

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

**Geração Genética de Classificador Fuzzy Intervalar
do Tipo-2**

Adinovam Henriques de Macedo Pimenta

São Carlos
Setembro/2009

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM

CIÊNCIA DA COMPUTAÇÃO

**Geração Genética de Classificador Fuzzy
Intervalar do Tipo-2**

Adinovam Henriques de Macedo Pimenta

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

Orientação: Profa. Dra. Heloísa de Arruda Camargo

São Carlos
Setembro/2009

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

P644gg

Pimenta, Adinovam Henriques de Macedo.

Geração genética de classificador fuzzy intervalar do tipo-2 / Adinovam Henriques de Macedo Pimenta. -- São Carlos : UFSCar, 2010.

91 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2009.

1. Inteligência artificial. 2. Sistemas Fuzzy. 3. Algoritmos genéticos. 4. Sistemas Fuzzy-genético. I. Título.

CDD: 006.3 (20^a)

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

**“Geração Genética de Classificador
Fuzzy Intervalar do Tipo 2”**

ADINOVAM HENRIQUES DE MACEDO PIMENTA

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)



Prof. Dr. Maurício Fernandes Figueiredo
(DC/UFSCar)



Prof. Dr. Fernando Antonio Campos Gomide
(FEEC/UNICAMP)

São Carlos
Outubro/2009

Agradecimentos

Ao meu pai Ernane, à minha mãe Alaíde, à minha irmã Djanane e à minha avó Francisca, pelo apoio, incentivo e cuidado que sempre me deram.

À Profa. Heloisa de Arruda Camargo por sua orientação, pela disposição e paciência.

Ao meu grande amigo Rodolfo Meneguette, pela companhia, pelas risadas e pelo apoio em momentos muito difíceis.

Aos meus queridos irmãos de mestrado Priscilla Lopes, Fábio Justo, Marcos Cintra e Tatiane Nogueira. Todo tipo de ajuda que eu tive de vocês foi muito importante para mim. Muito obrigado!

Aos meus “primos” de mestrado Flávio Montoro, Élen Tomazela e Ageu Duarte pelos momentos agradáveis e amizade durante todo esse tempo em que estivemos juntos. Ageu, obrigado pela ajuda e dicas com os Algoritmos Genéticos.

A todos que contribuíram de forma direta ou indireta para a realização deste trabalho.

Resumo

O objetivo deste trabalho é estudar, expandir e avaliar o uso de conjuntos fuzzy intervalares tipo-2 na representação do conhecimento em sistemas de inferência fuzzy, mais especificamente para os classificadores fuzzy, bem como sua geração automática a partir de conjuntos de dados, por meio de algoritmos genéticos. Esse trabalho investiga o uso de tais conjuntos com enfoque na questão de balanceamento entre o acréscimo de custo da representação e os ganhos em interpretabilidade e precisão, ambos decorrentes da complexidade de representação e processamento dos conjuntos fuzzy intervalares do tipo-2. Com este intuito, foi proposto e implementado um modelo evolutivo composto por três etapas. Na primeira etapa é gerada a base de regras, na segunda é otimizada a base de dados e, por fim, na terceira etapa o número de regras da base gerada é otimizado. O modelo desenvolvido foi avaliado em diversos conjuntos de dados benchmark e os resultados obtidos foram comparados com outros dois classificadores fuzzy, sendo um deles gerados pelo mesmo modelo, porém, utilizando conjuntos fuzzy do tipo-1 e, o outro, gerado pelo método de Wang&Mendel. Métodos estatísticos de comparação usualmente aplicados em contextos semelhantes mostraram aumento significativo na taxa de classificação do classificador fuzzy intervalar do tipo-2 gerado pelo modelo em relação aos outros dois classificadores utilizados para comparação.

Palavras-chave: Sistemas Fuzzy Intervalares do Tipo-2, Algoritmos Genéticos, Algoritmos Genéticos Auto-Adaptativos, Sistemas Fuzzy Genéticos, Geração Automática de Regras Fuzzy.

Abstract

The objective of this work is to study, expand and evaluate the use of interval type-2 fuzzy sets in the knowledge representation for fuzzy inference systems, specifically for fuzzy classifiers, as well as its automatic generation from data sets, by means of genetic algorithms. This work investigates the use of such sets focussing the issue of balance between the cost addition in representation and the gains in interpretability and accuracy, both deriving from the representation and processing complexity of interval type-2 fuzzy sets. With this intent, an evolutionary model composed of three stages was proposed and implemented. In the first stage the rule base is generated, in the second stage the data base is optimized and finally, the number of rules of the rule base obtained is optimized in the third stage. The model developed was evaluated using several benchmark data sets and the results obtained were compared with two other fuzzy classifiers, being one of them generated by the same model using type-1 fuzzy sets and the other one generated by the Wang&Mendel method. Statistical methods usually applied for comparisons in similar contexts demonstrated a significant improvement in the classification rates of the intervalar type-2 fuzzy set classifier generated by the proposed model, with relation to the other methods.

Keywords: Interval Type-2 Fuzzy Set Systems, Genetic Algorithms, Self-Adaptative Genetic Algorithms, Genetic Fuzzy Systems, Automatic Generation of Fuzzy-Rules.

Sumário

RESUMO	III
ABSTRACT	IV
1. INTRODUÇÃO.....	1
1.1 MOTIVAÇÕES E OBJETIVOS	3
1.2 ORGANIZAÇÃO DO TRABALHO	4
2. SISTEMAS FUZZY.....	5
2.1 CONJUNTOS FUZZY	5
2.2 REGRAS FUZZY	14
2.3 SISTEMAS DE INFERÊNCIA FUZZY	17
2.4 SISTEMA DE CLASSIFICAÇÃO FUZZY	19
2.5 CONSIDERAÇÕES FINAIS	22
3. CONJUNTOS FUZZY DO TIPO-2	23
3.1 DEFINIÇÕES	24
3.2 TIPOS DE CONJUNTOS FUZZY DO TIPO-2.....	28
3.3 OPERAÇÕES DE CONJUNTOS FUZZY DO TIPO-2	31
3.4 SISTEMAS DE INFERÊNCIA FUZZY DO TIPO-2	33
3.5 SISTEMAS DE INFERÊNCIA FUZZY INTERVALARES DO TIPO-2 MAMDANI	37
3.6 CONSIDERAÇÕES FINAIS	39
4. ALGORITMOS GENÉTICOS.....	40
4.1 FUNCIONAMENTO DOS ALGORITMOS GENÉTICOS	40
4.2 REPRESENTAÇÃO	41
4.3 SISTEMAS FUZZY GENÉTICOS	47
4.4 CONSIDERAÇÕES FINAIS	51
5. TRABALHOS RELACIONADOS.....	52
5.1 CONSIDERAÇÕES FINAIS	56
6. GERAÇÃO GENÉTICA DE CLASSIFICADOR FUZZY INTERVALAR DO TIPO-2.	57
6.1 ALGORITMO GENÉTICO AUTO-ADAPTATIVO	66
6.2 CONSIDERAÇÕES FINAIS	68
7. EXPERIMENTOS E ANÁLISE DOS RESULTADOS.....	69
7.1 DESCRIÇÃO DOS CONJUNTOS DE DADOS.....	69
7.2 EXPERIMENTOS REALIZADOS	70
7.3 CONSIDERAÇÕES FINAIS	81
8. CONCLUSÕES.....	82
9. REFERÊNCIAS BIBLIOGRÁFICAS.....	83

Lista de Figuras

Figura 3.1 - Função de pertinência do tipo-1.....	24
Figura 3.2 - Função de pertinência do tipo-1 manchada	24
Figura 3.3 - Exemplo de função de pertinência do tipo-2	26
Figura 3.4 – Exemplo de corte vertical para a função de pertinência do tipo-2 da figura 3.3	26
Figura 3.5 - FOU para a função de pertinência primária Gaussiana com incerteza na média.	29
Figura 3.6 - FOU para a função de pertinência primária Gaussiana com incerteza no desvio padrão.....	29
Figura 3.7 - Funções de pertinência superior e inferior do conjunto fuzzy do tipo-2.	30
Figura 3.8 - Sistema genérico de inferência fuzzy do tipo-2, baseado em (KARNIK; MENDEL; LIANG, 1999).....	33
Figura 4.1 - Cromossomo binário (a), cromossomo inteiro (b) e cromossomo real (c). 42	
Figura 4.2 - Método de seleção por roleta	44
Figura 4.3 - Cruzamento de um ponto	45
Figura 4.4 - Cruzamento de dois pontos.....	45
Figura 4.5 - Cruzamento uniforme	46
Figura 5.1 - Estrutura cromossômica.....	53
Figura 6.1 - Índices de medidas para a construção da função de pertinência intervalar do tipo-2.....	60
Figura 6.2 - Representação cromossômicas da base de regras.....	61
Figura 6.3 - Representação cromossômica da base de dados com funções de pertinência do tipo-1.....	63
Figura 6.4 - Representação cromossômica da base de dados com funções de pertinência do tipo-2.....	64
Figura 6.5 - Codificação de um cromossomo da etapa 3.	65

Lista de Tabelas

Tabela 7-1 - Conjuntos utilizados nos experimentos	69
Tabela 7-2 - Parâmetros dos AGs.....	70
Tabela 7-3 - Parâmetros dos AGs auto-adaptativos.	71
Tabela 7-4 - Taxa de classificação correta	72
Tabela 7-5 - Taxa de classificação correta.	72
Tabela 7-6 - Taxa de classificação correta.	72
Tabela 7-7 - Número de regras geradas.....	73
Tabela 7-8 - Número de regras geradas.....	73
Tabela 7-9 - Número de regras geradas.....	73
Tabela 7-10 - Ranking de classificação.....	74
Tabela 7-11 - Diferença entre média de ranking.	76
Tabela 7-12 - Ranking de classificação.....	77
Tabela 7-13 - Diferença entre média de ranking.	78
Tabela 7-14 - Ranking de em relação ao número de regras	79
Tabela 7-15 - Diferença entre média de ranking.	80

Lista de Abreviaturas

BC	Base de Conhecimento
BD	Base de Dados
BR	Base de Regras
DG	Diversidade Genética
FCM	Fuzzy C-Means
MI	Mecanismo de Inferência
RN	Redes Neurais
SCF	Sistemas de Classificação Fuzzy
SF	Sistema Fuzzy
SFBR	Sistema Fuzzy Baseado em Regras
SFG	Sistema Fuzzy-Genético
TCC	Taxa de Classificação Correta
WM	Wang & Mendel

1. Introdução

Em um sentido bastante abrangente, um Sistema Fuzzy (SF) é um sistema baseado na teoria de conjuntos fuzzy, a qual pode ser usada como base para a representação de conhecimento ou para modelar as interações e relacionamentos entre as variáveis do sistema. Os SFs têm provado serem ferramentas importantes para a modelagem de sistemas nos quais, devido à complexidade ou imprecisão, as abordagens clássicas são limitadas. Na grande maioria desses sistemas a representação do conhecimento se baseia na definição de categorias com limites imprecisos, implementadas pelo conceito de função de pertinência de conjuntos fuzzy, que permitem que elementos pertençam a um conjunto com graus que variam dentro de um intervalo pré-estabelecido (KLIR; YUAN, 1995)(PEDRYCZ; GOMIDE, 1998).

Com a constatação de que o poder de representação dos conjuntos fuzzy poderia ser mais bem explorado se ao invés de graus de pertinência definidos como valores numéricos e precisos, fossem utilizados graus de pertinência que são eles próprios conjuntos fuzzy, os conjuntos fuzzy tipo-2 foram introduzidos por Zadeh (1975) e, em seguida, estudados por diversos autores (MIZUMOTO; TANAKA, 1976) (MIZUMOTO; TANAKA, 1981)(KLIR;& FOLGER, 1988). A terminologia referente aos conjuntos fuzzy tipo-2 e algumas tentativas de classificação dos tipos desses conjuntos foram introduzidas em (MENDEL; JOHN, 2002)(MITCHELL, 2006).

Conjuntos fuzzy do tipo-2 nos permitem modelar vários tipos de incertezas, as quais nós não poderíamos modelar apropriadamente se utilizássemos conjuntos fuzzy do tipo-1 (conjuntos fuzzy ordinários) (MENDEL, 2007). Naturalmente, o emprego de conjuntos fuzzy do tipo-2 aumenta a complexidade computacional em comparação aos conjuntos fuzzy do tipo-1, pois eles possuem uma dimensão a mais em sua função de pertinência. Para cada pertinência primária, há um grau secundário para ser computado. De forma a simplificar esta complexidade computacional usam-se conjuntos fuzzy intervalares do tipo-2, permitindo que as funções de pertinência secundárias possam distribuir uniformemente a incerteza nas pertinências primárias.

Os pesquisadores divergem quanto à conveniência de uso dessa forma de representação. Enquanto alguns afirmam que o alto poder de representatividade obtido justifica o custo do aumento da complexidade computacional, outros afirmam que a

obtenção de formas mais simples de representação da função de pertinência é fundamental para o bom desempenho do sistema.

Os conjuntos fuzzy intervalares do tipo-2 vêm sendo empregados nos mais diversos campos de aplicação: aproximação, agrupamento de dados (clustering), controle, banco de dados, tomada de decisão, cuidados médicos, qualidade de controle, wireless, classificadores (foco deste trabalho) entre outros. Coleções representativas de sistemas e aplicações usando conjuntos fuzzy tipo-2 podem ser encontradas em (MENDEL, 2001)(MENDEL, 2007).

Os avanços constatados recentemente podem ser atribuídos às abordagens que propõem métodos que permitem que as operações entre conjuntos tipo-2 da forma geral, que envolvem um problema combinatório intratável, sejam reduzidas a operações mais simples. Como exemplos de simplificações bem sucedidas, podemos citar a proposta de Mendel & John (2002) para o cálculo de operações de intersecção, união e complemento de conjuntos fuzzy intervalares do tipo-2, baseadas em aritmética de intervalos e a proposta de Coupland (2007) de um defuzificador geométrico.

No problema de classificação de padrões, tanto os atributos quanto as hipóteses possuem incertezas, o que sugere ser adequado o uso de conjuntos fuzzy intervalares do tipo-2 nos classificadores para obter melhores resultados quanto à generalização de conceitos e quanto à taxa de classificação. Os classificadores fuzzy intervalares do tipo-2 obtiveram bons resultados em diversos domínios do conhecimento como, por exemplo, na classificação de veículos de batalha (WU & MENDEL, 2007), no reconhecimento de fala (ZENG; LIU, 2006a) (ZENG; LIU, 2004a)(ZENG; LIU, 2004b) e no reconhecimento de caractere chinês escrito à mão (ZENG; LIU, 2004) (ZENG; LIU, 2006b) (ZENG; LIU, 2006c).

Independentemente da forma de representação utilizada, a geração automática de Bases de Conhecimento Fuzzy tornou-se nos últimos anos um tópico de grande interesse para pesquisadores das áreas que investigam o grande tema de Inteligência Computacional. As metodologias fundamentadas na lógica fuzzy, que fornecem o aparato matemático para representação e processamento do conhecimento nos Sistemas Fuzzy oferecem mecanismos flexíveis para tratamento da imprecisão inerente a diversos domínios do conhecimento. A construção ou otimização automática de tais sistemas ou de partes dele geralmente utiliza-se de metodologias complementares, equipadas com recursos para aprendizado de máquina como, por exemplo, métodos de agrupamento de dados, Redes Neurais (RN) e Algoritmos Genéticos (AGs), entre outros (JANG et al.

1997)(NAUCK; KRUSE, 1997)(CASTELLANO; FANELLI, 2001)(CORDÓN et al, 2001)(CORDÓN et al. 2004)(CORDÓN et al., 2007).

Os crescentes esforços empregados nas duas últimas décadas para explorar o uso de algoritmos genéticos no projeto de SFs deram origem ao campo dos Sistemas Fuzzy Genéticos (SFGs), que tem atraído a atenção de pesquisadores de diferentes áreas. Um SFG é basicamente um SF acrescido de um processo de aprendizado proveniente de um AG (Cordón et al. 2004).

Este trabalho focaliza o uso de conjuntos fuzzy intervalares do tipo-2 no contexto dos SFG.

1.1 Motivações e objetivos

O objetivo deste trabalho é estudar, expandir e avaliar o uso de conjuntos fuzzy intervalares do tipo-2 na representação do conhecimento em sistemas de inferência fuzzy, mais especificamente para os classificadores fuzzy, bem como sua geração automática a partir de conjuntos de dados, por meio de AGs. Esse trabalho investiga o uso de tais conjuntos com enfoque na questão de balanceamento entre o acréscimo de custo da representação e os ganhos em interpretabilidade e precisão, ambos decorrentes da complexidade de representação e processamento dos conjuntos fuzzy intervalares do tipo-2.

Para tanto, foi criado um modelo genético baseado na abordagem de Pittsburg composta por três etapas:

1. **Geração da Base de Regras** - Nesta etapa um AG é usado para gerar regras fuzzy capazes de representar o conhecimento existente no conjunto de exemplos. Aqui a Base de Dados é gerada através do algoritmo Fuzzy C-Means.
2. **Otimização da Base de Dados** - Nesta etapa um AG é usado para otimizar a Base de Dados que foi usada na etapa anterior.
3. **Otimização da Base de Regras** - Esta etapa visa excluir, usando também AG, as regras redundantes e desnecessárias que por ventura foram geradas nas etapas anteriores, resultando numa base de regras compacta.

A classe de problemas abordada é a de classificação de padrões multi-dimensional. É necessário que exista um conjunto de dados numéricos que representam amostras do problema na forma atributos-valor, pelos quais o AG consegue extrair as regras fuzzy.

Para avaliar o desempenho da metodologia apresentada sob diversas condições, diferentes testes foram realizados utilizando alguns conjuntos de dados obtidos no UCI Repository of Machine Learning Databases (BLAKE & MERZ, 1998).

1.2 Organização do trabalho

A organização deste trabalho se distribui da seguinte maneira: no capítulo 2 são abordados os conceitos de conjuntos fuzzy do tipo-1, ao passo que o capítulo 3 abordará os conceitos de conjuntos fuzzy do tipo-2. No capítulo 4 é feita uma breve abordagem sobre os AGs. No capítulo 5 serão apresentados alguns trabalhos relacionados. No capítulo 6 a proposta do trabalho é apresentada. No capítulo 7 são apresentados os resultados e o capítulo 8 conclui o trabalho.

2. Sistemas Fuzzy

O conceito de conjuntos fuzzy foi inicialmente desenvolvido por Zadeh (1965) com a idéia de capturar, representar e processar dados associados a noções linguísticas ou objetos com fronteiras mal-definidas de maneira formal. O resultado foi a formalização de uma plataforma de processamento de informações imprecisas inspirada no conhecimento humano. Desta maneira os sistemas fuzzy representam uma possível ponte entre o processamento linguístico e o processamento numérico (DUBOIS; PRADE, 1998).

A lógica fuzzy pode ser entendida como uma extensão da lógica binária e da lógica multi-valorada. Na lógica binária uma proposição sempre assume valores verdadeiro ou falso. Em uma lógica multi-valorada, além disso, o valor da proposição pode assumir alguns valores intermediários. A lógica fuzzy, em contrapartida, assume que os valores-verdade são conjuntos fuzzy permitindo a utilização de um raciocínio aproximado linguisticamente interpretável.

Na seção 2.1 será apresentado o conceito de conjuntos fuzzy. Serão abordados tópicos como funções de pertinência e operadores fuzzy. Na seção 2.2 será introduzido o conceito de relações e regras fuzzy além do mecanismo clássico de raciocínio fuzzy. Na seção 2.3 serão apresentados os sistemas de inferência e na seção 2.4 serão apresentados os sistemas de classificação fuzzy.

2.1 Conjuntos Fuzzy

Os conjuntos clássicos (crisp sets) encontram aplicações na matemática e na ciência da computação. Entretanto, este tipo de conjunto não é adequado para tratar a natureza do pensamento humano, que tende a ser imprecisa e abstrata (JANG; SUN; MIZUTANI, 1997).

Na teoria clássica de conjuntos um elemento de um universo pertence ao conjunto ou não pertence ao conjunto. Entretanto, o mundo real apresenta situações em que os conjuntos clássicos não são capazes de definir de forma adequada certas classes de elementos por não possuírem flexibilidade suficiente e forçarem uma transição abrupta entre a pertinência absoluta e a não pertinência absoluta.

Já os conjuntos fuzzy, considerados uma generalização dos conjuntos clássicos, possibilitam a transição gradual, permitindo que cada elemento pertença parcialmente a um ou mais conjuntos.

Seja X um conjunto de objetos, denominado universo, e x um elemento de X . Um conjunto fuzzy A é definido como:

$$A = \{(x, \mu_A(x)) \mid x \in X\} \quad (2.1)$$

onde μ_A é uma função de pertinência definida como um mapeamento da forma:

$$\mu_A : X \rightarrow [0,1] \quad (2.2)$$

e $\mu_A(x)$ indica o grau de pertinência de x no conjunto fuzzy A . Nos casos extremos temos que $\mu_A(x) = 1$ indica que o elemento pertence totalmente ao conjunto e $\mu_A(x) = 0$ indica que o elemento não pertence ao conjunto A .

2.1.1 Funções de pertinência

A base da teoria de conjuntos fuzzy é o conceito de pertinência. Os formatos mais comuns de funções de pertinência são: triangular, trapezoidal e Gaussiana (KLIR; YUAN, 1995).

Função triangular: A função triangular pode ser apresentada na forma parametrizada como:

$$\mu_A(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ \frac{c-x}{c-b}, & b < x \leq c \\ 0, & x > c \end{cases} \quad (2.3)$$

Alternativamente, pode-se reescrever a expressão 2.3 como:

$$\mu_A(x) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \quad (2.4)$$

onde b é o valor modal de μ e a e c representam os limites inferior e superior dos elementos compatíveis com o conjunto fuzzy respectivamente. Graficamente, pode-se interpretar o conjunto $\{a, b, c\}$ como as abscissas dos três vértices do triângulo formado pela função de pertinência.

Função trapezoidal : A função trapezoidal é especificada por um conjunto de quatro parâmetros: $\{a, b, c, d\}$, conforme a expressão 2.5:

$$\mu_A(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ 1, & b < x \leq c \\ \frac{d-x}{d-c}, & b < x \leq c \\ 0, & x > c \end{cases} \quad (2.5)$$

Alternativamente, pode-se reescrever a expressão 2.5 como:

$$\mu_A(x) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right) \quad (2.6)$$

Graficamente, pode-se interpretar o conjunto $\{a, b, c, d\}$ como as abscissas dos quatro vértices do trapézio formado pela função de pertinência. Para o caso especial em que $b=c$ a função trapezoidal se reduz a uma função triangular. Para o caso especial $a=b$ e $c=d$ tem-se um conjunto clássico.

Função Gaussiana: Uma função de pertinência Gaussiana é especificada por dois parâmetros, o valor modal (ω) e a dispersão (σ), conforme definida a seguir:

$$\mu_A(x) = e^{-\frac{1}{2}\left(\frac{x-\omega}{\sigma}\right)^2} \quad (2.7)$$

A função Gaussiana é simétrica ao redor de ω e sua dispersão é especificada pelo parâmetro σ . Podemos também afirmar que o suporte¹ de *Gaussiana*(x, ω, σ) é infinito.

Devido à sua formulação simples e eficiência computacional, as funções do tipo triangular e trapezoidal são amplamente utilizadas, especialmente em implementações com requisitos de funcionamento em tempo real.

A especificação de um formato para uma função de pertinência nem sempre é óbvia, podendo inclusive não estar ao alcance do conhecimento de um especialista para a aplicação desejada. No entanto existem sistemas fuzzy cujos parâmetros das funções de pertinência são completamente definidos pelo especialista. Nestes casos, a escolha de funções triangulares e trapezoidais é mais comum porque a idéia de definição de regiões de pertinência total, média e nula é mais intuitiva que a especificação de valor modal e dispersão, associados ao projeto de funções Gaussianas. Entretanto, em trabalhos mais recentes existe a tendência ao projeto automático de sistemas fuzzy nos quais os parâmetros das funções de pertinência são ajustados no sentido de otimizar alguma função objetivo. Em geral essa otimização é realizada a partir de um conjunto significativo de dados de treinamento, ou seja, um conjunto de dados de treinamento que descreva adequadamente o comportamento do sistema.

Dentre os vários tipos de conjuntos fuzzy, um de significância especial são os conjuntos fuzzy que são definidos sobre os conjuntos dos números reais, chamados números fuzzy. As funções de pertinência desses conjuntos possuem um significado quantitativo e são usadas para modelar a aproximação de quantidades como, por exemplo, “perto de 5”, “abaixo de 100”, etc.

2.1.2 Operações básicas sobre conjuntos Fuzzy

As operações sobre conjuntos são essenciais para a representação e processamento do conhecimento. A seguir são descritas as operações fuzzy de complemento, intersecção e união.

¹ O conjunto suporte do *conjunto fuzzy* A é o conjunto numérico que contém todos os elementos de X que possuem grau de pertinência diferente de zero em A.

Complemento fuzzy

Se A é um conjunto fuzzy em X , o complemento fuzzy de A é denotado por cA e o valor de $c\mu_A(x)$ para $x \in X$ é interpretado tanto como o grau com que x pertence a cA quanto como o grau com o qual x não pertence a A e pode ser definido pela função:

$$c : [0,1] \rightarrow [0,1] \quad (2.8)$$

que atribui o valor $c(\mu_A(x))$ a cada grau de pertinência $\mu_A(x)$ de qualquer conjunto fuzzy A . O valor $c(\mu_A(x))$ é interpretado como o valor $c\mu_A(x)$, ou seja, para todo $x \in X$, por definição, $c(\mu_A(x)) = c\mu_A(x)$.

Para que os complementos fuzzy obtidos possam ser significativos e possam efetivamente expressar a idéia de complemento, é importante que a função c satisfaça os axiomas:

- Condições contorno: $c(0) = 1$ e $c(1) = 0$
- Monotonicidade: $c(a) \geq c(b)$, se $a \leq b$

Todas as funções que satisfaçam esses requisitos formam a classe geral de complementos fuzzy. O primeiro requisito garante que os complementos fuzzy podem ser utilizados para conjuntos clássicos. A segunda condição garante que um aumento do grau de pertinência de um conjunto indica uma diminuição da pertinência de sua negação. Um requisito opcional, que não costuma ser satisfeito pela maioria dos complementos fuzzy, é a condição de involução da negação fuzzy: $c(c(a)) = a$ que garante que a negação dupla do conjunto é o próprio conjunto.

Os tipos mais comuns de negação fuzzy encontrados na literatura (KLIR; YUAN, 1995) são:

- Complemento: $c_1(a) = 1 - a$
- Sugeno: $c_2(a, s) = \frac{1 - a}{1 + s \cdot a}$, se $s > -1$
- Yager: $c_3(a, w) = (1 - a^w)^{\frac{1}{w}}$, se $w > 0$

Podemos observar que $c_2(a,0) = c_1(a)$ e também $c_3(a,1) = c_1(a)$.

Intersecção

A intersecção de dois conjuntos fuzzy é definida pela operação binária $[0,1]^2 \rightarrow [0,1]$ que associa um grau de pertinência a um par de graus de pertinência.

Sejam A e B dois conjuntos fuzzy definidos no universo X , e $\mu_A(x)$ e $\mu_B(x)$ suas funções de pertinência, respectivamente, então:

$$\mu_{A \cap B}(x) = \mu_A(x) \mathbf{t} \mu_B(x) \quad (2.9)$$

De maneira semelhante ao complemento fuzzy, para que uma função \mathbf{t} assim definida possa ser considerada e adotada como intersecção fuzzy, deve satisfazer às seguintes propriedades que garantam que os conjuntos fuzzy resultantes da aplicação de \mathbf{t} sejam significativos como intersecções fuzzy:

- Condição de contorno: $0 \mathbf{t} x = 0 = 0$ e $1 \mathbf{t} x = x$
- Monotonicidade: Se $x \leq y$ e $w \leq z$, então $x \mathbf{t} w \leq y \mathbf{t} z$
- Comutatividade: $x \mathbf{t} y = y \mathbf{t} x$
- Associatividade: $x \mathbf{t} (y \mathbf{t} z) = (x \mathbf{t} y) \mathbf{t} z$

As funções conhecidas como t -normas possuem tais propriedades e, geralmente, são adotadas como intersecções fuzzy.

Algumas das t -normas mais utilizados na literatura (KLIR; YUAN, 1995) são:

- mínimo: $a \mathbf{t} b = \min\{a, b\}$
- produto algébrico: $a \mathbf{t} b = a.b$
- produto limitado: $a \mathbf{t} b = \max\{0, a+b-1\}$
- produto drástico: $a \mathbf{t} b = \begin{cases} a, & \text{se } b = 1 \\ b, & \text{se } a = 1 \\ 0, & \text{se } a \text{ e } b \neq 1 \end{cases}$
- Schweizer e Sklar: $a \mathbf{t} b = [\max\{0, (a^{-p} + b^{-p} - 1)\}]^{\frac{-1}{p}}, p > 0$

- Yager: $a \mathbf{t} b = 1 - \min\{1, [(1-a)^q + (1-b)^q]^{\frac{1}{q}}\}, q > 0$
- Dubois e Prade: $a \mathbf{t} b = \frac{ab}{\max\{a, b, \alpha\}}, \alpha \in [0, 1]$
- Hamacher: $a \mathbf{t} b = \frac{ab}{\gamma + (1-\gamma)(a+b-ab)}, \gamma > 0$
- Frank: $a \mathbf{t} b = \log_s \left[1 + \frac{(s^a - 1)(s^b - 1)}{s - 1} \right], s > 0$
- Sugeno: $a \mathbf{t} b = \max\{0, (\lambda + 1)(a + b - 1) - \lambda \cdot a \cdot b\}, \lambda \geq -1$
- Dombi: $\frac{1}{1 + [(a^{-1} - 1)^\lambda + (b^{-1} - 1)^\lambda]^{\frac{1}{\lambda}}}, \lambda > 0$

União

De maneira semelhante à intersecção fuzzy, a união fuzzy de dois conjuntos fuzzy é definida pela operação binária $[0, 1]^2 \rightarrow [0, 1]$ que associa um grau de pertinência a um par ordenado de graus de pertinência. Se A e B são os conjuntos fuzzy em questão e X o conjunto universo em que estão definidos, para todo $x \in X$ temos:

$$\mu_{A \cup B}(x) = \mu_A(x) \mathbf{s} \mu_B(x) \quad (2.10)$$

De maneira semelhante ao complemento e à intersecção fuzzy, para que uma função s , assim definida, possa ser considerada e adotada como união fuzzy, deve satisfazer propriedades que garantam que os conjuntos fuzzy resultantes da aplicação de s sejam significativos como uniões fuzzy. As funções conhecidas como t-conormas, ou s-normas, possuem tais propriedades e, geralmente, são adotadas como uniões fuzzy.

As propriedades das s-normas citadas são:

- Condição de contorno: $0 \mathbf{s} x = x$ e $1 \mathbf{s} x = 1$
- Monotonicidade: Se $x \leq y$ e $w \leq z$, então $x \mathbf{s} w \leq y \mathbf{s} z$
- Comutatividade: $x \mathbf{s} y = y \mathbf{s} x$
- Associatividade: $x \mathbf{s} (y \mathbf{s} z) = (x \mathbf{s} y) \mathbf{s} z$

Os tipos de s-normas triangulares mais utilizados na literatura (KLIR; YUAN, 1995) são:

- máximo: $a \text{ s } b = \max\{a, b\}$
- soma algébrica: $a \text{ s } b = a+b-a.b$
- soma restrita: $a \text{ s } b = \max\{1, a+b\}$
- soma drástica: $a \text{ s } b = \begin{cases} a, & \text{se } b = 0 \\ b, & \text{se } a = 0 \\ 1, & \text{se } a \text{ e } b \neq 0 \end{cases}$
- Schweizer e Sklar: $a \text{ s } b = 1 - [\max\{0, ((1-a)^{-p} + (1-b)^{-p} - 1)\}]^{\frac{-1}{p}}, p > 0$
- Yager: $a \text{ s } b = \min\{1, [a^q + b^q]^{\frac{1}{q}}\}, q > 0$
- Dubois e Prade: $a \text{ s } b = \left[a + b - ab - \frac{\min\{a, b, (1-\alpha)\}}{\max\{1-a, 1-b, \alpha\}} \right], \alpha \in [0, 1]$
- Hamacher: $a \text{ s } b = \frac{a + b - (\gamma - 2)ab}{1 + (\gamma - 1)ab}, \gamma > 0$
- Frank: $a \text{ s } b = 1 - \log_s \left[1 + \frac{(s^{1-a} - 1)(s^{1-b} - 1)}{s - 1} \right], s > 0$
- Sugeno: $a \text{ s } b = \min\{1, a + b - \gamma ab\}, \gamma \geq -1$
- Dombi: $\frac{1}{1 + [(a^{-1} - 1)^{-\lambda} + (b^{-1} - 1)^{-\lambda}]^{\frac{1}{\lambda}}}, \lambda > 0$

2.1.3 Relações Fuzzy

Sejam X e Y dois universos. Pode-se definir o conjunto fuzzy R como:

$$R: X \times Y \rightarrow [0, 1] \quad (2.11)$$

$$R: \{(x, y), \mu_R(x, y) \mid (x, y) \in X \times Y\} \quad (2.12)$$

sendo que a expressão 2.12 é uma relação fuzzy binária no universo $\mathbf{X} \times \mathbf{Y}$. Por exemplo, seja $\mathbf{X} \equiv \mathbf{Y} \equiv \mathfrak{R}^+$ e $R = "x \text{ é muito maior que } y"$. Supondo que a função de pertinência de R possa ser aproximada por:

$$\mu_R(x, y) = \begin{cases} \frac{y-x}{x+y-2}, & y > x \\ 0, & y \leq x \end{cases} \quad (2.13)$$

Para os universos $\mathbf{X} = \{3,4,5\}$ e $\mathbf{Y} = \{3,4,5,6,7\}$, pode-se expressar a mesma relação fuzzy R através de uma matriz de relação:

$$R = \begin{bmatrix} 0 & 0.200 & 0.333 & 0.429 & 0.500 \\ 0 & 0 & 0.143 & 0.250 & 0.333 \\ 0 & 0 & 0 & 0.111 & 0.200 \end{bmatrix} \quad (2.14)$$

onde o elemento na linha i e coluna j indica o grau de pertinência no conjunto “muito maior” do elemento (i, j) sendo $i \in \mathbf{X}$ e $j \in \mathbf{Y}$.

2.1.3.1 Composição de relações Fuzzy

As relações fuzzy definidas em produtos cartesianos distintos podem ser combinadas utilizando operadores de composição. Existem diversos operadores de composição possíveis sendo que o mais conhecido é a composição max-min proposta por Zadeh (1965).

Composição Max-min

Sejam R_1 e R_2 duas relações fuzzy definidas em $\mathbf{X}_1 \times \mathbf{X}_2$ e $\mathbf{X}_2 \times \mathbf{X}_3$ respectivamente. A composição Max-min de R_1 e R_2 é um conjunto fuzzy definido por:

$$R_1 \circ R_2 = \{((x_1, x_3), \mu_{R_1 \circ R_2}(x_1, x_3)) \mid x_1 \in \mathbf{X}_1, x_3 \in \mathbf{X}_3\} \quad (2.15)$$

Onde:

$$\mu_{R_1 \circ R_2}(x_1, x_3) = \max_{x_2 \in X_2} \min(\mu_{R_1}(x_1, x_2), \mu_{R_2}(x_2, x_3)) \quad (2.16)$$

A composição Max-prod foi proposta como uma alternativa à composição Max-min (JANG; SUN; MIZUTANI, 1997), entre outras coisas, para alcançar maior tratabilidade matemática.

Composição Max-Prod

Assumindo a mesma notação utilizada na definição da composição Max-min, pode-se definir a composição Max-prod como:

$$\mu_{R_1 \circ R_2}(x_1, x_3) = \max_{x_2 \in X_2} [\mu_{R_1}(x_1, x_2) \cdot \mu_{R_2}(x_2, x_3)] \quad (2.17)$$

Composição genérica

Analogamente pode-se utilizar o mesmo raciocínio para definir uma composição através de quaisquer pares de s-normas e t-normas (NGUYEN; GANESH; GONG, 1998).

$$\mu_{R_1 \oplus R_2}(x_1, x_3) = s_{x_2 \in X_2} [\mu_{R_1}(x_1, x_2) t \mu_{R_2}(x_2, x_3)] \quad (2.18)$$

ou ainda:

$$\mu_{R_1 \oplus R_2}(x_1, x_3) = t_{x_2 \in X_2} [\mu_{R_1}(x_1, x_2) s \mu_{R_2}(x_2, x_3)] \quad (2.19)$$

2.2 Regras Fuzzy

Regras fornecem uma maneira formal de representar diretrizes e estratégias, sendo muitas vezes apropriadas quando o conhecimento do domínio resulta da experiência ou de associações empíricas. Sistemas baseados em regras são construídos sobre um conjunto de regras e usa uma coleção de fatos para fazer inferências.

Regras são úteis para a representação de conhecimento sobre um determinado domínio. A modelagem de sentenças em linguagem natural ou artificial é relativamente simples por meio de regras fuzzy. Além disso, as regras permitem a representação do

conhecimento impreciso, o que facilita a modelagem de problemas do mundo real. As regras, fuzzy ou não, podem ser usadas por sistemas computacionais em seus mecanismos de raciocínio para inferir conclusões a partir de fatos conhecidos.

Esta seção descreve os conceitos de variáveis e regras fuzzy, bem como os fundamentos dos processos de inferência que utilizam regras fuzzy.

2.2.1 Variáveis linguísticas

Conforme apontado por Zadeh, técnicas convencionais para análise de sistemas são essencialmente inadequadas para o tratamento de sistemas baseados no conhecimento humano, cujo comportamento é influenciado pela percepção, julgamento e emoções. Essa é uma manifestação do princípio da incompatibilidade enunciado por Zadeh: "Com o aumento da complexidade do sistema, nossa habilidade de realizar indicações precisas e significativas sobre seu funcionamento diminui até que é alcançado um nível em que precisão e significado se tornam características quase mutuamente exclusivas" (ZADEH, 1973). Isso motivou Zadeh a propor o conceito de variáveis linguísticas como uma alternativa na modelagem do pensamento humano em que a informação é processada através de conjuntos fuzzy.

Formalmente uma variável linguística é caracterizada por uma quintupla:

$$\langle \chi, T(\chi), X, G, M \rangle \quad 2.20)$$

onde:

- χ é o nome da variável linguística;
- $T(\chi)$ é o conjunto de termos (ou valores) linguísticos. Cada elemento de $T(\chi)$ representa um rótulo dos termos que a variável χ pode assumir;
- X é o universo da variável linguística χ ;
- G é a gramática para geração dos rótulos;
- M é a regra que associa cada rótulo a um conjunto fuzzy no universo X , representando o seu significado.

Por exemplo, seja a variável linguística *velocidade* ($\chi = \text{velocidade}$) no universo $X = [0,150]$ e uma variável base x . Um conjunto de termos linguísticos associados à variável *velocidade* poderia ser: $T(\text{velocidade}) = \{\text{“baixa”}, \text{“média”}, \text{“alta”}\}$. Seja $P(X)$ uma coleção de conjuntos fuzzy definidos no universo X . A semântica de uma variável linguística é definida por um mapeamento:

$$M: T(\chi) \rightarrow P(X) \quad 2.21)$$

que associa um elemento do conjunto de termos $T(\chi)$ a um conjunto fuzzy de $P(\chi)$.

A gramática G define como os termos linguísticos não-primários serão produzidos, a partir dos termos primários, possivelmente associados à modificadores $\{\text{“muito”}, \text{“pouco”}, \text{etc}\}$. Usualmente as funções de pertinência associadas aos termos linguísticos primários têm formatos conhecidos. Os conjuntos fuzzy associados aos termos não primários, em contrapartida podem ser obtidos através de modificadores pré-especificados (PEDRYCZ; GOMIDE, 1998).

O processo de particionar um universo de uma variável em conjuntos fuzzy define uma partição fuzzy. Uma partição fuzzy pode ser uniforme (os termos linguísticos têm funções de pertinência de formato semelhante transladadas e equidistantes) ou não-uniforme. A granularidade da partição é definida pelo número de termos linguísticos. Um número baixo de termos linguísticos define uma partição esparsa ou grossa, enquanto um número alto resulta em uma partição fina.

Não existe uma metodologia consistente para a determinação da partição fuzzy ideal. Em geral essa tarefa é realizada através da intervenção de um especialista ou utilizando um método de particionamento a partir de dados de treinamento. Na ausência de um especialista ou de abordagens de ajuste automático, partições utilizando de 5 a 7 termos linguísticos uniformemente distribuídos são utilizadas com frequência (PEDRYCZ; GOMIDE, 1998).

2.2.2 Sintaxe e semântica das regras Fuzzy

As regras fuzzy são também conhecidas como proposições condicionais fuzzy. Elas são apropriadas quando o conhecimento do problema é resultado de associações empíricas e

experiências de um especialista, ou quando se deseja uma representação linguística do conhecimento adquirido.

Em geral, as regras fuzzy assumem a forma:

$$SE\ x\ é\ A\ ENTÃO\ y\ é\ B \quad (2.22)$$

onde A e B são rótulos linguísticos de conjuntos fuzzy definidos nos universos X e Y , respectivamente, e x e y são variáveis linguísticas. Frequentemente “ x é A ” é denominado de antecedente ou premissa, enquanto “ y é B ” é denominado de consequente ou conclusão (JANG; SUN; MIZUTANI, 1997).

A regra fuzzy 2.22 pode ser abreviada por $A \rightarrow B$ e pode ser definida como uma relação fuzzy R no produto cartesiano $X \times Y$, onde X e Y são conjuntos convencionais. Seja μ_R a função de pertinência associada à relação fuzzy R , a semântica da regra fuzzy pode ser especificada através de uma função $f: [0,1] \times [0,1] \rightarrow [0,1]$:

$$\mu_R(x, y) = f(\mu_A(x), \mu_B(y)) \quad (2.23)$$

De acordo com Pedrycz e Gomide (1998), as relações fuzzy induzidas por regras fuzzy são derivadas de três classes principais de funções: conjunções fuzzy, disjunções fuzzy e implicações fuzzy, sendo as conjunções e implicações as mais comuns. As conjunções e as disjunções podem ser vistas como generalizações duais do produto cartesiano fuzzy através de t-normas e s-normas, enquanto implicações fuzzy representam generalizações das implicações da lógica multi-valorada de Lukasiewicz (1970).

2.3 Sistemas de Inferência Fuzzy

As regras fuzzy são ferramentas eficientes na modelagem de sentenças em linguagem natural (JANG; SUN; MIZUTANI, 1997). Esta modelagem utiliza o conceito de variáveis linguísticas para gerar regras fuzzy que são mapeadas através de relações fuzzy permitindo a investigação de diferentes esquemas de raciocínio aproximado. Nestes esquemas, procedimentos de inferência baseados no conceito de regra

composicional de inferência são utilizados para derivar conclusões, a partir de uma base de conhecimento contendo regras e fatos conhecidos.

Um sistema de inferência fuzzy é uma ferramenta de computação baseada nos conceitos de conjuntos fuzzy, regras fuzzy e raciocínio fuzzy. Eles encontram aplicações em diversas áreas como controle automático, classificação de dados, tomada de decisões, predição de séries temporais, robótica e reconhecimento de padrões (JANG; SUN; MIZUTANI, 1997). Outras nomenclaturas para os sistemas fuzzy encontradas na literatura incluem: sistemas fuzzy baseados em regras, sistemas especialistas fuzzy (KANDEL, 1992), modelos fuzzy (TAKAGI; SUGENO, 1985)(SUGENO; KANG, 1988), memórias associativas fuzzy(KOSKO, 1992) e controlador lógico fuzzy (KANDEL, 1992)(LEE, 1990).

A estrutura básica de um sistema de inferência fuzzy é constituída de três componentes principais:

- **Base de regras** que contém um conjunto de regras fuzzy;
- **Base de dados** que define as funções de pertinência utilizadas nas regras fuzzy;
- **Mecanismo de inferência** que executa o procedimento de inferência sobre as regras fuzzy e retorna fatos para formar uma saída ou conclusão.

Observa-se que o sistema básico de inferência fuzzy pode receber como entrada tanto conjuntos fuzzy quanto entradas não-fuzzy (que são interpretadas como conjuntos unitários). As saídas produzidas são geralmente conjuntos fuzzy ou valores numéricos, dependendo do modelo de inferência utilizado. Em alguns casos é necessário obter como saída um valor escalar, mesmo nos casos em que as regras produziram uma saída não-fuzzy. Portanto é necessário um método de transformação da saída fuzzy em não-fuzzy para extrair um valor escalar que melhor represente o conjunto fuzzy. O tipo de sistema fuzzy utilizado ao longo desse trabalho, entretanto, não necessita de tal procedimento que não será profundamente detalhado.

Existem vários tipos de sistemas fuzzy. Na maioria dos casos, o antecedente é formado por proposições linguísticas e a distinção entre os modelos se dá no conseqüente das regras fuzzy. Entre os modelos mais conhecidos pode-se destacar:

Linguístico ou Mamdani: utiliza conjuntos fuzzy nos conseqüentes das regras fuzzy (MAMDANI; ASSILIAN, 1975). A saída final é representada por um conjunto fuzzy resultante da agregação da saída inferida de cada regra. Para se obter uma saída

final não fuzzy adota-se um dos métodos de transformação da saída fuzzy em não fuzzy(JANG; SUN; MIZUTANI, 1997)(PEDRYCZ; GOMIDE, 1998)(CORDÓN et al., 2001).

Funcional: o consequente é representado por uma função das variáveis de entrada (TAKAGI; SUGENO, 1993). A saída final é obtida pela média ponderada das saídas inferidas de cada regra. Os coeficientes de ponderação são dados a partir do grau de ativação de cada regra.

Tsukamoto: utiliza funções de pertinência monotônicas no consequente (TSUKAMOTO, 1993). A saída do sistema é um valor não-fuzzy obtido da média ponderada das saídas inferidas de cada regra.

2.4 Sistemas Fuzzy Baseados em Regras

Um Sistema Fuzzy é qualquer sistema que incorpora algum mecanismo derivado da lógica fuzzy e que tenha pelo menos uma variável que possa assumir valores linguísticos definidos por conjuntos fuzzy.

O tipo particular de Sistema Fuzzy de interesse para este trabalho são os Sistemas Fuzzy Baseados em Regras (SFBR), ou apenas Sistema de Regras Fuzzy. Os SFBR possuem dois componentes principais: a Base de Conhecimento (BC) e o Mecanismo de Inferência (MI).

A BC é composta pela Base de Dados Fuzzy, que armazena o conjunto de regras fuzzy que modelam um determinado problema.

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

Na próxima seção descreveremos com detalhes os sistemas de classificação fuzzy, que são o tipo de sistema de interesse para este trabalho.

2.5 Sistema 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 Baseados em Regras Fuzzy projetados com a finalidade específica de realizar a tarefa de classificação. Assim, os SCF utilizam métodos de raciocínio próprios para essa tarefa.

A seguir serão apresentados o formato usual de regras para classificação fuzzy e os respectivos mecanismos de inferência apropriados a essa tarefa, por serem de interesse para o trabalho desenvolvido nesta dissertação.

O formato da regra fuzzy de classificação segue o formato geral das regras fuzzy descrito na seção 2.3 e pode ser expressa por:

$$SE \ x_1 \ \acute{e} \ A_{l_1} \ E \ \dots \ x_n \ \acute{e} \ A_{l_n} \ ENT\tilde{A}O \ Classe = C_j \quad 2.24)$$

no qual x_1, \dots, x_n são os atributos do padrão considerado no problema (representados aqui por variáveis linguísticas), A_{l_1}, \dots, A_{l_n} são os valores linguísticos usados para representar os valores de tais n atributos, sendo que l_n representa um conjunto fuzzy da partição i e C_j ($j=1, \dots, M$) é a classe a que pertence o padrão.

Métodos de raciocínio fuzzy para classificação são procedimentos de inferência que tiram conclusões de um conjunto de regras SE...ENTÃO e um padrão de entrada, com o objetivo de classificar esse padrão. Os métodos mais utilizados são: Método de Raciocínio Fuzzy Clássico e Método de Raciocínio Fuzzy Geral.

O método de Raciocínio Fuzzy Clássico (CHI et al., 1996)(GONZÁLEZ & PÉREZ, 1999) classifica um padrão de entrada de acordo com a regra que possuir maior grau de compatibilidade com o mesmo.

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 $\mu_{A_i}(a_{pi}), i = 1, \dots, n$, 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 com cada regra R_K , para $K = 1, \dots, S$.

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

no qual t denota uma t-norma.

2. Encontrar a regra $R_{K_{\max}}$ que possui o maior grau de compatibilidade com o padrão.

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

3. Atribuir a classe C_j ao padrão e_p , onde C_j é a classe do consequente da regra $R_{k_{\max}}$.

O método de Raciocínio Fuzzy Geral combina as informações providas por todas as regras de uma mesma classe para classificar o padrão. Ele é definido a seguir para classificar o padrão e_p :

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

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

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 :

$$Classe_C = f\{Compat(R_K, e_p) | C \text{ é a classe do consequente de } R_K\}$$

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

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

2.6 Considerações Finais

Neste capítulo foram apresentados os principais conceitos sobre conjuntos fuzzy, regras fuzzy e Sistemas Fuzzy, em especial os Sistemas de Classificações Fuzzy, que é o alvo de estudo deste trabalho. No próximo capítulo serão apresentados os conceitos básicos sobre conjuntos fuzzy do tipo-2.

3. Conjuntos fuzzy do tipo-2

Os sistemas reais são inerentemente ruidosos e não lineares, sendo que qualquer elemento de um sistema pode contribuir com incertezas. Incerteza é uma parte inerente a Sistemas de Inferência Fuzzy (SIF) usados em aplicações reais. As seguintes fontes de incerteza podem estar presentes em um SIF (MENDEL, 2001b)(MENDEL; JOHN, 2002):

- Incerteza em relação ao significado das palavras usadas nos antecedentes e consequentes de regras linguísticas: diferentes indivíduos podem exprimir seus pensamentos de forma distinta. Este tipo de incerteza é denominado *imprecisão linguística*;
- Incerteza em relação ao consequente de uma regra: os consequentes podem ter um histograma de valores associados a eles, especialmente quando o conhecimento é extraído de um grupo de especialistas que não estão totalmente de acordo. Desta forma, diferentes consequentes podem ser obtidos para uma mesma regra. Esta incerteza é denominada *conflito*;
- Incerteza em relação às informações que ativam o SIF e que são usadas para ajustar os seus parâmetros. Estas informações podem ser deficientes e, portanto, gerar incerteza. Esta incerteza é denominada *imprecisão numérica*;
- Os dados que são usados para ajustar os parâmetros do SIF também podem ser ruidosos.

Todas essas incertezas se traduzem em incertezas sobre as funções de pertinência dos conjuntos fuzzy. Os conjuntos fuzzy discutidos no capítulo anterior, chamados a partir de agora de conjuntos fuzzy do tipo-1, não são apropriados para modelar diretamente tais incertezas porque suas funções de pertinência são exatas, ou seja, o valor retornado por uma função de pertinência de um conjunto fuzzy do tipo-1 é um número real. Por outro lado, conjuntos fuzzy do tipo-2 são úteis para modelar essas incertezas porque suas funções de pertinência são fuzzy, ou seja, o valor retornado por uma função de pertinência de um conjunto fuzzy do tipo-2 é um número fuzzy.

O conceito de conjunto fuzzy do tipo-2 (*type-2 fuzzy set*) foi introduzido por Zadeh em (ZADEH, 1975) como uma extensão do conceito do conjunto fuzzy do tipo-1. O conjunto fuzzy do tipo-2 é caracterizado pela sua função de pertinência. A função de pertinência dos conjuntos fuzzy do tipo-1 é modelada em duas dimensões, ao passo que a função de pertinência dos conjuntos fuzzy do tipo-2 é modelada em três dimensões. É esta terceira dimensão dos conjuntos fuzzy do tipo-2 que permite um grau adicional de liberdade que torna possível modelar essas incertezas diretamente sendo que o grau de pertinência para cada elemento deste conjunto é um número fuzzy no intervalo $[0, 1]$.

3.1 Definições

Nesta seção serão apresentadas algumas definições (MENDEL; JOHN, 2002) sobre conjuntos fuzzy do tipo-2 e alguns conceitos importantes associados a estes conjuntos.

Na figura 3.1 é ilustrada uma função de pertinência do tipo-1. Imagine agora se nós deslocarmos alguns pontos do triângulo para a esquerda ou para a direita não necessariamente nas mesmas proporções gerando, assim, uma mancha conforme mostrado na figura 3.2.

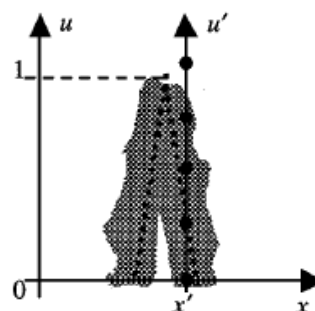
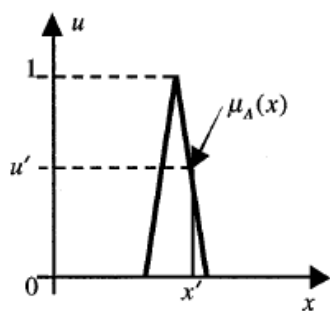


Figura 3.1 - Função de pertinência do tipo-1 **Figura 3.2** - Função de pertinência do tipo-1 manchada

Desta forma, considerando um valor específico para x , por exemplo, $x = x'$, este ponto não terá um único valor u' na função de pertinência, conforme podemos ver na figura 3.2. Ao invés disso, este ponto assume vários valores na linha vertical u' que intercepta esta mancha. Estes valores não precisam ser ponderados da mesma forma e, sendo assim, podemos aplicar uma distribuição para todos estes pontos. Fazendo isso para todos os pontos $x \in X$, nós criamos uma função de pertinência tridimensional (função de pertinência do tipo-2) que caracteriza um conjunto fuzzy do tipo-2.

Um conjunto fuzzy do tipo-2, denotado por \tilde{A} , é caracterizado por uma função de pertinência do tipo-2 $\mu_{\tilde{A}}(x,u)$, onde $x \in X$ e $u \in J_x \subseteq [0,1]$, onde J_x denota a função de pertinência primária e é definida como o domínio da função de pertinência secundária. Formalmente temos:

$$\tilde{A} = \{(x,u), \mu_{\tilde{A}}(x,u) \mid \forall x \in X, \forall u \in J_x \subseteq [0,1]\} \quad (3.1)$$

onde $0 \leq \mu_{\tilde{A}}(x,u) \leq 1$.

Nós também podemos representar \tilde{A} como:

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x,u) / (x,u) \quad J_x \subseteq [0,1] \quad (3.2)$$

onde \int denota a união sobre todo x e u admissíveis. Para o universo discreto o símbolo \int é substituído por \sum .

Na expressão 3.1, a expressão $\forall u \in J_x \subseteq [0,1]$ significa que ela é consistente com a restrição do tipo-1 $0 \leq \mu_A(x) \leq 1$ representada aqui pela expressão 2.2. Quando incertezas desaparecem, a função de pertinência do tipo-2 deve se reduzir para uma função de pertinência do tipo-1 e a variável u se torna igual a $\mu_A(x)$ e $0 \leq \mu_A(x) \leq 1$. A restrição $0 \leq \mu_{\tilde{A}}(x,u) \leq 1$ é consistente com o fato dos valores da função de pertinência estarem no intervalo $[0, 1]$.

A figura 3.3 ilustra $\mu_{\tilde{A}}(x,u)$ para X e U discretos, sendo $X = \{1, 2, 3, 4, 5\}$ e $U = \{0, 0.2, 0.4, 0.6, 0.8\}$. Para cada valor de $x = x'$, o plano 2-D formado pelos eixos u e $\mu_{\tilde{A}}(x',u)$ é chamado de *corte vertical* da função $\mu_{\tilde{A}}(x,u)$. Uma *função de pertinência secundária* é um corte vertical de $\mu_{\tilde{A}}(x = x', u)$ para $x \in X$ e $\forall u \in J_{x'} \subseteq [0,1]$. Desta forma, uma função de pertinência secundária pode ser definida como:

$$\mu_{\tilde{A}}(x' = x, u) \equiv \mu_{\tilde{A}}(x') = \int_{u \in J_{x'}} f_{x'}(u) / u \quad J_{x'} \subseteq [0,1] \quad (3.3)$$

onde $0 \leq f_{x'}(u) \leq 1$.

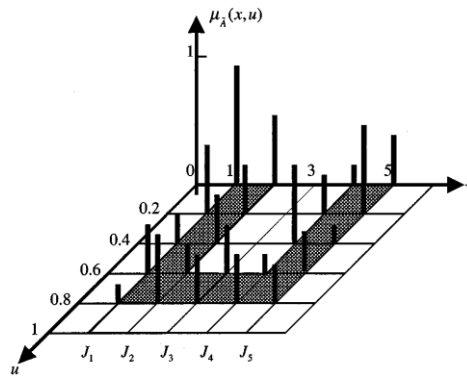


Figura 3.3 - Exemplo de função de pertinência do tipo-2

A função de pertinência secundária é um conjunto fuzzy do tipo-1 e é também chamada de *conjunto secundário*.

A função de pertinência do tipo-2 que é mostrada na figura 3.3 possui cinco cortes verticais. Um deles corresponde a $x=2$, mostrado na figura 3.4. A função de pertinência secundária para $x=2$ é $\mu_{\tilde{A}}(2)=0.5/0 + 0.3/0.2 + 0.3/0.4 + 0.15/0.6 + 0.4/0.8$. A função de pertinência secundária para $x=4$ é $\mu_{\tilde{A}}(4)=0.35/0 + 0.7/0.2 + 0.6/0.4 + 0.3/0.6 + 0.2/0.8$.

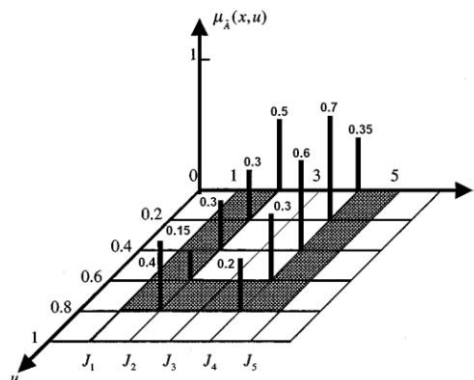


Figura 3.4 – Exemplo de corte vertical para a função de pertinência do tipo-2 da figura 3.3

Com base no conceito de conjuntos secundários, podemos interpretar um conjunto fuzzy do tipo-2 como sendo a união de todos os conjuntos secundários no universo X . Podemos, então, expressar \tilde{A} como:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid \forall x \in X\} \tag{3.4}$$

ou, ainda, como:

$$\tilde{A} = \int_{x \in X} \mu_{\tilde{A}}(x) / x = \int_{x \in X} \left[\int_{u \in J_x} f_x(u) / u \right] / x \quad J_x \subseteq [0,1] \tag{3.5}$$

O domínio de uma função de pertinência secundária é chamada de *pertinência primária* de x . Na expressão 3.5, J_x é a pertinência primária de x , onde $J_x \subseteq [0,1]$ para $x \in X$. Na figura 3.3, a pertinência primária de x , considerando $\mu_{\tilde{A}}(x,u) \neq 0$, é dada por:

$$J_1 = J_2 = J_4 = J_5 = \{0, 0.2, 0.4, 0.6, 0.8\} \quad e \quad J_3 = \{0.6, 0.8\}$$

A *amplitude* de uma função de pertinência secundária é chamada de *grau secundário*. Na expressão 3.5, $f_x(u)$ é um grau secundário. Na expressão 3.1, $\mu_{\tilde{A}}(x',u')$, $x' \in X$, $u' \in J_{x'}$ é um grau secundário.

Se X e J_x são ambos discretos, então o lado mais à direita de 3.5 pode ser expresso como:

$$\begin{aligned} \tilde{A} &= \sum_{x \in X} \left[\sum_{u \in J_x} f_x(u)/u \right] / x \\ &= \sum_{i=1}^N \left[\sum_{u \in J_{x_i}} f_{x_i}(u)/u \right] / x_i \\ &= \left[\sum_{k=1}^{M_1} f_{x_1}(u_{1k})/u_{1k} \right] / x_1 + \dots + \left[\sum_{k=1}^{M_N} f_{x_N}(u_{Nk})/u_{Nk} \right] / x_N \end{aligned} \quad (3.6)$$

Nesta expressão, o operador $+$ denota união, sendo que o universo X foi discretizado em N valores e para cada um desses valores x_i o universo U foi discretizado em M_i valores.

A incerteza na pertinência primária de um conjunto fuzzy do tipo-2, \tilde{A} , consiste em uma região limitada denominada *mancha de incerteza* (*footprint of uncertainty*) ou simplesmente FOU. Ela é a união de todas as pertinências primárias e é definida como:

$$FOU(\tilde{A}) = \bigcup_{x \in X} J_x \quad (3.7)$$

A mancha de incerteza representa a incerteza inerente em uma função de pertinência do tipo-2. A FOU descreve o domínio inteiro que suporta todos os graus secundários de uma função de pertinência do tipo-2 e permite representar o conjunto

fuzzy do tipo-2 graficamente em duas dimensões ao invés de três. A mancha de incerteza indica que existe uma distribuição disposta acima desta mancha; esta distribuição é a terceira dimensão dos conjuntos fuzzy do tipo-2. A aparência desta função de pertinência do tipo-2 em duas dimensões vai depender da escolha feita para os graus secundários. A região sombreada na figura 3.4 é um exemplo de FOU.

Quando a função de pertinência secundária é um *conjunto intervalar*, ou seja, quando o grau de pertinência secundário é 0 ou 1, o conjunto fuzzy do tipo-2 é chamado de *conjunto fuzzy intervalar do tipo-2*. Funções de pertinência secundária intervalares refletem uma incerteza uniforme na função de pertinência primária. A grande vantagem deste tipo de função de pertinência é que sua computação torna-se significativamente simplificada.

3.2 Tipos de conjuntos Fuzzy do tipo-2

Os conjuntos fuzzy do tipo-2 e suas respectivas FOUs mais utilizadas são (LIANG; MENDEL, 2000):

Função de Pertinência Primária Gaussiana com Incerteza no valor modal: a função de pertinência primária Gaussiana tem um desvio padrão σ fixo e uma incerteza na média que assume valores no intervalo $[m_1, m_2]$, isto é:

$$\mu_A(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] \quad m \in [m_1, m_2] \quad (3.8)$$

Assim, construindo-se o gráfico de $\mu_A(x)$ para diferentes valores de m , obtêm-se diferentes curvas de pertinência. Isto é ilustrado na figura 3.5, onde o FOU uniforme denota conjuntos intervalares para as funções de pertinência secundárias.

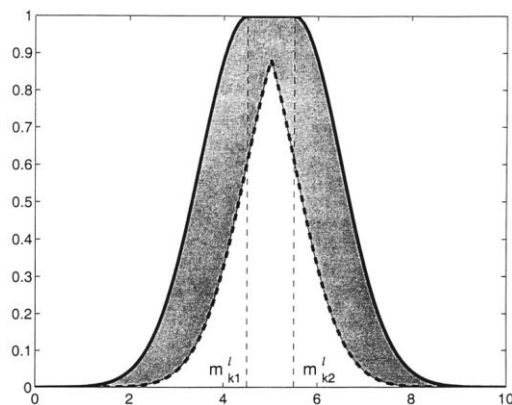


Figura 3.5 - FOU para a função de pertinência primária Gaussiana com incerteza no valor modal.

Função de pertinência primária Gaussiana com incerteza na dispersão: a função de pertinência primária Gaussiana tem uma média m fixa e uma incerteza na dispersão que assume valores em $[\sigma_1, \sigma_2]$, isto é:

$$\mu_A(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] \quad \sigma \in [\sigma_1, \sigma_2] \quad (3.9)$$

Fazendo a ilustração gráfica de $\mu_A(x)$ para diferentes valores de σ resulta em diferentes curvas de pertinência. Isto é ilustrado na figura 3.6, onde novamente o FOU uniforme denota conjuntos intervalares para as funções de pertinência secundárias.

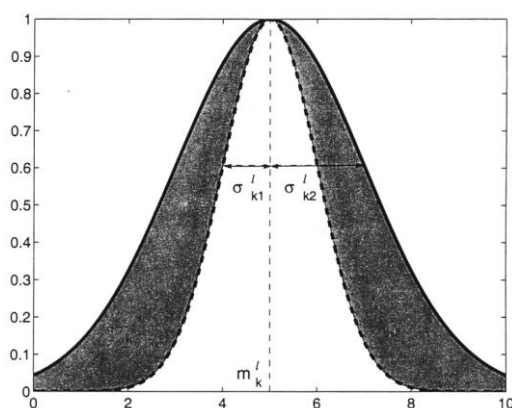


Figura 3.6 - FOU para a função de pertinência primária Gaussiana com incerteza na dispersão.

Funções de pertinência inferior e superior.

As funções de pertinência superior e inferior são duas funções de pertinência do tipo-1 que limitam a FOU de um conjunto fuzzy \tilde{A} do tipo-2. A função de pertinência superior é associada ao limite superior de $FOU(\tilde{A})$, e é denotada por $\bar{\mu}_A(x), \forall x \in X$. A função de pertinência superior é dada por:

$$\bar{\mu}_A(x) \equiv \overline{FOU(\tilde{A})} = \bigcup_{x \in X} \bar{J}_x \quad (3.10)$$

A função de pertinência inferior é associada ao limite inferior de $FOU(\tilde{A})$, e é denotada por $\underline{\mu}_A(x), \forall x \in X$. A função de pertinência inferior é dada por:

$$\underline{\mu}_A(x) \equiv \underline{FOU(\tilde{A})} = \bigcup_{x \in X} \underline{J}_x \quad (3.11)$$

onde \bar{J}_x e \underline{J}_x representam o limite superior e inferior de J_x , respectivamente.

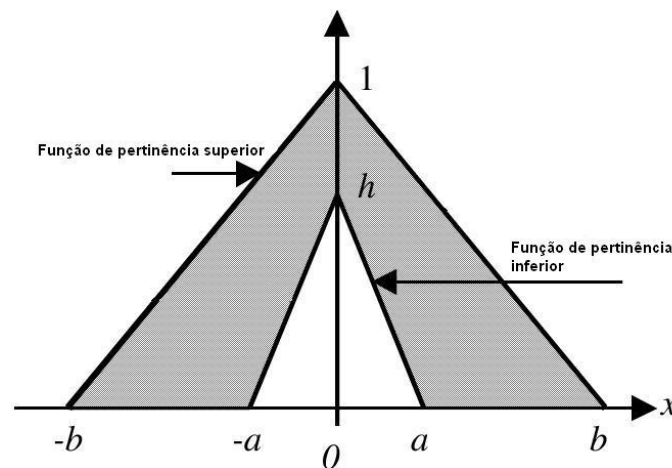


Figura 3.7 - Funções de pertinência superior e inferior do conjunto fuzzy do tipo-2.

Representação em termos das funções de pertinência Superior e Inferior

Um conjunto fuzzy do tipo-2 pode ser representado pelas suas funções de pertinência superior e inferior da seguinte forma:

$$\tilde{A} = \mu_{\tilde{A}}(x, u) = \int_{x \in X} \mu_{\tilde{A}}(x) / x = \int_{x \in X} \left[\int_{u \in J_x} f_x(u) / u \right] / x \quad (3.12)$$

$$\tilde{A} = \mu_{\tilde{A}}(x, u) = \int_{x \in X} \left[\int_{u \in [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]} f_x(u) / u \right] / x \quad (3.13)$$

Das equações 3.12 e 3.13 pode-se observar que a função de pertinência secundária $\mu_{\tilde{A}}(x)$ pode ser representada em termos das suas funções de pertinência superior e inferior como:

$$\mu_{\tilde{A}}(x) = \int_{u \in [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]} f_x(u) / u \quad (3.14)$$

No caso de conjuntos fuzzy intervalares do tipo-2, tem-se que $f_x(u) = 1$ (grau secundário). Portanto, \tilde{A} é representado pelas suas funções de pertinência superior e inferior da seguinte forma:

$$\tilde{A} = \mu_{\tilde{A}}(x, u) = \int_{x \in X} \left[\int_{u \in J_x} 1 / u \right] / x = \int_{x \in X} \left[\int_{u \in [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]} 1 / u \right] / x \quad (3.15)$$

3.3 Operações de conjuntos fuzzy do tipo-2

Nesta seção serão descritas as duas principais operações de conjuntos fuzzy do tipo-2: a união e a intersecção.

Para computar a união e a intersecção de conjuntos fuzzy do tipo-2 é necessário estender as operações binárias mínimo (ou produto), que retorna o menor dentre dois números e máximo que retorna o maior dentre dois números.

3.3.1 União de conjuntos fuzzy do tipo-2

Assim como a união entre os conjuntos fuzzy do tipo-1 A e B resulta em um conjunto fuzzy do tipo-1, a união entre os conjuntos fuzzy do tipo-2 \tilde{A} e \tilde{B} resulta em um conjunto fuzzy do tipo-2. Mais formalmente nós temos a seguinte expressão:

$$\tilde{A} \cup \tilde{B} = \mu_{\tilde{A} \cup \tilde{B}}(x) = \mu_{\tilde{A}}(x) \amalg \mu_{\tilde{B}}(x) = \int_{u \in J_x^u} \int_{w \in J_x^w} (f_x(u) \wedge g_x(w)) / (u \vee w) \quad (3.16)$$

onde \wedge é uma t-norma, \vee é uma s-norma e \amalg é chamado de operador *join*.

A união entre duas funções de pertinência secundária deve ser aplicada para todo par possível de pertinência primária. Considere o seguinte exemplo para um $x = x'$ particular cuja função de pertinência secundária nos conjuntos fuzzy do tipo-2 \tilde{A} e \tilde{B} são $\mu_{\tilde{A}}(x) = 0.5/0.1 + 0.8/0.2$ e $\mu_{\tilde{B}}(x) = 0.4/0.5 + 0.9/0.9$ a união entre eles será:

$$\begin{aligned} \mu_{\tilde{A}}(x) \amalg \mu_{\tilde{B}}(x) &= (0.5/0.1 + 0.8/0.2) \amalg (0.4/0.5 + 0.9/0.9) \\ &= (0.5 \wedge 0.4) / (0.1 \vee 0.5) + (0.5 \wedge 0.9) / (0.1 \vee 0.9) + \\ &\quad (0.8 \wedge 0.4) / (0.2 \vee 0.5) + (0.8 \wedge 0.9) / (0.2 \vee 0.9) \\ &= 0.4/0.5 + 0.5/0.5 + 0.4/0.5 + 0.8/0.9 \end{aligned}$$

3.3.2 Intersecção de conjuntos fuzzy do tipo-2

Assim como a intersecção entre os conjuntos fuzzy do tipo-1 A e B resulta em um conjunto fuzzy do tipo-1, a intersecção entre os conjuntos fuzzy do tipo-2 \tilde{A} e \tilde{B} também resulta em um conjunto fuzzy do tipo-2. Mais formalmente nós temos a seguinte expressão:

$$\tilde{A} \cap \tilde{B} = \mu_{\tilde{A} \cap \tilde{B}}(x) = \mu_{\tilde{A}}(x) \amalg \mu_{\tilde{B}}(x) = \int_{u \in J_x^u} \int_{w \in J_x^w} (f_x(u) \wedge g_x(w)) / (u \wedge w) \quad (3.17)$$

onde \wedge é uma t-norma e \amalg é chamado de operador *meet*.

A intersecção entre duas funções de pertinência secundária deve ser aplicada para todo par possível de pertinência primária. Como exemplo, considere os mesmos conjuntos fuzzy do tipo-2 do exemplo anterior. A intersecção entre eles é feita da seguinte forma:

$$\begin{aligned} \mu_{\tilde{A}}(x) \amalg \mu_{\tilde{B}}(x) &= (0.5/0.1 + 0.8/0.2) \amalg (0.4/0.5 + 0.9/0.9) \\ &= (0.5 \wedge 0.4) / (0.1 \wedge 0.5) + (0.5 \wedge 0.9) / (0.1 \wedge 0.9) + \\ &\quad (0.8 \wedge 0.4) / (0.2 \wedge 0.5) + (0.8 \wedge 0.9) / (0.2 \wedge 0.9) \\ &= 0.4/0.1 + 0.5/0.1 + 0.4/0.2 + 0.8/0.2 \end{aligned}$$

3.4 Sistemas de inferência fuzzy do tipo-2

Na literatura existem diferentes tipos de sistemas de inferência fuzzy do tipo-2. Entre os principais, tem-se:

- SIF do tipo-2 Mamdani com entradas numéricas;
- SIF do tipo-2 TSK com entradas numéricas;

A figura 3.8 mostra um sistema genérico de inferência fuzzy do tipo-2 (KARNIK; MENDEL; LIANG, 1999). Os SIF do tipo-2 apresentam quatro componentes: fuzzificador, base de regras, mecanismo de inferência, e módulo processador de saída. O módulo processador de saída é composto pelos módulos redutor de tipo e de defuzzificação.

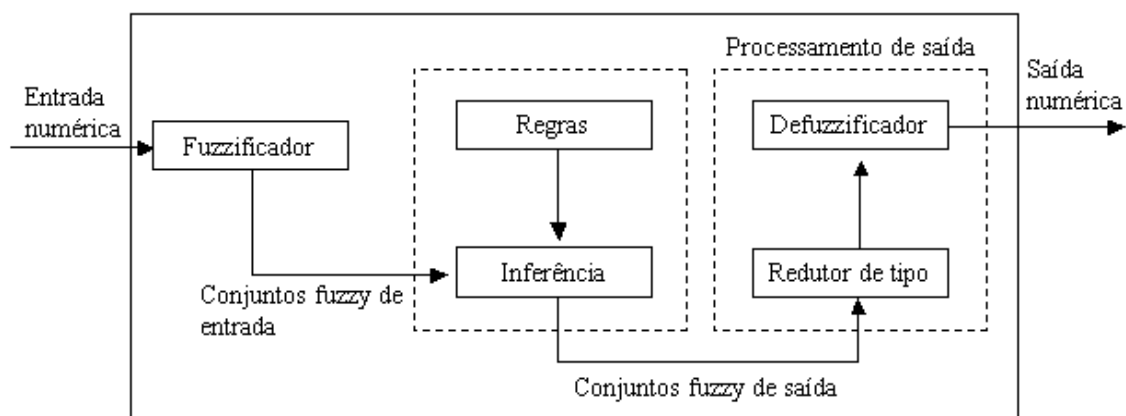


Figura 3.8 - Sistema genérico de inferência fuzzy do tipo-2, baseado em (KARNIK; MENDEL; LIANG, 1999)

Na figura 3.8, entradas numéricas são fuzzificadas em conjuntos fuzzy do tipo-2 de entrada, os quais ativam o bloco de inferência. O bloco de inferência combina regras do tipo-2 e mapeia conjuntos fuzzy do tipo-2 de entrada em conjuntos fuzzy do tipo-2 de saída. O conjunto fuzzy do tipo-2, resultante do bloco de inferência, é resultado da combinação de vários conjuntos de saída, onde cada conjunto de saída, da mesma forma que em sistemas de inferência fuzzy do tipo-1, é o resultado da ativação de uma regra.

Em um SIF genérico do tipo-2 o bloco processador de saída apresenta dois componentes. O primeiro, chamado redutor de tipo, mapeia o conjunto fuzzy do tipo-2 de saída do bloco de inferência em um conjunto fuzzy do tipo-1. O segundo bloco, chamado defuzzificador, executa a defuzzificação deste último conjunto.

3.4.1 Regras

Em um SIF genérico do tipo-2 com p entradas $x_1 \in X_1, \dots, x_p \in X_p$ e uma saída $y \in Y$, suponha M regras, então, e l -ésima regra é dada por:

$$R^l : SE \ x_1 \ \acute{e} \ \tilde{F}_1^l \ e \ \dots \ e \ x_p \ \acute{e} \ \tilde{F}_p^l, \ ENT\tilde{A}O \ y \ \acute{e} \ \tilde{G}^l \quad (3.18)$$

onde $l = 1, \dots, M$, sendo X_i o domínio da variável x_i e \tilde{F}_i^l é o valor linguístico.

Nesta regra do tipo-2, todos os conjuntos envolvidos são do tipo-2.

3.4.2 Mecanismo de inferência

O mecanismo de inferência combina regras e mapeia conjuntos fuzzy do tipo-2 de entrada em conjuntos fuzzy do tipo-2 de saída. Para este fim, calcula uniões e interseções de conjuntos fuzzy do tipo-2.

Em um SIF do tipo-2 cada regra é interpretada como uma relação fuzzy do tipo-2, isto é:

$$R^l : \tilde{F}_1^l \times \dots \times \tilde{F}_p^l \rightarrow \tilde{G}^l = \tilde{A}^l \rightarrow \tilde{G}^l \quad (3.19)$$

onde $l = 1, \dots, M$, $\tilde{F}_1^l \times \dots \times \tilde{F}_p^l = \tilde{A}^l$ e a função de pertinência da relação R^l é dada por:

$$\mu_{R^l}(\mathbf{x}, y) = \mu_{R^l}(x_1, \dots, x_p, y) = \mu_{\tilde{A}^l \rightarrow \tilde{G}^l}(\mathbf{x}, y) \quad (3.20)$$

onde (MENDEL, 2000):

$$\mu_{R^l}(\mathbf{x}, y) = \mu_{\tilde{F}_1^l}(x_1) \prod \dots \prod \mu_{\tilde{F}_p^l}(x_p) \prod \mu_{\tilde{G}^l}(y) = \left[\prod_{i=1}^p \mu_{\tilde{F}_i^l}(x_i) \right] \prod \mu_{\tilde{G}^l}(y) \quad (3.21)$$

\prod = operador *meet* (usando o mínimo ou o produto como t-norma)

p = número de variáveis de entrada.

De forma geral a entrada p -dimensional para a regra R^l é dada por um conjunto fuzzy do tipo-2 (representado por \tilde{A}_x), cuja função de pertinência é:

$$\mu_{\tilde{A}_x}(\mathbf{x}) = \mu_{F_1}(x_1) \prod \mu_{F_2}(x_2) \prod \dots \prod \mu_{F_p}(x_p) = \prod_{i=1}^p \mu_{F_i}(x_i) \quad (3.22)$$

onde $\tilde{F}_i (i=1, \dots, p)$ são rótulos (labels) dos conjuntos fuzzy do tipo-2 associados às entradas.

Cada regra R^l determina um conjunto fuzzy do tipo-2 da forma $\tilde{B}^l = \tilde{A}_x \circ R^l$.

Para um SIF do tipo-2 tem-se:

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{A}_x \circ R^l}(y) = \coprod_{x \in X} \left[\mu_{\tilde{A}_x}(\mathbf{x}) \prod \mu_{R^l}(\mathbf{x}, y) \right] \quad y \in Y \quad l = 1, \dots, M \quad (3.23)$$

onde: \coprod = operador *join* (união de conjuntos fuzzy do tipo-2) e

\prod = operador *meet* (usando o mínimo ou o produto como t-norma)

Substituindo as equações 3.21 e 3.23 obtêm-se:

$$\begin{aligned} \mu_{\tilde{B}^l}(y) &= \coprod_{x \in X} \left[\mu_{\tilde{A}_x}(\mathbf{x}) \prod \mu_{R^l}(\mathbf{x}, y) \right] = \coprod_{x \in X} \left\{ \prod_{i=1}^p \mu_{\tilde{X}_i}(x_i) \prod \left[\prod_{i=1}^p \mu_{\tilde{F}_i}(x_i) \prod \mu_{G^l}(y) \right] \right\} \\ \mu_{\tilde{B}^l}(y) &= \coprod_{x \in X} \left\{ \left[\prod_{i=1}^p \mu_{\tilde{X}_i}(x_i) \prod \mu_{\tilde{F}_i}(x_i) \right] \prod \mu_{G^l}(y) \right\} \end{aligned} \quad (3.24)$$

Da propriedade de associatividade do operador *meet* (usando o mínimo ou o produto para *t-norma*) tem-se:

$$\mu_{\tilde{B}^l}(y) = \mu_{G^l}(y) \prod \left\{ \left[\prod_{x_1 \in X_1} \mu_{\tilde{X}_1}(x_1) \prod \mu_{\tilde{F}_1}(x_1) \right] \prod \dots \prod \left[\prod_{x_p \in X_p} \mu_{\tilde{X}_p}(x_p) \prod \mu_{\tilde{F}_p}(x_p) \right] \right\} \quad (3.25)$$

$y \in Y$

A expressão 3.25 representa a relação entrada-saída entre os conjuntos fuzzy do tipo-2 que excitam uma regra no bloco de inferência e o conjunto fuzzy do tipo-2 na saída do bloco de inferência.

3.4.3 Fuzzificador

Conforme mencionado anteriormente, o fuzzificador usado em um SIF do tipo-2 Mamdani com entradas numéricas é o fuzzificador singleton do tipo-2. O fuzzificador mapeia o ponto $\mathbf{x} = (x_1, \dots, x_p)^T \in \mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_p \equiv \mathbf{X}$ em um conjunto fuzzy do tipo-2 \tilde{A}_x em \mathbf{X} .

Neste caso, \tilde{A}_x é um conjunto fuzzy do tipo-2 unitário, onde, $\mu_{\tilde{A}_x}(x) = 1/1$ para $x = x'$ e $\mu_{\tilde{A}_x}(x) = 1/0$ para $x \neq x'$. Como usamos funções de pertinência separáveis, então $\mu_{x_i}(x_i) = 1/1$ quando $x_i = x_i'$ e $\mu_{x_i}(x_i) = 1/0$ quando $x_i \neq x_i'$ para $i = 1, \dots, p$.

3.4.4 Processo de redução de tipo

Se o SIF do tipo-2 é visto como uma versão “perturbada” de um SIF do tipo-1 devido às incertezas nas suas funções de pertinência, então o conjunto tipo reduzido em um SIF do tipo-2 é uma representação das incertezas na saída *crisp* devido a essas “perturbações”. Portanto, alguma medida de largura do conjunto tipo reduzido pode ser usada para indicar a possível variação da saída *crisp* devido a essa “perturbação” (KARNIK; MENDEL, 1998)(MENDEL, 2001a)(GARIBALDI; MUSIKASUWAN; OZEN, 2005).

3.4.5 Defuzzificação

O segundo módulo do Processador de Saída é a Defuzzificação. Neste processo, o conjunto tipo reduzido é defuzzificado para conseguir a saída precisa (*crisp*) do SIF do tipo-2. O método mais natural para efetuar a defuzzificação parece ser calculando o centróide do conjunto tipo reduzido. Maiores detalhes são encontrados em (MENDEL, 2001), (COUPLAND; JOHN, 2005).

3.5 Sistemas de inferência fuzzy intervalares do tipo-2 Mamdani

Em geral, os SIF do tipo-2 são complexos devido aos seguintes aspectos:

- a complexidade computacional das operações *meet*, que deve ser executada para cada regra ativada;
- a execução do processo de redução de tipo é bastante complexo computacionalmente e
- não se tem uma base racional para se escolher a forma das funções de pertinência secundárias.

Para simplificar a complexidade computacional dos SIF do tipo-2, usam-se conjuntos fuzzy intervalares do tipo-2. Assim, o cálculo das operações *meet*, *join* e do processo de redução de tipo fica simplificado. Conjuntos intervalares do tipo-2 permitem que as suas funções de pertinência secundárias possam distribuir uniformemente a incerteza nas pertinências primárias. Portanto, a complicação em relação à escolha da forma das funções de pertinência secundárias fica solucionada.

A seguir serão descritos o bloco de inferência e os processos redução de tipo e defuzzificação dos SIF intervalares do tipo-2 Mamdani com entradas numéricas.

Bloco de inferência

Em um SIF intervalar do tipo-2 Mamdani com entradas numéricas, com *meet* calculado usando o mínimo ou o produto como t-norma, tem-se o resultado das operações da entrada e os antecedentes é um conjunto intervalar do tipo-1 contido no conjunto

ativado $\prod_{i=1}^p \mu_{\tilde{F}_i^l}(x_i) \equiv F^l(x')$ dado por (MENDEL, 2007):

$$F^l(x') = \left[\underline{f}^l(x'), \overline{f}^l(x') \right] \equiv \left[\underline{f}^l, \overline{f}^l \right] \quad (3.26)$$

$$\underline{f}^l(x') = \underline{\mu}_{\tilde{F}_1}(x_1) \wedge \cdots \wedge \underline{\mu}_{\tilde{F}_p}(x_p) \quad (3.27)$$

$$\overline{f}^l(x') = \overline{\mu}_{\tilde{F}_1}(x_1) \wedge \cdots \wedge \overline{\mu}_{\tilde{F}_p}(x_p) \quad (3.28)$$

onde $F^l(x')$ é a força de ativação ou conjunto intervalar ativado, $\underline{f}^l(x')$ e $\overline{f}^l(x')$ são o ponto limite esquerdo e direito, respectivamente, do domínio do conjunto intervalar ativado. O grau de pertinência de x_1 na função de pertinência superior do conjunto antecedente F_1 é dado por $\overline{\mu}_{\tilde{F}_1}(x_1')$ e o grau de pertinência de x_1 na função de pertinência inferior do conjunto antecedente F_1 é dado por $\underline{\mu}_{\tilde{F}_1}(x_1')$.

Usando a representação em termos das funções de pertinência Superior e Inferior, o conjunto consequente de saída da l -ésima regra ativada da expressão 3.25, $\mu_{B^l}^-(y)$, é um conjunto fuzzy do tipo-1 dado por (MENDEL, 2007):

$$\mu_{B^l}^-(y) = \int_{b^l \in [\underline{f}^l * \underline{\mu}_{\tilde{G}^l}(y), \overline{f}^l * \overline{\mu}_{\tilde{G}^l}(y)]} 1/b^l \quad y \in Y \quad (3.29)$$

onde $\underline{\mu}_{\tilde{G}^l}(y)$ é o grau de pertinência inferior de $\mu_{G^l}^-(y)$ e $\overline{\mu}_{\tilde{G}^l}(y)$ é o grau de pertinência superior de $\mu_{G^l}^-(y)$.

Supondo que N regras das M regras existentes no SIF do tipo-2 são ativadas, onde, $N \leq M$, e supondo que o conjunto fuzzy combinado do tipo-1 de saída é obtido combinando os conjuntos consequentes de saída ativados, isto é, $\mu_B^-(y) = \prod_{l=1}^N \mu_{B^l}^-(y)$, então:

$$\mu_B^-(y) = \int_{b \in [\underline{f}^1 * \underline{\mu}_{\tilde{G}^1}(y) \vee \dots \vee \underline{f}^N * \underline{\mu}_{\tilde{G}^N}(y), \overline{f}^1 * \overline{\mu}_{\tilde{G}^1}(y) \vee \dots \vee \overline{f}^N * \overline{\mu}_{\tilde{G}^N}(y)]} 1/b \quad y \in Y \quad (3.30)$$

onde:

\vee = s-norma máximo

$$\underline{b}^1(y) \vee \underline{b}^2(y) \vee \dots \vee \underline{b}^N(y) = \underline{b}(y)$$

$$\overline{b}^1(y) \vee \overline{b}^2(y) \vee \dots \vee \overline{b}^N(y) = \overline{b}(y)$$

Redução de tipo e defuzzificador

Em um SIF intervalar do tipo-2 Mamdani com entradas numéricas, o conjunto tipo reduzido é um conjunto intervalar dado por: $Y_{TR} = [y_l, y_r]$. Num SIF intervalar do tipo-2 Mamdani, o processo de redução de tipo centro de conjuntos é o mais usado, neste tipo de SIF do tipo-2 a saída do processo de redução de tipo centro de conjuntos, $Y_{COS}(x)$, é dada por (KARNIK; MENDEL, 1998).

$$Y_{COS}(x) = [y_l, y_r] = \int_{y^1 \in [y_l^1, y_r^1]} \cdots \int_{y^M \in [y_l^M, y_r^M]} \int_{f^1 \in [f^1, \bar{f}^1]} \cdots \int_{f^M \in [f^M, \bar{f}^M]} 1 / \frac{\sum_{i=1}^M f^i y^i}{\sum_{i=1}^M f^i} \quad (3.31)$$

onde:

$$y_l = \frac{\sum_{i=1}^M f_l^i y_l^i}{\sum_{i=1}^M f_l^i} \quad e \quad y_r = \frac{\sum_{i=1}^M f_r^i y_r^i}{\sum_{i=1}^M f_r^i} \quad (3.32)$$

onde:

$Y_{COS}(x)$ = conjunto intervalar de pontos limites $[y_l, y_r]$.

$Y^i = [y_l^i, y_r^i]$ = centróide do conjunto consequente intervalar do tipo-2, \tilde{G}^i . Estes valores devem ser calculados antes de se calcular $Y_{COS}(x)$.

M = número de regras.

f_r^i, y_r^i = valores de f^i, y^i que estão associados a y_r respectivamente.

f_l^i, y_l^i = valores de f^i, y^i que estão associados a y_l respectivamente.

Sendo que $f^i \in F^i = [f^i, \bar{f}^i]; y^i \in Y^i$

3.6 Considerações Finais

Neste capítulo foram apresentados os principais conceitos sobre conjunto fuzzy do tipo-2, conjunto fuzzy intervalar do tipo-2 e sistemas de inferência fuzzy do tipo-2. No próximo capítulo serão apresentados os conceitos básicos sobre AGs.

4. Algoritmos Genéticos

Neste capítulo serão apresentados os conceitos básicos do modelo computacional mais conhecido e utilizado da Computação Evolutiva: os AGs. Por ser um assunto bastante conhecido, serão considerados os principais pontos necessários para a construção de AGs. Serão apresentados também os SFGs, que é a abordagem utilizada no trabalho.

4.1 Funcionamento dos Algoritmos Genéticos

Os AGs fazem parte de uma área particular da Computação Evolutiva que tem como base as estratégias evolutivas formuladas por Darwin em sua Teoria da Evolução Natural (DARWIN, 1872).

Os AGs são frequentemente usados em problemas de busca, tendo encontrado ampla aplicação em diversas áreas. Estes algoritmos utilizam os princípios de seleção natural análogo ao modelo biológico, e têm provado ser uma abordagem de sucesso para muitos problemas, incluindo o problema de otimização (DAVIS, 1998)(GOLDBERG, 1989). Eles utilizam uma estratégia de busca probabilística, paralela e estruturada visando encontrar a melhor solução dentro de um espaço de soluções.

O funcionamento de um AG dá-se a partir do processo de *inicialização*, onde é criada uma *população* de indivíduos representando soluções iniciais que são submetidas ao processo de evolução que acontece pelas seguintes etapas:

- *Avaliação*: a partir de uma *função de aptidão* (fitness) é atribuído um valor ao *indivíduo* que representa quão bem este indivíduo resolve o problema em questão. Um indivíduo é a unidade fundamental dos AG, pois ele representa uma possível solução para o problema;
- *Seleção*: nesta etapa são escolhidos os indivíduos que irão se reproduzir para gerar uma nova população. Quanto maior a aptidão do indivíduo, maior é a sua probabilidade de ser escolhido;
- *Cruzamento*: feita a escolha dos indivíduos, suas características são recombinadas, dando origem a novos indivíduos;

- **Mutação:** a mutação altera algumas características dos indivíduos resultantes do cruzamento. Isso é feito para manter a variedade da população e evitar os máximos e os mínimos locais.
- **Atualização:** neste ponto todos os indivíduos criados são inseridos na população;
- **Encerramento:** caso a condição de parada do AG seja satisfeita, a execução do AG é encerrada. Caso contrário, o fluxo é retornado para a etapa de avaliação.

Apesar da aleatoriedade ser uma característica intrínseca aos AG, eles fazem uma busca direcionada. Os novos indivíduos gerados possuem características da população anterior, e essas características são usadas como histórico que direcionam a busca para um resultado melhor que o resultado gerado pela população anterior.

Abaixo podemos ver um pseudocódigo que representa o fluxo básico de um AG, onde $P(t)$ significa “população da geração t ”.

Início

$t = 0$

inicializa $P(t)$

avalia $P(t)$

Enquanto (condição_de_parada = falso)

$t = t + 1$

seleciona $P(t)$ de $P(t - 1)$

aplica_cruzamento $P(t)$

aplica_mutação $P(t)$

avalia $P(t)$

Fim-enquanto

end

Algoritmo 1 – Pseudocódigo de um AG

4.2 Representação

Uma possível solução para um determinado problema é chamada de *indivíduo* ou simplesmente *cromossomo*. O cromossomo pode ser representado de diversas formas, sendo que as mais comuns são: codificação binária, inteira ou real. A escolha da representação do cromossomo é uma das partes mais importantes durante o projeto de

um AG, visto que sua representação é um dos principais fatores responsáveis pelo seu desempenho. Exemplos de cromossomos são mostrados na figura 4.1.

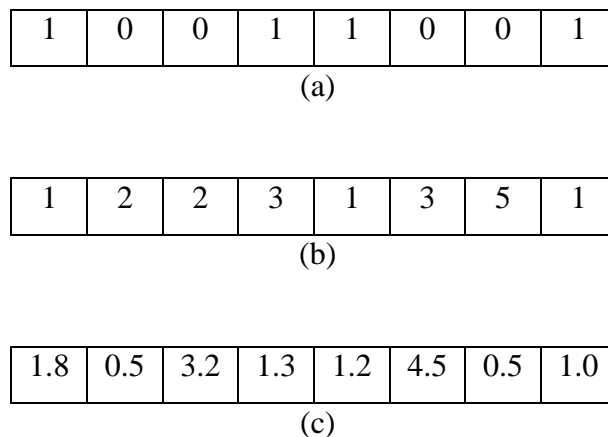


Figura 4.1 - Cromossomo binário (a), cromossomo inteiro (b) e cromossomo real (c).

A representação de cada parâmetro, de acordo com a representação utilizada, é chamada de *gene*. No exemplo da figura 4.1 cada posição do vetor é um gene.

4.2.1 Estrutura dos AG

A população de um AG é um conjunto de indivíduos que são soluções candidatas para o problema específico. O tamanho da população influencia diretamente o desempenho do AG. Se a população for pequena, pode-se não manter a diversidade entre os indivíduos, o que influencia a convergência a uma boa solução por cobrir uma pequena parte do espaço de busca do problema. Por outro lado, se a população for muito grande, a eficiência do AG ficará comprometida, haja vista que terá um número maior de indivíduos para se calcular a função de aptidão (GOLDBERG, 1989)(MICHALEWICZ, 1996).

Inicialização

É na inicialização que é gerada a população inicial em que serão aplicadas todas as ações subsequentes do processo. Geralmente a população inicial é gerada de forma aleatória. Isso permite que a diversidade da população seja mantida.

As funções de inicialização mais utilizadas são (GOLDBERG, 1989)(GEYER-SCHULTZ, 1997):

- **Inicialização aleatória uniforme:** cada gene do cromossomo receberá aleatoriamente um valor dentre o conjunto de possíveis valores para este gene;
- **Inicialização aleatória não-uniforme:** semelhante ao método acima. Porém, um valor de cada gene tende a ser escolhido com mais frequência.

Avaliação (Fitness)

É nesta fase que é dado o primeiro passo para a seleção. Cada indivíduo recebe uma avaliação de acordo com o seu grau de aptidão, ou seja, mede-se quão bom é o indivíduo para solucionar o problema em questão.

Na maioria dos casos, fazer o cálculo exato do grau de aptidão dos indivíduos não é uma tarefa fácil, e levando-se em conta que esta tarefa é executada para cada indivíduo a cada geração, podemos dizer que o seu custo computacional é relativamente alto.

Seleção

Na etapa de seleção, os indivíduos são selecionados para a próxima etapa, o cruzamento. Utiliza-se o grau de aptidão de cada indivíduo para fazer um sorteio no qual os indivíduos mais aptos possuem maior probabilidade de serem escolhidos.

A seguir são listados alguns métodos de seleção.

- **Roleta:** Neste método, os indivíduos de uma geração (ou população) são selecionados para a próxima geração utilizando uma roleta. Cada indivíduo da população é representado na roleta por uma fatia proporcional ao seu índice de aptidão. Assim, indivíduos com maior aptidão ocupam fatias maiores da roleta, enquanto indivíduos de menor aptidão ocupam fatias menores. Na Figura 4.2 é apresentada a criação de uma roleta a partir dos

valores de aptidão dos indivíduos de uma população. Para a seleção dos indivíduos, a roleta é girada N vezes, onde N é número da população inicial. A cada vez que a roleta parar de girar, o cromossomo selecionado pelo marcador será copiado para a próxima geração. Cromossomos com maior espaço na roleta terão maior chance de serem selecionados.

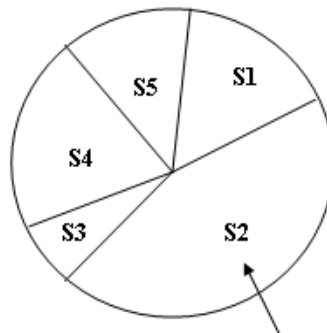


Figura 4.2 - Método de seleção por roleta

- **Torneio:** Quando este método é utilizado, um número de n indivíduos da população é escolhido de modo aleatório e com mesma probabilidade. O cromossomo com maior aptidão dentre estes n cromossomos é selecionado para a população intermediária. Este processo se repetirá até que a população intermediária seja preenchida.

Podemos, ainda, adicionar a esses métodos uma técnica chamada *elitismo*. Nesta técnica todos os indivíduos são ordenados de acordo com o seu valor de aptidão e os indivíduos que ficarem abaixo do valor médio são descartados. Esta técnica resulta em uma população que terá indivíduos com valor de aptidão acima da média da população anterior. Esta técnica também garante que o melhor indivíduo obtido na geração atual estará presente na próxima geração que está sendo gerada, preservando suas características genéticas.

Cruzamento (Crossover)

Depois que os indivíduos são selecionados, eles passam com uma probabilidade pré-estabelecida pelo processo de cruzamento. Essa probabilidade é chamada de *taxa de cruzamento*. Neste processo são selecionados dois indivíduos para o cruzamento, gerando dois novos indivíduos que irão compor a próxima geração.

Dentre os métodos de cruzamento podemos destacar (GOLDBERG, 1989):

- Cruzamento de um ponto:** neste método, é escolhido um ponto de corte p qualquer onde $1 < p < w$, onde w é o comprimento do cromossomo. O primeiro filho receberá os mesmos genes do pai 1 compreendidos entre o índice 1 e p . Os genes seguintes serão preenchidos com os genes do pai 2 compreendidos entre o índice $p+1$ e w . O segundo filho terá o preenchimento pela forma contrária, ou seja, receberá os mesmos genes do pai 2 compreendidos entre o índice 1 e p . Os genes seguintes serão preenchidos com os genes do pai compreendidos entre o índice $p+1$ e w . A figura 4.3 ilustra um exemplo deste tipo de cruzamento.

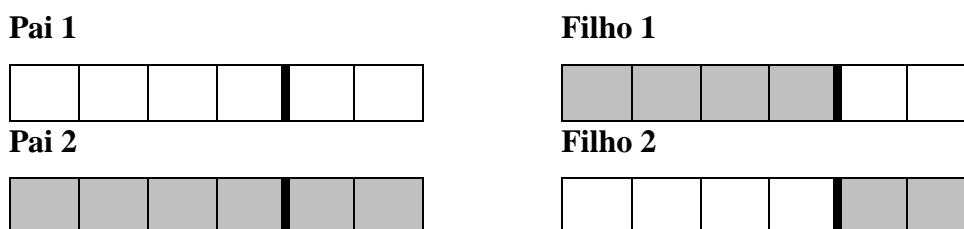


Figura 4.3 - Cruzamento de um ponto

- Cruzamento de dois pontos:** no cruzamento de dois pontos, são escolhidos aleatoriamente dois pontos, $p1$ e $p2$, tal que $p1 > p2$. O primeiro filho terá os mesmos genes do cromossomo do pai 1, entretanto, a região que compreende $p1$ e $p2$ terão os genes da pai 2. O segundo filho terá os mesmos genes do cromossomo da pai 2, entretanto, a região que compreende $p1$ e $p2$ terão os genes do pai 1. A figura 4.4 ilustra um exemplo deste tipo de cruzamento.

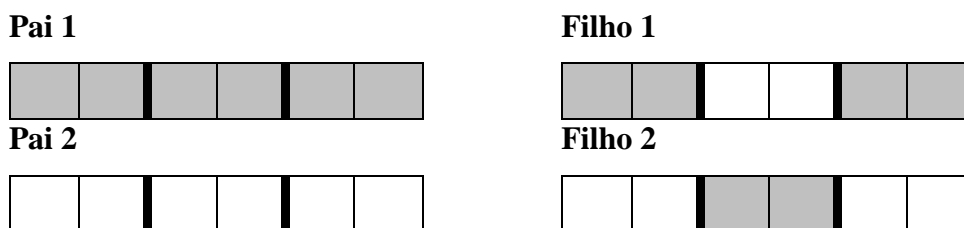


Figura 4.4 - Cruzamento de dois pontos

- Cruzamento uniforme:** neste cruzamento, é utilizada uma máscara binária. No caso do primeiro filho, a posição da máscara que tiver o valor 0 terá o gene igual ao do pai e quando o valor for 1 terá o gene da mãe. Para o segundo filho, a

posição da máscara que tiver o valor 0 terá o gene igual ao da mãe e quando o valor for 1 terá o gene do pai. A figura 4.5 ilustra um exemplo deste tipo de cruzamento.

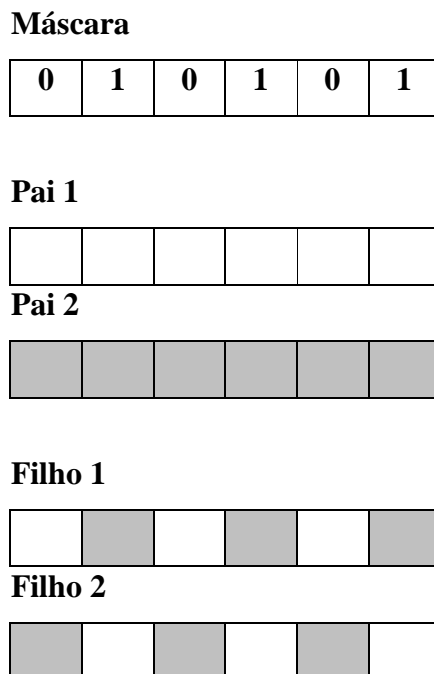


Figura 4.5 - Cruzamento uniforme

Mutação

A mutação é aplicada sobre os indivíduos resultantes do cruzamento, com uma taxa de probabilidade pré-definida chamada *taxa de mutação*. A mutação nada mais é do que algum tipo de mudança que deve ser realizada sobre um gene. Essa mudança depende de como o gene foi projetado. Em um cromossomo com codificação binária, por exemplo, a mutação alteraria um gene com valor 1 para valor 0 e vice versa. Em um cromossomo com codificação real ou inteira, o gene é alterado trocando o valor atual do gene por algum outro valor dentre os possíveis valores válidos para o gene em questão.

A mutação é um processo que garante a exploração de diversas alternativas, mantendo um nível mínimo de abrangência nas buscas.

Atualização

Nesta fase, os indivíduos resultantes do cruzamento e da mutação são inseridos na nova população que está sendo gerada. Essa inserção depende da forma como o AG é

implementado. Em algumas implementações, a população mantém sempre o mesmo número de indivíduos, e é sempre gerado o mesmo número de indivíduos que substituem por completo os indivíduos da população anterior. Em outras implementações, o tamanho da população pode ser menor a cada geração e o critério de inserção pode variar, como por exemplo, inserir apenas indivíduos na nova população se possuírem aptidão maior que o cromossomo a ser substituído.

Finalização

Nesta etapa verifica-se se o critério de parada é satisfeito. Se for satisfeito, a evolução para de prosseguir e o AG é finalizado, caso contrário o fluxo é desviado para o processo de avaliação. Podemos tomar como um critério de parada, dentre vários, o número de gerações alcançadas ou o grau de convergência da população atual, ou seja, o grau de proximidade dos valores da avaliação de cada indivíduo das populações geradas.

4.3 Sistemas Fuzzy Genéticos

Nos últimos anos, vários artigos e aplicações combinando conceitos fuzzy e AG surgiram na literatura e existe uma preocupação crescente sobre a integração entre eles. Muitas destas publicações exploram o uso do AG para a geração automática de um Sistema Fuzzy. A essas abordagens deu-se o nome de Sistemas Fuzzy Genéticos (SFG) (CORDON et al., 2004). Os SFG incluem Redes Neurais Fuzzy Genéticas, Algoritmos de Agrupamento Fuzzy Genéticos e Sistemas Fuzzy Genéticos Baseados em Regras (SFGBR), que será discutido nesta seção. As primeiras modalidades e tendências de SFGBR podem ser encontradas em (CORDON et al., 2004).

As abordagens de SFGBR se baseiam em dois aspectos básicos:

- Qual componente da Base de Conhecimento será otimizada pelos AG – a Base de Regras (BR), a Base de Dados (BD) ou a Base de Conhecimento (BC) completa;

- Qual a classe de problemas será abordada: *adaptação* de elementos da base de conhecimento já definidos, ou a *construção* de partes de base de conhecimentos sem conhecimento prévio.

Existem dois grupos que combinam as abordagens fuzzy e genética para a geração de bases de conhecimento: a *Adaptação Genética*, que ajustam componentes da base de regras, e a *Construção Genética*, que constrói os componentes da base de conhecimento.

Estes dois grupos são discutidos a seguir.

4.3.1 Adaptação genética

Neste grupo estão incluídos os métodos que utilizam AG para, a partir de uma base de dados ou base de regras já existente, aperfeiçoar o desempenho do sistema fazendo ajuste ou adaptação de uma ou mais partes da BC. Esses métodos podem ser subdivididos em grupos de acordo com o enfoque adotado:

Adaptação Genética dos Conjuntos Fuzzy

Neste grupo encontram-se os métodos que focalizam a sintonia de conjuntos fuzzy (KARR, 1991) (HERRERA et al., 1995) (BONISSONE et al., 1996) (CORDÓN; HERRERA, 1997) (CASILLAS et al., 2001). Nestes métodos a BR e a BD são definidas previamente. A BR é mantida inalterada, ao passo que a BD é alterada pelo AG, que altera os parâmetros das funções de pertinência de todas as partições fuzzy envolvidas no problema.

Otimização Genética do Número de Regras

Nesta categoria estão incluídos os métodos que otimizam o número de regras de uma BR, eliminando redundâncias (CASTRO; CAMARGO, 2004) (CORDÓN; HERRERA, 1997). Nestes métodos, um conjunto de regras é inicialmente obtido manualmente ou gerado através de algum mecanismo automático e posteriormente os AG são usados

para selecionar, a partir deste conjunto inicial de regras, um subconjunto de regras. Seu objetivo é reduzir o número de regras eliminando redundâncias e conflitos mantendo o mesmo desempenho do conjunto de regras inicial.

4.3.2 Construção genética

Essa abordagem focaliza a construção das regras fuzzy, considerando um conjunto de funções de pertinência já definido e fixo. Usualmente a definição da Base de Dados é feita escolhendo-se um número de valores linguísticos para as variáveis linguísticas e distribuindo uniformemente os valores, que são rótulos dos conjuntos fuzzy, no domínio de cada variável (HWANG;THOMPSON, 1994) (ISHIBUCHI et al., 1994) (HUFFMAN;PFISTER, 1997) (CASILLAS et al., 2001). Nesse grupo enquadram-se algumas das abordagens mais utilizadas e conhecidas:

- **A Abordagem de Pittsburgh** (HUFFMAN;PFISTER, 1997) (SPIEGEL;SUDKAMP, 2003) que se caracteriza pela representação de todo um conjunto de regras como um cromossomo. A principal desvantagem deste tipo de abordagem é que o crescimento do tamanho do cromossomo para problemas com número elevado de variáveis pode inviabilizar sua utilização.
- **A Abordagem de Michigan** (BONARINI, 1996) (ISHIBUCHI et al., 1999) onde cada cromossomo representa apenas uma regra e a base de regras será formada por todos os indivíduos da população. Em oposição à abordagem anterior, neste caso não ocorre o problema de aumento excessivo do tamanho do cromossomo, porém, a questão da competição entre as regras (indivíduos) deve ser tratada com cuidado, uma vez que as regras de uma base de regras devem cooperar entre si e não competir.
- **A Abordagem Iterativa** (CORDÓN;HERRERA, 1997) (JESUS et al., 2004) (HOFFMANN, 2004) da mesma forma que a abordagem de Michigan, o cromossomo representa apenas uma regra e o AG retorna um único cromossomo, regra, por iteração. Neste caso, porém, o processo genético é intercalado com um processo iterativo em que, a cada nova população gerada

pelo AG, a melhor regra da população é selecionada para fazer parte da base de regras. Assim, serão executadas quantas iterações forem necessárias até obter o número de regras desejado.

4.3.2.1 Construção Genética da Base de Dados

Nesta abordagem, proposta por (CORDÓN et al., 2001) e investigada em (PIRES;CAMARGO, 2004), o AG é usado para gerar a Base de Dados Fuzzy sem nenhum tipo de definição prévia, possivelmente modificando a cada passo o número de conjuntos fuzzy em cada partição bem como os parâmetros que controlam a definição das funções. O conjunto de regras também não é definido previamente, mas é reconstruído por meio de um método de geração automática de regras para cada indivíduo de cada população, em todas as gerações. Cada indivíduo representa uma possível partição para o problema e, a cada geração, um novo conjunto de regras é definido para cada um desses indivíduos pelo processo de geração de regras adotado, para que a função de aptidão (fitness) do AG possa ser calculada. Essa abordagem não é uma abordagem clássica como as anteriores e difere da adaptação genética dos conjuntos fuzzy na qual o que é feito é um ajuste de conjuntos considerando um conjunto fixo de regras.

4.3.2.2 Construção Genética da Base de Conhecimento

Os métodos nesse subgrupo usam AG para definir as regras e os conjuntos fuzzy simultaneamente (CORDÓN;HERRERA, 2001) (KIM et al., 2006). A complexidade do processo de busca nesse caso é alta, uma vez que o espaço no qual a busca vai acontecer é uma composição do espaço das possíveis regras com o espaço das possíveis funções de pertinência. Tanto as regras quanto os conjuntos fuzzy são representados na codificação.

Além das abordagens clássicas descritas nos grupos apresentados, um grande número de outras propostas surgiu nos últimos anos e continuam a surgir até os dias de hoje, trazendo abordagens que representam refinamentos ou combinações das categorias já detalhadas. Devido à grande variedade de propostas e o curto período de tempo transcorrido desde seu surgimento, é bastante difícil traçar uma distinção clara entre características de cada método, não sendo adequado agrupá-las em categorias como foi feito com as abordagens mais conhecidas.

4.4 Considerações Finais

Neste capítulo foram apresentados os conceitos fundamentais sobre AGs relacionados a este trabalho de pesquisa, bem como noções das principais abordagens de SFGs.

No próximo capítulo serão apresentados alguns trabalhos relacionados à proposta desenvolvida aqui.

5. Trabalhos relacionados

O número de trabalhos que utilizam conjuntos fuzzy do tipo-2 vem crescendo muito durante os últimos anos, sendo que alguns desses trabalhos buscam métodos de ajuste dos parâmetros das funções de pertinência de conjuntos fuzzy intervalares do tipo-2.

Em (Tan et al., 2007) os autores utilizaram conjuntos fuzzy intervalares do tipo-2 para construir um classificador baseado em regras fuzzy aplicado ao problema de classificação de arritmia cardíaca em registros de eletrocardiogramas extraídos da base de dados MIT-BIH. A base de regras possui nove regras que foram geradas por pelos conhecimentos e experiências de especialistas. Esse classificador possui duas variáveis de entrada e uma de saída.

Nesta abordagem, foram utilizadas funções de pertinência triangulares e para a construção dos conjuntos fuzzy do tipo-2, os autores utilizaram o algoritmo de agrupamento Fuzzy C-Means. Este algoritmo é executado em cada variável de entrada com o objetivo de encontrar três centros de cluster. Cada centro de cada cluster será usado como o centro do triângulo que representa a função de pertinência superior. O ponto inicial e o ponto final do triângulo serão os valores limite de cada cluster. A estratégia utilizada para a construção da função de pertinência inferior é baseada na comparação da variância dos dados de treinamento em cada cluster.

O treinamento foi realizado usando 62 instâncias e os testes foram executados com 30 instâncias. O classificador usando conjuntos fuzzy intervalares do tipo-2 foi comparado com um classificador fuzzy do tipo-1 e com um classificador que utiliza redes neurais artificiais. Analisando a média da taxa de classificação para cada uma das três classes, observou-se que o classificador fuzzy intervalar do tipo-2 obteve uma taxa de classificação 11,11% superior ao classificador fuzzy do tipo-1 e 21,79% superior ao classificador que faz uso de redes neurais.

Em (Chua & Tan, 2008) foi utilizado um AG para a geração da base de regras de um classificador fuzzy intervalar do tipo-2. Neste trabalho, os autores utilizaram a abordagem de Pittsburgh, sendo que cada cromossomo representa a base de regras inteira. Cada cromossomo é dividido em 3 partes e a sua estrutura é ilustrada na figura 5.1.

Parâmetros da MF	Classe	Relação
------------------	--------	---------

Figura 5.1 - Estrutura cromossômica

A primeira parte do cromossomo codifica os parâmetros da função de pertinência superior. O número de genes nesta parte do cromossomo é o número de conjuntos fuzzy multiplicado pelo número de parâmetros necessários para representar a função de pertinência. Se a função de pertinência for triangular, serão necessários três parâmetros (o centro e as pontas do triângulo). Caso sejam usadas funções Gaussianas, serão necessários dois parâmetros (média e desvio padrão).

A segunda parte representa o rótulo da classe e o número de genes nesta parte é igual ao número de regras na base de regras. A terceira parte representa a relação entre a função de pertinência superior e a função de pertinência inferior. Esta relação diz respeito apenas à distância entre o centro da função de pertinência superior e o centro da função de pertinência inferior. Esse valor está contido no intervalo (0,1]. Já os pontos laterais da função de pertinência inferior coincidem com os pontos laterais da função de pertinência superior.

Este classificador foi aplicado ao conjunto de dados automotivo Ford, cujo objetivo é identificar a presença de um humano no veículo a partir de dados coletados por sensores instalados no carro. Os resultados foram comparados com um classificador que utiliza conjuntos fuzzy do tipo-1 gerados a partir de um AG semelhante ao ilustrado na figura 1. Os autores concluíram com o trabalho que utilizar conjuntos fuzzy do tipo-2 para problemas de classificação do mundo real é uma opção melhor do que utilizar conjuntos fuzzy do tipo-1. Entretanto, esta abordagem não se preocupou com o tamanho da base de regras, sendo que o número de regras da base de regras gerada não foi discutido no trabalho.

Em (Wagner & Hagrais, 2007) os autores utilizaram conjuntos fuzzy intervalares do tipo-2 em controladores lógicos fuzzy aplicados à navegação de robôs móveis. Neste trabalho foram usadas funções de pertinências Gaussianas. Essas funções foram geradas através de AGs. A população inicial é gerada aleatoriamente e cada cromossomo representa uma base de regras completa, onde cada gene representa um parâmetro da função Gaussiana.

Em (Cao et al., 2008) foram utilizados conjuntos fuzzy intervalares do tipo-2 em controladores do sistema de suspensão de veículos. O formato das funções de

pertinência escolhido foi o formato triangular. Essas funções foram distribuídas na partição fuzzy de forma igualmente espaçada, tendo a FOU definida de forma empírica.

Em (Choia et al., 2009) foram propostos três métodos para a geração de funções de pertinência para conjuntos fuzzy intervalares: um baseado em heurística, um baseado em histograma e outro baseado em fuzzy C-Means intervalar do tipo-2. No método que utiliza a heurística, uma função de pertinência fuzzy do tipo-1 é escolhida para representar a função de pertinência fuzzy intervalar superior do tipo-2. A função de pertinência inferior é gerada multiplicando por $\alpha \in (0,1]$ os graus de pertinência de cada elemento com pertinência na função superior.

Na segunda abordagem, obtêm-se um histograma a partir de um conjunto de dados de treinamento. Para extrair a função de pertinência, uma função parametrizada é escolhida para modelar o histograma. Essa função é usando o método dos mínimos quadrados para encontrar uma função polinomial que represente este histograma. Calcula-se então o ponto máximo e mínimo dessa função. A altura da função Gaussiana será o valor do ponto máximo da função polinomial. O valor de dispersão será o menor valor entre a média da Gaussiana e o valor mínimo da função Gaussiana. De forma semelhante ao método heurístico, a função de pertinência inferior é obtida através de uma medida de proporção em relação a função de pertinência superior.

Na terceira abordagem, o fuzzy C-Means intervalar do tipo-2 é executado separadamente para cada classe e, após encontrar a menor distância entre cada classe e os padrões de entrada, as funções de pertinência superior e inferior são geradas usando a expressão modificada (proposta no artigo) do fuzzy C-Means intervalar do tipo-2.

Na classificação de padrões, tanto as características dos padrões quanto o espaço de hipóteses têm incertezas. Isso motivou alguns pesquisadores a empregar os conjuntos fuzzy do tipo-2 aos classificadores tradicionais para alcançar uma performance melhor em relação a sua robustez, a sua capacidade de generalização ou taxa de classificação (ZENG; LIU, 2007). Em (MITCHELL, 2005) Mitchell mostrou que a classificação de padrões é uma medida de similaridade entre dois sistemas fuzzy do tipo-2, no qual um conjunto considera a incerteza na característica do padrão e o outro conjunto na incerteza das hipóteses. Mitchell definiu esta medida de similaridade como:

$$S(\tilde{A}, \tilde{B}) = \sum_{m=1}^M \sum_{n=1}^N w_{mn} S(A_e^m, B_e^n) \quad (5.1)$$

sendo w_{mn} é o valor do grau secundário com o m -ésimo e o n -ésimo conjunto fuzzy embutido do tipo-1, tendo ao todo M e N conjuntos embutidos do tipo-1 em \tilde{A} e \tilde{B} respectivamente. Foi implementado um classificador fuzzy do tipo-2 utilizando esse grau de similaridade para a classificação de imagens radiográficas. Este classificador obteve uma taxa de erro 1.8% menor do que o um classificador fuzzy do tipo-1 implementado com o algoritmo K-NM proposto em (LIAO; LI, 1997).

Em (JOHN et al., 2000) foi utilizado um sistema fuzzy do tipo-2 para tratar a incerteza na interpretação de danos clínicos na tíbia. O sistema fuzzy do tipo-2 atuou no pré-processamento destas imagens que eram classificadas através dos algoritmos de agrupamento neuro-fuzzy fuzzyART e fuzzyMINMAX. Foi implementado também o tratamento da incerteza através de um sistema fuzzy do tipo-1. Os algoritmos que utilizaram tratamento de imagem com o sistema fuzzy do tipo-2 obtiveram uma taxa de classificação 5% maior do que o que utilizou o sistema fuzzy do tipo-1.

Em (LIANG; MENDEL, 2001) foram implementados classificadores fuzzy do tipo-1 com entradas singleton e entradas não-singleton, classificadores fuzzy do tipo-2 com entradas singleton e entradas não-singleton e classificadores Bayesianos. Esses classificadores foram usados para a classificação de tráfego de vídeo MPEG VBR, sendo que os classificadores fuzzy do tipo-2 obtiveram resultado superior aos classificadores fuzzy do tipo-1 e Bayesianos. Os testes mostraram que o classificador fuzzy do tipo-2 obteve uma taxa de alarme falso de 8.03% ao passo que o classificador fuzzy tipo-1 registrou uma taxa de 9.17%. Eles concluíram neste trabalho que classificadores fuzzy do tipo-2 são consideravelmente superiores que o classificador fuzzy do tipo-1 no que tange a robustez e taxa de erro de classificação.

De forma semelhante, Wu e Mendel (WU; MENDEL, 2006) usaram um classificador fuzzy do tipo-2 para classificar veículos de batalha através de suas características acústicas. Foram realizados 800 experimentos e o classificador fuzzy do tipo-2 obteve uma média de taxa de erro de classificação de 9.13%, ao passo que o classificador fuzzy do tipo-1 manteve sua média de taxa de erro de classificação em 12.8%. Além disso, eles realizaram os mesmos testes com um classificador bayesiano e observou-se que para este problema em específico, o classificador Bayesiano obteve desempenho inferior aos dois tipos de classificadores fuzzy.

Em (PARK; LEE-KANG, 2001) foi desenvolvido um sistema fuzzy do tipo-2 aplicado ao modelo de previsão de série caótica de Mackey-Glass. Neste método foi utilizado um AG para otimizar a função de pertinência do conjunto fuzzy do tipo-2. O método foi comparado ao método proposto em (JANG; SUN, 1993), que utiliza um sistema fuzzy do tipo-1 também aplicado à previsão de série caótica de Mackey-Glass. O método de Park e Lee-Kang conseguiu obter uma aproximação maior à série temporal real do que o método utilizado por Jang e Sun.

5.1 Considerações Finais

Neste capítulo foram apresentados alguns dos trabalhos de pesquisa que visam a geração e/ou otimização da base de dados em sistemas de classificação fuzzy do tipo-2.

No próximo capítulo será apresentado o método proposto por este trabalho.

6. Geração genética de classificador fuzzy intervalar do tipo-2.

Este capítulo tem por objetivo apresentar a metodologia proposta neste trabalho para geração de classificador fuzzy intervalar do tipo-2 usando AGs. O trabalho desenvolvido visa investigar os benefícios e custos obtidos com o uso de conjuntos fuzzy do tipo-2 na representação e processamento de conhecimento no problema de classificação, com relação a outros métodos baseados em conjuntos fuzzy do tipo-1. A metodologia utilizada é composta por três etapas:

Geração da base de regras: Nesta etapa um AG que utiliza a abordagem Pittsburgh é responsável por gerar regras fuzzy capazes de representar o conhecimento existente no conjunto de exemplos. Os conjuntos fuzzy são gerados com o auxílio do algoritmo Fuzzy C-Means (FCM) (BARALDI & BLONDA, 1999a) (BARALDI & BLONDA, 1999b) e permanecem fixos durante o processo genético.

Otimização da base de dados: Nesta etapa um AG é utilizado para otimizar os conjuntos fuzzy utilizados na etapa anterior.

Otimização da base de regras: Esta etapa visa excluir, usando também AG, as regras redundantes e desnecessárias que por ventura foram geradas na etapa anterior, resultando numa base de regras compacta.

Com essas três etapas o algoritmo é capaz de gerar uma base de regras compacta com alta habilidade para classificação.

A escolha da abordagem Pittsburgh foi motivada pela idéia de explorar a vantagem que ela oferece sobre as demais abordagens: possibilidade de avaliar toda a base de regras na função de aptidão, eliminando o problema de competição e cooperação entre as regras da mesma base.

Essas três etapas poderiam ser feitas em um único passo, usando um AG que fizesse simultaneamente a geração das regras e o ajuste dos conjuntos fuzzy.

Entretanto, a complexidade da busca gerada por tal junção certamente iria comprometer o desempenho da base de regras obtida. Outro motivo para dividir as atividades é a possibilidade de aplicá-las separadamente. Por exemplo, o AG para otimização da base de regras pode ser aplicado como pós-processamento para outros métodos que geram base de regras. Desta forma pode-se analisar seu comportamento quando as regras são criadas por outros métodos de aprendizado.

A fim de se obter um parâmetro de comparação, essas etapas foram implementadas duas vezes: uma utilizando conjuntos fuzzy do tipo-1 e outra utilizando conjuntos fuzzy intervalares do tipo-2. Os processos diferem na segunda etapa, em que a codificação dos conjuntos fuzzy do tipo-1 e do tipo-2 são diferentes.

A metodologia a ser apresentada se baseia na existência de uma base de dados previamente definida que contenha os conjuntos fuzzy associados a cada variável do problema, bem como suas funções de pertinência.

A seguir são descritas cada etapa do processo de geração genética de classificadores fuzzy.

Etapa 1: Geração da base de regras.

Neste trabalho, as funções de pertinência associadas aos conjuntos fuzzy foram definidas pelo algoritmo de agrupamento FCM. Agrupamento é uma técnica utilizada para reunir em grupos dados ou objetos que possuem algumas semelhanças. O algoritmo clássico de agrupamento gera partições em que cada objeto pertence a um único grupo. Já o agrupamento fuzzy permite que um objeto pertença a vários grupos simultaneamente, com diferentes graus de pertinência.

O processo para geração de conjuntos fuzzy utilizando FCM foi aplicado no domínio de cada atributo individualmente da seguinte forma: suponha um conjunto de dados com L exemplos descritos por $E = \{e_1, e_2, \dots, e_L\}$ e cada exemplo $e_p = (a_{p1}, a_{p2}, \dots, a_{pn})$ é composto por n atributos. Define-se para cada atributo o número de conjuntos fuzzy a ser gerado. Aplica-se então o FCM sobre cada domínio de atributo separadamente, gerando os grupos, em número igual ao de conjuntos fuzzy desejado para o domínio. O centro de cada grupo é usado para gerar os conjuntos fuzzy utilizados no restante do processo.

Para este trabalho, foi utilizada a forma triangular para representar as funções de pertinência. Para conjuntos fuzzy do tipo-1 o centro de cada grupo encontrado pelo FCM será o centro de cada triângulo. A base esquerda do triângulo será o centro do triângulo anterior e a base direita do triângulo será o centro do triângulo posterior.

Para conjuntos fuzzy do tipo-2 utilizou-se o método proposto por (TAN et al., 2007) para a definição das funções de pertinência superior e inferior de conjuntos fuzzy intervalares do tipo-2. Primeiramente aplica-se o FCM em cada atributo da mesma forma que é feito para conjuntos fuzzy do tipo-1. O centro da função de pertinência superior do conjunto J é o centro do grupo J encontrado pelo FCM. A base da função de pertinência superior J depende da dispersão dos dados do grupo J . Essa dispersão é calculada pela seguinte expressão:

$$d_j = \sum_{p=1}^L \sqrt{(a_{p1} - C_{1j})^2 + (a_{p2} - C_{2j})^2 + \dots + (a_{pn} - C_{nj})^2} \quad (6.1)$$

onde a_{pn} é o atributo n do exemplo p e $\langle C_{1j}, C_{2j}, \dots, C_{nj} \rangle$ é o centróide do grupo J .

A distância entre o ponto esquerdo da função de pertinência superior e o ponto esquerdo da função de pertinência inferior, β_{2j} e β_{1j} respectivamente, é dada por:

$$\beta_{2j} = \alpha_j \times \gamma_{2j} \quad (6.2)$$

$$\beta_{1j} = \alpha_j \times \gamma_{1j} \quad (6.3)$$

onde γ_{2j} corresponde à distância entre o centro da função de pertinência e sua base esquerda e γ_{1j} corresponde à distância entre o centro da função de pertinência e sua base direita. A representação gráfica dessas medidas pode ser vista na figura 6.1.

Ainda em (TAN et al., 2007), conclui-se através de experimentos que a altura da função de pertinência inferior não influi na taxa de classificação de classificadores que utilizam conjuntos intervalares do tipo-2. Por este motivo, a fim de diminuir a quantidade de parâmetros codificados no cromossomo, a altura da função de pertinência inferior foi fixada em 0,6.

Após a construção dos conjuntos fuzzy, que permanecerão fixos até o término desta etapa, inicia-se o processo de geração de regras.

A partir de um conjunto de dados que representa amostras ou exemplos do problema, o AG procura por uma base de regras mais apropriada para classificar corretamente estes exemplos.

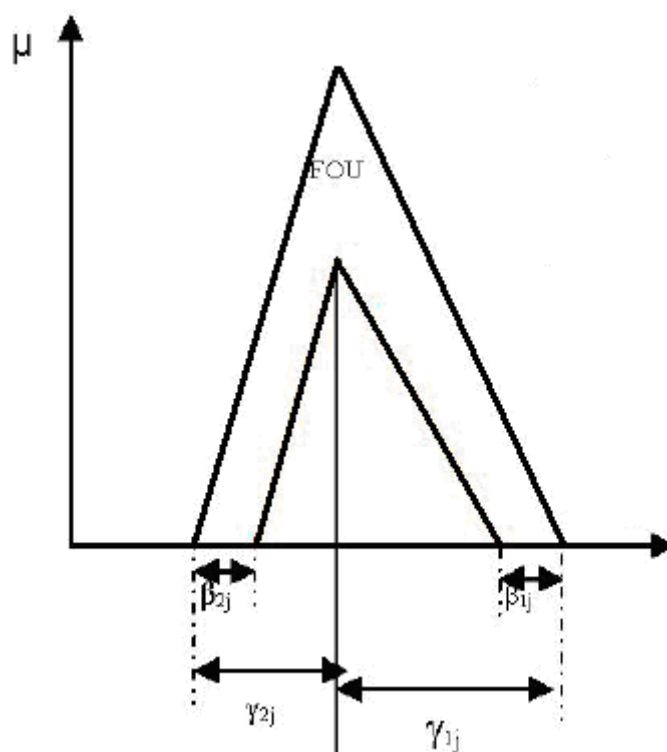


Figura 6.1 - Índices de medidas para a construção da função de pertinência intervalar do tipo-2.

O AG desta etapa é responsável por evoluir uma população de bases de regras a fim de encontrar uma base de regras apropriada, que consiga classificar o maior número possível de exemplos. O algoritmo obedece ao seguinte esquema:

1. Gerar uma população inicial com N bases de regras.
2. Classificar todos os exemplos do conjunto de treinamento usando as bases de regras da população corrente e calcular o valor de aptidão para cada uma.
3. Aplicar as operações genéticas nas bases de regras.

4. Se atingir o critério de parada, retorna a base de regras com maior valor de aptidão. Caso contrário, volta ao passo 2.

A seguir serão expostos em detalhes os principais componentes do AG utilizado nesta etapa, como codificação, função de aptidão, operadores genéticos e condição de parada do algoritmo.

- **Codificação das regras**

Primeiramente é preciso estabelecer a forma de codificar as possíveis soluções para que o AG possa manipulá-las. Foi definido que os cromossomos possuem tamanho fixo e é necessário estabelecer previamente a quantidade de regras que eles irão codificar. As regras são codificadas com números inteiros, os quais representam os índices dos conjuntos fuzzy que aparecem na sua parte antecedente e consequente (quando o conjunto de dados possui classe contínua). Se o conjunto de dados não possui classe contínua, cada classe receberá um número que a identificará. As classes contínuas foram discretizadas, por meio de um processo explicado junto com a apresentação da função de aptidão.

Como estamos utilizando a abordagem de Pittsburg, cada indivíduo da população representa uma base de regras completa. A figura 6.2 ilustra a representação genética de um cromossomo C_i que está codificando k regras, onde cada regra possui três antecedentes e um consequente.

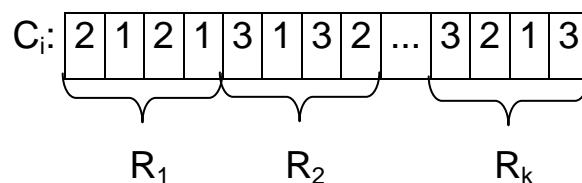


Figura 6.2 - Representação genética da base de regras.

O cromossomo C_i ilustrado na figura 6.2 representa a seguinte base de regras:

R_1 : SE x_1 é A_{21} E x_2 é A_{12} E x_3 é A_{23} , ENTÃO Classe=1

R_2 : SE x_1 é A_{31} E x_2 é A_{12} E x_3 é A_{33} , ENTÃO Classe=2

...

R_k : SE x_1 é A_{31} E x_2 é A_{22} E x_3 é A_{13} , ENTÃO Classe=3

A população inicial é gerada de forma aleatória.

- **Função de aptidão**

A função de aptidão, responsável por avaliar cada cromossomo C_i , $i=1...N$ é definida com base no desempenho do conjunto de regras nele codificado, medido pela Taxa de Classificação Correta (TCC). No caso do consequente da regra possuir uma classe contínua, calcula-se a classe esperada e a classificação será considerada correta se esta classe for igual a classe retornada pela regra. O objetivo do processo genético será maximizar a função:

$$\text{valor_aptidão}(C_i) = TCC(C_i) \quad (6.4)$$

- **Operadores genéticos**

A seleção é a primeira operação a ser empregada e utiliza o método da roleta para escolher os indivíduos. A operação de seleção é combinada com o elitismo, onde uma porcentagem dos melhores indivíduos da população sempre sobrevive para compor a próxima. A segunda operação a ser realizada é o cruzamento. Neste trabalho foi utilizado o cruzamento de um ponto. Os pares de cromossomos a serem cruzados e o ponto de cruzamento são escolhidos aleatoriamente. Por fim, a mutação padrão é aplicada para alterar um gene c_{ij} para um novo valor escolhido aleatoriamente dentro dos possíveis valores $\{0, 1, \dots, G\}$, sendo G o número máximo que o gene pode assumir. Em outras palavras, G é o número máximo de conjuntos fuzzy definido para a variável representada pelo gene c_{ij} .

- **Condição de parada do algoritmo**

O critério de parada adotado para este AG é o número máximo de gerações. Ao atingir esse número o AG retornará como solução o cromossomo que tiver o maior valor de aptidão dentre todos os indivíduos da população final.

Etapa 2: Otimização da base de dados.

Ao término da etapa 1, o melhor indivíduo da população terá a melhor taxa de classificação encontrada pelo AG. Esta base de regras será utilizada de maneira fixa nesta segunda etapa.

O AG desta etapa é responsável por evoluir uma população de base de dados a fim de encontrar o melhor ajuste para cada conjunto de cada partição fuzzy. O algoritmo obedece ao seguinte esquema:

1. Gerar uma população inicial com N indivíduos.
2. Classificar todos os exemplos do conjunto de treinamento usando as bases de dados da população corrente e calcular o valor de aptidão para cada uma.
3. Aplicar as operações genéticas nas bases de dados.
4. Se atingir o critério de parada, retorna a base de dados com maior valor de aptidão. Caso contrário, volta ao passo 2.

A seguir serão expostos em detalhes os principais componentes do AG utilizado nesta etapa.

- **Codificação da base de dados**

Foi definido para o AG desta etapa que os cromossomos possuem tamanho fixo e é necessário estabelecer previamente a quantidade de conjuntos que as partições terão. Os conjuntos são codificados com números reais, os quais representam os parâmetros de cada função de pertinência de cada partição.

Cada indivíduo da população representa uma base de dados completa. A figura 6.3 ilustra a representação de um cromossomo que está codificando nP partições, onde cada partição possui nC conjuntos fuzzy do tipo-1.

q_{11}^1	q_{12}^1	q_{13}^1	\dots	q_{nC1}^1	q_{nC2}^1	q_{nC3}^1	\dots	q_{11}^{nP}	q_{12}^{nP}	q_{13}^{nP}	\dots	q_{nC1}^{nP}	q_{nC2}^{nP}	q_{nC3}^{nP}
------------	------------	------------	---------	-------------	-------------	-------------	---------	---------------	---------------	---------------	---------	----------------	----------------	----------------

Figura 6.3 - Representação cromossômica da base de dados com funções de pertinência do tipo-1.

A figura 6.4 ilustra a representação de um cromossomo que está codificando nP partições, onde cada partição possui nC conjuntos fuzzy do tipo-2. Nesta representação,

$q_{S_{ab}}^P$ é o gene correspondente ao parâmetro $b \in \{1, 2, 3\}$ da função de pertinência superior a da partição P e $q_{I_{ab}}^P$ é o gene correspondente ao parâmetro $b \in \{1, 2\}$ da função de pertinência inferior a da partição P .

$q_{S_{11}}^1$	$q_{S_{12}}^1$	$q_{S_{13}}^1$	$q_{I_{11}}^1$	$q_{I_{12}}^1$...	$q_{S_{nC1}}^1$	$q_{S_{nC2}}^1$	$q_{S_{nC3}}^1$	$q_{I_{nC1}}^1$	$q_{I_{nC2}}^1$...	$q_{S_{11}}^{nP}$	$q_{S_{12}}^{nP}$	$q_{S_{13}}^{nP}$	$q_{I_{11}}^{nP}$	$q_{I_{12}}^{nP}$...	$q_{S_{nC1}}^{nP}$	$q_{S_{nC2}}^{nP}$	$q_{S_{nC3}}^{nP}$	$q_{I_{nC1}}^{nP}$	$q_{I_{nC2}}^{nP}$
----------------	----------------	----------------	----------------	----------------	-----	-----------------	-----------------	-----------------	-----------------	-----------------	-----	-------------------	-------------------	-------------------	-------------------	-------------------	-----	--------------------	--------------------	--------------------	--------------------	--------------------

Figura 6.4 - Representação cromossômica da base de dados com funções de pertinência do tipo-2.

A população inicial é gerada introduzindo um cromossomo que representa as funções de pertinência utilizadas de forma fixa na etapa 1. Os outros cromossomos são gerados aleatoriamente.

- **Função de aptidão**

A função de aptidão deste AG é definida com base no desempenho do conjunto de dados nele codificado, medido pela taxa de classificação correta (TCC). O melhor cromossomo será aquele que maximizar a função 6.4.

Assim como na etapa anterior, o método de seleção do AG desta etapa é o método da roleta e o operador de cruzamento é o de um ponto. O critério de parada adotado para este AG é o número máximo de gerações. Ao atingir esse número o AG retornará como solução o cromossomo que tiver o maior valor de aptidão dentre todos os indivíduos da população final.

Etapa 3: Otimização da base de regras.

Ao término da etapa 2, a base de regras contém algumas regras conflitantes e redundantes. Considera-se também a possibilidade de retirar da base de regras aquelas regras que não classificam nenhum padrão, ou seja, regras que possuem baixo grau de compatibilidade com os padrões de treinamento em relação a outras regras. Assim, um processo para elimina-las é requerido.

O processo de otimização da base de regras tem por objetivo formar subconjuntos de regras, através de diferentes combinações das regras previamente obtidas na etapa anterior, que possuam um número reduzido de regras e alta habilidade para classificação.

Tal algoritmo obedece ao seguinte esquema:

1. Gerar uma população inicial com N subconjuntos de regras a partir das regras previamente obtidas.
2. Classificar todos os exemplos de treinamento usando os subconjuntos de regras representados em cada indivíduo da população e calcular o valor de aptidão para cada um.
3. Aplicar as operações genéticas nas bases de dados.
4. Se atingir o critério de parada, retorna o subconjunto de regras com maior valor de aptidão. Caso contrário, volta ao passo 2.

A seguir serão expostos em detalhes os principais componentes do AG utilizado nesta etapa.

- **Codificação da base de dados**

Nesta etapa cada subconjunto (combinação) de regras será representado por cromossomos $C_i=(c_{i1},c_{i2},\dots,c_{im})$ que são codificados com sequências de dígitos binários e possuem tamanho fixo m , sendo m o número de regras geradas pela etapa anterior. Cada gene c_{ij} (um dígito binário) é associado a uma regra. Se $c_{ij}=1$, então a regra correspondente a este gene está ativada e faz parte do subconjunto representado pelo cromossomo C_i , caso contrário a regra está desativada e não faz parte do subconjunto em questão.

Suponha que na etapa anterior foi gerada uma base de regras com 5 regras, R_1, R_2, \dots, R_5 , e que esta precisa ser otimizada. O cromossomo da Figura 6.5 representa um possível subconjunto composto somente pelas regras R_1, R_4 e R_5 . Logo, estas regras estão ativadas e os exemplos de treinamento serão classificados somente por elas.

1	0	0	1	1
---	---	---	---	---

Figura 6.5 - Codificação de um cromossomo da etapa 3.

A população inicial é formada introduzindo um cromossomo que representa uma base de regras com todas as regras geradas pela etapa 2, ou seja, um cromossomo com todos os genes iguais a 1. Os outros cromossomos são gerados aleatoriamente.

- **Função de aptidão**

A função de aptidão utilizada nesta etapa avalia cada cromossomo (subconjunto de regras candidatas) com base em dois critérios (Castro & Camargo, 2005): taxa de classificação correta e o número de regras contido nesse subconjunto. Serão privilegiados os cromossomos que apresentarem a maior taxa de classificação correta com o menor número de regras. Desta forma, a função de aptidão a ser maximizada é definida como:

$$valor_aptidão(C_i) = TCC(C_i) * w_1 - NR(C_i) * w_2 \quad (6.5)$$

O objetivo é maximizar o $TCC(C_i)$ e minimizar $NR(C_i)$, que representa o número de regras contidas no cromossomo C_i . Os pesos w_1 e w_2 são números reais positivos que são especificados na preferência do usuário para um problema de classificação específico.

6.1 Algoritmo Genético auto-adaptativo

Para evitar o problema de convergência prematura do AG, foi utilizado um Algoritmo Genético Auto-Adaptativo (AGAA) (CASTRO & CAMARGO, 2004) para ajustar dinamicamente as taxas de cruzamento e mutação a cada nova geração, mantendo a Diversidade Genética (DG) das populações. Este algoritmo foi utilizado nos AGs de cada etapa que compõem o processo de geração genética de classificadores.

A DG foi calculada usando a média entre o valor médio e o melhor valor de aptidão encontrados em cada geração, como ilustrado na expressão a seguir:

$$DG = \frac{aptidão_média}{aptidão_máxima} \quad (6.6)$$

O valor *aptidão_média* representa a média do valor de aptidão de todos os cromossomos da população, ao passo que *aptidão_máxima* representa a maior aptidão encontrada na população.

Quando a DG se aproxima de 1, isto indica que os indivíduos da população estão muito semelhantes. Para evitar a convergência prematura, a taxa de cruzamento deve ser diminuída e a taxa de mutação deve ser aumentada. Desta forma, gera-se mais indivíduos diferentes entre si. Quando a DG se aproxima da média, isto significa que os indivíduos estão muito diferentes e o processo de busca pode se tornar aleatório e sem direção. Assim, torna-se necessário diminuir a DG aumentando a taxa de cruzamento e diminuindo a taxa de mutação para explorar os melhores indivíduos de cada geração. O procedimento de ajuste dinâmico das taxas de cruzamento e mutação é detalhado no algoritmo 6.1.

Algoritmo 6.1 – Algoritmo Genético auto-adaptativo

01: **INÍCIO**

02: **SE** $DG \geq VMax$

03: {

04: taxa_de_mutação = $\min(\text{taxa_de_mutação} * v, Msup)$;

05: taxa_de_cruzamento = $\max(\text{taxa_de_cruzamento} / v, Cinf)$;

06: } **SENÃO SE** $DG \leq VMin$ {

07: taxa_de_mutação = $\max(\text{taxa_de_mutação} / v, Minf)$;

08: taxa_de_cruzamento = $\min(\text{taxa_de_cruzamento} * v, Csup)$;

09: }

10: **FIM**

Os parâmetros *VMax* e *VMin* indicam os limites do intervalo de valores que DG pode assumir e $v \geq 1$ é uma constante. *Msup* e *Minf* são os limites superior e inferior, respectivamente, do intervalo que a taxa de mutação pode assumir e *Csup* e *Cinf* são os limites superior e inferior, respectivamente, do intervalo que a taxa de cruzamento pode assumir. O símbolo * nas linhas 4 e 8 representa a operação matemática de multiplicação.

O uso de AGAAs permitiu o ajuste dinâmico das taxas de cruzamento e mutação, mantendo a DG das populações dentro dos valores considerados normais. Isto permitiu evitar a convergência prematura do AG, ao mesmo tempo em que foram explorados os melhores indivíduos de cada geração, evitando que a busca se tornasse aleatória e sem direção.

No próximo capítulo serão apresentados e discutidos os resultados obtidos nos experimentos realizados com diversos conjuntos de dados, considerando a abordagem proposta para conjuntos fuzzy do tipo-1 e do tipo-2 e o método Wang&Mendel.

6.2 Considerações Finais

Neste capítulo foi apresentado o método para a geração genética de classificadores fuzzy intervalares do tipo-2, incluindo o algoritmo auto-adaptativo para o cálculo do valor de aptidão dos indivíduos no AG.

No próximo capítulo serão apresentados os experimentos realizados e a análise dos resultados obtidos.

7. Experimentos e análise dos resultados

Neste capítulo são apresentados os experimentos realizados, os parâmetros de configuração do AG e são discutidos os resultados obtidos para a avaliação do método.

7.1 Descrição dos conjuntos de dados

Com a finalidade de testar o método previamente apresentado de geração automática de classificadores fuzzy intervalar do tipo-2, foram realizados diferentes experimentos utilizando alguns conjuntos de dados obtidos no UCI Repository of Machine Learning Databases (BLAKE & MERZ, 1998).

Os conjuntos de dados utilizados durante os testes são apresentados na Tabela 7.1.

Tabela 7-1 - Conjuntos utilizados nos experimentos

Conjunto de Dados	Quantidade de exemplos	Número de atributos	Número de classes
Hayes-Roth	160	5	3
Íris	150	5	3
Servo	167	4	Contínua
Glass	214	10	7
Wine	178	13	3
Statlog (Heart)	690	14	2
Liver-disorder	345	6	2
Ecoli	336	8	8
AUTO-MPG	398	8	Contínua
Sonar	208	60	2
Balance-Scale	625	4	3
Boston Housing	506	14	Contínua
Ionosphere	351	34	2
Pima	768	8	2
Japanese Credit Screening	125	15	2
Breast Cancer Wisconsin	569	31	2
Vowel	990	14	10
Statlog-Vehicle	946	18	4
Car Evaluation	1728	6	4

O domínio de cada um dos atributos dos conjuntos de dados estão representados por valores numéricos contínuos e precisam ser granularizados em um número de conjuntos fuzzy que representem os valores linguísticos que cada atributo pode assumir, a fim de permitir a generalização do conhecimento.

Os conjuntos Boston Housing, Auto-mpg e Servo possuem valores numéricos contínuos na variável de saída (de classificação) e precisam ser granularizados da mesma forma que os atributos. Neste trabalho, todas as variáveis de saída contínuas foram granularizadas, gerando-se 3 conjuntos fuzzy para cada uma. Dessa forma foi definido que as classes utilizadas para representar a variável “consumo” no conjunto Auto-mpg são “baixo”, “médio” e “alto”. Seguindo o mesmo raciocínio, as classes para representar a variável “preço” das casas no conjunto Boston Housing são “baixo”, “médio” e “alto”, Para a variável Servo o tempo de resposta pode ser “baixo”, “médio” e “alto”. Durante o processo de inferência, as classes destas variáveis serão consideradas discretas e a classificação é feita encontrando o conjunto de maior pertinência para o valor contido na variável de saída para o exemplo atual a ser classificado. Se o índice do conjunto com maior pertinência for igual ao valor retornado pela regra vencedora, então o exemplo foi classificado corretamente. Caso contrário o exemplo foi classificado de forma incorreta.

As variáveis dos conjuntos de dados Iris, Vinho e Glass possuem valores nominais. Para este trabalho esses valores nominais são transformados em valores numéricos e representados como 1, 2, 3, etc, por exemplo.

7.2 Experimentos realizados

A Tabela 7.2 apresenta os valores iniciais para alguns dos parâmetros dos AGs utilizados no método proposto. Os valores iniciais utilizados foram escolhidos com base em experimentos preliminares que variaram as taxas utilizadas em experimentos similares. O método de validação foi o 10-fold cross validation (CHANG et al., 1992; KOHAVI, 1995).

Tabela 7-2 - Parâmetros dos AGs.

Parâmetro	Valor
Tamanho da população	200
Taxa de cruzamento inicial	0.7
Taxa de mutação inicial	0.05
Elitismo	0.05
Número de gerações	1000

A Tabela 7.3 apresenta os valores dos parâmetros do AGAA mostrado no Algoritmo 6.1. Estes valores foram definidos de forma empírica. Desta forma, as taxas de cruzamento e de mutação foram ajustadas dinamicamente como forma de controlar e manter a diversidade genética das populações geradas.

Tabela 7-3 - Parâmetros dos AGs auto-adaptativos.

Parâmetro	Valor
Vmax	0.8
Vmin	0.3
Cmax	0.9
Cmin	0.1
Mmax	0.25
Mmin	0.01
v	1.3

Na presente metodologia é necessário estabelecer previamente a quantidade de regras a serem codificadas em cada cromossomo. Foi observado que quando este número era muito pequeno, havia pouca diversidade genética e o algoritmo convergia rapidamente, deixando de explorar alguns pontos do espaço de busca. Como solução foi definido que o número de regras codificadas seria de 25% do total de exemplos de treinamento de cada conjunto de dados. Assim, os cromossomos codificam de 38 a 432 regras, dependendo do conjunto de dados. Isso não significa que todas farão parte da base de regras final, pois algumas delas, desnecessárias, serão excluídas na etapa de otimização.

Foram implementadas duas versões do método de geração genética de classificadores fuzzy: uma utilizando conjuntos fuzzy do tipo-1 e uma utilizando conjuntos fuzzy intervalares do tipo-2. Segundo (CORDÓN et al, 2001), a partição fuzzy exerce uma influência significativa no desempenho da base regras gerada. Com base nisso, foi examinado o desempenho do algoritmo mediante diferentes partições fuzzy para as variáveis. Sendo assim, cada versão do método foi testada com particionamento de 3, 5 e 7 conjuntos.

O classificador fuzzy do tipo-1 particionado em 3 conjuntos será aqui identificado como CT1_3. De forma semelhante serão utilizados os rótulos CT1_5 e CT1_7 para identificar os classificadores fuzzy do tipo-1 particionados em 5 e 7 conjuntos, respectivamente. Seguindo o mesmo esquema temos CT2_3, CT2_5 e CT2_7 para identificar os classificadores fuzzy intervalares do tipo-2 particionados em 3, 5 e 7 conjuntos respectivamente. O método também foi comparado com o método de

Wang & Mendel (WM) (WANG & MENDEL, 1992) utilizando 3, 5 e 7 partições (WM_3, WM_5 e WM_7 respectivamente).

Os resultados com a média da taxa de classificação para cada versão do método são mostrados na Tabela 7.4, na Tabela 7.5 e na Tabela 7.6. O número entre parênteses ao lado de cada taxa de classificação corresponde ao desvio padrão.

Tabela 7-4 - Taxa de classificação correta

	Hayes-Roth	Iris	Servo	Glass	Wine	Statlog (Heart)	Liver-disorder
WM_3	0.6154(0.4883)	0.9133(0.2822)	0.9063(0.2924)	0.5286(0.5003)	0.9059(0.2928)	0.6778(0.4682)	0.5235(0.5001)
WM_5	0.5077(0.5018)	0.9400(0.2382)	0.7813(0.4147)	0.5714(0.4960)	0.8647(0.3430)	0.5852(0.4936)	0.5794(0.4943)
WM_7	0.5000(0.5019)	0.8933(0.3097)	0.6938(0.4623)	0.5905(0.4929)	0.4235(0.4955)	0.5778(0.4948)	0.5647(0.4965)
CT1_3	0.8291(0.3350)	0.9788(0.4965)	0.9541(0.3854)	0.6845(0.4852)	0.8956(0.3102)	0.8741(0.4325)	0.8680(0.4852)
CT1_5	0.7843(0.3965)	0.9654(0.4165)	0.8358(0.4126)	0.7833(0.3254)	0.8685(0.4852)	0.8344(0.4284)	0.8571(0.5136)
CT1_7	0.7711(0.3823)	0.9622(0.4298)	0.9715(0.3957)	0.8310(0.3714)	0.8016(0.5023)	0.8286(0.4952)	0.8607(0.4658)
CT2_3	0.9254(0.2852)	0.9980(0.1205)	0.9754(0.2798)	0.8771(0.3225)	0.9920(0.3325)	0.9364(0.4159)	0.9157(0.4852)
CT2_5	0.8457(0.3156)	0.9784(0.2348)	0.9398(0.3215)	0.8645(0.2865)	0.9532(0.3551)	0.8854(0.4236)	0.9412(0.3932)
CT2_7	0.8654(0.3085)	0.9832(0.2514)	0.9876(0.3354)	0.8915(0.4125)	0.9580(0.3154)	0.8626(0.4074)	0.9487(0.3845)

Tabela 7-5 - Taxa de classificação correta.

	Ecoli	AUTO-MPG	Sonar	Balance-Scale	Boston Housing	Ionosphere	Pima
WM_3	0.7152(0.4520)	0.8256(0.3799)	0.6050(0.4900)	0.8145(0.3890)	0.8200(0.3845)	0.6514(0.4772)	0.6803(0.4666)
WM_5	0.7909(0.4072)	0.7436(0.4372)	0.5550(0.4982)	0.4581(0.4986)	0.6300(0.4832)	0.8457(0.3617)	0.6829(0.4656)
WM_7	0.7697(0.4216)	0.6026(0.4900)	0.4650(0.5000)	0.4710(0.4995)	0.5540(0.4975)	0.6657(0.4724)	0.6776(0.4676)
CT1_3	0.8739(0.3954)	0.9067(0.4251)	0.7929(0.4852)	0.8440(0.3754)	0.8617(0.4221)	0.7806(0.4215)	0.7912(0.4524)
CT1_5	0.9431(0.4321)	0.8401(0.4752)	0.7463(0.4952)	0.8700(0.3726)	0.8212(0.3952)	0.9189(0.5021)	0.8170(0.4952)
CT1_7	0.9278(0.3898)	0.7860(0.4632)	0.6629(0.5020)	0.8380(0.4215)	0.7941(0.4352)	0.9006(0.4852)	0.7808(0.3214)
CT2_3	0.9198(0.4632)	0.9640(0.3521)	0.8547(0.4721)	0.8571(0.3824)	0.9325(0.4554)	0.8349(0.4321)	0.8211(0.5021)
CT2_5	0.9645(0.4852)	0.8612(0.3265)	0.8123(0.4852)	0.9259(0.4580)	0.9346(0.4522)	0.9321(0.5214)	0.8268(0.4852)
CT2_7	0.9322(0.4125)	0.8354(0.4125)	0.7365(0.4125)	0.8469(0.4685)	0.8198(0.4752)	0.9089(0.4852)	0.8169(0.3751)

Tabela 7-6 - Taxa de classificação correta.

	Japanese CS	Breast CW	Vowel	Statlog Vehicle	Car Evaluation
WM_3	0.7908(0.4070)	0.9103(0.2859)	0.6111(0.4877)	0.5845(0.4931)	0.8192(0.3849)
WM_5	0.7477(0.4346)	0.6941(0.4611)	0.9202(0.2711)	0.6071(0.4886)	0.7000(0.4583)
WM_7	0.6846(0.4650)	0.6853(0.4647)	0.9354(0.2460)	0.5738(0.4948)	0.6547(0.4756)
CT1_3	0.8354(0.4325)	0.9254(0.4158)	0.7452(0.4125)	0.8647(0.4125)	0.8752(0.5001)
CT1_5	0.8064(0.3852)	0.7453(0.4120)	0.9432(0.3822)	0.8718(0.4055)	0.8414(0.4652)
CT1_7	0.7091(0.4521)	0.7362(0.4158)	0.9483(0.3758)	0.8475(0.4652)	0.8407(0.4705)
CT2_3	0.8954(0.4421)	0.9754(0.3854)	0.8354(0.3655)	0.8810(0.3852)	0.9632(0.3952)
CT2_5	0.8952(0.4215)	0.9637(0.3154)	0.9522(0.3742)	0.8966(0.4721)	0.8954(0.4122)
CT2_7	0.8611(0.4125)	0.9322(0.3652)	0.9655(0.4125)	0.8709(0.4215)	0.9370(0.3945)

O número médio de regras de cada classificador é mostrado na Tabela 7.7, na Tabela 7.8 e na Tabela 7.9. O número dentro dos parênteses ao lado do número médio de regras corresponde ao desvio padrão.

Tabela 7-7 - Número de regras geradas.

	Hayes-Roth	Iris	Servo	Glass	Wine	Statlog (Heart)	Liver-disorder
WM_3	70.3(0.1766)	14.7(0.4106)	37.5(0.4077)	42.0(0.3215)	127.0(0.3845)	221.1(0.3246)	45.7(0.3659)
WM_5	99.2(0.3647)	45.4(0.3879)	151.0(0.2434)	87.8(0.2179)	160.2(0.2234)	235.7(0.1532)	121.0(0.2543)
WM_7	102.7(0.2935)	66.3(0.4039)	151.0(0.3506)	126.4(0.2355)	161.0(0.3152)	241.4(0.4222)	186.1(0.3974)
CT1_3	16.3(0.4290)	14.4(0.2044)	16.6(0.2974)	23.2(0.3311)	17.4(0.3014)	23.4(0.3556)	19.8(0.3109)
CT1_5	16.5(0.2893)	17.8(0.1186)	19.3(0.2294)	25.4(0.4815)	15.2(0.2734)	27.8(0.4606)	23.4(0.4158)
CT1_7	19.8(0.3677)	18.3(0.3399)	18.7(0.3683)	22.1(0.3606)	15.8(0.4534)	29.7(0.3269)	23.6(0.3471)
CT2_3	12.4(0.3049)	8.7(0.3088)	15.3(0.3297)	10.6(0.3027)	12.7(0.2884)	19.5(0.2915)	15.7(0.2200)
CT2_5	14.8(0.2608)	13.9(0.1484)	15.8(0.3091)	12.8(0.4088)	14.4(0.3612)	25.8(0.3749)	21.9(0.3298)
CT2_7	17.2(0.4017)	15.3(0.2302)	17.9(0.1996)	13.8(0.2634)	14.6(0.2084)	26.6(0.3284)	22.4(0.3140)

Tabela 7-8 - Número de regras geradas.

	Ecoli	AUTO-MPG	Sonar	Balance-Scale	Housing	Ionosphere	Pima
WM_3	53.4(0.3239)	73.8(0.2255)	188.0(0.2735)	80.9(0.3482)	164.6(0.3723)	233.0(0.3945)	111.1(0.3434)
WM_5	120.4(0.2554)	171.7(0.4088)	188.0(0.3974)	563.0(0.3222)	315.6(0.3893)	274.3(0.4058)	426.9(0.4664)
WM_7	198.2(0.3130)	241.4(0.3854)	188.0(0.2138)	563.0(0.2178)	391.7(0.3556)	293.9(0.4617)	600.2(0.3151)
CT1_3	15.3(0.4587)	16.3(0.2483)	16.4(0.2594)	18.3(0.2121)	22.1(0.2539)	23.4(0.3428)	18.9(0.3312)
CT1_5	17.1(0.2444)	20.1(0.3120)	18.7(0.3396)	54.3(0.2978)	23.8(0.3837)	25.9(0.4940)	37.3(0.4189)
CT1_7	19.9(0.3895)	25.9(0.3575)	25.3(0.4519)	68.7(0.4455)	27.3(0.3671)	25.7(0.3494)	58.7(0.4989)
CT2_3	12.3(0.4433)	11.6(0.2143)	13.4(0.2450)	16.6(0.3840)	13.8(0.4893)	18.5(0.3563)	11.6(0.2502)
CT2_5	18.9(0.2593)	17.3(0.3865)	19.6(0.3245)	34.8(0.4002)	13.6(0.2269)	21.3(0.4162)	23.8(0.4127)
CT2_7	19.3(0.2460)	21.7(0.3158)	21.7(0.4077)	52.2(0.4920)	16.1(0.3836)	20.9(0.3531)	31.5(0.2862)

Tabela 7-9 - Número de regras geradas.

	Japanese CS	Breast CW	Vowel	Statlog Vehicle	Car Evaluation
WM_3	321.4(0.3918)	231.9(0.3842)	287.9(0.3128)	411.5(0.3901)	704.1(0.3406)
WM_5	512.1(0.2754)	296.5(0.4695)	649.4(0.4380)	708.2(0.2751)	1556.0(0.3112)
WM_7	553.2(0.3924)	358.4(0.4524)	748.3(0.4899)	756.3(0.2194)	1556.0(0.3556)
CT1_3	28.7(0.4005)	16.7(0.3146)	21.7(0.4644)	34.7(0.3652)	59.3(0.2443)
CT1_5	49.6(0.4426)	22.4(0.4566)	62.8(0.3981)	68.9(0.4494)	73.8(0.3309)
CT1_7	53.1(0.3153)	27.8(0.3720)	65.3(0.2500)	73.4(0.3400)	83.1(0.4524)
CT2_3	19.8(0.2355)	12.5(0.3298)	16.4(0.3410)	25.3(0.3212)	47.4(0.3357)
CT2_5	32.4(0.3417)	16.7(0.3666)	42.3(0.3867)	51.9(0.3955)	61.9(0.4157)
CT2_7	39.8(0.3991)	21.3(0.4448)	49.8(0.3697)	68.7(0.4758)	63.5(0.4602)

Os resultados demonstram que os classificadores fuzzy intervalares do tipo-2 possuem taxa média de classificação superior aos classificadores fuzzy intervalares do tipo-1 e superiores ao método de Wang & Mendel. Percebemos também que o número médio de regras do classificador fuzzy intervalar do tipo-2 é menor para a maioria dos conjuntos de dados.

Com o objetivo de fazer uma comparação mais rigorosa, foi utilizado o teste de Friedman (DEMSAR, 2006). O teste de Friedman é um teste não paramétrico que utiliza o ranking dos algoritmos para cada conjunto de dados separadamente, segundo a sua taxa de classificação, para verificar se a diferença entre diversos classificadores é significativa.

A primeira comparação feita foi entre o classificador fuzzy intervalar do tipo-2 e o método de Wang & Mendel. Para esta comparação foi feito o ranking das três versões de cada tipo de classificador conforme podemos ver na Tabela 7.10, que contém a taxa média de classificação de cada classificador para cada conjunto de dados. O número em negrito refere-se à posição de cada classificador no ranking. O classificador com a maior taxa de classificação recebe valor 1, o segundo melhor classificador recebe valor 2 e assim por diante.

Tabela 7-10 - Ranking de classificação.

	CT2_3	CT2_5	CT2_7	WM_3	WM_5	WM_7
Hayes-Roth	0.9254 1	0.8457 3	0.8654 2	0.6154 4	0.5077 5	0.5000 6
Iris	0.9980 1	0.9884 2	0.9832 3	0.9133 5	0.9400 4	0.8933 6
Servo	0.9754 2	0.9398 3	0.9876 1	0.9063 4	0.7813 5	0.6938 6
Glass	0.8771 2	0.8645 3	0.8915 1	0.5286 6	0.5714 5	0.5905 4
Wine	0.9920 1	0.9532 3	0.9580 2	0.9059 4	0.8647 5	0.4235 6
Statlog(Heart)	0.9364 1	0.8854 2	0.8626 3	0.6778 4	0.5852 5	0.5778 6
Liver-Disorder	0.9412 2	0.9157 3	0.9487 1	0.5235 6	0.5794 4	0.5647 5
Ecoli	0.9322 2	0.9645 1	0.9198 3	0.7152 6	0.7909 4	0.7697 5
AUTO-MPG	0.9640 1	0.8612 2	0.8354 3	0.8256 4	0.7436 5	0.6026 6
Sonar	0.8547 1	0.8123 2	0.7929 3	0.6050 4	0.5550 5	0.4650 6
Balance-Scale	0.8571 2	0.9259 1	0.8469 3	0.8145 4	0.4581 6	0.4710 5
Housing	0.9325 2	0.9346 1	0.8198 3	0.8200 4	0.6300 5	0.5540 6
Ionosphere	0.9145 2	0.9321 1	0.9089 3	0.6514 6	0.8457 4	0.6657 5
Pima	0.8211 2	0.8268 1	0.8169 3	0.6803 5	0.6829 4	0.6776 6
Japanese Credit Screening	0.8954 1	0.8952 2	0.8611 3	0.7908 4	0.7477 5	0.6846 6
Breast Cancer Wisconsin	0.9754 1	0.9637 2	0.9322 3	0.9103 4	0.6941 5	0.6853 6
Vowel	0.9522 2	0.8354 3	0.9655 1	0.6111 6	0.9202 5	0.9354 4
Statlog Vehicle	0.8810 2	0.8966 1	0.8709 3	0.5845 5	0.6071 4	0.5738 6
Car	0.9632 1	0.8954 3	0.9370 2	0.8192 4	0.7000 5	0.6547 6
Média do ranking	1,53	2,05	2,42	4,68	4,74	5,58

Na comparação feita pelo teste de Friedman consideramos como hipótese nula todos os algoritmos serem equivalentes, ou seja, a diferença entre a taxa média de classificação de cada classificador não ser significativa.

O primeiro passo do teste é encontrar o *valor crítico* que será usado como parâmetro de comparação. Esse valor é encontrado consultando a tabela estatística de distribuição-F usando a coluna $k-1$ e linha $(k-1)*(N-1)$, onde k é o número de algoritmos comparados e N é o número de conjunto de dados. Com 6 algoritmos em comparação e 19 conjuntos de dados, o valor crítico encontrado foi $F(5, 90) = 2,32$ utilizando $\alpha = 0,05$.

Para que a hipótese nula seja rejeitada, o valor crítico deve ser menor do que o valor encontrado pelo teste de Friedman, que é calculado pela expressão 7.1.

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1) - \chi_F^2} \quad (7.1)$$

onde χ_F^2 é calculado da seguinte forma:

$$\chi_F^2 = \frac{12N}{k(k+1)} \left[\sum_j R_j^2 - \frac{k(k+1)^2}{4} \right] \quad (7.2)$$

Na expressão 7.2 a variável R_j representa a média de ranking do algoritmo j . Este valor é mostrado na última linha da Tabela 7.10.

A resolução dos cálculos é ilustrada abaixo:

$$\begin{aligned} \chi_F^2 &= \frac{12 \cdot 19}{6(6+1)} \left[(1,53^2 + 2,05^2 + 2,42^2 + 4,68^2 + 4,74^2 + 5,58^2) - \frac{6(6+1)^2}{4} \right] \\ \chi_F^2 &= \frac{228}{42} [87,90 - 73,5] = 78,17 \\ F_F &= \frac{(19-1) \cdot 78,17}{19(6-1) - 78,17} = 83,6 \end{aligned}$$

Como 83,6 é maior que o valor crítico, a hipótese nula é rejeitada e concluímos que a diferença entre o grupo de classificadores comparados pela Tabela 7.10 é significativa. Entretanto, faz-se necessário um segundo teste para verificar o real desempenho entre pares de classificadores. Para isso, vamos comparar estes 6 classificadores entre si utilizando o método Bonferroni–Dunn (Dunn, 1961), que trata

de forma satisfatória os chamados erros do tipo-1 (ou falso positivo), onde a hipótese nula é rejeitada quando na verdade ela é verdadeira. Neste método, cada classificador é comparado dois a dois e dizemos que o classificador A tem diferença significativa em relação ao classificador B se a diferença entre a média de ranking entre A e B for maior que a Diferença Crítica (DC), que é dada pela expressão:

$$DC = q_{\alpha} \sqrt{\frac{k(k+1)}{6N}} \quad (7.3)$$

onde q_{α} é o índice da linha da tabela de valores críticos (DEMSAR, 2006) para testes post-hoc aplicados após o teste de Friedman utilizando o método Bonferroni–Dunn, com α podendo ser 5% ou 10%. Neste teste o valor de α é 5%. O número de algoritmos comparados é usado como índice para consultar a coluna da tabela de valores críticos. Com $q_{0,05}$ e 6 algoritmos, o valor crítico é 2,58 e a diferença crítica será:

$$DC = 2,58 \sqrt{\frac{6(6+1)}{6*19}} = 1,57$$

A Tabela 7.11 é uma tabela simétrica que mostra a diferença entre a média de ranking calculada entre todos os 6 classificadores. Nesta tabela, as células com fundo cinza indicam diferenças significativas (o valor da célula é maior que 1,57).

Tabela 7-11 - Diferença entre média de ranking.

	CT2_3	CT2_5	CT2_7	WM_3	WM_5	WM_7
CT2_3	-	0,52	0,89	3,15	3,21	4,05
CT2_5	0,52	-	0,37	2,63	2,69	3,53
CT2_7	0,89	0,37	-	2,26	2,32	3,16
WM_3	3,15	2,63	2,26	-	0,06	0,90
WM_5	3,21	2,69	2,32	0,06	-	0,84
WM_7	4,05	3,53	3,16	0,90	0,84	-

Podemos concluir após estes testes que os classificadores fuzzy intervalares do tipo-2 possuem taxa de classificação média superior ao método Wang & Mendel. Porém, quando as três versões dos classificadores fuzzy intervalares do tipo-2 são comparadas entre si pode-se verificar que a diferença entre sua taxa média de classificação não é significativa.

O mesmo teste de comparação foi realizado entre os classificadores fuzzy intervalar do tipo-2 e os classificadores fuzzy do tipo-1. Para esta comparação foi feito o ranking das três versões de cada tipo de classificador conforme podemos ver na Tabela 7.12.

Na comparação feita pelo teste de Friedman consideramos como hipótese nula todos os algoritmos serem equivalentes, ou seja, a diferença entre a taxa média de classificação de cada classificador não ser significativa.

Como a quantidade de algoritmos comparados e o número de conjuntos de dados utilizados não mudaram, o valor crítico da tabela de distribuição-F para este teste continua inalterado.

Tabela 7-12 - Ranking de classificação.

	CT2_3	CT2_5	CT2_7	CT1_3	CT1_5	CT1_7
Hayes-Roth	0.9254 1	0.8457 3	0.8654 2	0.8291 4	0.7843 5	0.7711 6
Iris	0.9980 1	0.9884 2	0.9832 3	0.9788 4	0.9654 5	0.9622 6
Servo	0.9754 2	0.9398 5	0.9876 1	0.9541 4	0.8358 6	0.9715 3
Glass	0.8771 2	0.8645 3	0.8915 1	0.6845 6	0.7833 5	0.8310 4
Wine	0.9920 1	0.9532 3	0.9580 2	0.9156 4	0.8685 5	0.8016 6
Statlog(Heart)	0.9364 1	0.8854 2	0.8626 4	0.8741 3	0.8344 5	0.8286 6
Liver-Disorder	0.9412 2	0.9157 3	0.9487 1	0.8571 6	0.8680 4	0.8607 5
Ecoli	0.9322 3	0.9645 1	0.9198 5	0.8739 6	0.9431 2	0.9278 4
AUTO-MPG	0.9640 1	0.8612 3	0.8354 5	0.9067 2	0.8401 4	0.7860 6
Sonar	0.8547 1	0.8123 2	0.7929 3	0.7365 5	0.7463 4	0.6629 6
Balance-Scale	0.8571 3	0.9259 1	0.8469 4	0.8440 5	0.8700 2	0.8380 6
Housing	0.9325 2	0.9346 1	0.8198 5	0.8617 3	0.8212 4	0.7941 6
Ionosphere	0.9145 3	0.9321 1	0.9089 4	0.7806 6	0.9189 2	0.9006 5
Pima	0.8211 2	0.8268 1	0.8169 4	0.7912 5	0.8170 3	0.7808 6
Japanese Credit Screening	0.8954 1	0.8952 2	0.8611 3	0.8354 4	0.8064 5	0.7091 6
Breast Cancer Wisconsin	0.9754 1	0.9637 2	0.9322 3	0.9254 4	0.7453 5	0.7362 6
Vowel	0.9522 2	0.8354 5	0.9655 1	0.7452 6	0.9432 4	0.9483 3
Statlog Vehicle	0.8810 2	0.8966 1	0.8709 4	0.8647 5	0.8718 3	0.8475 6
Car	0.9632 1	0.8954 3	0.9370 2	0.8752 4	0.8414 5	0.8407 6
Média do ranking	1,68	2,32	3,00	4,53	4,11	5,37

A resolução dos cálculos para o teste de Friedman é ilustrado abaixo:

$$\chi_F^2 = \frac{12*19}{6(6+1)} \left[(1,68^2 + 2,32^2 + 3,00^2 + 4,53^2 + 4,11^2 + 5,37^2) - \frac{6(6+1)^2}{4} \right]$$

$$\chi_F^2 = \frac{228}{42} [83,45 - 73,5] = 54,01$$

$$F_F = \frac{(19-1)*54,01}{19(6-1) - 54,01} = 23,71$$

Como 23,71 é maior que o valor crítico, a hipótese nula é rejeitada e concluímos que a diferença entre este grupo de classificadores é significativa.

Para analisarmos melhor a performance entre os 6 classificadores deste grupo vamos usar novamente o método Bonferroni–Dunn para fazer comparações em pares de classificadores.

Como as variáveis da expressão 7.3 não mudaram para este segundo teste, o valor de DC continuará valendo 1,57.

A Tabela 7.13 mostra a diferença entre a média de ranking calculada entre todos os 6 classificadores. Nesta tabela, as células com fundo cinza indicam diferenças significativas (o valor da célula é maior que 1,57).

Tabela 7-13 - Diferença entre média de ranking.

	CT2_3	CT2_5	CT2_7	CT1_3	CT1_5	CT1_7
CT2_3	-	0,64	1,32	2,85	2,43	3,69
CT2_5	0,64	-	0,68	2,21	1,79	3,05
CT2_7	1,32	0,68	-	1,53	1,11	2,37
CT1_3	2,85	2,21	1,53	-	0,42	0,84
CT1_5	2,43	1,79	1,11	0,42	-	1,26
CT1_7	3,69	3,05	2,37	0,84	1,26	-

Pela Tabela 7.4, Tabela 7.5 e Tabela 7.6 podemos ver que a taxa média de classificação dos classificadores fuzzy intervalares do tipo-2 são maiores do que a taxa média de classificação dos classificadores fuzzy do tipo-1 com o mesmo número de conjuntos em cada partição. Porém, observando a Tabela 7.13 concluímos que essa diferença só é significativa em alguns casos.

Classificadores fuzzy intervalares do tipo-2 com 3 e 5 conjuntos possuem taxa média de classificação significativamente maior do que classificadores fuzzy do tipo-1 com 3, 5 e 7 conjuntos. Entretanto, classificadores fuzzy intervalares do tipo-2 com 7 conjuntos não possui uma melhora significativa em relação a taxa média de classificação dos classificadores fuzzy do tipo-1 com 3 e 5 conjuntos, mas possui aumento significativo em relação aos classificadores fuzzy do tipo-1 também com 7 conjuntos.

O mesmo tipo de comparação foi realizado focando o desempenho dos classificadores em relação ao número de regras de cada um. A comparação foi feita apenas entre os classificadores fuzzy gerados pelos Algoritmos Genéticos. O método de

Wang & Mendel foi descartado para este teste por possuir um número muito elevado de regras em comparação aos outros dois tipos de classificadores utilizados neste trabalho.

Na comparação feita pelo teste de Friedman consideramos como hipótese nula todos os algoritmos serem equivalentes, ou seja, a diferença entre o número médio de regras de cada classificador não ser significativa.

Como a quantidade de algoritmos comparados e o número de conjuntos de dados utilizados não mudaram, o valor crítico para o teste de Friedman continua 2,32.

Tabela 7-14 - Ranking de em relação ao número de regras

	CT2(3)	CT2(5)	CT2(7)	CT1(3)	CT1(5)	CT1(7)
Hayes-Roth	12.4 1	14.8 2	17.2 5	16.3 3	16.5 4	19.8 6
Iris	8.7 1	13.9 2	15.3 5	14.4 3	17.8 6	18.3 7
Servo	15.3 1	15.8 2	17.9 4	16.6 3	19.3 6	18.7 5
Glass	10.6 1	12.8 2	13.8 3	23.2 5	25.4 6	22.1 4
Wine	12.7 1	14.4 2	14.6 3	17.4 6	15.2 4	15.8 5
Statlog(Heart)	19.5 1	25.8 3	26.6 4	23.4 2	27.8 5	29.7 6
Liver-Disorder	15.7 1	21.9 3	22.4 4	19.8 2	23.4 5	23.6 6
Ecoli	12.3 1	18.9 4	19.3 5	15.3 2	17.1 3	19.9 6
AUTO-MPG	11.6 1	17.3 3	21.7 5	16.3 2	20.1 4	25.9 6
Sonar	13.4 1	19.6 4	21.7 5	16.4 2	18.7 3	25.3 6
Balance-Scale	16.6 1	34.8 3	52.2 4	18.3 2	54.3 5	68.7 6
Housing	13.8 2	13.6 1	16.1 3	22.1 4	23.8 5	27.3 6
Ionosphere	18.5 1	21.3 3	20.9 2	23.4 4	25.9 6	25.7 5
Pima	11.6 1	23.8 3	31.5 4	18.9 2	37.3 5	58.7 6
Japanese Credit Screening	19.8 1	32.4 3	39.8 4	28.7 2	49.6 5	53.1 6
Breast Cancer Wisconsin	12.5 1	16.7 3	21.3 4	16.7 2	22.4 5	27.8 6
Vowel	16.4 1	42.3 3	49.8 4	21.7 2	62.8 5	65.3 6
Statlog Vehicle	25.3 1	51.9 3	68.7 4	34.7 2	68.9 5	73.4 6
Car	47.4 1	61.9 3	63.5 4	59.3 2	73.8 5	83.1 6
Média do ranking	1,05	2,74	4,00	2,74	4,84	5,79

A resolução dos cálculos é ilustrada a seguir:

$$\chi_F^2 = \frac{12 \cdot 19}{6(6+1)} \left[(1,05^2 + 2,74^2 + 4,00^2 + 2,74^2 + 4,84^2 + 5,79^2) - \frac{6(6+1)^2}{4} \right]$$

$$\chi_F^2 = \frac{228}{42} [89,06 - 73,5] = 84,5$$

$$F_F = \frac{(19-1) \cdot 84,5}{19(6-1) - 84,5} = 144,86$$

Como 144,86 é maior que o valor crítico, a hipótese nula é rejeitada e concluímos que a diferença entre este grupo de classificadores é significativa.

Para analisarmos melhor a performance entre os 6 classificadores deste grupo vamos usar novamente o método Bonferroni–Dunn para fazer comparações em pares de classificadores.

Como as variáveis da expressão 6.3 não mudaram para este segundo teste, o valor de DC continuará valendo 1,57.

A Tabela 7.15 mostra a diferença entre a média de ranking calculada entre todos os 6 classificadores. Nesta tabela, as células com fundo cinza indicam diferenças significativas (o valor da célula é maior que 1,57).

Tabela 7-15 - Diferença entre média de ranking.

	CT2_3	CT2_5	CT2_7	CT1_3	CT1_5	CT1_7
CT2_3	-	1,69	2,95	1,69	3,79	4,74
CT2_5	1,69	-	1,26	-	2,10	3,05
CT2_7	2,95	1,26	-	1,26	0,84	1,79
CT1_3	1,69	-	1,26	-	2,10	3,05
CT1_5	3,79	2,10	0,84	2,10	-	0,95
CT1_7	4,74	3,05	1,79	3,05	0,95	-

Pela Tabela 7.7, Tabela 7.8 e Tabela 7.9 podemos ver que o número médio de regras dos classificadores fuzzy intervalares do tipo-2 são menores do que o número médio de regras dos classificadores fuzzy do tipo-1 com o mesmo número de conjuntos em cada partição. Porém, observando a Tabela 7.15 podemos perceber que há casos em que essa diferença não é significativa.

Classificadores fuzzy intervalares do tipo-2 com 3 conjuntos possuem número médio de regras significativamente menor do que os outros classificadores utilizados neste trabalho. Já o classificador fuzzy intervalar do tipo-2 com 5 conjuntos possui ranking médio igual ao classificador fuzzy do tipo-1 com 3 conjuntos e a diferença entre o número médio de regras destes classificadores não é significativa.

Classificadores fuzzy intervalares do tipo-2 com 7 conjuntos não possuem uma melhora significativa em relação ao número médio de regras dos classificadores fuzzy do tipo-1 com 3 e 5 conjuntos, mas possui aumento significativo em relação aos classificadores fuzzy do tipo-1 também com 7 conjuntos.

7.3 Considerações Finais

Neste capítulo foram discutidos os experimentos realizados com as duas versões do método e com o método de Wang & Mendel. No capítulo seguinte serão apresentadas as conclusões do trabalho.

8. Conclusões

Este trabalho apresentou um método para a geração de classificadores fuzzy intervalares do tipo-2. O processo de geração é composto por três etapas. Na primeira etapa são geradas as regras utilizando algoritmos genéticos. O melhor conjunto de regras gerado é utilizado na segunda etapa, onde outro AG é utilizado para otimizar o formato das funções de pertinência. Por fim, na terceira etapa outro AG é utilizado para otimizar o número de regras da base de regras.

Cada AG utilizado no processo de geração de classificadores fuzzy intervalar do tipo-2 é auto-adaptativo, evitando a convergência prematura dos algoritmos.

Os testes evidenciam um bom desempenho dos classificadores fuzzy intervalares do tipo-2 gerados pelo método. A eficiência destes classificadores foi verificada nos experimentos apresentados, mostrando diferença significativa entre classificadores fuzzy intervalares do tipo-2 e classificadores fuzzy do tipo-1.

Embora a eficiência da abordagem proposta tenha sido comprovada, o desenvolvimento deste trabalho evidenciou a necessidade de aprofundar a investigação de diversos aspectos que deverão ser considerados em trabalhos futuros, tais como:

- Utilizar métodos diferentes para a geração inicial da base de dados, como o uso de histogramas, método proposto por (CHOIA et al., 2009) e comentado no capítulo 5;
- Fazer uso de variáveis do tipo “don’t care” nas regras fuzzy com o objetivo de reduzir o antecedente das regras e assim aumentar a interpretabilidade;
- Testar o uso de outros métodos de inferência.

9. Referências Bibliográficas

BARALDI, A.; BLONDA, P. “A survey of fuzzy clustering algorithms for patterns recognition - Part I”. IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics 29, pp. 778–785, 1999a.

BARALDI, A.; BLONDA, P. “A survey of fuzzy clustering algorithms for patterns recognition - Part II”. IEEE Trans. on Systems, Man, and Cybernetics - Part B: Cybernetics 29, pp. 786–801, 1999b.

BERTHOLD, M. Intelligent Data Analysis: An Introduction. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1999, 525p.

BISHOP, C. M. Neural Networks for Pattern Recognition. New York, USA: Oxford University Press, Inc., 1995, 482 p.

BONARINI, A. “Evolutionary Learning of Fuzzy rules: competition and cooperation”. Kluwer Academic Press, Norwell, MA, pp. 265–284, 1996.

BONISSONE, P.; KHEDKAR, P. S.; CHEN, Y. T. “Genetic algorithms for automated tuning of fuzzy controllers, a transportation application”. IEEE International Conference on Fuzzy Systems, pp. 674–680, 1996.

CAO, J.; LIU, H.; LI, P.; BROWN, D. "Adaptive Fuzzy Logic Controller for Vehicle Active Suspensions with Interval Type-2 Fuzzy Membership Functions". Proceedings of the Institution of Mechanical Engineers, v. 222, pp. 1361-1373, 2008.

CASILLAS, J. et al. “Genetic feature selection in a fuzzy rule-based classification system learning process for high-dimensional problems”. Information Sciences. v. 136, pp. 135–157, 2001.

CASTELLANO, G.; FANELLI, A. “Information granulation via neural network-based learning”. Proceedings of the 9th IFSA World Congress and the 20th NAFIPS Int. Conference, Vancouver, Canada, pp. 3059-3064, 2001.

CASTRO, P. A. D.; CAMARGO, H. A. “Learning and optimization of fuzzy rule base by means of self adaptative genetic algorithms”. IEEE International Conference on Fuzzy Systems, v. 2, pp. 1037-1042, 2004.

CASTRO, P. A. D., CAMARGO, H. A., NICOLETTI, M. C., SANTORO. “Improving a Pittsburg learnt fuzzy rule base using feature subset selection”. 4th Internacional Conference on Hybrid Intelligent Systems, pp. 180-185, 2004.

CASTRO, P. A. D.; CAMARGO, H. A. “Improving the genetic optimization of fuzzy rule base by imposing a constraint condition on the number of rules”. V Encontro Nacional de Inteligência Artificial, pp. 972-981, 2005.

CEYLAN, R., OZBAY, Y., KARLIK, B. "A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network". Expert systems with applications, v. 3, pp. 6721-6726, 2009.

CHI, Z., YAN, H., PHAM, T. "Fuzzy algorithms with applications to image processing and pattern recognition". World Scientific, 1996.

CHOIA, B-IN., RHEE, F. C-HOON. "Interval type-2 fuzzy membership function generation methods for pattern recognition", Information Sciences, v. 179, pp. 2102-2122, 2009.

CHUA, T. W., TAN, W. W. "Genetically evolved fuzzy rule-based classifiers and application to automotive classification", SEAL 2008, v. 5361, pp. 101-110, 2008.

CINTRA, M. E.; CAMARGO, H. A. "Fuzzy Rules Generation using Genetic Algorithms with Self-adaptive Selection". IEEE International Conference on Information Reuse and Integration v. 1, pp. 261-266, 2007.

CORDÓN, O. et al. "Genetic fuzzy systems – evolutionary tuning and learning of fuzzy knowledge bases". Advances in Fuzzy Systems - Applications and Theory. v. 19, 2001.

CORDÓN, O. et al. "Ten years of genetic fuzzy systems: Current framework and new trends". Fuzzy Sets and Systems. v. 141, pp. 5–31, 2004.

CORDÓN, O.; HERRERA, F. "A three stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples". International Journal of Approximate Reasoning. v. 7, pp. 369–407, 1997.

CORDÓN, O.; HERRERA, F. "Hybridizing genetic algorithms with sharing scheme and evolution strategies for designing approximate fuzzy rule-based systems". Fuzzy Sets and Systems. v. 118, pp. 235–255, 2001.

CORDÓN, O.; HERRERA, F.; HOFFMANN, F.; MAGDALENA, L. "Genetic Fuzzy Systems". Evolutionary tuning and learning of fuzzy knowledge bases. World Scientific, 2001.

CORDÓN, O.; ALCALÁ, R.; ALCALÁ-FDEZ, J.; ROJAS, I. "Special section on genetic fuzzy systems". IEEE Transactions on Fuzzy Systems, v. 15, pp. 533-592, 2007.

COUPLAND, S.; JOHN, R. "A new and efficient method for the type-2 meet operation". IEEE International Conference on Fuzzy Systems, pp. 959-964, 2004.

COUPLAND, S.; JOHN, R. "Towards more efficient type-2 fuzzy logic systems". Proceedings of IEEE FUZZ Conference, pp. 236–241, 2005.

COUPLAND, S. "Type-2 Fuzzy Sets: Geometric Defuzzification and Type-Reduction". IEEE Symposium on Foundations of Computational Intelligence, v. 1, pp. 622-629, 2007.

DARWIN, Charles Robert. "The origin of species: by means of natural selection". London, 1872, 512 p.

DAVIS, L.D. "Handbook of Genetic Algorithms, Artificial Intelligence", v. 100, 1998, pp. 325-330, 1998.

DELGADO, M. R. "Projeto Automático de Sistemas Fuzzy: Uma Abordagem Co-Evolutiva". Tese — Universidade Estadual de Campinas, 2002.

DRIANKOV, D., HELLENDORRN, H., REINFRANK, M. "An Introduction to Fuzzy Control. Berlin", Germany: Springer-Verlag, 1993.

DUBOIS, D.; PRADE, H. "Fuzzy Sets and Systems: Theory and applications". Academic Press, 1980.

DUBOIS, D.; PRADE, H. "Soft computing, fuzzy logic and artificial intelligence". Soft Computing, v. 2, pp. 7–11, 1998.

DUDA, R. O.; HART, P. E. "Pattern classification and scene analysis". John Wiley, 1973, 482p.

GARIBALDI, J. M.; MUSIKASUWAN, S.; OZEN, T. "The association between non-stationary and interval type-2 fuzzy sets: a case study". IEEE FUZZ Conference, pp. 224–229, 2005.

GEYER-SCHULTZ, A. "Fuzzy rule-based expert systems and genetic machine learning". Heidelberg: Physica-Verlag, 1997, 432p.

GOLDBERG, D. E. "Genetic Algorithms in Search, Optimization and Machine Learning". Addison-Wesley Professional, 1989, 432 p.

GONZÁLEZ, A., PÉREZ, R. "SLAVE: A genetic learning system based on an iterative approach". IEEE Transactions on Fuzzy Systems, v. 7, pp. 176-191, 1999.

HAGRAS, H. "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots", IEEE Transaction on Fuzzy Systems. v. 12, pp. 524–539, 2004.

HERRERA, F.; LOZANO, M.; VERDEGAY, J. L. "Tuning fuzzy logic controllers by genetic algorithms". International Journal of Approximate Reasoning. v. 12, pp. 299–315, 1995.

HOFFMANN, F. "Combining boosting and evolutionary algorithms for learning of fuzzy classification rules". Fuzzy Sets and Systems, v. 141, pp. 47-58, 2004.

HUFFMAN, F.; PFISTER, G. "Evolutionary design of a fuzzy knowledge base for a mobile robot". International Journal of Approximate Reasoning. v. 17, pp. 447–469, 1997.

HWANG, C. F. Rhee, "An interval type-2 fuzzy spherical shells algorithm". IEEE International Conference Fuzzy Systems, 2004.

HWANG, W. R.; THOMPSON, W. E. "Design of fuzzy logic controllers using genetic algorithms". *International Conference Fuzzy Systems*, v. 3, pp. 1708–1712, 1994.

ISHIBUCHI, H., NOZAKI, K., TANAKA, H., "Distributed representation of fuzzy rules and its application to pattern classification". *Fuzzy Sets and Systems*, v. 52, pp. 21-23, 1992.

ISHIBUCHI, H., NOZAKI, K., YAMAMOTO, N., TANAKA, H., "Selecting fuzzy if-then rules for classification problems using genetic algorithms". *IEEE Transaction on Fuzzy Systems*. v. 3, pp. 260-270, 1995.

ISHIBUCHI, H. et al. "Empirical study on learning in fuzzy systems by rice taste analysis". *Fuzzy Sets and Systems*, v. 64, pp. 129–144, 1994.

ISHIBUCHI, H.; NAKASHIMA, T., "Effect of rule weights in fuzzy rule based classification systems". *IEEE Transaction on Fuzzy Systems*, v. 9, pp. 506–515, 2001.

ISHIBUCHI, H.; NAKASHIMA, T.; MURATA, T. "Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems". *IEEE Transaction On Fuzzy Systems*, v. 29, pp. 601–618, 1999.

ISHIBUCHI, H.; YAMAMOTO, T. "Rule Weight Specification in Fuzzy Rule-Based Classification Systems". *IEEE Trans. Fuzzy Syst.*, v. 13, pp. 428–435, 2005.

JAIN, L. C.; MARTIN, N. M., "Fusion of Neural Networks, Fuzzy Sets, and Genetic Algorithms - Industrial Applications". *The CRC Press International Series on Computational Intelligence*, 1998.

JANG, J. R., SUN, C. T., "Predicting chaotic time series with fuzzy IF-THEN Rules", *IEEE International Conference on Fuzzy Systems*, pp. 1079-1084, 1993.

JANG, J.-S. R.; SUN, C.-T.; MIZUTANI, E. "Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence". Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1997.

JESUS, M. J. et al. "Induction of fuzzy-rule-based classifiers with evolutionary boosting algorithms". *IEEE Transactions on Fuzzy Systems*. v. 12, pp. 296–308, 2004.

JOHN R. I., INNOCENT, P. R., BARNES, M. R. "Neuro-fuzzy clustering of radiographic tibia image data using type-2 fuzzy sets," *Information Sciences*, v. 125, pp. 65–82, 2000.

KANDEL, A. "Fuzzy expert systems". Boca Raton, FL, USA: CRC Press, 1992.

KARNIK, N. N.; MENDEL, J. M. "Type-2 Fuzzy Logic Systems: Type-Reduction". *IEEE International Conference on Systems, Man, and Cybernetics*, v. 2, pp. 2046-2051, 1998.

KARNIK, N. N.; MENDEL, J. M.; LIANG, Q. "Type-2 Fuzzy Logic Systems". IEEE Transactions on Fuzzy Systems, v. 7, pp. 327-348, 1999.

KARR, C. L. "Design of an adaptive fuzzy logic controller using a genetic algorithm". International conference on genetic algorithms, v. 4, pp. 450-457, 1991.

KECMAN, V. "Learning and Soft Computing: Support Vector Machines, Neural Networks and Fuzzy Logic Models". MIT Press, 2001, 608 p.

KIM, M.; KIM, C.; LEE, J. "Evolving compact and interpretable takagi-sugeno fuzzy models with a new encoding scheme". IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics. v. 36, pp. 1006-1023, 2006.

KLEMENT, E. P.; MESIAR, R.; PAP, E. "Triangular Norms: Kluwer, Dordrecht", 2000. (Trends in Logic, v. 8).

KLIR, G. J.; FOLGER, T. A. Fuzzy Sets, "Uncertainty and Information. Englewood Cliffs", NJ. Prentice Hall, 1988.

KLIR, G. J.; YUAN, B. "Fuzzy Sets and Fuzzy Logic: theory and applications", 1. ed. Prentice-Hall, 1995, 574 p.

KOSKO, B. "Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence". Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1992.

LARSEN, P. M. "Industrial applications of fuzzy logic control". Int. J. Man-Machine Stud., v. 12, pp. 3-10, 1980.

LEE, C. C. "Fuzzy logic in control systems: fuzzy logic controller Part I and Part II," IEEE Transaction on Systems, Man, Cybernetics, v. 20, pp. 404-435, 1990.

LEE, C. C. "Fuzzy logic in control systems: Fuzzy logic controller: Parts 1 and 2". IEEE Transactions on Systems, Man and Cybernetics, v. 20, pp. 404-435, 1990.

LEONDES, C. T. "Fuzzy Theory Systems: Techniques and Applications", v. 1, pp. 1-4, 1999.

LIANG, Q., MENDEL, J. M. "MPEG VBR video traffic modeling and classification using fuzzy technique". IEEE Transaction on Fuzzy Systems, v. 9, pp. 183-193, 2001.

LIANG, Q., MENDEL, J. M. "Interval type-2 fuzzy logic systems: theory and design". IEEE International Conference on Fuzzy Systems, v. 8, pp. 535-550, 2000.

LIANG, Q.; MENDEL, J. M. "MPEG VBR video traffic modeling and classification using fuzzy technique". IEEE Transaction on Fuzzy Systems, v. 9, pp. 183-193, 2001.

LIAO, T.W., LI, D. "Two manufacturing applications of the fuzzy K-NN algorithm". Fuzzy Sets and Systems, v. 92, pp. 289-303, 1997.

LIN, C.-T.; LEE, C. S. G. "Neural fuzzy systems: a neuro-fuzzy synergism to intelligent systems". Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.

LUKASIEWICZ, J. "Philosophical remarks on many-valued systems of propositional logic". North Holland, 1970.

MAMDANI, E. H.; ASSILIAN, S. "An experiment in linguistic synthesis with a fuzzy logic controller". *Int. J. Hum.-Comput. Stud.*, v. 51, pp. 135–147, 1999.

MAMDANI, E. H.; ASSILIAN, S. "An experiment in linguistic synthesis with a fuzzy logic controller". *International Journal of Man-Machine Studies*, v. 7, pp. 1–13, 1975.

MENDEL, J. M. "On the importance of interval sets in type-2 fuzzy logic systems". 9th IFSA World Congress and 20th NAFIPS Conf., pp. 1647–1652, 2001.

MENDEL, J. M. "Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions", Ed. Prentice Hall, USA, 2000.

MENDEL, J. M. "The Perceptual Computer: An Architecture for Computing With Words". *IEEE International Conference on Fuzzy Systems*, v. 1, pp. 35-38, 2001.

MENDEL, J. M.; JOHN, R. I. B. "Type-2 Fuzzy Sets Made Simple", *IEEE Transactions on Fuzzy Systems*, v. 10, pp. 117-127, 2002.

MENDEL, J. M. "Advances in type-2 fuzzy sets and systems". *Information sciences*, v. 177, pp. 84-110, 2007.

MICHALEWICZ, Z. "Genetic Algorithms + Data structures = Evolution Programs". Springer, 1996.

MITCHELL, H. "Pattern recognition using type-II fuzzy sets". *Information Sciences*, v. 170, pp. 409–418, 2005.

MITCHELL, H. "Ranking type-2 fuzzy numbers". *IEEE Transaction On Fuzzy Systems*, v. 14, pp. 287-294, 2006.

MIZUMOTO, M.; TANAKA, K. "Some properties of fuzzy sets of type-2". *Information and Control*, v. 31, pp. 312-340, 1976.

MIZUMOTO, M.; TANAKA, K. "Fuzzy sets of type-2 under algebraic product and algebraic sum". *Fuzzy Sets and Systems*, v. 5, pp. 277-290, 1981.

NAUCK, D.; KRUSE, R. "A neuro-fuzzy method to learn fuzzy classification rules from data. *Fuzzy Sets and Systems*", v. 89, pp. 277–288, 1997.

NGUYEN, C. T.; GANESH, C.; GONG, K. F. "A fuzzy logic-based intelligent controller for contact management data integration". *IEEE World Congress on Computational Intelligence*. v. 2, pp. 879–884, 1998.

NICOLETTI, M. C.; CAMARGO, H. A. “Fundamentos da Teoria de Conjuntos Fuzzy”. São Carlos: EdUFSCar, 2004, 65 p. (Série Apontamentos).

PARK, S., LEE-KANG, H., “A designing method for type-2 fuzzy logic systems using genetic algorithms”. IFSA World Congress and 20th NAFIPS International Conference, v. 5, pp. 2567-2572, 2001.

PEDRYCZ, W.; GOMIDE, F. “An Introduction to Fuzzy Sets: Analysis and Design”. MIT Press, 1998.

PIRES, M. G.; CAMARGO, H. C. “Genetic learning and optimization of fuzzy sets in fuzzy rule-based systems”. IEEE International Conference on Information Reuse and Integration. pp. 623–628, 2004.

QUINLAN, J. R. “Induction of decision trees”. Machine Learning. v. 1, pp. 81–106, 1986.

SANCHEZ, E., SHIBATA, T., ZADEH, L. A., “Genetic Algorithms and Fuzzy Logic Systems”. World Scientific, 1997.

SPIEGEL, D.; SUDKAMP, T. “Sparse data in the evolutionary generation of fuzzy models”. Fuzzy Sets and Systems. v. 138, pp. 363–379, 2003.

SUGENO, M. “An introductory survey of fuzzy control”. Information Science, v. 36, pp. 59–83, 1985.

SUGENO, M.; KANG, G. T. “Structure identification of fuzzy model”. Fuzzy Sets and Systems, v. 28, pp. 15–33, 1988.

TAKAGI, T.; SUGENO, M. “Fuzzy identification of systems and its applications to modeling and control”. IEEE Transactions on Systems, Man and Cybernetics, v. 15, pp. 116–132, 1985.

TAKAGI, T.; SUGENO, M. “Fuzzy identification of systems and its applications to modeling and control”. Readings in Fuzzy Sets for Intelligent Systems. pp. 387–403, 1993.

TAN W. W., FOO C. L., CHUA T. W. “Type-2 Fuzzy System for ECG Arrhythmic Classification”, FUZZ-IEEE 2007, pp. 859 - 864, 2007.

TSUKAMOTO, Y. “An approach to fuzzy reasoning method”. Readings in Fuzzy Sets for Intelligent Systems. pp. 523–529, 1993.

WAGNER, C., HAGRAS, H. “A Genetic Algorithm Based Architecture for Evolving Type-2 Fuzzy Logic Controllers for Real World Autonomous Mobile Robots”, FUZZ-IEEE, pp. 1-6, 2007.

WANG, L., MENDEL, J. “Generating fuzzy rules by learning from examples”. IEEE Transaction on SMC, v. 22, pp. 414–427, 1992.

WONG, C.-C., CHEN, C.-C., LIN, B.-C., “Design of fuzzy classification system using genetic algorithms”. v. 1, pp. 297 – 301, 2000.

WONG, C.C.; CHEN, C.C.; LIN, B.C. “Design of fuzzy classification system using genetic algorithms”. IEEE International Conference on, v. 1, pp. 297-301, 2000.

WU, H., MENDEL, J. M. “Classification of battlefield ground vehicles using acoustic features and fuzzy logic rule-based classifiers”, IEEE Transaction on Fuzzy Systems, v. 15, pp. 56-72, 2006.

ZADEH, L. A. “Outline of a new approach to the analysis of complex systems and decision processes”. IEEE Transaction Systems, Man and Cybernetics, SMC-3, pp. 28–44, 1973.

ZADEH, L. A. “The Concept of A Linguistic Variable and Its Application to Approximate Reasoning-I”. Information Sciences, v. 8, pp. 199 – 249, 1975.

ZADEH, L. “Fuzzy Sets”. Information and Control, v. 3, pp. 338–353, 1965.

ZADEH, L. “The concept of a linguistic variable and its application to approximate reasoning”. Information Sciences, v. 1, pp. 119–249, 1975.

ZEND, J.; LIU, Z.-Q. “Type-2 fuzzy sets for pattern classification: A Review”. IEEE Symposium on Foundations of Computational Intelligence, pp. 192-200, 2007.

ZENG, J., LIU, Z.-Q., “Type-2 Fuzzy Sets for Pattern Classification: A Review”. Foundations of Computational Intelligence. IEEE Symposium on. pp. 193-200, 2007.

ZENG, J.; LIU, Z.-Q. “Interval type-2 fuzzy hidden Markov models”. Proc. FUZZ-IEEE, pp. 1123–1128, 2004a.

ZENG, J.; LIU, Z.-Q. “Type-2 fuzzy hidden Markov models to phoneme recognition”. International Conference on Pattern Recognition, v. 1, pp. 192–195, 2004b.

ZENG, J.; LIU, Z.-Q. “Type-2 fuzzy hidden Markov models and their application to speech recognition”. IEEE Transaction on Fuzzy Systems, v. 14, pp. 454–467, 2006a.

ZENG, J.; LIU, Z.-Q. “Type-2 fuzzy Markov random fields and their application to handwritten Chinese character recognition”. IEEE Transaction on Fuzzy Systems, v. 16, pp. 747-760, 2006b.

ZENG, J.; LIU, Z.-Q. “Type-2 fuzzy Markov random fields to handwritten character recognition”. International Conference on Pattern Recognition, pp. 1162–1165, 2006c.

ZENG, J.; LIU, Z.-Q. “Type-2 fuzzy sets for handling uncertainty in pattern recognition”. Proc. FUZZ-IEEE, pp. 6597–6602, 2006.

ZIMMERMANN, H.; ZYSNO, P. “Latent connectives in human decision making”. Fuzzy Sets and Systems, v. 4, pp. 37–51, 1980.

ZIMMERMANN, H. J. "Fuzzy set theory and its applications", 2. ed. Boston, USA: Kluwer Academic Publishers, 1991, 399 p.