sont rassemblés dans un sous-ensemble que, pour simplifier, l'on note aussi A . On réalise un histogramme des fréquences des éléments du sous-ensemble A sur le domaine de référence de la caractéristique A , et on le normalise par le maximum des fréquences :

$$\forall x \in X \quad \mu_A(x) = \text{fréq}_A(x) / \max_{x \in X} \text{fréq}_A(x).$$

Puis on lisse cet histogramme (par un filtrage gaussien, par exemple). Cet histogramme normalisé lissé représente notre fonction d'appartenance.

Cette même méthode ne peut pas être appliquée à tous les types de fonctions d'appartenance. En effet, la fonction représentant les personnes de grande taille issue de cette méthode indiquera que l'appartenance d'un individu de 2,10 m à l'ensemble des « grands » est petite, car il existe moins de personnes mesurant 2,10 m que de personnes mesurant 1,80 m. Or intuitivement, cet individu de 2,10 m devrait justement appartenir complètement au sous-ensemble flou des « grands ».

On peut se demander si, avec chacune des ces trois méthodes, on a bien modélisé une fonction d'appartenance par une mesure d'imprécision, ou bien si l'on a prouvé la variabilité des réponses à une question sur l'interprétation des mots du langage.

Méthode d'Ellen Hisdal

Cette méthode est proposée par Hisdal [Hi 88]. La première étape consiste à demander à l'expert, comme dans la méthode Oui-Non, de répondre par « Oui » ou « Non » à la question « Cet élément x de X appartient-il à A (sous-ensemble de X) ? ». On lui soumet différents éléments de X jusqu'à ce que l'on puisse obtenir une courbe non floue à seuil, qui pourrait être celle indiquée sur la figure 3 pour l'exemple sur l'ensemble X des tailles.

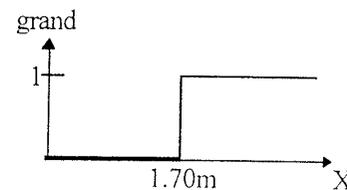


Figure 3. Courbe à seuil.

L'idée de Hisdal est qu'il est probable que l'on se soit trompé sur le choix de l'élément frontière (1,70 m sur l'exemple), mais que plus on s'éloigne de cet élément frontière, plus la probabilité que l'on se soit trompé est faible. Cette probabilité de l'erreur est matérialisée par une courbe d'erreur $E(u)$, qui est une distribution de probabilité gaussienne.

Sur notre exemple des tailles, cette courbe $E(u)$ est telle que l'intégrale de $E(u)$ soit égale à 1 (pour rester fidèle aux lois des probabilités).

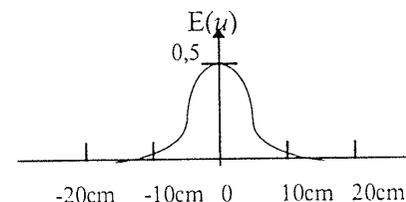


Figure 4. Probabilité d'erreur sur le choix de l'élément frontière dans la courbe à seuil.

Finalement, $\mu_{\text{grand}}(x)$ est la fonction à seuil « arrondie » par la fonction d'erreur; plus précisément, c'est la résultante de la convolution de la courbe à seuil avec la courbe de l'erreur estimée, ce qui donne :

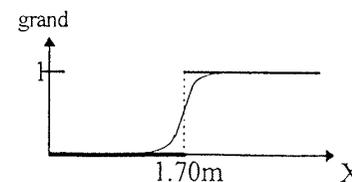


Figure 5. Fonction d'appartenance résultant de la méthode d'E. Hisdal.

Le choix de la forme gaussienne pour la fonction d'erreur n'est pas justifié. Le taux d'erreur pourrait tout aussi bien décroître linéairement au fur et à mesure que l'on s'éloigne de l'élément frontière, ce qui nous donnerait une fonction d'erreur triangulaire, et donc une fonction d'appartenance à pente linéaire.

Fonctions de densité de probabilité

On veut construire une fonction d'appartenance à partir d'une densité de probabilité. L'idée est que, plus la probabilité pour qu'un élément soit caractérisé par A est grande, plus cet élément appartient au sous-ensemble flou A .

Dans cette méthode exposées par Civanlar et Trüssel [Ci 86], la probabilité que les éléments de X soient caractérisés par A est connue. La fonction d'appartenance est donc proportionnelle (avec un coefficient λ à déterminer) à la densité de probabilité, en sachant qu'à partir d'un certain seuil de probabilité, la fonction d'appartenance sera égale à 1 (si la probabilité est suffisamment grande, l'appartenance à A sera totale).

$$\begin{aligned} \mu(x) &= \lambda p(x) & \text{si } \lambda p(x) < 1, \\ \mu(x) &= 1 & \text{si } \lambda p(x) \geq 1. \end{aligned}$$

Le coefficient λ est choisi tel que :

$$\int_{\lambda p(x) < 1} p^2(\eta) d\eta + \int_{\lambda p(x) \geq 1} p(\eta) d\eta - c = 0$$

où $c \in [0, 1]$ est le niveau de confiance dans le fait que les éléments x de X appartiennent à A . Plus le niveau de confiance c est grand, plus le degré d'appartenance de chaque x à A est élevé, et plus les éléments appartenant au noyau de A sont nombreux.

Le paragraphe suivant se concentre sur l'interrogation d'un expert. On préférera la qualité de la réponse (méthodes psychologiques) à la quantité de réponses (méthodes statistiques).

V. MÉTHODES DE PSYCHOMÉTRIE

Ici, ce sont les différentes façons d'interroger directement un expert qui sont étudiées. Il s'agit de lui faire numériser un concept linguistique vague. C'est un exercice de psychométrie.

Méthode élémentaire « Noyau-Support »

La méthode la plus élémentaire consiste à interroger un expert de manière à lui faire définir le *noyau* (éléments de X pour lesquels l'appartenance à A est certaine) et le *support* (X privé des éléments dont la non appartenance à A est certaine) de la fonction d'appartenance de A . Il suffit, pour cela, de faire produire quatre paramètres à l'expert (p_1, p_2, p_3 et p_4). À l'aide de ces données, il est aisé de construire une fonction de maximum 1 atteint sur $[p_2, p_3]$ et de minimum 0 à l'extérieur de $[p_1, p_4]$, la forme (linéaire, à base d'arcs de paraboles...) de cette fonction étant choisie arbitrairement sur l'ensemble $\text{supp}(A) - \text{noy}(A)$, non décroissante entre p_1 et p_2 , non croissante entre p_3 et p_4 (voir figure 1).

Quantification structurelle

Cette méthode est proposée par Zhang [Zh 93]. Il partitionne dans un premier temps l'ensemble X en trois sous-ensembles X_0, X_1 , et X_f . X_1 est l'ensemble des éléments de X appartenant à A avec une totale certitude, $\{x | \mu_A(x) = 1\}$, c'est-à-dire $\text{noy}(A)$; X_0 est l'ensemble des éléments de X n'appartenant pas à A , avec une totale certitude, soit $\{x | \mu_A(x) = 0\}$ ou encore $X - \text{supp}(A)$; et X_f est le sous-ensemble de X contenant tous les éléments n'appartenant ni à X_1 , ni à X_0 , $\{x | 0 < \mu_A(x) < 1\}$, c'est-à-dire l'ensemble : $\text{supp}(A) - \text{noy}(A)$. On a alors :

$$\mu_A(x) = \frac{\text{Card}(\{y \in X_f | x >_A y\})}{\text{Card}(X_f)} = \frac{\text{nombre de } y_A < x \text{ dans } X_f}{\text{nombre d'éléments de } X_f} \quad \forall x \in X_f,$$

où $x >_A y$ signifie que « $x \in A$ » est plus vrai que « $y \in A$ ».

La véracité de « grand » pour l'élément x dépend de la quantité de personnes qui sont moins grandes que x .

L'idée sous-jacente à la quantification structurelle est qu'un sous-ensemble flou est défini par rapport aux autres sous-ensembles; par exemple : « grand » est « de taille supérieure aux tailles moyennes », « bon » est un résultat « supérieur à moyen ». Cette méthode est limitée aux ensembles discrets.

Méthodes de Norwich et Turksen

Pour créer une fonction d'appartenance, Norwich et Turksen [No 84] décèlent la nécessité de l'existence d'un ordre faible sur les éléments de

$X \geq_A$, et d'un ordre faible sur les intervalles entre deux des éléments de $X \geq'_A$.

$x_1 \geq_A x_2 \Leftrightarrow \mu_A(x_1) \geq \mu_A(x_2)$, signifie que « x_1 est au moins autant A que x_2 est A » ou « " x_1 est A " est au moins aussi vrai que " x_2 est A " ».

$|x_2 x_1| \geq'_A |x_4 x_3| \Leftrightarrow \mu_A(x_2) - \mu_A(x_1) \geq \mu_A(x_4) - \mu_A(x_3)$, signifie que « " x_2 est plus A que x_1 " est au moins aussi vrai A que " x_4 est plus A que x_3 " ».

Ils considèrent aussi deux bornes x_m et x_M telles que

$x \geq_A x_m \forall x \in X$ et x_m est un objet dont on est sûr qu'il n'appartient pas à A .

$x_M \geq_A x \forall x \in X$ et x_M est un objet qui appartient de façon certaine à A . x_m et x_M sont choisis arbitrairement parmi tous les x satisfaisant ces définitions. x_m et x_M appartiennent respectivement à X_0 , et X_1 au sens de Zhang [Zh 93].

$$\mu_A(x_m) = 0 \quad \text{et} \quad \mu_A(x_M) = 1.$$

De même que Zhang [Zh 93], Norwich et Turksen [No 84] reconnaissent l'existence d'une partition du domaine de valeurs en trois : X_0 , X_1 , et X_f , où la fonction d'appartenance est monotone par parties sur X_f .

Norwich et Turksen présentent un axe unique au sujet interrogé (voir figure 6) en lui expliquant que x_m est un objet dont on est sûr qu'il n'appartient pas à A , que x_M est un objet qui appartient de façon certaine à A , et que le segment entre ces bornes désigne l'échelle des degrés de vérité pour la phrase « x est A ».



Figure 6. Échelle de vérité dans la méthode de Norwich et Turksen.

On va, ici, poser plusieurs fois la même question à chaque sujet, puis faire une moyenne de ses réponses, plutôt que de poser une fois la question à chacun et de faire la moyenne sur le groupe; ceci permet d'avoir une interprétation possible en fonction de chaque individu. Trois méthodes sont alors envisageables.

1°) *Évaluation directe* : on présente les éléments de X un par un. L'expert (ou sujet) interrogé doit placer un pointeur sur le segment $[x_m, x_M]$ de

façon à ce que la valeur choisie représente le degré de vérité associé au fait « x est grand ».

2°) *Évaluation inverse* : le pointeur est placé aléatoirement à une cote c sur le segment $[x_m, x_M]$. Puis l'ensemble des éléments de X est parcouru dans l'ordre croissant; l'expert doit arrêter cette énumération lorsqu'il estime que l'élément de X atteint correspond à la valeur de vérité représentée par la cote c . On remarque que dans le cas de fonctions d'appartenance trapézoïdales, une même cote c peut avoir 2 éléments de X qui lui sont associés; dans ce cas l'énumération de l'ensemble des éléments de X est poursuivie jusqu'à ce que le sujet l'arrête une seconde fois sur l'élément x_2 de valeur de vérité c .

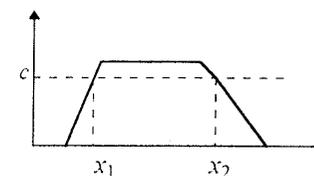


Figure 7. Obtention de la fonction d'appartenance par évaluation inverse.

En pratique, cette évaluation inverse sert de vérification à l'évaluation directe.

3°) *Comparaison d'intervalles* : finalement, on demande aux sujets de comparer les intervalles deux à deux, de façon répétitive et pour chaque paire aléatoire d'éléments de X . Cette démarche apporte une justification supplémentaire aux degrés d'appartenance des éléments de X à A . Elle permet aussi, en cas de doute, de confirmer une valeur litigieuse.

Ces trois méthodes, présentées ensemble, peuvent, bien sûr, être utilisées individuellement. Les expériences pratiquées montrent que les individus produisent des fonctions d'appartenance similaires par évaluation directe ou par évaluation inverse.

Grille répertoire

La méthode de « grille répertoire » ([Bo 85], [Ha 88], [Ha 92], [PI 86]) est une méthodologie d'interrogation d'un expert afin de lui faire préciser, à l'aide de termes linguistiques eux-mêmes vagues, la signification qu'il accorde à d'autres termes vagues faisant partie de l'expertise. Les résultats linguistiques récupérés dans la grille sont ensuite traduits de façon numérique.

Pour chaque terme flou, on demande à l'expert quel est le terme antagoniste ; par exemple pour « grand » l'expert répond « petit ». On crée ainsi une échelle bipolaire que l'on gradue de 1 à 5 de telle sorte que les chiffres aient la représentation suivante pour l'expert :

1 : très grand ; 2 : grand ; 3 : taille moyenne ; 4 : petit ; 5 : très petit.

On peut indifféremment laisser l'expert manipuler les valeurs symboliques (« très grand ») ou directement les chiffres en lui expliquant que plus le chiffre est grand, plus la taille est petite.

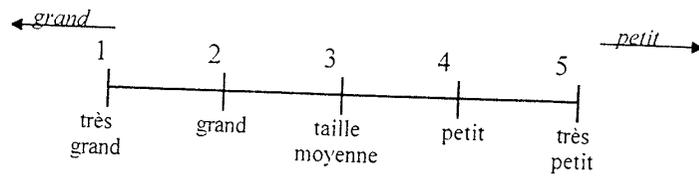


Figure 8. Échelle bipolaire de la grille répertoire.

On peut aussi graduer l'échelle de 1 à 3, ou encore de 1 à 9, etc.

Cette méthode peut être vue comme une façon subtile d'extraire le noyau et le support d'une fonction d'appartenance.

Alpha-sets

Ralescu suggère ([Ra 79]) de déterminer les sous-ensembles de niveau α (ou α -sets), puis de reconstituer la fonction d'appartenance à partir de ces α -sets (au lieu d'essayer de définir directement la fonction d'appartenance $\mu : X \rightarrow [0, 1]$).

Un sous-ensemble de niveau α pour A est un sous-ensemble de X dont les éléments appartiennent, avec un degré au moins égal à α , au sous-ensemble $A : \{x | \mu_A(x) \geq \alpha, x \in X\}$.

On se donne au départ un ensemble fini de degrés α classés par ordre croissant : $\{\alpha_1, \alpha_2, \dots\}$, $\alpha_1 \leq \alpha_2 \leq \dots$. On crée un sous-ensemble X_{α_1} de X qui soit un sous-ensemble de niveau α_1 de X , puis on crée un sous-ensemble X_{α_2} , etc. On obtient ainsi $X_1 \subseteq \dots \subseteq X_{\alpha_2} \subseteq X_{\alpha_1} \subseteq X$, où X_1 est l'ensemble des éléments de X appartenant tout à fait à A .

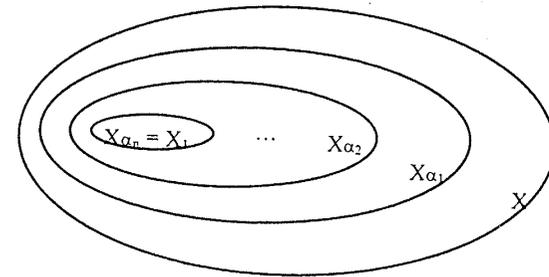


Figure 9. Emboîtement de α -coups.

VI. MÉTHODE DE CONSTRUCTION PAR INTERPOLATION

Pour construire une fonction d'appartenance sur un ensemble de définition X continu, il est nécessaire de connaître un nombre raisonnable de points appartenant à cette fonction. Lorsque le nombre de points connus est petit, il est possible de construire une fonction par interpolation à partir de ces points, vérifiant néanmoins certaines contraintes pour que la courbe obtenue corresponde à une fonction d'appartenance correcte. Ces contraintes sont formalisées par Chen et Otto dans [Ch 95], où ils proposent une méthode de construction des fonctions d'appartenance les vérifiant. En faisant l'hypothèse qu'il existe au moins un point appartenant totalement au sous-ensemble flou et au moins un point n'y appartenant pas du tout, et que l'on est capable de les identifier, voici les propriétés mathématiques auxquelles les fonctions d'appartenance $\mu : R \rightarrow [0, 1]$ répondent :

- $\mu(x) \in [0, 1] \forall x$.
- μ est différentiable.
- $\mu(x_i) = \mu_i$ pour un ensemble fini de couples $\{(x_1, \mu_1), \dots, (x_n, \mu_n)\}$.

Si l'ensemble des couples connus $\{(x_1, \mu_1), \dots, (x_n, \mu_n)\}$ est un ensemble flou convexe, alors μ est convexe.

Cette méthode permet de construire une fonction d'appartenance à partir de quelques points seulement. La fonction construite satisfait l'idée intuitive que l'on se fait d'une fonction d'appartenance : continue, convexe, avec une forme régulière comme si on l'avait dessinée d'un seul coup de crayon, c'est-à-dire continûment dérivable.

VII. QUELQUES REMARQUES GÉNÉRALES

Toutes les méthodes, aussi bien statistiques que psychologiques, présentées, comportent des points communs : on construit généralement des fonctions d'appartenance à noyau non vide ; les données de l'ensemble de référence X sont ordonnées par une structure d'ordre : $x_1 \geq x_2$ signifie que « $x_1 \in A$ » est au moins aussi vrai que « $x_2 \in A$ ».

Agrégation d'informations

Lorsque deux experts différents ne peuvent s'accorder sur une même et unique solution, nous calculons une solution intermédiaire proche de toutes les solutions fournies par tous les experts par des méthodes d'agrégation d'informations, afin de ne pas favoriser l'avis d'un expert par rapport à l'autre, ou au contraire de pondérer les différents avis recueillis suivant leur provenance [PI 86].

L'intérêt d'utiliser une représentation numérique des termes linguistiques est de pouvoir faire un consensus sur les données. Alors qu'il était impossible d'agrèger les différents avis d'experts « moyen » et « grand », il est possible, en revanche, d'agrèger leurs fonctions d'appartenance à l'aide d'opérateurs mathématiques.

L'agrégation de données représente un domaine de recherche dans lequel beaucoup de travaux ont déjà été effectués.

L'agrégation est particulièrement facile lorsqu'on utilise une des méthodes automatiques présentées au paragraphe 3 (réseaux de neurones ou algorithmes génétiques) car, lorsque deux experts différents fournissent des données d'apprentissage contradictoires, leurs deux avis participent à l'apprentissage avec la même importance. Le consensus est automatique.

Choix de la forme d'une fonction d'appartenance

Dans de multiples méthodes de construction de fonctions d'appartenance, on choisit *a priori* la forme des fonctions (forme trapézoïdale, gaussienne, forme de cloche [Bo 95], [Co 94]...) puis on ajuste les paramètres de la fonction (par exemple les bornes du noyau et du support pour les fonctions trapézoïdales. Il est aujourd'hui certain que ce choix *a priori* n'a pas de conséquences majeures sur les résultats des inférences des systèmes experts. Mais dans le cas des systèmes de commande floue, il existe des relations

entre la forme des fonctions d'appartenance et le choix des opérateurs de commande floue.

Dépendance du contexte

Les termes imprécis soulèvent un problème de généralité des données. Chacun de ces termes est relatif à son contexte. Par exemple, les mots « grand » et « vieux » n'ont pas la même signification si c'est un enfant qui les prononce ou si c'est un adulte ; ils n'ont pas non plus la même signification si l'on évalue la taille et l'âge des hommes du XIII^e siècle ou si l'on parle des hommes d'aujourd'hui. D'un côté, ceci pose un problème de réutilisabilité des données par plusieurs systèmes différents (et donc dans des contextes différents) ; d'un autre côté, il suffit de modifier les fonctions d'appartenance pour que le même système fonctionne dans deux contextes différents. Par exemple, un système médical pourra fonctionner en Belgique et en France à condition d'ajuster le vocabulaire vague à l'utilisation des termes médicaux qui est faite en Belgique ou en France.

CONCLUSION

La modélisation de termes linguistiques par des sous-ensembles flous n'est un problème que partiellement résolu. La prolifération des méthodes présentées montre qu'aucune n'est universelle.

Nous nous sommes concentrées, dans cet article, sur des méthodes relativement générales d'obtention des fonctions d'appartenance. Il existe bien d'autres méthodes d'obtention de fonctions d'appartenance propres au système auquel elles appartiennent ou au domaine pour lequel elles ont été conçues. C'est le cas, par exemple des méthodes proposées par Zimmermann et Zysno [Zi 85] dans le cadre de systèmes décisionnels.

Les interrogations liées au choix d'une méthode de mise en évidence des fonctions d'appartenance laissent autant de voies ouvertes à la recherche.

Références

- [Bä 94] T. BACK et F. KURSAWE, Evolutionary algorithms for fuzzy logic: a brief overview, *Proceedings of the fifth international conference IPMU. Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 1994, p. 659-664.

- [Bo 91] O. BOBROWICZ, C. CHOULET, A. HAURAT, F. SANDOZ et M. TEBAA, A method to build membership functions application to numerical/symbolic interface building, *Proceedings of the fifth international conference IPMU. Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 1990, p. 136-142.
- [Bo 85] J. H. BOOSE, A knowledge acquisition program for expert systems based on personal construct psychology, *International Journal of Man-Machine Studies*, 23, 1985, p. 495-525.
- [Bo 94] B. BOUCHON-MEUNIER, *La logique floue*, Que sais-je? n° 2702, Presse Universitaire de France, 2^e édition, 1994.
- [Bo 95] B. BOUCHON-MEUNIER, *La logique floue et ses applications*, Addison-Wesley, 1995.
- [BI 37] M. BLACK, Vagueness, *Philosophy of Science*, 4, 1937, p. 427-455.
- [Ch 95] J. E. CHEN et K. N. OTTO, Constructing membership functions using interpolation and measurement theory, *Fuzzy Sets and Systems*, 73, 1995, p. 313-327.
- [Ci 86] M. R. CIVANLAR et H. J. TRUSSELL, Constructing membership functions using statistical data, *Fuzzy Sets and Systems*, 18, 1986, p. 1-13.
- [Ch 95] S. CHAUVIN, Évaluation des théories de la décision appliquées à la fusion de capteurs en imagerie satellitaire, *Thèse de Doctorat d'Université*, Nantes, 1995.
- [Co 94] E. COX, *The fuzzy systems handbook*, 1994, Academic Press, Cambridge, MA.
- [Du 80] D. DUBOIS, H. PRADE, Fuzzy sets and systems: Theory and applications, *Mathematics in science and engineering*, Vol. 144, Academic Press, 1980.
- [Du 87] D. DUBOIS et H. PRADE, *Théorie des possibilités, applications à la représentation des connaissances en informatique*, Masson, 1987.
- [Ga 87] B. R. GAINES, Fuzzy and probability uncertainty logics, *Information and Control*, 38, 1978, p. 297-323.
- [Gi 88] R. GILES, The concept of grade of membership, *Fuzzy Sets and Systems*, 25, 1988, p. 297-323.
- [Go 89] D. E. GOLDBERG, *Genetic algorithms in search optimization and machine learning*, Addison-Wesley, 1989.
- [Ha 88] A. HART, *Acquisition du savoir pour les systèmes experts*, Masson, 1988.
- [Ha 92] A. HART, *Knowledge acquisition for expert systems*, 2nd edition, McGraw-Hill, 1992.
- [He 94] H. HEIDER, V. TRYBA et E. MÜHLENFELD, Automatic design of fuzzy systems by genetic algorithms, *Proceedings of the fifth international conference IPMU. Information Processing and Management of Uncertainty in Knowledge-Based System*, 1994, p. 665-670.
- [He 79] A. HERSCH, A. CARAMAZZA et H. H. BROWNELL, Effects of context on fuzzy membership function, in: M. M. GUPTA, R. K. RAGADE et R. R. YAGER, *Advances in fuzzy set theory and applications*, NorthHolland, Amsterdam, 1979.
- [Hi 88] E. HISDAL, Are grades of membership probabilities?, *Fuzzy Sets and Systems*, 25, 1988, p. 325-348.
- [Hi 93] E. HISDAL, Reconciliation of the Yes-No versus grade of membership dualism in human thinking, in: D. DUBOIS, H. PRADE et R. R. YAGER, *Readings in Fuzzy Sets for Intelligent Systems*, 1993, p. 854-860.

- [Kr 94] R. KRISHNAPURAM, Generation of membership functions via possibilistic clustering, *Proceedings of the third IEEE Conference on Fuzzy Systems. IEEE World Congress on Computational Intelligence*, Vol. 2, 1994, p. 902-908.
- [Ma 92] A. MAEDA, R. SOMEYA et M. FUNABASHI, A fuzzy-based expert system building tool, *Fuzzy Engineering Toward Human Friendly Systems '92*, 1992, p. 1134-1135.
- [Me 93] X. MENAGE et R. HARTANI, Synthèse des méthodes d'association des techniques neuronales et des techniques floues, rapport technique LAFORIA 93/23, 1993.
- [Mi 92] Z. MICHALEWICZ, *Genetic algorithms+Data Structures=Evolution programs*, Artificial Intelligence, Springer, Berlin, 1992.
- [No 84] A. M. NORWICH, I. B. TURKSEN, A model for the measurement of membership and the consequences of its empirical implementation, *Fuzzy Sets and Systems*, 12, 1984, p. 1-25.
- [No 79] M. NOWAKOWSKA, Fuzzy concepts: their structure and problems of measurement, in: M. M. GUPTA, R. K. RAGADE et R. R. YAGER, *Advances in fuzzy set theory and applications*, 1979.
- [PI 86] E. PLAZA, C. ALSINA, R. LOPEZ de MANTARAS, J. AGUILAR et J. AGUSTI, Consensus and knowledge acquisition, *Proceedings of the fifth international conference IPMU. Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 1986, p. 294-306.
- [Ra 79] D. RALESCU, A survey of the representation of fuzzy concepts and its applications, in: M. M. GUPTA, R. K. RAGADE et R. R. YAGER, *Advances in fuzzy set theory and applications*, North-Holland, 1979.
- [Tu 90] I. B. TURKSEN, Measurement of membership functions and their acquisition, *Fuzzy Sets and Systems*, 40, 1990, p. 5-38.
- [Wa 83] P. Z. WANG, From the fuzzy statistics to the falling random subsets, in: P. P. WANG, *Advances in Fuzzy Sets, Possibility Theory and Applications*, 1983, p. 81-96.
- [Za 65] L. A. ZADEH, Fuzzy sets, *Information and Control*, 8, 1965, p. 338-353.
- [Za 75] L. A. ZADEH, The concept of a linguistic variable and its application to approximate reasoning-II, *Information Sciences*, 8, 1975, p. 301-357.
- [Zh 93] L. ZHANG, Structural and functional quantization of vagueness, *Fuzzy Sets and Systems*, 55, 1993, p. 51-60.
- [Zi 85] H.-J. ZIMMERMAN et P. ZYSNO, Quantifying vagueness in decision models, *European Model of Operational Research*, 22, 1985, p. 148-158.

REVUE INTERNATIONALE DE SYSTÉMIQUE

BULLETIN D'ABONNEMENT

À renvoyer à votre libraire spécialisé ou à la S.P.E.S. - Société de Périodiques Spécialisés

Nom _____ Organisme _____

Adresse _____

Pays _____ Date _____

Tarifs 1997 (5 numéros par an)

France	1 050 FF
Export	1 430 FF

- Je désire m'abonner pour 1997
 Je désire recevoir une facture pro-forma
 Paiement joint
 Veuillez débiter ma CB (VISA / EUROCARD / MASTERCARD)

N° _____

Date d'expiration : _____ Signature : _____

S.P.E.S. - Société de Périodiques Spécialisés, BP 22, 41354 VINEUIL cedex - France

Tél. : (33) 02 54 50 46 12 - Fax : (33) 02 54 50 46 11

En application de l'article 27 de la Loi 78-17 Informatique et Liberté vous disposez d'un droit d'accès et de rectification pour toute information vous concernant sur notre fichier. Dunod Éditeur peut être amené à communiquer ces informations aux organismes qui lui sont liés contractuellement, sauf opposition de votre part notifiée par écrit.

Reproduction in whole or in part without the permission of the author or his representative is prohibited (law of March 11, 1957, Article 40, line 1). Such reproduction by whatever means, constitutes an infringement forbidden by Article 425 and those following it of the Penal Code. The law of March 11, 1975, line 2 and 3 of Article 41, authorizes only those copies or reproductions made for the exclusive use to the copyist, and not intended for collective use and such analyses and short quotations as are made for the purposes of an example or illustration.

Toute représentation ou reproduction intégrale ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause, est illicite (loi du 11 mars 1957, alinéa 1^{er} de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait une contrefaçon sanctionnée par les articles 425 et suivants du Code Pénal. La loi du 11 mars 1957 n'autorise, aux termes des alinéas 2 et 3 de l'article 41, que les copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective d'une part et d'autre part que les analyses et les courtes citations dans un bout d'exemple et d'illustration.

The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per-copy fee through the Copyright Clearance Center, Inc., Operations Center, 21 Congress St., Salem, Mass. 01970, U.S.A. for copying beyond that permitted by Sections 107 or 108 or the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purpose, for creating new collective works, or for resale.

© AFCET Gauthier-Villars 1997

Revue Internationale de systémique

RECOMMANDATIONS AUX AUTEURS

La Revue Internationale de Systémique publie essentiellement des articles de langue française. Toutefois, des articles rédigés dans d'autres langues peuvent également être publiés. Ces articles sont consacrés à la systémique (théorie, fondements et épistémologie, sciences de la cognition, application, archives...).

Toute proposition d'article, de note ou de compte rendu, doit être adressée en double exemplaire sous forme dactylographiée (30 pages maximum, en double interligne ou sous forme de disquette), à Revue Internationale de Systémique, B. Paulré, Rédacteur en Chef, 54, rue Bonaparte, 75006 Paris (France).

Chaque article doit être précédé d'un résumé en français et d'un résumé en anglais, de 6 lignes chacun au plus. Les noms des auteurs et leurs adresses professionnelles doivent être indiqués après le titre. Les figures, sur des feuilles séparées et numérotées au crayon, doivent pouvoir être reproduites sans modification. Toute référence est signalée dans le texte par les noms des auteurs suivis de l'année de publication. Les références sont rassemblées dans la bibliographie par ordre alphabétique.

Chaque projet d'article, ou de note, est soumis à deux rapporteurs. Les auteurs d'articles acceptés reçoivent un seul jeu d'épreuves qu'ils doivent retourner, corrigé, dans un délai d'une semaine. Les auteurs reçoivent 50 tirés à part à titre gracieux, ils peuvent en obtenir un plus grand nombre à leurs frais. Ils doivent signaler l'adresse à laquelle "épreuves et tirés à part" doivent être envoyés. La Revue n'est pas responsable des manuscrits.

NOTICE TO AUTHORS

The Revue Internationale de Systémique publishes papers in French and occasionally in other languages. The papers are devoted to systems science (theory, foundations and epistemology, cognitive sciences, applications, archives...).

Every proposed paper, note or review, must be sent in duplicate in type-written form (no more than 30 pages with double spacing or as a disquette) to Revue Internationale de Systémique, B. Paulré, Rédacteur en Chef, 54, rue Bonaparte, 75006 Paris (France).

Each paper must begin with an abstract in French and an abstract in English of no more than 6 lines each. The names of the authors and their professional addresses must be given after the title. Figure, on separate sheets and pencil numbered, must be reproduced without modification. Each reference is cited in the text with the names of the authors and followed by the year of publication. Reference should be listed, in alphabetic order, in the bibliography.

Each proposed paper, or note, is submitted to two referees. The authors of accepted papers receive only one proof which they must send back, after correction, within a week. Authors receive 50 reprints free of charge, they can purchase additional ones. They must give the address to which proofs and reprints must be sent. The Journal is not responsible of the manuscripts.

REVUE INTERNATIONALE DE SYSTÉMIQUE

DUNOD Éditeur, Revues Scientifiques et techniques,

120, bd Saint-Germain, 75280 Paris Cedex 06, France. Tél. : (33) 01.40.46.62.00. Fax : (33) 01.40.46.62.01.
E.mail : gauthier.villars.publisher@mail.sgip.fr Internet: <http://www.gauthier-villars.fr>

Dunod, American Promotion Office. **Gauthier-Villars North America Inc.**
875-81 Massachusetts Avenue, Cambridge, MA 02139, USA. Tel. (1) 617.354.78.75. - Fax (1) 617.354.68.75.

Abonnements/Subscriptions, Société de Périodiques Spécialisés (SPES),
BP 22, 41354 Vineuil Cedex, France. Tél. : (33) 02.54.50.46.12. Fax : (33) 02.54.50.46.11.

Abonnements/Subscription rates: 1997, Vol. 11, un an 5 numéros - one year 5 issues
France : 1 050 F; Autres pays/Foreign countries : 1 430 FF.

Envoi gratuit de spécimen sur demande/*Specimen copy sent on request.*

Réclamations : Les réclamations pour les numéros manquants ne sont recevables que dans un délai de six mois après la date de la publication. *Claims : Claims for missing issues should be made within six months of the publication date.*

Collections et numéros anciens selon disponibilité, sur demande à/Back issues, series annual volumes are available from: Librairie de la Faculté des Sciences, Route de la Fontaine, 89130 TOUCY, France.

AFCET - DUNOD