

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
CIÊNCIA DA COMPUTAÇÃO

**Modelagem Fuzzy usando Agrupamento Condicional**

Tatiane Marques Nogueira

São Carlos  
Agosto/2008

**Universidade Federal de São Carlos**

**Centro de Ciências Exatas e de Tecnologia**  
**Programa de Pós-Graduação em Ciência da Computação**

*“Modelagem Fuzzy usando  
Agrupamento Condicional”*

TATIANE MARQUES NOGUEIRA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação

Membros da Banca:



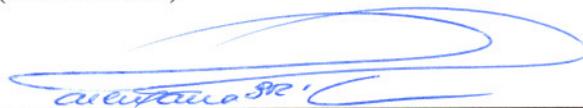
---

Profa. Dra. Heloisa de Arruda Camargo  
(Orientadora - DC/UFSCar)



---

Profa. Dra. Maria do Carmo Nicoletti  
(DC/UFSCar)



---

Profa. Dra. Marley Maria B. Rebuzzi Vellasco  
(PUC-Rio)

São Carlos  
Agosto/2008

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária da UFSCar**

N778mf

Nogueira, Tatiane Marques.  
Modelagem fuzzy usando agrupamento condicional /  
Tatiane Marques Nogueira. -- São Carlos : UFSCar, 2008.  
85 f.

Dissertação (Mestrado) -- Universidade Federal de São  
Carlos, 2008.

1. Geração automática de regras Fuzzy. 2. Fuzzy logic. 3.  
Método de agrupamento. 4. Sistema Fuzzy. I. Título.

CDD: 004.3 (20<sup>a</sup>)

# Agradecimentos

---

---

Primeiramente, agradeço a Deus pela força que tanto precisei, por ter me mantido firme nos momentos em que pensei em desistir e por me fazer um pouco melhor depois disto.

Agradeço aos meus pais pelo apoio e pelo amor incondicional, por terem compreendido que eu estava distante fisicamente, mas sempre perto de coração, e que por isto nenhuma cidade é tão distante assim que possamos realmente nos afastar. Pelo contrário, minha ausência se transformou em festa na minha chegada, que por sua vez se transformou em motivação para continuar. Obrigada mainha! Obrigada painho!

Agradeço aos meus irmãos que me ajudaram a cuidar dos meus pais na minha ausência e fizeram com que eles compreendessem que um filho não substitui o outro, mas pode ajudar na hora da saudade. Obrigada Ti! Obrigada Digo!

Agradeço à Ricardo, meu noivo, por ter sido namorado, pai, irmão, amigo, colega, enfim, um grande companheiro nesta jornada! Por ter sido o exemplo de motivação que eu precisava, acreditando mais em mim do que eu mesma e me mostrando que realmente vale a pena se dedicar um pouco mais. Obrigada, meu amor, por revisar o código, por revisar o texto, por me dar broncas na hora certa, pelo abraço quando eu sofria de saudade de casa, pelo carinho demonstrado nos pequenos detalhes.

Agradeço à Professora Heloisa Camargo, minha orientadora, por ter me ensinado muito mais do que alguns cálculos matemáticos, por ter me mostrado por meio da sua serenidade e tranquilidade que realmente "a pressa é inimiga da perfeição". Por, talvez sem perceber, sempre me acalmar naqueles momentos em que estive "ligada no 220". Obrigada, Heloisa, por suprir todos os requisitos que considero essenciais para um excelente orientador: calma, segurança e confiança.

Agradeço às colegas da república jardim (as que passaram e as que ficaram): Val, Amanda, Luana, Marília, Tatiana, Eliane, Kamila, Deyse, Renata, Paula e Cris. Por terem sido minha família aqui em São Carlos e por terem feito dos almoços de domingo uma grande festa, reunião de negócios, debates, enfim, tudo que se espera de uma mulherada

reunida. Obrigada, meninas!

Agradeço à Cris, por ter permitido que eu fosse tia e me mostrado o quanto uma criança alegre uma casa! Por me fazer enxergar as coisas por um outro ângulo e compreender que o mestrado não foi tão difícil assim para mim. Obrigada ao seu filho, Gui, por permitir à tia Tatinha a felicidade de se jogar no tapete e esquecer que o dia foi difícil e que aquele código que passei semanas fazendo, simplesmente não funcionou. Obrigada, meu bb!

Agradeço ao Marcos Cintra, colega e irmão pela orientadora, por guiar meus passos no mundo fuzzy, por ter me recebido no mestrado e ter passado todo seu código para que eu começasse a "botar a mão na massa", e por não me deixar desistir do inglês. Thank you very much!

Agradeço aos colegas do GEMe (Grupo de Estudos do Mestrado): Débora Corrêa, Eliane Nascimento, Reginaldo Gotardo, Simone Borges, Leonardo Botega, Vinícios Durelli, Ricardo Rios, Alexandre Mello, Márcio Roberto e Cristiane Santana. Quantas noites na salinha de estudos! Valeu mesmo a pena: ganhei uma dissertação e amigos para sempre. Obrigada, pessoal!

Agradeço à todos os professores do Departamento de Computação da UFSCar, em especial à Maria do Carmo Nicoletti e Estavam Hruschka por terem passado segurança em suas aulas e na banca de qualificação, transmitindo os seus conhecimentos da melhor maneira possível. Por terem sido realmente grandes exemplos de professores.

Agradeço à FAPESB pelo apoio financeiro.

Enfim, só tenho mesmo que muito agradecer! Obrigada, mais uma vez, à todos que direta ou indiretamente contribuíram para a conclusão do meu mestrado.

# Resumo

---

---

A combinação de sistemas fuzzy com algoritmos de agrupamento tem grande aceitação na comunidade científica devido; principalmente, a sua aderência ao princípio de balanceamento de vantagens da inteligência computacional, no qual metodologias diferentes colaboram entre si, potencializando a utilidade e aplicabilidade dos sistemas resultantes.

A modelagem fuzzy usando algoritmos de agrupamento apresenta a transparência e facilidade de compreensão típica dos sistemas fuzzy lingüísticos ao mesmo tempo em que se beneficia das possibilidades de redução da dimensionalidade por intermédio do agrupamento.

Neste trabalho é apresentado o método Fuzzy-CCM (Fuzzy Conditional Clustering based Modeling), que consiste de uma nova abordagem de Modelagem Fuzzy baseada no algoritmo de Agrupamento Fuzzy Condicional, cujo objetivo é prover novos meios de tratar a questão da interpretabilidade de bases de regras fuzzy.

Com o método Fuzzy-CCM, o balanço entre interpretabilidade e acuidade de regras fuzzy é tratado por meio da definição de contextos formados com um pequeno número de variáveis de entrada e a geração de grupos condicionados por estes contextos. As regras são geradas em um formato diferente, que contém variáveis lingüísticas e grupos no seu antecedente. Alguns experimentos foram executados usando diferentes domínios de conhecimento a fim de validar a abordagem proposta, comparando os resultados obtidos usando a nova abordagem com os resultados obtidos usando os métodos Wang&Mendel e Fuzzy C-Means.

A fundamentação teórica, as vantagens do método, os experimentos e os resultados obtidos são apresentados e discutidos.

**Palavras-chave:** Sistemas Fuzzy, Algoritmo de Agrupamento Condicional, Interpretabilidade, Modelagem Fuzzy.

# Abstract

---

---

The combination of fuzzy systems with clustering algorithms has great acceptance in the scientific community mainly due to its adherence to the advantage balance principle of computational intelligence, in which different methodologies collaborate with each other potentializing the usefulness and applicability of the resulting systems.

Fuzzy Modeling using clustering algorithms presents the transparency and comprehensibility typical of the linguistic fuzzy systems at the same time that benefits from the possibilities of dimensionality reduction by means of clustering.

In this work is presented the Fuzzy-CCM method (Fuzzy Conditional Clustering based Modeling) which consists of a new approach for Fuzzy Modeling based on the Fuzzy Conditional Clustering algorithm aiming at providing new means to address the topic of interpretability of fuzzy rules bases.

With the Fuzzy-CCM method the balance between interpretability and accuracy of fuzzy rules is dealt with through the definition of contexts defined by a small number of input variables and the generation of clusters induced by these contexts. The rules are generated in a different format, with linguistic variables and clusters in the antecedent. Some experiments have been carried out using different knowledge domains in order to validate the proposed approach by comparing the results with the ones obtained by the Wang&Mendel and conventional Fuzzy C-Means methods.

The theoretical foundations, the advantages of the method, the experiments and results are presented and discussed.

**Keywords:** Fuzzy Systems, Conditional Clustering Algorithm, Interpretability, Fuzzy Modeling.

# Sumário

---

---

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação e Objetivos . . . . .	2
1.2	Organização do Trabalho . . . . .	4
<b>2</b>	<b>Sistemas Fuzzy</b>	<b>5</b>
2.1	Conjuntos Fuzzy . . . . .	5
2.1.1	Funções de Pertinência . . . . .	6
2.1.2	Operações sobre Conjuntos Fuzzy . . . . .	8
2.1.3	Relações Fuzzy . . . . .	10
2.2	Computação com Regras Fuzzy . . . . .	12
2.2.1	Variáveis Linguísticas . . . . .	13
2.2.2	Regras Fuzzy . . . . .	14
2.2.3	Inferência Fuzzy . . . . .	15
2.3	Sistemas Fuzzy . . . . .	16
2.3.1	Sistemas Fuzzy Baseados em Regras . . . . .	16
2.3.2	Tipos de Sistemas Fuzzy . . . . .	18
2.3.3	Defuzificação . . . . .	19
2.3.4	Sistemas de Classificação Fuzzy . . . . .	21
2.4	Considerações Finais . . . . .	23

<b>3</b>	<b>Algoritmos de Agrupamento</b>	<b>24</b>
3.1	Medidas de Proximidade . . . . .	25
3.2	Algoritmos de Agrupamento . . . . .	27
3.2.1	Algoritmo K-Means . . . . .	29
3.2.2	Algoritmo Fuzzy C-Means . . . . .	30
3.2.3	Algoritmo Fuzzy C-Means Condicional . . . . .	31
3.3	Considerações finais . . . . .	35
<b>4</b>	<b>Modelagem Fuzzy usando Agrupamento Condicional</b>	<b>36</b>
4.1	Interpretabilidade $\times$ Acuidade . . . . .	37
4.2	Abordagens de Modelagem Fuzzy usando Agrupamento . . . . .	39
4.3	Método Proposto – Fuzzy-CCM . . . . .	43
4.3.1	Construção de Regras Fuzzy usando FCMC . . . . .	43
4.3.2	Método de Inferência . . . . .	48
4.4	Considerações Finais . . . . .	51
<b>5</b>	<b>Resultados Experimentais</b>	<b>52</b>
5.1	Considerações Preliminares . . . . .	52
5.2	Experimentos, Resultados e Análises . . . . .	53
5.2.1	Concepção da Abordagem . . . . .	54
5.2.2	Ampliação da Abordagem . . . . .	59
5.3	Considerações Finais . . . . .	66
<b>6</b>	<b>Conclusão e Trabalhos Futuros</b>	<b>67</b>
<b>A</b>	<b>Saída ReliefF</b>	<b>69</b>
A.1	Iris . . . . .	69
A.2	Bupa . . . . .	70
A.3	Diabetes . . . . .	71
A.4	Wine . . . . .	72
A.5	Vehicle . . . . .	73
A.6	AutoMpg . . . . .	74
A.7	Pollution . . . . .	75



# Lista de Figuras

---

---

2.1	Funções de Pertinência: (a) Triangular, (b) Trapezoidal e (c) Gaussiana . . . . .	7
2.2	Operações sobre Conjuntos Fuzzy: (a) Intersecção e (b) União . . . . .	8
2.3	Formas de representação de uma relação fuzzy binária: (a) Matriz e (b) Dígrafo	11
2.4	Variável Lingüística: Velocidade . . . . .	13
2.5	Base de Conhecimento de um SFBR . . . . .	17
2.6	Sistema Fuzzy . . . . .	18
2.7	Centro de Máximos Ponderado . . . . .	20
2.8	Média de Elementos Máximos . . . . .	21
3.1	Etapas do Processo de Agrupamento . . . . .	28
3.2	Distribuição dos Padrões . . . . .	33
3.3	Agrupamento não-Condiciona: (a) Dados pertencentes ao grupo A; (b) Dados pertencentes ao grupo B. . . . .	34
3.4	Agrupamento Condiciona: (a) Dados pertencentes ao grupo A; (b) Dados pertencentes ao grupo B. . . . .	34
4.1	Possibilidades da Modelagem Fuzzy (CASILLAS et al., 2003) . . . . .	37
4.2	Construção de Regras Fuzzy: (a) Seleção de Atributos, (b) Geração de Regras usando o FCMC, (c) Base de Regras . . . . .	44
4.3	Divisão do conjunto de dados com classe discreta . . . . .	46

4.4	Divisão do conjunto de dados com classe contínua . . . . .	47
4.5	Inferência para domínios com classe discreta . . . . .	49
4.6	Inferência para domínios com classe contínua . . . . .	50

# Lista de Algoritmos

---

---

1	K-Means . . . . .	29
2	Fuzzy C-Means . . . . .	32

# Lista de Tabelas

---

---

3.1	Características dos Padrões . . . . .	33
3.2	Centros dos grupos . . . . .	34
5.1	Características dos Domínios . . . . .	53
5.2	Taxas de Classificação Correta - Iris . . . . .	55
5.3	Taxas de Classificação Correta - Bupa . . . . .	56
5.4	Taxas de Classificação Correta - Diabetes . . . . .	56
5.5	Taxas de Classificação Correta - Wine . . . . .	58
5.6	Taxas de Classificação Correta - Vehicle . . . . .	59
5.7	AutoMpg: Taxas de Erro médio e desvio padrão usando os métodos WM e FCM para geração das regras e MEM para defuzificação . . . . .	60
5.8	AutoMpg: Taxas de Erro médio e desvio padrão usando os métodos WM e FCM para geração das regras e CMP para defuzificação . . . . .	60
5.9	AutoMpg: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e MEM para defuzificação . . . . .	61
5.10	AutoMpg: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e CMP para defuzificação . . . . .	61
5.11	Pollution: Taxas de Erro usando os métodos WM e FCM para geração das regras e MEM para defuzificação . . . . .	62

5.12 Pollution: Taxas de Erro usando os métodos WM e FCM para geração das regras e CMP para defuzificação . . . . .	63
5.13 Pollution: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e MEM para defuzificação . . . . .	64
5.14 Pollution: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e CMP para defuzificação . . . . .	65

# Lista de Abreviaturas

---

---

- AA:** *Algoritmos de Agrupamento*
- BC:** *Base de Conhecimento*
- BD:** *Base de Dados*
- BR:** *Base de Regras*
- CMP:** *Centro de Máximos Ponderado*
- FCM:** *Fuzzy C-means*
- FCMC:** *Fuzzy C-means Condicional*
- MFL:** *Modelagem Fuzzy Lingüística*
- MEM:** *Média de Elementos Máximos*
- RF:** *Regras Fuzzy*
- SCF:** *Sistemas de Classificação Fuzzy*
- SF:** *Sistema Fuzzy*
- SFBR:** *Sistema Fuzzy Baseado em Regras*
- WM:** *Wang&Mendel*

---

# Introdução

---

A Computação Flexível (Soft Computing) é um paradigma abrangente que agrega, tradicionalmente, as metodologias de redes neurais artificiais, dos sistemas fuzzy e da computação evolutiva. A evolução observada nesse campo nos últimos anos reforça a tendência cada vez mais evidente de tratar seus componentes como metodologias complementares, que dispõem de recursos adequados ao desenvolvimento de sistemas híbridos com as mais diversas finalidades. As potencialidades dos componentes envolvidos e as suas possibilidades de combinação oferecem uma nova visão para resolução de problemas complexos que muitas vezes não poderiam ser abordados de outra forma, particularmente no que diz respeito ao tratamento da incerteza e imprecisão (PEDRYCZ, 1998).

O conceito de Modelagem Fuzzy, inicialmente introduzido por ZADEH (1971, 1979) e intensivamente investigado desde então, tem sido usado com sucesso em vários domínios. Muitas abordagens híbridas para gerar automaticamente modelos fuzzy a partir de conjuntos de dados podem ser encontradas na literatura, sendo as técnicas mais bem sucedidas aquelas baseadas em Algoritmos de Agrupamento (LIAO et al., 1997), Redes Neurais (JANG et al., 1997) e Algoritmos Genéticos (CORDÓN et al., 2004).

Os Sistemas Fuzzy (SF) incorporam metodologias e técnicas fundamentadas nos con-

ceitos de conjuntos fuzzy e da lógica fuzzy (ZADEH, 1965) para representar e processar informações lingüísticas, e fornecer mecanismos para tratar a imprecisão e a incerteza inerentes ao conhecimento do mundo real. Por isso têm sido aplicados com sucesso em diversas áreas, como: controle automático, classificação e reconhecimento de padrões, tomada de decisão, previsão de séries temporais e robótica (PEDRYCZ, 1996b) (PEDRYCZ;GOMIDE, 1998) (DUMITRESCU et al., 2000).

Embora a utilidade dos SF seja reconhecida, eles não possuem mecanismos próprios para se efetuar o aprendizado automático, ou seja, para se construir o sistema automaticamente a partir de conjuntos de dados, motivo pelo qual se investiga a combinação dos SF com outras metodologias que ofereçam essa capacidade.

Este trabalho propõe um método para gerar automaticamente regras fuzzy usando o Algoritmo de Agrupamento Fuzzy C-means Condicional proposto por PEDRYCZ (1996a), a fim de obter como resultado sistemas interpretáveis pela redução no tamanho das regras, sem prejuízo do desempenho do sistema obtido.

Assim, aplicando-se algoritmos de agrupamento condicional para geração de regras de uma base de conhecimento fuzzy, obteve-se como resultado deste trabalho sistemas fuzzy que apresentam bom grau de desempenho, além de melhorias quanto à clareza, pela redução do número de variáveis que aparecem nas regras.

## 1.1 Motivação e Objetivos

Um tipo de sistema fuzzy resultante de técnicas de modelagem fuzzy são os Sistemas Fuzzy Baseados em Regras (SFBR), os quais são compostos por dois componentes principais: a Base de Conhecimento e o Mecanismo de Inferência. A Base de Conhecimento é composta pela Base de Regras, que armazena o conjunto de regras fuzzy que modelam o conhecimento sobre um determinado problema, e pela Base de Dados, que contém as definições dos conjuntos fuzzy relacionados aos termos lingüísticos usados nas regras fuzzy. O Mecanismo de Inferência é responsável pelo processamento das regras, o qual é realizado por um método de raciocínio fuzzy.

Uma das tarefas mais importantes e difíceis no desenvolvimento de um SFBR é gerar a sua Base de Conhecimento. Nos últimos anos, métodos automáticos têm sido utiliza-

dos para geração da Bases de Conhecimento, que consistem em extrair automaticamente o conhecimento necessário a partir de dados numéricos que representam amostras ou exemplos do problema.

A interpretabilidade ou transparência de regras fuzzy tem recebido crescente atenção da comunidade científica devido a sua importância em alguns domínios. Geralmente, a interpretabilidade de sistemas fuzzy depende de três aspectos (JIN, 2000):

1. A quantidade de conjuntos fuzzy das variáveis não deve ser grande demais nem pequena demais. Deve ser um número razoável de maneira a permitir a compreensão do significado de cada uma;
2. As regras fuzzy na base de regras devem ser consistentes, significando que não deve haver regras conflitantes;
3. O número de variáveis na premissa da regra deve ser tão pequeno quanto possível.

Muitas definições e maneiras de prover a interpretabilidade de SFBR são relatadas na literatura. Diferentes metodologias baseadas na otimização da estrutura e parâmetros podem ser encontradas em (JIN, 2000) (GUILLAUME, 2001) (CASTELLANO et al., 2002) (CASTRO;CAMARGO, 2005). Agrupamento tem sido usado na modelagem fuzzy de diferentes maneiras (HADJULI;WERTZ, 2002) (GOMEZ-SKARMETA et al., 1999) (ABONYI et al., 2002). Métodos enfatizando o balanço entre interpretabilidade e desempenho foram relatados em (SETNES et al., 1998) (PEDRYCZ;VASILAKOS, 1999) (PEDRYCZ;KWAK, 2006b) (PEDRYCZ;KWAK, 2006a).

Será apresentado neste trabalho o método Fuzzy-CCM (*Fuzzy Conditional Clustering based Modeling*) que consiste de uma nova abordagem para modelagem fuzzy usando agrupamento condicional. Neste método, o agrupamento condicional é executado a partir de um contexto definido por uma ou mais variáveis no espaço de entrada agrupando os dados com base nos demais valores das variáveis que não foram usadas na definição do contexto. As regras fuzzy resultantes deste método tem um formato diferente, o qual mistura a especificação de variáveis lingüísticas e seus valores com grupos de dados na premissa das regras. O processo de geração das regras é repetido para subconjuntos de dados separados pelos valores de saída, sejam eles contínuos ou discretos. O valor de saída é usado

para definir o conseqüente das regras geradas pela aplicação do método ao grupo de dados correspondente.

O método Fuzzy-CCM permite que a base de regras fuzzy seja interpretada sob diferentes pontos de vista, uma vez que diferentes versões da mesma base de conhecimento podem ser construídas nas quais o número de variáveis de entrada e as variáveis propriamente ditas usadas para definir o contexto podem variar.

## 1.2 Organização do Trabalho

Assim, visando alcançar o objetivo proposto, a pesquisa foi organizada com base nos principais assuntos focalizados: Lógica Fuzzy, Algoritmos de Agrupamento, Interpretabilidade e Acuidade de Sistemas Fuzzy. Com o subsídio formal e teórico dos assuntos anteriores, foi desenvolvida a abordagem e avaliada por meio de experimentos.

Este trabalho descreve o principal resultado deste trabalho de pesquisa, i.e., o método Fuzzy-CCM e está organizado da seguinte maneira:

*Capítulo 2 - Sistemas Fuzzy.* Neste capítulo são descritos os principais conceitos da Lógica Fuzzy, destacando os métodos para inferência e classificação fuzzy.

*Capítulo 3 - Algoritmos de Agrupamento.* Neste capítulo são abordados os principais conceitos sobre agrupamento de dados, destacando algumas possíveis medidas de similaridade, e posteriormente descritos os algoritmos relevantes para o trabalho: K-means, Fuzzy C-means e Fuzzy C-means Condicional.

*Capítulo 4 - Modelagem Fuzzy usando Agrupamento Condicional.* Neste capítulo são apresentados os princípios e abordagens utilizadas para prover o balanço interpretabilidade-acuidade de Sistemas Fuzzy que usam agrupamento no processo de modelagem, bem como o método Fuzzy-CCM, que é uma nova proposta de abordagem para modelagem fuzzy desenvolvida neste trabalho, com suas características e metodologia.

*Capítulo 5 - Resultados Experimentais.* Este capítulo apresenta os resultados obtidos na validação do método proposto em vários domínios de conhecimento.

*Capítulo 6 - Conclusão e Trabalhos Futuros.* Este capítulo apresenta as conclusões obtidas com as análises/avaliação feitas a partir dos experimentos e apresenta algumas sugestões de melhorias com o objetivo de refinar a proposta.

---

# Sistemas Fuzzy

---

---

Sistemas que fazem uso da teoria de conjuntos fuzzy são chamados genericamente de Sistemas Fuzzy (SF). Neste capítulo serão apresentados os principais conceitos acerca da teoria de conjuntos fuzzy, das regras fuzzy e do raciocínio aproximado, a fim de compreender o funcionamento de um sistema fuzzy.

Na teoria de conjuntos fuzzy, ao invés de afirmar que uma pessoa com certeza está ou não doente, podemos dizer, dados os sintomas, que é “possível” que a pessoa esteja doente, “talvez” a pessoa não esteja doente e outras afirmações. Os conjuntos fuzzy também permite-se a representação de conceitos vagos expressos por meio de termos lingüísticos, como temperatura alta, custo baixo ou clima frio, por exemplo (KLIR;YUAN, 1995).

A seguir, são descritos alguns conceitos sobre conjuntos fuzzy, funções de pertinência, operações sobre conjuntos fuzzy e relações fuzzy com o objetivo de compreender melhor a proposta.

## 2.1 Conjuntos Fuzzy

Na teoria clássica, um conjunto  $A$  é definido por meio da função característica

$$A : X \longrightarrow \{0, 1\}$$

em que elementos do conjunto universo  $X$  pertencem totalmente ao conjunto  $A$ :  $A(x) = 1$  ou não pertencem ao conjunto:  $A(x) = 0$ .

Já os conjuntos fuzzy são usados para modelar informação imprecisa, e podem ser abordados como uma generalização da noção clássica de conjuntos. Nos conjuntos fuzzy, os elementos pertencem ao conjunto com um certo grau, que usualmente é um valor entre 0 e 1. Quando os graus de pertinência de um elemento assumem os valores 0 ou 1, tem-se o caso clássico de não pertinência total ou pertinência total do elemento ao conjunto, respectivamente. Assim, a definição de um conjunto fuzzy é obtida ampliando-se o contradomínio da função característica  $\{0, 1\}$  para o intervalo  $[0, 1]$ , de forma a atribuir o grau com que um elemento pertence a um conjunto fuzzy. A generalização da função característica passa a ser chamada de função de pertinência, a qual definirá o conjunto fuzzy, associando elementos de um dado conjunto universo  $X$  a números reais do intervalo  $[0, 1]$ .

Logo, um conjunto fuzzy  $A$  no conjunto universo  $X$  é caracterizado pela função

$$A : X \longrightarrow [0, 1]$$

e  $A(x)$  define o grau com que o elemento  $x$  pertence ao conjunto fuzzy  $A$  (KLIR;YUAN, 1995).

### 2.1.1 Funções de Pertinência

A forma de um conjunto fuzzy depende da função de pertinência que as define. Algumas formas de funções de pertinência tornaram-se mais conhecidas e são utilizadas frequentemente, pela sua simplicidade e facilidade de implementação. Os três tipos de funções de pertinência mais comuns são (Figura 2.1):

- **Função Triangular (Figura 2.1(a))**, definida por meio dos parâmetros  $a$ ,  $m$  e  $b$ , tal que  $a \leq m \leq b$ .

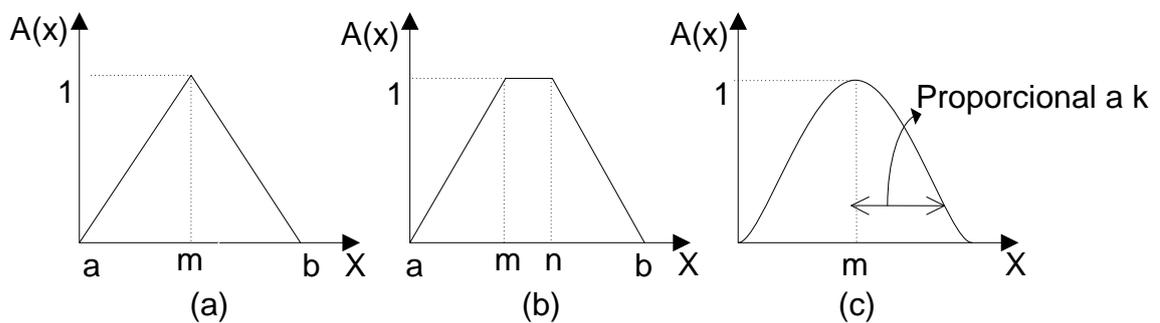
$$A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } x \in (a, m) \\ 1 & \text{se } x = m \\ \frac{b-x}{b-m} & \text{se } x \in (m, b) \\ 0 & \text{se } x \geq b \end{cases}$$

- **Função Trapezoidal (Figura 2.1(b))**, definida por meio dos parâmetros  $a, m, n$  e  $b$ , tal que  $a \leq m < n \leq b$ .

$$A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } x \in (a, m) \\ 1 & \text{se } x \in [m, n] \\ \frac{b-x}{b-n} & \text{se } x \in (n, b) \\ 0 & \text{se } x \geq b \end{cases}$$

- **Função Gaussiana (Figura 2.1(c))**, definida por meio dos parâmetros  $m$  e  $k$ , tal que  $k > 0$ .

$$A(x) = e^{-k(x-m)^2}$$



**Figura 2.1:** Funções de Pertinência: (a) Triangular, (b) Trapezoidal e (c) Gaussiana

### 2.1.2 Operações sobre Conjuntos Fuzzy

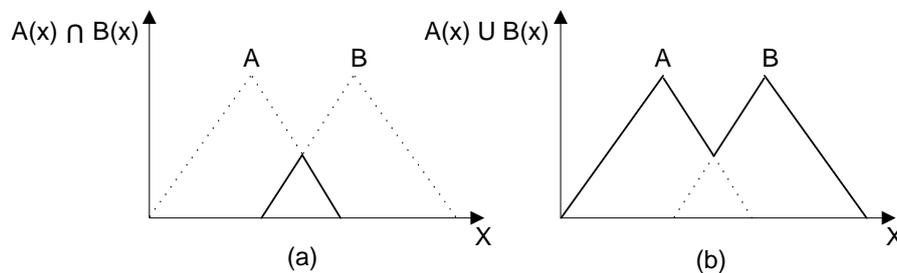
As operações de complemento, intersecção e união existentes na teoria de conjuntos convencionais podem ser estendidas aos conjuntos fuzzy, as quais podem ser implementadas por meio de uma classe de operadores. Uma forma particular de definir essas operações, conhecidas como operações-padrão, destaca-se na teoria de conjuntos fuzzy por ser a mais utilizada para essas definições.

Seja  $X$  o conjunto universo e  $x$  um elemento de  $X$ . O complemento padrão, intersecção padrão (Figura 2.2(a)), e união padrão (Figura 2.2(b)), de conjuntos fuzzy  $A$  e  $B$  sobre  $X$ , são definidos para todo  $x \in X$ , respectivamente, por:

$$\bar{A}(x) = 1 - A(x) \quad (2.1)$$

$$(A \cap B)(x) = \text{minimo} \{A(x), B(x)\} \quad (2.2)$$

$$(A \cup B)(x) = \text{maximo} \{A(x), B(x)\} \quad (2.3)$$



**Figura 2.2:** Operações sobre Conjuntos Fuzzy: (a) Intersecção e (b) União

Para cada uma das três operações-padrão, existe uma classe de operadores que podem ser utilizados na generalização dessas operações. Os operadores que podem ser utilizados na generalização das operações de intersecção e união são referidos na literatura como t-normas e s-normas, respectivamente.

Uma t-norma  $t$  é uma operação binária da forma  $t : [0, 1]^2 \rightarrow [0, 1]$  que deve satisfazer os axiomas abaixo, para todo  $a, b, c \in [0, 1]$  (KLIR;YUAN, 1995):

$$a_1. \text{ Condição Limite: } t(a, 1) = a, t(0, 0) = 0$$

$a_2$ . *Monotonicidade*:  $b \leq c$  implica  $t(a, b) \leq t(a, c)$

$a_3$ . *Comutatividade*:  $t(a, b) = t(b, a)$

$a_4$ . *Associatividade*:  $t(a, t(b, c)) = t(t(a, b), c)$

Algumas t-normas freqüentemente utilizadas, são:

*Intersecção padrão*:  $t(a, b) = \min(a, b)$

*Produto algébrico*:  $t(a, b) = ab$

*Diferença limitada*:  $t(a, b) = \max(0, a + b - 1)$

*Intersecção drástica*:  $t_{\min}(a, b) = \begin{cases} a, & \text{quando } b = 1; \\ b, & \text{quando } a = 1; \\ 0, & \text{caso contrário} \end{cases}$

A intersecção fuzzy generalizada ( $A \cap B$ ) dos conjuntos  $A$  e  $B$  é definida pela Equação (2.4), na qual  $t$  denota uma t-norma.

$$(A \cap B)(x) = A(x)tB(x) \quad (2.4)$$

Uma s-norma  $s$ , também chamado de t-conorma, é uma operação binária da forma  $s : [0, 1]^2 \rightarrow [0, 1]$  que deve satisfazer os axiomas abaixo, para todo  $a, b, c \in [0, 1]$  (KLIR;YUAN, 1995):

$a_1$ . *Condição Limite*:  $s(a, 0) = a, s(1, 1) = 1$

$a_2$ . *Monotonicidade*:  $b \leq c$  implica  $s(a, b) \leq t(a, c)$

$a_3$ . *Comutatividade*:  $s(a, b) = s(b, a)$

$a_4$ . *Associatividade*:  $s(a, s(b, c)) = s(s(a, b), c)$

Algumas s-normas freqüentemente utilizadas, são:

*União padrão*:  $s(a, b) = \max(a, b)$

*Soma algébrica*:  $a + b - ab$

*Soma limitada*:  $s(a, b) = \min(1, a + b)$

*União drástica*:  $s_{\max}(a, b) = \begin{cases} a, & \text{quando } b = 0; \\ b, & \text{quando } a = 0; \\ 1, & \text{caso contrário} \end{cases}$

A união fuzzy generalizada  $(A \cup B)$  dos conjuntos  $A$  e  $B$  é definida pela Equação (2.5), na qual  $s$  denota uma  $s$ -norma.

$$(A \cup B)(x) = A(x)sB(x) \quad (2.5)$$

Denota-se  $cA$  o complemento generalizado fuzzy de  $A$  do tipo  $c$ , no qual  $c$  é definida pela função  $c : [0, 1] \rightarrow [0, 1]$ . Logo,  $cA(x)$  pode ser interpretado não somente como o grau com que  $x$  pertence a  $cA$ , mas também como o grau no qual  $x$  não pertence a  $A$  (KLIR;YUAN, 1995). Para tanto, a função  $c$  deve satisfazer os seguintes axiomas:

- $a_1$ . *Condições Limites:*  $c(0) = 1$  e  $c(1) = 0$
- $a_2$ . *Monotonicidade:* para todo  $a, b \in [0, 1]$ , se  $a \leq b$ , então  $c(a) \geq c(b)$
- $a_3$ . *Continuidade:*  $c$  é uma função contínua
- $a_4$ . *Involução:*  $c$  é involutiva, ou seja,  $c(c(a)) = a$ , para cada  $a \in [0, 1]$

### 2.1.3 Relações Fuzzy

Relações fuzzy generalizam o conceito de relações da teoria de conjuntos convencionais e representam o grau da associação entre elementos de dois ou mais conjuntos fuzzy. Estas relações são conjuntos fuzzy definidos em conjuntos universos que são produtos cartesianos de conjuntos convencionais (NICOLETTI;CAMARGO, 2004).

A relação fuzzy binária  $R(X, Y)$ , na qual  $X$  e  $Y$  são conjuntos convencionais, é um conjunto fuzzy caracterizado pela função de pertinência  $R(x, y)$  na qual  $x \in X$  e  $y \in Y$ , e  $R(x, y)$  é um valor no intervalo  $[0, 1]$  que define o grau com que  $x$  e  $y$  estão relacionados pela relação  $R$ , ou seja:

$$R : X \times Y \rightarrow [0, 1]$$

Quando o conjunto universo é discreto, uma relação fuzzy binária pode ser representada por uma lista de pares ordenados precedidos dos graus de pertinência relativos a cada par, ou por matrizes.

Por exemplo, a representação da relação fuzzy binária  $R$  entre os conjuntos  $A = \{a_1, a_2\}$  e  $B = \{b_1, b_2, b_3\}$  é descrita pela notação de lista a seguir, na qual os números antes da barra representam o grau de pertinência de cada par (KLIR;YUAN, 1995) em  $R$ :

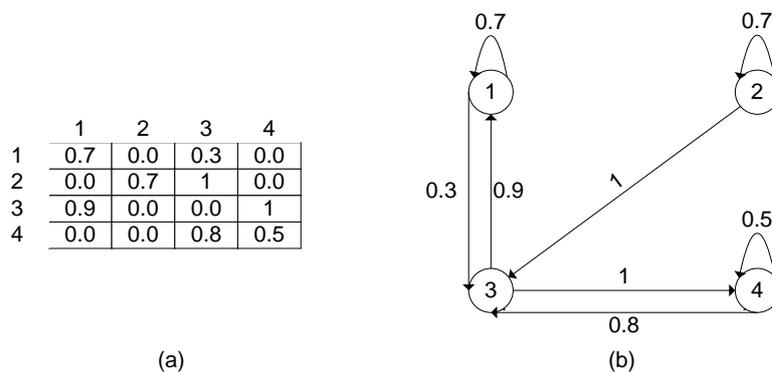
$$R(A, B) = 1/(a_1, b_1) + 0/(a_1, b_2) + 0.6/(a_1, b_3) + 0.9/(a_2, b_1) + 0.7/(a_2, b_2) + 0.3/(a_2, b_3)$$

A representação por matriz da relação entre os conjuntos  $X = \{x_1, x_2, x_3, \dots, x_n\}$  e  $Y = \{y_1, y_2, y_3, \dots, y_m\}$  é expressa da seguinte maneira (NICOLETTI;CAMARGO, 2004):

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{bmatrix}$$

na qual  $r_{ij} = R(x_i, y_j)$ , para cada  $i = 1, 2, \dots, n$  e  $j = 1, 2, \dots, m$ , representa o grau de pertinência do par  $(x_i, x_j)$  na relação  $R$ .

Relações fuzzy binárias sobre um mesmo conjunto  $X$  são relações da forma  $R(X, X)$  ou  $R(X^2)$ . Relações deste tipo podem ser representadas, além das duas formas descritas anteriormente, por dígrafos, cujos nós são os elementos de  $X$ . Os arcos representam os pares da relação  $R$  e os rótulos dos arcos representam o grau de pertinência de cada par em  $R$ . Podemos verificar, por exemplo, a Figura 2.3, a qual apresenta a relação  $R$  entre os elementos do conjunto  $X = \{1, 2, 3, 4\}$  representada por matriz e dígrafo (KLIR;YUAN, 1995).



**Figura 2.3:** Formas de representação de uma relação fuzzy binária: (a) Matriz e (b) Dígrafo

As operações de união, intersecção e complemento definidas para os conjuntos fuzzy

são também definidas para as relações fuzzy. Considerando-se duas relações fuzzy  $R$  e  $S$  definidas em um mesmo espaço  $X \times Y$ , as funções de pertinência resultantes da união e intersecção de relações fuzzy binárias são (ZIMMERMANN, 1991):

$$R \cup S(x, y) = \max\{R(x, y), S(x, y)\}, (x, y) \in X \times Y \quad (2.6)$$

$$R \cap S(x, y) = \min\{R(x, y), S(x, y)\}, (x, y) \in X \times Y \quad (2.7)$$

Relações fuzzy podem ser combinadas por meio da operação de Composição. Sejam  $R$  e  $S$  relações fuzzy definidas respectivamente nos produtos cartesianos  $X \times Y$  e  $Y \times Z$ . A relação  $T_1$ , denominada max-t, é definida pela Equação (2.8), na qual  $t$  é uma t-norma. A relação  $T_2$ , denominada min-s, é definida pela Equação (2.9), na qual  $s$  é uma s-norma. Sendo assim,  $T_1$  e  $T_2$  representam a composição de relações fuzzy  $R$  e  $S$ , cujo par  $(x, z) \in X \times Z$  (NICOLETTI;CAMARGO, 2004).

$$T_1(x, z) = (R \circ S)(x, z) = \max_{y \in Y} t[R(x, y), S(y, z)] \quad (2.8)$$

$$T_2(x, z) = (R \circ S)(x, z) = \min_{y \in Y} s[R(x, y), S(y, z)] \quad (2.9)$$

As definições anteriores desempenham um papel relevante na definição da semântica das regras fuzzy. Na seção seguinte encontra-se uma breve discussão sobre Computação com regras fuzzy, descrevendo os conceitos de Variável Linguística, Regras Fuzzy e Inferência Fuzzy, que são fundamentais para o entendimento deste trabalho.

## 2.2 Computação com Regras Fuzzy

As regras são utilizadas como um mecanismo formal para representar o conhecimento de um domínio. As regras fuzzy, mais especificamente, permitem a representação de conhecimento impreciso, o que muitas vezes torna a modelagem do problema mais fiel e adequada ao mundo real. Todo processo computacional envolvendo regras consiste em aplicar os chamados mecanismos de raciocínio que usam as regras fuzzy juntamente com os fatos, para fazer inferências.

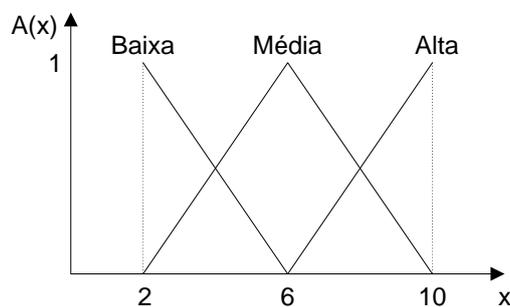
### 2.2.1 Variáveis Lingüísticas

Com o conceito de variáveis lingüísticas, problemas naturalmente imprecisos e complexos passam a ser manipuláveis por computadores. A interpretação de um determinado conhecimento expresso na forma lingüística torna-se passível de uma representação matemática por meio destas variáveis.

Variáveis lingüísticas são variáveis cujos valores são palavras ou sentenças em linguagem natural em vez de números (ZIMMERMANN, 1991). Elas são definidas sobre um determinado domínio, o qual é granularizado em termos lingüísticos definidos por conjuntos fuzzy. O processo de granularização de um domínio de uma variável em conjuntos fuzzy define a chamada partição fuzzy.

Algumas das motivações para o uso de variáveis lingüísticas é que as mesmas podem ser usadas como forma de compactar a informação e como meio de caracterização aproximada de fenômenos que são poucos definidos ou muito complexos para permitir um definição exata. Além disto, elas possuem meios para traduzir uma descrição lingüística em uma descrição numérica e computável.

Na Figura 2.4, pode-se observar uma partição fuzzy do domínio da variável *velocidade*  $[2, 10]$ , granularizado nos termos lingüísticos *alta*, *média* e *baixa*. Além desses termos básicos, modificadores como muito, mais ou menos, muitíssimo e outros, poderiam ser usados para gerar novos valores lingüísticos compostos para a variável velocidade como *muitíssimo alta*, *muito baixa*, *mais ou menos alta*, etc.



**Figura 2.4:** Variável Lingüística: Velocidade

### 2.2.2 Regras Fuzzy

O uso de regras como mecanismo formal para representação do conhecimento remonta às primeiras décadas da criação da área de Inteligência Artificial, quando as regras de produção foram propostas como um par condição-ação que define uma porção de conhecimento para solução de um problema (LUGER, 2004).

As regras são uma das técnicas mais antigas para representação de conhecimento e permanecem sendo largamente utilizadas. Sua popularidade vem do fato de possibilitarem a expressão clara de diretivas e estratégias, captarem o conhecimento experimental de especialistas humanos e possuírem um formato lingüístico de fácil compreensão.

As regras fuzzy, de forma geral, têm um formato semelhante ao das regras convencionais:

*SE antecedente ENTÃO conseqüente*

Este formato tem a finalidade de estabelecer relações entre as variáveis que aparecem no antecedente, também chamado de condição ou premissa, e as que aparecem no conseqüente, também chamado de conclusão ou ação. Por exemplo, a regra

*SE a velocidade é alta ENTÃO o risco\_de\_acidente é grande*

estabelece uma relação entre as variáveis *velocidade* e *risco\_de\_acidente* usando os termos *alta* e *grande*, que são inerentemente imprecisos.

Uma proposição fuzzy é a parcela de informação básica que pode aparecer em uma regra e seu formato mais simples é

$X \text{ é } A$

na qual  $X$  é uma variável lingüística e  $A$  é um termo lingüístico que representa um conjunto fuzzy, o qual é caracterizado por uma função de pertinência. Por exemplo, se  $X$  for a variável velocidade e  $A$  o termo alta, a proposição  $X \text{ é } A$  representa formalmente a parcela de informação do exemplo anterior que diz “*a velocidade é alta*”.

As proposições fuzzy também podem ser compostas, isto é, construídas a partir de proposições simples pelos operadores de conjunção ou disjunção, resultando em proposições com os formatos:

$$X_1 \text{ é } A \text{ E } X_2 \text{ é } B$$

ou

$$X_1 \text{ é } C \text{ OU } X_2 \text{ é } D$$

As proposições fuzzy compostas podem representar conhecimento tal como: “A temperatura é alta E a umidade é baixa” e “O tempo está nublado OU o trabalho é difícil”.

Uma regra fuzzy assume o formato de uma proposição condicional, que no caso mais simples fica:

$$\text{SE } X \text{ é } A \text{ ENTÃO } Y \text{ é } B,$$

na qual  $X$  e  $Y$  são variáveis lingüísticas sobre os domínios  $X$  e  $Y$ , respectivamente, e  $A$  e  $B$  são termos lingüísticos, que rotulam os conjuntos fuzzy definidos sobre os domínios  $X$  e  $Y$ , respectivamente.

Em uma regra fuzzy, tanto o antecedente quanto o conseqüente podem ser formados por proposições compostas, o que dá origem, por exemplo, a regras da forma:

$$\text{SE } X_1 \text{ é } A_1 \text{ E } X_2 \text{ é } A_2 \text{ E } X_3 \text{ é } A_3 \text{ ENTÃO } Y_1 \text{ é } B_1 \text{ OU } Y_2 \text{ é } B_2,$$

na qual  $A_1, A_2, A_3$  e  $B_1, B_2$  são conjuntos fuzzy nos universos  $X_1, X_2, X_3, Y_1, Y_2$ , respectivamente, e  $X_1, X_2, X_3, Y_1, Y_2$  são variáveis sobre os domínios  $X$  e  $Y$ .

### 2.2.3 Inferência Fuzzy

A computação com regras fuzzy envolve a aplicação de um método de raciocínio, o qual está baseado em regras de inferência. A regra de inferência composicional é a regra de inferência básica para os processos de raciocínio nos sistemas fuzzy. A regra de inferência composicional, também chamada de regra da composição, será apresentada para o caso simples, isto é, quando a regra tem apenas uma variável no antecedente e uma no conseqüente.

O processo de inferência é aplicado a partir de uma regra, a qual estabelece um conhecimento sobre algum domínio, e de um fato, o qual estipula uma informação específica sobre a situação da qual deseja-se obter alguma conclusão.

Seja  $A$  um conjunto fuzzy sobre  $X$ , e  $B$  um conjunto fuzzy sobre  $Y$ . A regra *SE  $X$  é  $A$  então  $Y$  é  $B$*  induz uma relação fuzzy  $R(x, y)$  sobre  $X$  e  $Y$ . A definição da relação  $R$  consiste na definição da semântica da regra. Essa relação é definida pela função  $f$  da forma:

$$R(x, y) = f((A(x), B(y))),$$

sendo que  $f$  pode ser de uma das três classes de funções: conjunção fuzzy, disjunção fuzzy e implicação fuzzy (detalhes em (PEDRYCZ;GOMIDE, 1998)).

Para usar essa regra em uma inferência, é necessário conhecer algum fato que estabeleça um valor específico para a variável  $X$  do antecedente. Dado o fato  $X$  é  $A'$ , pode-se inferir a conclusão  $Y$  é  $B'$ , sendo que  $B'$  é definida pela Equação (2.10) (PEDRYCZ;GOMIDE, 1998), para todo  $y \in Y$ , na qual  $\sup$  denota o operador de supremo e  $t$  uma t-norma.

$$B'(y) = \sup_{x \in X} [A'(x)tR(x, y)] \quad (2.10)$$

A aplicação da regra de inferência composicional pode ser facilmente estendida para regras com múltiplas variáveis, e também para situações em que o conhecimento do problema é descrito por um conjunto de regras. O processo geral de computação usando regras fuzzy é baseado na regra de inferência composicional.

Alternativamente a este, pode ser aplicado o chamado método de inferência escalonada, que envolve um número pequeno de cálculos numéricos com relação ao método geral. Esses dois processos gerais de computação com regras estão descritos em (PEDRYCZ;GOMIDE, 1998).

A seguir são abordados conceitos sobre Sistemas Fuzzy (SF), os quais fazem uso dos conceitos descritos anteriormente.

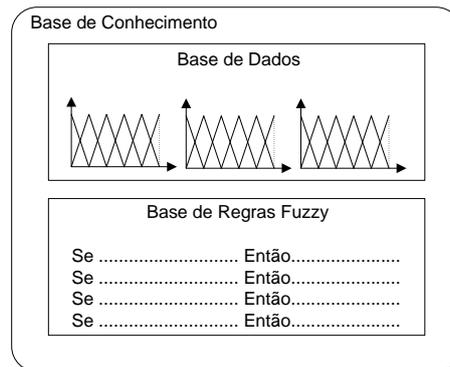
## 2.3 Sistemas Fuzzy

Como comentado anteriormente, os SF utilizam a lógica fuzzy em seu processo de raciocínio e as variáveis do problema são representadas por valores lingüísticos definidos por conjuntos fuzzy (KLIR;YUAN, 1995). O tipo específico de SF de interesse para este trabalho são os Sistemas Fuzzy Baseados em Regras (SFBR), descritos a seguir.

### 2.3.1 Sistemas Fuzzy Baseados em Regras

Os SFBR são compostos por dois componentes principais: a Base de Conhecimento (BC) e o Mecanismo de Inferência (MI).

A Base de Conhecimento, Figura 2.5, é composta pela Base de Dados (BD), a qual contém as definições dos conjuntos fuzzy relacionados aos termos lingüísticos usados nas regras fuzzy e pela Base de Regras (BR), que armazena o conjunto de regras que modelam um determinado problema.

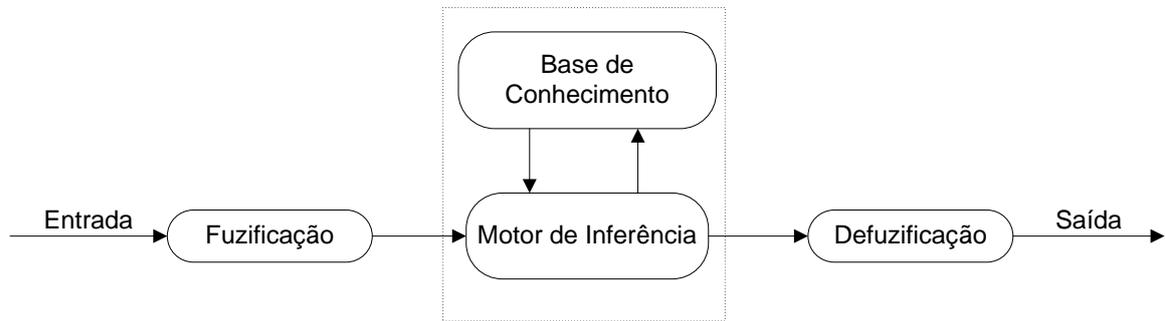


**Figura 2.5:** Base de Conhecimento de um SFBR

O Motor de Inferência é responsável pelo processamento das regras, o qual é realizado por algum método de raciocínio. Este consiste da aplicação de um procedimento de inferência para derivar conclusões a partir das regras e de fatos conhecidos.

O modo de operação de um SFBR, Figura 2.6, geralmente, consiste das seguintes etapas:

1. Transformação dos valores de entrada em conjuntos fuzzy, ou fuzificação. Para entradas numéricas este passo consiste no cálculo do grau de pertinência de cada valor de entrada no conjunto correspondente, e indica a compatibilidade da entrada com o antecedente de cada regra.
2. Agregação de antecedentes de cada regra por meio de operadores para o cálculo de conjunção fuzzy. O resultado obtido é chamado de Grau de Disparo da regra;
3. Aplicação da inferência com a derivação de resultados individuais para cada regra;
4. Combinação de possíveis saídas fuzzy;
5. Transformação do resultado fuzzy em um resultado preciso, processo conhecido como defuzificação. Esta etapa nem sempre ocorre, pois depende do tipo de problema em que será aplicado o SF.



**Figura 2.6:** Sistema Fuzzy

A interpretação e utilização dos conjuntos fuzzy resultantes dos processos de inferência podem ser feitas de forma distintas, dependendo do tipo de sistema e de aplicação. Assim, pode-se traduzir a saída para um valor linguístico, converter para um valor numérico (defuzificar) ou usar os graus de disparo das regras isoladamente ou de forma combinada para graduar a saída.

### 2.3.2 Tipos de Sistemas Fuzzy

Os SFBR podem ser de tipos diferentes, de acordo com o formato das regras. Entre os mais conhecidos estão os do tipo Mamdani (MAMDANI, 1977) e Takagi-Sugeno (TAKAGI;SUGENO, 1985). Nesses dois tipos de sistema, o antecedente das regras é formado por conjunções de proposições simples. O formato do conseqüente é diferente nos dois sistemas.

Os sistemas do tipo Mamdani possuem proposições fuzzy no antecedente e no conseqüente da regra, como mostrado a seguir, no qual  $X_1, \dots, X_n$  representam as variáveis lingüísticas de entrada,  $Y$  é uma variável lingüística de saída e  $A_{i_1}, \dots, A_{i_n}$  e  $B$  representam os termos lingüísticos.

$$SE X_1 \text{ é } A_{i_1} \text{ E } \dots \text{ E } X_n \text{ é } A_{i_n} \text{ ENTÃO } Y \text{ é } B.$$

As regras disparadas geram resultados parciais que são agregados para formar uma única saída final na forma de conjunto fuzzy. Esse conjunto é então defuzificado para a obtenção de uma saída numérica precisa.

Os sistemas do tipo Takagi-Sugeno possuem proposições fuzzy no antecedente da regra e o conseqüente da regra é uma função  $f$  aplicada sobre os valores de entrada  $x_1, \dots, x_n$ , como mostrado a seguir, no qual  $X_1, \dots, X_n$  representam as variáveis lingüísticas e  $A_{i_1}, \dots, A_{i_n}$  representam os termos lingüísticos.

$$SE X_1 \text{ é } A_{i_1} \text{ E } \dots \text{ E } X_n \text{ é } A_{i_n} \text{ ENTÃO } f(x_1, \dots, x_n).$$

Os sistemas do tipo Takagi-Sugeno já possuem uma saída numérica precisa como resultado da aplicação de uma função sobre os valores de entrada, não sendo necessária, portanto, a defuzificação.

### 2.3.3 Defuzificação

Vários métodos de defuzificação são utilizados a fim de converter o resultado lingüístico da inferência em um valor real que melhor o represente. Entre eles estão o Centro de Área, Centro de Máximos, Média de Máximos, Centro de Máximos Ponderado, Valor mais plausível e outros. Todos estes métodos vêm sendo largamente utilizados e estudados, a fim de conhecer suas particularidades e descobrir qual deles se adequa mais a determinados tipos de sistemas, como pode ser visto em (KENDEL;FRIEDMAN, 1998), (L. RONDEAU;LAMOTTE, 1997), (ROVENTA;SPIRCU, 2003), (RUNKLER, 1997).

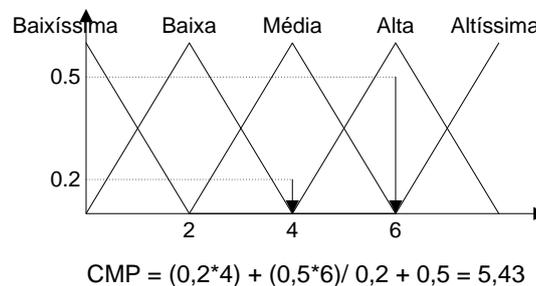
Para este trabalho, no entanto, apenas dois métodos foram estudados: Centro de Máximos Ponderado (ALTROCK, 1995) e Média de Elementos Máximos (adaptado de (ALTROCK, 1995)). Esses métodos, na forma como são apresentados aqui, pressupõem o uso de conjuntos fuzzy triangulares no domínio de saída.

O Centro de Máximos Ponderado (CMP) converte um conjunto fuzzy de saída em um único valor real calculando o melhor compromisso pela média ponderada dos valores dos elementos típicos dos termos envolvidos no conjunto de saída, sendo que os pesos são os graus de disparo das regras que geraram cada termo. Entende-se por elemento típico de um conjunto fuzzy o elemento do conjunto base cujo grau de pertinência no conjunto fuzzy é máximo.

A Equação (2.11) mostra o cálculo do método CMP, na qual  $K$  é o conjunto de regras disparadas (grau de disparo  $> 0$ ),  $DoF(R_k)$  é o grau de disparo da regra  $R_k$  e  $Vtipico(R_k)$  é o valor típico do conjunto de saída da Regra  $R_k$ .

$$CMP = \frac{\sum_{R_k \in K} V_{tipico}(R_k) * DoF(R_k)}{\sum_{R_k \in K} DoF(R_k)} \quad (2.11)$$

Por exemplo, a Figura 2.7 representa os conjuntos fuzzy definidos para a variável de saída Y de um sistema fuzzy que usa o CMP como método de defuzificação no qual V1 e V2 são as variáveis lingüísticas de entrada. Os valores típicos de cada termo são 0 para o termo baixíssima, 2 para o termo baixa, 4 para o termo média, 6 para o termo alta e 8 para o termo altíssima.



**Figura 2.7:** Centro de Máximos Ponderado

Supondo que, para determinados valores de entrada as regras:

$R_1$ : SE V1 é Baixa e V2 é Alta ENTÃO Y é Média

$R_2$ : SE V1 é Baixa e V2 é Baixa ENTÃO Y é Alta

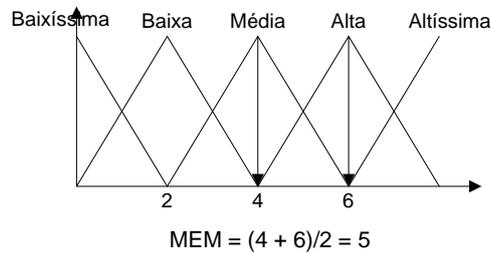
disparam com graus 0,2 e 0,5, respectivamente, o valor resultante da defuzificação é 5,43.

A Média de Elementos Máximos (MEM) é semelhante ao método anterior. Porém, ele calcula o melhor compromisso pela média dos valores típicos máximos dos termos envolvidos no conjunto de saída, convertendo um conjunto fuzzy de saída em um único valor real conforme Equação (2.12), na qual  $\sum k$  representa a quantidade de regras disparadas e  $V_{tipico}(R_k)$  é o valor típico do conjunto de saída da Regra  $R_k$ .

$$MEM = \frac{\sum_{R_k \in K} V_{tipico}(R_k)}{\sum k} \quad (2.12)$$

Seguindo as mesmas condições do exemplo anterior, a Figura 2.8 mostra o resultado

obtido pela defuzificação pelo método MEM.



**Figura 2.8:** Média de Elementos Máximos

Um tipo particular de SFBR e de interesse para este trabalho são os Sistemas de Classificação Fuzzy, os quais serão descritos a seguir.

### 2.3.4 Sistemas de Classificação Fuzzy

Classificação é uma importante tarefa encontrada nas áreas de reconhecimento de padrões, tomada de decisão, mineração de dados e modelagem. Dado um objeto, também chamado de padrão, o objetivo da classificação é atribuir uma classe, de um conjunto conhecido de classes, a esse objeto, o qual é descrito por  $n$  atributos.

Muitos métodos têm sido utilizados para a tarefa de classificação de padrões, tais como árvores de decisão (QUINLAN, 1986), métodos estatísticos (DUDA;HART, 1973), redes neurais (BISHOP, 1995) e sistemas fuzzy (KECMAN, 2001).

Sistemas de Classificação Fuzzy (SCF) são Sistemas Fuzzy Baseados em Regras projetados com a finalidade específica de realizar a tarefa de classificação. Assim, os SCF utilizam métodos de raciocínio próprios para essa tarefa.

O formato da regra fuzzy de classificação segue o formato geral das regras fuzzy descrito na Seção 2.2.2 e pode ser expressa por:

$$\text{SE } X_1 \text{ é } A_{i1} \text{ E } \dots \text{ E } X_n \text{ é } A_{in} \text{ ENTÃO } \textit{Classe} = C_j$$

na qual  $X_1, \dots, X_n$  são os atributos do padrão considerado no problema (representados aqui por variáveis lingüísticas),  $A_{i1}, \dots, A_{in}$  são os valores lingüísticos usados para representar os valores de tais atributos e  $C_j$  é a classe, fuzzy ou não, a que pertence o padrão.

Os métodos de raciocínio fuzzy para classificação mais utilizados são: Método de Raciocínio Fuzzy Clássico (CHI et al., 1996) (GONZÁLEZ;PÉREZ, 1999) (CORDÓN et al.,

1997) e Método de Raciocínio Fuzzy Geral (ISHIBUCHI et al., 1999b) (ISHIBUCHI et al., 1999a).

O método de Raciocínio Fuzzy Clássico classifica um padrão de entrada de acordo com a regra que possuir maior grau de compatibilidade com o mesmo conforme o procedimento a seguir.

Seja  $e_p = (a_{p1}, a_{p2}, \dots, a_{pn})$  um padrão a ser classificado e  $R_1, R_2, \dots, R_S$  um conjunto de  $S$  regras do sistema de classificação, cada uma com  $n$  antecedentes. Seja  $A_i(a_{pi})$ ,  $i = 1, \dots, n$ , o grau de pertinência do atributo  $a_{pi}$  no  $i$ -ésimo conjunto fuzzy da regra  $R_k$ . O Método de Raciocínio Fuzzy Clássico segue as seguintes etapas para a classificação do padrão  $e_p$ :

1. Calcular o grau de compatibilidade entre o padrão  $e_p$  e cada regra  $R_k$ , para  $k = 1, \dots, S$ , por:

$$Compat(R_k, e_p) = t(A_1(a_{p1}), A_2(a_{p2}), \dots, A_n(a_{pn}))$$

no qual  $t$  denota uma  $t$ -norma.

2. Encontrar a regra  $R_{kmax}$  que possui o maior grau de compatibilidade com o padrão, isto é,

$$Compat(R_{kmax}, e_p) = \max \{Compat(R_k, e_p)\}, k = 1, 2, \dots, S$$

3. Atribuir a classe  $C_j$  ao padrão  $e_p$ , na qual  $C_j$  é a classe do conseqüente da regra  $R_{kmax}$  que possui o maior grau de compatibilidade com o padrão, encontrada no passo anterior.

O método de Raciocínio Fuzzy Geral, o qual combina as informações providas por todas as regras de uma mesma classe para classificar o padrão, é descrito a seguir para a classificação do padrão  $e_p$ .

1. Calcular o grau de compatibilidade entre o padrão  $e_p$  e cada regra  $R_k$ , para  $k = 1, \dots, S$ , por:

$$Compat(R_k, e_p) = t(A_1(a_{p1}), A_2(a_{p2}), \dots, A_n(a_{pn}))$$

na qual  $t$  denota uma  $t$ -norma.

2. Calcular para cada classe  $C$  o valor de  $Classe_C$ , ou seja, o grau de classificação do padrão na classe, agregando os graus de compatibilidade do passo anterior de todas as regras cuja classe predita é  $C$ :

$$Classe_C = f \{ Compat(R_k, e_p) \mid C \text{ é a classe de } R_k \}$$

sendo  $f$  um operador de agregação tal que  $min \leq f \leq max$ .

3. A classe  $C_j$  será atribuída ao padrão  $e_p$ , sendo  $C_j$  a classe correspondente ao valor máximo de  $Classe_C$  encontrado no passo anterior.

Um método alternativo de raciocínio fuzzy é descrito em (CORDÓN et al., 1997), cujas duas primeiras etapas seguem o padrão dos métodos anteriores, sendo diferenciado pela maneira como é definida a função  $f$  para cálculo dos graus de classificação do passo 2.

## 2.4 Considerações Finais

Neste capítulo foram abordados os conceitos básicos da teoria de conjuntos fuzzy, raciocínio fuzzy e sistemas fuzzy. A teoria que fundamenta os Sistemas Fuzzy, em particular os Sistemas de Classificação Baseado em Regras Fuzzy, é muito ampla. Sendo assim, este capítulo abordou apenas os principais conceitos necessários à compreensão deste trabalho de pesquisa. O capítulo seguinte destaca os principais fundamentos acerca dos Algoritmos de Agrupamento.

---

# Algoritmos de Agrupamento

---

Agrupamento é a separação de padrões em grupos, também conhecidos como *clusters* (JAIN et al., 1999). Padrões representam os objetos a serem agrupados, podendo ser objetos físicos, como carros ou imóveis, ou noções abstratas, como características sociais de uma população. Padrões são formados por vetores de características (atributos) em que cada característica assume um dado valor.

O agrupamento de dados tem como finalidade identificar dados com características semelhantes e agrupá-los de acordo com essas características. Existem vários métodos e algoritmos de agrupamento com características diferentes que se adequam a diversos tipos de problemas. No contexto deste trabalho, foi estudado o agrupamento fuzzy, no qual um dado pode pertencer a mais de um grupo com diferentes graus de pertinência, e uma extensão desse, o agrupamento fuzzy condicional.

A seguir são destacadas algumas medidas de proximidade utilizadas para identificar dados semelhantes no processo de agrupamento relevantes para este trabalho.

### 3.1 Medidas de Proximidade

A medida de proximidade entre os padrões é de vital importância para a constituição de cada grupo porque é por meio dela que é estabelecida a caracterização dos mesmos. A escolha desta medida deve ser adequada a cada tipo de problema, pois existem vários tipos e escalas de medidas (JAIN et al., 1999).

Para que uma dada medida seja considerada uma medida de proximidade, ela precisa satisfazer algumas propriedades. Dados os padrões  $\mathbf{x}_i$ ,  $\mathbf{x}_j$  e  $\mathbf{x}_k$ , uma medida  $S$  é uma medida de proximidade se satisfaz as três primeiras propriedades a seguir (GORDON, 1999):

1.  $S_{ij} \geq 0$ , para todo  $i$  e  $j$  (positividade);
2.  $S_{ii} = 0$ , para todo  $i$  e  $j$  ( $i = j$ );
3.  $S_{ij} = S_{ji}$ , para todo  $i$  e  $j$  (simetria);
4.  $S_{ij} = 0$ , somente se  $i = j$ ;
5.  $S_{ik} \leq S_{ij} + S_{jk}$ , para todo  $i$ ,  $j$  e  $k$  (desigualdade triangular);

Se a medida de proximidade satisfaz também as propriedades 4 e 5, ela é considerada uma métrica de proximidade.

Existem duas definições comuns para medidas de proximidade: Medida de Similaridade, medida pela igualdade entre dois objetos, e Medida de Dissimilaridade, medida pela diferença entre dois objetos.

Para cada tipo de atributo, diferentes medidas são empregadas. Os atributos podem ser do tipo:

- **Binário.** Quando os atributos possuem apenas dois valores. Por exemplo, sexo (feminino ou masculino), doente (sim ou não), etc.
- **Nominal.** Quando os atributos possuem mais de dois tipos de valores. Por exemplo, situação do livro (disponível, emprestado, reservado), cor da pele (branca, negra, morena), etc.
- **Discreto.** Quando os atributos assumem valores fixos. Por exemplo, idade, quantidade de livros, quantidade de salas, etc.

- **Contínuo.** Quando os atributos assumem valores reais. Por exemplo, temperatura, peso, preço, etc.

Neste trabalho em particular, foram utilizados domínios apenas com atributos contínuos. Portanto, neste capítulo são apresentadas apenas medidas para este tipo de atributo. Para medir a dissimilaridade entre atributos contínuos, as distâncias Euclidiana e Manhattan, baseadas na métrica de Minkowski, são bastante utilizadas (KAUFMAN;ROUSSEEUW, 1990). A métrica de Minkowski é uma generalização da distância normal entre pontos em um espaço euclidiano<sup>1</sup>. Esta métrica é definida pela Equação (3.1), na qual  $r$  é um parâmetro que diferencia o tipo de distância utilizado,  $d$  é a dimensão do padrão,  $x_{ik}$  e  $x_{jk}$  são, respectivamente, o  $k$ -ésimo componente do  $i$ -ésimo e  $j$ -ésimo padrão,  $\mathbf{x}_i$  e  $\mathbf{x}_j$ .

$$P_{ij} = \left( \sum_{k=1}^d |x_{ik} - x_{jk}|^r \right)^{1/r} \quad (3.1)$$

**Distância Euclidiana.** Este tipo de medida trabalha melhor na busca de um conjunto de dados que possui grupos mais compactos ou isolados, sendo utilizada para padrões cujos atributos são contínuos. Este tipo de distância é definida pela Equação (3.2), a qual assume o parâmetro  $r = 2$  da métrica de Minkowski.

$$d_2(\mathbf{x}_i, \mathbf{x}_j) = \left( \sum_{k=1}^d (x_{ik} - x_{jk})^2 \right)^{1/2} \quad (3.2)$$

**Distância de Manhattan.** Este tipo de distância é também conhecida como distância bloco-cidade, a qual é definida pela Equação (3.3), assumindo o parâmetro  $r = 1$  da métrica de Minkowski.

$$d_1(\mathbf{x}_i, \mathbf{x}_j) = \sum_{k=1}^d |x_{ik} - x_{jk}| \quad (3.3)$$

O inconveniente em se usar a métrica de Minkowski é que atributos com uma escala maior possam dominar outros. Duas soluções para este problema são: normalizar os atributos contínuos para um intervalo comum, ou a atribuição de pesos entre os atributos de acordo com sua relevância no conjunto (JAIN et al., 1999).

A correlação linear entre atributos também pode prejudicar a medida de distância, a

---

<sup>1</sup>Espaço euclidiano é definido como um conjunto de pontos que satisfazem determinadas relações, expressas em termos de distância e de ângulo.

qual pode ser diminuída utilizando-se distância de Mahalanobis, definida pela Equação (3.4), na qual os padrões  $\mathbf{x}_i$  e  $\mathbf{x}_j$  são assumidos como vetores, e  $\Sigma$  é a matriz de covariância<sup>2</sup> dos padrões. A distância  $d_M$  atribui pesos diferentes para atributos diferentes baseada nas suas variâncias e a correlação linear entre pares de padrões (JAIN et al., 1999). Sendo assim, a matriz de covariância reduz as influências relativas de cada variável no cálculo das distâncias entre os padrões.

$$d_M(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i - \mathbf{x}_j)\Sigma^{-1}(\mathbf{x}_i - \mathbf{x}_j)^T \quad (3.4)$$

Na Seção seguinte é descrito o processo de execução realizado pelos algoritmos de agrupamento e destacados três algoritmos de agrupamento: K-Means, Fuzzy C-Means e Fuzzy C-Means Condicional.

## 3.2 Algoritmos de Agrupamento

Para a execução dos Algoritmos de Agrupamento (AA) é necessário que sejam definidos um método para particionar os dados de acordo com algum critério de agrupamento e uma medida de similaridade, de tal forma que dados semelhantes sejam alocados ao mesmo grupo e dados distintos alocados a grupos diferentes.

Os principais tipos de AA são os algoritmos do tipo Partição, os quais criam partições e avaliam padrões pertencentes a cada grupo de acordo com algum critério, e os algoritmos do tipo Hierárquico, os quais criam uma decomposição hierárquica de um conjunto de dados.

Os AA são utilizados para explorar as relações entre os dados de acordo com as etapas descritas a seguir, conforme Figura 3.1 adaptada de (JAIN et al., 1999).

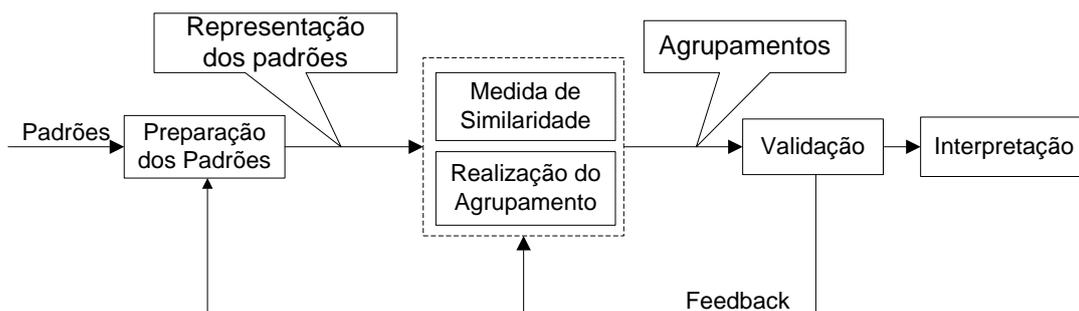
1. *Preparação dos padrões.* Nesta etapa é determinada a forma como os padrões serão representados no conjunto de dados durante o processo de agrupamento. Com o objetivo de obter uma boa representação de padrões, muitas vezes torna-se necessária a realização de normalizações, conversão de tipos e a redução do número de atributos

---

<sup>2</sup>Matriz de Covariância é uma matriz quadrada, simétrica, cuja diagonal contém a variância do padrão e em cada interseção linha  $i$  coluna  $j$  a covariância dos padrões  $\mathbf{x}_i$  e  $\mathbf{x}_j$ . A variância de um vetor mede o afastamento dos seus elementos em torno da sua média e a covariância é calculada de maneira análoga à variância mas envolve dois vetores de igual dimensão.

por meio de seleção ou extração de características. A preparação dos dados é especialmente importante para evitar *outliers*<sup>3</sup>, pois os mesmos podem influenciar nos grupos que são encontrados pelo algoritmo de agrupamento.

2. *Seleção da Medida de similaridade.* Esta etapa consiste na escolha de alguma das medidas descritas na Seção 3.1 ou outras medidas descritas na literatura.
3. *Execução do Algoritmo de Agrupamento.*
4. *Validação.* Nesta etapa é medida a “qualidade” dos grupos como, por exemplo, por meio da comparação dos resultados obtidos de diversos algoritmos de agrupamento ou determinando o valor mais apropriado de algum parâmetro do algoritmo (ex, número de grupos). Existem algumas medidas numéricas, juntamente com métodos estatísticos, que são aplicadas para avaliar os vários aspectos da validação de agrupamentos. Assim, se os grupos não satisfizer as medidas de validação, o processo retorna à fase de preparação de padrões (feedback) e um novo agrupamento é feito.
5. *Interpretação.* Nesta etapa ocorre a avaliação dos grupos, a qual permite verificar se a formação dos grupos condiz com os exemplos que levaram a sua formação.



**Figura 3.1:** Etapas do Processo de Agrupamento

Nas subseções seguintes são apresentados os algoritmos K-Means, Fuzzy C-Means e o Fuzzy C-Means Condicional, os quais são essenciais para o entendimento deste trabalho.

<sup>3</sup>Um *outlier* é uma observação que está situada fora do padrão de distribuição dos dados (MOORE;MCCABE, 1999).

### 3.2.1 Algoritmo K-Means

Também chamado de K-Médias, este é um algoritmo proposto por FORGY (1965) do tipo permuta, ou seja, a partição inicial é dada e as novas partições são obtidas pela permuta de um objeto de um grupo a outro, com a terminação do algoritmo quando a permutação satisfaz algum critério. Algoritmos deste tipo são relativamente rápidos na execução mas são prejudicados pela incerteza da partição inicial. Existe sempre a possibilidade de que diferentes partições iniciais possam levar a diferentes partições finais (HARTIGAN, 1975).

O K-Means é bastante conhecido devido à sua capacidade de sempre convergir para uma solução.

O número  $k$  de grupos é definido previamente e os centros dos grupos, centróides, são ajustados ao longo do processo de acordo com os dados mais próximos; a distância entre dados com atributos numéricos pode ser medida pela Distância Euclidiana. O algoritmo K-Means de uma forma geral pode ser descrito como descrito no Algoritmo 1.

---

**Algoritmo 1** K-Means

---

- 1: Definir (aleatoriamente) um número  $k$  de centróides.
  - 2: Para cada instância, determinar o centróide mais próximo e, a partir deste, escolher seu grupo.
  - 3: Redefinir os centróides pela média de cada grupo.
  - 4: Repetir os passos 2 e 3 até a convergência.
- 

Um passo básico para a execução do K-Means é a escolha adequada dos centróides iniciais. Esta escolha pode ser feita randomicamente, mas os resultados podem não ser satisfatórios. Portanto, é conveniente fazer uso de algumas técnicas que encontrem os centróides iniciais de maneira melhor do que randomicamente como, por exemplo, por meio da escolha de regiões densas, nas quais os padrões são bem separados, de forma que um centróide não seja escolhido por mais de um grupo.

O K-Means possui a vantagem de ser um algoritmo simples, obtendo bons resultados quando se conhece o número correto de grupos. No entanto, o K-Means necessita saber o número de grupos a serem formados, todas as instâncias são forçadas a pertencerem a um grupo, bem como é muito sensível a outliers.

Uma variação do K-Means é o K-Medóide, no qual a média de cada grupo na atualização dos centróides é substituída pela mediana. Esta variação permite a redução de outliers.

### 3.2.2 Algoritmo Fuzzy C-Means

Para o agrupamento fuzzy, uma das possibilidades é a utilização do Fuzzy C-Means (FCM) proposto inicialmente por BEZDEK (1981), o qual é o K-Means fuzificado, ou seja, um dado pertencerá a um certo grupo com um grau de pertinência. Este algoritmo é extremamente poderoso para o agrupamento de dados porque a idéia no algoritmo FCM é de ponderar a distância de um ponto ao centro do grupo por meio de um valor de pertinência e os limites entre os grupos são imprecisos.

Ao iniciar sua execução, o algoritmo FCM assume que uma determinada quantidade de grupos é definida previamente, bem como o grau de fuzificação  $m$  e um número pequeno  $\epsilon$  como critério de parada.

Uma característica deste algoritmo é que um dado, durante o processo iterativo, poderá pertencer a mais de um grupo, de acordo com o seu valor de pertinência. Além disso, todo dado deverá pertencer ao menos a um grupo e nenhum grupo poderá conter todos os dados. A atualização dos centros de grupos, centróides, é baseada na otimização de um critério de erro.

Considere um conjunto de dados  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ , no qual  $\mathbf{x}_k$  é um vetor de valores  $p$ -dimensional de atributos  $\mathbf{x}_k = [x_{k_1}, x_{k_2}, \dots, x_{k_p}] \in \mathbb{R}^p$ , para  $1 \leq k \leq n$ .

Uma pseudo-partição fuzzy, também conhecida como  $c$ -partição fuzzy, é uma família de conjuntos fuzzy de  $X$  denotados por  $P = \{A_1, A_2, \dots, A_c\}$ , que satisfaz as Equações (3.5) e (3.6), para todo  $k = 1 \dots n$  (KLIR; YUAN, 1995).

$$\sum_{i=1}^c A_i(\mathbf{x}_k) = 1 \quad (3.5)$$

$$0 < \sum_{k=1}^n A_i(\mathbf{x}_k) < n, \text{ para } i = 1, \dots, c \quad (3.6)$$

O FCM é um algoritmo iterativo que atualiza os centróides dos grupos definidos previamente com a definição de uma partição. Após cada atualização dos centróides, a partição é redefinida. Logo, a performance do FCM depende da forma de atualização destes centros e da redefinição das partições.

Os  $c$  vetores de centros,  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c$ , dos grupos são calculados pela Equação (3.7), para

todo  $i \in N_c$ , na qual  $m > 1$  é um número real que controla a influência dos graus de pertinência, ou seja,  $m$  é o grau de fuzificação (KLIR;YUAN, 1995) e  $A_i(\mathbf{x}_k)$  é o grau de pertinência do vetor  $\mathbf{x}_k$  ao grupo  $i$ .

$$v_i = \frac{\sum_{k=1}^n [A_i(\mathbf{x}_k)]^m \mathbf{x}_k}{\sum_{k=1}^n [A_i(\mathbf{x}_k)]^m} \quad (3.7)$$

Cada elemento da partição é redefinido pela Equação (3.8).

$$A_i(\mathbf{x}_k) = \frac{1}{\sum_{j=1}^c \left( \frac{\|\mathbf{x}_k - \mathbf{v}_i\|^2}{\|\mathbf{x}_k - \mathbf{v}_j\|^2} \right)^{\frac{1}{m-1}}} \quad (3.8)$$

A performance do FCM é baseada na otimização da função objetivo descrita pela Equação (3.9) sobre a pseudopartição  $P$ , na qual  $\|\mathbf{x}_k - \mathbf{v}_i\|$  representa a distância entre  $\mathbf{x}_k$  e  $\mathbf{v}_i$ , sendo que  $\mathbf{v}_i$  deve ser atualizada conforme Equação (3.7) (KLIR;YUAN, 1995).

$$J_m(P) = \sum_{k=1}^n \sum_{i=1}^c [A_i(\mathbf{x}_k)]^m \|\mathbf{x}_k - \mathbf{v}_i\|^2 \quad (3.9)$$

O objetivo do FCM é minimizar a função objetivo  $J_m$  descrita pela Equação (3.9), ou seja, minimizar a distância entre os padrões e os centros dos grupos. Assim, este problema de agrupamento é um problema de otimização. O Algoritmo 2 a seguir descreve as etapas do FCM.

### 3.2.3 Algoritmo Fuzzy C-Means Condicional

O algoritmo de agrupamento Fuzzy C-Means Condicional (FCMC) proposto como uma extensão do Fuzzy C-Means (FCM) em (PEDRYCZ, 1996a) é um método guiado por uma variável auxiliar, a qual permite que o agrupamento seja orientado por algum contexto. Ou seja, ao invés de simplesmente agrupar os dados  $X$ , serão agrupados os dados  $X$  no contexto  $A$ , contexto este que é determinado pela variável auxiliar.

No agrupamento FCMC, também chamado de agrupamento dependente de contexto ou sensível a contexto, cada vetor de dados  $\mathbf{x}_k$  tem uma variável auxiliar  $f_k$ . Assumindo que  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  são padrões  $p$ -dimensionais definidos no  $\mathbb{R}^p$ , e  $f_1, f_2, \dots, f_N$  são os valores das variáveis auxiliares associadas aos respectivos padrões, o aspecto condicionante do mecan-

**Algoritmo 2** Fuzzy C-Means

- 
- 1: Iteração  $t=0$ ;
  - 2: Selecionar uma partição inicial fuzzy  $P^{(0)}$ ;
  - 3: Calcular os centróides ( $\mathbf{v}_i$ ), usando a Equação (3.7);
  - 4: Atualizar  $P^{(t+1)}$ :
  - 5:     **Se**  $\|\mathbf{x}_k - \mathbf{v}_i^{(t)}\|^2 > 0$ , para todo  $i = 1 \dots c$      **Então**  $A_i^{(t+1)}(\mathbf{x}_k) = \left[ \sum_{j=1}^c \left( \frac{\|\mathbf{x}_k - \mathbf{v}_i^{(t)}\|^2}{\|\mathbf{x}_k - \mathbf{v}_j^{(t)}\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}$
  - 6:     **Fim     Se**
  - 7:     **Se**  $\|\mathbf{x}_k - \mathbf{v}_i^{(t)}\|^2 = 0$ , para todo  $i \in I \subseteq \{1 \dots c\}$      **Então** Definir  $A_i^{(t+1)}(\mathbf{x}_k)$  por um número real não-negativo que satisfaz:
 
$$\sum_{i \in I} A_i^{(t+1)}(\mathbf{x}_k) = 1$$
  - 8:     **Fim     Se**
  - 9: Comparar  $P^{(0)}$  e  $P^{(t+1)}$ :
  - 10:    **Se**  $\|P^{(t+1)} - P^{(t)}\| \leq \epsilon$ , para todo  $i = 1 \dots c$      **Então** Pare
  - 11:    **Caso contrário** Incremente  $t$  e retorne para o passo 2;
  - 12:    **Fim     Se**
- 

ismo de agrupamento é introduzido no agrupamento fazendo com que  $f_k$  seja distribuído aditivamente sobre as entradas da  $k$ -ésima coluna da partição  $u_{ik}$ , ou seja,

$$\sum_{i=1}^c u_{ik} = f_k, k = 1, 2, \dots, N, \quad (3.10)$$

na qual  $f_k$  denota o nível de envolvimento de  $\mathbf{x}_k$  na construção dos grupos e representa a condição adicionada ao processo de agrupamento e  $c$  é a quantidade de grupos.

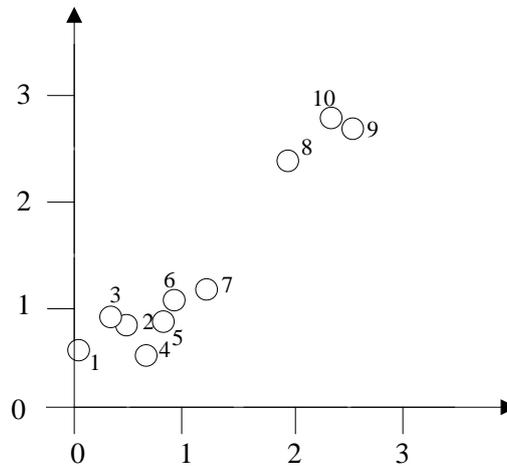
Assim, a iteração do FCMC otimizará o problema reformulando a Equação (3.8) para:

$$A_i(\mathbf{x}_k) = \frac{f_k}{\sum_{j=1}^c \left( \frac{\|\mathbf{x}_k - \mathbf{v}_i\|^2}{\|\mathbf{x}_k - \mathbf{v}_j\|^2} \right)^{\frac{1}{m-1}}} \quad (3.11)$$

Sendo assim, o comportamento do processo de agrupamento, bem como os grupos gerados serão influenciados pelo contexto no qual estarão inseridos. A seguir é apresentado um exemplo adaptado de (PEDRYCZ, 1996a) pelo qual é possível visualizar a diferença de comportamento entre os algoritmos FCM e FCMC.

A Figura 3.2 apresenta a distribuição dos padrões em um espaço bi-dimensional e a Tabela 3.1 apresenta os valores  $x_1$  e  $x_2$  dos padrões, bem como o valores  $f$  condicionantes do agrupamento.

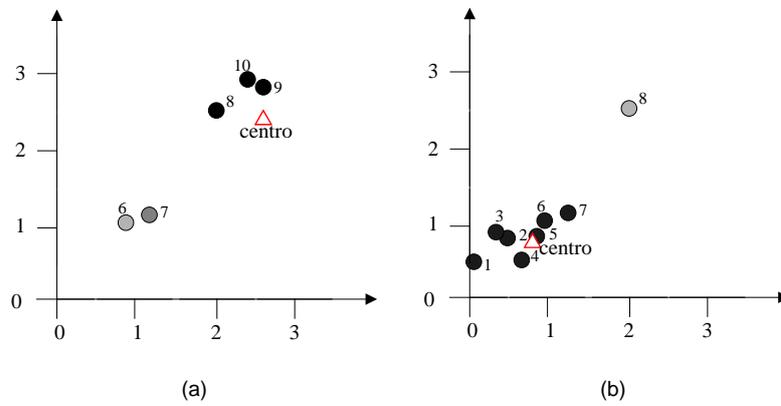
A seguir, a Figura 3.3 apresenta o resultado do agrupamento em 2 grupos utilizando o

**Figura 3.2:** Distribuição dos Padrões**Tabela 3.1:** Características dos Padrões

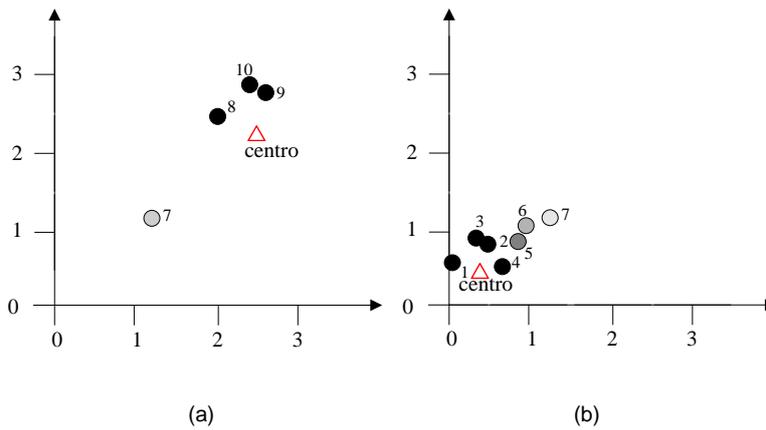
Padrões	$x_1$	$x_2$	$f$
1	0,00	0,50	0,96
2	0,40	0,67	0,84
3	0,30	0,70	0,70
4	0,67	0,45	0,82
5	0,87	0,95	0,21
6	0,97	1,05	0,12
7	1,20	1,10	0,18
8	2,00	2,50	0,96
9	2,40	2,67	0,84
10	2,30	2,70	0,70

algoritmo de agrupamento Fuzzy C-Means e a Figura 3.4 apresenta o resultado do agrupamento utilizando o algoritmo de agrupamento Fuzzy C-Means Condicional. Os pontos apresentam diferenças na tonalidade de cinza de acordo com o grau com que pertencem ao grupo, sendo os mais escuros aqueles com mais alto grau de pertinência. As Figuras (a) e (b) apresentam a distribuição dos padrões nos grupos A e B, respectivamente.

Observando as Figuras 3.3 e 3.4 é notável a diferença na organização dos padrões. Quando não considerado o valor  $f$ , Figura 3.3, os grupos são quase totalmente distinguíveis graficamente. Quando o valor  $f$  é considerado, os grupos se aproximam mais dos valores definidos pelo contexto, o que confirma a influência do contexto no agrupamento. É interessante observar que os padrões 5, 6 e 7 mesmo condicionados por um valor baixo ainda é determinante para a formação dos grupos.



**Figura 3.3:** Agrupamento não-Condicional: (a) Dados pertencentes ao grupo A; (b) Dados pertencentes ao grupo B.



**Figura 3.4:** Agrupamento Condicional: (a) Dados pertencentes ao grupo A; (b) Dados pertencentes ao grupo B.

A Tabela 3.2 apresenta os centros dos grupos gerados pelo algoritmo FCMC e FCM, respectivamente.

**Tabela 3.2:** Centros dos grupos

Fuzzy C-Means Condicional		
Centro(a)	2,33	2,57
Centro(b)	0,33	0,57
Fuzzy C-Means		
Centro(a)	2,63	2,77
Centro(b)	0,62	0,77

Sendo assim, o contexto atua como um filtro de dados concentrando o foco do processo de agrupamento em alguns subconjuntos específicos do conjunto de dados.

Quando o contexto é definido por valores lingüísticos, diz-se que o agrupamento é orientado lingüisticamente e o vetor  $\mathbf{x}_k$  é agrupado baseado em algum termo lingüístico definido em uma variável auxiliar  $y_k$ . Estes termos lingüísticos são tratados como conjuntos fuzzy (ou relações fuzzy quando  $y$  é um vetor). Se denotar a função de pertinência deste conjunto fuzzy como  $A(y)$ , então haverá o correspondente  $f_k = A(y_k) \in [0, 1]$  para cada vetor de dados  $\mathbf{x}_k$  (LESKI, 2003).

### 3.3 Considerações finais

Este capítulo abordou conceitos relacionados aos Algoritmos de Agrupamento, definindo conceitos como os significados de agrupamento e medidas de similaridade, dando maior ênfase nas medidas de distâncias que foram estudadas para o desenvolvimento deste trabalho.

Foram também apresentados os algoritmos K-Means, Fuzzy C-Means e Fuzzy C-Means Condicional, os quais foram usados para a implementação da abordagem utilizada neste trabalho.

Os Algoritmos de Agrupamento (AA) têm sido utilizados em conjunto com outras “metodologias inteligentes” na busca por sistemas mais eficientes, otimizando seus processos de execução. Com esta motivação e buscando relacionar este trabalho com outros trabalhos já existentes, no capítulo a seguir são apresentadas algumas maneiras pela qual os AA vem sendo utilizados juntamente com Sistemas Fuzzy e a abordagem proposta e implementada neste trabalho.

---

# Modelagem Fuzzy usando Agrupamento Condicional

---

O objetivo do método proposto neste trabalho é gerar regras de inferência fuzzy. Um conjunto de regras de inferência apropriado deve satisfazer dois requisitos que em muitas situações são conflitantes: ser facilmente interpretável e manter boas taxas de acuidade. A principal motivação para usar Sistemas Fuzzy Baseado em Regras (SFBR) é o fato de que eles são facilmente compreensíveis pelos seres humanos. Entretanto, esta habilidade pode ser prejudicada quando os sistemas fuzzy são automaticamente construídos a partir de um conjunto de dados, se não houver a preocupação em conservar tal compreensão durante o processo de geração das regras (CASTRO;CAMARGO, 2005).

Sendo assim, a modelagem de um sistema é a causa e efeito da abordagem para geração do modelo, isto é, geração de um esquema teórico que simplifica e facilita a compreensão de um sistema real e complexo. Favorecidos por este modelo, um sistema real pode ser explorado, controlado, simulado, predito e, inclusive, melhorado. O desenvolvimento de modelos compreensíveis é o objetivo principal na modelagem de um sistema. Uma vez

que sem compreensibilidade o modelo geralmente perde sua usabilidade (CASILLAS et al., 2003).

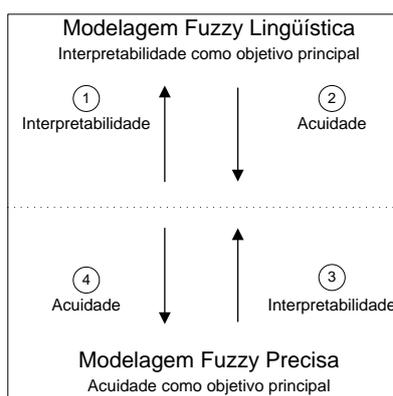
Neste capítulo é apresentada uma nova proposta de abordagem que utiliza o agrupamento condicional (PEDRYCZ, 1996a), apresentado no Capítulo 3, para a Modelagem Fuzzy (MF). Antes de ser apresentado o método proposto, é feita uma breve discussão sobre a questão do balanço interpretabilidade-acuidade dos sistemas fuzzy, bem como uma revisão da literatura sobre os trabalhos que usam agrupamento na MF.

## 4.1 Interpretabilidade x Acuidade

Dependendo dos requisitos para a construção de um sistema fuzzy, o campo da MF pode ser dividido em duas diferentes áreas (CASILLAS et al., 2003):

- Modelagem Fuzzy Lingüística (MFL), cujo principal objetivo é obter modelos fuzzy com boa interpretabilidade;
- Modelagem Fuzzy Precisa (MFP), cujo principal objetivo é obter modelos fuzzy com boa acuidade.

Esta divisão resulta em 4 diferentes possibilidades, conforme Figura 4.1: MFL com melhoria na interpretabilidade, MFL com melhoria na acuidade, MFP com melhoria na interpretabilidade, e MFP com melhoria na acuidade.



**Figura 4.1:** Possibilidades da Modelagem Fuzzy (CASILLAS et al., 2003)

Obter altos graus de interpretabilidade e acuidade pelo processo de MF é, muitas vezes, um propósito contraditório e, na prática, uma das propriedades prevalece sobre a outra.

No entanto, há uma tendência da comunidade científica em buscar pelo balanço entre interpretabilidade e acuidade (MONTES et al., 2006) (CASILLAS et al., 2003) (CASILLAS et al., 2000) (SETNES;ROUBOS, 1999) (SETNES et al., 1998).

Para obter este balanço, um esquema muito usado na literatura para a modelagem fuzzy segue as seguintes etapas:

1. Primeiramente é escolhido o objetivo principal (interpretabilidade ou acuidade) que definirá a estrutura de um modelo específico a ser usado, ou seja, é escolhida a abordagem da MF.
2. Em seguida, os componentes da modelagem (estrutura do modelo e/ou processo de modelagem) são melhorados por diferentes mecanismos para definir o balanço interpretabilidade-acuidade.

Uma possibilidade para obter um bom equilíbrio entre interpretabilidade e acuidade é desenvolver a MFL tentando obter modelos iniciais precisos, e em seguida aplicar um processo para melhorar a interpretabilidade do modelo obtido mesmo com a diminuição da acuidade. No entanto, sempre existe a chance de indiretamente melhorar a acuidade em termos de capacidade de generalização quando inconsistências e redundâncias existentes são removidas (CASILLAS et al., 2003).

Algumas abordagens existentes atualmente buscam o balanço interpretabilidade-acuidade dos sistemas fuzzy gerados automaticamente, utilizando algoritmos de agrupamento em sua modelagem. Segundo SETNES;ROUBOS (1999), a interpretabilidade, ou a transparência, e a acuidade dos sistemas fuzzy construídos usando agrupamento deve satisfazer os seguintes critérios:

1. **Moderado número de regras.** Quando o agrupamento é utilizado para geração de regras representadas por meio de grupos, o agrupamento deve assegurar que o tamanho da base de regras seja o suficientemente para tornar a base compreensível e, portanto, interpretável; e que as regras descrevam importantes regiões nos dados.
2. **Habilidade de distinção (*distinguishability*).** Um baixo número de grupos distintos induz funções de pertinência com pouca similaridade e com uma interpretação semântica clara. Isto pode ser conseguido por meio da otimização da base de regras inicial usando, inclusive, outros métodos que não sejam agrupamento.

3. **Cobertura.** A cobertura ponderada de grupos, que representam regras ou não, e suas posições no espaço de entrada e saída deve assegurar que o modelo é habilitado para derivar uma saída apropriada para todas as ocorrências de entrada.

Neste trabalho, especificamente, embora o objetivo principal seja a interpretabilidade, o balanço interpretabilidade-aciudade é garantido pela escolha prévia dos atributos considerados mais relevantes, os quais contribuem diretamente na precisão do sistema, e as regras são compactadas utilizando-se grupos condicionados a estes atributos relevantes.

Algumas abordagens que utilizam agrupamento em seu processo de modelagem fuzzy são apresentadas na seção a seguir.

## 4.2 Abordagens de Modelagem Fuzzy usando Agrupamento

Os algoritmos de agrupamento Fuzzy mais utilizados na modelagem fuzzy são os algoritmos Fuzzy C-Means (FCM) (BEZDEK, 1981), Guztafson-Kessel (GK) (GUZTAFSON;KESSEL, 1979), Gath-Geva (GG) (GATH;GEVA, 1989) e suas variações.

As abordagens de modelagem fuzzy que usam agrupamento seguem uma forte tendência para a geração de sistemas fuzzy do tipo Takagi-Sugeno (TS) (TAKAGI;SUGENO, 1985), já que no seu formato de regras o conseqüente é um valor numérico e, portanto, pode ser beneficiado pelo processo de geração de regras por agrupamento em que cada grupo representa uma regra (ZHAO et al., 1994) (KIM et al., 1997) (SETNES et al., 1998) (MONTES et al., 2006) (HADJULI;WERTZ, 2002). Para este tipo de abordagem, vários algoritmos de agrupamento são usados juntamente com outros métodos para a otimização das regras. Entre os métodos mais comuns pode-se destacar Algoritmos Genéticos e Gradiente Descendente.

Existe ainda a tentativa de melhorar os próprios algoritmos de agrupamento fuzzy, pois a maneira como esses algoritmos de se comportam influencia na modelagem fuzzy; esta tendência é aqui citada como otimização dos algoritmos de agrupamento fuzzy (PEDRYCZ;VUKOCIVH, 2004) (HIROTA;PEDRYCZ, 1995) (PEDRYCZ, 1996a) (HIROTA;PEDRYCZ, 1996) (TIMM, 2001) (ABONYI et al., 2002) (ALEXIEV;GEORGIEVA, 2004).

Por fim, uma promissora abordagem que usa o algoritmo de agrupamento na modelagem fuzzy é chamada de Modelagem Granular (PEDRYCZ;VASILAKOS, 1999), que utiliza

o algoritmo de agrupamento Fuzzy Condicional pelo qual contextos são definidos sobre o domínio de saída e esta definição direciona o processo de agrupamento estabelecido pelo domínio de entrada.

Importantes trabalhos que seguiram de alguma forma estas tendências são brevemente apresentados a seguir.

### **Agrupamento para geração completa das regras**

Para modelagem fuzzy usando agrupamento cujos grupos representam regras, geralmente é feita a integração entre a identificação da estrutura, a qual diz respeito à determinação do número de regras fuzzy e as partições do espaço de entrada em conjuntos fuzzy, e a determinação de parâmetros que é a estimativa do melhor valor para determinar o formato das funções de pertinência e o coeficiente das combinações lineares.

Desta maneira, os autores citados a seguir usaram algoritmos de agrupamento e outros métodos na tentativa de identificar uma boa estrutura de regra mantendo a acuidade das regras geradas por especialistas.

ZHAO et al. (1994) propõem uma nova abordagem para identificação de modelo fuzzy por meio do agrupamento de variáveis que aparecem no conseqüente, cujos centros dos grupos são hiperplanos. As regras fuzzy implícitas são extraídas a partir do resultado do agrupamento fuzzy, utilizando o algoritmo GK por ser um dos mais apropriados algoritmos de agrupamento fuzzy para grupos lineares ou planos.

KIM et al. (1997) propõem a modelagem fuzzy com um refinamento inicial, a qual determina os valores dos parâmetros do conseqüente aproximadamente, usando um agrupamento que é uma versão modificada do FCM chamada Fuzzy C-Regression Model (FCRM); mais apropriado para a formação de hiperplanos correspondendo a uma regra fuzzy, já que o FCM retorna grupos no formato de hiperesferas, e uma afinação mais precisa (*fine tuning*), o qual ajusta os parâmetros da premissa e do conseqüente das regras pelo algoritmo de Gradiente Descendente.

Em (SETNES et al., 1998) é elaborado um modelo fuzzy que utiliza agrupamento para gerar regras e uma descrição de uma melhoria desse modelo pode ser encontrado em (SETNES;ROUBOS, 1999). Nestes trabalhos, as regras de um modelo inicial são derivadas por meio de um método de agrupamento fuzzy supervisionado (SETNES, 1999) que, em

um certo grau, assegura a transparência da base de regras resultante. Este agrupamento é iniciado com um grande número de grupos e, em seguida, ocorre a remoção sucessiva dos grupos menos importantes por meio do *orthogonal-least squares rule reduction* (WANG;MENDEL, 1992b), a fim de determinar automaticamente o número de regras. Cada grupo identificado no espaço de entrada e saída deste sistema representa uma certa região e corresponde a uma regra fuzzy da base de regras. Por fim, um Algoritmo Genético (AG) é proposto para otimizar as regras geradas previamente.

Em (MONTES et al., 2006) é proposta também uma nova metodologia de construção de modelos fuzzy interpretáveis a partir de dados de entrada e saída. Nesta abordagem, as regras são determinadas pelos grupos e o método de mínimos quadráticos ajusta os seus conseqüentes e, para garantir a precisão, o Gradiente Descendente ajusta os valores modais dos conjuntos que formam o antecedente. Este trabalho também busca prover um sistema fuzzy com boa precisão sem perder a interpretabilidade.

Em (HADJULI;WERTZ, 2002) é proposta uma maneira de reduzir a complexidade dos modelos TS, com uma estrutura simples e com uma acuidade suficiente, pela estimativa de um número ótimo de regras fuzzy e pela seleção de entradas relevantes como variáveis do antecedente independente da seleção do conseqüente. Nesta abordagem a parte **SE** (função de pertinência dos antecedentes) é desacoplada da parte **ENTÃO** (parâmetros do conseqüente). O agrupamento Fuzzy C-Means é utilizado para gerar a partição fuzzy do espaço de entrada bem como estimar o número ótimo de regras.

### Melhoria dos algoritmos de agrupamento fuzzy

Segundo PEDRYCZ;VUKOCIVH (2004), como o uso de agrupamento fuzzy se tornou predominante na modelagem fuzzy, existe uma necessidade de melhorias na taxonomia das técnicas de agrupamento. De acordo com esta necessidade é possível considerar dois aspectos em relação a forma como as técnicas de agrupamento suportam o desenvolvimento de modelos fuzzy:

1. Natureza relacional do agrupamento fuzzy, a qual significa que no processo de busca por estruturas nos dados, não há distinção entre variáveis de entrada e saída, ou seja, o foco da atividade de agrupamento é revelar uma estrutura dentro de um conjunto de dados apenas e não há nada que estipule qual ou quanto uma variável de entrada

influencia numa variável de saída.

2. Natureza direcional do agrupamento fuzzy, a qual incorpora dentro do agrupamento a distinção entre as variáveis de entrada (independentes) e saída (dependente). Como a direcionalidade é dominante em modelos fuzzy, já que deve-se tomar uma decisão em relação à seleção das variáveis de entrada e saída em um estágio prévio da modelagem do sistema antes do desenvolvimento do modelo, faz-se necessário incorporar a característica de supervisão nos algoritmos de agrupamento fuzzy.

De acordo com o aspecto direcional, inúmeros algoritmos novos e melhorias de algoritmos bem conhecidos, como FCM, GK e GG, têm sido propostos a fim de melhorar a modelagem fuzzy usando agrupamento (HIROTA;PEDRYCZ, 1995) (PEDRYCZ, 1996a) (HIROTA;PEDRYCZ, 1996) (TIMM, 2001) (ABONYI et al., 2002) (ALEXIEV;GEORGIEVA, 2004).

### **Modelagem Granular**

Uma abordagem promissora para o desenvolvimento de modelos lingüísticos baseado no algoritmo de agrupamento Fuzzy Condicional foi proposto em (PEDRYCZ;VASILAKOS, 1999) e posteriormente investigado em (PEDRYCZ;KWAK, 2006a) e (PEDRYCZ;KWAK, 2006b), nos quais contextos são definidos sobre o domínio de saída e esta definição direciona o processo de agrupamento realizado sobre o domínio de entrada. As ligações entre os grupos e os contextos revelados durante o processo refletem a relação entrada-saída do modelo.

É enfatizado por PEDRYCZ;KWAK (2006b) que o agrupamento tem um papel crucial na modelagem granular. Primeiro, porque ajuda a converter dados numéricos em informação granular formando um esqueleto do modelo. Segundo, porque o agrupamento ajuda a gerenciar problemas de dimensionalidade, que é um ponto crítico na modelagem baseada em regras trazendo sérios problemas durante o processo de desenvolvimento, pois os grupos não são dependentes do número de variáveis, e eles não são explicitamente apresentados no desenvolvimento da informação granular.

O trabalho de PEDRYCZ;VASILAKOS (1999) desenvolve uma técnica genérica de modelagem lingüística guiada por grânulos (*linguistic granules-oriented modeling*) que revela e amarra os grânulos lingüísticos (conjuntos ou relações fuzzy) de maneira explícita. Para

este trabalho, é necessário compreender que modelos lingüísticos são vistos como uma associação de informações granular (termos lingüísticos) definidos sobre algumas variáveis do sistema. É exatamente nesta tendência em que se insere a proposta deste trabalho descrita a seguir.

### 4.3 Método Proposto – Fuzzy-CCM

O método proposto neste trabalho é chamado de Fuzzy-CCM (Fuzzy Conditional Clustering based Modeling) e consiste de uma nova abordagem para modelagem fuzzy usando agrupamento condicional. O método Fuzzy-CCM é utilizado para a construção de regras fuzzy mais compactas, uma vez que valores lingüísticos dos atributos no antecedente das regras não são especificados para todas as variáveis, mas somente para um subconjunto delas.

Este formato das regras melhora a interpretabilidade porque não é necessário que todas as variáveis apareçam explicitamente na regra, mas apenas algumas variáveis, enquanto outras aparecem embutidas em grupos.

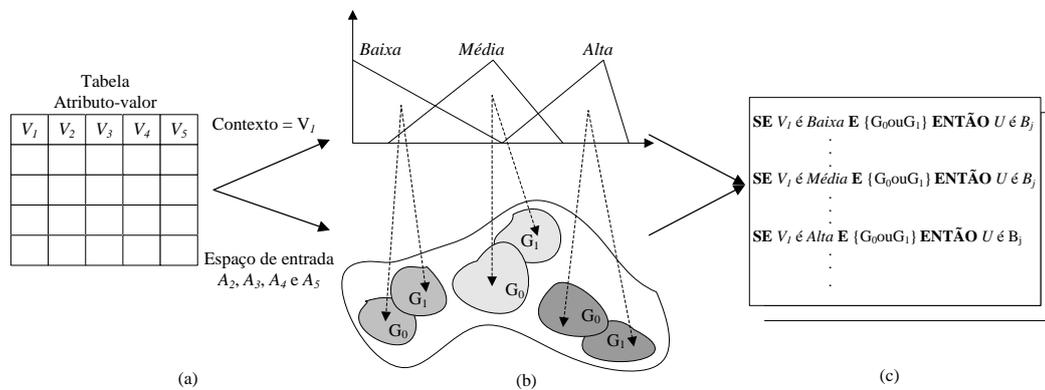
O fato de introduzir grupos no antecedente das regras não prejudica a compreensão porque se trata de um processamento do qual não é necessário ter um conhecimento prévio de todas as variáveis para que seja possível entender o que o grupo significa, basta considerar que as variáveis que não apareceram na regra contribuem para o disparo da regra unidas em forma de grupos.

A idéia básica do método proposto aqui é que os valores lingüísticos de algumas variáveis de entrada, consideradas mais relevantes, sejam usados para definir o contexto no qual os dados serão agrupados, considerando os valores das demais variáveis de entrada.

As subseções seguintes apresentam os mecanismos utilizados no desenvolvimento do método Fuzzy-CCM proposto neste trabalho que envolve a construção das regras fuzzy e geração do método de inferência.

#### 4.3.1 Construção de Regras Fuzzy usando FCMC

A abordagem desenvolvida para a construção de regras fuzzy envolve 2 etapas: Seleção de Atributos e Geração de Regras, as quais podem ser observadas na Figura 4.2.



**Figura 4.2:** Construção de Regras Fuzzy: (a) Seleção de Atributos, (b) Geração de Regras usando o FCMC, (c) Base de Regras

### Etapa 1. Seleção de Atributos

Esta etapa consiste da aplicação de um método de seleção de atributos para reduzir a dimensionalidade do problema, bem como identificar quais atributos devem ser considerados para definição do contexto. Os atributos selecionados para definir o contexto também serão aqueles que estarão explicitamente representados pelas variáveis lingüísticas na premissa das regras.

Para a seleção de atributos foi utilizado o algoritmo Relief-F (KONONENKO, 1994) que é uma otimização do Relief (KIRA;RENDELL, 1992), o qual é apropriado para estimar a relevância de atributos discretos e contínuos em dados que caracterizam apenas duas classes. O Relief-F é uma otimização do Relief porque trata o problema de múltiplas classes, avaliando a importância de um atributo pela amostragem repetida de um exemplo considerando o valor do atributo dado para a instância mais próxima com classe igual e diferente.

Foi proposta também por KONONENKO;R-SIKONIA (1996) uma extensão do Relief-F chamada RRelief-F (*Regression Relief-F*) para quando a classe é contínua, em que ao invés de considerar o conhecimento exato de que uma instância pertence a uma classe ou não, é introduzida uma probabilidade que define se duas instâncias são de classes diferentes.

Na seleção de atributos deste trabalho foi utilizada a implementação do algoritmo Relief-F disponibilizada na ferramenta WEKA (WITTEN;FRANK, 2000), a qual seleciona atributos

tanto de domínios com classes discretas quanto de domínios com classes contínuas, retornando uma lista de atributos ordenados pelo grau de relevância. Baseados nesta lista, os mais relevantes, selecionados pelo algoritmo, são selecionados para aparecer na regra e definir o contexto que orientará o agrupamento.

Os experimentos apresentados neste trabalho foram desenvolvidos usando como definições de contexto uma ou duas variáveis de entrada. Os domínios das variáveis selecionadas são granularizados por meio de conjuntos fuzzy e cada combinação de valores lingüísticos das variáveis de entrada define um contexto particular. O processo de geração de regras é repetido para cada um dos contextos obtidos.

Como pode ser visto no exemplo apresentado na Figura 4.2(a), apenas o atributo  $A_1$  foi considerado mais relevante e, portanto, o seu domínio foi granularizado nos conjuntos fuzzy *Baixa*, *Média* e *Alta*, sendo estes os contextos que guiarão o agrupamento.

Se mais de um atributo fosse considerado relevante, cada um seria granularizado individualmente e o agrupamento seria guiado pela combinação destes. Por exemplo, considere que, para um problema com  $n$  atributos, dois atributos  $V_1$  e  $V_2$  foram escolhidos como mais relevantes e cada um é granularizado em 3 conjuntos fuzzy: *Baixa*, *Média* e *Alta*. Os contextos serão gerados pela combinação dos conjuntos associados a cada atributo. Esta combinação pode ser feita utilizando a operação de conjunção fuzzy definida por alguma  $t$ -norma. Assim, para este exemplo serão obtidos 9 contextos:

$V_1$  é *Baixa* E  $V_2$  é *Baixa*

$V_1$  é *Baixa* E  $V_2$  é *Média*

$V_1$  é *Baixa* E  $V_2$  é *Alta*

$V_1$  é *Média* E  $V_2$  é *Baixa*

$V_1$  é *Média* E  $V_2$  é *Média*

$V_1$  é *Média* E  $V_2$  é *Alta*

$V_1$  é *Alta* E  $V_2$  é *Baixa*

$V_1$  é *Alta* E  $V_2$  é *Média*

$V_1$  é *Alta* E  $V_2$  é *Alta*

Dado um vetor de valores de entrada  $e_k = (a_{k1}, a_{k2}, \dots, a_{kn})$ , o valor condicionante  $f_k$  associado a esse vetor, que estipula o grau de envolvimento do vetor com o contexto definido, é calculado aplicando-se o operador de conjunção fuzzy nos graus de pertinência de  $a_{k1}$  e

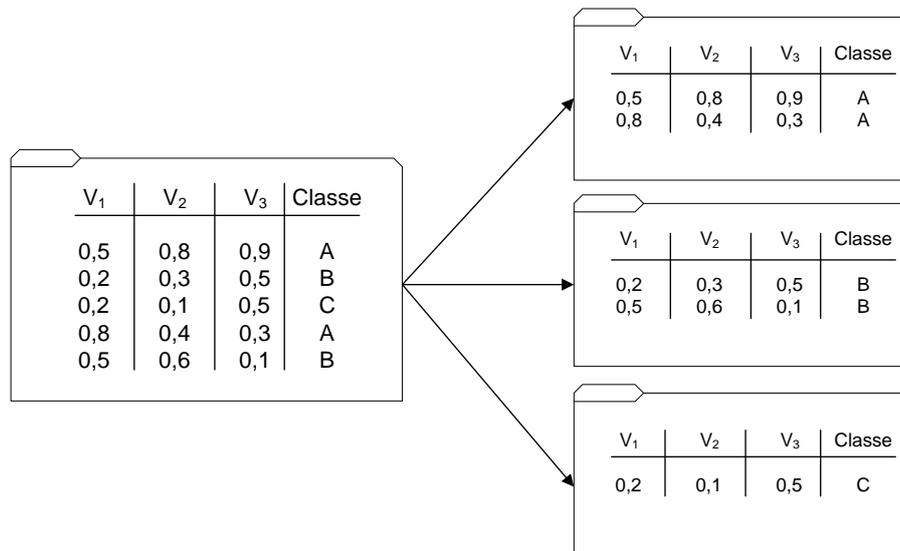
$a_{k2}$  nos conjuntos fuzzy correspondentes. Assim, se o contexto for  $V_1$  é *Baixa* E  $V_2$  é *Média*, temos

$$f_k = \min\{Baixa(a_{k1}), Média(a_{k2})\}$$

## Etapa 2. Geração de Regras Fuzzy usando FCMC

Nesta etapa, os atributos seleccionados previamente são considerados para definições de contexto para a aplicação do FCMC (PEDRYCZ, 1996a).

Para domínios com classes discretas, o conjunto de dados é dividido em um número de subconjuntos com base na quantidade de classes, conforme representado na Figura 4.3.

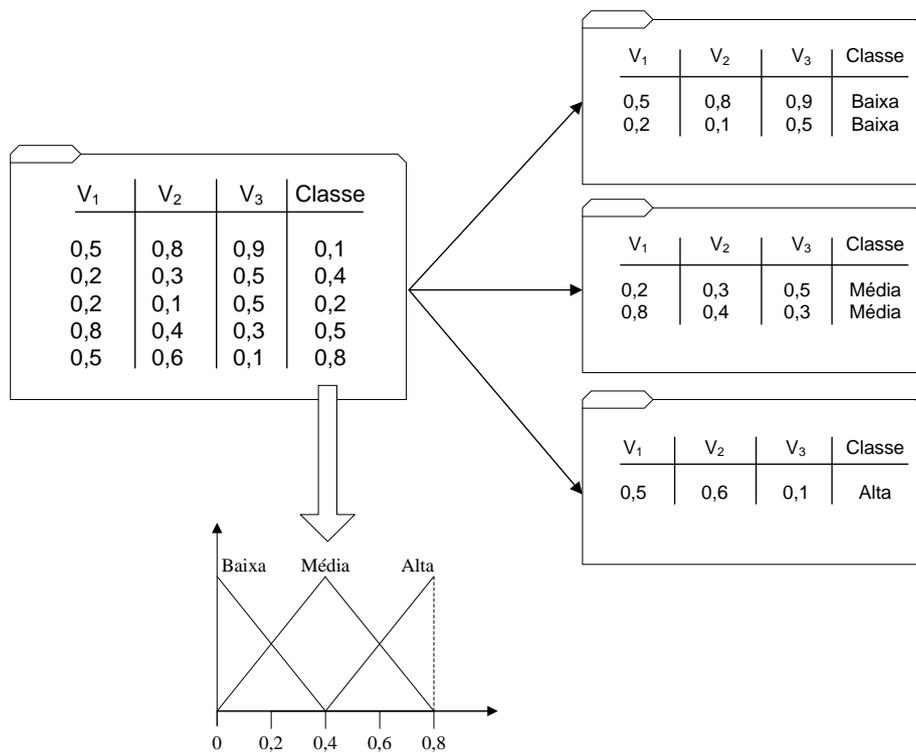


**Figura 4.3:** Divisão do conjunto de dados com classe discreta

Já para domínios com classes contínuas, o conjunto de dados é dividido em um número de subconjuntos com base na granularização prévia do domínio de saída conforme representado na Figura 4.4.

Após a separação do conjunto de dados em subconjuntos, a etapa de geração de regras é aplicada em cada um dos subconjuntos de dados para cada um dos contextos gerados e o conseqüente da regra será o valor do domínio de saída de cada subconjunto.

No geral, em um problema com  $n$  variáveis de entrada, se  $m$  ( $m \subset n$ ) variáveis de entrada são seleccionadas para definir os contextos, o agrupamento será aplicado às  $n - m$  variáveis



**Figura 4.4:** Divisão do conjunto de dados com classe contínua

restantes. Além da escolha do contexto, o número de grupos também deve ser escolhido previamente. Neste trabalho foram realizados experimentos utilizando 1, 2 e 3 grupos.

Como pode ser visto no exemplo apresentado na Figura 4.2(b), o atributo  $V_1$  guiará o agrupamento para gerar dois grupos que são representados por  $G_0$  e  $G_1$ . Isto significa que nas regras aparecerão apenas os termos da variável  $V_1$  considerados como contexto e estes dois grupos.

Esta combinação provê uma forma para gerar regras compactas com um pequeno número de variáveis de entrada explicitamente aparecendo na regra. Os valores restantes, ao invés de serem descartados, como usualmente é feito em alguns modelos lingüísticos nos quais são considerados apenas alguns atributos para geração das regras (TALON;CAMARGO, 2006) (CASTRO;CAMARGO, 2005), são também considerados escondidos em grupos de dados.

Considerando um domínio específico com 5 variáveis de entrada, sendo representados na regra pelas variáveis  $V_1, V_2, V_3, V_4$  e  $V_5$ , e que  $V_1$  é a mais relevante destas e foi selecionado

para definir os contextos definidos por  $V_1$  é *Baixo*,  $V_1$  é *Médio* e  $V_1$  é *Alto*, o processo de agrupamento pode ser descrito como:

1. *Faça o agrupamento de dados baseado nos valores das variáveis 2, 3, 4 e 5 considerando que "V<sub>1</sub> é Baixo".*
2. *Faça o agrupamento de dados baseado nos valores dos variáveis 2, 3, 4 e 5 considerando que "V<sub>1</sub> é Médio".*
3. *Faça o agrupamento de dados baseado nos valores dos variáveis 2, 3, 4 e 5 considerando que "V<sub>1</sub> é Alto".*

As regras usualmente assumirão o formato:

$$\mathbf{SE} V_1 \text{ é } A_{1l_1} \mathbf{E} \{G_0 \mathbf{OU} G_1\} \mathbf{ENTÃO} U \text{ é } B_j$$

na qual  $G_0$  e  $G_1$  representam os grupos formados pelos demais atributos do domínio no contexto definido pela variável  $V_1$ . A base de regras gerada por este método pode ser então observada na Figura 4.2(c).

### 4.3.2 Método de Inferência

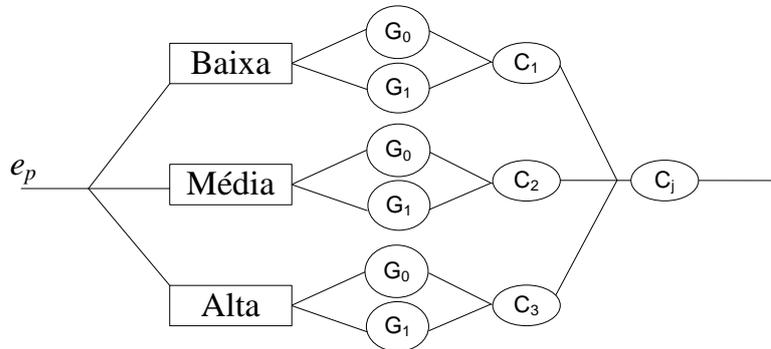
Para este trabalho foram desenvolvidos dois métodos de inferência, um para cada tipo de domínio: com classe discreta e com classe contínua.

#### Inferência para domínios com classe discreta

O método de inferência usado aqui é adaptado do método geral descrito na Seção 2.3.1 do Capítulo 2.

Seja  $e_p = (a_{p1}, a_{p2}, \dots, a_{pn})$  um vetor de entrada e  $\{R_1, R_2, \dots, R_S\}$  o conjunto de  $S$  regras no formato definido previamente na subseção anterior, cada uma com  $m$  variáveis de entrada com valores lingüísticos definidos explicitamente e  $k$  grupos definidos nas  $n - m$  variáveis de entrada restantes.

A Figura 4.5 representa o processo de inferência para uma base com três regras, em que cada uma aparece apenas uma variável explícita a qual foi granularizada nos termos *Baixa*, *Média* e *Alta*, e os grupos  $G_0$  e  $G_1$  guiados por cada termo.



**Figura 4.5:** Inferência para domínios com classe discreta

O processo de inferência para domínios com classes discretas utilizado para processar o vetor de entrada  $e_p$  (ilustrado na Figura 4.5) é:

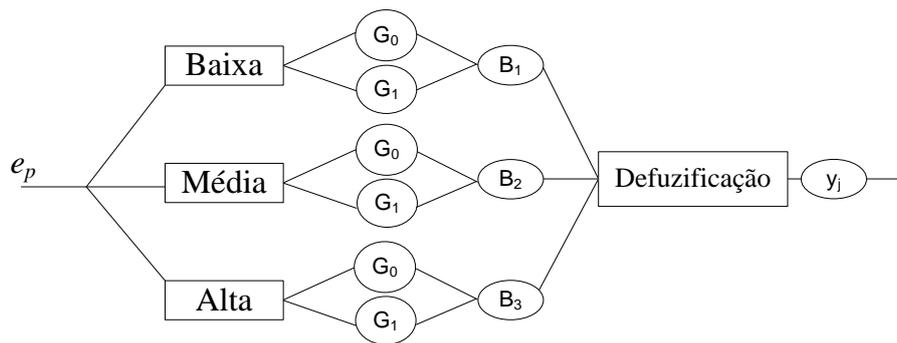
1. Calcular o grau de compatibilidade entre o vetor de entrada  $e_p$  e a “primeira parte” de cada regra  $R_k$ ,  $k = 1, \dots, S$ , baseado nas  $m$  variáveis de entrada que aparecem explicitamente no antecedente da regra;
2. Encontrar o conjunto de regras  $R_{set}$  com o maior grau de compatibilidade com o vetor de entrada. (Haverá uma regra para cada um dos subconjuntos de dados de entrada);
3. Para cada regra em  $R_{set}$ , calcular o grau de pertinência do vetor de entrada nos seus grupos e escolher o maior deles como o grau de disparo da regra;
4. Selecionar a regra com o maior grau de disparo e definir como saída do processo de inferência a classe da regra disparada.

#### Inferência para domínios com classe contínua

O método de inferência para domínios com saída contínua é adaptado do método geral descrito na Seção 2.3.1 do Capítulo 2 e inspirado no método de Mamdani (MAMDANI, 1977).

Seja  $e_p = (a_{p1}, a_{p2}, \dots, a_{pn})$  um vetor de entrada e  $\{R_1, R_2, \dots, R_S\}$  o conjunto de  $S$  regras no formato definido previamente na subseção anterior, cada um com  $m$  variáveis de entrada e  $k$  grupos definidos nas  $n - m$  variáveis de entrada restantes.

A Figura 4.6 a seguir representa o processo de inferência para uma base com três regras, em que cada regra aparece apenas uma variável explícita, a qual foi granularizada nos termos *Baixa*, *Média* e *Alta*, os grupos  $G_0$  e  $G_1$  guiados por cada termo, os conjuntos fuzzy  $B_1$ ,  $B_2$  e  $B_3$  como resultado da granularização do domínio de saída e o processo de defuzzificação, o qual transformará a saída fuzzy  $B_1$ ,  $B_2$  ou  $B_3$  em um resultado real  $y_j$



**Figura 4.6:** Inferência para domínios com classe contínua

O processo de inferência para domínios com classes contínuas utilizado para processar o vetor de entrada  $e_p$  (ilustrado na Figura 4.6) é:

1. Calcular o grau de compatibilidade entre o vetor de entrada  $e_p$  e a “primeira parte” de cada regra  $R_k$ ,  $k = 1, \dots, S$ , baseado nas  $m$  variáveis de entrada que aparecem explicitamente no antecedente da regra;
2. Encontrar o conjunto de regras  $R_{set}$  com o maior grau de compatibilidade com o vetor de entrada (Haverá uma regra para cada um dos subconjuntos de dados de entrada);
3. Para cada regra em  $R_{set}$ , calcular o grau de pertinência do vetor de entrada nos seus grupos e escolher o maior deles como o grau de disparo da regra;
4. Aplicar um método de defuzzificação, apresentado na Subseção 2.3.3 do Capítulo 2, considerando o valor de saída lingüístico de cada regra separados e seu correspondente grau de disparo, gerando o valor de saída  $y_j$ .

## 4.4 Considerações Finais

Neste capítulo foi apresentado o método Fuzzy-CCM proposto neste trabalho com suas definições e etapas de desenvolvimento. Para validar esta proposta foram feitos vários experimentos e análises.

Os experimentos são necessários não só para a validação do método, mas, principalmente, para buscar respostas para perguntas em relação à modelagem que surgem à medida que se deseja que o método cubra vários domínios e suas restrições. Para tanto, foram feitos experimentos a fim de responder questões relacionadas ao número de grupos, o número de conjuntos fuzzy, quantidade de atributos relevantes e à escolha do método de defuzzificação para domínios com classes contínuas. No capítulo a seguir serão apresentadas estas análises e experimentos.

---

# Resultados Experimentais

---

O objetivo deste capítulo é apresentar e analisar os resultados dos experimentos realizados com os sistemas fuzzy resultantes da aplicação do método Fuzzy-CCM, o qual é uma proposta de abordagem para modelagem fuzzy utilizando agrupamento condicional. Além da interpretabilidade, foco deste trabalho, será também analisada a acuidade dos sistemas resultantes.

A Seção 5.1 descreve os parâmetros, conjuntos de dados e algumas considerações iniciais acerca da metodologia utilizada nos experimentos e a Seção 5.2 os experimentos e resultados obtidos.

## 5.1 Considerações Preliminares

Os resultados apresentados neste capítulo foram obtidos por meio de sistemas de classificação fuzzy utilizando conjuntos de dados com classes discretas e inferência fuzzy utilizando conjuntos de dados com valores contínuos de saída.

Os conjuntos de dados foram extraídos do repositório UCI (BLAKE et al., 1998), os quais são apresentados na Tabela 5.1 ordenados pela quantidade de atributos, apresentando

também a quantidade de instâncias e tipo de classe. A escolha de cada conjunto de dados foi baseada nos tipos de atributos dos domínios, os quais são todos numéricos.

**Tabela 5.1:** Características dos Domínios

Domínio	Atributos	Instâncias	Tipo de Classe
Íris	4	150	Discreta
Bupa	6	290	Discreta
Diabetes	8	768	Discreta
AutoMpg	8	398	Contínua
Wine	13	90	Discreta
Pollution	16	60	Contínua
Vehicle	18	796	Discreta

A fim de estimar o erro mais próximo do erro real foi utilizado o método *10-fold cross validation* MITCHELL (1997) em todos os experimentos, pelo qual os padrões foram divididos em 10 partições independentes e em cada execução uma das partições foi usada para teste e as restantes foram usadas como dados de treinamento para a geração das regras.

Os parâmetros utilizados pelo agrupamento Fuzzy C-Means Condicional (FCMC) foram 1,25 para taxa de fuzificação e 0,01 para taxa de convergência. Para a seleção de atributos de todos os domínios foi utilizado o ReliefF, apresentado na Subseção 4.3.1 do capítulo anterior.

## 5.2 Experimentos, Resultados e Análises

Para avaliação do método Fuzzy-CCM apresentado no capítulo anterior, foram realizados dois tipos de experimento:

1. Concepção da abordagem, nos quais foram utilizados domínios com classes discretas, não havendo necessidade da etapa de defuzificação, gerando regras com apenas 2 grupos e atributos granularizados com apenas 3 conjuntos fuzzy.
2. Ampliação da abordagem, nos quais foram utilizados domínios com classes contínuas, testados com 2 diferentes métodos de defuzificação, regras com 1, 2 e 3 grupos, e granularização com 3, 5 e 7 conjuntos fuzzy no domínio da variável de saída.

### 5.2.1 Concepção da Abordagem

Este tipo de experimento visa conceber a idéia, testar a implementação e validar a proposta. Para tanto, foram utilizadas as bases Iris, Bupa, Diabetes, Wine e Vehicle. Para validação do desempenho por meio da comparação dos resultados obtidos com o método Fuzzy-CCM foi utilizado o método Wang&Mendel (WM) (WANG;MENDEL, 1992a). Também foram feitos testes com o WM e o método Fuzzy C-Means (FCM) convencional utilizando todos os atributos do domínio, o que permite a análise em relação ao balanço interpretabilidade-aciudade.

É importante notar que, neste trabalho, a avaliação da melhoria na interpretabilidade é feita com base na redução do número de variáveis explícitas no antecedente das regras.

A fim de minimizar o esforço computacional neste tipo de experimento foram geradas regras com apenas 2 grupos e os domínios dos atributos de entrada foram granularizados em apenas 3 conjuntos fuzzy.

Cada atributo considerado relevante aparece na regra como um variável explícita, já que os demais aparecerão escondidos no agrupamento como explicado no capítulo anterior.

Para a análise de desempenho do conjunto de regras gerado foi utilizada a Taxa de Classificação Correta (TCC) definida por:

$$TCC = \frac{N^{\circ} \text{ de acertos}}{\text{Quantidade de padrões}} \quad (5.1)$$

A seguir são apresentados os resultados obtidos por cada conjunto de dados escolhido quando realizados experimentos para concepção da abordagem.

#### Iris

Esse conjunto de dados envolve três espécies (classes) de flores (Iris setosa, Iris versicolor e Iris virginica) e 4 atributos. Para este conjunto de dados, o ReliefF retornou a seguinte ordem de importância dos atributos:  $x_3$ ,  $x_4$ ,  $x_2$  e  $x_1$ . Baseado nesta ordem, foram escolhidos os atributos comprimento da pétala da flor ( $x_3$ ) e largura da pétala da flor ( $x_4$ ), individualmente e combinados, para a definição dos contextos na realização dos experimentos.

A Tabela 5.2 apresenta as taxas de classificação correta médias e o desvio padrão obti-

das pelo uso das variáveis explícitas mais agrupamento no método Fuzzy-CCM e o uso apenas das variáveis explícitas no método WM. A última linha da tabela mostra o resultado obtido usando os métodos WM e FCM com todos os atributos/variáveis.

**Tabela 5.2:** Taxas de Classificação Correta - Iris

Variáveis Explícitas	Fuzzy-CCM	WM
$V_3$	0,64(0,14)	0,92(0,08)
$V_4$	0,72(0,13)	0,89(0,08)
$V_3V_4$	0,62(0,15)	0,95(0,04)
Média	0,66(0,04)	0,92(0,02)
	FCM	WM
Todas	0,75(0,12)	1,0(0,001)

Neste experimento foi possível observar que o método Fuzzy-CCM não teve um desempenho consideravelmente melhor que o método WM, devido ao resultado da influência da pequena quantidade de atributos a serem trabalhados com agrupamento. Quando os métodos WM e FCM são avaliados com todos os atributos, é possível confirmar esta análise, uma vez que o FCM também apresentou desempenho inferior ao WM.

### Bupa

Esse conjunto de dados busca avaliar sintomas de doenças do fígado, classificando pessoas em doentes ou não. Para este conjunto de dados, o ReliefF retornou a seguinte ordem de importância dos atributos:  $x_3$ ,  $x_5$ ,  $x_6$ ,  $x_4$ ,  $x_1$  e  $x_2$ . Baseado nesta ordem, foram escolhidos os atributos (nomes técnicos para alguns exames de sangue) *sgpt alamine aminotransferase* ( $x_3$ ) e *gammagt gamma-glutamyl transpeptidase* ( $x_5$ ), individualmente e combinados, para a definição dos contextos na realização dos experimentos.

A Tabela 5.3 apresenta as taxas de classificação correta médias e o desvio padrão obtidas pelo uso das variáveis explícitas mais agrupamento no método Fuzzy-CCM e o uso apenas das variáveis explícitas no método WM. A última linha da tabela mostra o resultado obtido usando os métodos WM e FCM com todos os atributos/variáveis.

Com este conjunto de dados foi possível observar um ganho considerável com o uso do método Fuzzy-CCM, tanto em relação à interpretabilidade, quanto à acuidade. O ganho em termos de interpretabilidade deve-se ao fato das regras serem mais compactas com o uso de apenas algumas variáveis explícitas. Já o ganho em relação à acuidade resulta do

**Tabela 5.3:** Taxas de Classificação Correta - Bupa

Variáveis Explícitas	Fuzzy-CCM	WM
$V_3$	0,610(0,06)	0,500(0,00)
$V_5$	0,560(0,09)	0,520(0,04)
$V_3V_5$	0,580(0,12)	0,520(0,04)
Média	0,58(0,02)	0,51(0,01)
	FCM	WM
Todos	0,62(0,06)	0,57(0,09)

fato do método Fuzzy-CCM alcançar taxas corretas de classificação usando apenas alguns atributos, superior às taxas encontradas com o uso de todos os atributos pelo WM.

### Diabetes

Este conjunto de dados apresenta o diagnóstico (classes), positivo ou negativo, de pessoas que apresentam sintomas relacionados à diabetes. Para este conjunto de dados, o ReliefF retornou a seguinte ordem de importância dos atributos:  $x_2$ ,  $x_6$ ,  $x_4$ ,  $x_1$ ,  $x_8$ ,  $x_7$ ,  $x_3$  e  $x_5$ . Baseado nesta ordem, foram escolhidos os atributos quantidade de vezes de gestação ( $x_1$ ), concentração de glucose ( $x_2$ ), flacidez do tríceps ( $x_4$ ) e massa corporal ( $x_6$ ), individualmente e combinados 2 a 2, para a definição dos contextos na realização dos experimentos.

A Tabela 5.4 apresenta as taxas de classificação correta médias e o desvio padrão obtidas pelo uso das variáveis explícitas mais agrupamento no método Fuzzy-CCM e o uso apenas das variáveis explícitas no método WM. A última linha da tabela mostra o resultado obtido usando os métodos WM e FCM com todos os atributos/variáveis.

**Tabela 5.4:** Taxas de Classificação Correta - Diabetes

Variáveis Explícitas	Fuzzy-CCM	WM
$V_1$	0,54(0,08)	0,49(0,06)
$V_1V_2$	0,42(0,07)	0,52(0,07)
$V_1V_4$	0,55(0,05)	0,50(0,06)
$V_1V_6$	0,51(0,07)	0,46(0,09)
$V_2$	0,46(0,04)	0,38(0,13)
$V_2V_4$	0,47(0,04)	0,45(0,06)
$V_2V_6$	0,43(0,06)	0,58(0,06)
$V_4$	0,58(0,04)	0,35(0,08)
$V_4V_6$	0,58(0,05)	0,48(0,06)
$V_6$	0,59(0,07)	0,57(0,05)
Média	0,51(0,06)	0,48(0,07)
	FCM	WM
Todos	0,59(0,08)	0,56(0,08)

Com este conjunto de dados foi possível observar ganhos do Fuzzy-CCM em relação ao método WM, o que pode ser afirmado pelo fato de que o FCM convencional também apresenta taxas melhores quando utilizados todos os atributos. Esta vantagem do FCM pôde ser refletida na nova abordagem.

### Wine

Esta conjunto de dados é o resultado da análise química de vinhos produzidos numa mesma região da Itália mas derivados de três diferentes cultivadores (classes). A análise determina a quantidade dos 13 constituintes encontrados nos três tipos de vinho.

Para este conjunto de dados, o ReliefF retornou a seguinte ordem de importância dos atributos:  $x_{12}$ ,  $x_7$ ,  $x_{13}$ ,  $x_{10}$ ,  $x_1$ ,  $x_{11}$ ,  $x_6$ ,  $x_8$ ,  $x_2$ ,  $x_9$ ,  $x_4$ ,  $x_5$  e  $x_3$ . Baseado nesta ordem, foram escolhidos os atributos álcool ( $x_1$ ), flavonoides ( $x_7$ ), intensidade da cor ( $x_{10}$ ), tonalidade ( $x_{11}$ ), OD280/OD315 de vinhos diluídos  $x_{12}$  e Proline (um tipo de aminoácido) ( $x_{13}$ ), individualmente e combinados 2 a 2, para a definição dos contextos na realização dos experimentos.

A Tabela 5.5 apresenta as taxas de classificação correta médias e o desvio padrão obtidas pelo uso das variáveis explícitas mais agrupamento no método Fuzzy-CCM e o uso apenas das variáveis explícitas no método WM. A última linha da tabela mostra o resultado obtido usando os métodos WM e FCM com todos os atributos/variáveis.

É possível concluir como análise deste conjunto de dados que não houve um alto ganho em acuidade do Fuzzy-CCM em relação ao WM porque métodos de agrupamento não têm um bom desempenho quando o número de instâncias é reduzido, o que pode ser comprovado com o uso do FCM com todos os atributos. Mesmo assim, o método Fuzzy-CCM apresentou bom desempenho em relação ao WM principalmente utilizando o atributo  $x_7$ .

### Vehicle

Este conjunto de dados classifica uma dada silhueta de carro como um dos quatro tipos (classes) de veículos usando um conjunto de características extraídas da silhueta. O veículo pode ser visto de um de muitos diferentes ângulos. Quatro modelos de veículos “Corgie” foram usados no experimento: a *double decker bus*, *Chevrolet van*, *Saab 9000* e um *Opel Manta 400*.

Para este conjunto de dados, o ReliefF retornou a seguinte ordem de importância dos

**Tabela 5.5:** Taxas de Classificação Correta - Wine

Variáveis Explícitas	Fuzzy-CCM	WM
$V_1$	0,63(0,09)	0,47(0,09)
$V_1V_7$	0,63(0,16)	0,74(0,17)
$V_1V_{10}$	0,63(0,09)	0,69(0,15)
$V_1V_{11}$	0,59(0,08)	0,48(0,07)
$V_1V_{12}$	0,69(0,14)	0,74(0,15)
$V_1V_{13}$	0,54(0,17)	0,60(0,08)
$V_7$	0,72(0,11)	0,34(0,04)
$V_7V_{10}$	0,74(0,13)	0,89(0,17)
$V_7V_{11}$	0,67(0,15)	0,54(0,10)
$V_7V_{12}$	0,62(0,17)	0,60(0,11)
$V_7V_{13}$	0,56(0,23)	0,59(0,11)
$V_{10}$	0,61(0,08)	0,67(0,15)
$V_{10}V_{11}$	0,66(0,15)	0,78(0,20)
$V_{10}V_{12}$	0,70(0,11)	0,87(0,20)
$V_{10}V_{13}$	0,52(0,15)	0,68(0,14)
$V_{11}$	0,65(0,13)	0,36(0,05)
$V_{11}V_{12}$	0,72(0,15)	0,63(0,13)
$V_{11}V_{13}$	0,52(0,17)	0,70(0,13)
$V_{12}$	0,69(0,13)	0,61(0,11)
$V_{12}V_{13}$	0,61(0,19)	0,44(0,15)
$V_{13}$	0,53(0,20)	0,54(0,18)
Média	0,63(0,06)	0,62(0,14)
	FCM	WM
Todas	0,63(0,14)	0,91(0,07)

atributos:  $x_8, x_{18}, x_7, x_{12}, x_9, x_3, x_{10}, x_{11}, x_2, x_{17}, x_1, x_{13}, x_4, x_6, x_{14}, x_{15}, x_5$  e  $x_{16}$ . Baseado nesta ordem, foram escolhidos os atributos taxa de dispersão ( $x_7$ ), alongamento ( $x_8$ ), *Pr.Axis Rectangularity* ( $x_9$ ), variância escalar ( $x_{12}$ ) e taxa da concavidade ( $x_{18}$ ), individualmente e combinados 2 a 2, para a definição dos contextos na realização dos experimentos.

A Tabela 5.6 apresenta as taxas de classificação correta médias e o desvio padrão obtidas pelo uso das variáveis explícitas mais agrupamento no método Fuzzy-CCM e o uso apenas das variáveis explícitas no método WM. A última linha da tabela mostra o resultado obtido usando os métodos WM e FCM com todos os atributos/variáveis.

O método Fuzzy-CCM apresentou um desempenho superior comparado ao WM em relação à acuidade utilizando apenas algumas variáveis explícitas e valida a ideia de que em alguns casos a acuidade pode ser beneficiada pela interpretabilidade.

Observa-se, pela Tabela 5.6, que o desempenho do método WM com todas as variáveis foi superior a todos os demais experimentos, que consideraram no máximo duas variáveis

**Tabela 5.6:** Taxas de Classificação Correta - Vehicle

Variáveis Explícitas	Fuzzy-CCM	WM
$V_7$	0,40(0,07)	0,29(0,02)
$V_7V_8$	0,43(0,06)	0,40(0,05)
$V_7V_9$	0,40(0,08)	0,30(0,02)
$V_7V_{12}$	0,44(0,05)	0,35(0,04)
$V_7V_{18}$	0,43(0,07)	0,39(0,04)
$V_8$	0,41(0,07)	0,44(0,06)
$V_8V_9$	0,40(0,05)	0,42(0,05)
$V_8V_{12}$	0,46(0,06)	0,46(0,07)
$V_8V_{18}$	0,39(0,05)	0,49(0,08)
$V_9$	0,37(0,05)	0,28(0,01)
$V_9V_{12}$	0,44(0,06)	0,33(0,03)
$V_9V_{18}$	0,42(0,05)	0,37(0,03)
$V_{12}$	0,45(0,06)	0,28(0,01)
$V_{12}V_{18}$	0,49(0,05)	0,37(0,01)
$V_{18}$	0,37(0,06)	0,25(0,00)
Média	0,42(0,03)	0,36(0,07)
	FCM	WM
Todas	0,365(0,03)	0,71(0,08)

explícitas. Esse resultado é compreensível devido ao alto número de atributos da base. Por outro lado, o uso de 18 atributos no antecedente da regra prejudica a compreensão das mesmas.

### 5.2.2 Ampliação da Abordagem

Este tipo de experimento visa o aprimoramento do método Fuzzy-CCM em relação aos experimentos anteriores e uma análise mais detalhada das suas restrições. Para tanto, foram utilizados os seguintes conjuntos de dados com classes contínuas: AutoMpg e Pollution.

Foram realizados testes usando dois métodos de defuzzificação: Média de Elementos Máximos e Centro de Máximos Ponderado, apresentados no Capítulo 2, além de testes alterando a quantidade de grupos na construção da regra, bem como a quantidade de conjuntos fuzzy pela qual o domínio da Variável de saída é granularizada.

A avaliação do desempenho foi feita utilizando a taxa de erro definida por:

$$err = \frac{1}{n} \sum_{i=1}^n (y_i - h(x_i))^2 \quad (5.2)$$

na qual,  $n$  é a quantidade de exemplos,  $y_i$  é o valor de saída esperado e  $h(x_i)$  é o valor de saída encontrado, ambos para o exemplo  $i$ .

A seguir são apresentados os conjuntos de dados e os testes realizados com cada um.

### AutoMpg

O conjunto de dados AutoMpg possui 8 atributos de carros de diversas marcas e modelos, e uma classe contínua, que indica o consumo. Para este conjunto de dados, o ReliefF retornou a seguinte ordem de importância dos atributos:  $x_6, x_5, x_1, x_4, x_3, x_2$  e  $x_7$ . Baseado nesta ordem, foram escolhidos os atributos aceleração ( $x_6$ ) e peso ( $x_5$ ), individualmente e combinados, para a definição dos contextos na realização dos experimentos.

A Tabela 5.7 apresenta as taxas de erro médio e desvio padrão para avaliação dos resultados obtidos utilizando o método de defuzzificação Média de Elementos Máximos (MEM) usando os métodos WM e FCM (F2-Means e F3-Means com 2 e 3 grupos, respectivamente) com todos os atributos com o domínio de saída granularizado com 3, 5 e 7 conjuntos fuzzy.

**Tabela 5.7:** AutoMpg: Taxas de Erro médio e desvio padrão usando os métodos WM e FCM para geração das regras e MEM para defuzzificação

Conjuntos	WM	F2-Means	F3-Means
3	0,044 (0,009)	0,022 (0,008)	0,018 (0,008)
5	0,038 (0,01)	0,009 (0,003)	0,01 (0,004)
7	0,04 (0,008)	0,006 (0,003)	0,005 (0,005)
Média	0,041 (0,003)	0,012 (0,008)	0,011 (0,007)

A Tabela 5.8 apresenta as taxas de erro para avaliação dos resultados obtidos utilizando o método de defuzzificação Centro de Máximo Ponderado (CMP) usando os métodos WM e FCM (F2-Means e F3-Means com 2 e 3 grupos, respectivamente) com todos os atributos com o domínio de saída granularizado com 3, 5 e 7 conjuntos fuzzy.

**Tabela 5.8:** AutoMpg: Taxas de Erro médio e desvio padrão usando os métodos WM e FCM para geração das regras e CMP para defuzzificação

Conjuntos	WM	F2-means	F3-means
3	0,008 (0,001)	0,015 (0,006)	0,013 (0,006)
5	0,007 (0,002)	0,017 (0,002)	0,016 (0,002)
7	0,007 (0,001)	0,016 (0,001)	0,015 (0,002)
Média	0,007 (0,001)	0,016 (0,001)	0,015 (0,002)

Os experimentos que resultaram nas Tabelas 5.7 e 5.8 anteriores permitiram a avali-

ação do uso de dois métodos de defuzzificação diferentes, a influência da quantidade de conjuntos fuzzy para o domínio de saída e da quantidade de grupos.

As Tabelas 5.9 e 5.10 a seguir, apresentam os mesmos testes feitos anteriormente, porém analisando os resultados obtidos com o método Fuzzy-CCM com 1, 2 e 3 grupos comparando com o WM. As variáveis utilizadas na definição de contexto indicadas na primeira coluna, foram também consideradas para o método WM.

**Tabela 5.9:** AutoMpg: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e MEM para defuzzificação

Variáveis explícitas	WM	1 Grupo	2 Grupos	3 Grupos
$V_5 - 3$	0,065 (0,021)	0,014 (0,006)	0,023 (0,004)	0,03 (0,008)
$V_5 - 5$	0,029 (0,009)	0,013 (0,005)	0,01 (0,005)	0,015 (0,003)
$V_5 - 7$	0,038 (0,01)	0,013 (0,005)	0,009 (0,004)	0,014 (0,005)
Média	0,044 (0,019)	0,013 (0,001)	0,014 (0,008)	0,019 (0,009)
$V_5 V_6 - 3$	0,069 (0,02)	0,021 (0,009)	0,024 (0,006)	0,027 (0,011)
$V_5 V_6 - 5$	0,034 (0,005)	0,018 (0,005)	0,01 (0,005)	0,021 (0,004)
$V_5 V_6 - 7$	0,045 (0,006)	0,017 (0,005)	0,017 (0,005)	0,019 (0,006)
Média	0,049 (0,018)	0,018 (0,002)	0,017 (0,007)	0,022 (0,004)
$V_6 - 3$	0,066 (0,022)	0,016 (0,009)	0,022 (0,01)	0,022 (0,012)
$V_6 - 5$	0,031 (0,007)	0,017 (0,006)	0,01 (0,002)	0,025 (0,006)
$V_6 - 7$	0,043 (0,009)	0,016 (0,004)	0,015 (0,003)	0,018 (0,004)
Média	0,047 (0,018)	0,016 (0,001)	0,016 (0,006)	0,022 (0,003)
Média Geral	0,047 (0,003)	0,016 (0,003)	0,016 (0,002)	0,021 (0,002)

**Tabela 5.10:** AutoMpg: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e CMP para defuzzificação

Variáveis Explícitas	WM	1 Grupo	2 Grupos	3 Grupos
$V_5 - 3$	0,052 (0,018)	0,067 (0,016)	0,016 (0,002)	0,013 (0,002)
$V_5 - 5$	0,017 (0,007)	0,062 (0,011)	0,014 (0,003)	0,012 (0,003)
$V_5 - 7$	0,025 (0,007)	0,064 (0,012)	0,015 (0,002)	0,013 (0,002)
Média	0,031 (0,018)	0,064 (0,003)	0,015 (0,001)	0,013 (0,001)
$V_5 V_6 - 3$	0,042 (0,012)	0,083 (0,023)	0,012 (0,003)	0,011 (0,003)
$V_5 V_6 - 5$	0,015 (0,002)	0,07 (0,014)	0,012 (0,003)	0,011 (0,003)
$V_5 V_6 - 7$	0,019 (0,003)	0,067 (0,013)	0,012 (0,003)	0,01 (0,002)
Média	0,025 (0,015)	0,073 (0,008)	0,012 (0)	0,01 (0)
$V_6 - 3$	0,049 (0,02)	0,072 (0,026)	0,012 (0,003)	0,011 (0,003)
$V_6 - 5$	0,016 (0,002)	0,073 (0,018)	0,013 (0,003)	0,01 (0,003)
$V_6 - 7$	0,024 (0,004)	0,069 (0,013)	0,012 (0,003)	0,011 (0,002)
Média	0,03 (0,017)	0,071 (0,002)	0,012 (0)	0,011 (0)
Média Geral	0,029 (0,003)	0,07 (0,005)	0,013 (0,001)	0,011 (0,001)

Com estes resultados é possível fazer diferentes análises:

1. Influência da quantidade de grupos: Os testes realizados com o conjunto de dados

AutoMpg, produziram resultados que variaram muito quando observada a quantidade de grupos (1, 2 ou 3) utilizada. Na média geral, este conjunto de dados obteve um melhor ganho do método Fuzzy-CCM em relação ao WM quando utilizado 3 grupos.

2. Influência da quantidade de conjuntos fuzzy no domínio da variável de saída: Para este conjunto de dados, a quantidade de conjuntos fuzzy entre 5 e 7 apresentou sempre melhores resultados. No entanto, não foi possível afirmar qual quantidade é realmente satisfatória.
3. Influência do método de defuzzificação: Embora o método CMP apresente os melhores resultados, não se pode afirmar que este seja o melhor método de defuzzificação para ser utilizado pelo método Fuzzy-CCM, pois existem um grande número de métodos que se adequam a cada tipo de problema.

### Pollution

Este conjunto de dados representa estimativas da taxa de mortalidade relacionada a qualidade do ar. Para este conjunto de dados, o ReliefF retornou a seguinte ordem de importância dos atributos:  $x_1, x_9, x_2, x_5, x_6, x_3, x_{15}, x_{12}, x_{13}, x_8, x_7, x_4, x_{11}, x_{14}$  e  $x_{10}$ . Baseado nesta ordem, foram escolhidos os atributos média anual de precipitação em polegadas ( $x_1$ ), média das temperaturas em Janeiro em °F ( $x_2$ ), média do tamanho da família ( $x_5$ ) e população negra em áreas urbanizadas  $x_9$ , individualmente e combinados 2 a 2, para a definição dos contextos na realização dos experimentos.

A Tabela 5.11 apresenta as taxas de erro médias e desvio padrão para avaliação dos resultados obtidos utilizando o método de defuzzificação Média de Elementos Máximos (MEM) usando os métodos WM e FCM (F2-Means e F3-Means com 2 e 3 grupos, respectivamente) com todos os atributos com o domínio de saída granularizado com 3, 5 e 7 conjuntos fuzzy.

**Tabela 5.11:** Pollution: Taxas de Erro usando os métodos WM e FCM para geração das regras e MEM para defuzzificação

Conjuntos	WM	F2-Means	F3-Means
3	0,142 (0,021)	0,008 (0,003)	0,015 (0,007)
5	0,138 (0,037)	0,003 (0,002)	0,01 (0,011)
7	0,134 (0,045)	0,003 (0,003)	0,005 (0,004)
Média	0,138 (0,004)	0,005 (0,003)	0,01 (0,005)

A Tabela 5.12 apresenta as taxas de erro médias e desvio padrão para avaliação dos resultados obtidos utilizando o método de defuzzificação Centro de Máximos Ponderado (CMP) usando os métodos WM e FCM (F2-Means e F3-Means com 2 e 3 grupos, respectivamente) com todos os atributos com o domínio de saída granularizado com 3, 5 e 7 conjuntos fuzzy.

**Tabela 5.12:** Pollution: Taxas de Erro usando os métodos WM e FCM para geração das regras e CMP para defuzzificação

Conjuntos	WM	F2-Means	F3-Means
3	0,0012 (0,0014)	0,0009 (0,0010)	0,0010 (0,0009)
5	0,0008 (0,0006)	0,0017 (0,0006)	0,0016 (0,0008)
7	0,0006 (0,0005)	0,0007 (0,0005)	0,0007 (0,0003)
Média	0,0008 (0,0003)	0,0011 (0,0005)	0,0011 (0,0004)

Assim como nos experimentos realizados com o conjunto de dados AutoMpg, os experimentos que resultaram nas Tabelas 5.11 e 5.12 permitiram a avaliação do uso de diferentes métodos de defuzzificação, a influência da quantidade de conjuntos fuzzy para os valores de saída e da quantidade de grupos.

As Tabelas 5.13 e 5.14 a seguir apresentam os mesmos testes feitos anteriormente, porém analisando o método Fuzzy-CCM com 1, 2 e 3 grupos comparando com o WM. As variáveis utilizadas na definição de contexto indicadas na primeira coluna, foram também consideradas para o método WM.

Com estes resultados é possível fazer diferentes análises:

1. Influência da quantidade de grupos: Os testes realizados com este conjunto de dados apresentaram melhor desempenho quando utilizado um grupo, o que valida a proposta de melhoria na interpretabilidade usando o método Fuzzy-CCM.
2. Influência da quantidade de conjuntos fuzzy: O uso de 5 conjuntos fuzzy apresentou os melhores resultados do Fuzzy-CCM em relação ao WM, assim como o conjunto de dados anterior. Com isto podemos afirmar que a granularização não deve ser nem tão pequena, a ponto de perder alguma informação importante, nem tão grande que aumente demais a informação a cerca dos dados.
3. Influência do método de defuzzificação: Ao contrário do conjunto de dados anterior, este apresentou melhores resultados utilizando o método MEM. É importante observar que os resultados obtidos utilizando este método de defuzzificação tanto para

**Tabela 5.13:** Pollution: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e MEM para defuzificação

Variáveis Explícitas	WM	1 Grupo	2 Grupos	3 Grupos
$V_1 - 3$	0,024 (0,013)	0,02 (0,009)	0,029 (0,013)	0,057 (0,02)
$V_1 - 5$	0,021 (0,014)	0,019 (0,012)	0,03 (0,015)	0,06 (0,02)
$V_1 - 7$	0,025 (0,022)	0,02 (0,009)	0,025 (0,017)	0,051 (0,024)
Média	0,023 (0,0019)	0,02 (0,0008)	0,028 (0,0024)	0,056 (0,0046)
$V_1V_2 - 3$	0,052 (0,009)	0,047 (0,008)	0,061 (0,011)	0,099 (0,015)
$V_1V_2 - 5$	0,051 (0,011)	0,051 (0,01)	0,05 (0,014)	0,095 (0,017)
$V_1V_2 - 7$	0,065 (0,014)	0,047 (0,008)	0,056 (0,014)	0,082 (0,02)
Média	0,056 (0,0077)	0,049 (0,0022)	0,056 (0,0056)	0,092 (0,0086)
$V_1V_5 - 3$	0,04 (0,013)	0,037 (0,006)	0,05 (0,011)	0,078 (0,009)
$V_1V_5 - 5$	0,046 (0,018)	0,037 (0,008)	0,051 (0,011)	0,084 (0,017)
$V_1V_5 - 7$	0,049 (0,012)	0,037 (0,006)	0,047 (0,012)	0,077 (0,017)
Média	0,045 (0,0047)	0,037 (0,0002)	0,049 (0,002)	0,08 (0,0041)
$V_1V_9 - 3$	0,043 (0,01)	0,042 (0,009)	0,052 (0,015)	0,099 (0,012)
$V_1V_9 - 5$	0,041 (0,009)	0,039 (0,009)	0,043 (0,009)	0,065 (0,019)
$V_1V_9 - 7$	0,051 (0,015)	0,042 (0,009)	0,05 (0,013)	0,072 (0,012)
Média	0,045 (0,0051)	0,041 (0,0018)	0,048 (0,005)	0,079 (0,0179)
$V_2 - 3$	0,041 (0,012)	0,041 (0,012)	0,056 (0,014)	0,08 (0,016)
$V_2 - 5$	0,039 (0,014)	0,039 (0,012)	0,047 (0,013)	0,072 (0,016)
$V_2 - 7$	0,057 (0,019)	0,041 (0,012)	0,05 (0,02)	0,066 (0,024)
Média	0,046 (0,0096)	0,041 (0,0016)	0,051 (0,0041)	0,073 (0,0066)
$V_2V_5 - 3$	0,055 (0,007)	0,047 (0,012)	0,063 (0,012)	0,091 (0,016)
$V_2V_5 - 5$	0,054 (0,015)	0,046 (0,01)	0,067 (0,016)	0,103 (0,015)
$V_2V_5 - 7$	0,068 (0,017)	0,047 (0,012)	0,061 (0,017)	0,081 (0,02)
Média	0,059 (0,0074)	0,047 (0,0009)	0,064 (0,0033)	0,092 (0,0115)
$V_2V_9 - 3$	0,052 (0,011)	0,053 (0,01)	0,071 (0,009)	0,083 (0,015)
$V_2V_9 - 5$	0,049 (0,015)	0,049 (0,01)	0,059 (0,009)	0,066 (0,018)
$V_2V_9 - 7$	0,076 (0,022)	0,053 (0,01)	0,066 (0,018)	0,085 (0,018)
Média	0,059 (0,0151)	0,052 (0,0023)	0,065 (0,0057)	0,078 (0,0105)
$V_5 - 3$	0,029 (0,008)	0,028 (0,006)	0,037 (0,007)	0,074 (0,014)
$V_5 - 5$	0,027 (0,017)	0,026 (0,007)	0,041 (0,012)	0,073 (0,01)
$V_5 - 7$	0,036 (0,009)	0,028 (0,006)	0,034 (0,012)	0,052 (0,013)
Média	0,031 (0,0051)	0,027 (0,0008)	0,037 (0,0038)	0,067 (0,0122)
$V_5V_9 - 3$	0,043 (0,005)	0,047 (0,005)	0,058 (0,01)	0,086 (0,009)
$V_5V_9 - 5$	0,037 (0,015)	0,04 (0,01)	0,052 (0,014)	0,055 (0,022)
$V_5V_9 - 7$	0,054 (0,016)	0,047 (0,005)	0,054 (0,022)	0,072 (0,018)
Média	0,045 (0,0087)	0,044 (0,0037)	0,055 (0,0028)	0,071 (0,0157)
$V_9 - 3$	0,028 (0,006)	0,03 (0,005)	0,046 (0,009)	0,061 (0,009)
$V_9 - 5$	0,024 (0,015)	0,028 (0,01)	0,038 (0,013)	0,049 (0,017)
$V_9 - 7$	0,034 (0,014)	0,03 (0,005)	0,043 (0,013)	0,058 (0,015)
Média	0,029 (0,0046)	0,029 (0,001)	0,042 (0,0037)	0,056 (0,0059)
MédiaGeral	0,044 (0,013)	0,039 (0,01)	0,05 (0,011)	0,074 (0,013)

Fuzzy-CCM quanto para WM apresentaram resultados bastante similares.

**Tabela 5.14:** Pollution: Taxas de Erro médias e desvio padrão usando Fuzzy-CCM para geração das regras e CMP para defuzificação

Variáveis Explícitas	WM	1 Grupo	2 Grupos	3 Grupos
$V_1 - 3$	0,0011 (0,0010)	0,0011 (0,0004)	0,0010 (0,0008)	0,0010 (0,0009)
$V_1 - 5$	0,0013 (0,0009)	0,0013 (0,0005)	0,0014 (0,0005)	0,0013 (0,0005)
$V_1 - 7$	0,0013 (0,0007)	0,0011 (0,0004)	0,0006 (0,0005)	0,0006 (0,0004)
Média	0,0012 (0,0001)	0,0012 (0,0001)	0,0010 (0,0004)	0,0010 (0,0003)
$V_1V_2 - 3$	0,0011 (0,0011)	0,0013 (0,0007)	0,0011 (0,0009)	0,0012 (0,0009)
$V_1V_2 - 5$	0,0016 (0,0019)	0,0013 (0,0008)	0,0014 (0,0006)	0,0014 (0,0007)
$V_1V_2 - 7$	0,0014 (0,0006)	0,0013 (0,0007)	0,0006 (0,0005)	0,0006 (0,0005)
Média	0,0014 (0,0003)	0,0013 (0,0000)	0,0010 (0,0004)	0,0010 (0,0004)
$V_1V_5 - 3$	0,0012 (0,0011)	0,0012 (0,0006)	0,0009 (0,0006)	0,0009 (0,0005)
$V_1V_5 - 5$	0,0019 (0,0019)	0,0012 (0,0004)	0,0014 (0,0005)	0,0012 (0,0005)
$V_1V_5 - 7$	0,0012 (0,0007)	0,0012 (0,0006)	0,0006 (0,0005)	0,0007 (0,0005)
Média	0,0014 (0,0004)	0,0012 (0,0000)	0,0010 (0,0004)	0,0009 (0,0003)
$V_1V_9 - 3$	0,0012 (0,0012)	0,0012 (0,0008)	0,0014 (0,0010)	0,0013 (0,0010)
$V_1V_9 - 5$	0,0013 (0,0009)	0,0010 (0,0006)	0,0012 (0,0005)	0,0012 (0,0004)
$V_1V_9 - 7$	0,0015 (0,0009)	0,0012 (0,0008)	0,0009 (0,0004)	0,0009 (0,0004)
Média	0,0013 (0,0002)	0,0011 (0,0001)	0,0011 (0,0003)	0,0011 (0,0002)
$V_2 - 3$	0,0012 (0,0010)	0,0015 (0,0006)	0,0013 (0,0005)	0,0014 (0,0006)
$V_2 - 5$	0,0014 (0,0006)	0,0014 (0,0006)	0,0017 (0,0008)	0,0017 (0,0008)
$V_2 - 7$	0,0020 (0,0005)	0,0015 (0,0006)	0,0007 (0,0006)	0,0008 (0,0005)
Média	0,0015 (0,0004)	0,0015 (0,0000)	0,0012 (0,0005)	0,0013 (0,0004)
$V_2V_5 - 3$	0,0011 (0,0010)	0,0015 (0,0009)	0,0013 (0,0008)	0,0013 (0,0006)
$V_2V_5 - 5$	0,0014 (0,0007)	0,0014 (0,0006)	0,0016 (0,0007)	0,0016 (0,0006)
$V_2V_5 - 7$	0,0013 (0,0008)	0,0015 (0,0009)	0,0007 (0,0005)	0,0007 (0,0005)
Média	0,0013 (0,0001)	0,0014 (0,0001)	0,0012 (0,0005)	0,0012 (0,0004)
$V_2V_9 - 3$	0,0011 (0,0011)	0,0012 (0,0008)	0,0012 (0,0007)	0,0014 (0,0007)
$V_2V_9 - 5$	0,0011 (0,0004)	0,0009 (0,0005)	0,0013 (0,0005)	0,0013 (0,0006)
$V_2V_9 - 7$	0,0013 (0,0006)	0,0012 (0,0008)	0,0009 (0,0005)	0,0010 (0,0004)
Média	0,0011 (0,0002)	0,0011 (0,0002)	0,0011 (0,0002)	0,0012 (0,0002)
$V_5 - 3$	0,0011 (0,0010)	0,0015 (0,0009)	0,0012 (0,0007)	0,0012 (0,0008)
$V_5 - 5$	0,0016 (0,0009)	0,0015 (0,0006)	0,0016 (0,0006)	0,0016 (0,0005)
$V_5 - 7$	0,0013 (0,0008)	0,0015 (0,0009)	0,0006 (0,0005)	0,0008 (0,0005)
Média	0,0013 (0,0002)	0,0015 (0,0000)	0,0011 (0,0005)	0,0012 (0,0004)
$V_5V_9 - 3$	0,0011 (0,0011)	0,0012 (0,0009)	0,0014 (0,0007)	0,0015 (0,0008)
$V_5V_9 - 5$	0,0011 (0,0005)	0,0011 (0,0005)	0,0013 (0,0005)	0,0013 (0,0004)
$V_5V_9 - 7$	0,0012 (0,0007)	0,0012 (0,0009)	0,0009 (0,0005)	0,0010 (0,0005)
Média	0,0011 (0,0001)	0,0012 (0,0001)	0,0012 (0,0003)	0,0013 (0,0003)
$V_9 - 3$	0,0011 (0,0010)	0,0012 (0,0008)	0,0012 (0,0008)	0,0015 (0,0009)
$V_9 - 5$	0,0012 (0,0004)	0,0010 (0,0005)	0,0013 (0,0005)	0,0013 (0,0004)
$V_9 - 7$	0,0016 (0,0007)	0,0012 (0,0008)	0,0010 (0,0004)	0,0009 (0,0005)
Média	0,0013 (0,0003)	0,0011 (0,0001)	0,0012 (0,0002)	0,0013 (0,0003)
MédiaGeral	0,0013 (0,0001)	0,0013 (0,0002)	0,0011 (0,0001)	0,0011 (0,0001)

### 5.3 Considerações Finais

Este capítulo apresentou dois diferentes tipos de experimentos que permitiram a análise detalhada da viabilidade do método Fuzzy-CCM: concepção da abordagem e ampliação da abordagem.

Com estes experimentos, além do ganho ou manutenção de acuidade em muitos testes, foi possível observar benefícios do método Fuzzy-CCM relacionados à interpretabilidade em dois diferentes aspectos:

1. Quando comparado ao método Wang&Mendel, as regras geradas pelo método Fuzzy-CCM têm muito menos variáveis no antecedente. Aliado a isto, quanto mais valores de entrada são agrupados, mais a dimensionalidade do problema é reduzida.
2. Com relação ao agrupamento fuzzy tradicional, o ganho na interpretabilidade é óbvio com a inclusão de um ou mais valores lingüísticos para domínios de entrada.

É importante notar que, neste trabalho, não foi abordada a questão do número de regras geradas pelo método, o que é também um dos parâmetros decisivos para avaliar a interpretabilidade da base de regras. Considerando que os contextos são definidos pela combinação de todos os valores lingüísticos das variáveis selecionadas, o número de regras cresce proporcionalmente ao número de variáveis e valores de contexto utilizados. A possibilidade de redução do número de regras geradas, por meio de seleção, será investigada futuramente.

No próximo capítulo será apresentada a conclusão do trabalho desenvolvido bem como sugestão de trabalhos futuros que possibilitem melhorias ao método Fuzzy-CCM proposto.

---

## Conclusão e Trabalhos Futuros

---

Este trabalho propõe um novo método para Modelagem Fuzzy usando o algoritmo de agrupamento Fuzzy C-Means Condicional, chamado Fuzzy-CCM (Fuzzy Conditional Clustering based Modeling), para obter regras compactas selecionando poucas variáveis para aparecerem explicitamente no antecedente da regra e agrupando os demais valores. Os experimentos mostraram que este é um método promissor para tratar problemas de interpretabilidade e dimensionalidade.

Para validar a proposta e obter conclusões que justifiquem o uso do método Fuzzy-CCM, inicialmente foi apresentada a idéia, para em seguida serem apresentados os capítulos de fundamentação teórica: Sistemas Fuzzy e Algoritmos de Agrupamento. No Capítulo 2, Sistemas Fuzzy, foram abordados os conceitos básicos da teoria de conjuntos fuzzy, raciocínio fuzzy e sistemas fuzzy, enquanto no Capítulo 3, Algoritmos de Agrupamento, foram apresentados os algoritmos K-Means, Fuzzy C-Means e Fuzzy C-Means Condicional, os quais foram usados para a implementação da abordagem utilizada neste trabalho.

Após a fundamentação teórica, foi apresentado o Capítulo 4, Modelagem Fuzzy usando Agrupamento Condicional, que destacou vários aspectos e abordagens que favorecem o desempenho de modelos fuzzy, bem como foram comentadas algumas abordagens exis-

tentes na literatura que utilizam agrupamento na modelagem fuzzy para que fosse possível compreender a relevância da proposta. Neste capítulo também foi detalhado o processo de construção de regras e inferência pelo método Fuzzy-CCM.

Por fim, o Capítulo 5, Resultados Experimentais, apresentou os experimentos realizados para validação da proposta, com os quais foi possível avaliar aspectos relacionados ao número de grupos, o número de conjuntos fuzzy, quantidade de atributos relevantes e à escolha do método de defuzificação para conjunto de dados com valores de saída contínuos.

Por meio dos experimentos, além do ganho ou manutenção de acuidade em muitos testes, foi apresentado no capítulo anterior benefícios relacionados a interpretabilidade em dois diferentes aspectos:

1. Quando comparado ao método de Wang&Mendel geral, as regras criadas pelo método Fuzzy-CCM têm um número bem menor de variáveis no antecedente. Aliado a isto, quanto mais valores de entrada são agrupados, mais a dimensionalidade do problema é reduzida.
2. Com relação ao agrupamento fuzzy tradicional, o ganho na interpretabilidade é óbvio com a inclusão de um valor lingüístico para algum domínio de entrada.

Apesar de muitas investigações acerca da proposta já terem sido feitas, algumas considerações são relevantes e podem ser abordadas em trabalhos futuros:

1. Utilizar técnicas para otimização da base de regras, eliminando regras redundantes e conflitantes. Para tanto, podem ser utilizados os Algoritmos Genéticos;
2. Fazer um comitê com mais de um algoritmo de seleção de atributos, para garantir a qualidade dos atributos escolhidos como relevantes;
3. Escolher mais atributos/variáveis e combiná-los, já que este trabalho apresentou apenas a combinação 2 a 2;
4. Além da comparação com o método de Wang&Mendel, comparar os resultados com outras abordagens.

---

# Saída ReliefF

---

Este apêndice apresenta os resultados da execução do algoritmo Relief-F (KONONENKO, 1994) disponibilizada na ferramenta WEKA (WITTEN;FRANK, 2000) para todos os conjunto de dados usados nesse trabalho de pesquisa, cujas características estão citadas no capítulo 4.

## A.1 Iris

=== Run information ===

Evaluator: weka.attributeSelection.ReliefFAttributeEval -M -1 -D 1 -K 10 -A 2

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: iris

Instances: 150

Attributes: 5 sepalwidth, sepalwidth, petalwidth, petalwidth, class

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method: Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 class):

ReliefF Ranking Filter

Instances sampled: all

Number of nearest neighbours (k): 10

Equal influence nearest neighbours

Ranked attributes:

0.376 4 petalwidth

0.359 3 petallength

0.14 1 sepallength

0.122 2 sepalwidth

Selected attributes: 4,3,1,2 : 4

## A.2 Bupa

=== Run information ===

Evaluator: weka.attributeSelection.ReliefFAttributeEval -M -1 -D 1 -K 10 -A 2

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: iris

Instances: 345

Attributes: 7 mvc, alkphos, sgpt, agot, gammagt, drinks, class,

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method: Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 7 class):

ReliefF Ranking Filter

Instances sampled: all

Number of nearest neighbours (k): 10

Equal influence nearest neighbours

Ranked attributes:

0.005446 3 sgpt

0.004876 5 gammagt

0.004855 6 drinks  
0.003343 4 agot  
0.002838 1 mvc  
-0.00062 2 alkphos  
Selected attributes: 3,5,6,4,1,2 : 6

### A.3 Diabetes

=== Run information ===

Evaluator: weka.attributeSelection.ReliefFAttributeEval -M -1 -D 1 -K 10 -A 2

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: pima\_diabetes

Instances: 768

Attributes: 9 preg, plas, pres, skin, insu, mass, pedi, age, class

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method: Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 9 class):

ReliefF Ranking Filter

Instances sampled: all

Number of nearest neighbours (k): 10

Equal influence nearest neighbours

Ranked attributes:

0.02753 2 plas

0.01554 6 mass

0.01251 4 skin

0.01156 1 preg

0.0099 8 age

0.00796 7 pedi

0.00563 3 pres

0.00426 5 insu

Selected attributes: 2,6,4,1,8,7,3,5 : 8

## A.4 Wine

=== Run information ===

Evaluator: weka.attributeSelection.ReliefFAttributeEval -M -1 -D 1 -K 10 -A 2

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: wine

Instances: 108

Attributes: 14, um, dois, tres, quatro, cinco, seis, sete, oito, nove, dez, onze, doze, treze,

class

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method: Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 14 class):

ReliefF Ranking Filter

Instances sampled: all

Number of nearest neighbours (k): 10

Equal influence nearest neighbours

Ranked attributes:

0.2149 12 doze

0.2017 7 sete

0.1788 13 treze

0.1612 10 dez

0.1502 1 um

0.132 11 onze

0.109 6 seis

0.0775 8 oito

0.0676 2 dois

0.065 9 nove

0.0582 4 quatro

0.0341 5 cinco

0.0287 3 tres

Selected attributes: 12,7,13,10,1,11,6,8,2,9,4,5,3 : 13

## A.5 Vehicle

=== Run information ===

Evaluator: weka.attributeSelection.ReliefFAttributeEval -M -1 -D 1 -K 10 -A 2

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: wine

Instances: 108

Attributes: 14 um, dois, tres, quatro, cinco, seis, sete, oito, nove, dez, onze, doze, treze,  
class

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method: Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 14 class):

ReliefF Ranking Filter

Instances sampled: all

Number of nearest neighbours (k): 10

Equal influence nearest neighbours

Ranked attributes:

0.2149 12 doze

0.2017 7 sete

0.1788 13 treze

0.1612 10 dez

0.1502 1 um

0.132 11 onze

0.109 6 seis

0.0775 8 oito

0.0676 2 dois

0.065 9 nove

0.0582 4 quatro

0.0341 5 cinco

0.0287 3 tres

Selected attributes: 12,7,13,10,1,11,6,8,2,9,4,5,3 : 13

## A.6 AutoMpg

=== Run information ===

Evaluator: weka.attributeSelection.ReliefFAttributeEval -M -1 -D 1 -K 10 -A 2

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: autoMpg

Instances: 398

Attributes: 8 cylinders, displacement, horsepower, weight, acceleration, model, origin,  
class

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method: Attribute ranking.

Attribute Evaluator (supervised, Class (numeric): 8 class):

ReliefF Ranking Filter

Instances sampled: all

Number of nearest neighbours (k): 10

Equal influence nearest neighbours

Ranked attributes:

0.05428 6 model

0.01723 5 acceleration

0.01415 1 cylinders

0.01383 4 weight

0.01158 3 horsepower

0.00434 2 displacement

-0.00454 7 origin

Selected attributes: 6,5,1,4,3,2,7 : 7

## A.7 Pollution

=== Run information ===

Evaluator: weka.attributeSelection.ReliefFAttributeEval -M -1 -D 1 -K 10 -A 2

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: pollution

Instances: 60

Attributes: 16, PREC, JANT, JULT, OVR65, POPN, EDUC, HOUS, DENS, NONW, WW-DRK, POOR, HC, NOX, SO@, HUMID, MORT

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method: Attribute ranking.

Attribute Evaluator (supervised, Class (numeric): 16 MORT):

ReliefF Ranking Filter Instances sampled: all

Number of nearest neighbours (k): 10

Equal influence nearest neighbours

Ranked attributes:

0.02667 1 PREC

0.01885 9 NONW

0.01622 2 JANT

0.01199 5 POPN

0.01147 6 EDUC

0.00759 3 JULT

0.00662 15 HUMID

0.00559 12 HC

0.00487 13 NOX

0.00378 8 DENS

0.00283 7 HOUS

0.00176 4 OVR65

0.00106 11 POOR

-0.00231 14 SO@

-0.00716 10 WWDRK

Selected attributes: 1,9,2,5,6,3,15,12,13,8,7,4,11,14,10 : 15

---

## Algoritmo Wang&Mendel

---

O método Wang&Mendel para geração de bases de regras fuzzy é bastante conhecido e utilizado devido ao fato de ser facilmente interpretado e executado, gerando bases de regras com bom poder de classificação e relativamente compactas, por não terem regras conflitantes ou redundantes.

Assumindo-se que a partição fuzzy para o domínio abordado é conhecida, a geração da base de regras fuzzy é realizada por meio dos seguintes passos (WANG;MENDEL, 1992a)(WANG, 2003):

1. Para cada exemplo  $E_h = (a_1, \dots, a_n, y)$ ,  $h = 1, 2, \dots, p$ , contendo  $n$  valores de atributos de entrada e valor para a classe de saída  $y$ , construir uma regra fuzzy que tenha o maior grau de compatibilidade com o exemplo. Isto é feito atribuindo, para cada valor de entrada de cada atributo do exemplo, o conjunto fuzzy em que ele tem maior grau de pertinência. Caso um determinado atributo tenha o mesmo valor em dois conjuntos fuzzy pode-se selecionar aleatoriamente apenas um dos conjuntos para criação da regra. A partir disto são geradas  $p$  regras.
2. Dar um grau de importância a cada uma das regras geradas. Seja:

$$R_1 : \text{Se } X_1 \text{ é } A_1 \text{ e } \dots \text{ e } X_n \text{ é } A_n, \text{ Então } Y \text{ é } B$$

a regra gerada pelo exemplo  $E_1$ . Logo, o grau de importância dessa regra será dado por:

$$G(R_k) = t(A(a_1), A(a_2), \dots, A(a_n), b(y))$$

na qual  $t$  representa uma  $t$ -norma,  $A(a_x)$  representa o grau de pertinência do valor  $a_i$  do exemplo  $E_k$  no conjunto fuzzy da variável  $i$  e  $b(y)$  representa o grau de pertinência do valor de saída do exemplo em questão no conjunto fuzzy de saída  $B$ .

3. Eliminar as regras redundantes, ou seja, regras iguais do conjunto de  $p$  regras geradas;
4. Se duas ou mais regras, dentre as regras restantes, possuírem antecedentes iguais, descartar as regras com menor grau de importância e manter somente a regra com maior grau de importância;
5. Compor a base de regras final adicionando as regras selecionadas.

# Referências Bibliográficas

---

---

- ABONYI, J.; BABUSKA, R.; SZEIFERT, F. Modified gath-geva fuzzy clustering for identification of takagi-sugeno fuzzy models. **IEEE Transactions on Systems, Man and Cybernetics-Part B**. v. 32, p. 612–621, 2002.
- ALEXIEV, K. M.; GEORGIEVA, O. I. Improved fuzzy clustering for identification of takagi-sugeno model. In: IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS, 2004. **Proceedings...**, 2004. p. 213–218.
- ALTROCK, C. V. **Fuzzy Logic and Neuro Fuzzy Applications Explained**. Prentice Hall, 1995., 368 p.
- BEZDEK, J. C. **Pattern Recognition with Fuzzy Objective Function Algorithms**. Norwell, MA, USA: Kluwer Academic Publishers, 1981., 256 p.
- BISHOP, C. M. **Neural Networks for Pattern Recognition**. New York, NY, USA: Oxford University Press, Inc., 1995., 482 p.
- BLAKE, C. et al. **UCI Repository of Machine Learning Databases**. University of California, Irvine, Dept. of Information and Computer Sciences.
- CASILLAS, J.; CORDON, O.; HERRERA, F. **Can linguistic modeling be as accurate as fuzzy modeling without losing its description to a high degree?**, 2000. Relatório técnico.

- CASILLAS, J.; CORDON, O.; HERRERA, F.; MAGDALENA, L. **Interpretability Issues in Fuzzy Modeling**. Springer, 2003., 643 p.
- CASTELLANO, G.; FANELLI, A. M.; GENTILE, E.; ROSELI, T. A GA-based approach to optimize fuzzy models learned from data. In: GECCO PROGRAM, 2002. **Proceedings...**, 2002. p. 5–8.
- CASTRO, P. A.; CAMARGO, H. Focusing on interpretability and accuracy of genetic fuzzy systems. In: IEEE INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS, 14, 2005. **Proceedings...**, 2005. p. 696–701.
- CHI, Z.; YAN, H.; PHAM, T. **Fuzzy Algorithms with Applications to Image Processing and Pattern Recognition**. World Scientific, 1996., v. 10, 225 p.
- CORDÓN, O.; GOMIDE, F. A. C.; HERRERA, F.; HOFFMAN, F.; MAGDALENA, L. Special issue on genetic fuzzy systems. **Fuzzy Sets and Systems**. v. 141, 2004.
- CORDÓN, O.; JESUS, M. J.; HERRERA, F. **A proposal on Reasoning Methods in Fuzzy Rule-Based Classification Systems**. Granada-Spain: E.T.S. de Ingeniería Informática, Universidad de Granada, Novembro, 1997. Relatório Técnico DECSAI-970126.
- DUDA, R. O.; HART, P. E. **Pattern classification and Scene Analysis**. John Wiley, 1973., 482 p.
- DUMITRESCU, D.; LAZZERINI, B.; JAIR, L. **Fuzzy Sets and Their Application to Clustering and Training**. International Series on Computational Intelligence. CBC Press, 2000., 664 p.
- FORGY, E. Cluster analysis of multivariate data: Efficiency vs. interpretability of classifications. In: BIOMETRICS SOCIETY MEETING, 3, 1965. **Proceedings...** Riverside-CA, 1965. p. 768.
- GATH, I.; GEVA, B. Unsupervised optimal fuzzy clustering. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. v. 7, p. 773–781, 1989.
- GOMEZ-SKARMETA, A. F.; DELGADO, M.; VILA, M. A. About the use of fuzzy clustering techniques for fuzzy model identification. **Fuzzy Sets and Systems**. v. 106, n. 2, p. 179–188, 1999.

- GONZÁLEZ, A.; PÉREZ, R. Slave: A genetic learning system based on an iterative approach. **IEEE Transactions on Fuzzy Systems**. v. 7, n. 2, p. 176–191, 1999.
- GORDON, A. D. **Classification**, 2. ed. CRC Press, 1999., 256 p.
- GUILLAUME, S. Designing fuzzy inference systems from data: An interpretability-oriented review. **IEEE Transactions on Fuzzy Systems**. v. 9, n. 3, p. 426–443, 2001.
- GUZTAFSON, E. E.; KESSEL, W. C. Fuzzy clustering with a fuzzy covariance matrix. In: IEEE CONFERENCE ON DECISION AND CONTROL, 1979. **Proceedings...**, 1979. p. 503–516.
- HADJULI, M. L.; WERTZ, V. Takagi-sugeno fuzzy modeling incorporating input variables selection. **IEEE Transactions on Fuzzy Systems**. v. 10, n. 6, p. 728–742, 2002.
- HARTIGAN, J. A. **Clustering Algorithms**. New York: John Wiley and Sons, 1975., 351 p.
- HIROTA, K.; PEDRYCZ, W. D-fuzzy clustering. **Pattern Recognition Letters**. v. 16, n. 2, p. 193–200, 1995.
- HIROTA, K.; PEDRYCZ, W. Directional fuzzy clustering and its application to fuzzy modeling. **Fuzzy Sets and Systems**. v. 3, p. 315–326, 1996.
- ISHIBUCHI, H.; NAKASHIMA, T.; MURATA, T. Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. **IEEE Transaction On Fuzzy Systems, Man And Cybernetics-Part B: Cybernetics**. v. 29, n. 5, p. 601–618, 1999a.
- ISHIBUCHI, H.; NAKASHIMA, T.; MURATA, T. Voting fuzzy rule-based systems for pattern classification problems. **Fuzzy Sets and Systems**. v. 103, p. 223–238, 1999b.
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **ACM Computing Surveys**, New York, USA. v. 31, n. 3, p. 264–323, 1999.
- JANG, S. R.; SUN, C. T.; MIZUTANI, E. **Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence**. Prentice Hall, 1997., 614 p.

- JIN, Y. Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement. **IEEE Transactions on Fuzzy Systems**. v. 8, n. 2, p. 212–221, 2000.
- KAUFMAN, L.; ROUSSEEUW, P. J. **Finding Groups in Data: An Introduction to Cluster Analysis**. Wiley-Interscience, 1990., 368 p.
- KECMAN, V. **Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models**. MIT Press, 2001., 608 p.
- KENDEL, A.; FRIEDMAN, M. Defuzzification using most typical values. **IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics**. v. 28, n. 6, p. 901–906, 1998.
- KIM, E.; PARK, M.; Ji, S. A new approach to fuzzy modeling. **IEEE Transactions on Fuzzy Systems**. v. 5, n. 3, p. 328–337, 1997.
- KIRA, K.; RENDELL, L. A. A practical approach to feature selection. In: INTERNATIONAL WORKSHOP ON MACHINE LEARNING, 1992. **Proceedings...**, 1992. p. 249–256.
- KLIR, G. J.; YUAN, B. **Fuzzy Sets and Fuzzy Logic: theory and applications**, 1. ed. Prentice-Hall, 1995., 574 p.
- KONONENKO, I. Estimating attributes: Analysis and extensions of RELIEF. In: EUROPEAN CONFERENCE ON MACHINE LEARNING, 1994. **Proceedings...**, 1994. p. 171–182.
- KONONENKO, I.; R-SIKONIA, M. **Relieff for estimation and discretization of attributes in classification**.
- L. RONDEAU, R. RUELAS, L. L.; LAMOTTE, M. A defuzzification method respecting the fuzzification. **Fuzzy Sets and Systems**. v. 86, p. 311–320, 1997.
- LESKI, J. M. Generalized weighted conditional fuzzy clustering. **IEEE Transactions On Fuzzy Systems**. v. 11, n. 6, p. 709–715, 2003.
- LIAO, T. W.; CELMINS, A. K.; HAMMELL, R. J. A fuzzy C-Means variant for the generation of fuzzy term sets. **Fuzzy Sets and Fuzzy Systems**. v. 135, n. 2, p. 241–257, 1997.

- LUGER, G. **Artificial intelligence: Structures and strategies for complex problem solving**, 5. ed. Addison Wesley Longman, 2004., 928 p.
- MAMDANI, E. H. Application of fuzzy logic to approximate reasoning using linguistic synthesis. **Transactions on Computers**. v. 26, n. 12, p. 1182–1191, 1977.
- MITCHELL, T. **Machine Learning**, 1. ed. McGraw-Hill Education, 1997., 352 p.
- MONTES, J. C.; LLORCA, R. M.; GRAU, J. P. Building interpretable fuzzy systems: a new approach to fuzzy modeling. In: ELETRONICS, ROBOTICS AND AUTOMOTIVE MECHANICS CONFERENCE, 2006. **Proceedings...**, 2006.
- MOORE, D. S.; MCCABE, G. P. **Introduction to the Practice of Statistics**, 5. ed. W. H. Freeman, 1999., 828 p.
- NICOLETTI, M. C.; CAMARGO, H. A. **Fundamentos da Teoria de Conjuntos Fuzzy**. São Carlos: EdUFSCar, 2004., 65 p. (Série Apontamentos).
- PEDRYCZ, W. Conditional fuzzy c-means. **Pattern Recognition Letter**, New York, USA. v. 17, n. 6, p. 625–631, 1996a.
- PEDRYCZ, W. **Fuzzy Modelling: Paradigms and Practice**. Kluwer Academic Press, 1996b., 416 p.
- PEDRYCZ, W. **Computational Intelligence: An Introduction**. Boca Raton, FL., 1998., 284 p.
- PEDRYCZ, W.; GOMIDE, F. **An introduction to fuzzy sets: Analysis and design**. MIT Press, 1998., 465 p.
- PEDRYCZ, W.; KWAK, K. Boosting of granular models. **Fuzzy Sets and Systems**. v. 157, p. 2934–2953, 2006a.
- PEDRYCZ, W.; KWAK, K. Linguistic models as a framework of user-centric system modeling. **IEEE Transactions on Systems, Man, and Cybernetics-Part A**. v. 36, p. 727–745, 2006b.

- PEDRYCZ, W.; VASILAKOS, A. V. Linguistic models and linguistic modeling. **IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics**. v. 29, n. 6, p. 745–757, 1999.
- PEDRYCZ, W.; VUKOCIVH, G. Fuzzy clustering with supervision. **Pattern Recognition**. v. 37, p. 1339–1349, 2004.
- QUINLAN, J. R. Induction of decision trees. **Machine Learning**. v. 1, p. 81–106, 1986.
- ROVENTA, E.; SPIRCU, T. Averaging procedures in defuzzification processes. **Fuzzy Sets and Systems**. v. 136, p. 375–385, 2003.
- RUNKLER, T. A. Selection of appropriate defuzzification methods using application specific properties. **IEEE Transactions on Fuzzy Systems**. v. 5, n. 1, p. 72–79, 1997.
- SETNES, M. Supervised fuzzy clustering for rule extraction. **IEEE Transactions on Fuzzy Systems**. v. 8, n. 4, p. 416–424, 1999.
- SETNES, M.; BABUSKA, R.; VERBRUGGEN, H. B. Rule-based modeling: Precision and transparency. **IEEE Transactions on System, Man, and Cybernetics - Part C: Applications and Review**. v. 28, n. 1, p. 165–169, 1998.
- SETNES, M.; ROUBOS, H. Transparent fuzzy modeling using fuzzy clustering and GA's. In: IEEE INTERNATIONAL CONFERENCE ON THE NORTH AMERICAN, 1999. **Proceedings...**, 1999. p. 198–202.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its application to modeling and control. **IEEE Transactions on Systems, Man, and Cybernetics**. v. 15, n. 1, p. 116–132, 1985.
- TALON, A. F.; CAMARGO, H. D. A. Co-evolution: An approach to automatic generation of fuzzy systems. In: INTERNATIONAL CONFERENCE ON HYBRID INTELLIGENT SYSTEMS, 6, 2006. **Proceedings...**, 2006. p. 35–35.
- TIMM, H. Fuzzy cluster analysis of classified data. In: 9° IFSA WORLD CONGRESS E 20° NAFIPS CONFERENCE, 2001. **Proceedings...**, 2001. p. 1431–1436.
- WANG, L. The WM method completed: a flexible fuzzy system approach to data mining. **IEEE Transactions on Fuzzy Systems**. v. 11, p. 768–782, 2003.

- WANG, L.; MENDEL, J. Generating fuzzy rules by learning from examples. **IEEE Transaction on Fuzzy Systems, Man and Cybernetics**. v. 22, n. 6, p. 1414–1427, 1992a.
- WANG, L. X.; MENDEL, J. M. Fuzzy basis functions, universal approximation, and orthogonal least-squares learning. **IEEE Transactions on Neural Networks**. v. 3, p. 807–813, 1992b.
- WITTEN, M.; FRANK, E. **Data Mining, Practical Machine Learning Tools and Techniques with Java Implementations**. Morgan Kaufmann, 2000. pp. 126.
- ZADEH, L. Fuzzy sets. **Information and Control**. v. 8, n. 3, p. 338–353, 1965.
- ZADEH, L. A., 1971. **Aspects of Network and System Theory**, chapter Toward a theory of fuzzy systems. Rinehart & Winston.
- ZADEH, L. A., 1979. **Advances in Fuzzy Set Theory and Applications**, chapter Fuzzy sets and information granularity, p. 3–18. Elsevier.
- ZHAO, J.; WERTZ, V.; GOREZ, R. A fuzzy clustering method for the identification of fuzzy models for dynamic systems. In: IEEE INTERNATIONAL SYMPOSIUM ON INTELLIGENT CONTROL, 1994. **Proceedings...**, 1994. p. 172–177.
- ZIMMERMANN, H. J. **Fuzzy set theory and its applications**, 2. ed. Boston, USA: Kluwer Academic Publishers, 1991., 399 p.