

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique

Ecole Nationale Polytechnique



PROJET DE FIN D'ETUDE

Présenté au laboratoire d'Automatique
Département d'Electronique
Pour l'obtention du diplôme

D'INGENIEUR D'ETAT EN ELECTRONIQUE

Réalisé Par :

Sofiane HELLALI

Elève Ingénieur en Electronique à l'E.N.P

THEME

MODELISATION FLOUE ET ANALYSE DE SIMILARITE POUR LA SIMPLIFICATION DES REGLES DE BASE FLOUES

Soutenu publiquement le 26/06/2000

Proposé par :

C.LARBES
AIT-CHEIKH

Promotion : juin 2000

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique

Ecole Nationale Polytechnique



PROJET DE FIN D'ETUDE

Présenté au laboratoire d'Automatique
Département d'Electronique
Pour l'obtention du diplôme

D'INGENIEUR D'ETAT EN ELECTRONIQUE

Réalisé Par :

Sofiane HELLALI

Elève Ingénieur en Electronique à l'E.N.P

THEME

MODELISATION FLOUE ET ANALYSE DE SIMILARITE POUR LA SIMPLIFICATION DES REGLES DE BASE FLOUES

Soutenu publiquement le 26/06/2000

Proposé par :

C.LARBES
AIT-CHEIKH

Promotion : juin 2000

E.N.P, 10, Avenue Hassen Badi, 16200 El-Harrach, Alger

Dédicace

Je dédie ce modeste travail à :

Ma mère, mon père, mes sœurs Nacéra, Souhila, Rabéa, Hakima, Lamia ainsi que mon frère Mohamed qui m'ont aidé et soutenu durant toute l'année.

Je tiens à remercier mes promoteurs, Monsieur le président du jury ainsi que Madame et Monsieur les examinateurs sans oublier tout particulièrement Monsieur M.TRABELSI.

Je remercie également mes amis Assia, Omar, Fouad, Mourad, Amar, Yassine, Sofiane, Khaled ainsi que tous ceux que j'ai connu à l'école Nationale Polytechnique .

BIZ

الهدف من هذا العمل هو استعراض تقنيات تجريد و تخفيض احكام القواعد الغامضة عن طريق استعمال مفهوم المماثلة ما بين المجموعات الغامضة.

في البداية قمنا بتقديم لحة عن المفاهيم الاساسية للانظمة الغامضة [المجموعات الغامضة، التحكم الغامض و النموذج الغامضة]. ثم قمنا باقتراح و دراسة تقنيات تجريد الاحكام الغامضة و كذا مفهوم "كلسترينغ" الذي يعطي تحقيقا ذاتيا و تركيبيا للنماذج و الذي لم يكن من السهل قبل ذلك. نتائج اخاكاة العمدية سمحت بتقييم هذه التقنيات.

الكلمات الاساسية: الانظمة الغامضة، النموذج الغامضة، مقاييس المماثلة و تخفيض القواعد الغامضة.

Abstract :

The purpose of this dissertation is to present fuzzy rule-base simplification techniques and their reduction by using the concept of similarity measures between fuzzy sets.

First, main notions of fuzzy systems (fuzzy sets, fuzzy contról and fuzzy modeling) are presented. Then, techniques for fuzzy rule-base simplification and clustering notions which gives identification to structural models are proposed and studied, thing which was difficult before.

The simulation results allows the evaluation of the proposed techniques.

Key words : Fuzzy systems, Fuzzy modeling, similarity measures, fuzzy rule-base simplification.

Résumé :

Le travail présenté dans ce mémoire concerne des techniques de simplification et de réduction des règles floues en utilisant le concept d'analyse de similarité entre ensembles flous.

En premier lieu, nous avons abordé des notions fondamentales des systèmes flous (ensembles flous, contrôle flou, modélisation floue), par la suite, nous avons proposé et étudié des techniques pour la simplification des règles floues, ainsi que la notion de clustering qui donne une identification structurelle des modèles, chose qui n'était pas facile auparavant.

Les résultats de simulation obtenus ont permis l'évaluation de ces techniques.

Mots clés : systèmes flous, modélisation floues, mesures de similarité, simplification des règles flous.

SOMMAIRE



Introduction générale.....	1
----------------------------	---

Chapitre 1 : Notions de base de la logique floue

Introduction.....	2
1.1 Définition d'un sous ensemble flou.....	2
1.2 Caractéristiques d'un ensemble flou.....	2
1.3 Opération sur les ensembles flous.....	3
1.4 Relations floues.....	4
1.4.1 Principe d'extension cylindrique.....	5
1.4.2 Principe de projection.....	6
1.4.3 Composition Sup-Star.....	6
1.5 Théorie de la logique floue appliquée à la représentation et raisonnement linguistique.....	7
1.5.1 Variables linguistiques.....	7
1.5.2 Fonctions d'appartenances.....	8
1.5.3 Propositions floues.....	9
1.6 Conjonction et disjonction des prépositions.....	10
1.6.1 Conjonction.....	10
1.6.2 Disjonction.....	10
1.6.3 Addition de conjonction et disjonction.....	10
1.6.4 Négation.....	11
1.7 Raisonnement flous.....	11
1.7.1 Implication floues.....	12
1.7.2 Inférence floues.....	13
Conclusion.....	13

Chapitre 2 : Méthodologies de commande flous.

Introduction.....	14
2.1 Principe du contrôle flous.....	14
2.1.1 Stratégies de fuzzification.....	16
2.1.2 Base de données.....	16
2.1.3 Discrétisation/Normalisation de l'univers de discours.....	16
2.1.4 Partition des espaces d'entrées et de sorties.....	18
2.1.5 Complétude.....	19
2.1.6 Fonctions d'appartenance d'un ensemble flou.....	19
2.2 Base de règles.....	20
2.2.1 Justification des règles de contrôle flous.....	22
2.2.2 Types de règles de contrôle flous.....	22
2.2.3 Consistance, interactivité et Complétude.....	23
2.3 Logique de prise de décision.....	24
2.3.1 Famille d'implication floues.....	24

2.3.2 Interprétation des connecteurs de phrases ‘Aussi’ et ‘Et’.....	25
2.3.3 Opérateurs compositionnels et mécanismes d’inférences.....	25
2.4 Stratégie de défuzzification.....	28
Conclusion.....	31

Chapitre 3 : Modélisation floue.

Introduction.....	32
3.1 Analyse et description des systèmes flous de Sugeno.....	32
3.1.1 Architecture du modèle flou de Sugeno.....	32
3.1.2 Adaptation des paramètres des systèmes flous de Sugeno.....	34
3.2 Backpropagation.....	34
3.3 Taux d’apprentissage.....	35
3.4 Gradient de la sortie.....	36
3.5 Objectif.....	37
3.6 Simulation et résultats.....	38
Conclusion.....	39

Chapitre 4 : Notions de similarité et clustering

Introduction.....	40
4.1 Notion de clustering et identification par clustering.....	40
4.2 Complexité des règles de base.....	41
4.3 Fonctions d’appartenance similaires.....	42
4.4 La similarité.....	44
4.4.1 Les divers similarités.....	44
4.4.2 Les approximation linguistique.....	45
4.4.3 La similarité et la perception.....	46
4.4.4 Similarité et fonctions d’appartenance.....	47
4.4.5 Les types et les cas.....	47
Conclusion.....	50

Chapitre 5 : Les mesures de similarités.

5.1 Mesures de distances.....	51
5.1.1 Mesure métrique de Minkowski.....	52
5.1.2 Mesure métrique de Hausdorff.....	54
5.1.3 Mesure de Goetschel Voxman.....	55
5.1.4 L’indice de dissemblance.....	57
5.1.5 Les caractéristiques basées sur les mesures de distances.....	58
5.2 Les mesures de similarités.....	60
5.2.1 Neighbourhoodness.....	60
5.2.2 Les mesures de similarités dérivées de la théorie des ensembles clasiques.....	61
5.2.3 L’indice d’égalité.....	63
5.2.4 Corrélation entre deux fonctions d’appartenance.....	66

Conclusion.....	67
<u>Chapitre 6 : Etudes et évaluations des mesures de similarité</u>	
6.1 Grouper ces fonctions de similarité.....	68
6.2 Evaluations.....	70
6.2.1 Les testes.....	70
6.2.2 Critères.....	72
6.3 Les résultats de l'évaluation.....	73
6.3.1 Minkowski.....	73
6.3.2 Hausdorff.....	74
6.3.3 Goetschel Voxman.....	76
6.3.4 Indice de dissemblance.....	76
6.3.5 S_1	77
6.3.6 S_2	77
6.3.7 S_3	78
6.3.8 S_4	78
6.4 Sélection d'une mesure de similarité.....	78
Conclusion.....	81
<u>Chapitre 7 : La simplification des règles de base</u>	
7.1 La fusion des fonctions d'appartenances.....	83
7.1.1 La simplification des règles de base.....	84
7.1.2 La réduction des règles de base.....	84
7.1.3 Comment fusionner.....	86
7.1.4 L'enlèvement des fonctions d'appartenance.....	87
7.2 La méthode.....	88
7.3 Application pour un système à une entrée et une sortie.....	90
7.4 Application à un moteur Asynchrone.....	92
7.5 Application à une enzyme de terre.....	94
7.6 Conclusion.....	96
Conclusion générale.....	97
<u>Annexe :</u>	
Annexe A.....	98
Annexe B.....	100
Annexe C.....	101
Annexe D.....	102
Références Bibliographiques.....	105

LISTE DES SYMBOLES



\mathbb{R}	ensemble de tous les nombres réels
x, y, z	variables du domaine
X, Y, Z, U	domaines/ univers de discours
A, B, C, \dots	ensembles (propres ou flous)
$A = B$	égalité d'ensembles
\bar{A}	complément de A
$A \cup B$	union d'ensemble
$A \cap B$	intersection d'ensemble
$A \subset B$	inclusion d'ensemble
\emptyset	ensemble vide
$A \times B$	produit cartésien des ensembles A et B
$ A $	cardinalité de l'ensemble A
$\ A\ $	cardinalité relative de A
$\{x, y, \dots\}$	ensemble des éléments x, y, \dots
$\{x / p(x)\}$	ensemble défini par la propriété p
i, j, k, m	indices arbitraires
$[a, b]$	intervalle fermé entre a et b
$[a, b)$	intervalle de nombres réels fermé en a et ouvert en b
μ_A	fonction d'appartenance de l'ensemble flou A
d	mesure de distance
$d(A, B)$	distance entre A et B
s	mesure de similarité
$s(A, B)$	similarité entre A et B
d_r	fonction de distance général de Minkowski
d_{nr}	fonction de distance normalisé de Minkowski
d_2	distance Euclidienne
S_{d_2}	similarité basée sur la distance Euclidienne
d_1	fonction distance City-Blok
S_{d_1}	similarité basée sur la distance City-Blok
d_∞	distance de Dominance
S_{d_∞}	similarité basée sur la distance de dominance
q	distance de Hausdorff
q_1	distance de Hausdorff 1
S_{q_1}	similarité basée sur la distance de Hausdorff 1
q_∞	distance de Hausdorff ∞
S_{q_∞}	similarité basée sur la distance de Hausdorff ∞
q^*	distance de Hausdorff *
S_{q^*}	similarité basée sur la distance d e Hausdorff *
Δ	distance entre intervalle propre utilisé pour l'indice dissemblance

Δ_1	indice de dissemblance 1
S_{Δ_1}	similarité basée sur l'indice de dissemblance 1
Δ_∞	indice de dissemblance ∞
S_{Δ_∞}	similarité basée sur l'indice de dissemblance ∞
Δ^*	indice de dissemblance *
Q	fonction distance basée sur la distance de Hausdorff
S_Q	similarité basée sur la distance Q
R	distance de Bhattacharyya
S_R	similarité basée sur la distance de Bhattacharyya
M	mesure du Neighbourhoodness de Chen
S_1	similarité S_1
S_2	similarité S_2
S_3	similarité S_3
S_4	similarité S_4
$(A \equiv B)$	indice d'égalité de Pedrycz
φ	pseudocomplément
$A \square B$	différence symétrique entre ensembles flous A et B.

Introduction Générale

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

La logique floue suscite actuellement un intérêt général de la part des chercheurs, des ingénieurs et des industriels, mais plus généralement de la part de tous ceux qui éprouvent le besoin de formaliser des méthodes empiriques, de généraliser des modes de raisonnement naturels, d'automatiser la prise de décision dans leur domaine, de construire des systèmes artificiels effectuant les tâches habituellement prises en charge par les humains.

En effet, des études ont montré que les systèmes à base de logique floue sont dotés de la propriété d'approximation et de la généralisation des fonctions non linéaires [25]. De ce fait, ces systèmes constituent un outil de plus, à côté des systèmes à réseaux de neurones, outil dans la commande des systèmes non linéaires complexes tels que les moteurs, machines et robots manipulateurs. Ainsi, des algorithmes d'adaptations des paramètres et d'identifications de structure du modèle flou peuvent être développés ce qui permet une modélisation de ces systèmes complexes.

L'objectif de notre travail présenté dans ce mémoire est d'étudier et de porter une modeste contribution au développement des techniques de modélisations et d'identifications structurelles en utilisant de nouvelles méthodes dites « mesures de similarités » en vue de la simplification et la réduction des règles floues.

A cet effet, notre travail est divisée en sept chapitres :

Le premier chapitre présente les concepts de base de la logique floue définissant les ensembles flous, leurs caractéristiques, les relations floues, le raisonnement linguistique ainsi que le raisonnement flou.

Dans le deuxième chapitre, nous développons les méthodologies de commandes floues sachant que la modélisation est faite en vue de commande des systèmes. Ainsi Ce chapitre, a aussi pour objectif de nous faire introduire le raisonnement interne d'un système flou (fuzzification, inférence, défuzzification, etc....).

Dans le chapitre trois, nous nous intéressons à l'extraction automatique des connaissances dans les systèmes flous de Takagi-Sugeno à partir de données numériques et à la construction incrementale de la base de règles, en utilisant l'algorithme de backpropagation qui assure convenablement l'apprentissage.

Le quatrième chapitre servant d'introduction aux chapitres qui le suivent, nous permet de développer les principales notions de similarité. Nous commencerons par présenter la notion de clustering flou, les différentes algorithmes de clustering ainsi que la façon d'appliquer ces algorithmes, sans oublier la notion d'identification des systèmes.

Dans le chapitre cinq, nous introduisant de nouvelles techniques dites « analyses de similarités », elles ont été développées afin de résoudre les problèmes rencontrés dans l'identification structurelle des systèmes.

Dans le chapitre six, nous présentons des études et des évaluations de ces mesures de similarité afin de pouvoir les utiliser dans la simplification des règles floues.

Dans le chapitre sept, nous développons une méthode pour la simplification et la réduction des règles floues basées sur les mesures de similarité et appliquées à des modèle de Takagi-Sugeno.

Enfin, notre mémoire est clôturée par une conclusion générale et une perspective pour une éventuelle continuité du développement de ce nouveau concept.

CHAPITRE 1

Notion de base de la logique floue

Introduction

Etant donné un ensemble de référence U , on peut indiquer les éléments de U qui appartiennent à une certaine classe de U et ceux qui n'appartiennent pas. Cette classe est alors (au sens habituel de la théorie des ensembles) un sous-ensemble de U . On le qualifie de classique ou d'ordinaire dans la suite. Si l'appartenance de certains éléments de U à une classe n'est pas absolue, on peut alors indiquer avec quel degré chaque élément appartient à cette classe. Celle-ci est alors un sous-ensemble flou de U .

1.1 Définition d'un sous-ensemble flou :

La notion de sous-ensemble flou provient du constat établie par L.Zadeh qui stipule que « *très souvent les classes des objets rencontrées dans le monde physique, ne possèdent pas de critères d'appartenance bien définis* » [36]. Le concept de l'ensemble flou permet de considérer des classes d'objets dont les frontières ne sont pas clairement déterminées, par l'introduction d'une fonction caractéristique (fonction d'appartenance des objets à la classe) prenant des valeurs courantes entre 0 et 1, contrairement aux ensembles booléens dont la fonction caractéristique ne prend que deux valeurs possibles 0 ou 1.

En logique classique, un élément x soit il appartient à un ensemble A ou non. Donc, on peut associer à la propriété appartenir la fonction caractéristique suivante :

$$\mu_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases} \quad (1.1)$$

En logique floue, un ensemble dit flou, est la généralisation de la notion de l'ensemble en logique classique. La propriété d'appartenance est définie par la fonction caractéristique suivante :

$$\begin{aligned} \mu_A: U_x &\rightarrow [0 \ 1] \\ x &\rightarrow \mu_A(x) \end{aligned} \quad (1.2)$$

Où, U_x est l'univers de discours de x . $\mu_A(x)$ est appelée fonction d'appartenance, l'ensemble flou noté A est défini par l'ensemble des couples $\{x, \mu_A(x)\}$. Donc :

$$A = \int_{U_x} \mu(x)/x \quad \text{si } U \text{ est infinie}$$

Le symbole \int_{U_x} indique la collection sur tout l'intervalle U_x .

$$\text{Si } U \text{ est fini, on a } A = \sum_{x \in U_x} \mu(x)/x$$

1.2 Caractéristiques d'un ensemble flou [1] [10] [11]:

Pour pouvoir décrire facilement un ensemble flou A de U , on utilise certaines de ses caractéristiques, essentiellement celles qui montrent dans quelle mesure il diffère d'un ensemble classique de U

1.2.1 Support de A :

C'est la première des caractéristiques que l'on peut citer, il définit l'ensemble des éléments de U qui appartiennent, au moins un peu à A. Il est noté $\text{Supp}(A)$ et c'est la partie de U sur laquelle la fonction d'appartenance de A n'est pas nulle :

$$\text{Supp}(A) = \{x \in U / \mu_A(x) \neq 0\}. \quad (1.3)$$

1.2.2 Hauteur de A :

Notée $h(A)$, c'est le plus fort degré avec lequel un élément de U appartient à A. c'est la plus grande valeur prise par sa fonction d'appartenance

$$h(A) = \text{Sup}_{x \in U} \mu_A(x)$$

L'ensemble A pour lequel il existe au moins un élément de U appartenant de façon absolue (avec un degré 1) à A correspond à ceux qui sont normalisés. Plus précisément, A est normalisé si sa hauteur $h(A)$ est égale à 1.

1.2.3 Noyau de A :

L'ensemble de tous les éléments appartenant de façon absolue (avec un degré 1) à A est appelé le noyau de A et noté $\text{noy}(A)$:

$$\text{noy}(A) = \{x \in U / \mu_A(x) = 1\} \quad (1.4)$$

si A est un ensemble ordinaire de U, il est normalisé et il est identique à son support et à son noyau.

1.2.4 Cardinalité :

Une dernière caractéristique de l'ensemble flou A de U (lorsque U est fini) est sa cardinalité, évaluant le degré global avec lequel les éléments de U appartiennent à A. Elle est définie par :

$$|A| = \sum_{x \in U} \mu_A(x) \quad (1.5)$$

si A est un ensemble ordinaire de U, sa cardinalité est le nombre d'éléments qui le composent, selon la définition classique.

1.3 Opérations sur les ensembles flous :

Puisque les fonctions caractéristiques des ensembles flous sont graduelles l'extension de certaines opérations définies sur les ensembles classiques au cas des ensembles flous pose certaines difficultés. Donc, il n'est pas possible d'appliquer directement la logique propositionnelle qui manipule des ensembles à fonction caractéristiques binaires (0 ou 1). Pour palier à ce problème, on introduit les définitions de trois outils nécessaires pour le traitement des ensembles flous.

Soient, les nombres réels : a, b, c et $d \in [0, 1]$.

a°/ Définition de norme triangulaire (T-norme) : [34]

La norme triangulaire de a et b , notée $a \otimes b$, dénote une classe de fonctions binaires qui peuvent représenter l'opération d'intersection entre ensembles. Elle satisfait les propriétés suivantes :

$$T.1 \quad a \otimes b = b \otimes a$$

$$T.2 \quad (a \otimes b) \otimes c = a \otimes (b \otimes c)$$

$$T.3 \ a \leq c \text{ et } b \leq d \Rightarrow a \otimes b \leq c \otimes d$$

$$T.4 \ a \otimes 1 = a$$

b°/ Définition de norme co-triangulaire (T-conorme ou S-norme)

La conorme triangulaire de a et b , notée $a \oplus b$, dénote une classe de fonctions binaires qui peuvent représenter l'opération d'union entre ensembles. Elle satisfait les propriétés suivantes :

$$S.1 \ a \oplus b = b \oplus a$$

$$S.2 \ (a \oplus b) \oplus c = a \oplus (b \oplus c)$$

$$S.3 \ a \leq c \text{ et } b \leq d \Rightarrow a \oplus b \leq c \oplus d$$

$$S.4 \ a \oplus 0 = a$$

c°/ Définition du complément

L'opérateur de complément noté $(\bar{\quad})$, doit au moins satisfaire les propriétés suivantes :

$$C.1 \ a(\bar{0}) = 1$$

$$C.2 \ a < b \Rightarrow \bar{a} > \bar{b}$$

$$C.3 \ (\bar{\bar{a}}) = a$$

d°/ Définitions de l'intersection d'ensembles flous

L'intersection de deux ensembles flous A_1 et A_2 est l'ensemble $A = A_1 \cap A_2$ défini sur l'univers U_x de fonction d'appartenance :

$$\mu_A(x) = \mu_{A_1 \cap A_2}(x) = \mu_{A_1}(x) \otimes \mu_{A_2}(x) \quad (1.6)$$

Exemple : Parmi les opérations d'intersection :

- opérateur noté « min » : $\mu_{A_1 \cap A_2}(x) = \min(\mu_{A_1}(x), \mu_{A_2}(x))$
- opérateur produit noté « x » : $\mu_{A_1 \cap A_2}(x) = \mu_{A_1}(x) \times \mu_{A_2}(x)$

e°/ Définition de l'union d'ensembles flous

L'union de deux ensembles flous A_1 et A_2 est l'ensemble $A = A_1 \cup A_2$ défini sur l'univers de discours U_x , de fonction d'appartenance :

$$\mu_A(x) = \mu_{A_1 \cup A_2}(x) = \mu_{A_1}(x) \oplus \mu_{A_2}(x) \quad (1.7)$$

Exemple : Parmi les opérations d'union :

- opérateur maximum noté « max » : $\mu_{A_1 \cup A_2}(x) = \max(\mu_{A_1}(x), \mu_{A_2}(x))$
- opérateur somme algébrique : $\mu_{A_1 \cup A_2}(x) = \mu_{A_1}(x) + \mu_{A_2}(x) - \mu_{A_1}(x) \times \mu_{A_2}(x)$

f°/ Définition de complément d'ensemble flou

le complément d'un ensemble flou A_1 et l'ensemble flou \bar{A}_1 défini sur l'univers de discours U_x dont la fonction d'appartenance est donnée par :

$$\mu_{\bar{A}_1}(x) = \overline{\mu_{A_1}(x)} \quad (1.8)$$

Exemple : opérateur de complément : $\mu_{\bar{A}_1}(x) = 1 - \mu_{A_1}(x)$

1.4. Relations floues [34] [36] [37] :

En logique classique, une relation représente la présence ou l'absence d'association, d'interaction ou d'interconnexion entre les éléments de deux ensembles ou plus. En logique floue, une relation traduit le degré de présence ou d'absence d'association, d'interaction ou d'interconnexion entre les éléments de deux ensembles flous ou plus.

Soit, $U_{x_i}, i=1 \dots n$: des univers de discours. Une relation floue R définie sur l'espace de produit $U_{x_1} \times U_{x_2} \times \dots \times U_{x_n} = \prod_{i=1}^n U_{x_i}$ est un ensemble flou caractérisé par la fonction d'appartenance μ_R telle que :

$$R = \int_{\prod_{i=1}^n U_{x_i}} \mu_R(x_1, x_2, \dots, x_n) / (x_1, x_2, \dots, x_n)$$

Exemple : Considérons la relation binaire R : "x est très proche de y" définie sur l'univers $U_x \times U_y$ tel que : $U_x = \{0.1, 0.2, 0.3\}$; $U_y = \{0, 0.1, 0.2\}$. la fonction d'appartenance de R peut être décrite par la table

y ▼ \ x ►	0.1	0.2	0.3
0	0.5	0.25	0
0.1	1	0.5	0.25
0.2	0.5	1	0.5

Table (1.1) : table de la fonction μ_R

Opérations sur les relations floues

Vu que les relations floues sont définies par des ensembles flous, il est possible d'utiliser les opérations d'intersection, d'union et de complément pour combiner plusieurs relations floues définies sur le même univers de discours.

Soit R et S deux relations floues définies sur l'univers $U_x \times U_y$. L'intersection de R et S est définie par l'ensemble flou : $R \cap S$ tel que :

$$R \cap S = \int_{U_x \times U_y} \mu_{R \cap S}(x, y) / (x, y) = \int_{U_x \times U_y} \mu_R(x, y) \otimes \mu_S(x, y) / (x, y) \quad (1.9)$$

L'union des deux relations R et S est définie par l'ensemble flou $R \cup S$ tel que :

$$R \cup S = \int_{U_x \times U_y} \mu_{R \cup S}(x, y) / (x, y) = \int_{U_x \times U_y} \mu_R(x, y) \oplus \mu_S(x, y) / (x, y) \quad (1.10)$$

\otimes et \oplus dénotent respectivement la T-norme et la S-norme.

1.4.1 Principe d'Extension cylindrique

Il permet l'extension de l'univers de discours d'un ensemble ou d'une relation floue à l'univers de dimension plus grande.

Soit l'univers de discours $U = \prod_{i=1}^n U_{x_{im}}$ et soit V un sous univers de U tel que :

$V = \prod_{m=1}^k U_{x_{im}} \quad i_m \leq n, k < n$, sur lequel on définit une relation R

$$R = \int_V \mu_R(x_{i1}, \dots, x_{ik}) / (x_{i1}, \dots, x_{ik})$$

L'extension cylindrique de R sur U est définie par :

$$Ce_U(R) = \int \mu_R(x_{i1}, \dots, x_{ik}) / (x_1, \dots, x_n) \tag{1.11}$$

Grâce à l'introduction de la notion de l'extension cylindrique, il sera possible de réaliser les opération d'union, d'intersection entre ensembles et relations flous. Il suffit d'effectuer des extensions sur un même univers contenant tous les univers des ensembles et relations, et d'appliquer l'opération envisagée.

1.4.2 Principe de projection

A partir d'une relation ou ensemble flou défini sur un univers donné, la projection permet d'extraire une relation définie sur un univers de discours de dimension inférieurs.

Soit la relation définie sur l'univers $U = \prod_{i=1}^n U_{x_i}$. la projection de R sur l'univers $V =$

$\prod_{m=1}^k U_{x_{ik}} \quad i_m \leq n, k < n$, est la relation $Proj_V(R)$ définie par :

$$Proj_V(R) = \int_{V, x_{j1}, \dots, x_{jl}} \text{Sup } \mu_R(x_1, \dots, x_n) / (x_{i1}, \dots, x_{ik}) \tag{1.12}$$

Où $\prod_{m=1}^l U_{x_{im}}$ est l'univers de discours complémentaire de V dans U ($l+k=n$)

1.4.3 Composition Sup-Star :

La composition Sup-Star est une opération indispensable pour effectuer des combinaisons entre ensembles et relations flous. Elle est basée sur les principes de l'Extension cylindrique et de la Projection.

La combinaison d'ensemble flou avec une relation floue et appelée *composition*. Soit un ensemble flou A défini sur $\prod_{i=1}^n U_{x_i}$ et R une relation définie sur

$\prod_{i=1}^n U_{x_i} \times \prod_{i=1}^m U_{y_i}$. La composition Sup-Star de A et R

Est l'ensemble flou B défini sur $\prod_{i=1}^m U_{y_i}$ tel que :

$$B = A \circ R = \text{Proj}_{\prod_{i=1}^m U_{y_i}} \left(Ce_{\prod_{i=1}^n U_{x_i} \times \prod_{i=1}^m U_{y_i}} (A) \cup R \right) \tag{1.13}$$

$$\text{Donc } B = \int_{\prod_{i=1}^m U_{y_i}} \text{Sup}_{(x_1, \dots, x_n)} (\mu_A(x_1, \dots, x_n) \otimes \mu_R(x_1, \dots, x_n, y_1, \dots, y_m)) / (y_1, \dots, y_m) \tag{1.14}$$

1.5 Théorie de la logique floue appliquée à la représentation et raisonnement linguistique [34] [10] [11] :

L'utilisation des ensembles et des relations flous offre une base pour une voie systématique de manipulation de concepts vagues et imprécis. En particulier, on peut représenter et raisonner sur des expressions linguistiques ou des mots du langage naturel.

Considérons l'expression verbale suivante :

“Si c'est le cas ou l'erreur actuelle a la propriété d'être négative grande et la variation de l'erreur actuelle a la propriété d'être positive grande, alors c'est une raison pour changer la commande en l'incrémentant pour qu'elle ait la propriété négative petite” [34].

Dans cette expression interviennent les notions clefs suivantes :

- des variables linguistiques : erreur, variation d'erreur.....
- propriétés des variations linguistiques : négative grande, positive grande....
- Des connecteurs linguistiques : “et “ , “ou“ “a la propriété “
- Propositions :
 - “Erreur actuelle... négative grande“
 - “Variation de l'erreur ... positive grande“
- Conditionnement : SI Alors.....

Dans le but de représenter l'expression précédente à l'aide de la logique floue, on introduit une représentation symbolique des différentes notions citées ci-dessus. L'association d'ensembles et de relations flous à la représentation symbolique permettra de quantifier l'expression.

1.5.1 Variables linguistiques

C'est l'unité fondamentale pour représenter les connaissances pour être traitées par le raisonnement à logique floue. Une variable linguistique est une variable dont les valeurs sont définies par des termes du langage naturel (mots ou phrases).

Chaque variable linguistique est déterminée par quatre caractéristiques (X , LX , U_x , M_x) où

1. X : le nom symbolique de la variable. Exemple : E pour l'erreur, ΔE pour la variation d'erreur.
2. LX : l'ensemble des valeurs de la variable X. Exemple : $LE=\{N, Z,P\}$ où N, Z et P indiquent respectivement, négative, nulle et positive.
3. U_x : l'univers de discours dans lequel la variable de symbole X prend ses valeurs. Si la variable est normalisée entre -1 et $+1$ alors $U_x = [-1, +1]$.
4. M_x : fonction sémantique qui attribue à chaque valeur de X son sens. Ce dernier est caractérisé par un ensemble flou défini sur l'univers de discours U_x , donc :

$$M_x \quad LX \longrightarrow LX$$

Où :

$$LX = \int_{U_x} \mu_{LX}(x_i) / x_i$$

Donc la variable X peut être considérée comme une variable floue dont les valeurs sont les éléments LX.

Exemple : Considérons le cas de la variable Erreur, on a $LE=\{N, Z, P\}$. la fonction M_x consiste à attribuer à ces valeurs des ensembles flous. La figure [1.1] montre une possibilité de choix d'ensembles.

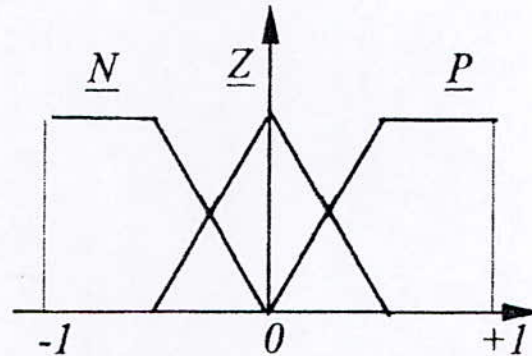


Figure 1.1 : Exemple d'ensemble flou associés à la variable erreur

1.5.2 Fonction d'appartenance

Afin de permettre un traitement numérique des variables linguistiques dans la prise de décisions floues sur ordinateur, une définition des variables linguistiques à l'aide de fonctions d'appartenance s'impose. Dans ce contexte, on associe à chaque valeur de la variable linguistique une fonction d'appartenance désignée par $\mu_A(x)$, où x est la variable linguistique, tandis que A indique l'ensemble concerné. Une valeur précise $\mu_A(x)$, sera désignée par le degré ou le facteur d'appartenance. Il est à noter que l'ensemble des éléments de x pour lesquels $\mu_A(x) > 0$, est appelé support de A . Le plus souvent, on utilise pour les fonctions d'appartenance les fonctions suivantes figure 1.2

- *fonction triangulaire* : Elle est définie par trois paramètres $\{a, b, c\}$, qui déterminent les coordonnées des trois sommets.

$$\mu(x) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \quad (1.15)$$

- *fonction trapézoïdale* : Elle est définie par quatre paramètres $\{a, b, c, d\}$

$$\mu(x) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right) \quad (1.16)$$

- *Fonction gaussienne* : Elle est définie par deux paramètres $\{\sigma, m\}$:

$$\mu(x) = \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (1.17)$$

- **Fonction sigmoïdale** : Une fonction sigmoïdale est définie par deux paramètres $\{a, c\}$

$$\mu(x) = \frac{1}{1 + \exp(-a(x - c))} \quad (1.19)$$

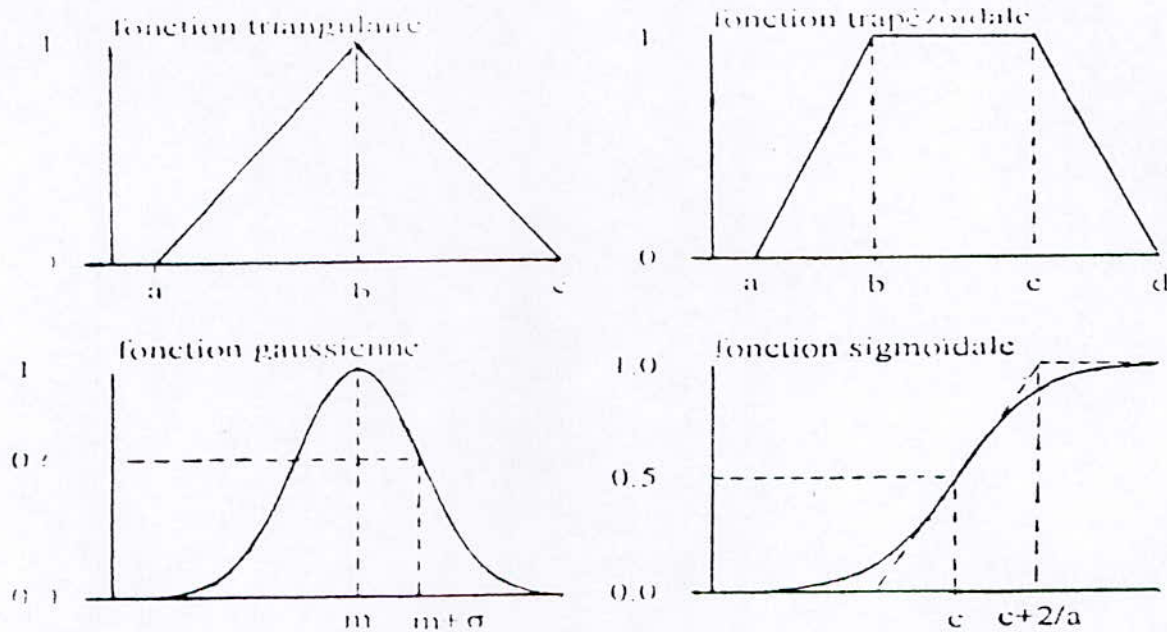


Figure [1.2] : Formes usuelles des fonctions d'appartenance

Remarque :

L'attribution de fonctions d'appartenance aux valeurs d'une variable linguistique est une opération subjective, donc elle diffère d'une personne à une autre. Cela est dû au fait que les concepts abstraits (tel que la description linguistique) sont de nature imprécise.

Comme indique la figure [1.1], dans les zones de chevauchement, la variable floue prend plus d'une valeur à la fois mais avec des degrés d'appartenance différents, ce qui reflète l'imprécision sur les frontières de ces valeurs. Dans le cas général, le passage entre des valeurs floues voisines (qui se chevauchent) est graduel (progressif).

1.5.3 Propositions floues :

on distingue deux types de propositions :

1.5.3.1 Proposition floue simple

Considérons l'expression : "Erreur a la propriété d'être négative grande". Une représentation symbolique possible est la suivante :

"E est NG"

où "est" réalise le sens de "la propriété d'être" ;

NG : le symbole d'ensemble flou associé à la valeur négative grande.

NG : le symbole d'ensemble flou associé à la valeur négative grande.
 Dans ce cas la représentation symbolique "E est NG", est appelée proposition floue simple.

1.5.3.2 Proposition floue composée

La combinaison de plusieurs propositions floues simples par les connecteurs "et", "ou" et "Si-Alors" permet de générer des propositions floues composées.

Exemple

Considérons les valeurs de l'erreur $LE=\{NG, NM, NP, Z, PP, PM, PG\}$ et de sa variation $\Delta E=\{N, Z, P\}$, où NG désigne négative grande, NM : négative moyenne, NP : négative petite, Z : nulle, PP : positive petite, PM : positive moyenne, PG : positive grande.

A partir de ces valeurs, on peut construire les propositions floues composées suivantes :

"E est NG ou E est PP" ; "E est NG et ΔE est P" ; "E n'est pas P".

En logique floue, de telles propositions peuvent être quantifiées en utilisant les notions : conjonction, disjonction et négation.

1.6 Conjonction et disjonction des propositions

1.6.1 Conjonction

La conjonction des propositions P_1, P_2, \dots, P_n définies par

P_1 : "X est A_1 ", ..., P_n : "X est A_n "

Est la proposition composée P telle que :

P : "X est A_1 et X est A_2 et ... X est A_n ".

Le sens de la proposition P est représenté par un ensemble flou A résultat de l'intersection des ensembles flous A_1, A_2, \dots, A_n . D'après (), on obtient :

$$\mu_A(x) = \mu_{A_1}(x) \otimes \mu_{A_2}(x) \otimes \dots \otimes \mu_{A_n}(x) \quad x \in U_x \quad (1.20)$$

on écrit : $\mu_A(x) \hat{=} \bigotimes_{i=1}^n \mu_{A_i}(x)$

1.6.2 Disjonction

La disjonction des propositions P_1, P_2, \dots, P_n définies par :

P_1 : "X est A_1 ", ..., P_n : "X est A_n "

Est la proposition composée P telle que :

P : "X est A_1 ou X est A_2 ou ... X est A_n ".

Le sens de la proposition P est représenté par un ensemble flou A résultat de l'union des ensembles flous A_1, A_2, \dots , et A_n . D'après (), on obtient :

$$\mu_A(x) = \mu_{A_1}(x) \oplus \mu_{A_2}(x) \oplus \dots \oplus \mu_{A_n}(x) \quad x \in U_x \quad (1.21)$$

on écrit : $\mu_A(x) \hat{=} \bigoplus_{i=1}^n \mu_{A_i}(x)$

1.6.3 Addition de conjonction et disjonction

Considérons la proposition composée suivante :

P : " x_1 est A_1 et x_2 est A_2 et ... x_n est A_n ".

Contrairement au sens de la conjonction décrit précédemment, dans ce cas les variables $\{x_i\}_{i=1..n}$ dans la proposition P sont définies sur des univers différents $\{U_{x_i}\}_{i=1..n}$. Le sens de la proposition P est représenté par une relation R définie

sur l'univers $\prod_{i=1}^n U_{x_i}$. Elle est obtenue en utilisant le principe d'extension cylindrique et l'opérateur d'intersection. D'après (I.1) et (I.3), on a :

$$R = \int_{\prod_{i=1}^n U_{x_i}} Ce(A_i) = \int_{i=1}^n \otimes \mu_{A_i}(x)/(x_1 \dots x_n) \tag{1.22}$$

Où :

$$\int_{i=1}^n \otimes Ce(A_i) \hat{=} \int_{i=1}^n Ce(A_1) \cap \int_{i=1}^n Ce(A_2) \cap \dots \cap \int_{i=1}^n Ce(A_n)$$

De la même façon, dans le cas de disjonction, le sens de la proposition P suivante :

P : "x₁ est A₁ ou x₂ est A₂ ou ... x_n est A_n"

Est représenté par la relation floue :

$$R = \int_{\prod_{i=1}^n U_{x_i}} Ce(A_i) = \int_{i=1}^n \oplus \mu_{A_i}(x)/(x_1 \dots x_n) \tag{1.23}$$

Où :

$$\int_{i=1}^n \oplus Ce(A_i) \hat{=} \int_{i=1}^n Ce(A_1) \cup \int_{i=1}^n Ce(A_2) \cup \dots \cup \int_{i=1}^n Ce(A_n)$$

1.6.4 Négation

La négation d'une proposition P donnée par :

P : "x est A" est la proposition P : "x n'est pas A"

Dont le sens est représenté par l'ensemble flou A tel que :

$$\bar{A} = \int_{U_x} \mu_{\bar{A}}(x)/x$$

Où \bar{A} est l'ensemble flou complément de l'ensemble flou A.

1.7. Raisonnement flou

En général, plusieurs valeurs de variables linguistiques convenablement définies par des fonctions d'appartenances sont liées entre elles par des règles afin de tirer des conclusions. Ces règles utilisent souvent les connecteurs Si-Alors. La procédure utilisée pour des conclusions à partir d'un ensemble de règles et de conditions, représentées par des ensembles et relations flous, est appelée raisonnement flou.

Dans le raisonnement flou intervient trois niveaux fondamentaux :

1. Implication floue ;
2. Inférence floue ;
3. Agrégation floue.

1.7.1 Implication floue

Les règles de la forme Si-Alors qui manipulent des variables linguistiques (ou flou), sont appelées règles floues. L'utilisation des opérations sur des ensembles et de relations flous nous a permis dans la section précédente (1.6), d'associer des ensembles flous à des propositions floues complexes. Donc, il sera toujours possible d'écrire une règle de la forme :

SI x_1 est A_1 ... et x_n est A_n Alors (Y est B)

Sous la nouvelle forme :

$$\text{SI } X \text{ est } A \text{ Alors } Y \text{ est } B \tag{1.24}$$

Où, $X=[x_1, \dots, x_n]$ vecteur des variables d'entrée.

A : ensemble flou défini sur l'univers de discours $\prod_{i=1}^n U_{x_i}$

Donc,

$$A = \int_{\prod_{i=1}^n U_{x_i}} \mu_A(x_1, \dots, x_n) / (x_1, \dots, x_n) = \int_{\prod_{i=1}^n U_{x_i}} \bigotimes_{i=1}^n \mu_{A_i}(x_i) / (x_1, \dots, x_n) \tag{1.25}$$

Le sens de (1.25) est représenté par une relation d'implication floue définie sur l'univers de discours $X \times U_y$. Cette relation mesure le degré de vérité de la règle floue, donc de l'implication entre antécédent et conséquence.

En logique classique une relation d'implication possède plusieurs représentation dont les résultats sont identiques. Considérons les deux variables logiques binaires p_1 et p_2 . La relation d'implication $p_1 \Rightarrow p_2$ est équivalente aux deux relations $(\bar{p}_1 \vee p_2)$ ou $[(p_1 \wedge p_2) \vee \bar{p}_1]$, avec les opérateurs de la logique classique $\vee, \wedge, \bar{}$, désignant respectivement les opérateurs "et", "ou", "complément".

Sur la base de ces deux relations, nous avons introduit plusieurs définitions de l'implication floue, parmi lesquelles nous citons :

- La norme de Klane-Dienes : basée sur la relation $\bar{p}_1 \vee p_2$, l'opérateur "et" est réalisé par T-norme "maximum". Elle est définie par la fonction d'appartenance suivante :

$$\mu_{A \rightarrow B}(x_1, \dots, x_n, y) = \max(1 - \mu_A(x_1, \dots, x_n), \mu_B(y))$$

- La norme de Lukasiewicz : basée sur la relation d'implication floue est donnée par la relation :

$$\mu_{A \rightarrow B}(x_1, \dots, x_n, y) = \min(1, 1 - \mu_A(x_1, \dots, x_n) + \mu_B(y))$$

- La norme de Zadeh : basée sur la relation $(p_1 \wedge p_2) \vee p_2$, l'opérateur "et" est utilisé par T-norme "somme-bornée", la relation d'implication floue est donnée par la relation :

$$\mu_{A \rightarrow B}(x_1, \dots, x_n, y) = \max(\min(\mu_A(x_1, \dots, x_n), \mu_B(y)), 1 - \mu_A(x_1, \dots, x_n))$$

La littérature rapporte deux autres normes très utilisées[36, 37] : la norme de Mamdani et la norme de Larsen. En réalité, elles ne sont pas des implications mais plutôt des conjonctions, car elles traduisent d'une certaine manière une notion de progressivité de la forme :

plus x est A plus y est B.

elles sont définies par les relations suivantes :

- La norme de Mamdani : $\mu_{A \rightarrow B}(x_1 \dots x_n, y) = \min(\mu_A(x_1 \dots x_n), \mu_B(y))$
- La norme de Larsen : $\mu_{A \rightarrow B}(x_1 \dots x_n, y) = \mu_A(x_1 \dots x_n) \times \mu_B(y)$

1.7.2 Inférence flou :

La procédure utilisée pour déduire des conclusions à partir d'un ensemble de règles floues et d'une ou plusieurs conditions est appelée *Inférence floue* (raisonnement flou, ou encore raisonnement approximatif)

En logique floue, il y a deux importantes règles d'inférence : *Modus Tollens Généralisé* "MTG" et le *Modus Ponens Généralisé* "MPG"

1.MTG :

condition : Y est B'
règle floue : Si x est A Alors Y est B
 ensemble de résultat : x est A'

2.MPG

condition : x est A'
règle floue : Si x est A Alors y est B
 ensemble de résultat : Y est B'

l'implication floue est basée sur la règle compositionnelle d'inférence pour le raisonnement approximatif proposé par Zadeh. Ici on a introduit les ensembles flous A, A', B et B' par des variables linguistiques x, y au lieu des ensembles ordinaires de la logique traditionnelle. Le MPG, qui se réduit au 'Modus ponens' quand A'=A et B'=B, est utilisé dans le raisonnement par chaînage avant (déduction successive de faits) et est particulièrement utile dans la commande floue. Quant au MTG, utilisé en chaînage inférentiel arrière (vérification d'hypothèse), il trouve son domaine d'application dans les systèmes experts et spécialement ceux du diagnostic médical.

Conclusion :

Ce chapitre a été consacré aux rappels sur la théorie de la logique floue en vue de son application pour la commande. La notion de sous ensemble flou a pour but de permettre des gradations dans l'appartenance d'un élément à une classe, c'est à dire d'autoriser un élément à appartenir plus au moins fortement à cette classe.

Nous avons présenté tout d'abord une définition accés claire des ensembles flous en donnant leurs différentes caractéristiques en passant du support à la hauteur sans oublier le noyau et la cardinalité des ensembles, puis il est évident qu'à tout ensemble on associe des opérateurs tels que l'intersection et l'union etc.... En introduisant dans la théorie de la logique floue le concept de raisonnement linguistique et variables linguistiques nous avons donné une interprétation plus avancée et plus proche du raisonnement humain.

Notons qu'on parle souvent d'ensemble flou et non de sous ensemble flou, par abus de langage et conformément à la traduction du terme original de « fuzzy set ».

CHPITRE 2

Méthodologies de Contrôle Flou

Introduction :

Le contrôle flou est le domaine dans lequel il existe le plus de réalisations effectives, en particulier industrielles. Son but est de traiter les problèmes de commande du processus, plus souvent à partir des connaissances des experts ou des opérateurs qualifiés

Ce chapitre, a aussi pour objectif de nous faire introduire le raisonnement interne d'un système flou (fuzzification, inférence, défuzzification, etc....)

2.1 Principe du contrôle flou [16] [17] [15] [18] :

Après avoir passer en revue les principaux concepts de la théorie des ensembles flous, nous pouvons à présent procéder au développement des notions importantes constituant la base des contrôleurs flous.

La figure 2.1 montre le schéma synoptique général d'un contrôleur flou appliqué à un processus.

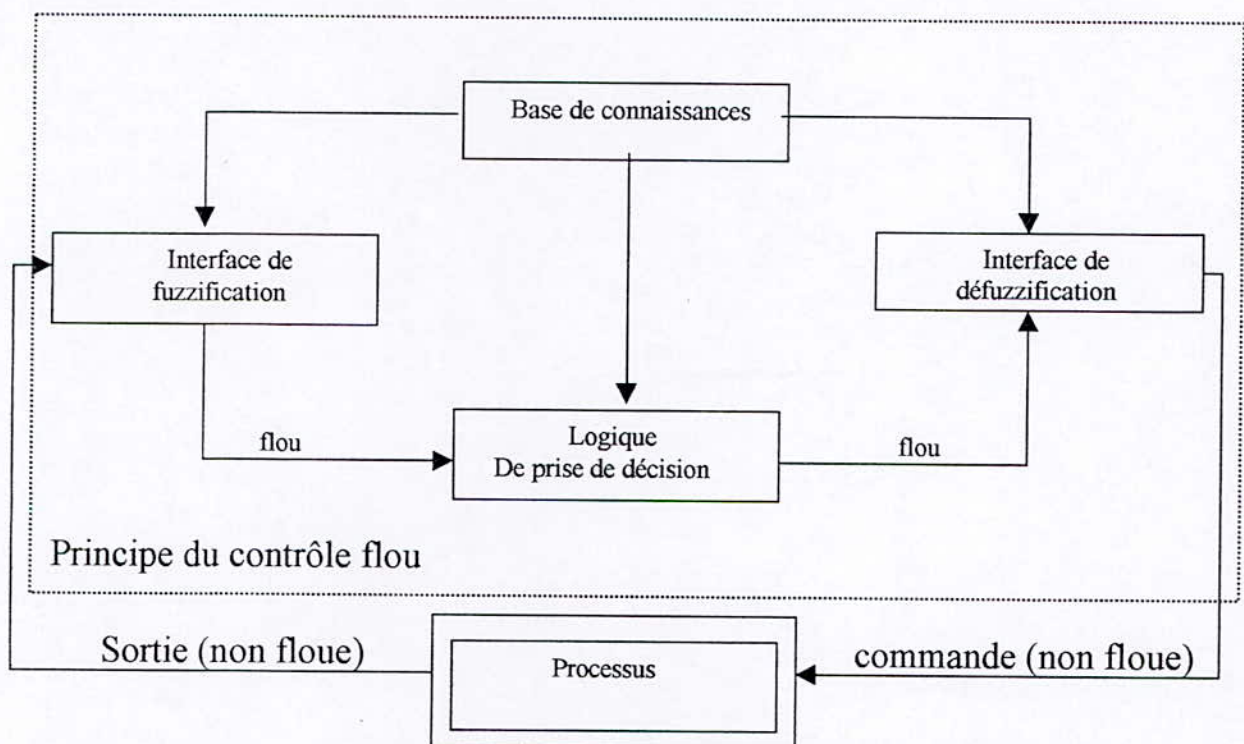


Figure [2.1] : schéma synoptique général d'un contrôleur flou

Ce schéma est composé comme suit :

- une interface de fuzzification à l'entrée
- une base de connaissances
- une logique de prise de décision
- Une interface de défuzzification à la sortie.

1) Interface de fuzzification :

L'interface de fuzzification inclue les fonctions suivantes :

- a) mesure les valeurs des variables d'entrée,
- b) réalise une cartographie d'échelle transférant la plage des variables d'entrée aux univers de discours correspondants,

- c) réalise la fonction de fuzzification qui convertit les données d'entrée en valeurs linguistiques pouvant être vues comme étiquettes d'ensembles flous.
- 2) Base de connaissance :
- Elle comprend une connaissance du domaine d'application et les buts du contrôle prévu. Elle consiste en une "base de données" et une "base de règles de contrôle linguistique flou" :
- a) La base de données fournit les définitions nécessaires utilisées pour définir les règles de contrôles linguistiques et la manipulation des données floues dans le contrôleur flou,
 - b) La base de règles caractérise les buts et la stratégie de la commande émis par les experts du domaine au moyen d'un ensemble de règles linguistiques de contrôle.
- 3) Logique de prise de décision (inférence) :
- La logique de prise de décision est le noyau du contrôleur flou ; elle a la capacité de simuler la prise de décision de l'être humain en se basant sur les concepts flous et d'inférer les actions de commande en employant l'implication floue et les règles d'inférence en logique floue.
- 4) Interface de défuzzification :
- L'interface de défuzzification réalise les fonctions suivantes :
- a) Une cartographie d'échelle qui convertie la plage des valeurs des variables de sortie aux univers de discours correspondant,
 - b) La défuzzification, qui fournit une action de commande non flous (physique) à partir d'une action inférée de commande floue.
- Nous pouvons maintenant faire apparaître les principaux paramètres de conception pour un contrôleur flou à savoir :
- 1- Les stratégie de fuzzification
 - 2- La base de données :
 - a) Discrétisation/normalisation des univers de discours
 - b) Partition floue des espaces d'entrée et de sortie,
 - c) complétude,
 - d) choix des fonctions d'appartenance.
 - 3- la base de règles :
 - a) choix des variables d'états (d'entrée) du processus et des variables de commande (de sortie) du contrôleur flou,
 - b) source et dérivation des règles de contrôle flou,
 - c) types de règles de contrôle flou,
 - d) consistance, interactivité, complétude des règles de contrôle flou.
 - 4- la logique de prise de décision :
 - a) définition d'une implication floue,
 - b) interprétation du connecteur de phrases ET,
 - c) interprétation du connecteur de phrases AUSSI,
 - d) définition d'un opérateur de composition des règles,
 - e) mécanisme d'inférence.
 - 5- les stratégies de défuzzification.

Dans ce qui suit, chacune de ces parties fera l'objet d'un développement détaillé.

2.1.1 Stratégies de fuzzification [1] [10] [11] [13] :

La fuzzification est liée au caractère vague et à l'imprécision dans un langage naturel. C'est une évaluation subjective qui transforme une mesure en une évaluation d'une valeur subjective, et donc, elle peut être définie comme une cartographie d'un espace d'entrée observé vers des ensembles flous en certains univers de discours d'entrée. La fuzzification joue un rôle important en relation avec l'information incertaine qui peut être objective ou subjective.

Dans les applications de commande floue, les données observées sont habituellement physique (réelles numériques). Or, le traitement de données dans un contrôleur flou est basé sur la théorie des ensembles flous ; ceci nécessite donc une fuzzification.

La fuzzification s'articule autour des axes suivants :

- 1) Un opérateur de fuzzification convertit une valeur numérique en un singleton flou à l'intérieur d'un certain univers de discours.
- 2) Les données observées sont affectées par un bruit aléatoire. Dans ce cas, un opérateur de fuzzification devrait convertir ces données probabilistes en nombres flous.

2.1.2 Base de données

La base de connaissances d'un contrôleur flou est composée de deux parties : une base de données et une base de règles de contrôle. Cette dernière fera l'objet de la prochaine section. Les concepts associés à la base de données sont utilisés pour caractériser les règles de contrôle flou et la manipulation des données floues dans un contrôleur flou. Ces concepts sont subjectivement définis et basés sur l'expérience et le jugement de l'expert en engineering. Dans cet ordre d'idée, il peut être noté que le choix correct des fonctions d'appartenance d'un ensemble de termes joue un rôle essentiel dans la réussite d'une application. On va donc chercher à déterminer les phases importantes de construction d'une base de données dans un contrôleur flou.

2.1.3 Discrétisation/normalisation des univers de discours

La représentation d'une information incertaine par des ensembles flous pose le problème de quantification d'une telle information pour la traiter sur calculateur. En général, cette représentation dépend de la nature de l'univers de discours. Dans un contrôleur flou, un univers de discours peut être continu ou discret. Si l'univers est continu, on peut toujours le discrétiser. De plus, un univers continu peut être normalisé comme on le verra par la suite.

a) discrétisation d'un univers de discours

La discrétisation d'un univers de discours est fréquemment associée à une quantification. En effet, la quantification discrétise un univers en un certain nombre de segments (niveaux de quantification). Chaque segment est étiqueté comme un élément générique et forme un univers discret. Un ensemble flou est alors défini par le degré d'appartenance attribués à chaque élément générique de ce nouvel univers de discours discret. Basée sur les univers discrets, une table de décision, définissant la sortie du contrôleur flou pour toutes les combinaisons possibles des signaux. Dans le cas des univers continus, le nombre des niveaux de quantification doit être assez grand pour fournir une approximation adéquate mais tout de même faible pour limiter

les besoins en mémoire. Le choix des niveaux a une importante influence sur la manière d'obtenir un contrôle fin (commande fine). Par exemple, si un univers est quantifié toutes les cinq unités de mesure au lieu de dix unités, alors le contrôleur est deux fois plus sensible aux variables observées.

La discrétisation nécessite une cartographie d'échelle transformant les variables mesurées en valeurs dans l'univers de discours. La cartographie peut être uniforme (linéaire) ou non uniforme (non linéaire) ou les deux en même temps. Le choix des niveaux de quantification reflète une connaissance a priori.

Un exemple de discrétisation est montré dans la table [T.1] où un univers de discours est discrétisé en treize niveaux de quantification et sept termes linguistiques (ensembles flous) y sont définis. En général, à cause de la discrétisation, la performance d'un contrôleur flou est moins sensible aux petites variations des valeurs des variables d'état du processus.

Table T.1 : quantification et ensembles flous utilisant une définition numérique

No niveau De Quantification	Plage de variation	Termes linguistiques						
		NG	NM	NP	EZ	PP	PM	PG
-6	$x_0 \leq -3.2$	1.0	0.3	0.0	0.0	0.0	0.0	0.0
-5	$-3.2 < x_0 \leq -1.6$	0.7	0.7	0.0	0.0	0.0	0.0	0.0
-4	$-1.6 < x_0 \leq -0.8$	0.3	1.0	0.3	0.0	0.0	0.0	0.0
-3	$-0.8 < x_0 \leq -0.4$	0.0	0.7	0.7	0.0	0.0	0.0	0.0
-2	$-0.4 < x_0 \leq -0.2$	0.0	0.3	1.0	0.3	0.0	0.0	0.0
-1	$-0.2 < x_0 \leq -0.1$	0.0	0.0	0.7	0.7	0.0	0.0	0.0
0	$-0.1 < x_0 \leq 0.1$	0.0	0.0	0.3	1.0	0.3	0.0	0.0
1	$0.1 < x_0 \leq 0.2$	0.0	0.0	0.0	0.7	0.7	0.0	0.0
2	$0.2 < x_0 \leq 0.4$	0.0	0.0	0.0	0.3	1.0	0.3	0.0
3	$0.4 < x_0 \leq 0.8$	0.0	0.0	0.0	0.0	0.7	0.7	0.0
4	$0.8 < x_0 \leq 1.6$	0.0	0.0	0.0	0.0	0.3	1.0	0.3
5	$1.6 < x_0 \leq 3.2$	0.0	0.0	0.0	0.0	0.0	0.7	0.7
6	$x_0 \geq 3.2$	0.0	0.0	0.0	0.0	0.0	0.3	1.0

B) Normalisation d'un univers de discours

La normalisation d'un univers de discours nécessite une discrétisation de cet univers en nombre fini de segments ayant chacun un segment correspondant dans l'univers normalisé. Dans ce cadre, un ensemble flou est alors défini explicitement par sa fonction d'appartenance. La normalisation d'un univers de discours repose aussi sur une connaissance a priori de l'espace entrée/sortie. La cartographie de l'échelle peut être uniforme ou non ou les deux. Un exemple est montré dans la table T.2, où l'univers de discours est $[-10 +10]$ est transformé en l'intervalle normalisé $[-1 +1]$. Cette normalisation introduit une grande flexibilité lors de l'implémentation d'un contrôleur flou.

Table T.2 : normalisation et ensembles flous utilisant une définition analytique

Univers normalisé	Intervalle normalisé	Intervalle nominale	Ensemble flous
[-1,+1]	[-1.000, -0.500]	[-10, -5]	NG
	[-0.500, -0.250]	[-5, -2.5]	NM
	[-0.250, 0.125]	[-2.5, -1.5]	NP
	[-0.125, +0.125]	[-1.5, +1.5]	EZ
	[+0.125, +0.250]	[+1.5, +2.5]	PP
	[+0.250, +0.500]	[+2.5, +5]	PM
	[+0.500, +1.000]	[+5,+10]	PG

2.1.4 Partition des espaces d'entrée et de sortie

Dans l'antécédent d'une règle de contrôle flou, une variable linguistique forme un espace d'entrée flou en respect à un certain univers de discours, tandis que celle dans la conséquence de la règle forme un espace de sorti flou. En général, une variable linguistique est associée à un ensemble de termes dont chacun est défini sur le même univers. Une partition floue détermine alors le nombre de termes, ceci est équivalent à la détermination du nombre d'ensembles flous. Ce nombre constitue la granularité de la commande réalisable par un contrôleur flou. Un exemple typique est montré dans la figure [2.2] décrivant deux partitions floues dans le même univers normalisé [-1 +1]. Des fonctions triangulaire et trapézoïdale sont utilisées ici.

Puisqu'un univers de discours implique la connaissance de l'espace entrée/ sortie via des cartographies d'échelle appropriées, un ensemble de termes bien formé peut être réalisé. Si ce n'est pas le cas, ou un univers non normalisé est utilisé, les termes pourraient être asymétriques et distribuées de façon non uniforme dans l'univers. En outre, la cardinalité d'un ensemble de termes dans un espace d'entrée flou détermine le nombre maximal de règles de contrôle flou qu'on peut construire. Dans le cas des systèmes flous à deux entrées et une sorite, si les cardinalités de $T(x)$ et $T(y)$ sont respectivement n et m , le nombre de règles est de $n \times m$. il peut être noté que la partition floue de l'espace d'entrée/sortie flou n'est pas déterministe et n'a pas de solution unique. Une procédure heuristique d'essai/erreur est habituellement nécessaire pour trouver la partition floue optimale.

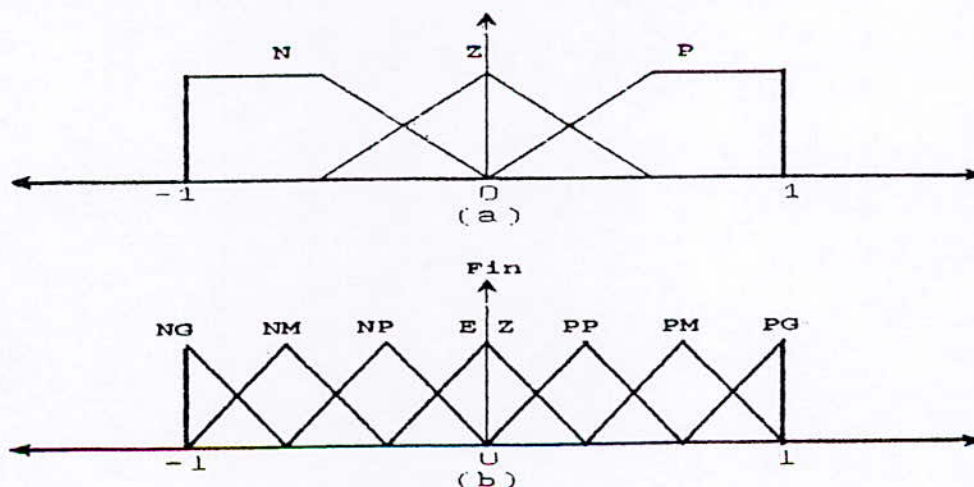


Figure 2.2 : deux partition floues dans le même univers de discours

2.1.5 Complétude

Intuitivement, un algorithme de contrôle flou devrait être toujours capable d'inférer une action de commande propre à tout état du processus. Cette propriété est appelée "complétude".

La complétude pour un contrôleur flou est relative à sa base de données, sa base de règles ou les deux.

a) *stratégie de la base de données*

La stratégie de la base de données concerne les supports sur lesquels les ensembles flous sont définis. L'union de ces supports devrait couvrir l'univers de discours relatif avec un degré ε . Cette propriété, dans un contrôleur flou, est appelée ε -complétude. En général, on choisit le degré ε au point de chevauchement ou commutation, impliquant que nous avons une forte croyance en la positivité du sens des règles du contrôle flou qui sont associées au contrôleur flou. Dans ce sens, une règle dominante existe toujours et est associée à un degré de croyance supérieur à 0.5. Dans le cas extrême, deux règles dominantes sont activées avec un degré commun.

b) *Stratégie de la base de règles*

La stratégie de la base de règles traite des règles de contrôle flou elles-mêmes. La propriété de complétude est incorporée dans ces règles selon l'expérience de conception et les connaissances d'experts en engineering. Une règle additionnelle est ajoutée si jamais une condition floue n'est pas incluse dans la base de règles ou si jamais le degré d'harmonie partielle entre quelques entrées et les conditions floues prédéfinis est inférieur à un certain niveau, par exemple 0.5, le premier cas montre qu'aucune action de commande ne résultera. Le dernier cas indique qu'aucune règle dominante ne sera activée.

2.1.6 Fonction d'appartenance d'un ensemble flou

Pour la définition des ensembles flous, il y a deux méthodes dépendant de ce que l'univers de discours est discret ou continu : numérique ou fonctionnelle (analytique).

a) *Définition numérique :*

dans ce cas, le degré de la fonction d'appartenance d'un ensemble flou est représenté comme un vecteur de nombre dont la dimension dépend du degré de discrétisation. Un exemple illustratif est montré dans la table T.3. dans ce cas, la fonction d'appartenance de chaque ensemble flou a la forme de

$$\mu_A(u) = \sum_{i=1}^5 \frac{a_i}{u_i} \quad \text{où } a_i = [0.3, 0.7, 1.0, 0.7, 0.3] \quad (2.1)$$

c) *Définition fonctionnelle (analytique) :*

Une approche analytique exprime la fonction d'appartenance d'un ensemble flou comme une forme fonctionnelle typique : en forme de cloche (gaussienne), de forme triangulaire, de forme trapézoïdale....

De telles fonctions sont utilisées dans les contrôleurs flous car elles mènent elles-mêmes à une manipulation selon une arithmétique floue. La définition fonctionnelle peut être adaptée à volonté à un changement dans la normalisation de l'univers. La table T.3 et la figure [2.3] montrent un exemple d'une définition analytique exprimée par :

$$\mu_A(x) = \exp\left\{\frac{-(x - \mu_a)^2}{2\sigma_a^2}\right\} \quad (2.2)$$

[T.3]

Univers normalisé	Variation de x, segments normalisés	Plage	Moyenne de x	Variance de x	Ensemble flous primaires
[-1, +1]	[-1,-0.6]	[-3.60,-2.16]	-1.0	0.4	NG
	[-1,-0.3]	[-3.60,-1.08]	-0.5	0.2	NM
	[-0.6,0.0]	[-2.16,-2.20]	-0.2	0.2	NP
	[-0.3,0.3]	[-2.20,-0.00]	0.0	0.2	EZ
	[0.0,0.6]	[+0.00,+1.00]	+0.2	0.2	PP
	[0.3,1.0]	[+1.00,+2.50]	+0.5	0.2	PM
	[0.6,1.0]	[+2.50,+4.50]	+1.0	0.4	PG

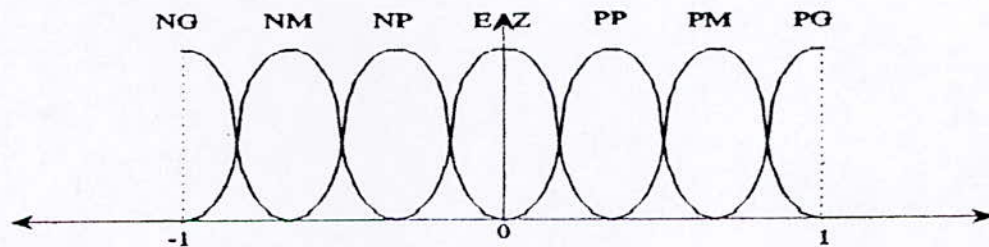


Figure 2.3 : Exemple de définition analytique d'ensembles flous

2.2 Base de règles [10], [11], [23]

Un système flou est caractérisé par un ensemble de déclarations ou d'expressions linguistiques basées sur une connaissance d'expert. Cette connaissance est habituellement sous la forme de règles "Si -Alors" qui sont faciles à implémenter par des expressions conditionnelles floues. La collection de ces règles de contrôle flou forme la base de règles ou l'ensemble de règles d'un contrôleur flou.

Dans cette partie, nous allons examiner les points suivants relatifs aux règles de contrôle flou :

- le choix des variables d'état (entrées) du processus et des variables de commande,
- la source et la dérivation des règles de contrôle flou,
- la justification des règles de contrôle flou,
- les types des règles de contrôle,
- les propriétés de : consistance, interactivité et complétude.

a) Choix des variables d'état du processus et des variables de contrôle des règles de contrôle flou

Le choix approprié des variables d'état du processus et des variables de contrôle est essentiel pour caractériser un système flou. De plus, la sélection des variables linguistiques a un effet important sur les performances d'un contrôleur flou. Durant cette étape de sélection, l'expérience et la connaissance de l'ingénieur jouent un rôle important. En particulier, les choix des variables linguistiques et leurs fonctions d'appartenance ont une grande influence sur la structure linguistique du contrôleur flou.

b) Source de dérivation des règles de contrôle flou

Il y a quatre modes de dérivation des règles de contrôle flou. Ces quatre modes ne sont pas mutuellement exclusifs, et il semble probable que leur combinaisons serait nécessaire pour la construction d'une méthode efficace pour la dérivation des règles de contrôle.

b.1) Expérience d'expert et connaissance d'ingénieur

Les règles de contrôle flou ont la forme d'expressions conditionnelles reliant les variables d'état comme antécédents et les variables de contrôle comme conséquences.

Dans la vie quotidienne, la plupart des informations sur lesquelles se basent nos décisions sont plutôt linguistiques que numériques dans leur nature. Vues, dans cette perspective, les règles de contrôle flou apportent un outil naturel pour la caractérisation du comportement humain et l'analyse décisionnelle.

b.2) Modélisation des actions de contrôle d'un opérateur

Dans plusieurs systèmes homme-machine de contrôle, les relations entrée/sortie ne sont pas connues avec une précision suffisante pour pouvoir employer la théorie de la commande classique pour la modélisation et la simulation. Or des opérateurs qualifiés peuvent tout à fait contrôler de tels systèmes avec succès sans avoir aucun modèle quantitatif en vue. En fait, un opérateur humain utilise consciemment ou subconsciemment un ensemble de règles "Si-Alors" floues pour contrôler le processus.

Il est avantageux d'exprimer les règles de contrôle de l'opérateur sous la forme de règles "Si-Alors" floues en employant des variables linguistiques. En pratique de telles règles peuvent être déduites à partir de l'observation des actions du contrôleur humain en termes de données de fonctionnement entrée/sortie.

c) *Modélisation floue d'un processus*

Dans l'approche linguistique, la description linguistique des caractéristiques dynamiques d'un processus contrôlé peut être vue comme un modèle flou de ce dernier. Sur la base de ce modèle, on peut générer un ensemble de règles de contrôle flou pour atteindre des performances optimales d'un système dynamique. Cet ensemble constitue la base de règles d'un contrôleur flou.

d) *Apprentissage et auto-organisation*

Plusieurs contrôleurs flous ont été conçus pour émuler le comportement humain de prise de décision mais peu sont basés sur l'apprentissage humain, autrement dit, l'aptitude de créer des règles de contrôle flou et de les modifier en se basant sur l'expérience.

2.2.1 Justification des règles de contrôle flou

Il y a deux approches principales pour la dérivation des règles de contrôle flou. La première est une méthode heuristique dans laquelle une collection de règles de contrôle flou est formée par l'analyse du comportement du processus à contrôler. Ces règles de contrôle sont établies de telle sorte que l'écart par rapport à l'état désiré puisse être corrigé. L'établissement de ces règles est purement heuristique en nature et repose sur la connaissance qualitative du comportement du processus. Plusieurs méthodes ont été développées pour l'ajustement de la sélection des règles. La seconde approche est fondamentalement une méthode déterministe qui peut systématiquement déterminer la structure linguistique et/ou les paramètres des règles flou satisfaisants les objectifs et les contraintes du contrôle.

2.2.2 Types de règles de contrôle flou

Selon leur nature, deux types de règles de contrôle flou, un pour l'évaluation d'état et l'autre pour l'évaluation de l'objectif, sont couramment utilisés dans la conception des contrôleurs à logique floue.

a) *Règle de contrôle flou pour l'évaluation d'état*

La plupart des contrôleurs flous ont des règles de contrôle flou pour l'évaluation d'état qui, dans le cas des collection de règles de la forme :

R_1 : Si x est A_1, \dots , et y est B_1 Alors z est C_1

R_2 : Si x est A_2, \dots , et y est B_2 Alors z est C_2

.....

R_n : Si x est A_n, \dots , et y est B_n Alors z est C_n

Où :- x, \dots, y et z sont des variables linguistiques représentant les variables d'état du processus et la variable de commande.

- A_i, \dots, B_i et C_i sont des termes linguistiques des variables linguistiques x, \dots, y et z dans les univers de discours U, \dots, V et W respectivement avec $i=1, 2, \dots, n$.

D'une façon générale, la conséquence est représentée comme une fonction des variables d'état x, \dots, y du processus :

R_i : si x est A_i, \dots , et y est B_i Alors $z = f_i(x, \dots, y)$

b) Règle de contrôle pour l'évaluation de l'objectif

Yasunobu, Miyamoto et Ihra [18] ont proposé un autre algorithme qui prédit les actions de commande présentes et futures et évalue les objectifs de commande. Il est appelé "commande floue pour l'évaluation de l'objectif" ou "commande floue prédictive". Les règles en question, qui sont obtenues à partir de l'expérience d'un opérateur habile et expérimenté, une règle typique est décrite comme :

R_i : Si (u est $C_i \rightarrow (x \text{ est } A_i \text{ et } y \text{ est } B_i))$ alors u est C_i

Une relation de commande est déduite à partir d'une évaluation d'objectif du résultat d'une commande floue, satisfaisant les états et les objectifs désirés. une commande u prend une valeur réelle et x et y sont des indices de performance pour l'évaluation de la $i^{\text{ème}}$ règle en prenant des valeurs comme "bonne" ou "mauvaise". La règle de commande la plus adaptée est sélectionnée à travers la prédiction des résultats (x,y) correspondant pour chaque commande C_i .

En termes linguistiques, la règles est interprétée comme :

" Si l'indice de performance x est A_i et l'indice y est B_i quand une commande u est choisie égale à C_i . Alors cette règle est sélectionnée et la commande C_i est prise pour être la sortie du contrôleur".

Il est bien connu que la commande des systèmes rencontre des difficultés pour satisfaire des indices de performance multiples simultanément et pour réaliser un contrôle précis en présence de perturbations. Dans de telles circonstances, la commande floue fournit un outil effectif comme solution. Cependant, la commande floue d'évaluation d'état n'évalue pas les actions de commande floue calculées de la même manière qu'un opérateur humain. Par contre, la commande floue prédictive fournit un mécanisme d'évaluation pour que les états désirés et les objectifs de contrôle puissent être atteints plus facilement. Ce type de contrôle a été utilisé avec succès sur des trains et des grues.

2.2.3 Propriétés de consistance, interactivité et complétude [18]

a) Complétude

Ce thème a été déjà exposé et le lecteur est prié de se référer au paragraphe 2.1.5

b) Nombre de règle de contrôle flou

Pour cette première partie de notre travail nous pouvons insinué qu'il n'y a aucune procédure générale pour décider du nombre optimal de règles de contrôle flou, vu qu'un certain nombre de facteurs sont impliqués dans la décision tels que : la performance du contrôleurs, l'efficacité de calcul, le comportement de l'opérateur humain et le choix des variables linguistiques. Chose qui sera démentit par la suite dans notre seconde partie de travail.

c) Consistance des règles de contrôle flou

Si la dérivation des règles de contrôle flou est basée sur l'expérience de l'opérateur humain, les règles peuvent être soumises à différents critères de performance. Dans la pratique, il est important de vérifier la consistance des règles de contrôle flou pour minimiser la possibilité de contradiction.

d) Interactivité des règles de contrôle flou

En supposant qu'une collection de règles de contrôle flou a la forme suivante :

R_i : Si x est A_i Alors z est C_i , $i=1,2,\dots,n$

Si une entrée x_0 est A_i , on s'attendrait que l'action de contrôle z soit C_i . en fait, l'action de contrôle z peut être un sous ensemble ou un ensemble tout entier de C_i , dépendant de la définition de l'implication floue et de la composition "sup-star". Ceci peut se produire comme conséquence de l'interaction entre les règles. Le problème d'interaction est complexe et non encore bien compris. Les recherches effectuées indiquent que l'inactivité des règles peut être contrôlée par le choix de l'implication floue et la composition "sup-star". La consistance des règles peut être améliorée à travers l'utilisation du concept de noyau de règles de contrôle flou. Dans cet ordre d'idée, il est à noter que l'algorithme de raisonnement et d'identification de Sugeno apportent une solution alternative à ces problèmes.

2.3 Logique de prise de décision [16], [17], [23], [26]

Comme il a déjà été signalé, un contrôleur flou peut être vu comme un moyen d'émulation d'un opérateur humain habile. Plus généralement, l'utilisation d'un contrôleur à logique floue peut être considérée comme un pas dans la direction de la modélisation de la prise de décision humaine avec l'aide de la logique floue et du raisonnement approximatif.

2.3.1 Famille d'implication floue

En général, une règle de contrôle flou est une relation exprimée sous forme d'implication floue. En logique floue il y a plusieurs voies dans lesquelles une implication floue peut être définie. La définition d'une implication floue peut être exprimée comme une fonction d'implication floue.

a) Familles de fonctions d'implication floue

A la suite de l'introduction de la règle compositionnelle d'inférence en raisonnement approximatif par Zadeh [36, 37], un nombre de chercheurs ont proposé différentes sorte d'implications floues dans lesquelles les antécédents et les conséquences contiennent des variables floues. En effet, près de 40 fonctions d'implications floues distinctes ont été décrites dans la littérature.

Plusieurs fonctions d'implications floues peuvent être générées en employant les "normes" et les "co-normes" triangulaires. Les implications floues qui sont souvent adoptées dans un contrôleur flou ont été données dans le chapitre précédent.

b) Choix des fonction d'implication floue

Pour choisir une fonction d'implication, un ensemble de critères intuitifs construisant les relations entre les antécédents et les conséquences d'une proposition conditionnelle a été proposée.

Comme il a été signalé précédemment (1.7.2), il y a deux règles d'inférence floue :

- Modus Ponens Généralisé (MPG)
- Modus Tollens Généralisé (MTG)

Dans les applications de commande à logique floue, une action de contrôle est déterminée par les entrées observées et les règles de commande, sans que la

conséquence d'une règle soit l'antécédent d'une autre. En effet, les fonctions du contrôleur à logique floue ont un seul niveau de chaînage avant (MPG).

2.3.2 Interprétation des connecteurs de phrases "AUSSE" et "ET"

Dans la plupart des contrôleurs flous existants, le connecteur de phrases " ET" est d'habitude implémenté comme une conjonction floue dans un espace de produit cartésien dans lequel les variables prennent des valeurs dans des univers de discours différents.

Comme illustration, dans "Si (A et B) Alors C", l'antécédent est interprété comme un ensemble flou dans l'espace produit $U \times V$ avec la fonction d'appartenance suivante :

$$\mu_{A \times B}(u, v) = \min\{\mu_A(u), \mu_B(v)\} \text{ ou } \mu_{A \times B}(u, v) = \mu_A(u) \times \mu_B(v) \quad (2.3)$$

Où U et V sont les univers de discours associés respectivement à A et B .

Quand un système flou est caractérisé par un ensemble de règles de contrôle flou, l'ordre des règles est sans importance. Ceci implique que le connecteur de phrases "AUSSE" doit avoir les propriétés de commutativité et d'associativité.

2.3.3 Opérateurs compositionnels et mécanismes d'inférence

La stratégie de réglage dépend essentiellement des grandeurs mesurées qui sont les variables d'entrée (transformées en variables linguistiques par fuzzification), à la variable de sortie ou de commande qui est également exprimée comme variable linguistique. Dans les inférences du contrôleur flou interviennent les opérateurs ET et OU. L'opérateur ET s'applique aux variables à l'intérieur d'une règle, tandis que l'opérateur OU lie les différentes règles. A cause du chevauchement des fonctions d'appartenance, en général deux ou plusieurs règles sont activées en même temps. Ce fait doit être pris en considération lors de la réalisation de l'opérateur OU.

Une règle d'inférence est activée lorsque le degré d'appartenance lié à la condition de cette règle est non nul. Il existe plusieurs possibilités pour réaliser ces opérateurs qui s'appliquent aux fonctions d'appartenance. on introduit alors la notion de méthode ou mécanisme d'inférence. Elle détermine la réalisation des différents opérateurs dans une inférence, permettant ainsi un traitement numérique de cette dernière.

En commande floue, on utilise en général une des méthodes suivantes

- méthode d'inférence *maximum-minimum* (MAX-MIN)
- méthode d'inférence *maximum-produit* (MAX-PROD)
- méthode d'inférence *somme-produit* (SOM-PROD)

Elles permettent un traitement numérique des inférences, à partir de certains valeurs (degrés d'appartenance) pour les fonctions d'appartenance des variables d'entrée (données par fuzzification), nous obtenons l'allure de la fonction d'appartenance de la variable de sortie.

Ces trois méthodes d'inférences seront décrites aux paragraphes suivants. Afin de mettre en évidence le traitement numérique, nous ferons appel à un exemple à deux variables d'entrée x_1 et x_2 et une variable de sortie x_r . Chacune est décomposée en trois ensembles NG, EZ, PG et définie par des fonctions d'appartenance. Pour les variables d'entrée prenons comme valeurs numériques : $x_1=0.44$ et $x_2=-0.67$.

L'inférence est composée de deux règles :

R : Si (x_1 est PG ET x_2 est EZ), Alors x_r est EZ ou

(x_1 est EZ ou x_2 est NG), Alors x_r est NG

a) Méthode d'inférence MAX-MIN

La méthode d'inférence *max-min* réalise, au niveau de la condition, l'opérateur OU par la formation du maximum et l'opérateur ET par la formation du minimum. La conséquence dans chaque règle, introduit par Alors, lie le degré d'appartenance de la condition avec la fonction d'appartenance de la variable de sortie par l'opérateur ET, réalisé par la formation du maximum. Enfin, l'opérateur OU qui lie les différentes règles est réalisé par la formation du maximum.

MAX-MIN : OU \Rightarrow MAXIMUM
Alors \Rightarrow MINIMUM

La figure 2.4 représente graphiquement le principe de la méthode d'inférence max-min.

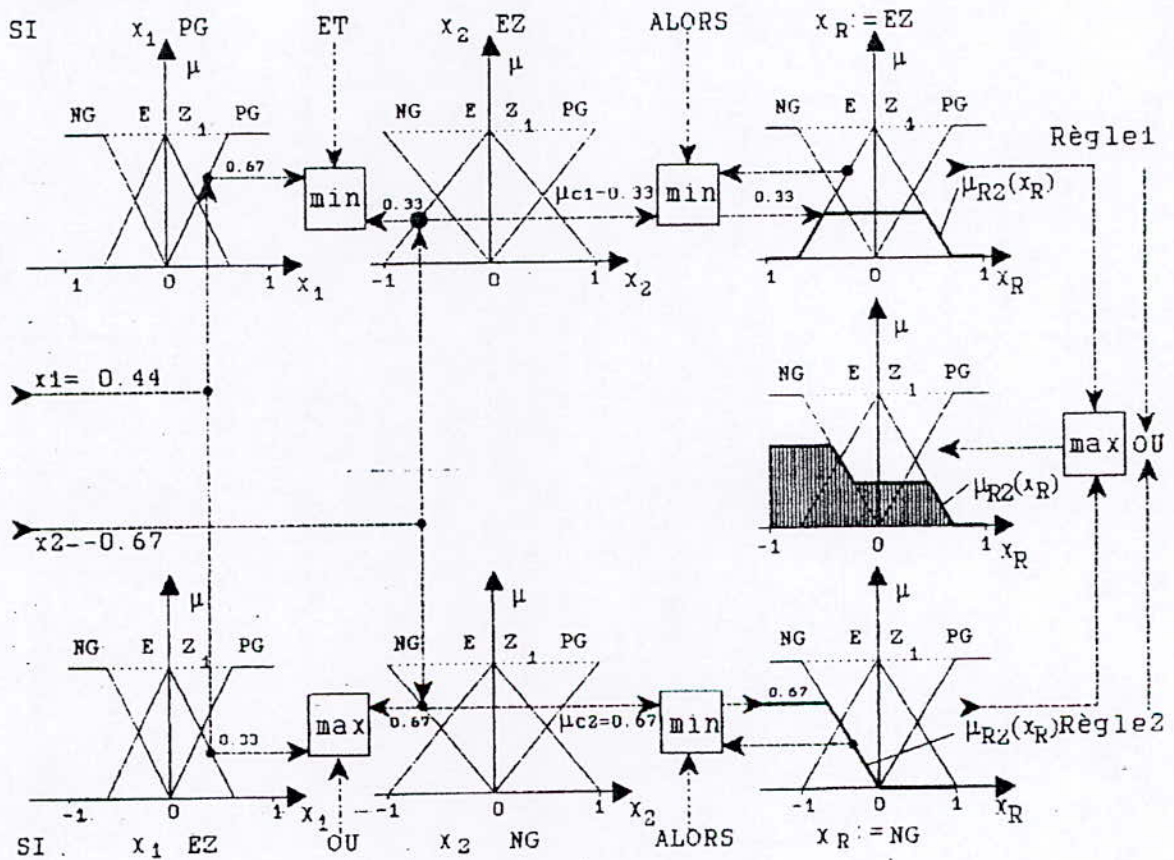


Figure 2.4 : Méthode d'inférence MAX-MIN

b) Méthode d'inférence MAX-PROD

Cette méthode d'inférence réalise en général au niveau de la condition, l'opérateur OU par la formation du maximum et l'opérateur ET par la formation du minimum. En revanche, la conclusion dans chaque règle, introduite par Alors, liant le

degré d'appartenance de la condition à la fonction d'appartenance de la variable de sortie par l'opérateur ET, est réalisée par la formation du produit. L'opérateur OU qui lie les différentes règles est réalisé par la formation du maximum.

MAX-PROD : OU \Rightarrow MAXIMUM
ALORS \Rightarrow PRODUIT

La figure 2.5 représente graphiquement le principe de la méthode d'inférence *max-prod*.

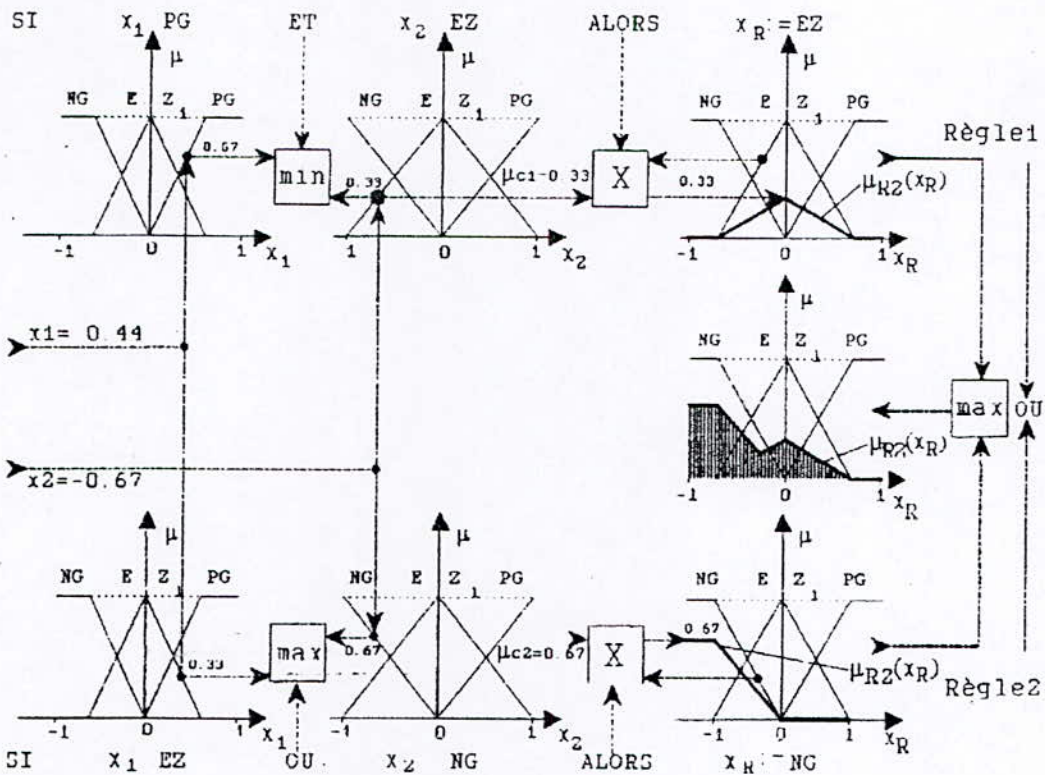


Figure 2.5 : Méthode d'inférence MAX-PROD

c) Méthode d'inférence SOM-PROD

Par opposition aux méthodes d'inférence précédentes, la méthode d'inférence *somme-prod* réalise, au niveau de la condition, l'opérateur OU par la formation de la somme, plus précisément par la valeur moyenne, tandis que l'opérateur est réalisé par la formation du produit. La conclusion de chaque règle, précédée par ALORS, liant le degré d'appartenance de la condition avec la fonction d'appartenance de la variable de sortie par l'opérateur ET, est réalisée par la formation du produit. L'opérateur OU qui lie les différentes règles est réalisé par la formation de la somme, donc de la valeur moyenne [10], [11], [18]

SOMME-PROD : OU \Rightarrow SOMME (MOYENNE)
ALORS \Rightarrow PRODUIT

La figure 2.6 représente graphiquement le principe de la méthode d'inférence *somme-prod*.

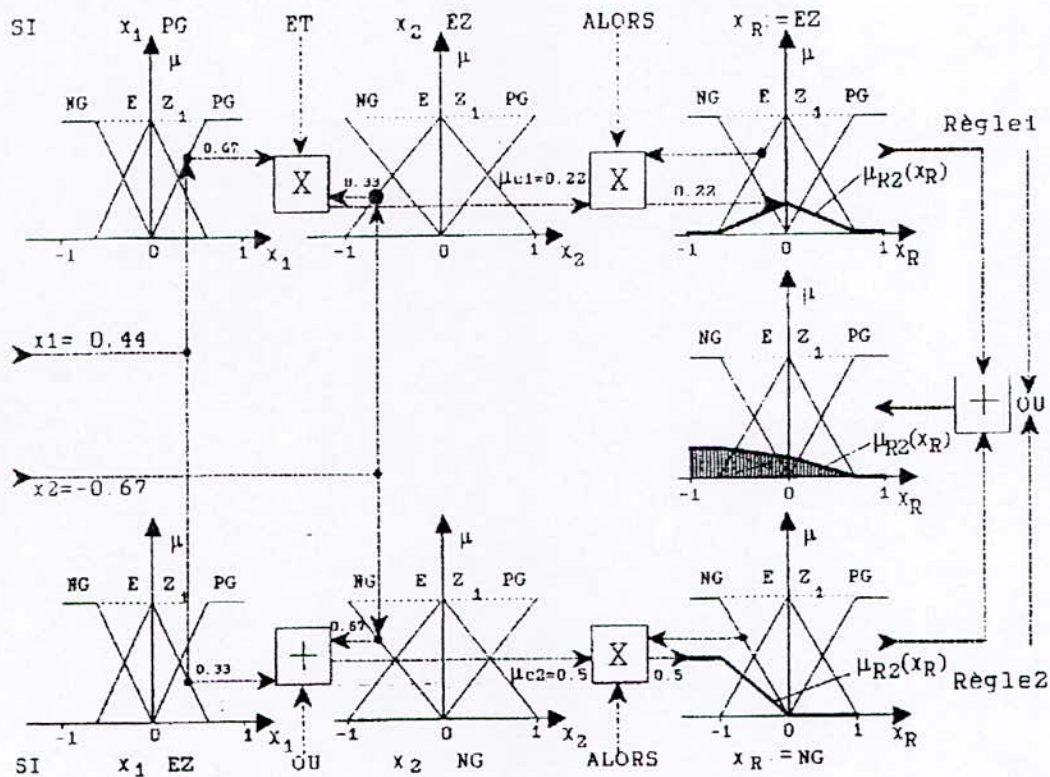


Figure 2.6 : méthode d'inférence SOM-PROD

2.4 Stratégie de défuzzification

Comme nous l'avons précédemment vu, les méthodes d'inférences fournissent une fonction d'appartenance résultante $\mu_{res}(x_r)$ pour la variable de sortie x_r , il s'agit donc d'une information floue. Etant donné que l'organe de commande nécessite un signal de commande u précis à son entrée. Il faut prévoir une transformation de cette information floue en une information déterminée. Cette transformation est appelée défuzzification.

Plusieurs stratégies de défuzzification existent [1], [2], [11], [13], [14]. Les plus utilisées sont :

- méthode du maximum,
- méthode de la moyenne des maxima,
- méthode du centre de gravité.

a) Méthode du maximum

Pour cette méthode, la commande est égale à la commande locale au point où la distribution de la fonction d'appartenance relative à l'action de contrôle est maximale.

b) Méthode de la moyenne des maxima

Cette stratégie génère une commande qui représente la valeur moyenne de toutes les actions de contrôle locales correspondant aux points ayant des fonctions d'appartenance maximales. Plus spécialement, dans le cas d'un univers de discours discret, l'action de commande peut être exprimée par :

$$u_0 = \sum_{i=1}^l \frac{w_i}{l} \quad (2.4)$$

Où w_i est la valeur de la commande locale pour laquelle la fonction d'appartenance atteignent la valeur maximale $\mu_u(w_i)$ et l est le nombre total des valeurs prises par la fonction d'appartenance.

c) Méthode du centre de gravité

Cette méthode est la plus utilisée dans les contrôleurs flous. Elle génère le centre de gravité de l'espace flou délimité par la distribution de la fonction d'appartenance résultante représentant l'action de commande. Dans le cas d'un univers discret, cette méthode donne :

$$u_0 = \frac{\sum_{i=1}^n \mu_u(w_i) \times w_i \times S_i}{\sum_{i=1}^n \mu_u(w_i) \times S_i} \quad (2.5)$$

Où n est le nombre de niveaux de quantification de la sortie du contrôleur et S_i est la surface de la $i^{\text{ème}}$ fonction d'appartenance de la variable de sortie. Dans le cas où ces fonctions d'appartenance sont de la forme symétrique et de distribution uniforme l'équation (2.5) devient :

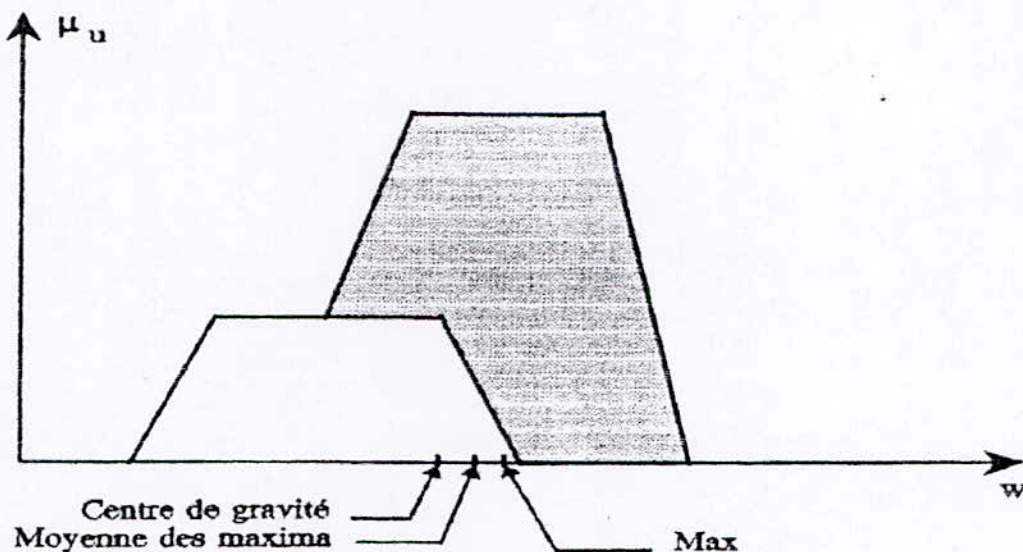
$$u_0 = \frac{\sum_{i=1}^n \mu_u(w_i) \times w_i}{\sum_{i=1}^n \mu_u(w_i)} \quad (2.6)$$

La figure [2.7] montre une interprétation graphique de ces différentes stratégies de défuzzification.

Dans [1], [10], [18], des analyses détaillées de plusieurs stratégies de défuzzification sont présentées et il en ressort que la méthode de centre de gravité donne des résultats supérieurs. Cependant, la méthode de la moyenne des maxima fournit une meilleure performance transitoire alors que celle du centre de gravité est plus performante en régime permanent.

A noter aussi que lorsque la méthode de la moyenne des maxima est utilisée, la performance du contrôleur flou est similaire à celle d'un système à relais et à plusieurs niveaux. Tandis que la méthode du centre de gravité produit des résultats similaires à ceux d'un PI conventionnel.

Un contrôleur flou basé sur la méthode du centre de gravité produit donc généralement une erreur plus faible qu'avec la méthode de la moyenne des maxima.



Figure[2.7] : Interprétation graphique des différentes méthodes de défuzzification

Remarque : Un problème posé par la méthode du maximum est le choix de la commande nécessaire quand plusieurs éléments atteignent cette valeur maximale(figure [2.8.a]). Choisir une valeur au hasard est bien sûr à éviter, la solution serait de

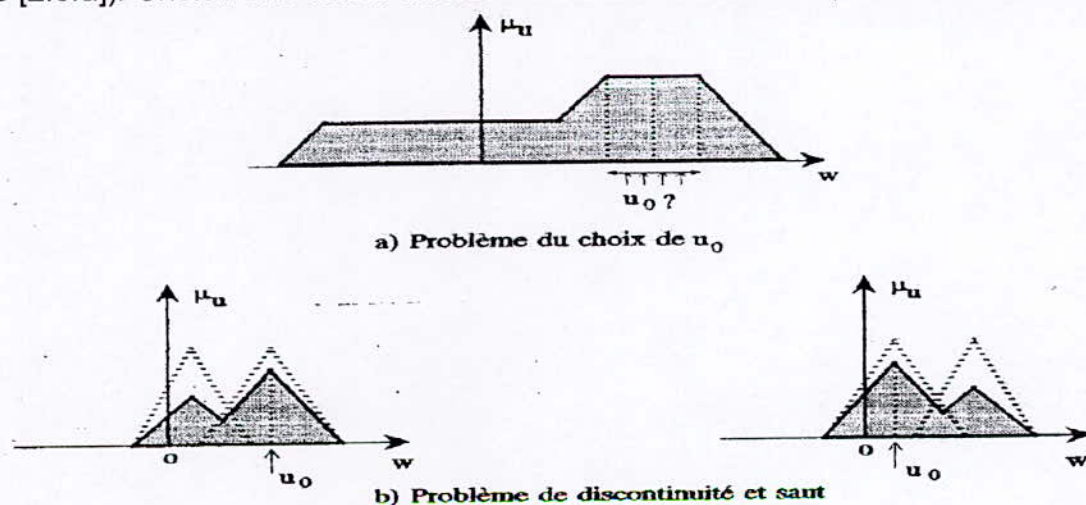


Figure2.8 : Problèmes posés par la méthode de défuzzification du maximum

Prendre la moyenne des maxima. Il y a un second inconvénient qui entâche cette méthode ; la valeur du signal de commande résultant saute si la dominance change d'une fonction d'appartenance partielle à une autre ce qui provoque des à-coups sur le signal de commande et conduit à un mauvais comportement d'un circuit de réglage (figure [2.8.b]). Par conséquent, cette méthode de défuzzification n'est pas recommandée.

Conclusion :

Ce chapitre présente une méthodologie de base du contrôle flou ainsi que les différentes stratégies utilisées en raisonnement par les systèmes flous, partant des mécanismes de fuzzification où s'effectue la transformation données propres (analytiques) en fonctions d'appartenance floues, après lequel on est confronté à un mécanisme d'inférence qui fait la déduction floue (raisonnement flou) puis l'opération inverse de la fuzzification dite défuzzification. Sans oublier les conditions sur la base de données (normalisations, dénormalisation, discrétisation, etc....).

CHAPITRE 3

Modélisation Floue

Introduction :

Les systèmes flous sont des approximateurs universels qui servent à la simulation et commande des processus réels. Ils sont utiles en particulier quand les techniques classiques de modélisation font défaut. C'est à dire, quand le processus est non linéaire ou trop complexe par le nombre de paramètres à considérer. Mais la phase d'identification des paramètres –prémisse/conséquence- est délicate ; elle nécessite plusieurs expérimentations pour disposer d'un bon échantillon et ceci pose des fois beaucoup de problème. Dans ce travail, nous nous intéressons à l'extraction automatique des connaissances dans les systèmes flous à partir de données numériques et à la construction incrementale de la base de règles, en utilisant l'algorithme de backpropagation qui assure convenablement l'apprentissage.

3.1 Analyse et description des systèmes flou de Sugeno

Il existe un grand nombre de possibilités de réalisation de systèmes flous. En effet, pour chaque module constitutif d'un système flou, il existe une multitude de choix différents, et chaque combinaison de ces choix engendre une classe de systèmes flous. Dans ces travaux, nous nous intéressons au modèle flou du type Sugeno, développé initialement par Sugeno et Takagi pour la modélisation des systèmes à partir de données numériques [4], [5], [7], [15], [20], [21], [24], [32], [33]. Une règle typique dans le système flou de Sugeno est de la forme suivante :

$$R_1: \text{Si } x_1 \text{ est } A_1^1 \text{ et } \dots \text{ et } x_n \text{ est } A_n^1 \quad \text{Alors } y_1 = f(x) \quad (3.1)$$

Où A_i^j , sont des ensembles flous dans l'espace d'entrée, tandis que $f(x)$ est une fonction numérique dans l'espace de sortie. En général, $f(x)$ est une fonction polynomiale en fonction des variables d'entrées, mais elle peut être aussi une fonction arbitraire tant qu'elle peut décrire convenablement le comportement du système étudié. Si $f(x)$ est un polynôme d'ordre un, on a affaire donc à un système de Sugeno d'ordre un. Si par contre, $f(x)$ est un polynôme d'ordre zéro, on a donc un système de Sugeno d'ordre zéro. Etant donnée que chaque règle possède une conclusion numérique, le temps consommé par la procédure de défuzzification est évité.

Comme notre but est de développer une loi d'adaptation des paramètres du système flou. Il est donc indispensable de donner l'expression analytique de la sortie du système flou de Sugeno, et de présenter ensuite l'algorithme d'estimation des paramètres du système flou pour approximer une fonction non linéaire quelconque à partir de données numériques.

3.1.1 Architecture du modèle flou de Sugeno [4], [5], [25]

Notons par x_1, \dots, x_n les entrées du modèle flou, et par y sa sortie. Pour chaque x_i est associé m_i ensembles flous A_i^j dans U_i , tel que pour x_i quelconque de U_i , il existe au moins un degré d'appartenance $\mu_{A_i^j}(x_i) \neq 0$, où $i=1,2,\dots,n$ et $j=1,2,\dots,m_i$. la base des règles du système flou comporte $M = \prod m_i$ règles telles que :

$$R_1: \text{SI } x_1 \text{ est } A_1^{l_1} \text{ et } \dots \text{ et } x_i \text{ est } A_i^{l_i} \text{ et } \dots \text{ et } x_n \text{ est } A_n^{l_n} \text{ Alors}$$

$$y_L(x) = a_0^l + a_1^l x_1 + \dots + a_n^l x_n. \quad (3.2)$$

Avec :

l : correspond au nombre de règles, ou $l=1, \dots, M$.

i : correspond au nombre d'entrées, ou $i=1, \dots, n$

$$1 \leq l_i \leq m_i.$$

La base de connaissance comporte toutes les combinaisons possible des ensembles flous des variables d'entrée, des lors, on aura, chaque règle flou R_l correspondra à une combinaison des $((A_1^{l_i}, \dots, A_i^{l_i}, \dots, A_n^{l_i}))$ ensembles flous.

La sortie du système flou sera donnée par l'expression suivante :

$$y = \frac{\sum_{l=1}^M \mu_l y_l}{\sum_{l=1}^M \mu_l} \quad \text{avec : } \mu_l = \prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) ; 1 \leq l_i \leq m_i. \quad (3.3)$$

μ_l : représente le degré d'activation de la règle R_l

Si on a affaire a des prémisses à paramètres fixe, alors on n'ajustera que ceux de la conclusion, ainsi la sortie finale peut être sous la forme suivante :

$$Y = w(x) A. \quad (3.4)$$

Avec :

A : le vecteur regroupant les paramètres a_i^l de dimension $(1, M^* (n + 1))$.

W : vecteur ligne de fonctions floues de dimension $(M^* (n + 1), 1)$.

La logique flou nous permet d'avoir un large choix quant au fonction d'appartenance, pour caractériser les ensembles flou A_i^j on choisies la fonction gaussienne définie par la relation :

$$\mu_{A_i^j}(x_i) = \exp \left\{ -0.5(v_i^j (x_i - c_i^j))^2 \right\} \quad (3.5)$$

Où c est la moyenne, v est la variance. Cette dernière est adopté afin d'éviter la division par zéro lors de l'adaptation de ce paramètre.

La figure 3.1 montre une représentation schématique sous forme d'un réseau d'un système flou de Sugeno à deux variable d'entrées et deux ensembles flous pour chaque variable.

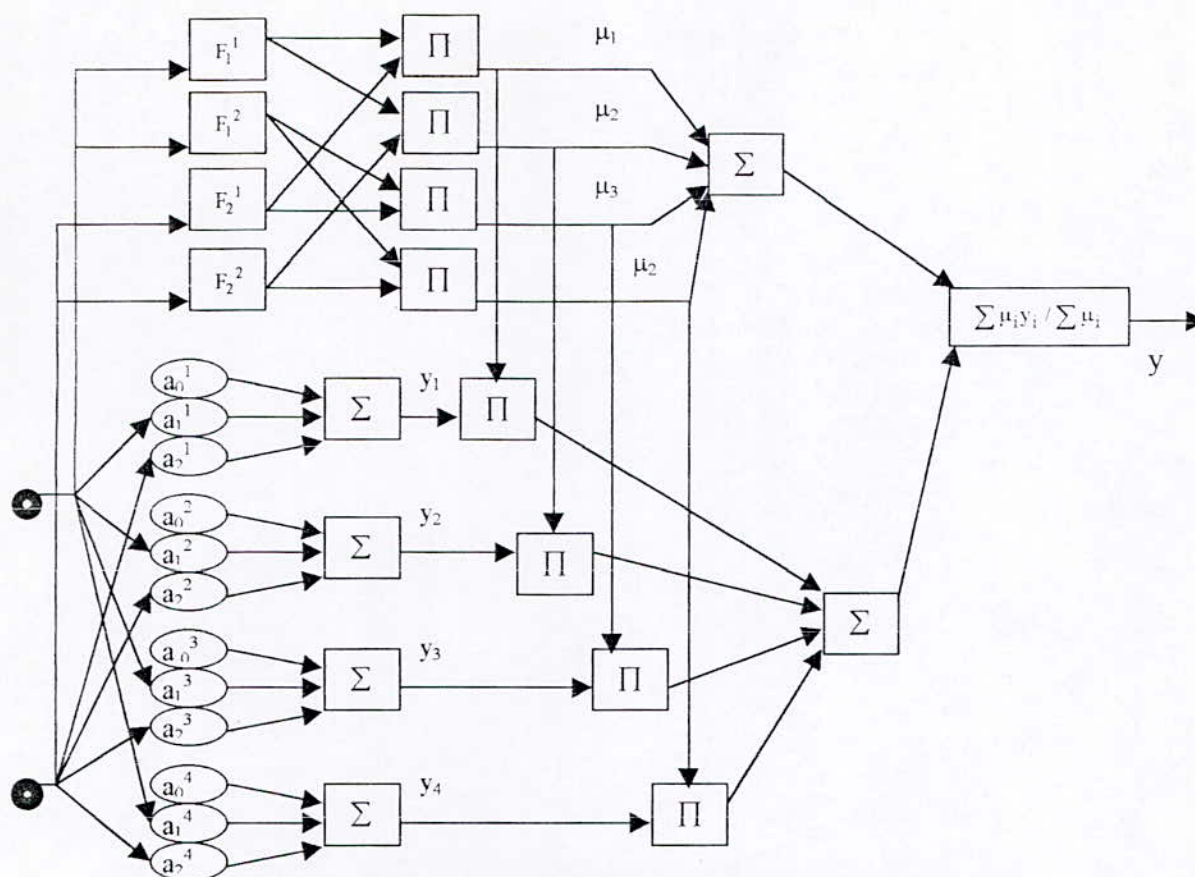


Figure 3.1 : Schéma du système flou de Sugeno

3.1.2 Adaptation des paramètres des systèmes flous de Sugeno [22], [25] :

Compte tenu du nombre important de paramètres à déterminer (paramètres des prémisses et des conséquents), la détermination des paramètres d'un système flou est la phase la plus délicate dans sa conception. De ce fait, un expert ne pourra garantir un réglage optimal du système flou qu'après un certain nombre d'essais (passage à travers notre fichier d'exemple), jusqu'à satisfaire les conditions de test d'arrêt (en général c'est à une erreur minimale qu'on doit s'attendre), et ce malgré une très grande rigueur et une méthodologie conséquente.

Des techniques d'optimisation et d'apprentissage ont été développées dont le but de suppléer ou d'aider tout utilisateur dans la phase de mise au point du système flou. La méthode d'identification est basée sur les travaux de Li-Xin Wang, dans laquelle un modèle flou peut être développé directement à partir de données numériques issues du système à modéliser.

3.2 Backpropagation :

C'est Werbos qui, en 1972 [31], a le premier mis en œuvre cet algorithme avec notamment, son fondement mathématique. Il avait appelé cette méthode The dynamic Algorithm feed-back. En 1986, un groupe de chercheurs de l'université de Stanford aux États-Unis, l'ont complètement adapté aux réseaux multicouches. Ils l'ont appelé la méthode backpropagation en référence à l'erreur qui se rétro-propage à travers les couches du réseau.

Dans notre cas, on a affaire à une estimation des paramètres du système flou θ , ce vecteur contient les paramètres des prémisses et des conclusions. C'est à ce stade qu'intervient l'algorithme d'adaptation en essayant de minimiser l'erreur de sortie e .

La forme bien connue de cette algorithme est donnée par :

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \eta(k)\psi(k)e(k) \\ e(k) = d(k) - y(x(k), \hat{\theta}(k-1)) \end{cases} \quad (3.6) \text{ et } (3.7)$$

$$\theta = [c, v, a]^T.$$

η : est le gain de l'algorithme d'estimation, ou taux d'apprentissage.

$\hat{\theta}$: est le vecteur estimé.

φ : est le gradient de la sortie estimée.

3.3 Taux d'apprentissage :

Les différentes méthodes d'identification récursive citées dans la littérature se distinguent essentiellement par le choix de ce gain qui doit être dans tous les cas pris égal à une matrice positive. Conventionnellement ce facteur doit être petit assurant la convergence mais d'une manière très lente, au voisinage de l'optimum, et vis versa. Grand quand on est loin de ce dernier pour accélérer la convergence. De ce fait, le choix de la valeur de ce paramètre est d'une importance capitale, pour cette raison le taux d'apprentissage doit être adaptatif.

Les méthodes les plus utilisées et qui sont très souvent citées dans la littérature [31,35] sont :

- **3.3.1 Méthode du gradient simple :**

C'est la méthode la plus simple à mettre en œuvre. Le gain $P(k)$ est choisi égal au produit d'un gain scalaire par la matrice identité :

$$\eta(k) = \alpha I, \quad \alpha > 0 \quad (3.8)$$

- **3.3.2 Méthode du gradient modifiée :**

Dans cette méthode, le gain $P(k)$ est adaptatif il est donné par :

$$\eta(k) = \frac{\alpha_1 I}{\alpha_2 + \psi^T(k)\psi(k)} \quad \alpha_1 > 0, \alpha_2 > 0 \quad (3.9)$$

- **3.3.3 Méthode des moindres carrés récursifs :**

Dans ce cas, le gain d'estimation $P(k)$ est donné par la relation récursive suivante :

$$\eta(k) = \eta(k-1) + \frac{\psi^T(k)\eta(k-1)\psi(k)}{1 + \psi^T(k)\eta(k-1)\psi(k)} \quad (3.10)$$

avec $P(0) = \alpha I$ ou $\alpha > 0$.

3.4 Gradient de la sortie :

Nous remarquons que l'algorithme de backpropagation nécessite le calcul du gradient $\varphi = \frac{\partial y}{\partial \theta}$ ou $\theta = [c, v, a]^T$, nous avons donc :

$$\frac{\partial y}{\partial \theta} = \left[\frac{\partial y}{\partial c}, \frac{\partial y}{\partial v}, \frac{\partial y}{\partial a} \right]^T \quad (3.11)$$

Pour calculer ce terme il faut user de quelques artifices mathématiques :

$$\frac{\partial y}{\partial c_i^j} = \sum_{\substack{k \in L \\ c_i^j}} \frac{\partial y}{\partial \mu_k} \frac{\partial \mu_k}{\partial \mu_{A_i^j}} \frac{\partial \mu_{A_i^j}}{\partial c_i^j} \quad (3.12)$$

Où $L_{c_i^j}$: est l'ensemble des indices L des règles floues (R_L) dont lesquelles apparaît l'ensemble flou A_i^j .

$$\frac{\partial y}{\partial \mu_k} = \frac{y_k - y}{\sum_{l=1}^M \mu_l} \quad (3.13)$$

$$\frac{\partial \mu_k}{\partial \mu_{A_i^j}} = \prod_{\substack{l=1 \\ l \neq i}}^n \mu_{A_l^{kl}}(x_l) \quad (3.14)$$

Pour une fonction gaussienne on a :

$$\frac{\partial \mu_{A_i^j}}{\partial c_i^j} = v_i^{j2} (x_i - c_i^j) \mu_{A_i^j}(x_i). \quad (3.15)$$

Après calcul, on a :

$$\frac{\partial y}{\partial c_i^j} = \frac{v_i^{j2} (x_i - c_i^j) \sum_{k \in L} \mu_k (y_k - y)}{\sum_{l=1}^M \mu_l} \quad (3.16)$$

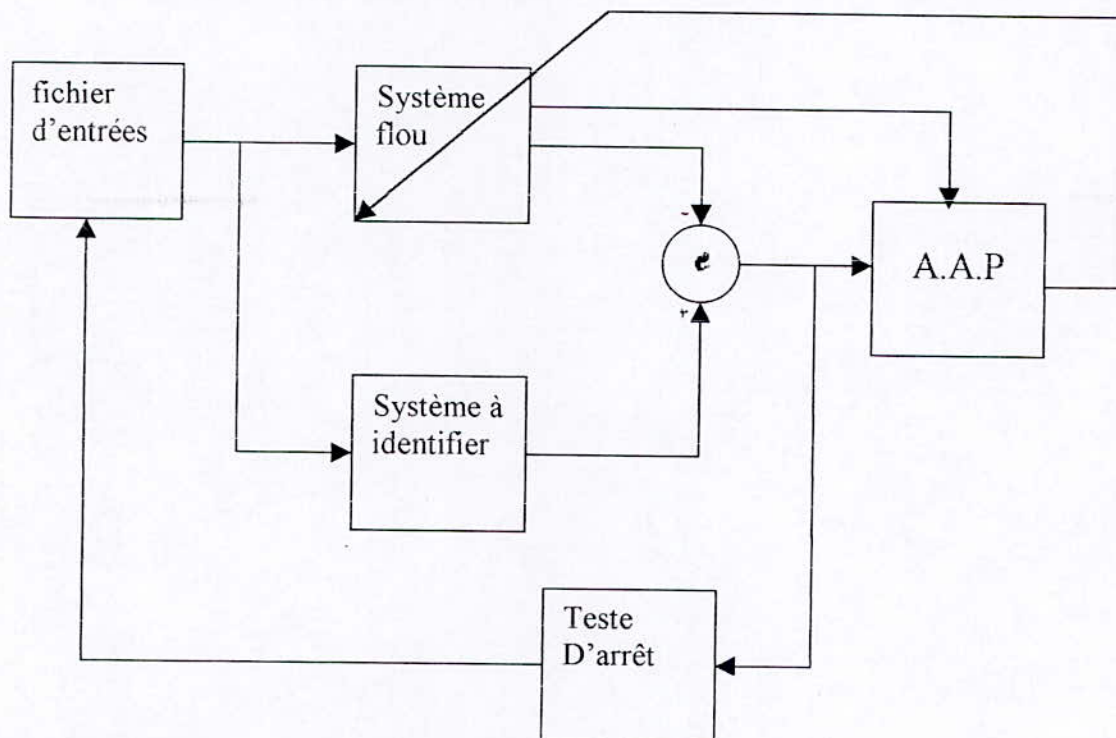
De la même manière, on trouve pour les autres paramètres :

$$\frac{\partial y}{\partial v_i^j} = \frac{-v_i^j (x_i - c_i^j)^2 \sum_{k \in I} \mu_k (y_k - y)}{\sum_{l=1}^M \mu_l} \quad (3.17)$$

La dérivée de la sortie par rapport aux paramètres des conséquences est donnée par :

$$\frac{\partial y}{\partial a_i^j} = \frac{\mu_j x_i}{\sum_{l=1}^M \mu_l} \quad (3.18)$$

On termine cette partie en donnant un schéma synthétisant des étapes de la backpropagation :



3.5 Objectif :

On veut réduire le nombre des règles tout en éliminant la procédure de défuzzification, ce qui est faisable en utilisant les systèmes flous de Sugeno. Nous allons procéder avec quatre méthodes :

- Nombre de fonctions d'appartenance de chaque entrée est de 3 et le nombre de règles est pris à 9
- Nombre de fonctions d'appartenance de chaque entrée est 7 et le nombre de règles est pris égal à 7.
- Nombre de fonctions d'appartenance de chaque entrée est de 5 et le nombre de règles est 5

- Nombre de fonctions d'appartenance est 3 et le nombre de règles est pris à 3

Une fois la base de connaissance choisie il nous reste à déterminer l'estimation des paramètres θ pour chacun des 4 cas. Pour cela nous allons utiliser l'algorithme de backpropagation.

3.6 Simulation et résultats :

Dans les quatre approches d'apprentissage, nous fixons le nombre d'échantillons N_{ex} et la somme des erreurs quadratiques sur tout l'horizon (fichier d'exemple) E_{rt} par :

$$N_{ex}=1000, E_{rt}=10^{-4}$$

Où les A sont les paramètres des conclusions et C et S le centre et l'inverse de la variance de la prémisse.

Nous obtenons les résultats suivants pour chaque approche:

- Nombre de fonctions d'appartenance de chaque entrée est 3 et le nombre de règles est pris égal à 9.

A

-1.0858e-1	-2.5727e-2	3.3425e-1
2.4521e-2	8.8368e-3	2.1094e-2
1.0013e-2	9.9999e-3	1.0009e-2
5.5266e-2	7.9754e-1	1.2349e+000
9.2521e-2	5.4564e-2	6.8136e-2
3.1334e-2	1.3949e-2	1.2921e-1
-5.9059e-2	9.3534e-4	1.9192e-1
1.9011e-2	9.6522e-3	1.9587e-2
3.1833e-2	1.4038e-2	1.3223e-1

C

-3.5799e-2	-5.9031e-3	-1.1731e-1
8.2290e-1	2.7177e+00	4.6442e+00

S

3.1473e+001	3.1473e+01	3.1473e+01
1.0440e+000	2.7177e+00	4.6442e+00

- Nombre de fonctions d'appartenance de chaque entrée est 7 et le nombre de règle est pris égal à 7.

A

.46492	.48062	.46044
.13886	.58245	.66809
-.41105	.63154	.63642
9.5628*10 ⁽⁻²⁾	.50315	1.0106
.4307	.11378	.88329
.37191	.36665	.34639
.25086	.25094	3.4168*10 ⁽⁻³⁾

C

-0.77238	-1.3625
-.57370	-.74888
-.35022	-0.15868
-.15519	.25501
.38842	.63427
.70450	1.4467
.99988	1.9994

S

3.65	1.8347
3.6738	1.8474
3.6955	1.7535
3.6539	1.8065
3.6386	1.6575
3.6349	1.8133
3.6378	1.8173

- Nombre de fonctions d'appartenance de chaque entrée est 5 et le nombre de règles est pris égal à 5.

A

0.18680	.39049	.22842
-0.21925	0.86856	1.1356
0.45976	0.37088	1.1834
.048726	-0.063816	-0.16829
0.35842	.36794	.17214

C

-0.72411	-1.2941
-0.38144	.51062
0.38296	2.0088
.88484	3.8867
1.0151	5.6071

S

2.5205	.57473
2.5971	.84613
2.6071	.73252
2.4632	.48668
2.4304	.62013

- Nombre de fonction d'appartenance de chaque entres est 5 et le nombre de règle est pris égal à 5.

A

$5.8009 \cdot 10^{-3}$	$-2.4638 \cdot 10^{(2)}$	$4.0479 \cdot 10^{-1}$
$-3.8152 \cdot 10^{(3)}$	1.7427	1.3579
$1.9496 \cdot 10^{-1}$	$2.6875 \cdot 10^{-3}$	$3.4226 \cdot 10^{-1}$

C

$-9.1583 \cdot 10^{-1}$	-1.4219
$-7.9943 \cdot 10^{-4}$	1.7643
1.0515	5.4695

S

1.0784	$6.6739 \cdot 10^{-1}$
2.5037	$4.4811 \cdot 10^{-1}$
1.1519	$5.8236 \cdot 10^{-1}$

Conclusion :

Les systèmes flous exigent la disponibilité d'une expertise. Par conséquent, les performances de ces derniers sont étroitement liées aux techniques d'acquisition de connaissances et la justesse des informations acquises.

Pour palier à ces inconvénients, ont été développées des méthodes d'apprentissage utilisables dans le cas où aucune expertise n'est disponible. De ce fait, les systèmes flous sont devenus des systèmes dotés d'une capacité d'apprentissage.

Notons que la réduction du nombre de règles a été faite en imposant le nombre des variables prémisses 9, 7, 5 et puis 3 pour l'identification du système flou. Pour palier à cet inconvénient, les notions de clustring et d'analyses de similarité ont été développées dans les chapitres précédents pour la réduction des règles de base.

CHAPITRE 4

Notions de Clustering et de Similarité

Introduction :

Ce chapitre servant d'introduction aux chapitres qui le suivent, nous permet de développer les principales notions de similarité. On présentera au début, la notion de clustering flou, les différentes algorithmes de clustering, ainsi que la façon d'appliquer ces algorithmes, sans oublier la notion d'identification des systèmes.

Nous parlerons aussi de la complexité des règles de base, la façon de réduire cette complexité en se servant des techniques de clustering.

Enfin, ce chapitre abordera les notions de 'similarité', ce dont il s'agit et la manière dont elles contribuent dans divers domaines. La dernière partie de ce chapitre se limitera à expliquer la similarité entre fonctions d'appartenance, du point de vue simplification des règles de base. Aussi ce chapitre a pour but d'inspirer le lecteur pour penser au sujet de similarité de différentes manières.

A noter, que les modèles considérés dans tout ce qui suit sont des modèles de Takagi-Sugeno.

4.1 Notion de clustering et Identification par Clustering [2] [3] [4] [5] [6] [7] :

La modélisation floue peut se faire par l'extraction de connaissances des experts humains et transformer ces expertises en règles et fonctions d'appartenance. Le système résultant peut être réglé par le contrôle de ces performances à travers des tâtonnements ou bien par divers procédures automatisées, par exemple l'adaptation et l'apprentissage. Cependant, en se basant sur les connaissances des experts et leurs expériences cela présente quelques problèmes. Le premier de tous, est que l'expert humain n'est pas toujours disponible, et quand il l'est, ces connaissances sont souvent incomplètes et épisodiques plutôt que systématiques. De plus, c'est un problème d'acquisition de connaissances effectives qui nous amène au réglage du nombre de paramètres chaque fois qu'on se rapporte au choix des variables, ainsi que la détermination du type de fonction d'appartenance, leurs formes et paramètres, et le nombre de règles dans leurs règles de base. Finalement, le système doit avoir l'habileté d'apprentissage pour se mettre à jour et s'autorégler, en se basant sur l'information devenue disponible durant l'opération. La modélisation automatique basée sur la formation de données numériques (mesures du système) est donc très souhaitable. Le résultat de chaque modélisation automatique de processus peut être un système de règles de bases comme donné en figure 4.1

La tâche de modélisation peut être divisée en deux parties similaires comme dans la théorie des systèmes ordinaires. L'identification de structures et l'identification de paramètres. L'une des méthodes qui peut être appliquée à la génération automatique du modèle flou est le Clustering flou. L'algorithme de clustering flou est un algorithme supervisé. Un modèle flou du système peut être obtenu. Cependant, avant l'application de l'algorithme de clustering flou, la structure du problème doit être déterminée, identifiant le rapport système variables. Alors, le nombre de groupes requis doit être sélectionné. Ce nombre détermine le nombre de règles dans les règles de base résultantes. Deux algorithmes de clustering flou connus sont le C-means flou Algorithme et le Gustafson-Kessel Algorithme (GK algorithme)[30].

Les conséquences dans le modèle de Sugeno-Takagi sont les hyperplans dans l'espace produit prémisses-conséquences. Pour chaque modèle, le GK Algorithme convient particulièrement à la génération automatique des règles, comme il cherche des hyperplans de dimension (n-1) dans un espace de dimension n. L'algorithme identifie le modèle par clustering de données à partir de mesures entrées/sorties

disponibles, ayant un nombre spécifique de groupes et de paramètres pour le clustering flou et le critère d'arrêt. Cependant, l'identification de structure [8], [9], c'est à dire détermination du nombre de groupes, est quelque peu complexe.

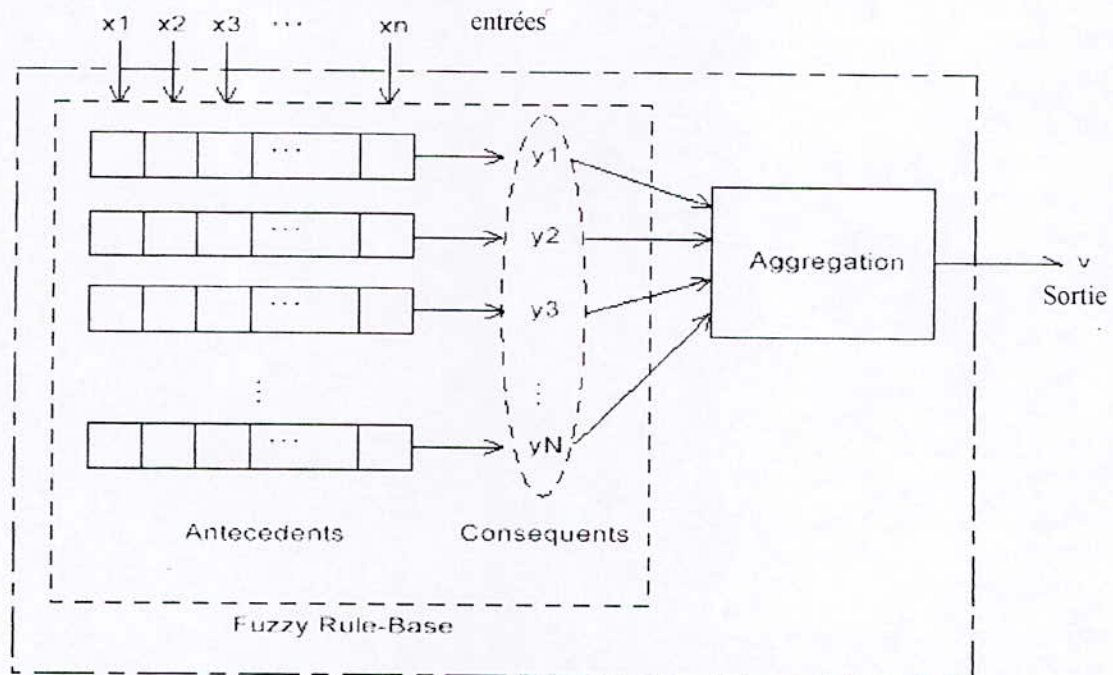


Figure 4.1 : système d'inférence flou.

Quand les groupes dans l'espace produit du modèle entrées et sorties sont trouvés, ils sont représentés par des fonctions d'appartenance qui peuvent être données en projetant les groupes sur les variables entrées/sorties et ensuite ajuster les ensembles flous convexes aux projections. Si on utilise le C-means flou méthode, les ensembles flous adaptés aux projections des variables entrées/sorties peuvent servir de fonctions d'appartenance dans les entrées et sorties antécédentes des règles dans les modèles de types Takagi-Sugeno. La projection des groupes générée par l'algorithme GK sur les variables d'entrées donnera les fonction d'appartenance dans les règles antécédentes. Les paramètres dans les combinaisons linéaires résultantes, c'est à dire, le modèle local (3.2), peuvent être obtenus par clustering ou par technique d'estimation des paramètres, tel que la méthode des moindres carrés et la méthode du gradient, utilisant les données formées.

4.2 Complexité des règles de base

Généralement appliqué, l'algorithme de clustering flou tel que le C-means flou ou le GK algorithme ne donne pas d'indication du nombre correct de groupes nécessaires. Ils trouvent un nombre optimal (spécifié) de groupes, peu importe si les groupes obtenus sont significatifs ou pas. Une spécification correcte du nombre de groupes est importante. Un nombre important de résultats inutiles va compliquer les règles de base, bien qu'un nombre insignifiant peut résulter dans un modèle non complexe. Par conséquent, des méthodes sont suggérées pour déterminer le nombre optimal de règles (groupes) dans les règles de base. L'une des méthodes est de démarrer le clustering avec deux groupes, puis évaluer la fonction de performance

et répéter le processus par addition des autres groupes jusqu'à ce que la fonction performance satisfasse un certain critère. L'autre méthode est l'algorithme de fusion de groupes compatibles qui démarre avec un nombre important de groupes puis le réduit en le fusionnant jusqu'à ce que aucun groupe ne peut être fusionné. Dans ce qui suit, la similarité entre les fonctions d'appartenance est étudié avec comme objectif la réduction de la complexité des règles de base.

4.3 Fonctions d'appartenance similaires [28], [30] :

Si une modélisation automatique est appliquée, le résultat peut être un modèle avec plusieurs fonctions d'appartenance. Quelques-unes de ces fonctions d'appartenance peuvent être similaires (quelque fois égales), ainsi introduisant la redondance dans les règles de base. Par fonctions d'appartenance similaires, on doit penser au grand chevauchement d'ensembles flous définis dans le même domaine dont les fonctions d'appartenances prennent approximativement les mêmes valeurs pour toutes les valeurs des variables du domaine.

La figure 4.2 donne un exemple de fonctions d'appartenances similaires ou pas. Une explication détaillé sur la similarité est donnée dans la partie 4.5

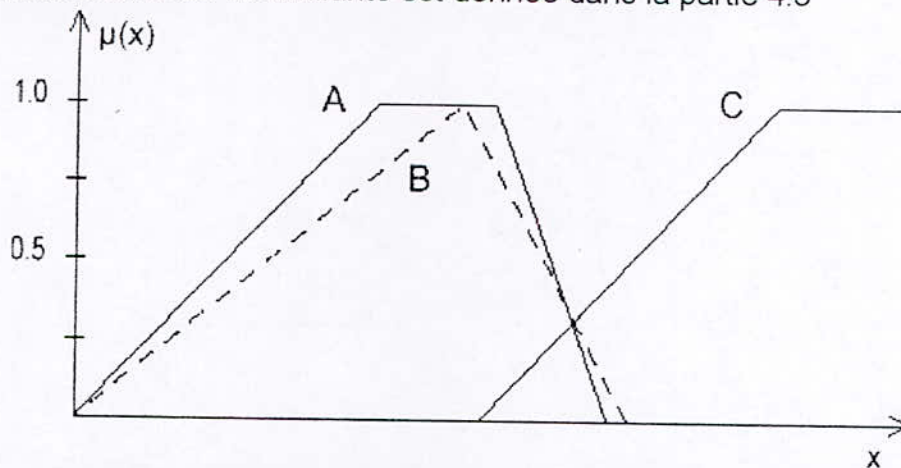
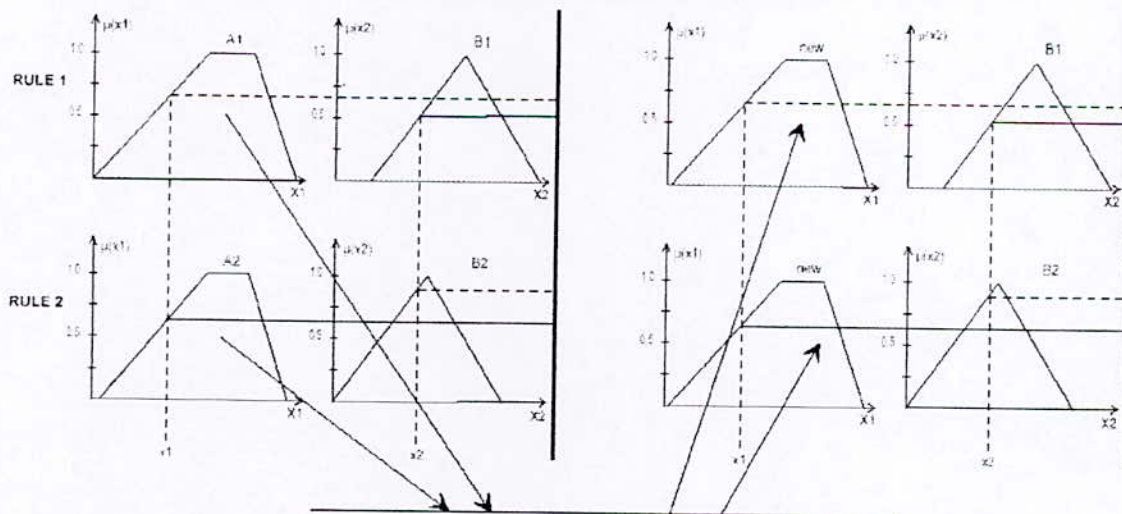


Figure 4.2 : A et B sont des fonctions d'appartenance très similaires, ce n'est pas le cas pour, (A et C) ainsi que (B et C)



A₁ et A₂ sont similaires, elles sont remplacées par 'new'

Figure 4.3 : fonctions d'appartenance A₁ et A₂ similaires remplacées par la fonctions 'new'

La redondance introduite dans les règles de base, peut être vu comme mariage d'entrée, contre approximativement les mêmes antécédents (fonctions d'appartenance similaires) dans deux règles ou plus. Chacune de ces fonctions d'appartenance donnera le même degré d'activation de l'entrée, donc ces fonctions d'appartenance peuvent être remplacées par une fonction d'appartenance commune donnant cette activation de l'entrée. Ainsi l'idée est de fusionner ces fonctions d'appartenance hautement similaires pour créer une fonction d'appartenance représentative pour cette fusion. Figure 4.3

Fusionner les fonctions d'appartenance similaires (c'est à dire, remplacer deux fonctions ou plus par une) réduit la complexité du modèle. Les règles de base résultantes seront sémantiquement plus étendues, l'approximation linguistique sera simplifiée, d'où la description qualitative du modèle sera facile a saisir, en abaissant le nombre total de fonctions d'appartenance mais aussi les règles de base dans le moins de mémoire possible quand on les implémentes.

Le concept de réduction de complexité des règles de base par fusion des fonctions d'appartenance, appelé simplification des règles de base, signifie que le nombre de fonctions d'appartenance est réduit, mais le nombre de règles reste le même. La réduction du nombre de règles de base sera appelée réduction des règles de base. Si tous les antécédents dans deux règles sont similaires, de façon que toutes les paires de fonctions d'appartenance dans les dimensions individuelles peuvent être fusionnés, c'est à dire, les prémisses des deux règles sont rendus égaux, alors leurs parties résultantes peuvent être combinées et les deux règles peuvent être remplacées par une.

Cette réduction des règles de base apportera bien sur les mêmes avantages que la simplification de règles de base.

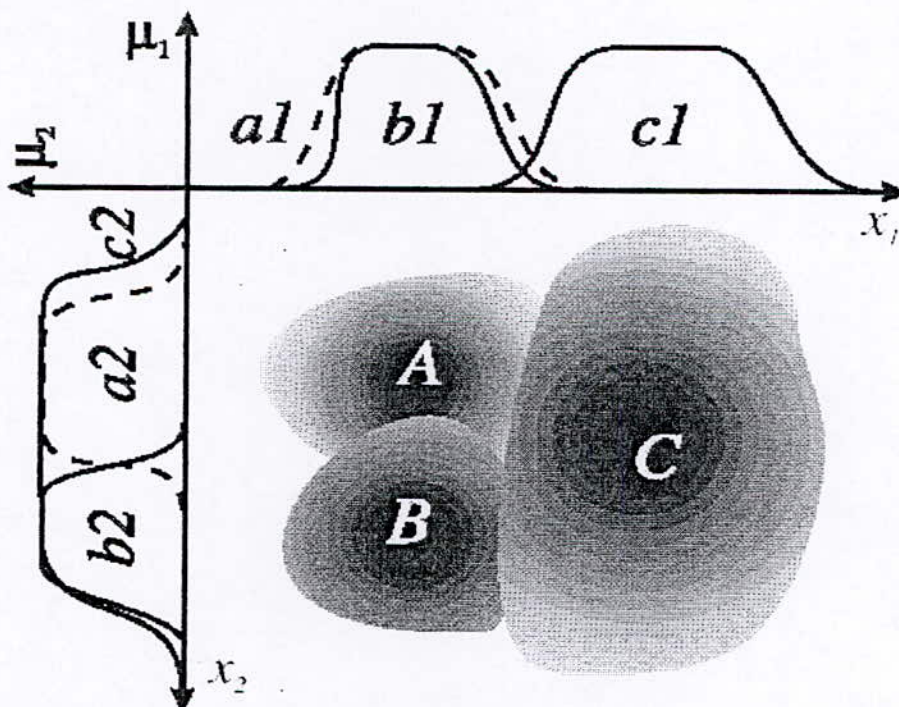


Figure 4.4 : les groupes flous projetés sur les axes x_1 et x_2

Dans l'identification des modèles [27], [28], [30], la similarité entre les fonctions d'appartenance résulte de la projection des groupes sur les axes de l'espace groupé. Ceci est illustré en figure 4.4.

4.4 La similarité

La similarité joue un rôle important dans le raisonnement humain et la prise de décision. Nous avons tendance à grouper des objets dans des classes basées sur comment ces derniers nous apparaissent, c.-à-d. nous vérifions la similarité entre les objets. La même chose est appliquée pour l'apprentissage. Un opérateur du processus nous a appris que si la pression dans un réservoir tombe, l'opérateur doit augmenter la température. Pendant les opérations du processus, l'opérateur cherche un comportement semblable à ses références. Tel qu'une chute de pression. La modélisation et le raisonnement flou sont souvent utilisées pour modéliser l'opérateur humain ou à l'aider à prendre des décisions, et il n'y a aucune raison pour que la similarité ne devrait pas occuper également un rôle important dans l'inférence floue. Aussi, le (fameux) problème de classement des nombres flous est en rapport avec la similarité.

Si nous classifions les nombres flous A, B et C comme $A < B < C$, nous disons en fait que A est plus proche de B que de C où le rapprochement est une mesure de similarité entre les nombres flous. Pour la modélisation et le contrôle flou, la question d'Assortir/comparer des valeurs floues et leurs concepts peuvent être vues comme une mesure de similarité. En dépit de cela, la similarité a reçue une attention limitée dans les publications sur la théorie des ensembles flous. Cependant, les psychologues ont longuement discuté de la similarité et de la distance entre objets ; et comme beaucoup d'autres aspects de la théorie des ensembles flous, la mesure de similarité pour les ensembles flous a été proposée, inspirée par les études dans les sciences des comportements. Les méthodes pour mesurer la similarité entre ensembles flous sont abordés dans Chapitre 5.

4.4.1 Les divers similarités :

Le mot « semblable » est utilisé dans beaucoup de contextes et la signification exacte du mot dépend du contexte actuel [28]. Quelquefois la similarité est utilisée pour voir comment les choses sont classées (par exemple un B-52 et un DC-10 peuvent être vu comme étant semblables) et d'autres fois nous avons tendance à l'utiliser comme une description des fonctionnalités (par exemple la charge d'un DC-10 n'est pas semblable à celle d'un B-52).

Des exemples d'expressions apparentées sont égaux, semblables, le même, approximativement et plus ou moins. Si nous consultons le dictionnaire [WEB], nous trouvons les interprétations suivantes de « semblable » et « égal »:

Semblable: - avoir des caractéristiques en commun

- pas différent dans la forme mais dans la dimension ou la position
- ressembler à mais pas le même
- du même genre ou même quantité

La similarité: - la qualité ou l'état d'existence de similitude

- un point dans le quel les choses sont semblables

Egale: - de la même mesure, quantité, somme, ou un nombre comme un autre

- identique dans la valeur

- avoir les mêmes droits ou les mêmes situations

Cependant, l'interprétation de similarité dans la théorie des ensembles est légèrement différente dans les langues parlées. Dans la théorie des ensembles, la similarité est basée sur combien d'éléments habituellement deux ensembles ont en commun ou sur les mesures métriques (Définition 5) définissant la distance entre les deux ensembles. Si les deux ensembles contiennent exactement les mêmes éléments, ils sont reportés comme étant égal. Pour les ensembles flou cela devient plus compliqué quelque peu, puisque chaque élément peut appartenir avec des degrés différents à plusieurs ensembles.

Les subdivisions suivantes illustrent l'usage de la similarité dans plusieurs champs d'applications. Dans chacun de ces cas, différentes mesures/méthodes peuvent être utilisés à la quantification/capture de similitude.

4.4.2 les approximation linguistique [8], [9], [29]

Un modèle linguistique est décrit ou est exprimé en utilisant des termes linguistiques dans la structure de logique floue au lieu d'équations mathématiques avec des valeurs numériques ou des formules logiques conventionnels avec les emblèmes de la logique.

Considérez un modèle flou utilisé dans le contrôle flou communément:

1. Si x est positif petit, ALORS y est négatif petit,
2. Si x est moyen positif, ALORS y est positif petit.

Ici les termes 'petit positif', 'moyen positif', 'etc..', sont des labels assignés conventionnellement aux ensembles flous qui jouent un rôle important contrairement aux labels qui peuvent être (A1, B2 etc., voir Figure 4.5). Lorsque l'approximation linguistique détient une proposition avec des attributs flous, elle essaie de trouver un mot ou une phrase de sortie d'un ensemble donné de mots (par exemple 'calme', 'doux', 'modéré', 'etc..') pour se rapprocher linguistiquement de la proposition. Donc la nature qualitative du modèle est améliorée en le remplaçant avec un modèle où l'attention est concentrée (focalisée) sur l'explication linguistique ou qualitative du comportement des systèmes. Cela est fait en mesurant la similarité entre les attributs flous et les fonctions d'appartenance représentant les mots qui sont utilisés pour l'approximation linguistique. Le mot qui a une fonction d'appartenance plus semblable à l'attribut flou remplace alors cet attribut dans la règle. Comme exemple, la figure 4.5 montre l'approximation des deux ensembles flous A et B, définis pour la force du vent, utilisant les termes linguistiques. En plus des mots eux-mêmes, on peut aller aussi à une étape supplémentaire et utiliser des superlatifs et des Connectives.

Les superlatifs sont utilisés pour modifier les mots afin qu'ils améliorent le rapprochement des attributs flous dans le modèle. Les superlatifs peuvent être des adverbes tel que par exemple 'très ' ou 'plus ou moins', et sont placés devant les mots dans le modèle linguistique: " Très froid ". Les superlatifs sont souvent modelés en appliquant des opérateurs aux ensembles flous qui décrivent les mots, tel qu'intensification, dilatation, concentration, puissance etc.. Le superlatif 'Très ' est typiquement modelé en rendant carré la fonction d'appartenance appartenant au mot: $\mu_{\text{très froid}} = \mu_{\text{froid}}^2$. Les Connecteurs sont des mots comme 'et ', 'mais ', et 'ou '. L'usage de superlatifs et Connecteurs peuvent mener aux règles souhaitées:

Si la température est 'plus ou moins chaude' ou 'très chaude' ALORS

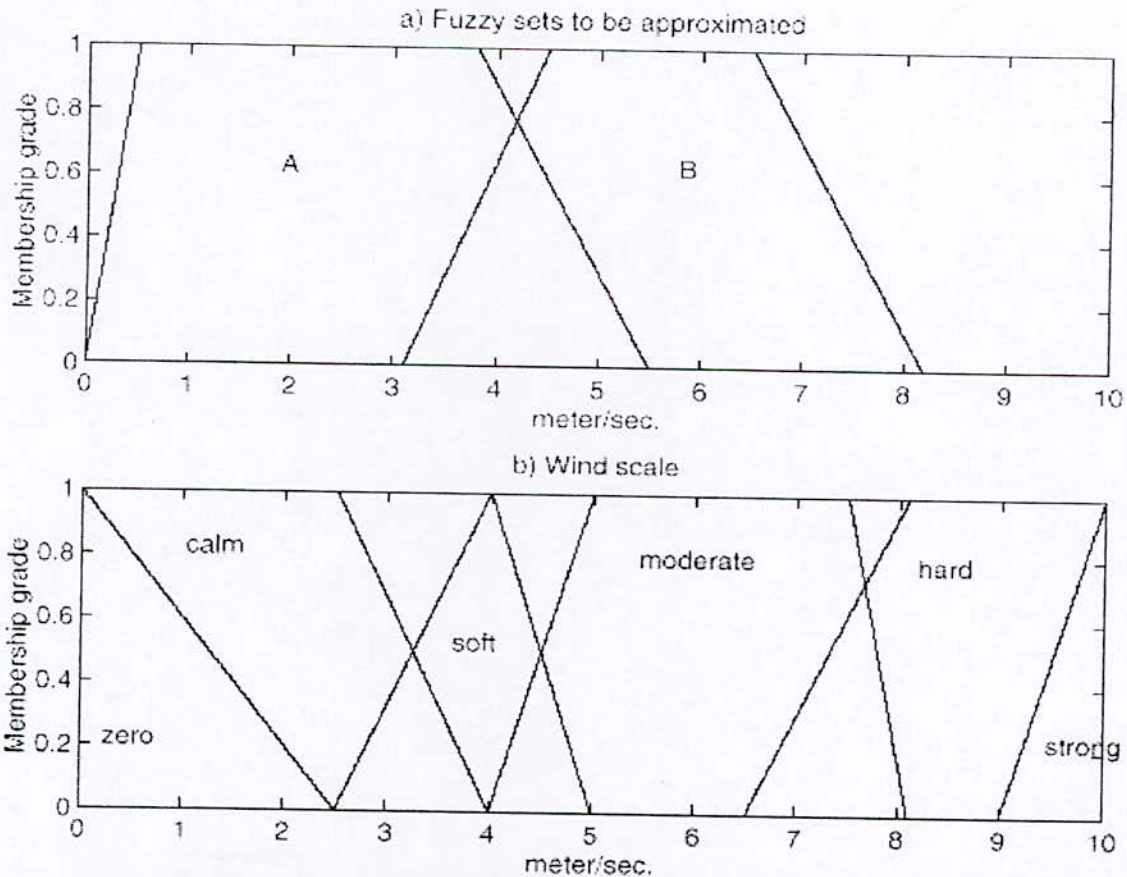


Figure 4.5 : a) ensembles flous, b) l'approximation linguistique

4.4.3 La similarité et la perception

La notion de similarité joue un rôle fondamental dans les théories de la connaissance et du comportement et a été négocié dans la littérature de la psychologie [28]. La similarité est importante pour savoir comment nous allons organiser, simplifier et interpréter les stimulus complexes présentés à nous par l'environnement. Le processus qui est en notre capacité et qui donne signification aux stimulus est appelé la perception. Les processus perceptuels de base opèrent d'après une série de principes qui décrit comment nous organisons les morceaux et les pièces d'information dont la totalité est pleine de sens. Ils sont connues comme lois d'organisation, d'après les études menées par les psychologues Allemands[48]:

La fermeture: - Habituellement, pour les Groupements on utilise les termes de clos ou figures complètes plutôt que des termes comme ouvert (dans Fig. 4.6a nous voyons une série de crochés qui nous illustrent ce principe).

La proximité: - Les éléments qui sont plus proches entre eux sont groupés ensemble (dans Fig. 4.6b nous avons tendance à voir des paires de points plutôt qu'une ligne de points seuls).

La similarité: - Les éléments qui sont semblables en apparence sont groupés ensemble (dans Fig. 4.6c nous voyons des lignes horizontales de cercles et de carrés au lieu de colonnes verticales mélangées).

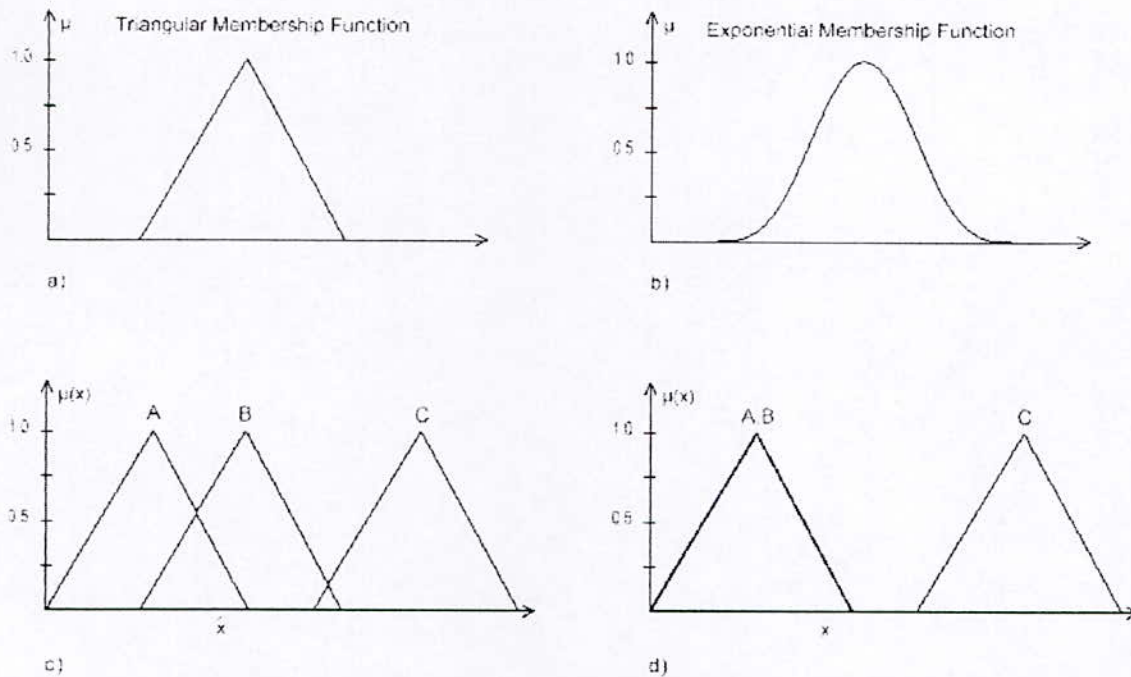


Figure 4.7 :types et cas de fonctions d'appartenance.

Une fonction d'appartenance est une fonction de la variable du domaine. Dans la Figure. 4.7c la variable du domaine est x , les trois fonctions d'appartenance sont $\mu_A(x)$; $\mu_B(x)$ et $\mu_C(x)$. elles sont toutes fondamentalement égales dans la forme, cependant, différentes du point de vu similarité: $s(A, B) > s(A, C)$. La plus proche valeur à $\mu_A(x)$ est $\mu_B(x)$, le plus semblable à $\mu_A(x)$ est $\mu_B(x)$. If $\mu_A(x) = \mu_B(x)$; $\forall x \in X$ alors $s(A; B) = 1$, on a $\mu_A(x)$ et $\mu_B(x)$ égaux, Fig. 4.7d.

Nous considérons que la similarité intervient dans plusieurs cas liés a des types bien déterminés. Cela peut être entre des cas différents du même type ou des types différents. Cependant, les types doivent appartenir à la même classe à quelque haut niveau d'abstraction (par exemple les types de fonctions d'appartenance Triangulaire et Exponentielle peuvent appartenir à la même classe d'ensembles Flous). Plus loin, les cas (de fonctions d'appartenance) doivent être définis sur le même domaine. Il est insensé de mesurer la similarité entre un ' Grands, Moyens' autobus et une dame ' Très chevelue '. pourtant les deux fonctions d'appartenance sont des cas du même type (par exemple fonction d'appartenance Triangulaire), voire Figure. 4.8.

Deux cas du même type n'ont pas nécessairement une grande similarité que deux cas de types différent. Cela veut dire qu'une fonction d'appartenance triangulaire μ_A peut être plus semblables à - par exemple- une fonction d'appartenance exponentiel μ_B qu'à une autre fonction d'appartenance triangulaire μ_C . Un exemple de cela est quand μ_A et μ_B ont la même place dans le domaine (supports à recouvrir) et que μ_A et μ_C ne chevauchent pas. Donc la signification de μ_A est plus semblable à la signification de μ_B qu'à la signification de μ_C (illustré dans Fig. 4.9). Il est clair que cette interprétation de la similarité est différent de la définition dans les dictionnaires.

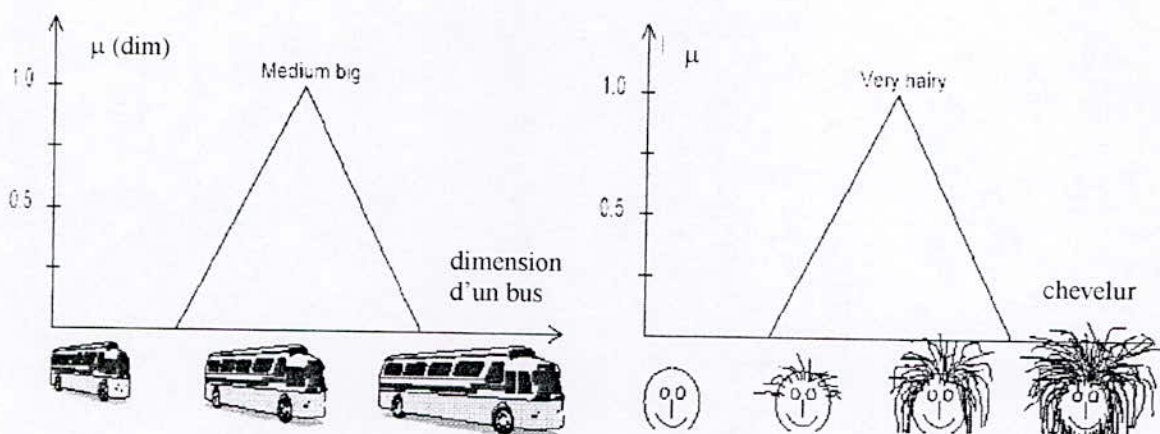


Figure 4.8 : la similarité doit être mesuré entre fonctions d'appartenance dans le même domaine

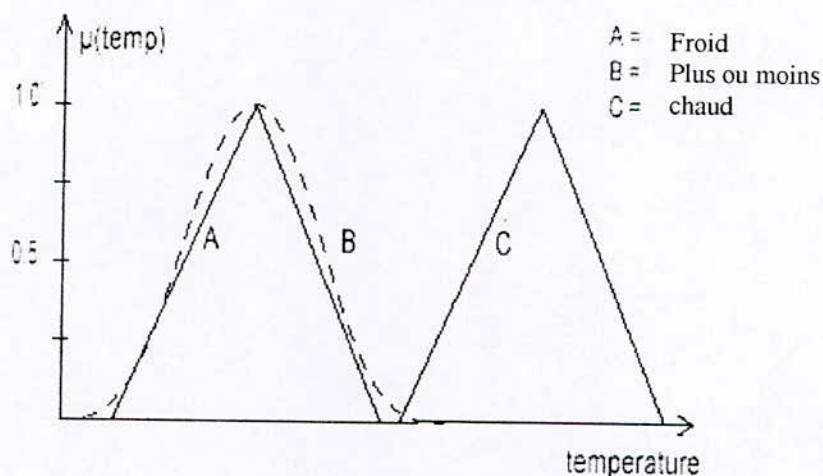


Figure 4.9 : la similarité entre différents types dans le même domaine.

Enfin, les définitions suivantes sont données pour compléter le concept de similarité, elles sont d'une très grande importance quand il s'agit d'évaluer les mesures de similarité dans le chapitre 6.

Définition 1 : Deux fonctions d'appartenance $\mu_A(x)$ et $\mu_B(x)$ sont dites égales si et seulement si $\mu_A(x) = \mu_B(x), \forall x \in X$, c'est à dire,

$$A=B \iff \mu_A(x) = \mu_B(x), \forall x \in X.$$

Définition 2 : Deux fonctions d'appartenance $\mu_A(x)$ et $\mu_B(x)$ sont dites dissemblables si elles ne se recouvrent pas pour toutes les valeurs des variables du domaine, c'est à dire :

$$\forall x \in \text{supp}(A), \mu_B(x) = 0.$$

Définition 3 : Deux fonctions d'appartenance $\mu_A(x)$ et $\mu_B(x)$ sont dites similaires si elles se recouvrent pour les mêmes valeurs de x , c'est à dire :

$$\exists x \in X, \min\{\mu_A(x), \mu_B(x)\} \neq 0.$$

Définition 4 : La similarité entre deux ensembles flous A et B est le degré d'égalité de $\mu_A(x)$ a $\mu_B(x)$.

Conclusion :

Dans ce chapitre la notion de similarité est rapportée, elle est utilisée dans différents domaines et son interprétation diffère selon le contexte dans lequel elle est utilisée. Quelques propriétés de la définition de similarité sont vues dans la dernière partie de ce chapitre et les termes égale, similaire et dissemblable y sont définis. Dans le contexte de simplification des règles de base les conclusions suivantes sont tirées :

- la similarité est mesurée entre les cas de fonctions d'appartenance et non entre les types.
- La similarité est toujours définie entre les paires de fonctions d'appartenance dans le même domaine.
- La similarité est définie comme un degré d'égalité de deux fonctions d'appartenance.
- Les positions des fonctions d'appartenance dans le domaine sont plus importantes que leurs formes.
- La valeur de la similarité donnée par les mesures de similarités ne doit pas être influencée par l'échelle du domaine.

CHAPITRE 5

Mesures de Similarités

Mesures de similarités

Dans les différentes documentations on trouve plusieurs méthodes pour mesurer la similarité [3], [5], [27], [28]. Quelques-unes d'elles sont des méthodes généralisées de celle équivalente dans la théorie des ensembles ordinaires. Très fréquemment la mesure de similarité est basée sur la mesure de distance. Par exemple, la méthode métrique de *Minkowski* et celle de *Hausdorff*. Dans ces cas là, la transformation est donnée de la distance à la similarité (c'est à dire entre distance et similarité). La compatibilité entre similarité et distance est basée sur le fait que la similarité peut être considérée comme une sorte d'inverse de la distance :

- Si la distance entre deux ensembles est grande, alors leurs similarité est moindre. Considérons deux ensembles flous A et B et la mesure de distance $d(A,B)$ tel que $d \in [0, \infty)$, la mesure de similarité entre les deux ensembles flous A et B peut être écrite comme suit :

$$s(A,B) = \frac{1}{1 + d(A,B)}; s \in [0,1] \quad (5.1)$$

La similarité peut être aussi mesurée sans utiliser la fonction distance. Les opérateurs théoriques des ensembles et les relations logiques (logique de connexion) peuvent être utilisés pour définir une mesure de similarité. L'indice de consistance (cohérence) (5.24) et l'indice d'égalité (5.25) proposé par Pedrycz [23] sont des exemples qui nous donnent une mesure de similarité entre les ensembles flous sans avoir recours à définir ou à utiliser la fonction distance.

5.1 Mesures de distances :

Le mot distance ne peut être utilisé arbitrairement en mathématique. Si x et y sont deux éléments d'un ensemble propre U entre lesquelles on voudrait définir une distance, Il est nécessaire que la mesure de distance satisfasse les conditions qui suivent :

Définition 5: $\forall x, y$ et $z \in U$

1. $d(x, y) \geq 0$ (non négative)
2. $d(x, y) = d(y, x)$ (symétrique)
3. $d(x, z) \leq d(x, y) + d(y, z)$ (inégalité triangulaire)
4. $d(x, y) = 0 \iff x = y$ (réflexion)

Quand une fonction d satisfait cette définition, elle est appelée une fonction de distance (ou une mesure de distance), et nous appelons (U, d) l'espace métrique. D'où d est souvent référé comme métrique.

Définition 6:

Un espace métrique est la paire (U, d) où U est un ensemble et d est une fonction de distance $d : U \times U \rightarrow [0, \infty)$

Comme le montre la figure [5.1] une mesure de distance peut être transformée en mesure de similarité. La transformation précise à être utiliser fait l'objet d'une discussion. Si $d \in [0, \infty)$ alors (5.1) est la transformation approprié. Si la distance est normalisée tel que $d \in [0, 1]$, alors $s = 1 - d$ semble approprié. Dans ce qui suit, les mesures de distances sont telles qu'elles sont appliquées aux ensembles flous.

5.1.1 Méthode métrique de Minkowski :

Une classe particulière (et même générale) de fonctions de distance est connue comme métrique de *Minkowski*. Cette métrique est une classe des fonctions de distance. La forme générale est définie comme suit :

$$d_r(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^r \right)^{\frac{1}{r}} ; r \geq 1 \quad (5.2)$$

Où x et y sont deux points dans un espace de dimension n avec les composantes $(x_i, y_i) ; i = 1, \dots, n$. ce modèle de distance est un modèle de distance géométrique et il se base sur les distances entre les points ou les éléments. Les modèles de distance géométrique représentent les objets comme étant des points dans quelques espaces de coordonnées, tel que la distance observée parmi les objets correspond à la distance métrique entre les points respectifs. Dans la théorie des ensembles ordinaires, la distance entre deux sous-ensembles X et Y d'un univers de discours U est souvent basée sur la distance géométrique entre les éléments x et y . Un exemple de ceci est vu dans la figure 5.1, illustrant la distance entre deux ensembles $d(x, y) = \inf\{d(x, y), x \in X \text{ et } y \in Y\}$.

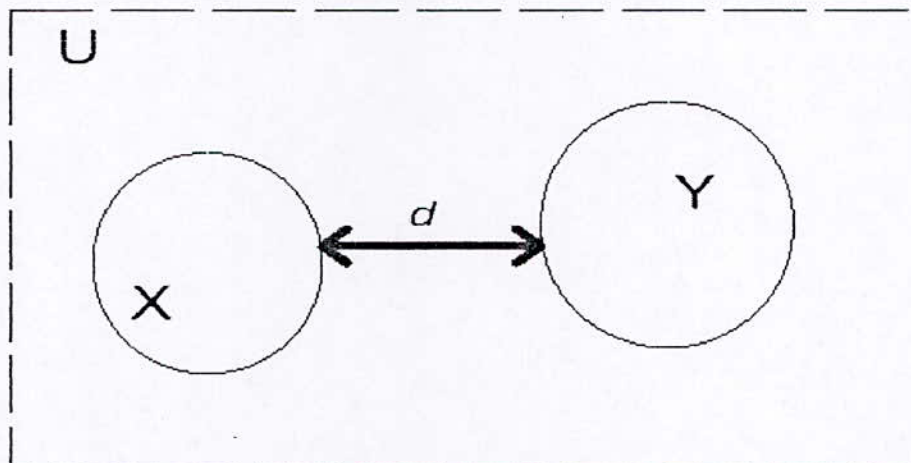


Figure 5.1 : Exemple de distance dans la théorie des ensembles

Bien sur le modèle de distance de *Minkowski* peut être généralisé aux ensembles flous. Donnant A et B comme deux sous-ensembles flous de l'ensemble $X = \{x_1, x_2, \dots, x_n\}$ qui est l'univers du discours. Ceci donne une classe d'un paramètre de mesure de distance entre A et B :

$$d_r(A, B) = \left(\sum_{i=1}^n |\mu_A(x_i) - \mu_B(x_i)|^r \right)^{\frac{1}{r}} ; \quad r \geq 1 \quad (5.3)$$

Où : $\mu_A(x_i)$, $x_i \in X$, est le degré d'appartenance de l'élément x_i à l'ensemble flou A et $\mu_B(x_i)$, $x_i \in X$, est le degré d'appartenance de l'élément x_i à l'ensemble flou B. si $X = \mathbb{R}$, le modèle de distance prend une forme continue :

$$d_r(A, B) = \left(\int_{-\infty}^{\infty} |\mu_A(x) - \mu_B(x)|^r \right)^{\frac{1}{r}} ; \quad r \geq 1 \quad (5.4)$$

Il est noté des équations plus haut Que la distance entre les ensembles flous est basée sur la différence dans le degré d'appartenance de x_i dans A et B. ceci est illustré dans le figure (5.2).

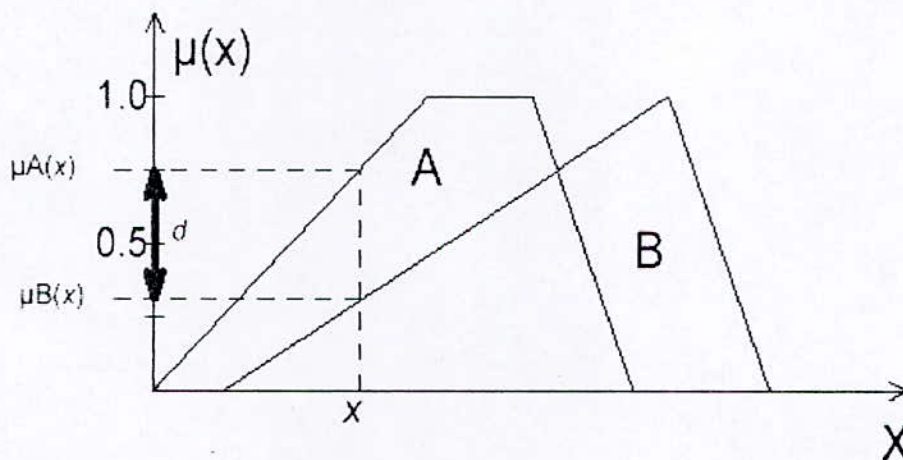


Figure 5.2 : Mesure de distance de Minkowski entre 2 ensembles flous basés sur la différence de leurs valeurs d'appartenance

La fonction distance dans (5.3) est aussi appelée la distance générale entre ensembles flous. Lorsque $r=2$, elle donne la fonction de distance euclidienne relative :

$$d_2(A, B) = \left(\sum_{i=1}^n |\mu_A(x_i) - \mu_B(x_i)|^2 \right)^{\frac{1}{2}} \quad (5.5)$$

5.1.2 Méthode métrique de Hausdorff :

On peut définir la distance entre deux sous-ensembles propres d'un espace métrique (U, d) en mettant $d(x, y) = \inf\{d_2(x, y) | x \in X, y \in Y\}$ où d_2 est la distance définie dans (5.2). Cette distance peut être utile dans plusieurs contextes, mais elle a le désavantage de distinguer que les ensembles en intersections ont toujours une distance nulle. Cependant, cette fonction de distance ne conduit pas aux mesures métriques pour des familles de sous-ensembles de U (c'est à dire qu'elle ne peut pas distinguer entre ensembles qui se chevauchent). La méthode métrique de Hausdorff tient compte de cela.

Définition 7 : Supposons que (U, d) est un espace métrique avec un diamètre fini, ayant une collection de sous-ensembles proches non vides de U .

Pour X et $Y \in H$ posons $d_x(y) = \sup_{y \in Y} d(y, x)$ et $d_y(x) = \sup_{x \in X} d(x, y)$

Alors $q: H \times H \rightarrow [0, \infty)$ défini par $q(x, y) = \max\{d_x(y), d_y(x)\}$ est la mesure métrique de Hausdorff pour H

Cette méthode est une généralisation de la distance entre deux points dans l'espace métrique de deux sous-ensembles non vides compactes. Si x et y sont de tels ensembles de nombres propres, alors la distance de Hausdorff est définie par :

$$q(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d_2(x, y), \sup_{y \in Y} \inf_{x \in X} d_2(x, y) \right\} \quad (5.6)$$

Dans le cas d'intervalle X et Y réels, la méthode de Hausdorff est décrite par

$$q(X, Y) = \max \{|x_1 - y_1|, |x_2 - y_2|\}$$

ou $X = [x_1, x_2]$ et $Y = [y_1, y_2]$

Ce n'est pas une mesure métrique unique qui étend la distance de Hausdorff à la théorie des ensembles flous. Ralescu et Ralescu[28] proposent les généralisations suivantes :

$$q_1(A, B) = \int_{\alpha=0}^1 q(A_\alpha, B_\alpha) d\alpha \quad (5.7)$$

$$q_\infty(A, B) = \sup_{\alpha > 0} q(A_\alpha, B_\alpha) \quad (5.8)$$

Où : A_α est la α -coupe de l'ensemble flou A (respectivement B_α à B). la mesure de distance q_1 est illustrée en figure [5.3]

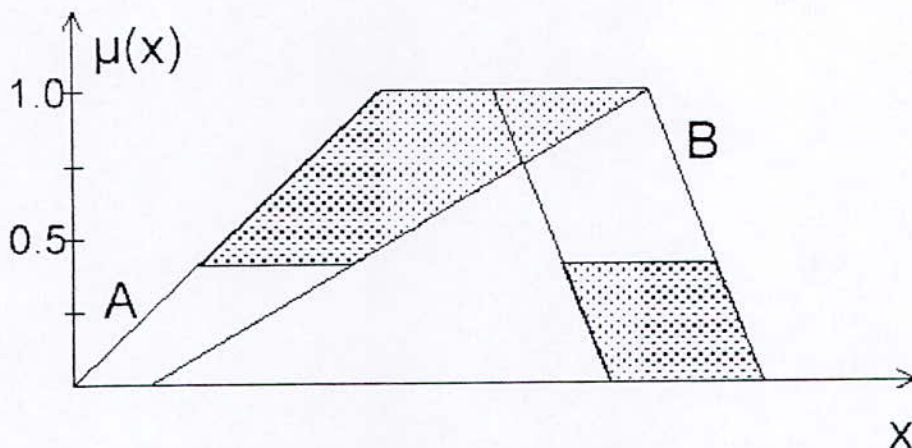


Figure 5.3 : distance de Hausdorff $q_1(A, B)$

Zwick et Al [27] Proposent la distance de Hausdorff entre ensembles flous propres tel que $\{x/\mu_A(x)=1\}$

$$q_*(A, B) = q(A_{1.0}, B_{1.0}) \tag{5.9}$$

où $A_{1.0}$ est l' α -coupe a $\alpha=1.0$ (c'est la 1-coupe) de l'ensemble flou A.

si A et B sont des intervalles propres, Alors $q_1(A, B) = q_\infty(A, B) = q_*(A, B) = q(A, B)$. figure 5.4

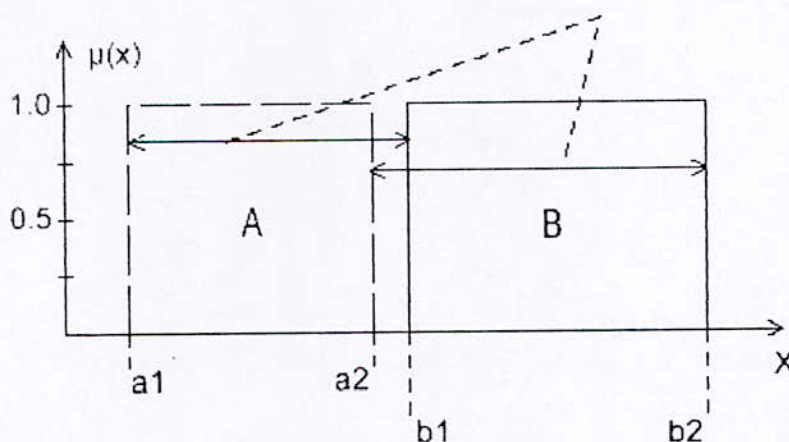


Figure 5.4 : Distance de Hausdorff quand A et B ont des intervalles propres.

5.1.3 La méthode métrique Goetschel et Voxman

La méthode métrique de Hausdorff est aussi utilisée par Goetschel et Voxman [7]. Ils redéfinissent les nombres flous et assignent les mesures métriques (basés sur la méthode métrique de Hausdorff) à cette classe d'ensembles flous. Dubois et Prade [13] ont introduit la notion de nombres flous (réels) et ont établis quelques propriétés de base de cette famille d'ensembles flous. Les définitions de Dubois et Prade d'un nombre flou sont comme suit:

Définition 8: Un nombre flou est un ensemble flou: $\mu : \mathbb{R} \rightarrow \mathbb{I}$ avec les propriétés:

1. μ est continu.
2. $\mu(x) = 0$ en dehors de quelques intervalle $[a; d]$.
3. les nombres b et c sont réels, on a : $a : a \leq b < c \leq d$, tel que μ augmente strictement sur $[a; b]$, et décroît strictement sur $[c; d]$, et $\mu(x) = 1$ pour chaque $x \in [b; c]$.

Donc, d'après cette définition, un nombre flou est un ensemble flou normalisé. Goetschel et Voxman changent cette définition légèrement en remplaçant les propriétés 1 et 3 par :

1' / μ est une fonction semi-continue supérieure.

3' / b et c sont des nombres réels, $a \leq b < c \leq d$ tel que μ est une fonction qui augmente sur l'intervalle $[a, b]$, et qui diminue sur l'intervalle $[c, d]$, avec $\mu(x) = 1$ pour chaque $x \in [b, c]$.

Goetschel et Voxman font ces changements pour deux raisons. En premier lieu, la définition de Dubois et Prade est quelque peu contradictoire depuis qu'ils autorisent la possibilité $a = d$ (le nombre flou résultant peut alors être associé à un nombre propre a), pendant que simultanément ils insistent sur la continuité de tous les nombres flous. Deuxièmement, la relaxation des propriétés 1 et 3 tient d'une définition de mesure métrique de la famille des nombres flous et permet l'étude de la topologie de cet espace métrique.

Soit A et B deux nombres flous, et posons $\text{supp}(A) = [a_1, a_2]$ (le support du nombre flou A) et le $\text{supp}(B) = [b_1, b_2]$, illustrés dans la Figure. [5.5]. Mettons $A = \min\{a_1, b_1\}$ et $b = \max\{a_2, b_2\}$ et les ensembles :

$$A^* = \{(x, y) | a \leq x \leq b, 0 \leq y \leq \mu_A(x)\}$$

$$B^* = \{(x, y) | a \leq x \leq b, 0 \leq y \leq \mu_B(x)\}$$

Alors la distance métrique définie par Goetschel et Voxman est:

$$Q(A, B) = q(A^*, B^*). \tag{5.10}$$

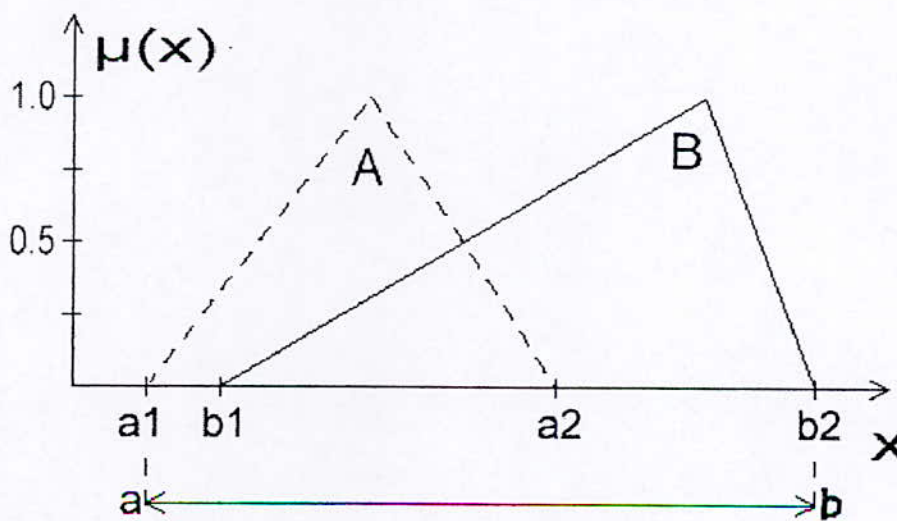


Figure 5.5 : Goetschel et Voxman prennent en considération le domaine

De la relation (5.9) nous remarquons que cette distance métrique est en fait celle de Hausdorff. Se conformant avec les définitions métriques de Hausdorff, et tenant compte de son usage, Goetschel et Voxman créent les ensembles propres de A^* et B^* , à partir des deux nombres flous A et B . La distance métrique de Hausdorff (5.6) est alors appliquée à ces deux ensembles et non aux nombres flous A et B directement. Elle calcule la distance de chaque point/élément $a^* \in A^*$ au point le plus proche $b^* \in B^*$ (Fig. 5.6), et la plus grande de ces distances est sélectionnée:

$$d_{A^*}(B^*) = \sup_{a^* \in A^*} \inf_{b^* \in B^*} d_2(a^*, b^*)$$

La même chose est faite pour chaque point de B . La plus grande de ces deux distances est donc sélectionnée comme une représentation de la distance entre les deux nombres flous A et B :

$$Q(A; B) = \max \{d_{A^*}(B^*), d_{B^*}(A^*)\}$$

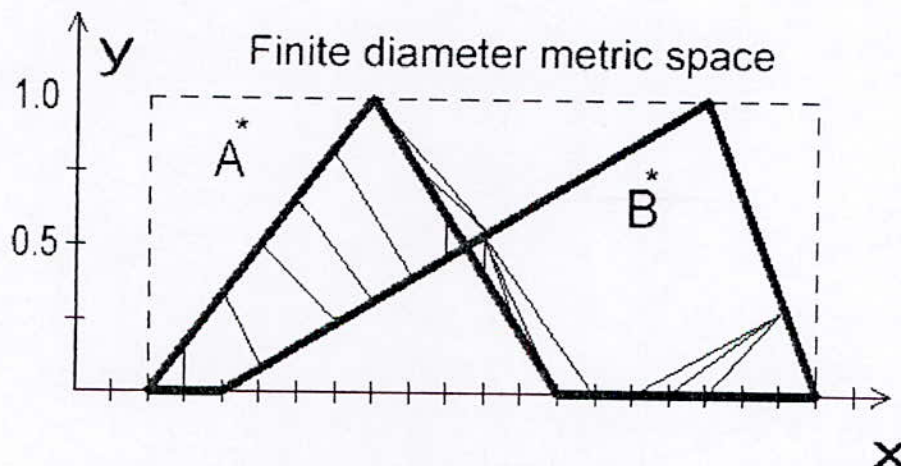


Figure 5.6 : Goetschel, Voxman appliquent la mesure de Hausdorff aux ensembles A^* et B^*

5.1.4 L'indice de Dissemblance :

Kaufman et Gupta [24] ont définis l'indice de dissemblance qui est une distance entre deux nombres flous. Leur concept de distance est légèrement différent du concept habituel de distance (métrique). Dans celui-ci ils changent la propriété de réflexivité :

$$d(x, y) = 0 - x = y$$

en

$$(X, Y) \Rightarrow d(X, Y) = 0$$

Kaufman et Gupta commencent avec une distance entre intervalles. Soit $A = [a_1, a_2]$ et $B = [b_1, b_2]$ deux les intervalles propres contenus dans $[\beta_1, \beta_2]$, i.e.

$\beta_1 \leq \min \{a_1, b_1\}$; $\beta_2 \leq \max\{a_2, b_2\}$ (Figure 5.7), et ont définis la distance entre ces intervalles comme :

$$\Delta(A, B) = \frac{|a_1 - b_1| + |a_2 - b_2|}{2(\beta_2 - \beta_1)} \quad (5.11)$$

Cette distance est normalisée, et $\Delta(A, B) \in [0; 1]$.

Alors, soient A et B deux nombres flous dans IR. Pour chaque niveau α (α -coupe) nous pouvons considérer $\Delta(A_\alpha, B_\alpha)$, où β_1 et β_2 sont donnés par toutes les valeurs convenables qui entourent A_α et B_α , pour tous les $\alpha \in [0; 1]$. Cependant, Kaufman et Gupta définissent l'indice de dissemblance pour les nombres flous comme étant:

$$\Delta_1(A, B) = \int_{\alpha=0}^1 \Delta(A_\alpha, B_\alpha) d\alpha \quad (5.12)$$

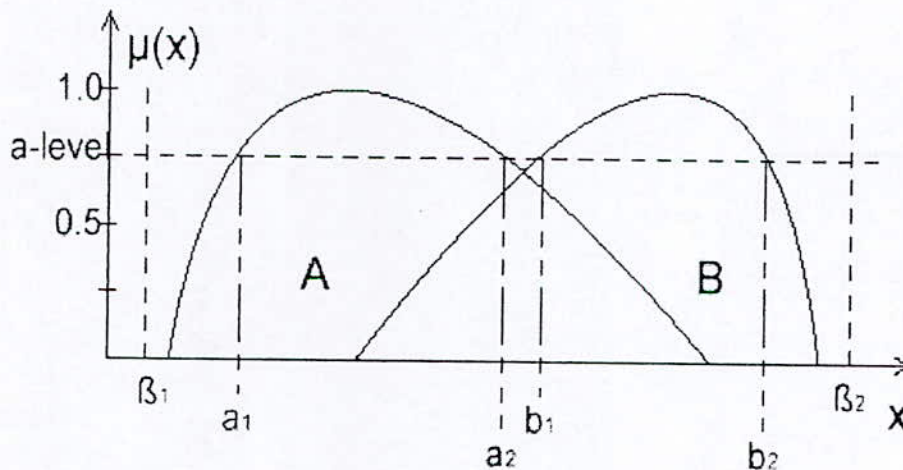


Figure 5.7 : l'indice de dissemblance est une intégration des distances calculés à chaque α -niveau

$$\Delta_\infty(A, B) = \sup_{\alpha > 0} \Delta(A_\alpha, B_\alpha) \quad (5.13)$$

$$\Delta^*(A, B) = \Delta(A_{1,0}, B_{1,0}) \quad (5.14)$$

5.1.5 Les caractéristiques basés sur les mesures de distance:

Bonissone [9] a proposé une approche pour mesurer la distance entre ensembles flous basée sur la distance entre les caractéristiques sélectionnés des ensembles. Bonissone avait l'approximation linguistique dans l'esprit, et dans cette approche, l'efficacité du processus d'approximation linguistique est d'une importance majeure. L'approche est composée de deux étapes.

Dans la première étape, l'ensemble des labels possibles (pour l'approximation linguistique) est resserré en utilisant une mesure de distance approximative qui, d'après Bonissone, a pour objectif de bien parfaire les ensembles flous qui sont

distants les uns des autres. Pour arriver à cela, chaque ensemble flou est représenté par un nombre limité de caractéristiques pour simplifier le calcul de la distance. Quatre caractéristiques ont été choisies par Bonissone:

1. La puissance de l'ensemble (l'air sous la courbe de la fonction d'appartenance ou la cardinalité scalaire (5.19).
2. Une mesure de fuzzification d'ensemble (entropie non probabilistique) définie comme [12]:

$$G(A) = \sum_{x \in X} g(\mu_A(x))$$

où $g(\mu_A(x)) = -\mu_A(x) \ln[\mu_A(x)] - (1 - \mu_A(x)) \ln[1 - \mu_A(x)]$.

3. Le premier moment de l'ensemble (centre de gravité de la fonction d'appartenance) est défini par :

$$FMO(A) = \frac{\sum_{x \in X} x \mu_A(x)}{\text{puissance}(A)}$$

4. Skewness, la quatrième caractéristique, est définie comme:

$$\text{skew}(A) = \sum_{x \in X} (x^4 - FMO(A))^3 \mu_A(x).$$

Avec ces quatre caractéristiques calculées, Bonissone définit la distance entre deux ensembles flous A et B, décrits par leurs vecteurs de caractéristiques respectives $V_{fb}(A) = [\text{Power}(A), G(A), FMO(A), \text{Skew}(A)]$ et $V_{fb}(B) = [\text{Power}(B), G(B), FMO(B), \text{Skew}(B)]$, comme des distances Euclidiennes entre ces deux vecteurs :

$$F_B = d_2(V_{fb}(A), V_{fb}(B)) \quad (5.15)$$

Après avoir resserré l'ensemble des labels possibles avec la mesure de distance approximative $F_B(A, B)$, on commence la seconde étape du processus de Bonissone.

Dans celle-ci, une distance de Bhattacharyya modifiée est calculée. Cette distance est donc significative pour bien différencier entre les ensembles qui sont proches les uns des autres. La distance Bhattacharyya est définie dans [23] comme:

$$R(A, B) = \left(1 - \int_{-\infty}^{+\infty} (\mu_A^*(x) \mu_B^*(x))^{\frac{1}{2}} dx \right)^{\frac{1}{2}} \quad (5.16)$$

Où la fonction d'appartenance a été normalisée : $\mu_A^*(x) = \frac{\mu_A(x)}{\text{puissance}(A)}$, resp $\mu_B^*(x)$

Wenstøp [47] a aussi proposé une approche de caractéristique basée sur une mesure de la distance pour les ensembles flous. Wenstøp utilisa les deux composantes du vecteurs de caractéristiques $V_{fw}(A) = [Power(A); FMO(A)]$ pour représenter un ensemble flou. La distance entre les deux ensembles flous A et B est définie comme la distance Euclidienne entre les deux vecteurs de caractéristiques correspondants :

$$F_w(A, B) = d_2(V_{fw}(A), V_{fw}(B)) \tag{5.17}$$

La figure [5.8] montre comment les deux ensembles flous A et B sont représentés par deux points dans un système de coordonnées FMO-Power pour l'approche de Wenstøps

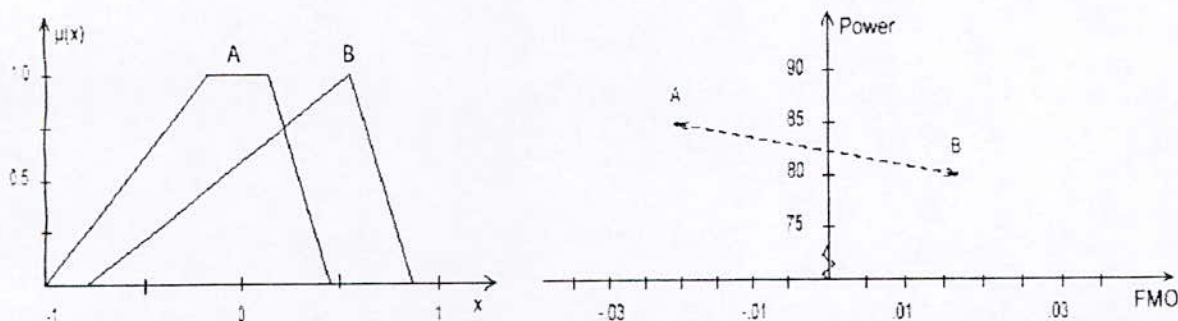


Figure 5.8 : montre comment les 2 ensembles flous A et B sont représentés par 2 points dans un système de coordonnées FMO-Power pour l'approche de wenstøps

5.2 Les mesures de similarités [27], [28], [30]

Les mesures de similarités entre ensembles flous peuvent aussi être définies sans l'usage explicite d'une fonction de distance. de telles mesures sont également connues dans les différentes recherches. Cependant, il n'y a pas beaucoup de mesures de similarité de ce type rapportées, comme est le cas pour celles qui sont transformées par une mesure de distance.

5.2.1 Neighbourhoodness

Dans le contexte des problèmes de prise de décision floue, Chen [28] a considéré la similarité entre les ensembles flous et propose une mesure du Neighbourhoodness de ce qu'il appelle 'vecteurs flous' :

$$M = \frac{V_1 V_2}{\max\{V_1 V_1, V_2 V_2\}}$$

où V_1 et V_2 sont des vecteurs, qui contiendront les valeurs d'appartenance correspondantes pour tout $x \in X$ (X est l'univers de discours). Donnant X fini, et dénotant chaque $x \in X$ comme x_i ; $i \in [1, 2, \dots, n]$. Le vecteur $V_1(A)$ représentant l'ensemble flou A devient $[A(x_1), A(x_2), A(x_3), \dots, A(x_n)]$ et son équivalent pour B est

$V_2(B) = [B(x_1), B(x_2), B(x_3), \dots, B(x_n)]$. La mesure proposée par Chen peut être alors transcrite comme suit :

$$M(A, B) = \frac{\sum_{i=1}^n \mu_A(x_i) \mu_B(x_i)}{\max \left\{ \sum_{i=1}^n \mu_A(x_i)^2, \sum_{i=1}^n \mu_B(x_i)^2 \right\}} \quad (5.18)$$

$M(A;B)$ est une fonction qui représente le rapprochement, où le Neighbourhoodness, des deux ensembles flous A et B. $M(A; B)$ prend les valeurs dans $[0, 1]$, où $M(A; B) = 1$ indique l'égalité des ensembles flous (identiques).

La figure. [5.9] montre la mesure de similarité de Chen appliquée aux deux ensembles flous A et B ($M(A; B) = 0.176$ pour ce cas).

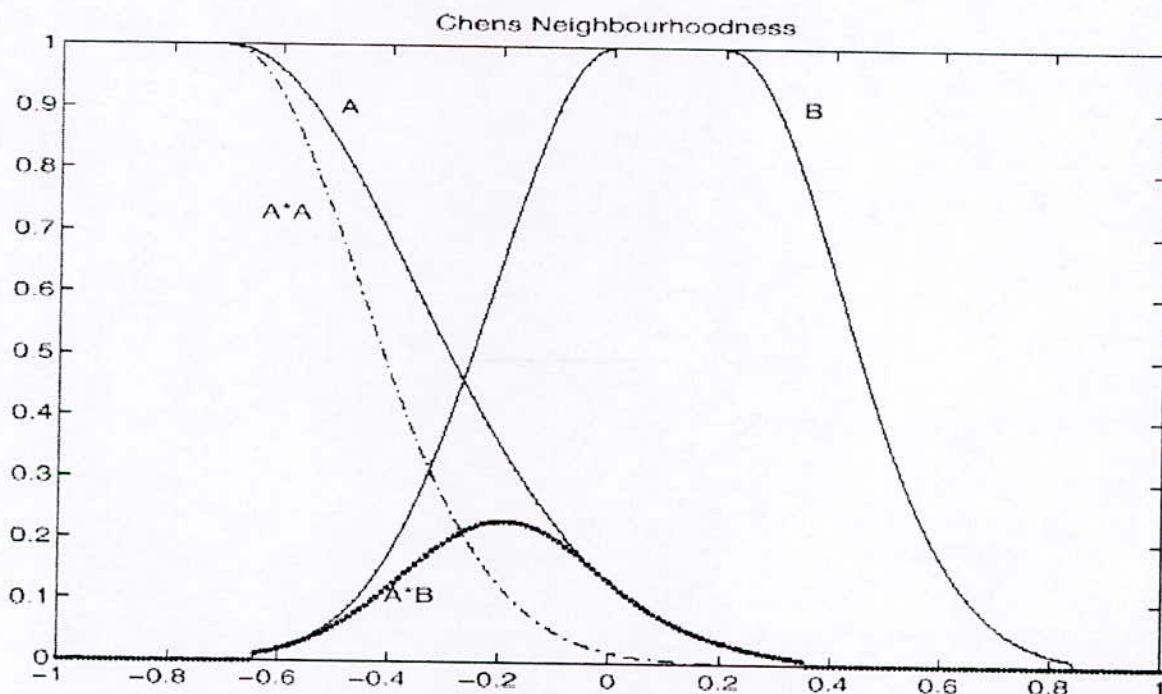


Figure 5.9 : mesure de Neighbourhoodness

5.2.2 Les mesures de similarité dérivées de la théorie des ensembles classiques:

Quelques mesures de similarité proposées pour les ensembles flous peuvent être considérées comme généralisations des fonctions d'égalités théoriques des ensembles classiques. Ces généralisations se basent essentiellement sur les définitions de cardinalité et de différence dans la théorie des ensembles flous. Les mesures de similarité présentées ici sont proposées par Dubois et Prade [14], et utilisent la cardinalité scalaire (puissance) des ensembles flous. La cardinalité scalaire d'un sous-ensemble flou A d'un ensemble universel fini X est définie comme:

$$|A| = \sum_{x \in X} \mu_A(x), \quad (5.19)$$

Quand le nombre d'éléments dans l'univers X n'est pas fini, $|A|$ n'existe pas toujours. Cependant, si l'ensemble A a un support fini,

$$|A| = \sum_{x \in \text{Supp}(A)} \mu_A(x) \quad , \text{de cela nous aurons la cardinalité relative :}$$

$$\|A\| = \frac{|A|}{|X|}. \tag{5.20}$$

De plus, nous utilisons les opérateurs théoriques des ensembles, tel que l'intersection et l'union entre ensembles flous sont définis comme les opérateurs max et min

$$\begin{aligned} \mu_{A \cap B}(x) &= \min\{\mu_A(x), \mu_B(x)\}, \quad \forall x \in X, \\ \mu_{A \cup B}(x) &= \max\{\mu_A(x), \mu_B(x)\}, \quad \forall x \in X. \end{aligned}$$

Comme différence, la différence symétrique est utilisée

Définition 9 : La différence symétrique notée $A \boxplus B$ est l'ensemble flou des éléments qui appartiennent à A et non à B ou bien à B et non à A :

$$\mu_{A \boxplus B} = \max(\min(\mu_A(x), 1 - \mu_B(x)), \min(1 - \mu_A(x), \mu_B(x))), \quad \forall x \in X. \tag{5.21}$$

La figure 5.10 illustre l'ensemble flou 'différence symétrique entre A et B ' (notons que ce n'est pas un ensemble flou convexe).

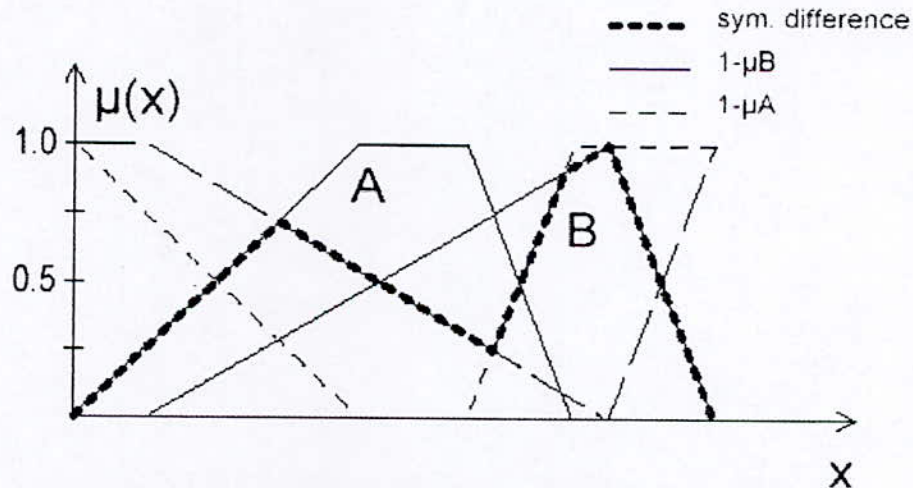


Figure 5.10 : différence symétrique entre A et B

Les mesures de similarité suivantes entre les ensembles flous A et B ont été proposées dans la documentation [28] :

- Basées sur l'intersection, l'union et la cardinalité :

$$S_1(A, B) = \frac{\|A \cap B\|}{\|A \cup B\|} \tag{5.22}$$

si $A=B$ Alors $S_1(A, B) = 1$.

- Basées sur la différence et la cardinalité :

$$S_2(A, B) = 1 - |A \ominus B| \quad (5.23)$$

Si $A \leftrightarrow B$ Alors $S_2(A, B) \geq 1/2$.

- Basées sur la différence :

$$S_3(A, B) = \sup_{x \in X} \mu_{A \ominus B}(x) \quad (5.24)$$

$A \leftrightarrow B$ SI et toujours Si $S_3(A, B) \geq 1/2$.

- L'indice de consistance :

$$S_4(A, B) = \sup_{x \in X} \mu_{A \cap B}(x) \quad (5.25)$$

$S_4(A, B) = 0$ veut dire que A et B sont séparés, par contre $S_4(A, B) = 1$ signifie qu'il est possible de trouver un élément $x \in X$ qui appartient totalement aux deux ensembles A et B.

L'indice de consistance est illustré dans la figure 5.11, $1 - S_4(A, B)$ est quelque fois utilisé comme un indice de séparation entre les deux ensembles flous A et B.

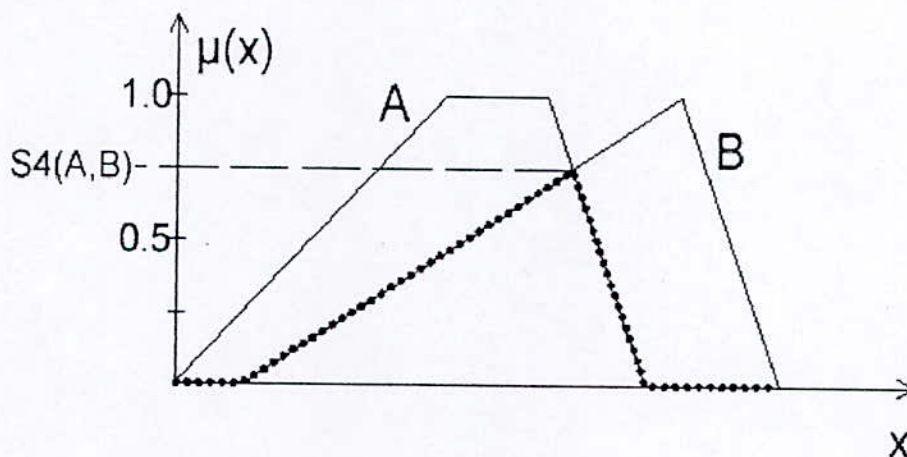


Figure 5.11 : L'indice de consistance

5.2.3 L'indice d'égalité :

Pedrycz a proposé une approche pour comparer les ensembles flous d'une manière logique [23]. Dans la théorie classique des ensembles, l'égalité de deux ensembles propres A et B peut être étudiée, en vérifiant que les deux conditions suivantes soient simultanément satisfaites :

- A est inclus dans B, c'est à dire $A \subset B$

- B est inclus dans A, c'est à dire $B \subset A$

Si une différence est observé, ces deux ensembles sont considérés inégaux. Dans la théorie des ensembles flous, ce cas peut être exprimé en notation logique comme suit :

$$A = B, \text{ Si } A \subset B \text{ et } B \subset A$$

Si nous modélisons l'inclusion \subset par un opérateur d'implication (ϕ) et la conjonction 'Et' par la t-norme correspondante, nous pouvant définir le degré pour lequel A et B sont égaux au point $x \in X$, comme :

$$(A \equiv B)(x) = [A(x) \phi B(x)] t [B(x) \phi A(x)]$$

Cet indice d'égalité n'est pas symétrique, et est modifié par addition d'une condition reliant les compléments de A et B :

$$\bar{A} \subset \bar{B} \text{ et } \bar{B} \subset \bar{A}$$

L'indice d'égalité peut alors être défini comme une moyenne de ces deux conditions :

$$(A \equiv B)(x) = \frac{1}{2} ([A(x) \phi B(x)] t [B(x) \phi A(x)] + [\bar{A}(x) \phi \bar{B}(x)] t [\bar{B}(x) \phi \bar{A}(x)]) \quad (5.26)$$

La moyenne est utilisée pour obtenir $(A \equiv B)(x) \in [0; 1]$, nous aurions $(A \equiv B)(x) \in [0; 2]$, du moment que les deux éléments dans la somme contribuent avec une valeur dans $[0,1]$.) Dans (5.25) la première partie de la somme décrit le degré d'égalité (similarité) entre chaque composante $A(x)$ et $B(x)$, ou un minimum de degré pour lequel $A(x)$ implique $B(x)$ et vice versa. De la même façon, la dernière partie dénote le degré d'égalité entre compléments des valeurs observées en $A(x)$ et $B(x)$. L'opérateur de l'implication (' ϕ -opérateur) qui modèle l'inclusion peut être envisagé comme un niveau d'inclusion d'ensemble flou A dans B au point x.

Le ϕ -opérateur est le *pseudocomplément*, associé avec la conjonction t-norme utilisée pour modéliser le 'Et' et est défini comme suit :

$$A(x) \phi B(x) = \sup\{c \in [0, 1] / A(x) t c \leq B(x)\} \quad (5.27)$$

L'opérateur d'implication utilisé dans l'indice d'égalité est donc induit par la t-norme utilisée dans cet indice, c'est à dire que les t-normes dans (5.25) et (5.26) sont identiques. La forme spécifique de l'indice d'égalité donc variera avec la t-norme choisie pour l'opérateur 'et'. Par exemple, quand le produit algébrique est utilisé pour l'opération 'et', le *pseudocomplément* (' ϕ -opérateur) devient:

$$A(x) \phi B(x) = \begin{cases} 1 & \text{pour } A(x) \leq B(x) \\ \frac{B(x)}{A(x)} & \text{pour } A(x) > B(x) \end{cases} \quad (5.28)$$

Et donc l'indice d'égalité devient :

$$(A(x) \equiv B(x))(x) = \begin{cases} \frac{1}{2} \left(\frac{A(x)}{B(x)} + \frac{1-B(x)}{1-A(x)} \right) \text{ pour } A(x) < B(x) \\ 1 \text{ pour } A(x) = B(x) \\ \frac{1}{2} \left(\frac{B(x)}{A(x)} + \frac{1-A(x)}{1-B(x)} \right) \text{ pour } A(x) > B(x) \end{cases} \quad (5.29)$$

Quand la t-norme utilisée est l'opérateur 'Et' de Lukasiewicz, l'indice d'égalité prend une forme linéaire :

$$(A(x) \equiv B(x))(x) = \begin{cases} 1 - B(x) + A(x) \text{ pour } A(x) < B(x) \\ 1 \text{ pour } A(x) = B(x) \\ 1 - A(x) + B(x) \text{ pour } A(x) > B(x) \end{cases} \quad (5.30)$$

L'indice d'égalité ressort la similarité d'un seul élément $x_i \in X$. La mesure total exprime la similarité de A et B qui devrait renvoyer toutes les valeurs obtenues pour les éléments individuels dans X. Cela est accompli en agrégeant les valeurs $(A \equiv B)(x_i)$ sur tous les éléments $x_i \in X$. Une des méthodes à utiliser est de calculer la moyenne:

$$(A(x) \equiv B(x))_{av} = \frac{1}{2} \sum_{i=1}^n (A \equiv B)(x_i), \quad (5.31)$$

Où n est le nombre d'éléments dans X.

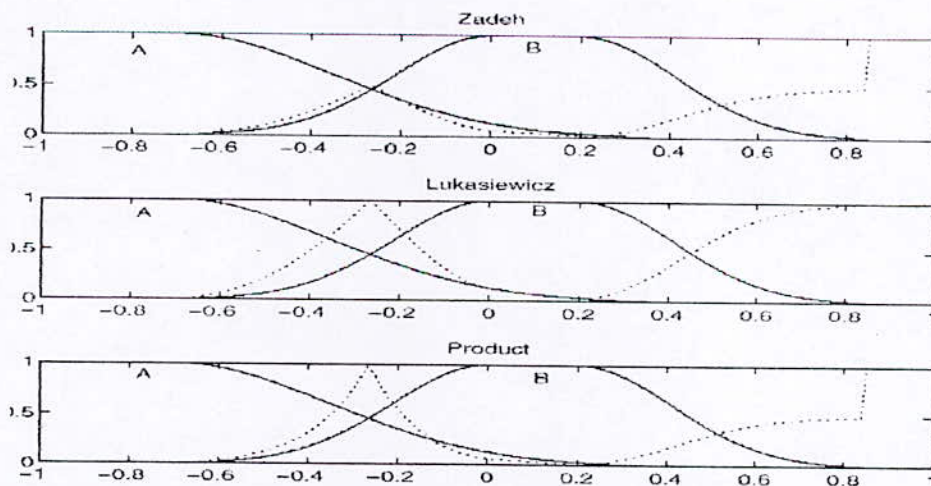


Figure 5.12 : L'indice d'égalité appliqué aux deux ensembles flous A et B utilisant les trois différents opérateurs

Deux autres méthodes d'agrégation bien connues sont l'agrégation optimiste et l'agrégation pessimiste:

$$(A \equiv B)(x)_{opt} = \sup\{(A \equiv B)(x)\},$$

$$(A \equiv B)(x)_{pes} = \inf\{(A \equiv B)(x)\}.$$
(5.32) et (5.33)

La figure 5.12 illustre l'indice d'égalité appliqué aux deux ensembles flous A et B utilisant trois différents opérateurs 'Et' : Zadeh, Lukasiewicz et le produit algébrique. Si nous utilisons l'agrégation de moyenne (3.38), l'indice d'égalité prend les valeurs 0.21, 0.41 et 0.24 pour les trois cas illustrés dans cette figure

5.2.4 Corrélation entre deux fonctions d'appartenance :

Murthy, Pal, et Majumder ont définis une mesure de rapport entre fonctions d'appartenance floues qu'ils appellent 'corrélacion' [30]. Dans Fig. 5.13 cinq fonctions d'appartenance décrivent la hauteur d'un homme. En général, si quelqu'un est 'grand', alors la valeur de 'très grand' n'est pas vraiment basse, alors que la valeur de 'petit' est basse. Donc la corrélation entre 'grand' et 'très grand' est plus grande que celle entre 'grand' et 'petit', la même chose est appliqué à la similarité.

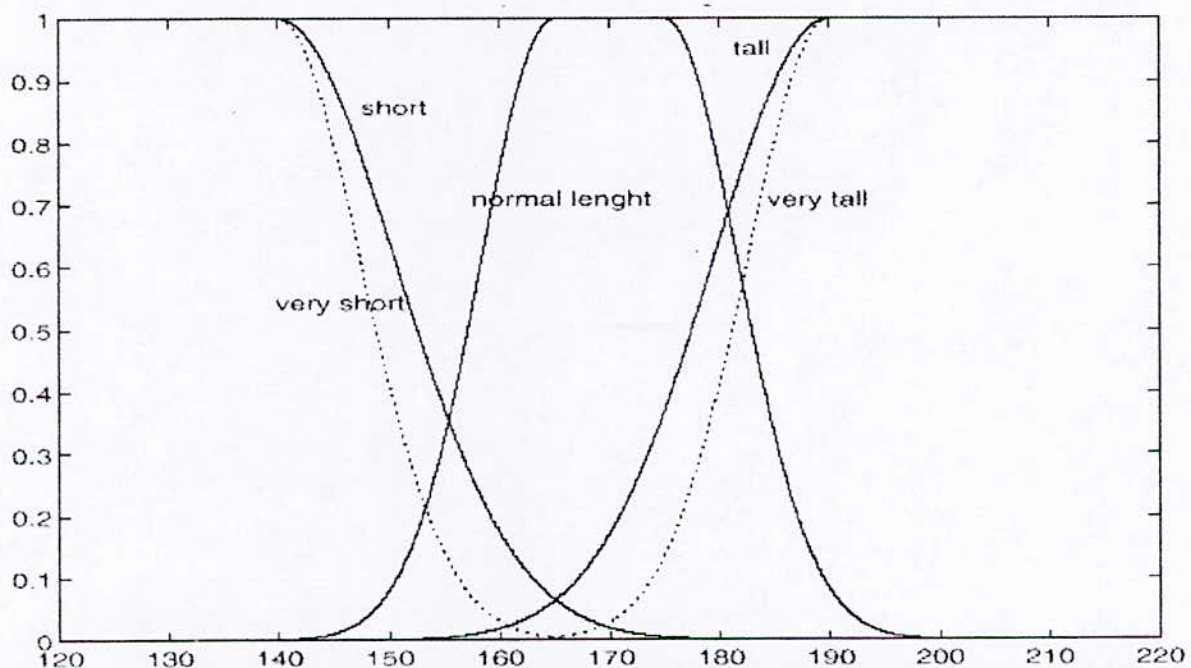


Figure 5.13 : fonctions d'appartenance décrivant la taille d'un homme

Dans les statistiques, le coefficient de corrélation r_{xy} entre les deux variables x et y est calculé avec l'usage de la moyenne des deux variables. Pour des variables X et Y aléatoires le coefficient corrélation est défini comme [29]

$$\rho_{xy} = \frac{COV(X,Y)}{\sigma_x \sigma_y} = \frac{E[(X - E[X])(Y - E[Y])]}{\sigma_x \sigma_y}$$

Où E dénote la moyenne, et σ_x et σ_y sont respectivement les déviations standards de X et Y . Lorsque nous développons une expression pour la corrélation entre fonctions d'appartenance floues, les principes appliqués dans les statistiques ne peuvent pas être pris en considération. Cependant, les concepts de "moyen" et "déviacion standard" ne sont pas directement applicables aux fonctions

d'appartenance floue. La corrélation entre fonctions d'appartenance doit être définie quelque peu différemment que dans les statistiques.

Définition 10: Soit Ω un intervalle fermé dans \mathbb{R} . soit A et B deux fonctions d'appartenance continue dans cet intervalle. on a :

$$X_A = \int_{\Omega} [\mu_A(x) - (1 - \mu_A(x))]^2 dx = \int_{\Omega} (2\mu_A(x) - 1)^2 dx$$

$$X_B = \int_{\Omega} [\mu_B(x) - (1 - \mu_B(x))]^2 dx = \int_{\Omega} (2\mu_B(x) - 1)^2 dx$$

La corrélation entre les deux fonctions d'appartenance μ_A et μ_B peut alors être transcrite comme suit :

$$C_{\mu_A, \mu_B} = 1 - \frac{4}{X_A + X_B} \int_{\Omega} (\mu_A(x) - \mu_B(x))^2 dx \quad (5.34)$$

C_{μ_A, μ_B} prend des valeurs dans $[-1, 1]$, et devient 1 quand μ_A et μ_B sont identiques et -1 quand μ_A et μ_B se complètent. Le coefficient de corrélation est une mesure de rapprochement de μ_A et μ_B mais pas exactement une mesure de distance. En donnant μ_A, μ_B, μ_C et μ_D , $C_{\mu_A, \mu_B} = C_{\mu_C, \mu_D}$ cela signifie que la différence entre μ_A et μ_B comparés avec (divisé par) la somme des différences entre μ_A et $1-\mu_A$ et μ_B et $1-\mu_B$ est la même que la différence entre μ_C et μ_D comparés avec la somme des différences entre μ_C et $1-\mu_C$ et μ_D et $1-\mu_D$.

Conclusion

Dans ce chapitre, plusieurs méthodes de mesures pour ressortir la distance ou la similarité entre ensembles flous sont présentées et étudiées du point de vue mathématique. En générale, on peut dire que les mesures de distances peuvent être transformées en mesures de similarité en utilisant une transformation appropriée. Notons que ces mesures ne sont pas toutes conformes aux propriétés souhaitables cités dans le chapitre 4, lesquelles sont proposées du point de vue de la simplification des règles de base. Ces mesures de similarité seront étudiées et évaluées dans le chapitre suivant.

CHAPITRE 6

Etudes et Evaluations des Mesures de Similarités

Etudes et Evaluation des mesures de similarité

Dans ce chapitre, les différentes méthodes présentées dans le chapitre précédent seront étudiées et évaluées du point de vue de leurs capacités à mesurer la similarité entre ensembles flous. L'objectif global est de trouver une mesure qui peut être utilisée dans le contexte de la simplification des règles de base et qui influencera l'évaluation. Les mesures de distance rapportées dans la section 5.2 seront transformées en mesures de similarité utilisant les transformations appropriées.

Les modèles géométriques de similarité ont prévalu dans l'analyse théorique de similarité entre ensembles flous (chapitre 5). Ces modèles représentent les objets comme étant des points dans quelques coordonnées de l'espace. Une courte distance correspond aux ensembles semblables, pendant qu'une plus grande indique la dissemblance. De tels mesures d'objets exposent toutes les propriétés d'une distance (ou mesure métrique) donnée dans la définition 5. Cependant, quelques psychologues [28] estiment que la similarité ne peut être une relation symétrique et que tous les stimulus ne peuvent pas être également semblables à eux mêmes. Cela suggère que la similarité peut être ressortie par une fonction qui n'est pas conceptuellement une distance géométrique, mais une opération d'ensembles théorique.

Dans la plupart des documentations sur les ensembles flous (par exemple [2], [3], [8], [9]) les mesures de distance et de similarité étudiées sont appliquées aux ensembles flous avec un haut degré de chevauchement. Mesurer la similarité entre deux fonctions d'appartenance dissemblables (qui ne chevauchent pas) paraît avoir peu d'intérêt. Les ensembles flous considérés dans la simplification des règles de base sont créés par projections de groupes dans un espace multidimensionnelle sur les axes de cet espace. Dans un processus automatisé, la méthode à utiliser pour les considérations de similarité peut être appliqué aux fonctions dissemblables aussi bien qu'aux fonctions d'appartenance semblables, voir Figure 6.1. L'étude des méthodes présentées dans le chapitre 5 démontrera que certaines de ces méthodes conviennent beaucoup plus pour mesurer la similarité entre les paires dissemblables que d'autres qui conviennent pour les paires semblables des ensembles flous.

6.1 Grouper ces fonctions de similarité [3], [27], [28]

En général, les méthodes pour mesurer la similarité entre ensembles flous peuvent être divisées en deux groupes principaux, qui dépendent de la manière utilisée pour ressortir la similarité:

1. Modèles de similarité géométrique
2. Modèles de similarité d'ensemble théorique.

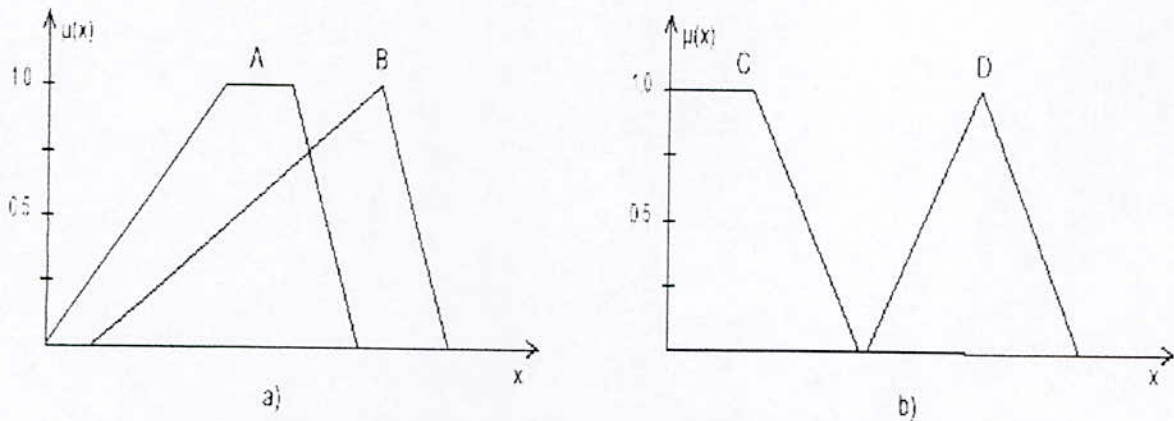


Figure 6.1 : a) fonctions d'appartenance similaire, et b) fonctions d'appartenance dissimilaire

Les modèles de similarité géométriques sont tirés de, ou dans un sens font usage de, fonctions de distance géométrique. La plupart de ces mesures qui appartiennent à ce groupe sont des mesures qui expliquent la distance entre les ensembles flous sur l'axe x (leur place relative dans l'univers de discours), comme la mesure métrique de Goetschel et Voxman (5.9), ou la représentation basée sur les mesures de similarité. Cette dernière est considérée comme un sous-groupe de la similarité géométrique et est trouvée à partir de la distance géométrique entre les points qui représentent les ensembles flous dans cette espace.

Les modèles Géométrique de similarité sont inscrits dans la Table 6.1.

Table 6.1 : modèles de similarité géométriques

Mesures de similarité	Basée sur	Equation
Mesure Euclidienne	L'appartenance	5.5
Mesure de dominance	L'appartenance	A.2
Mesure city-block	L'appartenance	A.3
Mesure de Kacprzyk	L'appartenance	
Mesure de Hausdorff 1	La position	5.7
Mesure de Hausdorff ∞	La position	5.8
Mesure de Hausdorff *	La position	5.9
Mesure de goetschel voxman	Position et appartenance	5.10
Mesure de dissemblance 1	La position	5.12
Mesure de dissemblance ∞	La position	5.13
Mesure de dissemblance*	La position	5.14
Corrélation	L'appartenance	5.34
Mesure de Wenstop	Multicritère	5.17
Mesure de Bonissone	Multicritère	5.15

Les modèles de similarité d'ensembles théoriques utilisent des opérations théoriques (par exemple t-normes) pour modéliser la similarité entre les ensembles flous. Ces modèles ne peuvent pas distinguer entre paires d'ensembles flous dissemblable, c.-à-d. qu'ils sont insensibles aux variations des degrés de dissemblance. La raison est qu'ils ignorent leurs ordres sur l'axe x . Leurs avantages, cependant, est qu'ils sont clairement définis même pour des fonctions

d'appartenance sur un espace réduit. Les modèles de similarité d'ensembles théoriques sont inscrits dans la Table 6.2.

Table 6.2 : modèle de similarité basé sur les ensembles théorique

Mesures de similarité	Equation
Mesure de Bhattacharyya	5.16
Mesure du Neighbourhoodness	5.18
S_1	5.22
S_2	5.23
S_3	5.24
L'indice de consistance, S_4	5.25
L'indice d'égalité	5.26

6.2 Evaluations

Pour évaluer les performances des mesures de similarité, elles seront appliquées à quelques cas de tests synthétiques. Ces cas refléteront le comportement des mesures en cas d'inclusions, égalités, chevauchement, et non chevauchement des fonctions d'appartenance. La sensibilité des mesures de similarité pour la définition du domaine (c.-à-d. les valeurs numériques des variables du domaine) est aussi illustrée par ces tests.

6.2.1 les tests [27], [28], [29]

Trois cas de tests synthétiques ont été sélectionnés pour illustrer les propriétés des mesures de similarité.

Premier cas : Deux fonctions d'appartenance triangulaires égales sont placées dans un univers de discours normalisé $X = [-1,+1]$. Graduellement, un des ensembles est défini à droite du deuxième ensemble, avec un pas fixe de 0.2. A chaque position la similarité entre les deux ensembles est mesurée. Ce cas est illustré dans la Figure 6.2, laquelle montre toutes les positions étudiées des deux ensembles.

Deuxième cas : Ce cas est le même que le précédent (cas 1), à l'exception de l'échelle. Les deux fonctions d'appartenance triangulaires égales (même forme comme dans le cas 1) sont placées dans un domaine (univers de discours non normalisé) $X = [-10,+10]$. A chaque position la similarité entre les deux ensembles est mesurée. Ce cas est illustré dans Fig. 6.3, laquelle montre toutes les positions étudiées des deux ensembles.

Troisième cas: Les ensembles flous 'exponentiels' sont inclus l'un dans l'autre. Chacun d'eux est décalé à gauche, et est comparé avec ceux décalés à droite (voir le Fig. 6.4). La première comparaison est celle de la fonction exponentielle 1 avec elle-même, puis l'exponentiel 1 avec la 2 etc. jusqu'à l'exponentiel 17. Ce cas est illustré dans Fig. 6.4 laquelle montre toutes les fonctions exponentielles étudiées.

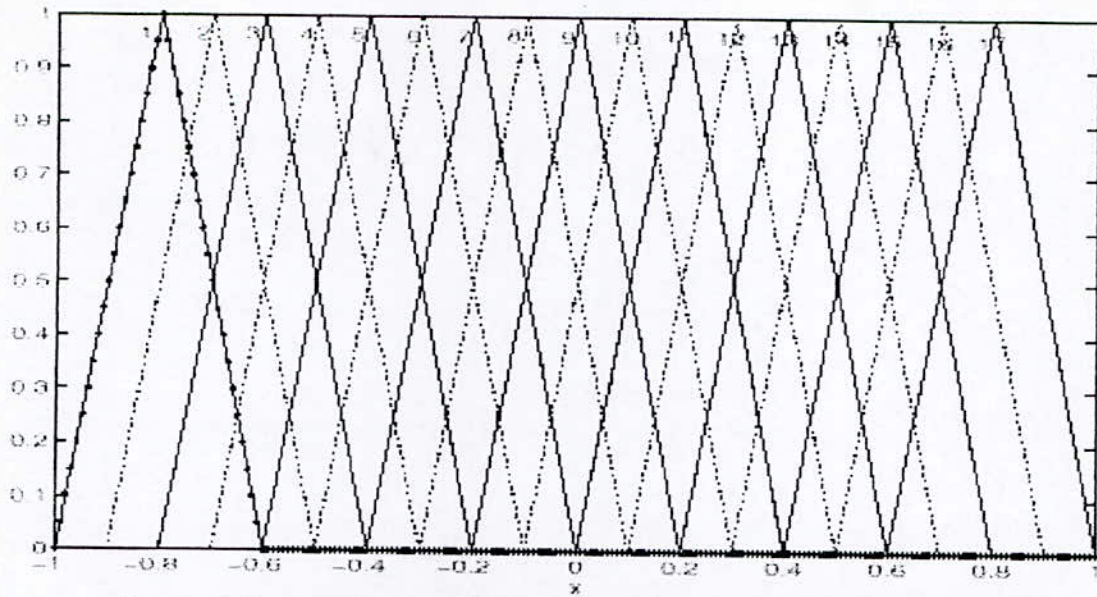


Figure 6.2 : cas 1, ensembles flous triangulaire dans un domaine normalisé

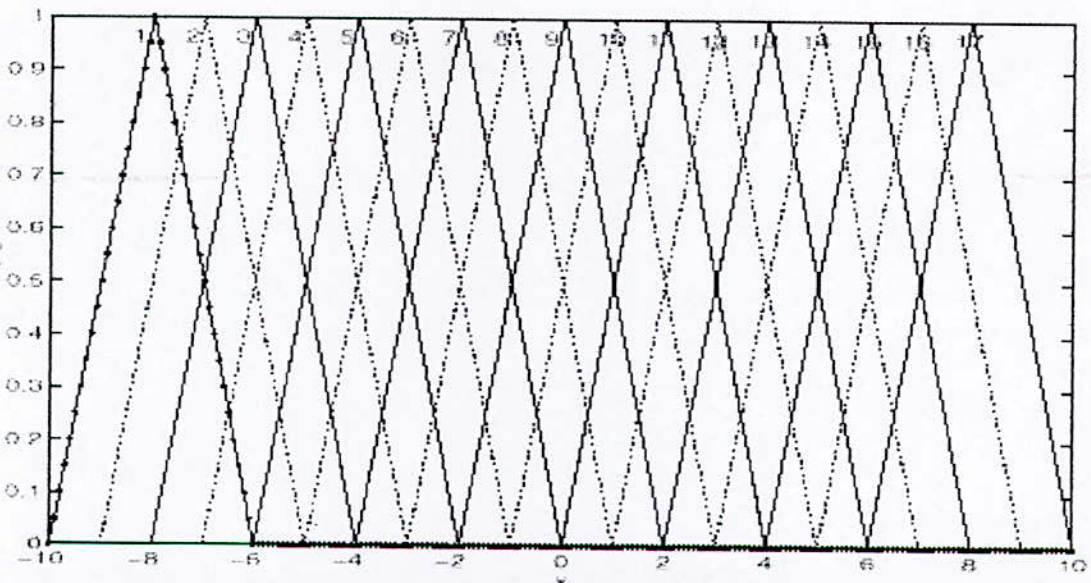


Figure 6.3 : cas 2, ensembles flous triangulaires dans un domaine non normalisé

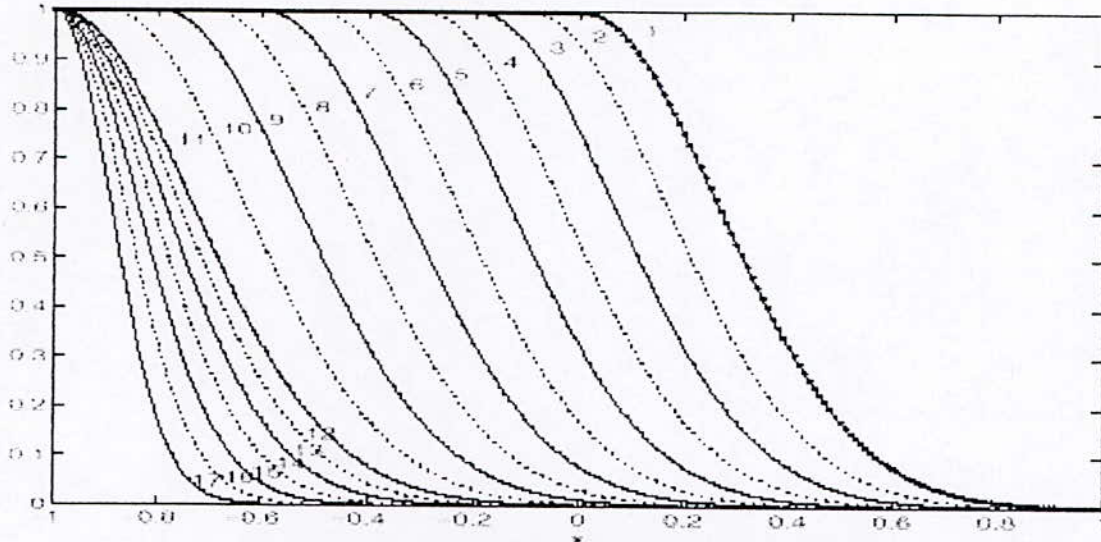


Figure [6.4] : cas 3, ensembles exponentiel.

6.2.2 Critères [28]

Toutes les mesures de similarité ont leurs propriétés spécifiques. Cependant, pas toutes ne sont bonnes du point de vue de la simplification des règles de base. Par exemple, la valeur de la similarité assignée entre deux fonctions d'appartenance ne devraient pas être basée uniquement sur leur valeurs d'appartenance pour un point du domaine, car cela peut conduire à une fusion de ces fonctions, lesquelles ont une grande valeur d'appartenance au point d'intersection, mais ne sont pas généralement très semblables. Aussi, comme affirmé dans le Chapitre 4, les ensembles flous qui se chevauchent devraient être considérés plus semblables que ceux qui ne le sont pas. Dans la simplification des règles de base, on s'intéresse essentiellement à la fusion des fonctions d'appartenance avec des supports de recouvrement, c.-à-d. des fonctions qui basculent la validation d'une règle pour plus ou moins les mêmes valeurs de la variable d'entrée. Une autre propriété souhaitable est que la mesure de similarité évalue $s \in [0; 1]$ où la valeur $s(A;B) = 1$ devrait indiquer que les deux ensembles flous A et B sont égaux.

Pour sélectionner une mesure de similarité adéquate pour la simplification des règles de base, quelques critères sont présentés. Mais nous tenons compte uniquement des mesures qui se conforment aux critères et que nous considérons comme convenables.

Critère 1 : La similarité $s(A, B)$ devrait être évaluée comme 1 si et seulement si A et B sont égaux :

$$S(A, B) = 1 - \mu_A(x) = \mu_B(x), \quad \forall x \in X.$$

Critère 2 : Si pour un point $x \in X$ nous avons $\mu_A(x) = 1$ et $\mu_B(x) = 0$, la similarité entre A et B ne devrait pas essentiellement être évaluée comme $s(A, B) = 0$:

$$\exists x \in X, \mu_A(x) - \mu_B(x) = 1 \Rightarrow s(A, B) = 0$$

Critère 3 : La similarité $s(A, B)$ entre les deux ensembles flous A et B ne devrait pas être influencée par l'échelle du domaine.

Critère 4 : La valeur de la similarité assignée aux paires d'ensembles flous désassimilées doit être plus basse que celle assignée aux paires d'ensembles flous similaires.

Avec ces critères, nous pouvons imaginer qu'une réponse optimale de la mesure de similarité pour les cas donnés ressemblerait à celle illustrée dans la Fig. 6.5

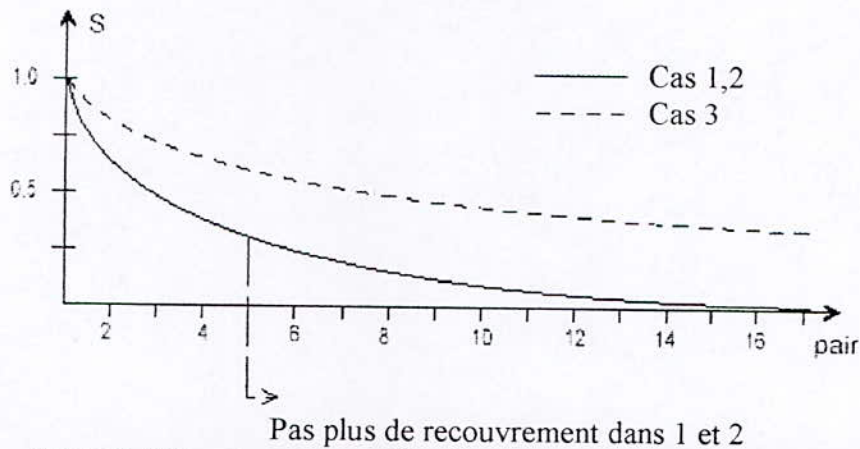


Figure 6.5 : réponse optimale des mesures de similarité

6.3 Les résultats de l'évaluation :

Les résultats donnés par les différentes mesures de similarité appliquées aux trois cas de tests abordés plus haut sont présentés dans cette section. Pour chaque mesure, trois réponses graphiques sont tracées, correspondant aux cas des tests donnés. Ces tracés peuvent être évalués en se basant sur les 4 critères présentés dans la section antérieure. A partir des tracés des cas 1 et 2, il est aisément noté comment la mesure de similarité manie paires d'ensembles flous dissemblables. Si elle prend une valeur constante après la 5^{ème} paire, elle ne différencie pas entre les différentes paires d'ensembles flous dissemblables.

Remarque : les résultats des autres mesures de similarité appliqués aux trois cas sont présentés dans l'annexe A.

6.3.1 Minkowski :

Les mesures de distance Euclidienne peuvent être transformées en mesure de similarité par la relation (5.1). Où $d \in [0, \infty)$ donnant ainsi $s \in (0, 1]$. Ces mesures de distances peuvent aussi être normalisées (définition 6), donnant $d_n \in [0, 1]$ Alors la transformation de similarité devient:

$$s_n(A, B) = 1 - d_n(A, B),$$

Avec $s_n \in [0, 1]$. Nous considérons la version normalisée, qui nous donne la mesure de similarité suivante pour la méthode de Minkowski (Euclidienne similarité) :

$$S_{d_2}(A, B) = 1 - \left(\frac{1}{n} \sum_{i=1}^n |\mu_A(x_i) - \mu_B(x_i)|^2 \right)^{1/2} \quad (6.1)$$

La mesure de similarité transformée par la mesure de distance Euclidienne (6.1) est appliquée aux trois cas présentés au-dessus, et les résultats sont donnés dans la Fig. 6.6.

Les mesures Minkowski ne sont pas capables de différencier entre paires d'ensembles flous qui ne chevauchent pas. Lorsqu'elle en est dotée elle prend une valeur dépendante de la forme des deux ensembles flous et est indépendante de leur positions relatives dans le domaine. Cela suppose qu'une paire de fonctions d'appartenance dissemblables peut avoir pour résultat une valeur de similarité plus grande que deux fonctions d'appartenance semblables.

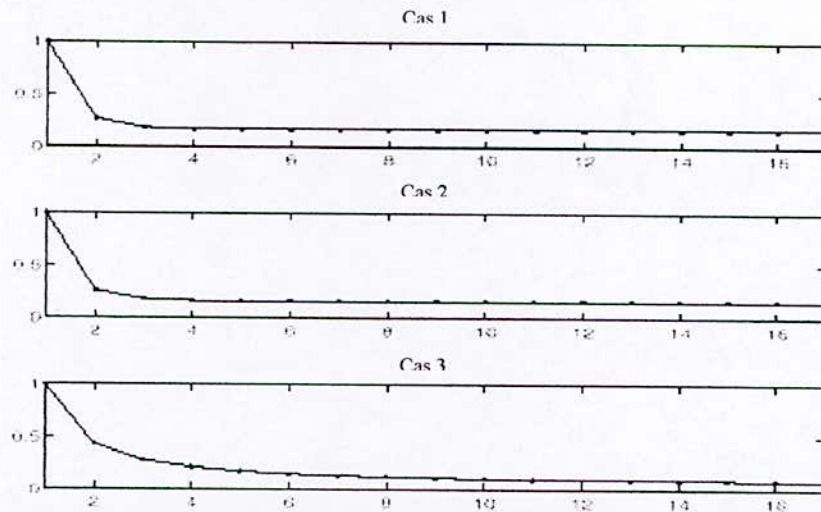


Figure 6.6 : résultats donnés par la mesure de similarité Euclidienne

6.3.2 Hausdorff :

Les mesures de distance de Hausdorff, q_1 ; q_∞ et q^* (respectivement l'équations 5.7, 5.8 et 5.9) pour les ensembles flous, mesurent la distance sur l'axe x , où ils prennent une valeur $d \in [0; 1)$. Leur transformation en similarité peut être alors donnée par (5.1), cela permet aux mesures de similarité suivantes à être identifiées par la transformation des trois mesures de distance de Hausdorff pour les ensembles flous:

$$S_{q_1}(A, B) = \frac{1}{1 + \int_{\alpha=0}^1 q(A_\alpha, B_\alpha) d\alpha} \quad (6.2)$$

$$S_{q_\infty}(A, B) = \frac{1}{1 + \sup_{\alpha > 0} q(A_\alpha, B_\alpha)} \quad (6.3)$$

$$S_{q^*}(A, B) = \frac{1}{1 + q(A_{1.0}, B_{1.0})} \quad (6.4)$$

Les résultats donnés par ces trois mesures de similarité appliquées aux trois cas de tests peuvent être vus dans la Figure. 6.7.

De ces résultats, il est évident que les mesures de Hausdorff sont influencées par la valeur numérique de la variable du domaine. Cependant, ils sont en mesure de distinguer entre différentes paires de fonctions d'appartenance qui ne chevauchent pas (c.-à-d. ils peuvent manier ensembles flous semblable aussi bien que les ensembles flous dissemblables). Les mesures de similarité Sq_1 et Sq_∞ considèrent la distance de Hausdorff entre α -coupes pour toutes les valeurs de $\alpha \in [0; 1]$. Sq_1 considère la distance de Hausdorff entre les 1-coupes seulement des ensembles flous. Cela la rend incapable de différencier entre les 7 dernières paires

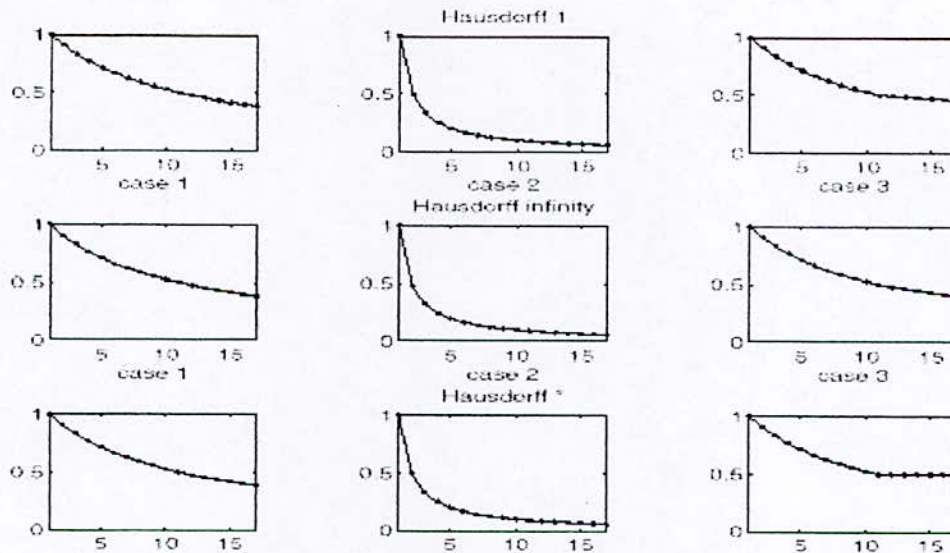


Figure 6.7 : résultats des mesures de Hausdorff

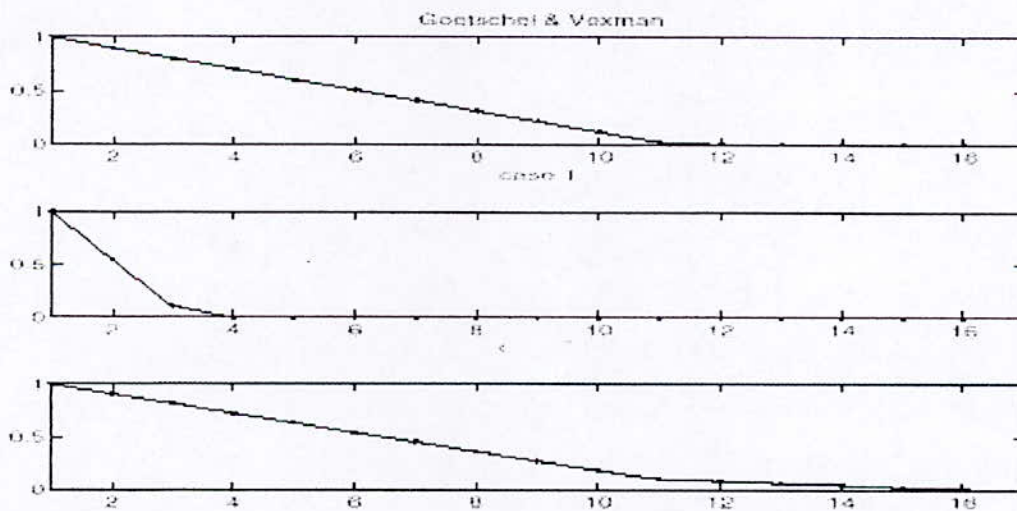


Figure 6.8 : résultats des mesures de Goetschel et Voxman

6.3.3 Goetschel et Voxman :

La mesure de distance proposée par Goetschel et Voxman (5.10) prend des valeurs $Q \in [0; 1]$, où 0 correspond aux ensembles égaux, et 1 est donné lorsque:

$$\text{Min}(\text{supp}(B)) - \text{max}(\text{supp}(A)) \geq 1.$$

C.-à-d. pour deux ensembles flous A et B qui ne chevauchent pas, dont la différence entre les deux membres les plus proches de leurs supports est ≥ 1 . Donc cette mesure de distance peut être transformée en mesure de similarité par $s = 1 - d$, $s \in [0, 1]$. La mesure de similarité transformé avec la méthode métrique de Groetchel et Voxman devient :

$$S_Q(A, B) = 1 - q(A^*, B^*). \tag{6.5}$$

La valeur $s = 1$ correspond aux ensembles flous égaux, et s devient égale à 0 pour des ensembles flous qui ne chevauchent pas. La figure 6.8 montre les résultats des trois cas de tests.

6.3.4 Indice de dissemblance :

Les trois versions de l'indice de Dissemblance rapportées dans la section 5.2 sont toutes normalisées, c.-à-d. qu'elles reviennent à une mesure de distance $d \in [0, 1]$, et la transformation en mesure de similarité est faite par $s = 1 - d$; $s \in [0, 1]$.

Les trois mesures de similarité sont données pas :

$$S_{\Delta_1}(A, B) = 1 - \int_{\alpha=1}^1 \Delta(A_\alpha, B_\alpha) d\alpha \tag{6.6}$$

$$S_{\Delta_\infty}(A, B) = 1 - \sup_{\alpha > 0} \Delta(A_\alpha, B_\alpha) \tag{6.7}$$

$$S_{\Delta^*}(A, B) = 1 - \Delta(A_{1.0}, B_{1.0}) \tag{6.8}$$

La figure 6.9 montre les résultats des trois mesures

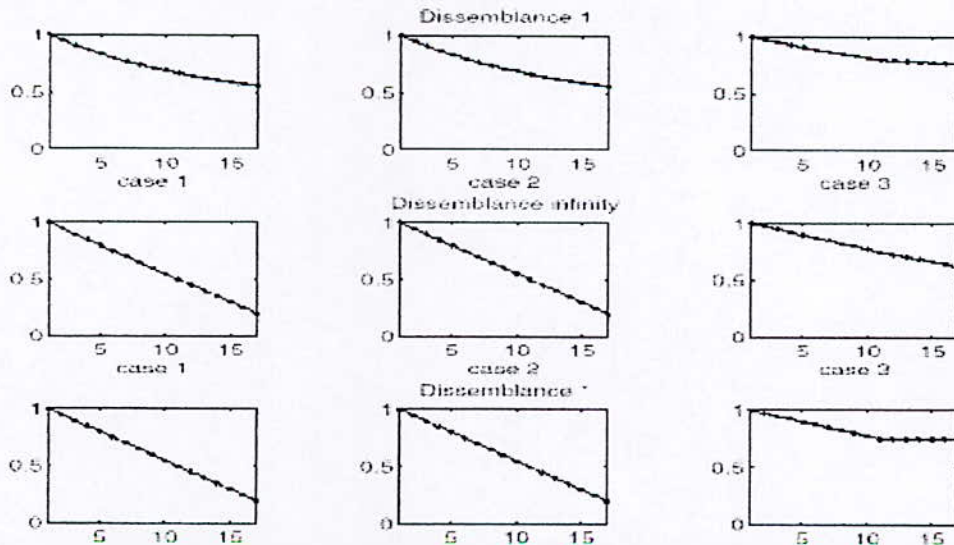


Figure 6.9 : résultats des mesures de similarité par indice de dissemblance

6.3.5 S1 :

La mesure de similarité S1 (5.21) divise l'intersection de deux ensembles par leur union. elle rend la mesure de similarité $s \in [0; 1]$ où $s = 1$ correspond à des ensembles égaux (l'intersection et l'union sont les mêmes) et $s = 0$ donné pour des ensembles qui ne se recouvrent pas (l'intersection est vide). Quand en l'applique aux trois cas de tests, S1 donne les valeurs illustrées dans Fig. 6.10.

6.3.6 S2 :

La mesure de similarité S2 (5.22) est basée sur la différence symétrique (Définition 9). Elle donne une valeur de similarité $s \in [0,1]$; Cependant, elle ne donnera pas la valeur de similarité $s = 1$ pour les ensembles flous égaux. La similarité devient 1 seulement pour des fonctions d'appartenance proprement égales. Pour des fonctions d'appartenance 'Faiblement égales' s devient $\geq 1/2$. Aussi la valeur $s = 0$ n'est pas réservé pour les fonctions d'appartenance qui ne se recouvrent pas comme est le cas avec les autres mesures de similarité basées sur la théorie des ensembles. Les résultats de l'application S2 aux tests peuvent être vus dans la Fig. 6.11.

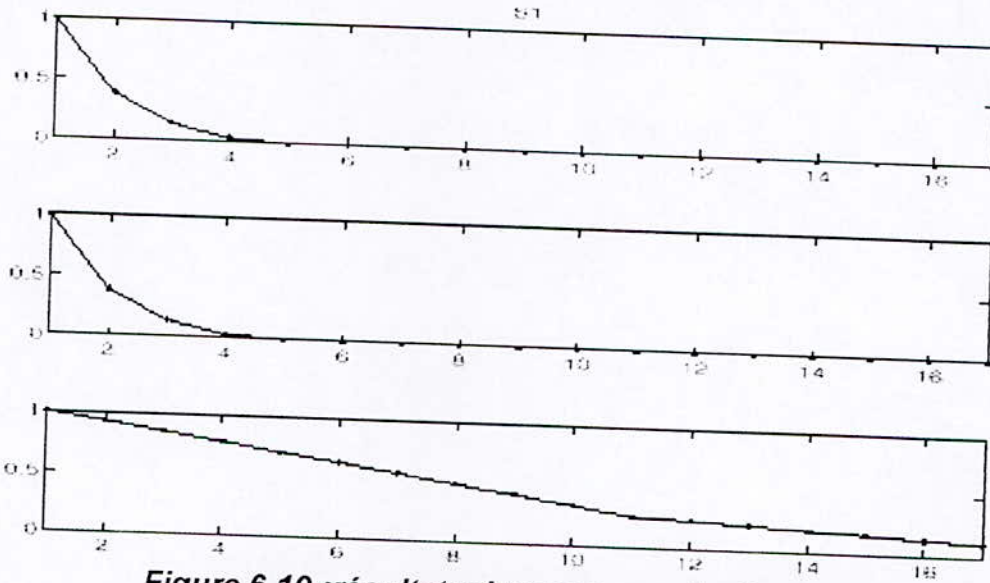


Figure 6.10 : résultats des mesures de similarité S₁

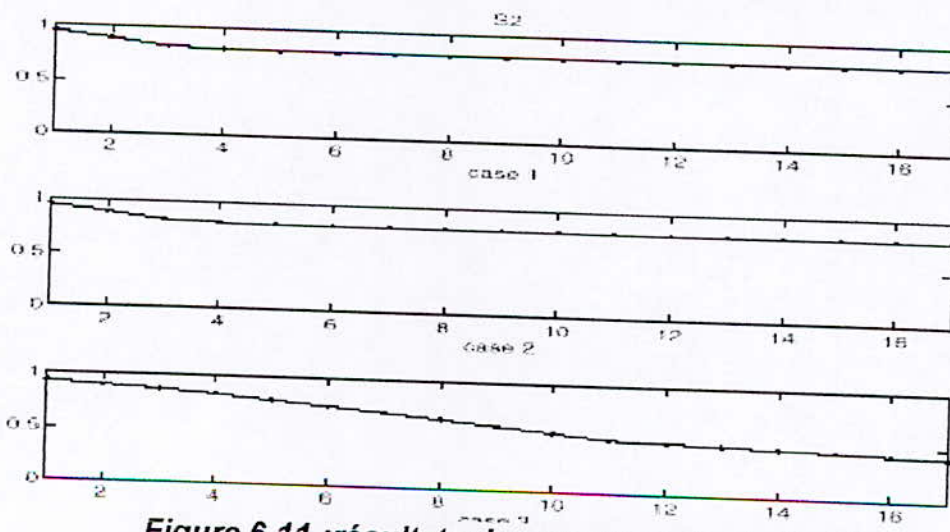


Figure 6.11 : résultats des mesures de similarité S₂

6.3.7 S3 :

S_3 (5.23) est aussi une mesure de similarité basée sur la différence symétrique. Si A et B sont deux ensembles flous égaux, alors $\sup_{x \in X} \mu_{A \ominus B}(x) = 0.5$ et $S_3(A, B)_{\text{égaux}} = 0.5$. Si les deux ensembles flous ne chevauchent pas, alors $S_3(A, B)_{\text{non recouvert}} = 0$. Quand elle est appliquée aux cas de tests, S3 donne les résultats tracés dans la Fig. 6.12 qui montre que les ensembles flous égaux sont assignés par une valeur de similarité $s = \frac{1}{2}$ et les ensembles qui ne se recouvrent pas ont la valeur $s = 0$. Donc, la mesure n'est pas capable de différencier entre ensembles flous dissemblables et est insensible à la valeur de la variable du domaine (résultats cas 1 = résultats cas 2). Légèrement modifiées, d'autres versions des deux S_2 et S_3 sont données dans l'annexe.

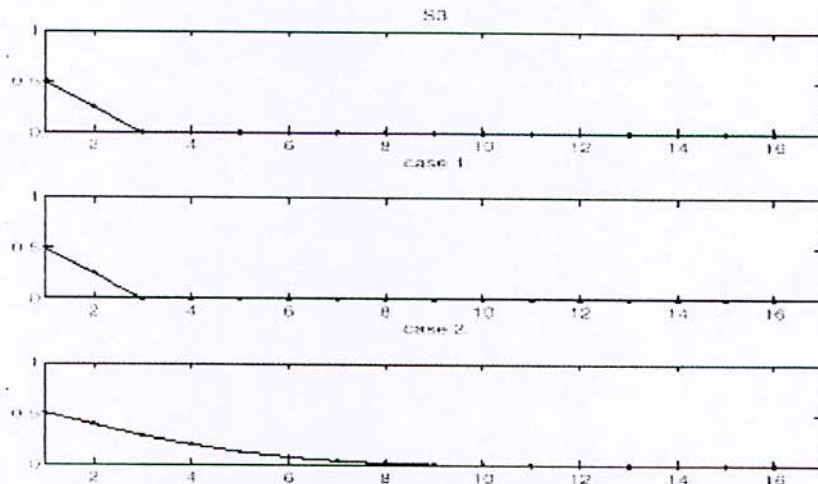


Figure 6.12 : résultats des mesures de similarité S_3

6.3.8 S4 :

La valeur de la similarité donnée par la méthode de mesure S_4 (indice de consistance) est basée sur un point seulement, à savoir le maximum de l'intersection des deux ensembles flous A et B (5.24). Il rend la valeur de la similarité $s \in [0; 1]$ où $s = 1$ est donné quand les deux ensembles flous en question ont un élément commun (point), c.-à-d. $\exists x \in X; \mu_A(x) = \mu_B(x) = 1 - S_4(A; B) = 1$. Pour des ensembles flous qui ne chevauchent pas, l'intersection est vide, et S_4 donne $s = 0$. La fig. 6.13 montre les résultats de l'application de S_4 aux trois cas de tests où il est noté que S_4 est insensible à la valeur numérique des variables du domaine. Cependant, S_4 ne peut pas manier l'inclusions, donnant le résultat $s = 1$ pour tout les ensembles flous dans cas 3.

6.4 Sélection d'une mesure de similarité:

Les quatre critères présentés au début de ce chapitre nous permettent d'en éliminer quelques-unes de ces mesures de similarité. la table 6.3 montre lesquelles de ces mesures sont éliminées. il ne sera pas si difficile d'en sélectionner une qui servira pour notre but. Cependant, il est possible d'en modifier quelques-unes de ces mesures de similarité pour se conformer a ces critères. Des Suggestions à de telles modifications sont présentées dans l'Annexe. L'élimination basée sur les quatre critères nous laisse avec 6 mesures qui peuvent être considérées comme candidates: S_1 , S_2 , la Neighbourhoodness de Chen, la Bhattacharyya, Dissemblance 1, et Dissemblance ∞ .

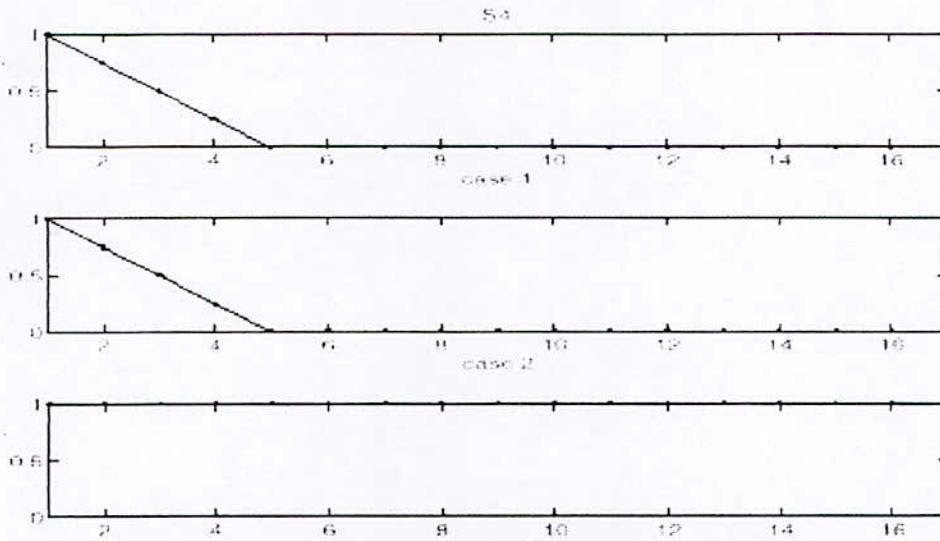


Figure 6.13 :résultats des mesures de similarité S_4

Table 6.3 :élimination de mesures de similarité

Mesures de similarité	Non choisie pour le critère	Equation
Euclidienne	4	6.1
Dominance	2	A.2
City-blok	4	A.1
Hausdorff1	3	6.2
Hausdorff ∞	3 ; 4	6.3
Hausdorff*	3	6.4
Goetschel et Voxman	3	6.5
Dissemblance	3	6.8
Corrélation	4	A.5
Wenstop	3 ; 4	A.3
S_3	2	5.24
S_4	1	5.25
L'indice d'égalité	4	5.26

Si nous examinons de plus près S_2 , il devient clair qu'on peut donner des résultats inattendus, en considérant les différentes paires d'ensembles flous données dans la Figure 6.14. Son utilisation pour mesurer la similarité entre les paires d'ensembles flou, nous permet d'obtenir les résultats suivants:

$A - F$	$C - G$	$D - E$	$D - H$	$J - K$
0.544	0.830	0.840	0.930	0.808

Clairement, la paire (J-K) est candidate évidente pour fusionner, car les ensembles flous sont presque égaux. Cependant, La mesure de similarité S_2 suggère la fusion de (C-G) et (D-E) avant (J-K) et (D-H). La raison est l'usage de la différence symétrique dans l'équation de S_2 (5.22). Deux ensembles flous sont considérés 'faiblement égaux' si leur similarité $S_2 \geq 0.5$, seulement les ensembles proprement égaux sont assignés de la valeur de similarité 1. De la Figure 6.14 il est noté que les fonctions d'appartenance C, G, D et E sont plus propres que les deux

fonctions d'appartenance J, K et H. Donc, la mesure S_2 assigne aux paires d'ensembles flous propres une plus grande similarité qu'à une paire de deux ensembles flous (J-K). Ce n'est pas une propriété positive dans le contexte de simplification des règles de base, donc la mesure de similarité S_2 est abandonnée.

Les deux mesures de similarité basé sur l'indice de Dissemblance $*$ et le Dissemblance ∞ (6.8 et 6.7) peuvent être assez risqué comme mesures de similarité pour indiquer si les deux ensembles flous devrait fusionner ou pas. Ils rendent les valeurs de similarité basées sur les distances mesurées à une α -coupe seulement et font l'usage d'un domaine normalisé ($[\beta_1; \beta_2]$, voir 5.10), lequel n'est pas régulièrement défini. Ces deux mesures de similarité sont écartées.

La mesure de similarité de Bhattacharyya (A.4) semble favoriser les inclusions. Elle donne une très grande valeur de similarité pour les cas où un ensemble flou est complètement inclus dans un autre (c.-à-d. $\mu_A(x) \leq \mu_B(x); \forall x \in X$). De plus, on peut prétendre qu'avoir une mauvaise résolution (par exemple toutes les valeurs données pour le test du cas 3 (Fig. A.2) sont très grandes) cela rend difficile de trouver une valeur du seuil pour décider si oui ou non les deux fonctions d'appartenance sont candidates pour fusionner. Cela nous laisse le choix entre deux mesure de similarité, la mesure S1 et la mesure du Neighbourhoodness (5.21 et 5.17). A partir des cas de tests donnés dans les Figures. 6.10 et A.3, il est noté que leur comportements est assez semblable. Cependant, la mesure de similarité S1 peut être appelée à être utilisée, étant basée sur des opérations d'ensembles théoriques plus intuitifs et bien connus (intersection et union). Dans ce qui suit, la mesure de similarité S1 sera utilisée pour la simplification des règles de base

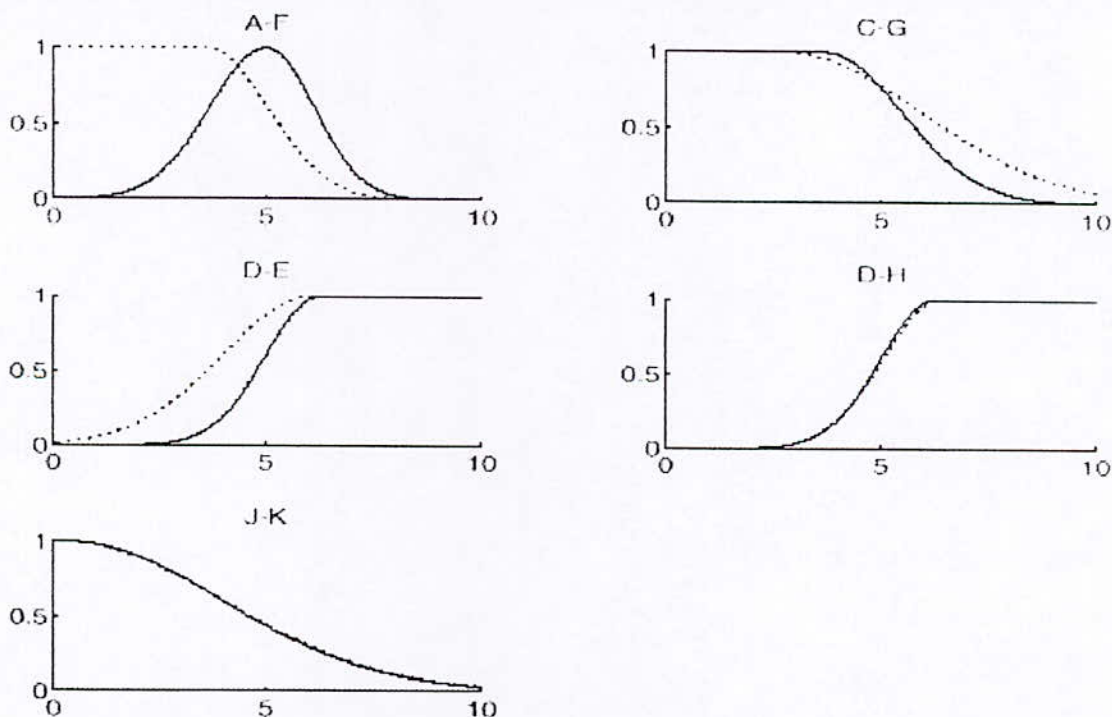


Figure 6.14 :différentes paires d'ensembles flous

6.5 Conclusion

Dans ce chapitre, similarité et mesures de distances rapportées dans le chapitre 5 ont été appliquée pour mesurer la similarité entre ensembles flous, ces mesures peuvent être divisées en deux groupes « les modèles de similarité géométriques » et « les modèles de similarités basées sur les ensembles théoriques ». Pour la sélection d'une mesure de similarité convenable à la simplification des règles de base, ces mesures seront testés et évaluées par des critères qui peuvent délimiter le nombre de méthodes à utiliser

Notons que recouvrement, normalisation de l'univers de discours, inclusion, similarité et dissemblabilité sont les critères imposés pour la sélection. En particulier, la mesure de similarité S_1 a été sélectionnée parmi les différentes méthodes proposées pour la simplification des règles de base.

Comme résumé la table [6.4] donne toutes les études et évaluations exposées dans ce chapitre ainsi que dans l'annexe A.

Similarité	Type	DS	Incl.	S>DS	s(A,A)-1	Echelle	Expression
Minkowsky	GM	Non	Oui	Non	Oui	Oui	$1 - \left(\frac{1}{n} \sum_{i=1}^n \mu_A(x_i) - \mu_B(x_i) ^r \right)^{\frac{1}{r}}$
Hausdorff 1	GM	Oui	Oui	Non	Oui	Non	$\frac{1}{1 + \int_{\alpha=0}^1 q(A_{\alpha}, B_{\alpha}) d\alpha}$
Hausdorff ∞	GM	Oui	Oui	Non	Oui	Non	$1/1 + \sup_{\alpha > 0} q(A_{\alpha}, B_{\alpha})$
Hausdorff *	GM	Oui	Non	Non	Oui	Non	$1/1 + q(A_{1,0}, B_{1,0})$
Goetschel	GM	Oui	Oui	Non	Oui	Non	$1 - q(A^*, B^*)$
Dissemblance 1	GM	Oui	Oui	Oui	Oui	Oui	$1 - \int_{\alpha=0}^1 \Delta(A_{\alpha}, B_{\alpha}) d\alpha$
Dissemblance ∞	GM	Oui	Oui	Non	Oui	Oui	$1 - \sup_{\alpha > 0} \Delta(A_{\alpha}, B_{\alpha})$
Dissemblance *	GM	Oui	Oui	Non	Oui	Non	$1 - \Delta(A_{1,0}, B_{1,0})$
Wenstøp	GM	Oui	Oui	Non	Oui	Non	$1/1 + d_2(V_{f_w}(A) - V_{f_w}(B))$
Bhattacharyya	ST	Non	Oui	Oui	Oui	Oui	$\left(\int_{-\infty}^{+\infty} (\mu_A^*(x)\mu_B^*(x))^{\frac{1}{2}} dx \right)^{\frac{1}{2}}$
Chen	ST	Non	Oui	Oui	Oui	Oui	$\frac{\sum_{i=1}^n \mu_A(x_i)\mu_B(x_i)}{\max \left(\sum_{i=1}^n \mu_A(x_i)^{\alpha}, \sum_{i=1}^n \mu_B(x_i)^{\alpha} \right)}$
S ₁	ST	Non	Oui	Oui	Oui	Oui	$\ A \cap B\ / \ A \cup B\ $
S ₂	ST	Non	Oui	Non	Non	Oui	$1 - \ A \square B\ $
S ₃	ST	Non	Oui	Oui	$s(A,A)=.5$	Oui	$1 - \sup_{\alpha \in x} \mu_A \square_B(x)$
S ₄	ST	Non	Oui	Oui	Non	Oui	$1 - \sup_{\alpha \in x} \mu_{A \cap B}(x)$
Index de qualité	ST	Non	Oui	Non	Oui	Oui	$\frac{1}{2} (\mu_A(x) \mu_B(x)) [\mu_B(x) \mu_A(x)] + [\mu_A(x) \mu_B(x)] [\mu_B(x) \mu_A(x)]$
Correction	GM	Non	Oui	Non	Oui	Oui	$(C_{\mu_A, \mu_B} + 1)/2$

Table 6.4 : Evaluations des Mesures de Similarités

CHAPITRE 7

Simplification des Règles de Base

La Simplification des règles de base

Le clustering flou est une méthode non supervisée utilisée pour obtenir les règles de base d'un modèle flou. Les algorithmes du clustering flou sont appliqués à des mesures de données d'un système et produit des groupes qui se recouvrent avec les limites floues. Avant l'application des techniques de clustering, la structure du problème doit être identifiée et les prémisses et les variables conséquentes (système de variables d'entrée-sortie) doivent être établies. Alors, l'algorithme de clustering est appliqué aux données (mesures des systèmes) dans l'espace produit prémisses/conséquentes. Chaque groupe représente une règle floue qui peut être vue comme une description locale des relations d'entrées-sorties dans leurs parties de l'espace produit prémisses - conséquentes. La projection des groupes identifiés sur les variables prémisses et conséquentes donne les fonctions d'appartenance, et la projection topographique définit les règles floues.

Les modèles présentés dans cette section ont des règles du type Sugeno-Takagi. Les parties conséquentes des règles ne sont pas des entités floues mais des entités propres, se sont des fonctions linéaires des variables prémisses. Les modèles Sugeno-Takagi identifient le système en utilisant des fonctions linéaires pour modéliser le comportement local du système dans un sous espace de l'espace prémisses-conséquentes. La technique de clustering flou qui convient le mieux pour l'identification de tels règles, est l'algorithme de clustering de Gustafson-Kessel qui identifie des hyperplans de dimension $(n-1)$ dans un espace prémisses conséquentes de dimension n , les parties conséquentes des règles de Takagi-Sugeno sont identifiées des groupes eux-mêmes, ou bien par techniques d'estimation des paramètres.

La plupart de ces fonctions d'appartenance ont tendance à être très semblables et ont un grand degré de recouvrement. Cela introduit la redondance dans les règles de base et rend la description linguistique des systèmes complexes difficiles, parce que les fonctions d'appartenance n'ont aucune signification intuitive directe. Donc il est préférable de fusionner les fonctions d'appartenance semblables pour réduire la redondance et améliorer l'interprétation sémantique des règles. La simplification des règles de base est le terme utilisé pour la réduction du nombre de fonctions d'appartenance dans une règle de base.

Dans les chapitres précédents, la similarité et les méthodes de mesure de la similarité entre ensembles flous ont été abordés. La mesure de similarité $S1$ (5.23) a été jugée comme une mesure adéquate pour la simplification des règles de base. Dans ce chapitre, cette mesure sera appliquée aux fonctions d'appartenance dans des modèles flous déjà identifiés et testés. Basée sur les valeurs de similarité obtenues entre paires de fonctions d'appartenance, la simplification des règles de base sera prise en considération. Il sera aussi démontré que la simplification des règles de base peut induire à la réduction du nombre total de règles comme est le cas pour le nombre de fonctions d'appartenance. Elle sera appelée *la réduction des règles de base*.

7.1 La fusion des fonctions d'appartenance :

Les modèles considérés sont des modèles Sugeno-Takagi. Cela veut dire que nous tiendrons compte uniquement de la partie prémisses des modèles, c.-à-d. les antécédents de la règle, en prenant en considération la similarité entre fonctions d'appartenance. Toutes les décisions concernant la simplification des règles de base

(c.-à-d. fusionner ou pas), et leur réduction quand elles sont applicables, seront uniquement basées sur la similarité des ensembles flous.

7.1.1 La simplification des règles de base [27], [28], [37]

La simplification des règles de base signifie la réduction totale du nombre des fonctions d'appartenance inégales. Les fonctions d'appartenance semblables définies pour la même variable prémisse, et Basées sur la mesure de similarité $S1$ sont sélectionnées pour fusionner. Cette Fusion d'une paire de fonctions d'appartenance moyennement semblables sera remplacée par une fonction d'appartenance qui sera représentatif aux deux. De ce fait, le nombre de ces fonctions décrivant des antécédents flous avec une grande similarité dans différentes règles est réduit. Cette réduction de la redondance dans les règles de base est aisément contrôlée en ce qui concerne la description sémantique du modèle comme est réduit le nombre des différents mots et fonctions d'appartenance nécessaire pour décrire le processus. Un modèle sémantique peut être utilisées par exemple pour la validation experte du modèle et pour l'instruction des opérateurs. Si plusieurs ensembles flous sont définis pour une variable, il peut être difficile de coller qualitativement des étiquettes linguistiques significatifs aux ensembles flous, la Figure [7.1] illustre l'idée.

7.1.2 La réduction des règles de base [27], [28], [37] :

Si il arrive que deux règles ont des prémisses égales après avoir fusionner, leurs parties conséquentes peuvent être combinées. Dans le modèles Sugeno-Takagi, où les conséquentes sont des modèles locaux linéaires, cela peut être fait en faisant la moyenne des deux conséquentes pour former une nouvelle conséquente. Cela conduit a la réduction des règles de base, c.-à-d que le nombre de règles dans les règles de base est réduit. Dans la Fig. 7.2 cela est illustré en utilisant deux règles avec un antécédent pour chacun. La similarité entre les deux antécédents est grande, donc ils sont fusionnés. Cela mène a des parties prémisses égales pour les règles 1 et 2, donc leurs conséquentes peuvent être combinées, conduisant à une règle qui remplace les deux règles originales. Pour les modèles étudiés dans ce chapitre et comme il a été décrit précédemment, quand les prémisses égaux se produisent comme un résultat de la fusion des fonctions d'appartenance, les conséquentes de la règle sont combinées. Cependant, la réduction des règles est un sujet qui nécessite des études approfondies. On peut dire que si les prémisses des deux règles sont très semblables, et que leur conséquentes ne le sont pas, il y a une contradiction dans les règles de base. Des prémisses semblables tel que ceux-ci peuvent être une indication d'une identification non optimale de la structure des systèmes.

De ce fait, le modèle est produit par inspection ou heuristiquement, les prémisses semblables dans les règles avec des conséquentes contradictoires peuvent indiquer une erreur dans la phase d'acquisition des connaissances.

On peut imaginer aussi que le seuil de similarité pour fusionner est mis très bas, aboutissant à des règles avec des prémisses égales. Si il existe un modèle de règles de base détaillé du système, alors on pourrait vouloir en produire d'autre, pas aussi détaillé, par exemple former des résultats ou pour usage dans des environnements en temps réels (raisonnement progressif) ou d'autres classifications hiérarchiques. Pour de tels résultats un seuil de similarité bas peut être utilisé.

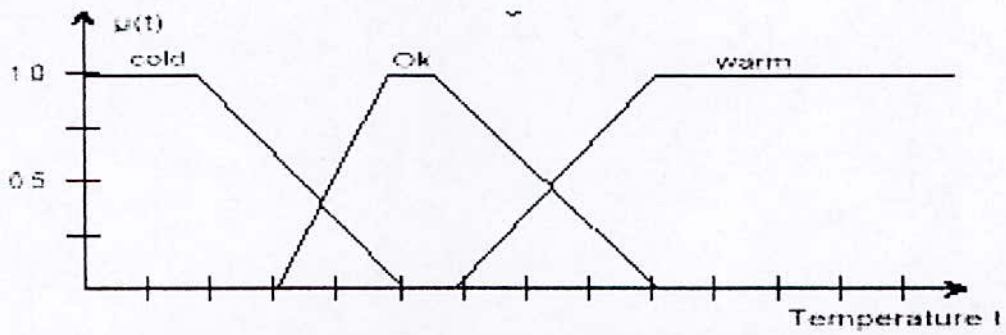
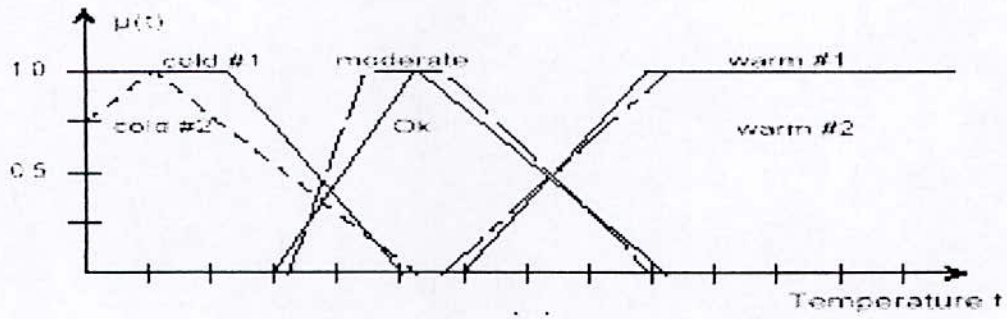


Figure 7.1 : Simplification des règles de base : fusionner les fonctions d'appartenance similaire

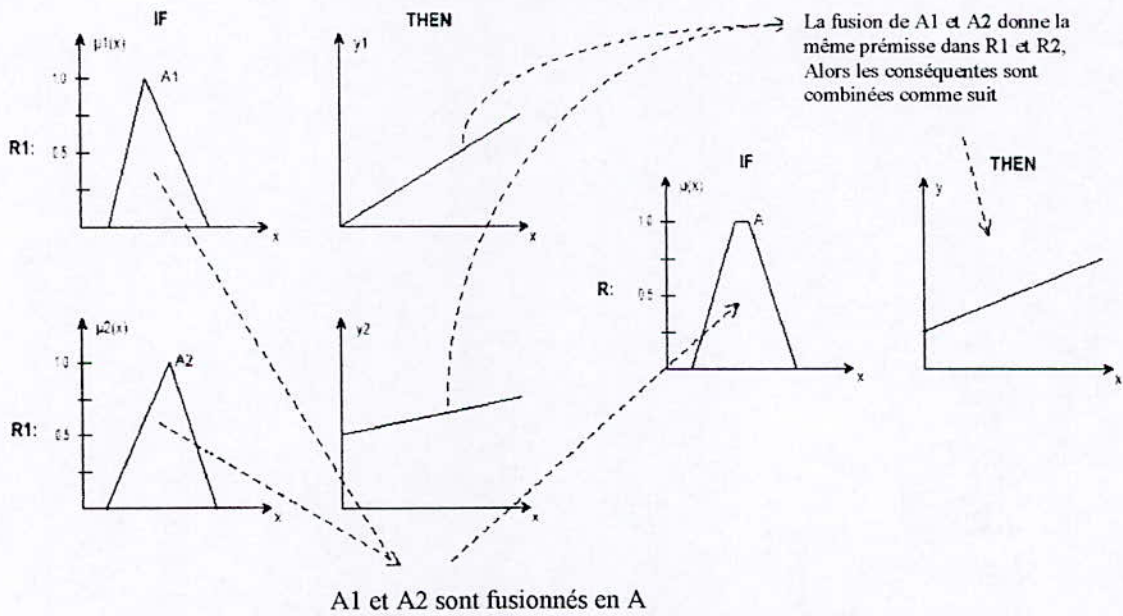


Figure 7.2 : Réduction des règles de base

7.1.3 Comment fusionner [28], [3]:

Quand les deux ensembles flous A et B sont considérés des candidats pour fusionner, la question qui survient est comment les fusionner ?. En général, il y a trois possibilités:

- Remplacer A par B
- Remplacer B par A.
- Remplacer les deux ensembles A et B par un nouvel ensemble flou C.

L'objectif total de la fusion des fonctions d'appartenance est de simplifier les règles dans le modèle identifié. Si nous supposons que le modèle identifié donne l'exactitude optimale, alors l'effet de la fusion sur la sortie devrait être aussi petit que possible. On peut minimiser l'expression du changement dans le modèle comme résultat du changement des fonctions d'appartenance et ceci d'après quelque critère d'erreur pour trouver la meilleur fonction d'appartenance pour remplacer A et B. Lorsque deux fonctions d'appartenance sont fusionnées, ceci affectera le degré de rapprochement des deux règles. L'équation suivante donne le changement dans la sortie du modèle de Sugeno-Takagi si une fonction d'appartenance ou plus dans la prémisse de la règle R m est changé:

$$\partial \hat{y} = \frac{\partial \omega^m}{\partial \omega^m + \sum_{i=1}^N \omega^i} \frac{y^j \sum_{i=1, i \neq m}^N \omega^i - \sum_{i=1, i \neq m}^N \omega^i y^i}{\sum_{i=1}^N \omega^i} \quad (7.1)$$

Pour le calcul de la fonction d'appartenance optimale qui remplace A et B il y a aussi la question d'utiliser le bon critère d'erreur. l'influence de l'erreur de sortie dans quelques cas est acceptable, en général on peut dire que si le modèle est plus sensible aux changements dans A qu'au changements dans B, alors A devrait remplacer B, ou bien le nouveau ensemble flou commun C devrait ressembler à A plus qu'il ne devrait ressembler à B. Cela indique qu'une fonction de sensibilité peut être appliqué pour trouver le meilleur ensemble flou C qui remplace A et B.

Si les ensembles flous A et B sont fusionnés, les conditions suivantes pour la fonction d'appartenance de l'ensemble flou résultante C doivent être vérifiées :

$$\begin{aligned} \min(a_1, b_1) &\leq c_1 \leq \max(a_1, b_1) \\ \min(a_2, b_2) &\leq c_2 \leq \max(a_2, b_2) \\ \min(a_3, b_3) &\leq c_3 \leq \max(a_3, b_3) \\ \min(a_4, b_4) &\leq c_4 \leq \max(a_4, b_4) \\ c_1 &\leq c_2 \leq c_3 \leq c_4 \end{aligned}$$

Quelques manières pour créer C a partir de A et B sont montrées dans la figure 7.5 (Ici aucune considération n'est donnée pour respecter la minimisation de quelques critères d'erreur.)

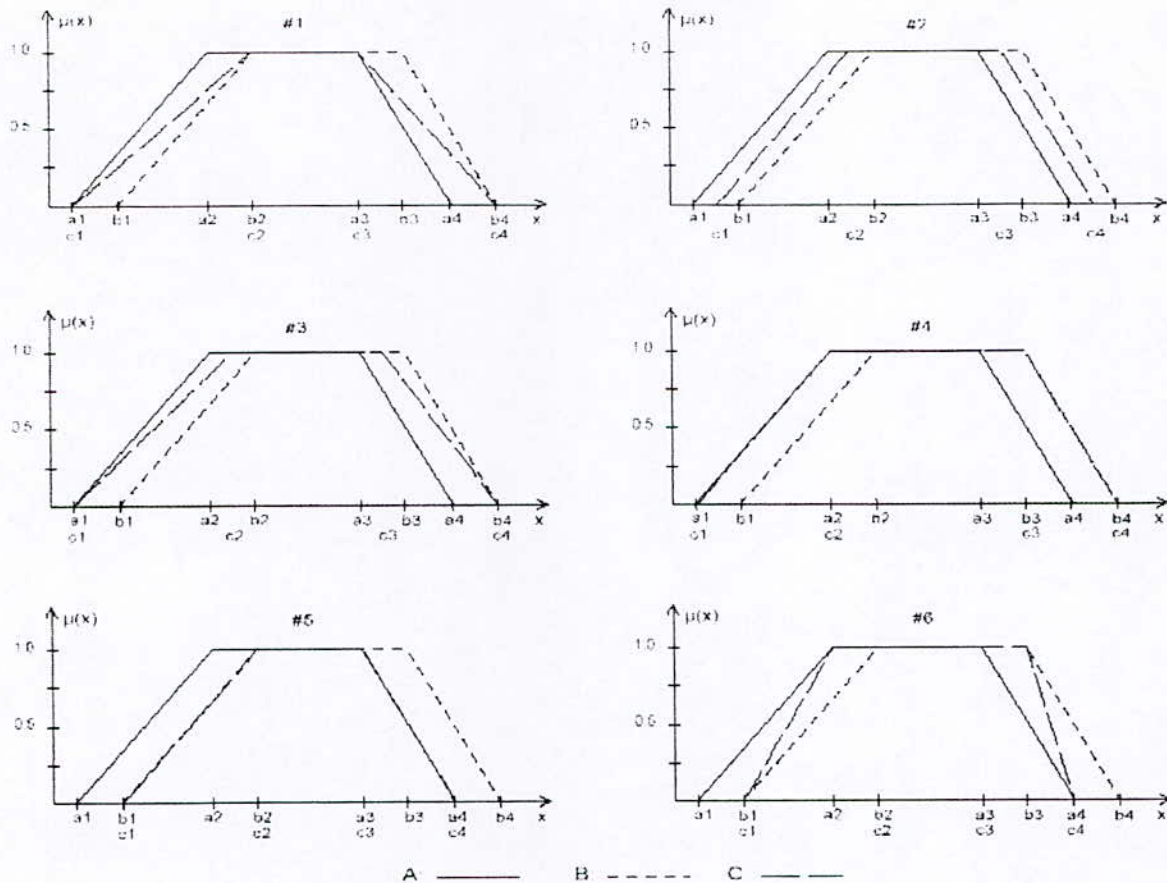


Figure 7.3 : Différentes façons de fusionner A et B pour avoir c

Le choix de la nouvelle fonction d'appartenance qui remplace A et B peut être assez important. Ce choix n'influencera pas seulement l'erreur dans la sortie, mais aussi le nombre total des fonctions d'appartenance fusionnées pour un seuil de similarité donné. Si, en se basant sur une mesure de similarité, on décide de fusionner A et B, alors le choix du nouveau ensemble flou C influencera une éventuelle fusion de C avec un autre ensemble flou dans une étape ultérieure. Si nous regardons les différentes solutions données dans la Figure 7.3, nous remarquerons que le risque d'avoir des vides dans les règles est apparent pour les méthodes 2, 5, et 6. Les Méthodes 5 et 6 présentent le plus grand risque de vides et donnent également le résultat le plus évident qui mène à la prédiction du plus petit nombre de fusion pour un modèle et un seuil donné. La méthode 2 résulte d'une fonction d'appartenance définie par la moyenne des paramètres qui sont fusionnés. Bien sûr, plus le seuil de similarité pour fusionner est inférieur, plus important est le choix de l'ensemble flou C qui remplace A et B. Quand le seuil est bas, le risque des vides dans les règles est plus grand. La méthode 1 n'assurera pas les vides dans ces règles. C'est le même cas pour les méthodes 3 et 4. Cependant, ces méthodes peuvent mener aux grands changements relatifs dans les valeurs d'appartenance. Pour ce qui suit, la méthode 2 est utilisée.

7.1.4 L'enlèvement des fonctions d'appartenance.

Quand le clustering est appliqué, les projections des groupes peuvent mener à des fonctions d'appartenance prenant une valeur $\mu(x) \approx 1; \forall x \in X$. Puisque cette

fonction d'appartenance donnera ≈ 1 pour chaque entrée, cela ne contribue pas à la sortie et n'a aucune signification. De telles fonctions d'appartenance, sont semblables à l'ensemble universel pour toutes les valeurs du domaine, elles peuvent être enlevées des règles de base sans pour autant changer le modèle. Avec l'utilisation d'une mesure de similarité, la similarité entre fonctions d'appartenance aux ensembles universel (c.-à-d. $\mu_U(x) = 1; \forall x \in X$) peut être mesuré. Ces fonctions d'appartenance qui sont semblables à l'ensemble universel sont enlevé des règles, ceci est illustré dans la Figure 7.6.

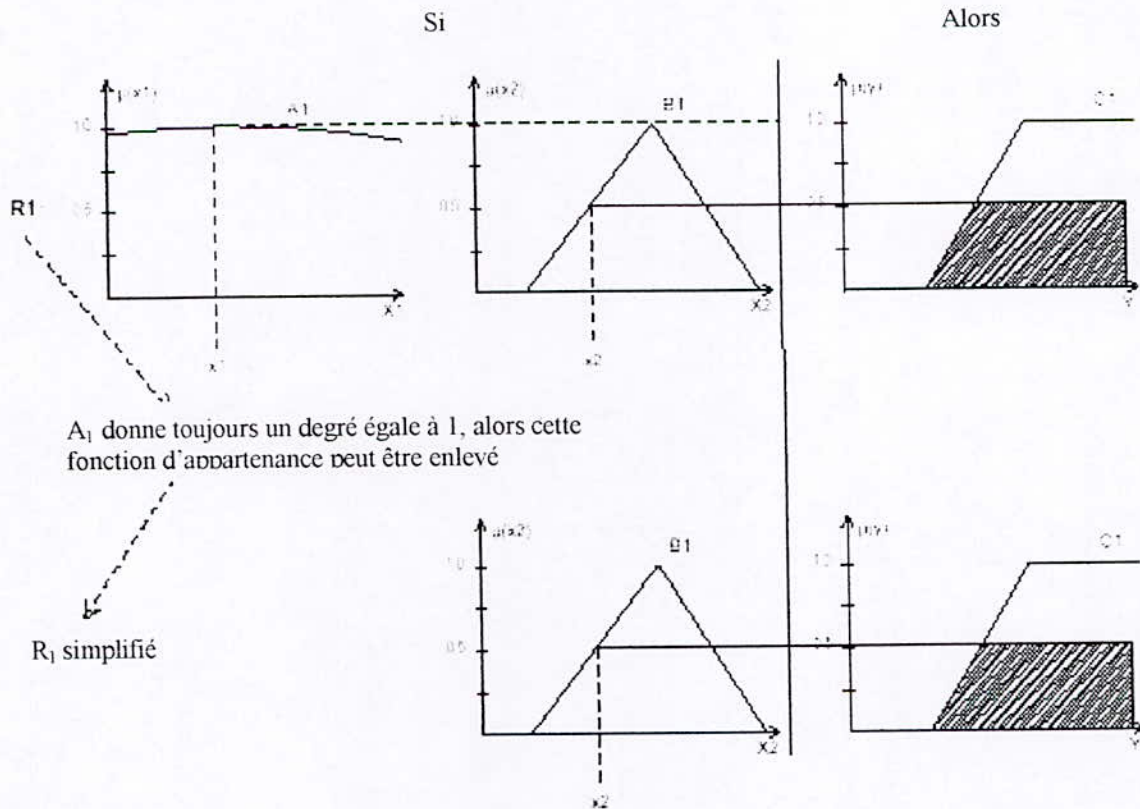


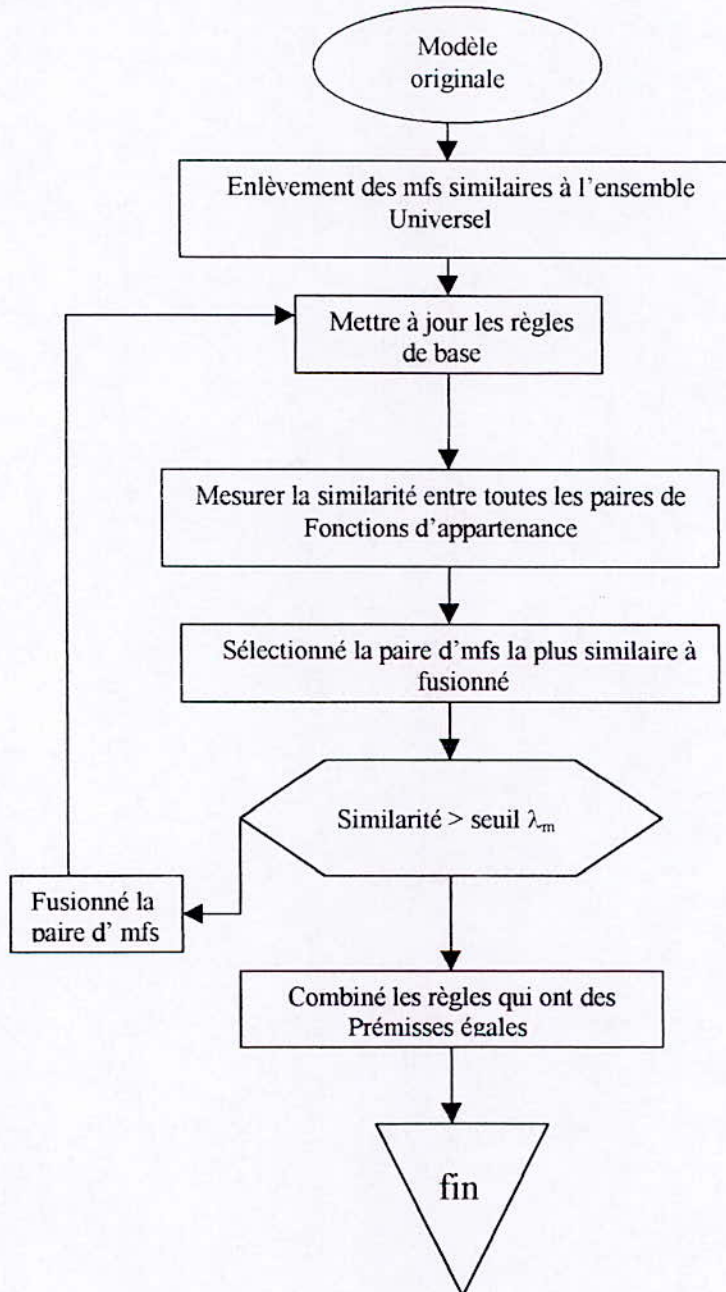
Figure 7.4 : Enlèvement de fonction d'appartenance.

7.2 La méthode

Dans ce qui suit, la mesure de similarité S1 sera appliquée aux fonctions d'appartenance des modèles identifiés afin de trouver des fonctions semblables qui peuvent être fusionner. L'algorithme peut être résumé comme suit:

1. Prendre un seuil de similarité λ_m pour fusionner les fonctions d'appartenance. (Si $S1(A; B) \geq \lambda_m$ alors A et B sont fusionnés). Par défaut λ_m est = 0:8.
2. Prendre un seuil de similarité λ_r pour enlever les fonctions d'appartenance semblable à l'ensemble universel U (Si $S2(A; U) \geq \lambda_r$ alors A est enlevé). Par défaut on a λ_r est = 0:9.

4. Mettre à jour les règles de base.
5. Mesurer la similarité entre toutes les fonctions d'appartenance pour chaque variable prémisse. Sélectionner la paire de fonctions d'appartenance qui a la plus grande similarité.
6. Si $S1(A; B) \geq \lambda_m$ alors A et B sont fusionnés. Si $S1(A; B) < \lambda_m$ alors FIN.
7. Mettre à jour les règles de base.
8. Aller à 5.



7.3 Application pour un système à une entrée et une sortie :

Comme simple application de l'algorithme de simplification et réduction des règles de base, considérons l'approximation d'une fonction sinusoïdale non linéaire simple, de façon à avoir un système de Takagi-Sugeno à une entrée x et une sortie y . Cette fonction est générée par la formule mathématique suivante:

$$y = \sin(0.0015x^2) \frac{x^{2.9}}{1000}, x \in [0,100]$$

Le résultat est un modèle flou de règles de base avec 11 règles, chaque règle a une fonction d'appartenance dans la partie prémisse et la conséquente est une fonction linéaire locale qui se rapproche du comportement de la fonction dans chaque région locale.

Les fonctions d'appartenance du modèle identifié sont tracées dans la Figure 7.1 et les matrices RLS et MFS qui nous donnent le nombre de règles et les fonctions d'appartenance sont données ci-dessous:

rls =

```

1 1
1 2
1 3
1 4
1 5
1 6
1 7
1 8
1 9
1 10
1 11

```

mfs =

```

[2.0000 -9.0830 2.9004 2.9004 48.4205]
[2.0000 0 0 1.0000 56.5706]
[2.0000 -4.8367 23.7613 23.7613 55.2600]
[2.0000 -1.0553 39.7559 39.7559 51.2447]
[2.0000 30.6846 46.4032 46.4032 91.4380]
[2.0000 32.5437 53.0898 53.0898 74.9714]
[2.0000 46.9023 63.6498 63.6498 83.6016]
[2.0000 55.7043 72.2026 72.2026 86.9231]
[2.0000 62.6048 80.1912 80.5537 98.9374]
[2.0000 74.5800 91.6989 91.8787 108.0956]
[2.0000 86.4368 99.6180 100.0000 100.0000]

```

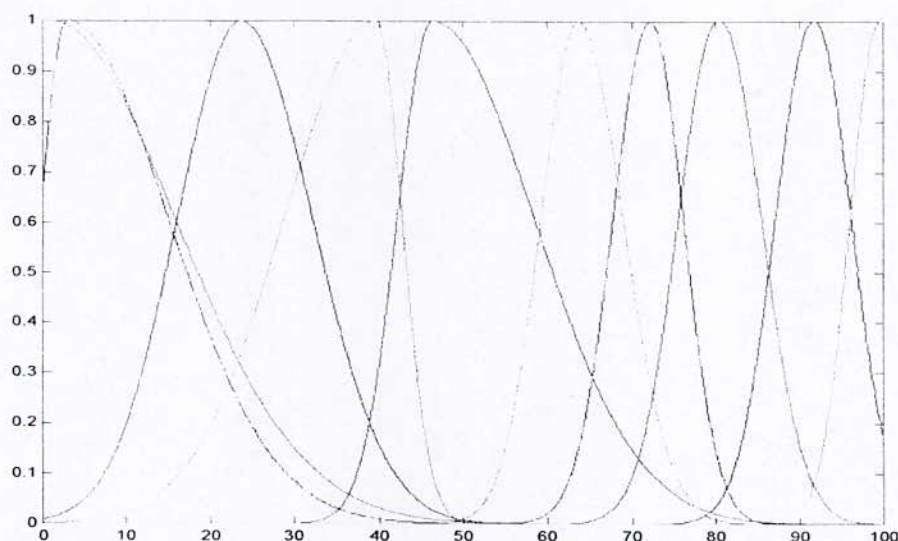



Figure 7.5 : Fonctions d'appartenance des règles de base de la fonction y

Après avoir appliqué l'algorithme de simplification et réduction des règles avec un seuil S de 0.2, nous obtenons un modèle simplifié avec des fonctions d'appartenance tracés en Figure. 7.2. Le nouveau modèle est constitué par deux matrices RLS NEW2, MFS NEW2 données par:

rls_new2 =

1 1
1 2
1 3
1 4
1 5
1 6

mfs_new2 =

[2.0000 -9.0830 0 2.9004 56.5706]
[2.0000 -2.9460 31.7586 31.7586 53.2524]
[2.0000 31.6142 49.7465 49.7465 83.2047]
[2.0000 51.3033 67.9262 67.9262 85.2624]
[2.0000 62.6048 80.1912 80.5537 98.9374]
[2.0000 80.5084 95.6584 95.9393 104.0478]

Comme prévu, le nouveau modèle est beaucoup plus simple que le modèle original. Cependant, utiliser une telle valeur basse du seuil S peut introduire quelques erreurs de plus, comme c'est le cas pour les grandes valeurs de la variable x . Le modèle de sinusoïde originellement identifié (avec 11 fonctions d'appartenance) introduit moins d'erreurs que la version réduite (avec 6 fonctions d'appartenance).

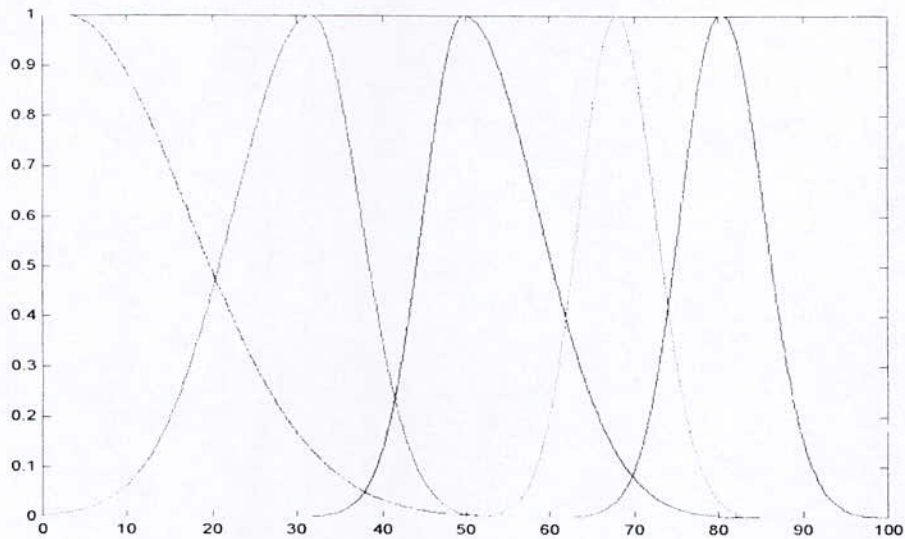


Figure 7.6 : fonctions d'appartenance simplifi  de la fonction y pour $\lambda_m=0.2$

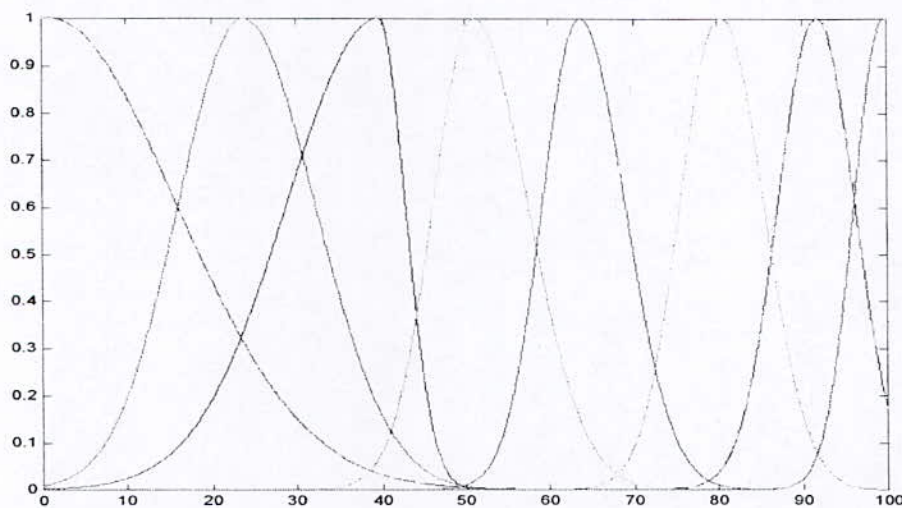


Figure 7.7 : fonctions d'appartenance simplifi es de la fonction y pour $\lambda_m=0.6$

7.4 Application   un moteur Asynchrone :

Comme application de l'algorithme de simplification et r duction des r gles de base   un syst me flou   deux variables pr misses et une variable cons quente (mod le Takagi-Sugeno), consid rons le mod le de la machine asynchrone aliment  en tension[12] donn  dans l'annexe D, o  les deux entr es sont les tensions normalis es V_q et V_d , la sortie est la vitesse.

Le r sultat de l'application du clustering est un mod le flou de base de r gles avec 14 r gles et 7 fonctions d'appartenance ; chaque r gle a deux fonction d'appartenance dont la partie pr misse et cons quente est une fonction lin aire locale qui se rapproche du comportement de la fonction vitesse dans chaque r gion locale.

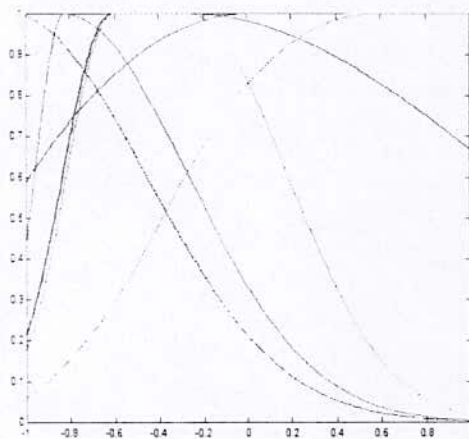
Les fonctions d'appartenance du modèle identifié sont tracées dans la Figure 7.7 et les matrices MFS qui nous donnent le nombre de fonctions d'appartenance sont données ci-dessous:

Pour V_q :

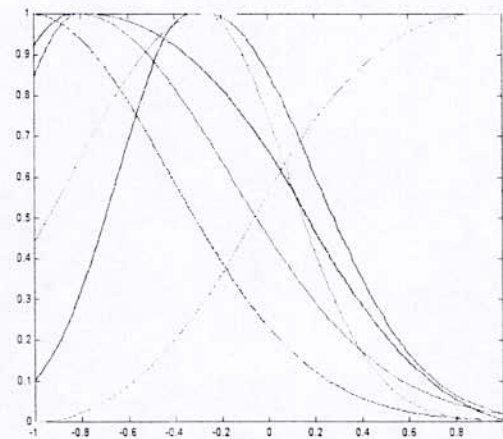
```
mfs=[2 -3.201 -1.064 -1.064 1.201;
      2 -1.337 -0.824 -0.824 1.254;
      2 -1.44 -0.611 -0.26 4.99;
      2 -1.40 -0.601 -0.262 1.4;
      3 -2.21 0.008 1.37 9.150;
      3 -3.03 0.54 2.07 9.83;
      3 -1.33 0.56 2.45 6.09];
```

Pour V_d :

```
Mfs =[2 -3.66 -1.01 -1.01 1.2;
      2 -2.03 -0.808 -0.808 1.6;
      3 -1.67 -0.83 -0.77 1.11;
      2 -2.4 -0.27 -0.27 1.06;
      2 -1.5 -0.33 -0.25 1.43;
      2 -1.2 -0.27 -0.15 1.66;
      3 -1 0.88 3.45 6.60];
```



tension V_q



tension V_d

Figure 7.7 : fonctions d'appartenance des prémisses V_q et V_d

Après avoir appliqué l'algorithme de simplification et réduction des règles avec un seuil λ_m de 0.2, nous obtenons un modèle simplifié avec des fonctions d'appartenance tracés en Figure 7.8. Le nouveau modèle est constitué par deux matrices MASNEWV_q, MASNEWV_d données par:

Pour MASNEWV_q :

```
mfs=[2 -3.00 -1.00 -0.94 1.12;
      2 -1.30 -0.30 -0.270 1.40;
      2 -1.96 0.11 2.66 4.08];
```

Pour MASNEWV_d :

```
mfs=[2 -1.05 -1.009 -0.88 1.21;
      2 -1.50 -0.33 -0.33 1.40;
      2 -1.009 0.84 1.15 2.333];
```

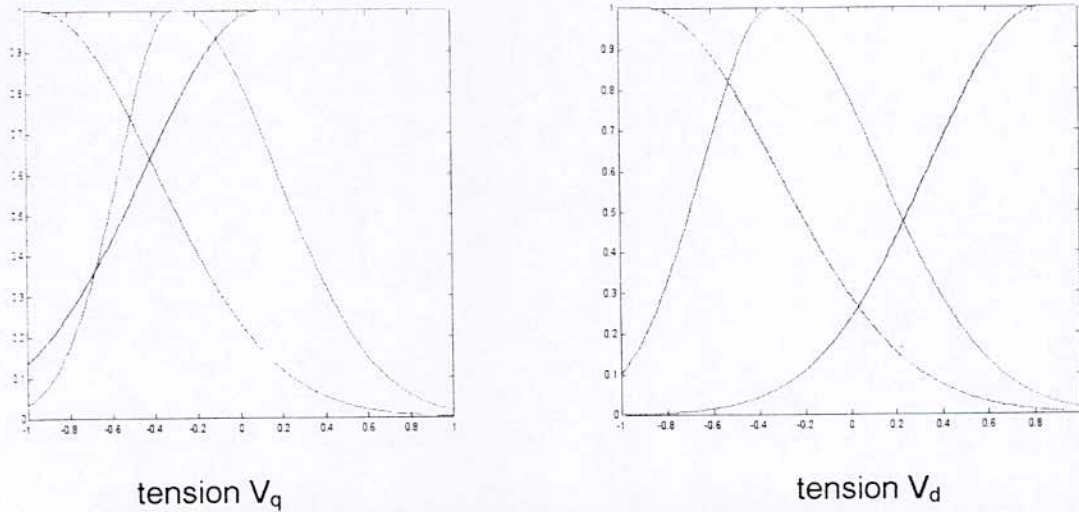


Figure 7.8 : fonctions d'appartenance de V_q et V_d pour $\lambda_m=0.3$

Comme prévu, le nouveau modèle est beaucoup plus simple que le modèle original. Cependant, utiliser une telle valeur basse du seuil S peut introduire quelques erreurs de plus. Le modèle de la machine originairement identifié (avec 7 fonctions d'appartenance, 14 règles) introduit moins d'erreurs que la version réduite (avec 3 fonctions d'appartenance et 9 règles).

7.5 L'application à une enzyme de terre [28] :

Dans le but de montrer le cas où on a une suppression d'une variable prémisses, nous tenons à exposer et à vérifier le modèle non linéaire de Takagi-Sugeno de la dégénérescence d'une enzyme de terre, ceci afin de montrer comment le passage d'un modèle de trois variables prémisses à un modèle de deux variables prémisses, en jouant sur les valeurs des seuils λ_m et λ_r .

La combinaison de la sortie en fonction des entrées prémisses dans le modèle de Takagi-Sugeno est donné par :

$$y = a_0 + a_1 \times \text{time} + a_2 \times \text{pH} + a_3 \times \text{température.}$$

Les fonctions d'appartenance de ces entrées dans le modèle original sont données dans la figure 7.9

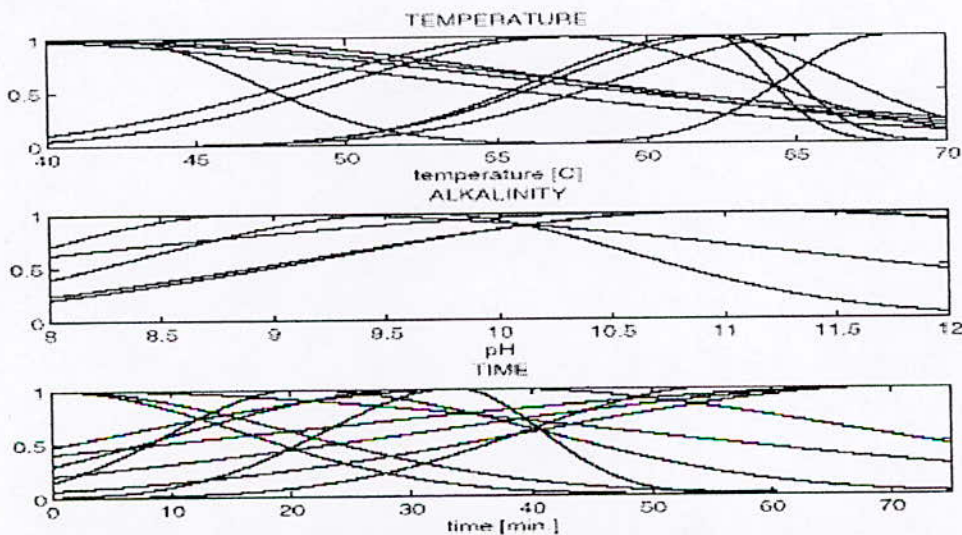


Figure 7.9 : fonctions d'appartenance du modèle originale

Notons que le nombre total de fonctions d'appartenance du modèle original est égale a 33, avec 11 règles de base.

Après l'application de l'algorithme de simplification et réduction, on notera qu'a différente valeur de λ_m on a un nombre de fonctions qui se réduit avec la diminution du seuil, ainsi pour les valeurs 0.8, 0.5 et 0.3 les fonctions d'appartenance sont données dans les figure ci-dessous.

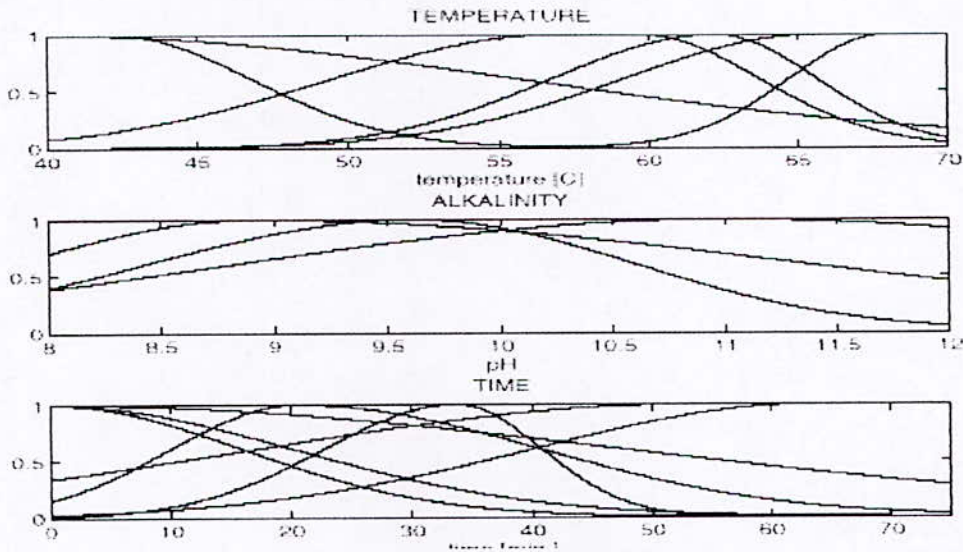


Figure 7.10 : fonctions d'appartenance à $\lambda_m=0.8$.

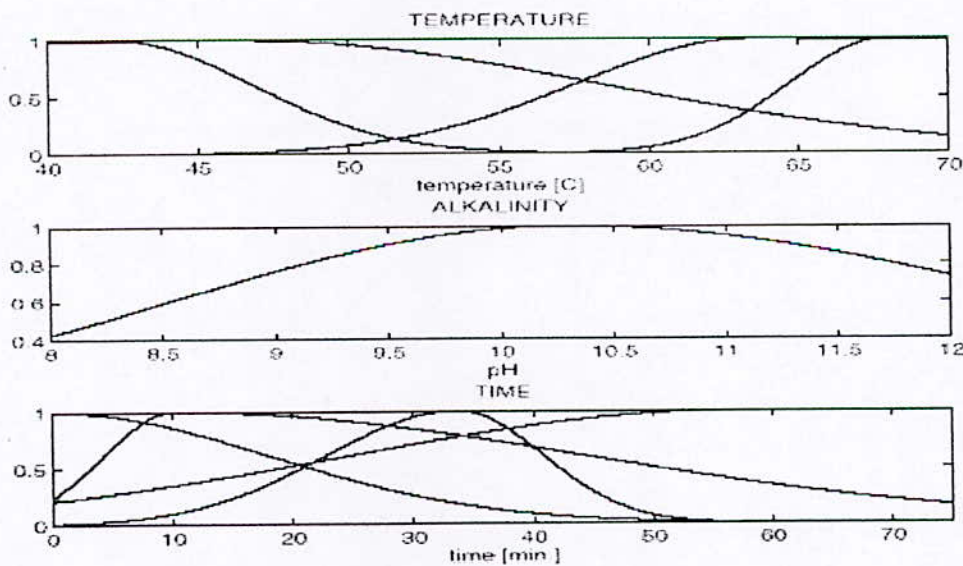


Figure 7.11 : fonctions d'appartenance à $\lambda_m=0.5$.

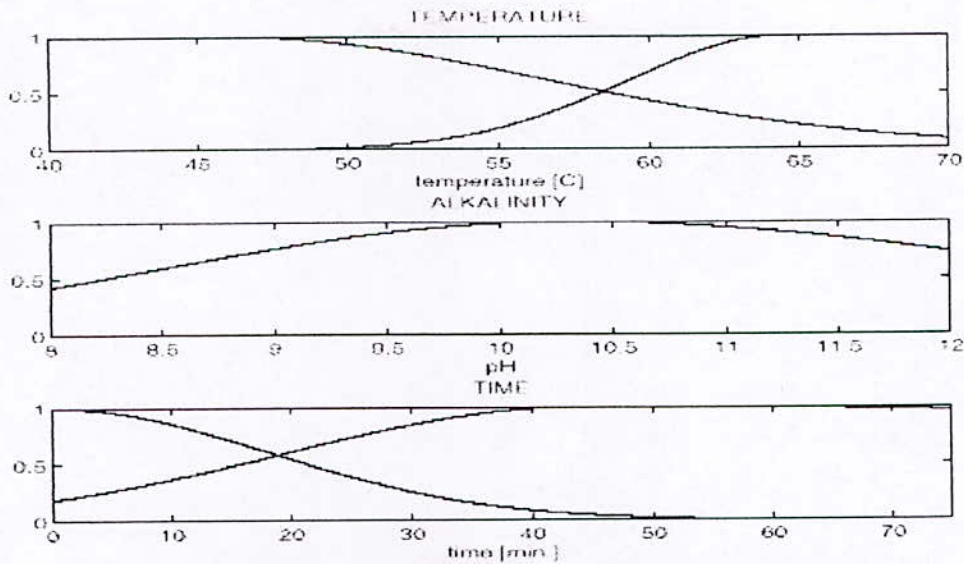


Figure 7.12 : fonctions d'appartenance à $\lambda_m=0.3$.

Dans le cas où le seuil est égale à 0.3, l'algorithme de simplification et de réduction va comparer la fonction Alcalinité à une fonction universel. En comparant la similarité entre eux, on en déduit que le modèle est indépendant de la fonction alcalinité donc, cette dernière n'est pas prise en compte par le système est le modèle devient a deux variables prémisses.

$$y^j = a_0 + a_1 \times \text{time} + a_2 \times \text{température.}$$

7.6 Conclusion :

Dans ce chapitre, les modèles flous existants ont été simplifiés et réduits, en se basent sur la similarité entre les fonctions d'appartenance dans la partie prémisses des règles. La mesure de similarité S_1 a été utilisée pour ressortir la similarité. En introduisant le seuil de fusion λ_m dans l'algorithme de simplification, on peut aisément noter que le degré de simplification peut être ajuster. Nous concluons que, si le seuil λ_m est réduit, le degré de simplification augmente, c'est à dire que le nombre de règles et de fonctions d'appartenance sera réduit.

Conclusion générale

Dans ce travail, nous nous sommes intéressés à la simplification et à la réduction des règles flous en utilisant les mesures de similarité pour l'obtention d'un modèle optimale de Takagi-Sugeno.

Nous avons présenté tout d'abord une définition détaillée des ensembles flous en donnant leurs différentes caractéristiques, et en introduisant dans la théorie de la logique floue le concept de variables et de raisonnement linguistiques nous avons essayé de donner une interprétation plus proche du raisonnement humain.

Nous avons présenté en outre, une méthodologie de base du contrôle flou ainsi que les différentes stratégies utilisées en raisonnement par les systèmes flous.

Les modèles flous exigent la disponibilité d'une expertise. Par conséquent, les performances de ces derniers sont étroitement liées aux techniques d'acquisition de connaissances et à la justesse des informations acquises.

La similarité est utilisée dans différents domaines, et son interprétation diffère selon le contexte dans lequel elle est utilisée. Dans le cas de la simplification des règles floues, plusieurs méthodes de mesures pour ressortir la distance ou la similarité entre ensembles flous sont présentées et étudiées du point de vue mathématique, mais du point de vue application (simplification et réduction des règles) on a été contraint d'imposer quelques critères liés aux ensembles flous pour pouvoir utiliser une mesure de similarité (S_1) qui nous convienne.

A l'issue de l'étude effectuée et des résultats obtenus présentés dans ce mémoire, nous envisageons les perspectives de continuité suivantes :

- L'extension de l'application des mesures de similarités aux autres disciplines de l'électronique (traitement du signal, de l'image, commande des systèmes, traitement de l'information, etc....), mais aussi dans d'autres domaines de l'Engineering (biologie, économie, agriculture, etc....).

Annexe A :

Nous développerons dans cette annexe les études et les évaluations des mesures de similarités qui n'ont pas été représentées dans le chapitre 6 . De ce fait, nous aurons a donner leurs équations ainsi que les réponses aux trois cas de tests donnés dans ce même chapitre.

Mesures de similarités :

1. La mesure de similarité City-block :

$$S_{d_1}(A, B) = 1 - \left(\frac{1}{n} \sum_{i=1}^n |\mu_A(x_i) - \mu_B(x_i)| \right) \quad (A.1)$$

2. La mesure de similarité Dominance :

$$S_{d_x}(A, B) = 1 - \sup_{x \in X} |\mu_A(x) - \mu_B(x)| \quad (A.2)$$

3. la mesure de similarité de Wenstop :

$$S_{F_{W'}}(A, B) = \frac{1}{1 + d_2(V_{fw}(A), V_{fw}(B))} \quad (A.3)$$

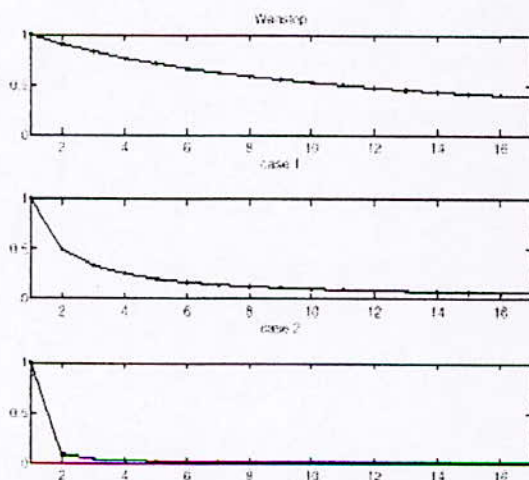
4. la mesure de similarité de Bhattacharyya :

$$S_B(A, B) = \left(\int_{-\infty}^{+\infty} (\mu_A^*(x) \mu_B^*(x))^{\frac{1}{2}} dx \right)^{\frac{1}{2}} \quad (A.4)$$

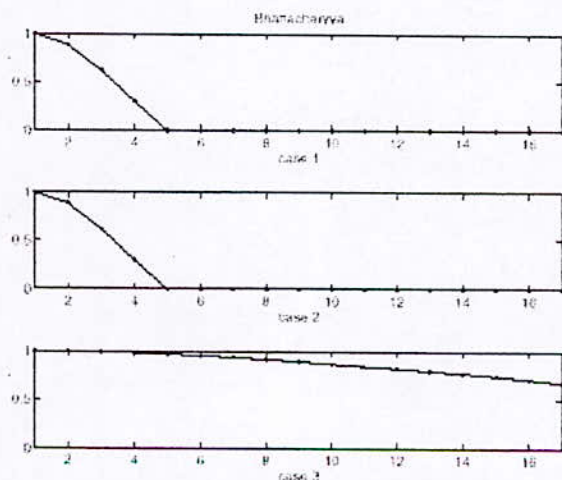
5. L'indice de corrélation :

$$S_c(A, B) = \frac{C_{\mu_A, \mu_B} + 1}{2} \quad (A.5)$$

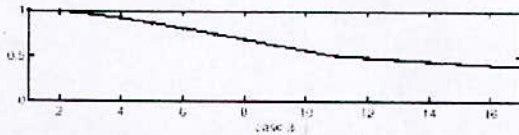
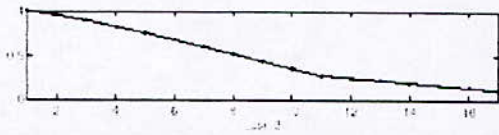
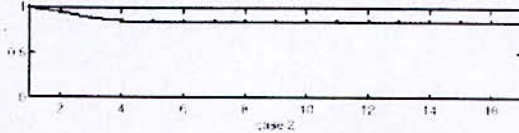
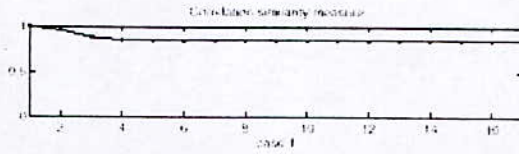
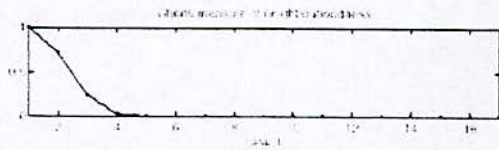
Les résultats des tests appliqués a ces méthodes ainsi que celles citées dans le chapitres 5(Indice d'égalité, Neighbourhoodness,...) sont donnés dans les figures qui suivent.



[A.1] : Résultats de la mesure de Wenstop

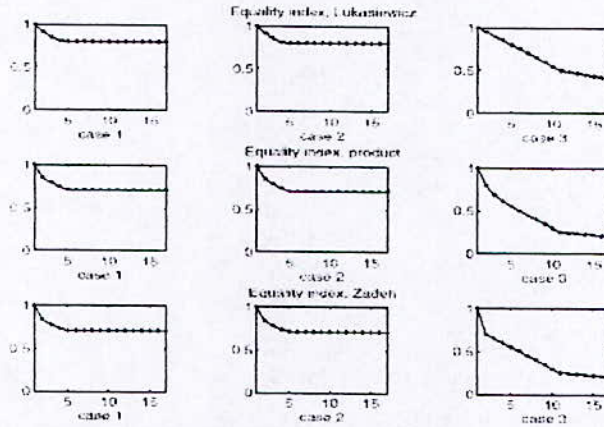


[A.2] : Résultats de la mesure de Bhattacharyya



[A.3] : Résultats de la mesure du Neighbourhoodness

[A.4] : Résultats de la mesure de corrélation



[A.5] Résultats de la mesure de l'indice d'égalité

Annexe B : L'algorithme de clustering de Gustafson-Kessel.

Le GK algorithme recherche des hyperplans de dimension (n-1) par le biais de donnés dans un espace de dimension (n). Posons $X = \{x_1, x_2, \dots, x_n\}$.

L'algorithme de clustering minimise la fonction objective J_m :

$$J_m(U, V, A) = \sum_{k=1}^n \sum_{i=1}^c \mu_{ik}^m \|x_k - v_i\|_{A_i}^2, m \geq 1. \quad (B.1)$$

Où m est le cluster fuzziness, n le nombre de points donnés, et c est le nombre de groupes. U est la matrice de partition floue.

L'algorithme peut être spécifié comme suit :

1. définir la matrice de partition U, sélectionné le critère d'arrêt ϵ .
2. Calcul des nouveaux groupes avec :

$$v_i = \frac{\sum_{k=1}^n \mu_{ik}^m x_k}{\sum_{k=1}^n \mu_{ik}^m}, i = 1, 2, \dots, c. \quad (B.2)$$

3. mise a jour de U en U^* avec :

$$\mu_{ik}^* = \left(\frac{c \|x_k - v_i\|_{A_i}^{2/m-1}}{\sum_{j=1}^c \|x_k - v_j\|_{A_j}^{2/m-1}} \right)^{-1} \quad (B.3)$$

$$A_i = (\rho_i \det(S_i))^{1/2} S_i^{-1}, i=1, 2, \dots, c, \quad (B.4)$$

où $S_i = \sum_{k=1}^n \mu_{ik}^m (x_k - v_i)(x_k - v_i)^T$.

4. SI $\max_{i,k} |\mu_{ik}^* - \mu_{ik}| \leq \epsilon$, Stop, sinon $U = U^*$, Aller à 2.

Le plus important paramètre dans cet Algorithme est c, le nombre de groupes. Ce paramètre détermine le nombre de règles dans un système de règles de base quand le clustering flou est utilisé pour la modélisation.

Annexe C : quelques méthodes de mesures de similarité modifiées.

Pour satisfaire les critères énoncés dans le chapitre 6, quelques méthodes de mesures de la similarité sont modifiées, créés des méthodes S_2 et S_3 , les mesures M_{S_2} , M_{S_3} sont écrites comme suit :

$$M_{S_2}(A, B) = 2 \frac{S_2(A, B)}{S_2(A, A) + S_2(B, B)} \quad (C.1)$$

$$M_{S_3}(A, B) = 2 \frac{S_3(A, B)}{S_3(A, A) + S_3(B, B)} \quad (C.2)$$

une autre méthode dite S_r qui est inspiré de la mesure métrique de Minkowski qui prend ces valeurs dans $[0, 1]$

$$S_r(A, B) = \frac{\|A + B\|_r - \|A - B\|_r}{\|A + B\|_r}, r \geq 1 \quad (C.3)$$

où r indique le type de norme à utiliser :

$$\|A\|_r = \left(\sum_{i=1}^n |\mu_A(x_i)|^r \right)^{\frac{1}{r}} \quad (C.4)$$

pour $r=2$, en utilise la norme Euclidienne.

ANNEXE D**Modélisation de la machine asynchrone**

Le modèle de la machine asynchrone est un système de six équations non linéaires et à coefficients périodiques du temps. Par conséquent la résolution d'un tel système s'avère très complexe. On introduira la transformation de Park pour réduire cette complexité pour trouver un système d'équation à cinq équations, dont les coefficients constants.

A-1- Equation électrique de la machine asynchrone :

la machine est alors représentée par les équations des phases statorique et rotoriques suivantes :

$$[V_{Sabc}] = [R_S][i_{Sabc}] + \left(\frac{d}{dt}\right)[\Phi_{Sabc}] \quad (D-2-1)$$

$$[V_{Rabc}] = [R_R][i_{Rabc}] + \left(\frac{d}{dt}\right)[\Phi_{Rabc}] \quad (D-2-2)$$

Les expressions des flux en fonction des courants statorique et rotoriques sont :

$$[\Phi_{Sabc}] = [I_S][i_{Sabc}] + [M_{SR}][i_{Rabc}] \quad (D-2-3)$$

$$[\Phi_{Rabc}] = [I_R][i_{Rabc}] + [M_{SR}]^T [i_{Sabc}] \quad (D-2-4)$$

Avec $[R_x]$ et $[I_x]$ et $[M_{SR}]$ sont respectivement les résistances, les inductances et les mutuelles, sachant que $x = S, R$ et S représentant le stator et R le rotor.

A-2- Transformation de Park :

La transformation de Park fait projeter les enroulements d'axes a, b et c sur deux axes fictifs d et q en quadrature. Cette transformation est définie par la matrice de Park :

$$P(\theta_x) = \sqrt{2/3} \begin{bmatrix} \cos(\theta_x) & \cos(\theta_x - 2\pi/3) & \cos(\theta_x + 2\pi/3) \\ -\sin(\theta_x) & -\sin(\theta_x - 2\pi/3) & -\sin(\theta_x + 2\pi/3) \\ 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \quad .x = S, R. \quad (D-2-5)$$

A-3- Transformation dq → αβ :

Elle est définie par la matrice
$$\begin{bmatrix} \cos(\theta_s) & -\sin(\theta_s) \\ \sin(\theta_s) & \cos(\theta_s) \end{bmatrix} \quad (D-2-6)$$

Où $\dot{\theta}_s = \omega_s$

A-4-Modelle de la machine alimenter en tension :

En appliquant La transformation de park au modèle de la machine et en adoptants la notation suivante $x = (x_1, x_2, x_3, x_4, x_5) = (i_{ds}, i_{qs}, \phi_{dR}, \phi_{qR}, \omega)$ on obtient les équations :

Repère dq :

$$\dot{x}_1 = -a_0 x_1 + b_0 x_3 + \omega_s x_2 + c_0 x_4 x_5 + e_0 V_{ds} \quad (D-2-7)$$

$$\dot{x}_2 = -a_0 x_2 + b_0 x_4 - \omega_s x_1 - c_0 x_3 x_5 + e_0 V_{qs} \quad (D-2-8)$$

$$\dot{x}_3 = -a x_3 + e x_1 + \omega_s x_4 - x_4 x_5 \quad (D-2-9)$$

$$\dot{x}_4 = -a x_4 + e x_2 - \omega_s x_3 + x_3 x_5 \quad (D-2-10)$$

$$\dot{x}_5 = -d x_5 + c (x_2 x_3 - x_1 x_4) - b C_r \quad (D-2-11)$$

Repère αβ :

$$\dot{x}_1 = -a_0 x_1 + b_0 x_3 + c_0 x_4 x_5 + e_0 V_{ds} \quad (D-2-12)$$

$$\dot{x}_2 = -a_0 x_2 + b_0 x_4 - c_0 x_3 x_5 + e_0 V_{qs} \quad (D-2-13)$$

$$\dot{x}_3 = -a x_3 + e x_1 - x_4 x_5 \quad (D-2-14)$$

$$\dot{x}_4 = -a x_4 + e x_2 + x_3 x_5 \quad (D-2-15)$$

$$\dot{x}_5 = -d x_5 + c (x_2 x_3 - x_1 x_4) - b C_r \quad (D-2-16)$$

Où la dernière équation définit l'équation mécanique et C_r le couple résistant

A-4-Modelle de la machine alimenter en courant :

Repère dq :

$$\dot{x}_3 = -a x_3 + \omega_s x_4 - x_4 x_5 + e i_{ds} \quad (D-2-17)$$

$$\dot{x}_4 = -a x_4 - \omega_s x_3 + x_3 x_5 + e i_{qs} \quad (D-2-18)$$

$$\dot{x}_5 = -d x_5 + c (i_{ds} x_3 - i_{qs} x_4) - b C_r \quad (D-2-19)$$

Repère $\alpha\beta$:

$$\dot{x}_3 = -a x_3 - x_4 x_5 + e i_{ds} \quad (D-2-20)$$

$$\dot{x}_4 = -a x_4 + x_3 x_5 + e i_{qs} \quad (D-2-21)$$

$$\dot{x}_5 = -d x_5 + c (i_{ds} x_3 - i_{qs} x_4) - b C_r \quad (D-2-22)$$

$$a_0 = (R_S / \sigma L_S) + R_R (1 - \sigma) / (\sigma L_R) \quad , \quad a = L_R / R_R \quad , \quad c = R_R / L_R \quad .$$

$$b_0 = (1 - \sigma) / (\sigma M \frac{L_R}{R_R}) \quad , \quad e = \frac{M R_R}{L_R} \quad , \quad c_0 = (1 - \sigma) / (\sigma M) \quad .$$

$$c = (p^2 M) / (J L_R) \quad , \quad e = 1 / (\sigma L_S) \quad , \quad b = p / J \quad .$$

REFERENCES BIBLIOGRAPHIQUES

- [1] ARAGO 14, «logique flou », observation Français des Techniques avancées, Masson, Paris, 1994.
- [2] R. BABUSKA, DATA – Driven Modeling : Transparency and Complerty Issues, 1998.
- [3] BABUSKA, M. Setnes, U. Kaymat and H.R Van Nanta ; Rule base simplification with similarity measure's. IEEE International Confana on Fuzzy System USA, pp. 1642 – 1647.
- [4] R. BABUSKA, « Fuzzy Modeling for Control », Boston, USA. (1998).
- [5] R. BABUSKA , « Promissing Fuzzy Modeling and control Methodologies for Industrial Application. (1999).
- [6] R. BABUSKA, « Fuzzy Logic for Engeneering Applications ». Tudelft, March, 99.
- [7] R. BABUSKA, « Fuzzy systems, Modeling and Identification ». Tudelft 1997.
- [8] R. BABUSKA, « Fuzzy Sets and Systems ». ITS. Tudelft 1999.
- [9] R. BABUSKA, « Fuzzy Sets and Fuzzy Logic » Tudelft February 2000.
- [10] B. BOUCHON, Meunnier, « La Logique Floue ». 1993.
- [11] H. Bühler, « Réglage par Logique Floue », PPR, 1994.
- [12] H. CHEKIREB, A. FARAH, « Réglage flou par la méthode de Sugeno du flux et de la vitesse d'une Machine Asynchrone », Conférence Internationale SSA2. Blida 1999.
- [13] P. FALL, N. Lassouaoui, « Application de la logique flou à la commande de bras manipulateur et réalisation d'un logiciel de simulation et de commande », Thèse PFE, ENP 1998.
- [14] M. JAMSHIDI, N. VADIEE, T.J. ROV, « Fuzzy Logic an Control ». 1993.
- [15] LABIOD, « Commade Adaptation par les systèmes flous, Application aux robots , manipulateurs », Thèse de Magister, ENP 1998.

- [31] J. SHING, R. IANG, « Self Learning Fuzzy Controllers Based on Temporal Back – Propagation ». IEEE, Transaction on Neural Networks, Vol3, Sep 1992.
- [32] M. SUGENO et T. TAKAGI, « Fuzzy identification of Systems and its Application to Modeling and Control », IEEE Trans Syst. Man and Cybern, Vol.SMC-15, n°1 pp 116 – 132, 1985.
- [33] M. SUGENO and G.T. KANG, « Structure Identification of Fuzzy Model ». Fuzzy Sets and Systems 28, (1988) pp. 15 – 33.
- [34] Z. TIDJANI, « Commande Décentralisée Floue Appliquée aux Robots Manipulateurs », Thèse Magister, ENP 1999.
- [35] H.B VERBREGGEN, « Fuzzy Control and Conventional Control », Fuzzy Sets and Systems 90 (1997) pp 151 – 160.
- [36] L. ZADEH, « Fuzzy Sets », Information and Control, Vol8, 1965.
- [37] L. ZADEH, « Fuzzy Logic », IEEE, pp 83 – 93, April 1988.

- [16] C.C.LEE, « Fuzzy Logic in Control System : Fuzzy Logic Controller - Part I » IEEE Trans. Syst., Man. and Cybern, Vol. 20, pp.404 – 418, Feb 1990.
- [17] C.C.LEE, « Fuzzy Logic in Control System : Fuzzy Logic Controller - Part II » IEEE Trans. Syst, Man. And Cybern, Vol. 20, pp.419 – 435, Feb 1990.
- [18] M. LOODINI, « Modelisation, Analyse et Méthodologies de Commande Linguistique Floue d'un bras manipulateur de Robot flexible », Magister, ENP 1997.
- [19] J.M. MENDEL, « Fuzzy Logic Systems for Engineering : A Tutorial ». IEEE. pp. 345 – 377. 1995.
- [20] S. MOLLOV, R. BABUSKA, « MIMO Predictive Control by Multipl – Step Linearization of Takagi-Sugeno Fuzzy Models ». TUDelft. IFAC 1998.
- [21] S. MOLLOV, « Extraction of Local Linear Models from Takagi-Sugeno Fuzzy Model with Application to Model – based predictive Control » TUDelft 1999.
- [22] Y. NAKAMORI, « Fuzzy Modeling for Adaptive Process Control ». Fuzzy Control Systems, pp 296 – 313. Konan University 8-9-1 . Okamoto, Japan.
- [23] PEDRYCZ. W, « Fuzzy Control and Fuzzy Systèms », 1988.
- [24] J.A. RUBOS. « Predictive Control by Local Linéarization of Takagi-Sugeno Fuzzy Model ». TUDelft 1998.
- [25] J.A . RUBOS, S. MOLLOV, R. BABUSKA, H.B. VERBRUGGEN, «Fuzzy Model Based Predictive Control Using Takagi-Sugeno Models ». International Journal of Approximate Reasonnig, March 1999.
- [26] H. SCHOLTEN, « Logique Floue et Régulateur PID », Publitronec, Paris.1994.
- [27] M. SETNES, R. BABUSKA and H.B VERBREGGEN : « Fuzzy Modeling and Similarity Analysis ». TUDelft 1998.
- [28] M. SETNES, « Fuzzy Rule Base Simplification using Similarity Measures », M. SC, Thesis, July 1995.
- [29] M. SETNES and R. BABUSKA, « Fuzzy Logic Control, Advances in Applications » TUDelft, April 1999.
- [30] M. SETNES, H. ROUBOS « Transparent Fuzzy Modeling Using Fuzzy Clustering and GA's ». TUDelft 1999.