

Universidade Estadual de Campinas

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO
CIENTÍFICA

Mestrado em Matemática Aplicada

Um Estudo das Ligações entre Memórias
Associativas Fuzzy Implicativas e Equações
Relacionais Fuzzy com Aplicações

Autor: Rodolfo Miyasaki

Orientador: Prof. Dr. Peter Sussner

Campinas - SP

Maio/2008

Um Estudo das Ligações entre Memórias Associativas Fuzzy Implicativas e Equações Relacionais Fuzzy com Aplicações.

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por **Rodolfo Miyasaki** e aprovada pela comissão julgadora.

Campinas, 14 de Maio de 2008.



Prof. Dr: Peter Sussner
Orientador

Banca Examinadora:

- 1- Fernando Gomide, Dr. DCA / FEEC / Unicamp
- 2- Laécio Carvalho Barros, Dr. DMA / IMECC / Unicamp
- 3- Peter Sussner, Dr. DMA / IMECC / Unicamp

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de MESTRE em Matemática Aplicada.

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**

Bibliotecária: Crislene Queiroz Custódio – CRB8a 162/2005

Miyasaki, Rodolfo.

M699e Um estudo das ligações entre memórias associativas fuzzy implicativas e equações relacionais fuzzy com aplicações / Rodolfo Miyasaki -- Campinas, [S.P. : s.n.], 2008.

Orientador : Peter Sussner.

Dissertação (Mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Memória associativa. 2. Morfologia matemática. 3. Conjuntos difusos. 4. Redes neurais.. I. Sussner, Peter. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Título em inglês: An investigation of the relationship between implicative fuzzy associative memories and fuzzy relational with applications

Palavras-chave em inglês (Keywords): 1. Associative memory. 2. Mathematical morphology. 3. Fuzzy sets. 4. Neural networks.

Área de concentração: Matemática Aplicada.

Titulação: Mestre em Matemática Aplicada.

Banca examinadora: Prof. Dr. Peter Sussner (IMECC/UNICAMP)
Prof. Dr. Fernando Antônio Campos Gomide. (FEEC/UNICAMP)
Prof. Dr. Laécio Carvalho Barros (IMECC/UNICAMP)

Data da defesa: 14/05/2008

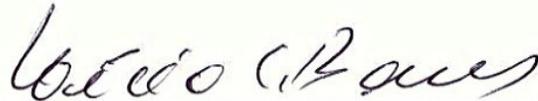
Programa de Pós-Graduação: Mestrado em Matemática Aplicada

Dissertação de Mestrado defendida em 14 de maio de 2008 e aprovada

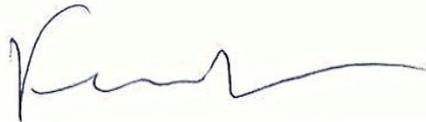
Pela Banca Examinadora composta pelos Profs. Drs.



Prof. (a). Dr (a). PETER SUSSNER



Prof. (a). Dr (a). LAÉCIO CARVALHO DE BARROS



Prof. (a). Dr (a). FERNANDO ANTONIO CAMPOS GOMIDE

Resumo

As *memórias associativas* (AMs - Associative Memories) permitem armazenar associações de padrões e recuperar desejados padrões de saída mesmo após a apresentação de possíveis versões incompletas e/ou distorcidas de um padrão de entrada. As *memórias associativas fuzzy* (FAMs - Fuzzy Associative Memories) são modelos de AMs cujos padrões de entrada e saída são conjuntos fuzzy. As FAMs mostraram-se poderosas ferramentas na implementação em sistemas de base de regras fuzzy. O fato de modelos de FAMs estarem relacionadas à *morfologia matemática* (MM) levou ao recente desenvolvimento das *memórias associativas morfológicas fuzzy* (FMAMs - Fuzzy Morphological Associative Memories), em particular as *memórias associativas fuzzy implicativas* (IFAMs - Implicative Fuzzy Associative Memories). Os neurônios da FMAM executam uma das operações elementares da MM, i.é, erosão, dilatação, anti-erosão ou anti-dilatação.

Essa dissertação relaciona a existência de soluções nos sistemas de *equações relacionais fuzzy* (FREs - Fuzzy Relational Equations) à recordação perfeita das IFAMs. Formulamos o problema de escolher um modelo apropriado de IFAM para uma dada aplicação através de um problema de otimização. Mais precisamente, determinamos o modelo de IFAM dado pela t-norma parametrizada de Yager que minimiza o erro entre os padrões recordados e os desejados padrões de saída. Uma imagem em tons de cinza pode ser expressa como uma *relação fuzzy* e dado uma família de conjuntos fuzzy, pode-se comprimi-la através de FREs. Assim, surge o problema inverso de encontrar uma reconstrução da imagem original a partir da imagem comprimida. Essa dissertação de mestrado determina a melhor aproximação por meio de uma IFAM.

Palavras chave: Memória Associativa, Morfologia Matemática, Conjuntos Fuzzy, Redes Neurais.

Abstract

Associative Memories (AMs) allow for the storage of pattern associations and the retrieval of the desired output patterns upon the presentation of a possibly noisy or incomplete version of an input pattern. Fuzzy Associative Memories (FAMs) are models of AMs whose input and output patterns are fuzzy sets. FAMs have proven to be a powerful tool for implementing fuzzy rule-based systems. The fact that FAMs models are related to mathematical morphology (MM) has led to the development of fuzzy morphological associative memories (FMAMs), in particular fuzzy implicative fuzzy associative memories (IFAMs). The neurons of an FMAM perform one of the elementary operations of MM which are erosion, dilation, anti-erosion and anti-dilation.

This thesis relates the existence of solutions in systems of fuzzy relational equations (FREs) to the perfect recall using IFAMs. We formulated the problem of choosing an appropriate IFAM model for a given application as an optimization problem. More precisely, we determined the IFAM model given by a parameterized Yager t-norm which minimizes the error between the recalled patterns and the desired output patterns. A gray-scale image can be expressed as a fuzzy relation and, given a family of fuzzy sets, it can be compressed by means of FREs. Thus, the inverse problem arises of finding a reconstruction of the image original based on the compression. This master thesis determines the best approximation by means of a IFAMs.

Keywords: Associative Memory, Mathematical Morphology, Fuzzy Sets, Neural Networks.

*À minha família,
Andréia Roberto Miyasaki e
Gustavo Roberto Miyasaki*

Agradecimentos

Aos meus pais Ywao Miyasaki e Luiza Mayumi Sakita Miyasaki, pela sólida formação dada até minha juventude, que me proporcionou a continuidade nos estudos até a chegada a este mestrado, meus eternos agradecimentos.

À minha família, pelo incentivo durante esta jornada, com os sacrifícios que só ela sabe quais foram. Em especial, à minha esposa Andréia Roberto Miyasaki, pela paciência e companherismo durante nossa, ainda, curta vida de casados. Aos meus irmãos Rafaela Miyasaki e Rodrigo Miyasaki e sua esposa Kayse Alvez Miyasaki, bem como minha avó Emiko Sakita que sempre estiveram presentes nas horas mais difíceis de minha vida. Não posso deixar de citar duas pessoas, Aparecido Roberto e Antônia dos Reis Roberto sempre dispostos a me ajudar. E aos mais novos membros da família, Gustavo Roberto Miyasaki e Pedro Henrique Alvez Miyasaki que tanto nos alegram.

A todos meus amigos, professores e funcionários da Unicamp que de alguma forma contribuíram para realização deste trabalho.

A sua disponibilidade irrestrita, sua forma exigente, crítica e criativa de argüir as idéias apresentadas, creio que deram norte a este trabalho, facilitando o alcance de seus objetivos. A Peter Sussner, sou grato pela orientação.

À FAPESP, pelo apoio financeiro a esta dissertação de mestrado, n. processo 05/03648-5.

A todos agradeço, profundamente, e dedico o resultado do trabalho.

Sumário

1	Introdução	1
2	Morfologia Matemática, Conjuntos Fuzzy e Morfologia Matemática Fuzzy	3
2.1	Introdução a Morfologia Matemática	3
2.1.1	Introdução	3
2.1.2	Conceitos Fundamentais de Reticulados Completos	4
2.1.3	Operadores Básicos da Morfologia Matemática	5
2.1.4	Relações de Dualidade - Adjunção e Negação	6
2.2	Conceitos Relevantes da Teoria dos Conjuntos Fuzzy	8
2.2.1	Introdução	8
2.2.2	Conectivos Básicos da Lógica Fuzzy	9
2.2.3	Relações de Dualidade entre Operadores Fuzzy	11
2.2.4	Produtos Matriciais Fuzzy	13
2.2.5	Equações Relacionais Fuzzy	14
2.3	Morfologia Matemática Fuzzy	16
2.3.1	Medidas de Inclusão e Intersecção Fuzzy	16
2.3.2	Erosão Fuzzy e Dilatação Fuzzy	17
3	Redes Neurais Artificiais e Memórias Associativas	19
3.1	Redes Neurais Artificiais	19

3.1.1	Contexto Histórico	19
3.1.2	Introdução	20
3.1.3	Topologia de uma Rede Neural Artificial	21
3.1.4	Aprendizagem	23
3.2	Redes Neurais Fuzzy Morfológicas	24
3.3	Memórias Associativas	24
4	Memórias Associativas Morfológicas e Memórias Associativas Morfológicas Fuzzy	27
4.1	Memórias Associativas Morfológicas	27
4.2	Memórias Associativas Morfológicas Fuzzy	35
4.2.1	Introdução	35
4.2.2	Neurônios Max- C , Min- D e Min- J	35
4.2.3	FMAMs Max- C	37
4.2.4	FMAMs Min- D , a Negação das FMAMs max- C	37
4.2.5	Aprendizado por Adjunção	38
4.3	Memórias Associativas Fuzzy Implicativas	41
4.3.1	Introdução	41
4.3.2	Mémória Associativa Fuzzy Implicativa	42
4.3.3	Mémória Associativa Fuzzy Implicativa Dual	44
4.3.4	Memórias Implicativas Fuzzy Autoassociativas	45
5	Existência de Soluções nos Sistemas de FREs em Termos das IFAMs	49
5.1	Introdução	49
5.2	Uma Comparação entre FREs e as Composições max- C , min- D e min- J	49
5.3	Relação entre FREs e IFAMs	51
6	Determinação do Melhor Modelo de IFAM de Yager para uma Aplicação	57
6.1	Introdução	57
6.2	Formulação do Problema	57

6.3	Aplicações de IFAMs de Yager em Séries Temporais	60
6.3.1	Previsão da Mão-de-obra em Indústrias Metalúrgicas	60
6.3.2	Previsão da Vazão Mensal de uma Usina Hidrelétrica	67
7	Aplicação de IFAMs no Processo de Descompressão de Imagens	75
7.1	Introdução	75
7.2	Processo de Compressão/Descompressão	75
7.3	Experimentos	83
7.4	O Processo de Reconstrução de Imagens e a Teoria das Memórias Associativas Fuzzy Implicativas	85
7.4.1	Considerações Iniciais	85
7.4.2	Formulação do Problema	86
7.5	Experimentos	91
7.6	Formulação do Problema de Otimização	93
8	Conclusão	101
	Referências Bibliográficas	102

CAPÍTULO 1

Introdução

A *Morfologia Matemática* (MM) foi desenvolvida por George Matheron e Jean Serra nos anos 60 com o objetivo de extrair informações de imagens binárias utilizando os operadores de dilatação e erosão a partir de elementos predefinidos, denominados elementos estruturantes [59, 42]. Posteriormente, os principais operadores da morfologia matemática binária foram estendidos para o caso de tons de cinza [20, 66, 73, 74]. Tais abordagens empregam a teoria dos conjuntos fuzzy, formulada por Lotfi Zadeh em 1965, como uma ferramenta para modelar a imprecisão [82]. As abordagens binárias e tons de cinza, possuem como base a teoria dos reticulados completos que representam o contexto mais geral onde a morfologia matemática pode ser conduzida [29, 56, 60].

Nas últimas décadas, diversos pesquisadores têm usado operações elementares da MM como regra de agragação nos neurônios, derivando uma nova classe de redes neurais artificiais (ANN - Artificial Neural Network), conhecidas na literatura como *redes neurais morfológicas* [53, 54, 55]. Confiando no fato que o intervalo $[0,1]$ e o hipercubo $[0, 1]^n$ constituem reticulados completos, a *morfologia matemática fuzzy* (FMM - Fuzzy Mathematical Morphological) possibilitou a criação das *redes neurais morfológicas fuzzy* (FMNNs - Fuzzy Morphological Neural Networks) cujos neurônios executam uma operação de erosão, dilatação, anti-erosão ou anti-dilatação de $[0, 1]^n$ em $[0,1]$. Uma FMNN que exerce uma função de memória associativa é denominada *memória associativa morfológica fuzzy* (FMAM - Fuzzy Morphological

Associative Memory) [73, 80]. As FMAMs pertencem à classe das *memórias associativas fuzzy* (FAMs - Fuzzy Associative Memories) que foram extensivamente estudadas desde a introdução dos modelos das FAMs de Kosko. Na verdade, as FMAMs abrangem diversos modelos de FAMs existentes na literatura, incluindo as *memórias associativas fuzzy implicativas* (IFAMs - Implicative Fuzzy Associative Memories) estabelecidas por Sussner e Valle, composta de neurônios morfológicos $\max\text{-}T$. A matriz dos pesos sinápticos é sintetizada através do *aprendizado R-implicativo* ou *aprendizado por adjunção*, onde a implicação fuzzy é adjunta a uma t-norma contínua que por sua vez representa uma dilatação em $[0,1]$. Neste contexto, também foram estabelecidos os modelos de *IFAMs duais*, compostas de neurônios morfológicos $\min\text{-}S$ que executam operações de erosão. Na verdade, existe uma relação de dualidade entre a IFAM e a IFAM dual com respeito à uma negação fuzzy [72].

Esta dissertação de mestrado está organizada da seguinte forma. Primeiro, serão exibidos os conceitos básicos de morfologia matemática, teoria dos conjuntos fuzzy e morfologia matemática fuzzy. Então, será dada uma introdução de redes neurais artificiais e memórias associativas. O Capítulo 3 apresenta os aspectos gerais de MAMs e FMAMs, dando ênfase as IFAMs. O Capítulo 4, relaciona a existência de soluções nos sistemas de equações relacionais fuzzy em termos da recordação perfeita da IFAM. O Capítulo 5, estabelece uma forma de escolher a melhor IFAM de Yager para uma dada aplicação. Finalmente, é apresentada uma relação no processo de compressão/descompressão de imagens à teoria das IFAMs. O Capítulo 8 conclui a dissertação.

CAPÍTULO 2

Morfologia Matemática, Conjuntos Fuzzy e Morfologia Matemática Fuzzy

2.1 Introdução a Morfologia Matemática

2.1.1 Introdução

Como já dito antes, a *Morfologia Matemática* (MM) foi desenvolvida por George Matheron e Jean Serra em 1964 para resolver problemas de análise de imagens, i.é, extrair informações de imagens a partir de transformações de formas através dos operadores elementares de dilatação e erosão que dependem de padrões predefinidos, denominados elementos estruturantes [42, 59, 61]. A dilatação e a erosão foram criadas a partir das noções de soma e subtração de Minkowski e Hadwiger [25, 44]. A MM é uma teoria com base nos conceitos de teoria de conjuntos, geometria e topologia para análise de estruturas geométricas em imagens [73]. Sua aplicação abrange segmentação e reconstrução de imagens [76], decomposição de sinais [8].

Originalmente, os operadores da MM desenvolvidos por Matheron e Serra foram voltados para imagens binárias. Posteriormente, Sternberg generalizou os principais operadores da morfologia matemática binária para o caso de tons de cinza empregando o noção de umbra [66]. Outra abordagem que trata o caso em tons de cinza está baseada na teoria dos conjuntos

fuzzy [20, 46, 73, 74]. As abordagens binárias, tons de cinza e fuzzy, possuem como base a teoria dos reticulados completos que representam o contexto mais geral onde a morfologia matemática pode ser conduzida [29].

2.1.2 Conceitos Fundamentais de Reticulados Completos

A teoria dos reticulados foi introduzida por Dedekind em 1987 [19] e seus conceitos estão difundidos em toda algebra moderna [5]. Essa subseção apresenta os conceitos necessários para o desenvolvimento da dissertação e uma descrição detalhada sobre a teoria dos reticulados pode ser encontrada em [5].

Um conjunto não vazio X equipado com uma ordem parcial \leq é chamado de conjunto parcialmente ordenado e denotado por (X, \leq) . Uma ordem parcial é uma relação binária \leq em X que satisfaz as seguintes propriedades para todo $x, y, z \in X$:

(O1) Reflexividade: $x \leq x, \forall x \in X$

(O2) Anti-Simetria: $x \leq y$ e $y \leq x$ então $x = y, \forall x, y \in X$

(O3) Transitividade: $x \leq y$ e $y \leq z$ então $x \leq z, \forall x, y, z \in X$

O conjunto (X, \leq) é *totalmente ordenado* se também satisfazer a propriedade:

(O4) $x \leq y$ ou $y \leq x, \forall x, y \in X$

Dado um conjunto X parcialmente ordenado (X, \leq) e um subconjunto $Y \subseteq X$. O elemento $l \in X$ é um *limite inferior* de Y se $l \leq y$ para todo $y \in Y$. O maior limite inferior $l_0 \in X$, quando existir, é chamado de *ínfimo* de Y , i.e, l_0 é um limite inferior de Y e $l \leq l_0$ para todos os limites inferiores l de Y . O ínfimo de Y é denotado por $\bigwedge Y$ ou equivalentemente por $\bigwedge_{j \in J} y_j$ quando $Y = \{y_j : j \in J\}$ para algum conjunto de índices J .

Similarmente, dado um conjunto parcialmente ordenado (X, \leq) e um subconjunto $Y \subseteq X$. O elemento $s \in X$ é um *limite superior* de Y se $y \leq s$ para todo $y \in Y$. O menor limite superior, quando existir, é chamado de *supremo* de Y . O supremo de Y é denotado por $\bigvee Y$ ou equivalentemente por $\bigvee_{j \in J} y_j$ quando $Y = \{y_j : j \in J\}$ para algum conjunto de índices J .

Definição 2.1.1. (Reticulado e Reticulado Completo [5]): *Um conjunto parcialmente ordenado X é um reticulado se todo subconjunto finito não vazio admite um supremo e um ínfimo em X , ou seja, para todo $Y \subseteq X$ finito tem-se $\bigwedge Y \in X$ e $\bigvee Y \in X$. Um reticulado X é completo se todo subconjunto, finito ou infinito, possui um ínfimo e um supremo em X .*

Exemplo 2.1.1. (Reticulados e Reticulados Completos):

1. *Todo conjunto totalmente ordenado é um reticulado. De fato, para todo subconjunto finito de elementos de um conjunto totalmente ordenado pode ser arranjado em ordem crescente, e portanto conter o menor e o maior elemento.*
2. *O intervalo fechado $[0,1]$ com a ordem parcial usual é um reticulado completo.*
3. *Nem todo reticulado é completo. De fato, tome o intervalo aberto $(0,1)$ com a ordem parcial usual é um conjunto totalmente ordenado, porém não é um reticulado completo, pois o conjunto $\{1/n : n \in \mathbb{N}\}$ não possui limite inferior.*

2.1.3 Operadores Básicos da Morfologia Matemática

Os operadores básicos da morfologia matemática são a erosão e dilatação, definidos a seguir:

Definição 2.1.2. (Erosão e Dilatação [29]): *Sejam \mathbb{L} e \mathbb{M} reticulados completos. O operador $\epsilon : \mathbb{M} \rightarrow \mathbb{L}$ que comuta com o operador de ínfimo é chamado erosão, ou seja, o operador ϵ satisfaz a seguinte equação para todo subconjunto $Y \subseteq \mathbb{M}$:*

$$\epsilon\left(\bigwedge Y\right) = \bigwedge_{y \in Y} \epsilon(y) \tag{2.1}$$

Similarmente, um operador $\delta : \mathbb{L} \rightarrow \mathbb{M}$ que comuta com o operador de supremo é chamado de dilatação, ou seja, o operador δ satisfaz a seguinte equação para todo subconjunto $X \subseteq \mathbb{L}$:

$$\delta(\bigvee X) = \bigvee_{x \in X} \delta(x) \quad (2.2)$$

Além dos operadores de dilatação e erosão, muitos pesquisadores também consideram os operadores de *anti-dilatação* e *anti-erosão* como operadores básicos da morfologia matemática.

Definição 2.1.3. (Anti-erosão e Anti-dilatação [3]): *Sejam \mathbb{L} e \mathbb{M} reticulados completos. O operador $\bar{\epsilon} : \mathbb{M} \rightarrow \mathbb{L}$ é chamado anti-erosão se satisfaz a seguinte equação para todo subconjunto $Y \subseteq \mathbb{M}$:*

$$\bar{\epsilon}(\bigwedge Y) = \bigvee_{y \in Y} \bar{\epsilon}(y) \quad (2.3)$$

Similarmente, um operador $\bar{\delta} : \mathbb{L} \rightarrow \mathbb{M}$ é chamado de anti-dilatação se satisfaz a seguinte equação para todo subconjunto $X \subseteq \mathbb{L}$:

$$\bar{\delta}(\bigvee X) = \bigwedge_{x \in X} \bar{\delta}(x) \quad (2.4)$$

Operadores de dilatação, erosão, anti-dilatação e anti-erosão são considerados operadores elementares da morfologia matemática, pois qualquer aplicação entre reticulados completos pode ser expressa em termos de combinações de supremos e ínfimos desses [3].

2.1.4 Relações de Dualidade - Adjunção e Negação

Os operadores de dilatação e erosão estão relacionados por meio de uma relação de dualidade. Muitos pesquisadores, tais como Bloch e Maître [7], Sinha e Dougherty [63], Nachtgael e Kerre [46] empregam o conceito de dualidade baseado no conceito de negação. Outros pesquisadores, como Maragos [41], Ronse [56], Deng e Heijmans [20] utilizam a dualidade baseada no conceito de adjunção.

Definição 2.1.4. (Adjunção [20]): *Sejam \mathbb{L} e \mathbb{M} reticulados completos e os operadores $\epsilon : \mathbb{M} \rightarrow \mathbb{L}$, $\delta : \mathbb{L} \rightarrow \mathbb{M}$. O par (ϵ, δ) representa uma adjunção em \mathbb{L} e \mathbb{M} se e somente se*

$$\delta(x) \leq y \Leftrightarrow x \leq \epsilon(y), \text{ para todo } x \in \mathbb{L} \text{ e } y \in \mathbb{M}. \quad (2.5)$$

A proposição a seguir mostra que o conceito de adjunção representa uma dualidade entre os operadores de erosão e dilatação.

Proposição 2.1.1. (Adjunção, Erosão e Dilatação [20]): *Sejam \mathbb{L} e \mathbb{M} reticulados completos e os operadores $\epsilon : \mathbb{M} \rightarrow \mathbb{L}$ e $\delta : \mathbb{L} \rightarrow \mathbb{M}$*

1. *Se (ϵ, δ) é uma adjunção, então ϵ é uma erosão e δ é uma dilatação.*
2. *Para toda dilatação δ , existe uma única erosão ϵ tal que (ϵ, δ) é uma adjunção. A erosão adjunta é dada por:*

$$\epsilon(y) = \bigvee \{x \in \mathbb{L} : \delta(x) \leq y\} \text{ para todo } y \in \mathbb{M} \quad (2.6)$$

3. *Para toda erosão ϵ , existe uma única dilatação δ tal que (ϵ, δ) é uma adjunção. A dilatação adjunta é dada por:*

$$\delta(x) = \bigwedge \{y \in \mathbb{M} : x \leq \epsilon(y)\} \text{ para todo } x \in \mathbb{L} \quad (2.7)$$

Definição 2.1.5. (Negação e Negação de um operador Ψ [20]): *Sejam \mathbb{L} e \mathbb{M} reticulados completos.*

- *Um operador $\nu : \mathbb{L} \rightarrow \mathbb{L}$ é denominado negação em \mathbb{L} se ν é uma bijeção involutiva ($\nu(\nu(x)) = x$) que reverte a ordem parcial de \mathbb{L} .*
- *Sejam $\Psi : \mathbb{L} \rightarrow \mathbb{M}$, e ν_L e ν_M negações em \mathbb{L} e \mathbb{M} , respectivamente. Considere as negações de \mathbb{L} e \mathbb{M} definidas. O operador $\Psi^\nu : \mathbb{L} \rightarrow \mathbb{M}$ é chamado de negação de Ψ com respeito à ν_L e ν_M e dado por:*

$$\Psi^\nu(x) = \nu_M(\psi(\nu_L(x))), \text{ para todo } x \in \mathbb{L}. \quad (2.8)$$

Exemplo 2.1.2. (Negação de Dilatação e Erosão): *Sejam \mathbb{L} e \mathbb{M} reticulados completos com negações ν_L e ν_M , respectivamente. O operador $\delta : \mathbb{L} \rightarrow \mathbb{M}$ é uma dilatação se e somente se $\epsilon = \delta^\nu : \mathbb{L} \rightarrow \mathbb{M}$ é uma erosão. Em outras palavras, a negação de uma dilatação é uma erosão e vice-versa.*

2.2 Conceitos Relevantes da Teoria dos Conjuntos

Fuzzy

2.2.1 Introdução

A teoria dos conjuntos fuzzy foi formada por Lotfi Zadeh em 1965 como uma ferramenta para modelar imprecisões [4, 82]. A palavra *fuzzy*, de origem inglesa, aparece na matemática brasileira como *nebuloso*. Um conjunto fuzzy em um universo \mathbf{X} é definido por uma função φ de um conjunto \mathbf{X} para o intervalo $[0,1]$. A função φ é denominada *função de pertinência* e o valor $\varphi(\mathbf{x})$ representa o *grau de pertinência* de \mathbf{x} no conjunto fuzzy φ .

Um conjunto clássico \mathbf{A} pode ser visto como um caso particular de um conjunto fuzzy. De fato, se \mathbf{X} é um conjunto clássico e \mathbf{A} um subconjunto de \mathbf{X} , então \mathbf{A} pode ser caracterizado por meio de sua *função característica*, definida segundo:

$$C_A(\mathbf{x}) = \begin{cases} 1 & \text{se } \mathbf{x} \in \mathbf{A} \\ 0 & \text{se } \mathbf{x} \notin \mathbf{A} \end{cases} \quad (2.9)$$

Desta forma, a definição de conjuntos fuzzy foi obtida ampliando o contra-domínio da função característica, isto é, do conjunto $\{0,1\}$ para o intervalo $[0,1]$. Neste momento, o símbolo $\mathcal{F}(\mathbf{X}) = [0,1]^{\mathbf{X}}$ denotará a classe dos conjuntos fuzzy em \mathbf{X} . A teoria dos conjuntos fuzzy pode ser usada para o desenvolvimento de operadores de imagens, uma vez que uma imagem $\mathbf{a} : \mathbf{X} \rightarrow [0,1]$ pode ser interpretada como um conjunto fuzzy em \mathbf{X} . Em particular, uma imagem fuzzy $\mathbf{a} \in \mathcal{F}(\mathbf{X})$ pode ser representada por um vetor sempre que $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ for um conjunto finito com n elementos. Nesse caso, tem-se

$\mathbf{a} = [a_1, \dots, a_n]^T \in [0, 1]$, onde $a_i = \mathbf{a}(x_i)$, para $i = 1, \dots, n$.

2.2.2 Conectivos Básicos da Lógica Fuzzy

A lógica fuzzy permite representar valores lógicos intermediários entre Verdadeiro e Falso, possibilitando o tratamento de atributos imprecisos e fundamenta-se na teoria dos conjuntos fuzzy [82]. Esta subseção apresenta as definições de conjunção, disjunção, implicação e negação fuzzy que serão empregados posteriormente no desenvolvimento das memórias associativas fuzzy.

Definição 2.2.1. (Conjunção Fuzzy e t-norma [33, 50]): *Uma conjunção fuzzy é uma função crescente $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$ que satisfaz $C(0, 0) = C(1, 0) = C(0, 1) = 0$ e $C(1, 1) = 1$. Em particular, uma conjunção fuzzy $T : [0, 1] \times [0, 1] \rightarrow [0, 1]$, comutativa $T(x, y) = T(y, x)$ e associativa $T(x, T(y, z)) = T(T(x, y), z)$ que satisfaz $T(x, 1) = x$ para todo $x, y, z \in [0, 1]$ é chamada de norma triangular ou simplesmente t-norma.*

Exemplo 2.2.1. *Exemplos de conjunções fuzzy são o mínimo, produto e de Lukasiewicz. Em particular, essas conjunções fuzzy também são exemplos de t-normas e são descritas, respectivamente abaixo:*

$$C_M(x, y) = x \wedge y \quad (2.10)$$

$$C_P(x, y) = xy \quad (2.11)$$

$$C_L(x, y) = 0 \vee (x + y - 1) \quad (2.12)$$

Definição 2.2.2. (Disjunção Fuzzy e s-norma [33, 50]): *Uma disjunção fuzzy é uma função crescente $D : [0, 1] \times [0, 1] \rightarrow [0, 1]$ que satisfaz $D(1, 0) = D(1, 1) = D(0, 1) = 1$ e $D(0, 0) = 0$. Em particular, uma disjunção fuzzy $S : [0, 1] \times [0, 1] \rightarrow [0, 1]$, comutativa $S(x, y) = S(y, x)$ e associativa $S(x, S(y, z)) = S(S(x, y), z)$ que satisfaz $S(x, 0) = x$ para todo $x, y, z \in [0, 1]$ é chamada de co-norma triangular ou simplesmente s-norma.*

Exemplo 2.2.2. *Exemplos de disjunções fuzzy são o máximo, soma probabilística e de Lukasiewicz. Em particular, essas disjunções fuzzy também são exemplos de s-normas e são descritas, respectivamente abaixo:*

$$D_M(x, y) = x \vee y \quad (2.13)$$

$$D_P(x, y) = x + y - xy \quad (2.14)$$

$$D_L(x, y) = 1 \wedge (x + y) \quad (2.15)$$

Definição 2.2.3. (Implicação Fuzzy e Implicação Fuzzy Reversa [33, 50, 79]): *Uma implicação fuzzy é uma função $I : [0, 1] \times [0, 1] \rightarrow [0, 1]$ decrescente no primeiro argumento e crescente no segundo argumento que estende a implicação crisp usual em $\{1, 0\} \times \{1, 0\}$, i.e., $I(0, 0) = I(1, 1) = I(0, 1) = 1$ e $I(1, 0) = 0$. A implicação fuzzy reversa $J : [0, 1] \times [0, 1] \rightarrow [0, 1]$ é obtida invertendo os argumentos da implicação fuzzy I , i.e., $J(x, y) = I(y, x)$ para todo $x, y \in [0, 1]$.*

Exemplo 2.2.3. *Exemplos de implicações fuzzy são de Gödel, de Lukasiewicz e de Keene-Dienes.*

$$I_G(x, y) = \begin{cases} 1 & , x \leq y \\ y & , x > y \end{cases} \quad (2.16)$$

$$I_L(x, y) = 1 \wedge (y - x + 1) \quad (2.17)$$

$$I_K(x, y) = (1 - x) \vee y \quad (2.18)$$

Definição 2.2.4. (Negação Fuzzy [33, 50]): *Uma negação fuzzy é uma bijeção involutiva $N : [0, 1] \rightarrow [0, 1]$ decrescente.*

Exemplo 2.2.4. *Exemplo de negação fuzzy é a negação fuzzy usual N_S . Outras negações fuzzy são apresentadas abaixo:*

$$N_S(x) = 1 - x \quad (2.19)$$

$$N_D(x) = \frac{1 - x}{1 + px}, \text{ para } p \in (-1, \infty) \quad (2.20)$$

$$N_R(x) = \sqrt[p]{1 - x^p}, \text{ para } p \in (0, \infty) \quad (2.21)$$

2.2.3 Relações de Dualidade entre Operadores Fuzzy

Uma conjunção fuzzy C e uma implicação fuzzy I formam uma *adjunção* se, e somente se, $C(x, \cdot)$ e $I(x, \cdot)$ formam uma adjunção para todo $x \in [0, 1]$. Se C e I formam uma adjunção então pela Proposição 2.1.1, tem-se que $I(x, \cdot)$ é uma erosão e $C(x, \cdot)$ é uma dilatação em $[0, 1]$ para todo $x \in [0, 1]$. E mais, se o par (I, C) forma uma adjunção então:

$$I(x, y) = \bigvee \{z \in [0, 1] : C(x, z) \leq y\} \text{ para todo } x, y \in [0, 1] \quad (2.22)$$

A implicação I que satisfaz a Equação (2.22) é denominada *R-implicação associada a conjunção C* [50]. Denota-se por I_T a R-implicação associada a t-norma contínua T . Como um operador contínuo representa uma dilatação, a Equação 2.22 está bem definida [29]. Por exemplo, os pares (I_L, C_L) e (I_G, C_M) representam uma adjunção [73].

Ressalta-se que existem implicações fuzzy que não representam erosões e existem conjunções fuzzy que não representam dilatações [73]. Considere a conjunção fuzzy C_{RM} no Exemplo 2.2.5.

Exemplo 2.2.5. *Seja a conjunção fuzzy C_{RM} definida pela Equação 2.23:*

$$C_{RM}(x, y) = \begin{cases} 1 & , x = 1 \quad e \quad y = 1 \\ 0 & , \text{ caso contrário} \end{cases} \quad (2.23)$$

Note que $C_{RM}(1, \cdot)$ não representa uma dilatação, pois não satisfaz a Equação 2.2, ou seja, não comuta com o operador supremo:

$$C_{RM} \left(1, \bigvee_{n \in \mathbb{N}} \left(1 - \frac{1}{n} \right) \right) = C_{RM}(1, 1) = 1 \quad (2.24)$$

e

$$\bigvee_{n \in \mathbb{N}} C_{RM} \left(1, 1 - \frac{1}{n} \right) = \bigvee_{n \in \mathbb{N}} 0 = 0 \quad (2.25)$$

Uma conjunção fuzzy pode estar associada a uma implicação fuzzy por meio de uma relação de dualidade com respeito à uma negação fuzzy [74]. Obtém-se uma implicação (conjunção) fuzzy a partir de uma conjunção (implicação) fuzzy como se segue:

$$I(x, y) = N(C(x, N(y))) \quad \text{ou} \quad C(x, y) = N(I(x, N(y))), \quad \forall x, y \in [0, 1]. \quad (2.26)$$

A *implicação fuzzy dual* é um operador $\bar{I} : [0, 1] \times [0, 1] \rightarrow [0, 1]$ derivado de uma implicação fuzzy I e de uma negação fuzzy N [72], segundo a Equação 2.27:

$$\bar{I}(x, y) = N(I(N(x), N(y))) \quad \forall x, y \in [0, 1]. \quad (2.27)$$

Apesar do nome, a implicação fuzzy dual \bar{I} não representa uma implicação fuzzy [72], segundo a Definição 2.2.3. Se os operadores $D(x, \cdot)$ e $\bar{I}(x, \cdot)$ formam uma adjunção para todo $x \in [0, 1]$, então $\bar{I}(x, \cdot)$ representa uma dilatação e $D(x, \cdot)$ uma erosão para todo $x \in [0, 1]$ (Proposição 2.1.1). E mais, a seguinte equação é válida, para todo $x, y \in [0, 1]$.

$$\bar{I}(x, y) = \bigwedge \{z \in [0, 1] : D(x, z) \geq y\} \quad (2.28)$$

A implicação \bar{I} que satisfaz a Equação (2.28) é denominada *R-implicação dual associada a disjunção D* [72]. Denota-se por \bar{I}_S a R-implicação dual associada a s-norma contínua S .

Uma conjunção fuzzy C e uma disjunção fuzzy D são duais com respeito à uma negação fuzzy N se e somente se $D(x, \cdot)$ for a negação de $C(N(x), \cdot)$ para todo $x \in [0, 1]$, ou seja, a Equação (2.29) é satisfeita para todo $x, y \in [0, 1]$ [73]:

$$C(x, y) = N(D(N(x), N(y))) \quad \text{ou} \quad D(x, y) = N(C(N(x), N(y))) \quad (2.29)$$

Caso $C(x, \cdot)$ realiza uma dilatação no segundo argumento para todo $x \in [0, 1]$, então a disjunção $D(x, \cdot)$ obtida a partir da Equação 2.29 realiza uma erosão no segundo argumento para todo $x \in [0, 1]$ e vice-versa [73].

2.2.4 Produtos Matriciais Fuzzy

Diversos modelos de memórias associativas morfológicas fuzzy são descritos a partir de produtos matriciais fuzzy, combinando o máximo e mínimo com os operadores fuzzy C , D e I [79]. Portanto, serão introduzidos três tipos de produtos matriciais fuzzy que serão extensamente utilizados nos próximos capítulos.

Considere os conjuntos $M = \{1, 2, \dots, m\}$, $N = \{1, 2, \dots, n\}$, $K = \{1, 2, \dots, k\}$ finitos. Sejam as matrizes A com elementos $a_{i\xi} \in [0, 1]$ e B com elementos $b_{\xi j} \in [0, 1]$, definidas sobre $M \times K$ e $K \times N$, respectivamente, para $1 \leq i \leq m$, $1 \leq j \leq n$ e $1 \leq \xi \leq k$. Por abuso de notação, as matrizes A e B serão denotadas por $A \in [0, 1]^{m \times k}$ e $B \in [0, 1]^{k \times n}$.

Define-se o *produto max-C* de $A \in [0, 1]^{m \times k}$ e $B \in [0, 1]^{k \times n}$, denotado por $E = A \circ B$, como se segue [71]:

$$e_{ij} = \bigvee_{\xi=1}^k C(a_{i\xi}, b_{\xi j}) \quad \forall i = 1, 2, \dots, m \quad \text{e} \quad j = 1, 2, \dots, n \quad (2.30)$$

Similarmente, define-se o *produto min-D* e o *produto min-J*, denotados por $G = A \bullet B$ e $H = A \otimes B$, respectivamente, para todo $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$ [71]:

$$g_{ij} = \bigwedge_{\xi=1}^k D(a_{i\xi}, b_{\xi j}) \quad \text{e} \quad h_{ij} = \bigwedge_{\xi=1}^k J(b_{\xi j}, a_{i\xi}) \quad (2.31)$$

Subscritos nos símbolos \circ , \bullet e \otimes indicam que tipo de operadores fuzzy usados nas Equações 2.30 e 2.31. Por exemplo, a matriz $E = A \circ_M B$ é dada por:

$$e_{ij} = \bigvee_{\xi=1}^k C_M(a_{i\xi}, b_{\xi j}) = \bigvee_{\xi=1}^k (a_{i\xi} \wedge b_{\xi j}) \quad (2.32)$$

2.2.5 Equações Relacionais Fuzzy

As *equações relacionais fuzzy* (FREs - Fuzzy Relational Equations) foram introduzidas por Sanchez [58] e muitos outros pesquisadores continuam a estudá-la, como: Pedrycz, Di Nola, Sessa e Perfilieva entre outros [21, 22, 51, 52]. Geralmente, uma FRE relaciona um conjunto fuzzy $\mathbf{x} : \mathbf{X} \rightarrow [0, 1]$ a um conjunto fuzzy $\mathbf{y} : \mathbf{Y} \rightarrow [0, 1]$, onde \mathbf{X} e \mathbf{Y} são conjuntos finitos. Por simplicidade, assume-se que $\mathbf{x} \in [0, 1]^n$ e $\mathbf{y} \in [0, 1]^m$.

Uma *relação fuzzy* indica o grau da relação entre dois objetos associados e por isso sua formalização matemática é dada pela teoria dos conjuntos fuzzy [4]. Uma relação fuzzy \mathcal{R} é definida por uma função de pertinência $\varphi : X_1 \times X_2 \times \dots \times X_n \rightarrow [0, 1]$. Nesta dissertação, serão apenas consideradas as relações fuzzy binárias, i.e, $\varphi : \mathbf{X} \times \mathbf{Y} \rightarrow [0, 1]$. Geralmente, representa-se uma relação fuzzy binária, quando X e Y são finitos, pela forma matricial. Em outras palavras, sejam os conjuntos $X = \{x_1, \dots, x_n\}$ e $Y = \{y_1, \dots, y_m\}$ e a relação fuzzy \mathcal{R} sobre $X \times Y$, com função de pertinência dada por $\varphi(x_j, y_i) = r_{ji}$ para $j = 1, \dots, n$ e $i = 1, \dots, m$. Por simplicidade de notação e para destacar a relação entre as FREs e as memórias associativas fuzzy implicativas (IFAMs - Implicative Fuzzy Associative Memories), assume-se que $\mathcal{R} \in [0, 1]^{n \times m}$.

As FREs tratam de encontrar a forma matricial de uma relação fuzzy binária $\mathcal{R} \in [0, 1]^{n \times m}$. Considere os conjuntos fuzzy $\mathbf{x} \in [0, 1]^n$ e $\mathbf{y} \in [0, 1]^m$. As equações relacionais fuzzy de interesse tem a forma:

$$\mathbf{y} = R \cdot \mathbf{x} \quad \mathbf{y} = R\Delta\mathbf{x} \quad \mathbf{y} = R\phi\mathbf{x} \quad (2.33)$$

As composições ”.”, ” Δ ” e ” ϕ ” que ocorre em 2.33 são definidas por 2.34, 2.35 e 2.36, respectivamente.

$$y_i = (R \cdot \mathbf{x})(i) = \bigvee_{j=1}^n T(x_j, r_{ji}) \text{ para } i = 1, 2, \dots, m. \quad (2.34)$$

$$y_i = (R\Delta\mathbf{x})(i) = \bigwedge_{j=1}^n S(x_j, r_{ji}) \text{ para } i = 1, 2, \dots, m. \quad (2.35)$$

$$y_i = (R\phi\mathbf{x})(i) = \bigwedge_{j=1}^n J(x_j, r_{ji}) \text{ para } i = 1, 2, \dots, m. \quad (2.36)$$

onde T representa uma t-norma, S uma s-norma e J uma implicação fuzzy reversa.

Exemplo 2.2.6. Dadas as matrizes $X \in [0, 1]^{4 \times 3}$ e $Y \in [0, 1]^{3 \times 3}$:

$$X = \begin{bmatrix} 0.2 & 0.4 & 0.5 \\ 0.0 & 0.3 & 0.9 \\ 0.8 & 0.0 & 0.2 \\ 1.0 & 0.7 & 0.0 \end{bmatrix} \quad e \quad Y = \begin{bmatrix} 0.7 & 0.6 & 0.8 \\ 0.3 & 0.4 & 0.9 \\ 1.0 & 0.7 & 0.2 \end{bmatrix} \quad (2.37)$$

Uma solução para o sistema de FREs do tipo $\mathbf{y}^\xi = R \cdot \mathbf{x}^\xi$, onde \mathbf{x}^ξ e \mathbf{y}^ξ são as colunas de X e Y , respectivamente, para $\xi = 1, 2, 3$ é dada pela relação fuzzy \mathcal{R}_0 , tomando a t-norma do mínimo. Note que esta solução não é única, pois a relação fuzzy \mathcal{R}_1 também é solução, e mais, \mathcal{R}_1 é a maior solução.

$$\mathcal{R}_0 = \begin{bmatrix} 0.0 & 0.4 & 0.0 \\ 0.8 & 0.9 & 0.2 \\ 0.7 & 0.3 & 0.0 \\ 0.6 & 0.0 & 1.0 \end{bmatrix} \quad \mathcal{R}_1 = \begin{bmatrix} 1.0 & 1.0 & 0.2 \\ 0.8 & 1.0 & 0.2 \\ 0.7 & 0.3 & 1.0 \\ 0.6 & 0.3 & 1.0 \end{bmatrix} \quad (2.38)$$

Condições que garantem a existência de soluções são complexas de deduzir [21, 22, 51, 52]. Caso não exista solução, uma alternativa é obter soluções aproximadas para o sistema de FREs através da omissão de um subconjunto de equações ou por meio de modificações dos conjuntos fuzzy \mathbf{x}^ξ e \mathbf{y}^ξ para estabelecer conexões entre eles e a relação fuzzy [21]. O objetivo desta dissertação é estudar a existência de soluções para os sistemas de FREs em termos das memórias associativas fuzzy implicativas por meio da recordação perfeita desses modelos de memória associativas [72].

2.3 Morfologia Matemática Fuzzy

2.3.1 Medidas de Inclusão e Intersecção Fuzzy

A *morfologia matemática fuzzy* (FMM - Fuzzy Mathematical Morphology) é uma extensão da morfologia binária para o caso em tons de cinza com base na teoria dos conjuntos fuzzy e portanto, consiste de todos os operadores morfológicos cujos domínio e contra-domínio são classes de conjuntos fuzzy. Uma abordagem da FMM está baseada nos conceitos de medida de inclusão e intersecção fuzzy [73].

Uma *medida de inclusão fuzzy* é uma função $Inc : \mathcal{F}(\mathbf{X}) \times \mathcal{F}(\mathbf{X}) \rightarrow [0, 1]$ cuja restrição a $\mathcal{P}(\mathbf{X}) \times \mathcal{P}(\mathbf{X})$ coincide com a inclusão para conjuntos clássicos. Se $\mathbf{a}, \mathbf{b} \in \mathcal{F}(\mathbf{X})$ são conjuntos fuzzy então $Inc(\mathbf{a}, \mathbf{b})$ é interpretado como o grau de inclusão do conjunto fuzzy \mathbf{a} no conjunto fuzzy \mathbf{b} [73, 74]. Formalmente, tem-se a seguinte equação para todo conjunto $\mathbf{A}, \mathbf{B} \in \mathcal{P}(\mathbf{X})$ e os conjuntos fuzzy $\mathbf{a}, \mathbf{b} \in \mathcal{F}(\mathbf{X})$ tais que $\mathbf{A}(\mathbf{x}) = \mathbf{a}(\mathbf{x})$ e $\mathbf{B}(\mathbf{x}) = \mathbf{b}(\mathbf{x})$, $\forall \mathbf{x} \in \mathbf{X}$

$$Inc(\mathbf{a}, \mathbf{b}) = Inc_c(\mathbf{A}, \mathbf{B}) = \begin{cases} 1, & \mathbf{A} \subseteq \mathbf{B} \\ 0, & \mathbf{A} \not\subseteq \mathbf{B} \end{cases} \quad (2.39)$$

Pode-se obter diretamente a medida de inclusão fuzzy da inclusão de conjuntos clássicos da seguinte forma. Sejam $\mathbf{A}, \mathbf{B} \subseteq \mathbf{X}$ conjuntos clássicos, sabe-se que $\mathbf{A} \subseteq \mathbf{B}$ se e somente se “ $\mathbf{x} \in \mathbf{A}$ implica que $\mathbf{x} \in \mathbf{B}$ para todo $\mathbf{x} \in \mathbf{X}$ ”, ou seja,

$$Inc_c(\mathbf{A}, \mathbf{B}) = \bigwedge_{\mathbf{x} \in \mathbf{X}} I_c(\mathcal{C}_A(\mathbf{x}), \mathcal{C}_B(\mathbf{x})) \quad (2.40)$$

onde $I_c : \{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}$ representa uma implicação clássica e $\mathcal{C}_A, \mathcal{C}_B : \mathbf{X} \rightarrow \{0, 1\}$ são as funções características definidas pela Equação 2.9. Agora, considerando os conjuntos fuzzy $\mathbf{a}, \mathbf{b} \in \mathcal{F}(\mathbf{X})$, a direta fuzzificação da Equação 2.40, define a *medida de inclusão fuzzy Inf-I* ou *medida de inclusão de Bandler-Kohout* [2, 73] dada por:

$$Inc(\mathbf{a}, \mathbf{b}) = \bigwedge_{\mathbf{x} \in \mathbf{X}} I(\mathbf{a}(\mathbf{x}), \mathbf{b}(\mathbf{x})) \quad (2.41)$$

Similarmente, obtem-se a *medida de intersecção fuzzy Sup-C*, definida por $Sec : \mathcal{F}(\mathbf{X}) \times \mathcal{F}(\mathbf{X}) \rightarrow [0, 1]$ através da Equação 2.42 cuja restrição a $\mathcal{P}(\mathbf{X}) \times \mathcal{P}(\mathbf{X})$ coincide com a intersecção para conjuntos clássicos.

$$Sec(\mathbf{a}, \mathbf{b}) = \bigvee_{\mathbf{x} \in \mathbf{X}} C(\mathbf{a}(\mathbf{x}), \mathbf{b}(\mathbf{x})) \quad (2.42)$$

Particulares escolhas de implicações e conjunções fuzzy produzem particulares medidas de inclusões e intersecções fuzzy, respectivamente. Dada uma particular medida de inclusão fuzzy Inf-I ou uma intersecção Sup-C, indica-se a implicação fuzzy ou a conjunção fuzzy utilizada por meio de um sub-índice. Por exemplo, Inc_L denota a inclusão fuzzy Inf-I baseada na implicação de Lukasiewicz.

2.3.2 Erosão Fuzzy e Dilatação Fuzzy

Sussner e Valle observaram que as principais abordagens da morfologia matemática fuzzy podem ser derivadas da morfologia matemática binária através dos conceitos de medidas de inclusão e intersecção fuzzy [73]. Várias outras propostas que envolvem a morfologia matemática e os conceitos da teoria dos conjuntos fuzzy são encontrados nos trabalhos de Sinha e Dougherty [62, 63], Bloch e Maître [7], De Baets [18], Deng e Heijmans [20] e Maragos [41].

Seja Inc a medida de inclusão fuzzy Inf-I tal que $Inc(s, \cdot)$ comuta com o operador ínfimo para todo $s \in \mathcal{F}(\mathbf{X})$. Dado uma imagem fuzzy $\mathbf{a} \in \mathcal{F}(\mathbf{X})$ e um elemento estruturante fuzzy $s \in \mathcal{F}(\mathbf{X})$, define-se a *erosão fuzzy* \mathcal{E} de \mathbf{a} por s referente à Inc por:

$$\mathcal{E}(\mathbf{a}, s)(\mathbf{x}) = Inc(\mathbf{s}_x, \mathbf{a}) = \bigwedge_{\mathbf{y} \in \mathbf{X}} I(\mathbf{s}_x(\mathbf{y}), \mathbf{a}(\mathbf{y})) \quad (2.43)$$

onde $\mathbf{s}_x(\mathbf{y}) = s(\mathbf{y} - \mathbf{x}), \forall \mathbf{y} \in \mathbf{X}$ é a translação de s por $\mathbf{x} \in \mathbf{X}$. A última afirmação é certamente verdadeira, ou seja, $\mathcal{E}(\cdot, s) : \mathcal{F}(\mathbf{X}) \rightarrow \mathcal{F}(\mathbf{X})$ é uma erosão para todo $s \in \mathcal{F}(\mathbf{X})$, se para todo $s \in [0, 1]$ a implicação fuzzy $I(s, \cdot)$ comuta com o ínfimo. Em outras palavras $I(s, \cdot)$ é uma erosão em $[0, 1]$. Por exemplo, as implicações I_G, I_L e I_K comutam com o ínfimo no segundo argumento e portanto, $\mathcal{E}_G, \mathcal{E}_L$ e \mathcal{E}_K representam erosões fuzzy [74].

Similarmente, uma dilatação fuzzy $\mathcal{D} : \mathcal{F}(\mathbf{X}) \times \mathcal{F}(\mathbf{X}) \rightarrow \mathcal{F}(\mathbf{X})$ baseada em uma medida de intersecção fuzzy Sup-C tal que $Sec(s, \cdot)$ comuta com o supremo para todo $s \in \mathcal{F}(\mathbf{X})$ é definida por:

$$\mathcal{D}(\mathbf{a}, s)(\mathbf{x}) = Sec(\bar{\mathbf{s}}_x, \mathbf{a}) = \bigvee_{\mathbf{y} \in \mathbf{X}} C(\bar{\mathbf{s}}_x(\mathbf{y}), \mathbf{a}(\mathbf{y})) \quad (2.44)$$

onde $\bar{\mathbf{s}}_x = s(\mathbf{x} - \mathbf{y})$ é a translação da reflexão para todo $\mathbf{x}, \mathbf{y} \in X$. Se a conjunção fuzzy $C(s, \cdot)$ é uma dilatação em $[0, 1]$ para todo $s \in [0, 1]$, então o operador $\mathcal{D}(\cdot, s) : \mathcal{F}(\mathbf{X}) \rightarrow \mathcal{F}(\mathbf{X})$ é uma dilatação para todo $s \in \mathcal{F}(\mathbf{X})$ [74].

CAPÍTULO 3

Redes Neurais Artificiais e Memórias Associativas

3.1 Redes Neurais Artificiais

3.1.1 Contexto Histórico

O cérebro humano é uma fonte de motivação para o desenvolvimento de máquinas com a capacidade de aprender e continuar a adaptar-se, realizando tarefas gradativamente mais eficientes. Inspirado nas características biológicas, acredita-se que surgirá em um futuro próximo, uma geração de novos sistemas computacionais, muito mais eficientes e inteligentes que os sistemas atuais.

Os primeiros estudos das redes neurais artificiais deram-se por Waren McCulloch e Walter Pitts em 1943 com a apresentação de um modelo matemático de um neurônio biológico [43]. Em 1949, o neurofisiologista Donald Hebb apresentou a primeira formulação de uma regra de aprendizagem, observando que as conexões sinápticas do cérebro são continuamente modificadas conforme um organismo aprende novas tarefas, criando, assim agrupamentos neurais. O postulado de aprendizagem (ou regra) de Hebb é descrito como: “A eficiência de uma sinapse é aumentada pela iteração entre dois neurônios através da sinapse” [28].

O primeiro neuro computador com capacidade de aprendizado (ajuste automático dos pesos entre as sinapses) foi desenvolvido por Minsky em 1951, que apesar de não executar qualquer função de processamento de informação interessante, serviu de inspiração para outros pesquisadores, como Frank Rosenblatt cujo interesse inicial da criação do *perceptron* era o reconhecimento de padrões [57]. Werbos, em 1974 lançou bases para o algoritmo de retropropagação (backpropagation) que posteriormente foi estudado por Rumelhart, Hinton e Williams em 1986 e emergiu como um algoritmo de aprendizado para o treinamento de perceptrons de multiplas camadas.

3.1.2 Introdução

As redes neurais artificiais (ANNs - Artificial Neural Networks) são descritas como modelos matemáticos que assemelham-se as estruturas neurais biológicas e que têm capacidade computacional por meio de aprendizado e generalização. Essa semelhança é dada por dois aspectos:

1. O conhecimento é adquirido pela rede a partir de sua iteração com o ambiente através de um processo de aprendizagem;
2. Conexões entre os neurônios, conhecidos como pesos sinápticos são utilizados para armazenar o conhecimento adquirido.

Uma rede neural artificial é caracterizada pelo *modelo do neurônio*, *topologia* ou *arquitetura da rede* e a *regra de treinamento* ou *regra de aprendizado*. O processamento da informação numa ANN é realizado por meio de estruturas neurais artificiais cujo o armazenamento e o processamento são computados de maneira paralela e distribuída por unidades processadoras, denominadas *neurônios* ou *nós*. A representação gráfica de um neurônio artificial é apresentada na Figura 3.1, consistindo dos seguintes elementos [27]:

- Um conjunto de *sinais de entrada* e um conjunto de *pesos sinápticos* ou *conexões sinápticas*;

- Uma *regra de agregação* que combina os sinais de entrada, ponderados pelos respectivos pesos sináptios;
- Uma *função de ativação* que restringe o sinal de saída em um dado intervalo ou de introduzir não-linearidade no modelo.

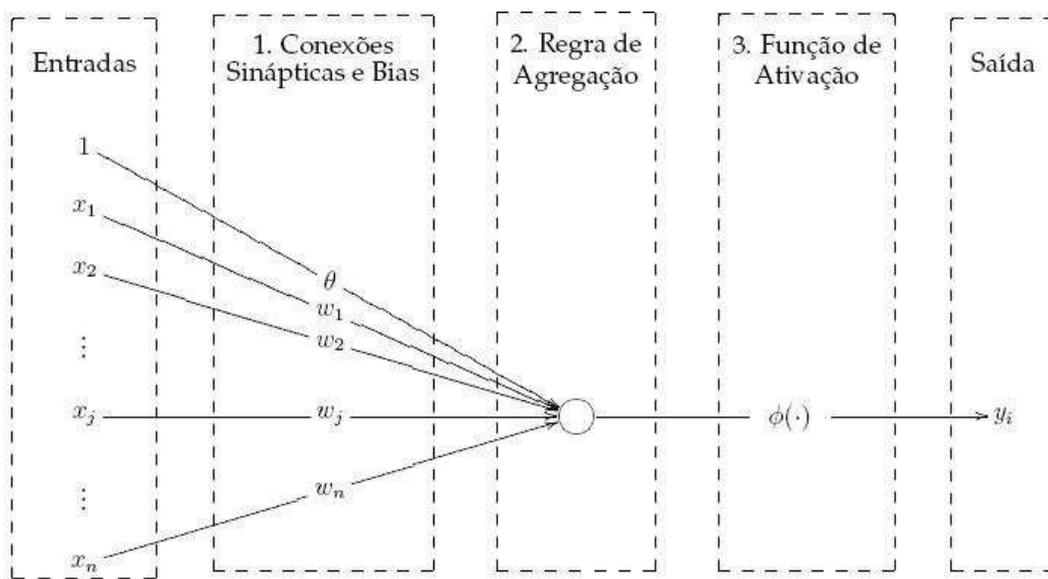


Figura 3.1: Modelo geral de um neurônio

Alguns modelos de neurônios apresentam um parâmetro externo, denominado *bias* e na Figura 3.1 designado pelo símbolo θ . Note que, o bias é introduzido como um peso sináptico conectado a uma entrada constante.

3.1.3 Topologia de uma Rede Neural Artificial

A topologia da ANN refere-se à sua arquitetura, i.e, a forma de interconexão dos neurônios que por sua vez estão organizados em camadas. As camadas são classificadas por suas funções ou localizações na rede neural [27]. A representação gráfica de uma ANN é mostrada na Figura 3.2 [67].

- *Camadas de Entrada:* encontram-se os neurônios que recebem os sinais externos;
- *Camadas Ocultas:* os neurônios desta camada recebem os sinais das unidades anteriores e os associa aos pesos das conexões por onde os sinais são repassados;
- *Camadas de Saída:* representam o último passo de processamento e seu resultado é apresentado através de um único terminal de saída.

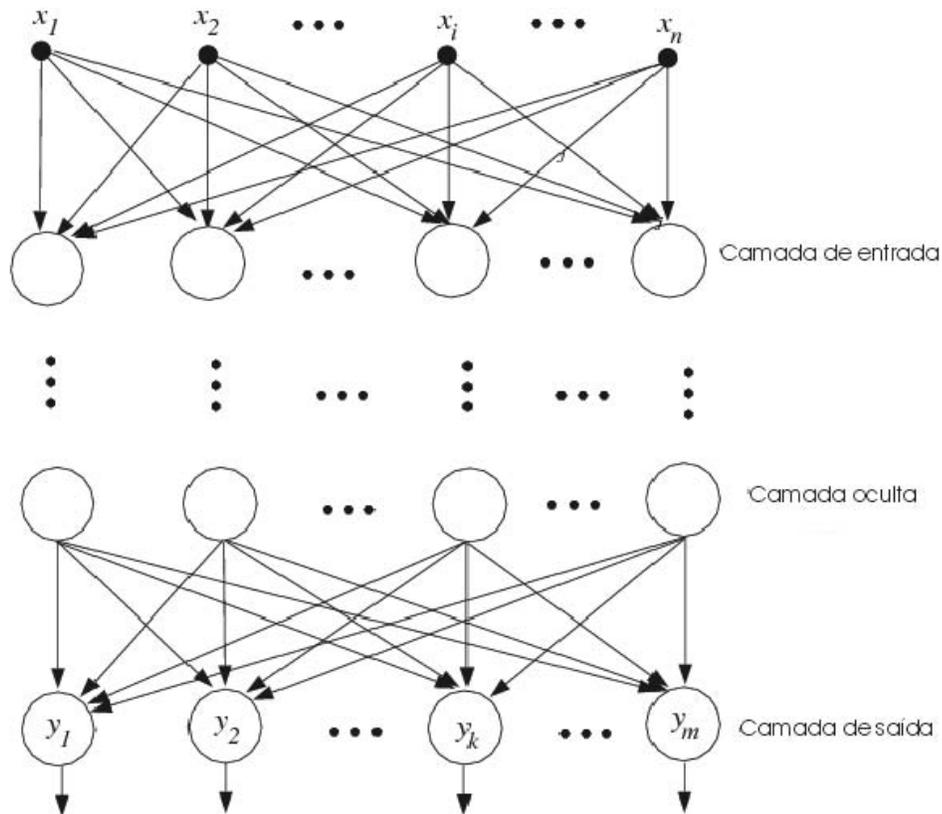


Figura 3.2: Rede neural artificial estática

De acordo com o esquema de interconexão podemos classificar as redes neurais como *estáticas* ou *recorrentes*:

- *Estática*: as conexões apontam para o mesmo sentido, i.e, das unidades de entrada em direção as unidades de saída.
- *Recorrente*: quando o sinal de saída de um neurônio serve como entrada para a mesma camada ou para alguma camada anterior, servindo como uma retroalimentação.

3.1.4 Aprendizagem

O aprendizado em redes neurais artificiais está associada à sua capacidade de as mesmas adaptarem os pesos de conexão como consequência da sua iteração do meio externo e, assim, melhorar seu desempenho gradativamente à medida que interage com as informações fornecidas [27]. As formas de aprendizado são:

- *Aprendizado supervisionado*: os algoritmos de aprendizagem supervisionado utilizam de informações de um conjunto de entradas e um correspondente conjunto de saídas desejadas. Durante o processo de aprendizagem o ajuste dos pesos assume o papel de minimizar a diferença (erro) entre os dados de saída desejados e a respectiva saída real aos dados de entrada.
- *Aprendizado não-supervisionado*: As informações sobre o ambiente são passadas a rede neural sem qualquer informação de classificação. Uma vez adaptada as regularidades estatísticas dos dados de entrada, a rede neural desenvolve a habilidade de formar representações internas para codificação das entradas e então criar novas classes automaticamente. Em outras palavras, o aprendizado está baseado em agrupamentos de padrões, onde os pesos são ajustados de modo que padrões semelhantes produzam a mesma saída.
- *Aprendizado por reforço*: O treinamento por reforço ocorre através de adaptação de pesos de acordo com um sinal de reforço. Este sinal, proveniente de um “crítico” externo, é responsável por sinalizar se uma resposta produzida pela rede neural é ou não satisfatória. Se a resposta produzida pela rede neural for satisfatória, um sinal de reforço favorável é produzido pelo crítico, o que faz a rede neural aceitar como boa

a resposta produzida. Caso o crítico produza um sinal de reforço não favorável, a resposta é considerada insatisfatória e a rede neural absorve este conhecimento através de seus pesos, com o intuito de não mais repetir essa resposta.

Os dados de entrada devem ser significativos e cobrir amplamente o domínio do problema. Pode-se criar também um conjunto de validação utilizado para verificar a eficiência da rede quanto a sua capacidade de *generalização*, ou seja, a sua capacidade de dar respostas coerentes para dados não apresentados a ela previamente durante o treinamento.

3.2 Redes Neurais Fuzzy Morfológicas

As *Redes Neurais Morfológicas* (MNNs - Morphological Neural Networks) foram introduzidas por Davidson e Ritter no início dos anos 90 como uma nova classe de redes neurais artificiais [17, 16]. Uma MNN pode ser definida como uma ANN cujos neurônios computam uma operação elementar da morfologia matemática (MM), i.e, uma operação de dilatação, erosão, anti-dilatação ou anti-erosão, seguida por uma função de ativação $f(\cdot)$ [53, 54]. Em contraste com o neurônio clássico, o neurônio morfológico efetua uma operação não-linear durante a regra de agregação. Nesta dissertação de mestrado, são tratados apenas os modelos de neurônios morfológicos onde a função identidade é empregada como função de ativação.

Uma *Rede Neural Fuzzy* (FNN - Fuzzy Neural Network) é uma rede neural cujos padrões de entrada, padrões de saída e os pesos sinápticos representam conjuntos fuzzy [24]. Finalmente, uma *Rede Neural Morfológica Fuzzy* (FMNN - Fuzzy Morphological Neural Network) é simplesmente uma rede neural fuzzy que realiza uma das operações elementares da morfologia matemática em cada neurônio [80].

3.3 Memórias Associativas

As *memórias associativas* (AMs - Associative Memories) foram desenvolvidas por Taylor na década de 50 [75] e posteriormente, Steinbuch introduziu o conceito de *matriz de apren-*

dizagem em 1961 [65]. Em 1972, Anderson, Kohonen e Nakano iniciaram os estudos sobre memórias associativas por matriz de correlação [1, 34, 47], baseada na regra de aprendizagem por produto externo que pode ser interpretada como uma generalização do postulado de aprendizagem de Hebb [28]. Hopfield em 1982, apresentou a famosa rede de Hopfield, sugerindo que um sistema dinâmico pode representar uma memória associativa dinâmica, onde cada estado estável do sistema é um padrão memorizado [31]. As memórias associativas são uma classe de redes neurais que utilizam um algoritmo de aprendizado/armazenamento mapeando as entradas e suas respectivas saídas. Cada par de entrada-saída $(\mathbf{x}^\xi, \mathbf{y}^\xi)$ armazenado na memória é chamada de *memória fundamental*. O processo para determinar uma memória associativa é conhecido como *fase de armazenamento*, que se reduz em determinar a matriz dos pesos sinápticos. Nesta fase, apresenta-se o conjunto das memórias fundamentais $\{(\mathbf{x}^\xi, \mathbf{y}^\xi), \xi = 1, 2, \dots, k\}$ cujo objetivo é encontrar um mapeamento que recupere cada um desses pares, i.e, uma função F tal que $F(\mathbf{x}^\xi) = \mathbf{y}^\xi$. Terminando a fase de armazenamento, inicia-se a *fase de recordação*, onde verifica se as memórias fundamentais foram corretamente armazenadas e sua respectiva capacidade de correção, ou seja, se $\tilde{\mathbf{x}}^\xi$ é uma versão ruidosa de \mathbf{x}^ξ , deseja-se que $\tilde{\mathbf{x}}^\xi$ e \mathbf{x}^ξ retorne a mesma saída, o que significa, $F(\tilde{\mathbf{x}}^\xi) = \mathbf{y}^\xi$. Os modelos de AMs que trabalham com esse tipo de associação são denominados *memórias heteroassociativas*. Quando o conjunto das memórias fundamentais é da forma $\{(\mathbf{x}^\xi, \mathbf{x}^\xi), \xi = 1, 2, \dots, k\}$, atribui-se o nome de *memórias autoassociativas*. Um dos maiores problemas encontrados nas memórias associativas são a criação de associações que não fazem parte do conjunto das memórias fundamentais e são armazenadas indevidamente, conhecidas na literatura como *memórias espúrias*. Por outro lado, as características desejáveis numa memória associativa se traduzem em: grande capacidade de armazenamento; tolerância a ruídos ou entradas incompletas; existência de poucas memórias espúrias; recordação rápida e de baixo custo computacional [49].

Os modelos de memórias associativas são classificados segundo o modelo de armazenamento (dinâmica ou estática), natureza das associações armazenadas (autoassociativas ou heteroassociativas) e o tipo dos padrões armazenados: bipolares ($\{-1, 1\}$), binários ($\{0, 1\}$),

discretos (\mathbb{Z}) ou contínuos (\mathbb{R}). Deve-se ainda considerar a complexidade e a capacidade do algoritmo de aprendizagem. Uma *memória associativa estática* é uma rede neural progressiva que associa o padrão de entrada \mathbf{x} ao padrão de saída \mathbf{y} . Por sua vez, estes modelos dividem-se em *memórias associativas lineares* (LAMs - Linear Associative Memories) desenvolvidas por Anderson, Kohonen e Nakano em 1972 [1, 34, 47] e *memórias associativas lineares ótimas* (OLAM - Optimal Linear Associative Memory) propostos por Kohonen e Ruohonen [35].

Uma rede neural *recorrente* usada como mapeamento associativo produz uma *memória associativa dinâmica* (DAM - Dynamic Associative Memory) [26, 47, 31]. A DAM atua como um filtro que corrige versões de entradas parciais e/ou ruidosas $\tilde{\mathbf{x}}$ de um padrão \mathbf{x} . Em outras palavras, inicia-se a memória com um vetor $\mathbf{x}(0) = \tilde{\mathbf{x}}$ e espera-se que a DAM convirja para \mathbf{x} , existindo uma região de atração ao redor de cada uma das memórias fundamentais. Porém, a complexa dinâmica desses modelos, permite que memórias espúrias também tornem-se atratoras, resultando na degradação de sua performance. Uma das DAM mais conhecidas na literatura é a *memória associativa de Hopfield*, desenvolvida pelo físico John Hopfield em 1982 [31]. Essa memória pode ser descrita por uma rede neural recorrente de camada única totalmente conexa com *função sinal* como função de ativação. A convergência da DAM de Hopfield pode ser analisada através de uma função de energia. Na verdade, a dinâmica da rede impõe um processo de minimização da função de energia, estabilizando-se necessariamente em um ponto de mínimo.

CAPÍTULO 4

Memórias Associativas Morfológicas e Memórias Associativas Morfológicas Fuzzy

4.1 Memórias Associativas Morfológicas

As memórias associativas morfológicas (MAMs) foram introduzidas por Ritter e Sussner como um modelo de memória associativa para armazenar e recordar padrões com valores reais [53, 54, 55]. As MAMs representam os primeiros modelos de redes neurais morfológicas e por isso seus fundamentos matemáticos encontram-se na morfologia matemática [29], geralmente no contexto dos reticulados completos [56].

As MAMs podem ser descritas em termos de produtos matriciais definidas em estruturas algébricas chamadas de *belts* e *blogs* (“bounded lattice ordered groups”) [69]. O conjunto dos números reais estendidos $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{-\infty, \infty\}$, representa tanto um belt como um blog, especificamente, $(\mathbb{R}_{\pm\infty}, \vee, +)$ e $(\mathbb{R}_{\pm\infty}, \wedge, +')$ representam belts e $(\mathbb{R}_{\pm\infty}, \vee, \wedge, +, +')$ representa um blog. As operações de $+$ e $+'$ comportam-se como a operação de soma usual e são idênticas sobre $\mathbb{R}_{\pm\infty}$, com exceção quando operam sobre ∞ e $-\infty$:

$$\begin{aligned}\infty + (-\infty) &= (-\infty) + \infty = -\infty \\ \infty +' (-\infty) &= (-\infty) +' \infty = \infty\end{aligned}\tag{4.1}$$

Seja $r \in \mathbb{R}_{\pm\infty}$, defini-se o *conjugado aditivo* r^* de r por:

$$r^* \equiv \begin{cases} -r & \text{se } r \in \mathbb{R} \\ -\infty & \text{se } r = +\infty \\ +\infty & \text{se } r = -\infty \end{cases} \quad (4.2)$$

As seguintes propriedades são válidas para todo $r, u \in \mathbb{R}_{\pm\infty}$:

$$(r^*)^* = r \quad \text{e} \quad r \wedge u = (r^* \vee u^*)^* \quad (4.3)$$

Para matrizes A e $B \in \mathbb{R}_{\pm\infty}^{n \times m}$, o máximo $C = A \vee B$ e o mínimo $D = A \wedge B$ entre as matrizes A e B são dadas, respectivamente por:

$$c_{ij} = a_{ij} \vee b_{ij} \quad \text{e} \quad d_{ij} = a_{ij} \wedge b_{ij} \quad \text{para} \quad 1 \leq i \leq n \quad \text{e} \quad 1 \leq j \leq m \quad (4.4)$$

e tem-se as seguintes propriedades:

$$A \wedge B = (A^* \vee B^*)^* \quad \text{e} \quad A \vee B = (A^* \wedge B^*)^* \quad (4.5)$$

Considere as matrizes $A \in \mathbb{R}_{\pm\infty}^{n \times k}$ e $B \in \mathbb{R}_{\pm\infty}^{k \times m}$. Denomina-se *produto max* de A por B a matriz $C = A \boxtimes B$, de forma similar, tem-se o *produto min* de A por B a matriz $D = A \boxdot B$, definidas em termos das seguintes equações:

$$c_{ij} = \bigvee_{l=1}^k (a_{il} + b_{lj}) \quad \text{e} \quad d_{ij} = \bigwedge_{l=1}^k (a_{il} + b_{lj}) \quad (4.6)$$

Sejam as matrizes $A \in \mathbb{R}_{\pm\infty}^{m \times p}$ e $B \in \mathbb{R}_{\pm\infty}^{p \times n}$. Existe uma elegante dualidade entre o produto max e o produto min, dada por:

$$A \boxtimes B = (B^* \boxdot A^*)^* \quad \text{e} \quad A \boxdot B = (B^* \boxtimes A^*)^* \quad (4.7)$$

Considere o conjunto das memórias fundamentais $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, 2, \dots, k\}$, onde $\mathbf{x}^\xi = [x_1^\xi, x_2^\xi, \dots, x_n^\xi]^t \in \mathbb{R}^n$ e $\mathbf{y}^\xi = [y_1^\xi, y_2^\xi, \dots, y_m^\xi]^t \in \mathbb{R}^m$. Por simplicidade, denota-se as matrizes X e Y por $[\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^k] \in \mathbb{R}^{n \times k}$ e $[\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^k] \in \mathbb{R}^{m \times k}$, respectivamente. Desta forma, obtem-se um par de matrizes (X, Y) como padrão associado e define-se a matriz W_{XY} em termos do *produto min* e a matriz M_{XY} do *produto max* de Y por X segundo as equações [69]:

$$W_{XY} = Y \boxtimes X^* \quad M_{XY} = Y \boxdot X^* \quad (4.8)$$

onde os elementos da matriz W_{XY} e da matriz M_{XY} , são descritos pelas equações:

$$w_{ij} = \bigwedge_{\xi=1}^k (y_i^\xi + (-x_j^\xi)) \quad m_{ij} = \bigvee_{\xi=1}^k (y_i^\xi + (-x_j^\xi)) \quad (4.9)$$

De acordo com 4.9, novos pares podem ser facilmente adicionadas nas memórias W_{XY} e M_{XY} . De fato, pode-se armazenar um novo par $(\mathbf{x}^{k+1}, \mathbf{y}^{k+1})$ ao conjunto das memórias fundamentais $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, 2, \dots, k\}$ da seguinte forma:

$$W_{XY}^{nova} = W_{XY} \wedge [\mathbf{y}^{k+1} + (-\mathbf{x}^{k+1})^t] \quad M_{XY}^{nova} = M_{XY} \vee [\mathbf{y}^{k+1} + (-\mathbf{x}^{k+1})^t] \quad (4.10)$$

As memórias associativas morfológicas são vistas como redes neurais progressivas de camada única e são descritas segundo as expressões, para $\xi = 1, \dots, k$:

$$\mathbf{y}^\xi = W_{XY} \boxtimes \mathbf{x}^\xi \quad \text{e} \quad \mathbf{y}^\xi = M_{XY} \boxdot \mathbf{x}^\xi \quad (4.11)$$

As expressões em 4.11 podem ser interpretadas como operadores de dilatação $\delta : \mathbb{R}_{\pm\infty}^n \rightarrow \mathbb{R}_{\pm\infty}^m$ e erosão $\epsilon : \mathbb{R}_{\pm\infty}^n \rightarrow \mathbb{R}_{\pm\infty}^m$, respectivamente, da morfologia matemática [59, 64]. De fato, seja $\delta(\mathbf{x}) = \bigvee_{j=1}^m (w_{ij} + x_j)$, para $i = 1, 2, \dots, m$. Mostraremos que o operador δ é uma dilatação, ou seja, δ comuta com o operador de máximo [29], para toda coleção $\mathbf{x}^l \in \mathbb{R}_{\pm\infty}^n$, $l \in L$ e $i = 1, \dots, m$:

$$\begin{aligned}
\delta \left(\bigvee_{l \in L} \mathbf{x}^l \right) &= \bigvee_{j=1}^n \left[w_{ij} + \left(\bigvee_{l \in L} \mathbf{x}^l \right)_j \right] = \bigvee_{j=1}^n \left[w_{ij} + \left(\bigvee_{l \in L} x_j^l \right) \right] = \\
&= \bigvee_{j=1}^n \left[\bigvee_{l \in L} (w_{ij} + x_j^l) \right] = \bigvee_{l \in L} \left[\bigvee_{j=1}^n (w_{ij} + x_j^l) \right] = \bigvee_{l \in L} \delta(\mathbf{x}^l)
\end{aligned} \tag{4.12}$$

De forma similar, mostra-se que o operador ϵ comuta com o operador de mínimo e portanto, representa uma erosão no sentido de Heijmans [29]. Uma imagem $\tilde{\mathbf{x}}$ é uma versão corrompida de \mathbf{x} com ruído *erosivo* (*dilatativo*) se $\tilde{\mathbf{x}} \leq \mathbf{x}$ ($\tilde{\mathbf{x}} \geq \mathbf{x}$). Desta forma, uma dilatação (erosão) é usada para revomer o ruídos erosivos (dilatativo) de uma imagem.

Note que existe uma importante relação de dualidade envolvendo as Equações 4.5 e 4.7, ou seja, a segunda metade das Equações 4.5 e 4.7 são duais da primeira metade. Assim, toda afirmação induz uma afirmação dual, trocando simplesmente cada símbolo “ \wedge ” por “ \vee ” e vice versa, e revertendo cada inequação [15]. Portanto, serão enunciados apenas as afirmações primais envolvendo as MAMs e omitidas as correspondentes afirmações duais. A seguir, serão apresentados os principais resultados das memórias associativas morfológicas. Denota-se o símbolo $O(W_{XY})$ como o conjunto de todos elementos $W_{XY} \boxtimes \mathbf{x}$ tal que $\mathbf{x} \in \mathbb{R}^n$. Analogamente, o símbolo $O(M_{XY})$ denota o conjunto de todos os elementos $M_{XY} \boxtimes \mathbf{x}$ tal que $\mathbf{x} \in \mathbb{R}^n$. A seguir, o Teorema 4.1.1 e o Corolário 4.1.1 caracterizam completamente a saída das memórias morfológicas heteroassociativas para qualquer padrão de entrada \mathbf{x} [70].

Teorema 4.1.1. *Para $X \in \mathbb{R}^{n \times k}$ e $Y \in \mathbb{R}^{m \times k}$, os conjuntos $O(W_{XY})$ e $O(M_{XY})$ coincidem. Se \mathcal{O} denota estes conjuntos, então \mathcal{O} consiste exatamente das seguintes expressões:*

$$\bigvee_{i=1}^n \bigwedge_{\xi=1}^k \left(a_i^\xi + \mathbf{y}^\xi \right), \text{ onde } a_i^\xi \in \mathbb{R}. \tag{4.13}$$

Alternativamente, o conjunto \mathcal{O} pode ser caracterizado como o conjunto:

$$\bigwedge_{i=1}^n \bigvee_{\xi=1}^k \left(c_j^\xi + \mathbf{y}^\xi \right), \text{ onde } c_j^\xi \in \mathbb{R}. \tag{4.14}$$

Além disso, para um vetor arbitrário $\mathbf{x} \in \mathbb{R}^n$, o padrão $W_{XY} \boxtimes \mathbf{x}$ é igual a menor expressão dada pela Equação 4.13 tal que $\bigvee_{i=1}^n \bigwedge_{\xi=1}^k (a_i^\xi + \mathbf{x}^\xi)$ é o supremo de \mathbf{x} no conjunto dos pontos fixos $\mathcal{F} = F(W_{XX})$ de $W_{XX} \in \mathbb{R}^{n \times n}$, ou seja, todo $\mathbf{x} \in \mathbb{R}^n$ tal que $W_{XX} \boxtimes \mathbf{x} = \mathbf{x}$.

Corolário 4.1.1. Se $X \in \mathbb{R}^{n \times k}$, $Y \in \mathbb{R}^{m \times k}$ e $\mathbf{x} \in \mathbb{R}^n$ tal que $W_{XX} \boxtimes \mathbf{x} = \mathbf{x}^\gamma$. Tem-se $W_{XY} \boxtimes \mathbf{x} = \mathbf{y}^\gamma$ se e somente se a seguinte implicação é válida para todo $c^\xi \in \mathbb{R}$, onde $\xi = 1, 2, \dots, k$.

$$\mathbf{x}^\gamma \leq \bigvee_{\xi=1}^k c^\xi + \mathbf{x}^\xi \Rightarrow \mathbf{y}^\gamma \leq \bigvee_{\xi=1}^k c^\xi + \mathbf{y}^\xi \quad (4.15)$$

As condições vistas no Teorema 4.1.1 e no Corolário 4.1.1 são difíceis de serem verificadas. Assim, um exemplo numérico será apresentado a seguir.

Exemplo 4.1.1. Considere as matrizes:

$$X = [\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & -3 \\ 0 & -4 & 0 \end{bmatrix} \quad e \quad Y = [\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3] = \begin{bmatrix} 0 & -1 & 0 \\ 1 & -1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.16)$$

Pela Expressão 4.8, tem-se:

$$W_{XY} = [\mathbf{y}^1 +' (-\mathbf{x}^1)^t] \wedge [\mathbf{y}^2 +' (-\mathbf{x}^2)^t] \wedge [\mathbf{y}^3 +' (-\mathbf{x}^3)^t] = \begin{bmatrix} -1 & 0 & 0 \\ -2 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.17)$$

$$M_{XY} = [\mathbf{y}^1 + (-\mathbf{x}^1)^t] \vee [\mathbf{y}^2 + (-\mathbf{x}^2)^t] \vee [\mathbf{y}^3 + (-\mathbf{x}^3)^t] = \begin{bmatrix} 0 & 3 & 3 \\ 1 & 1 & 3 \\ 0 & 3 & 4 \end{bmatrix} \quad (4.18)$$

Pode-se verificar que $W_{XY} \boxtimes \mathbf{x}^\xi = \mathbf{y}^\xi = M_{XY} \boxtimes \mathbf{x}^\xi$, para $\xi = 1, 2, 3$. Por exemplo:

$$W_{XY} \boxtimes \mathbf{x}^1 = \begin{bmatrix} -1 & 0 & 0 \\ -2 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \boxtimes \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \mathbf{y}^1 \quad (4.19)$$

$$M_{XY} \boxtimes \mathbf{x}^2 = \begin{bmatrix} 0 & 3 & 3 \\ 1 & 1 & 3 \\ 0 & 3 & 4 \end{bmatrix} \boxtimes \begin{bmatrix} 0 \\ -2 \\ -4 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} = \mathbf{y}^2 \quad (4.20)$$

Verifica-se que $W_{XY} \boxtimes \tilde{\mathbf{x}}^3 = \mathbf{y}^3$, então a Equação 4.15 do Corolário 4.1.1 é satisfeita, para todo $c^\xi \in \mathbb{R}$ e $\xi = 1, 2, 3$. De fato, tomando $c^1 = 1$, $c^2 = 2$ e $c^3 = 1$, tem-se:

$$W_{XY} \boxtimes \tilde{\mathbf{x}}^3 = \begin{bmatrix} -1 & 0 & 0 \\ -2 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \boxtimes \begin{bmatrix} -1 \\ -3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} = \mathbf{y}^3 \quad (4.21)$$

Pela Expressão 4.15, segue que

$$\mathbf{x}^3 \leq \bigvee_{\xi=1}^3 (c^\xi + \mathbf{x}^\xi) \Rightarrow \mathbf{y}^3 \leq \bigvee_{\xi=1}^3 (c^\xi + \mathbf{y}^\xi). \quad (4.22)$$

De fato,

$$\bigvee_{\xi=1}^3 (c^\xi + \mathbf{x}^\xi) = [(c^1 + \mathbf{x}^1) \vee (c^2 + \mathbf{x}^2) \vee (c^3 + \mathbf{x}^3)] = \quad (4.23)$$

$$\left[\left(\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \vee \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \vee \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right) \right] = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \geq \begin{pmatrix} 0 \\ -3 \\ 0 \end{pmatrix} = \mathbf{x}^3 \quad (4.24)$$

e

$$\bigvee_{\xi=1}^3 (c^\xi + \mathbf{y}^\xi) = [(c^1 + \mathbf{y}^1) \vee (c^2 + \mathbf{y}^2) \vee (c^3 + \mathbf{y}^3)] = \quad (4.25)$$

$$\left[\left(\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \vee \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \vee \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right) \right] = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} = \mathbf{y}^3 \quad (4.26)$$



Figura 4.1: Imagens originais utilizadas para construir as memórias morfológicas autoassociativas W_{XX} e M_{XX} .

De 4.9, pode-se estabelecer a seguinte relação:

$$W_{XY} \boxtimes X \leq Y \leq M_{XY} \boxtimes X \quad (4.27)$$

e se existem matrizes A e B tais que $A \boxtimes X = Y$ e $B \boxtimes X = Y$ então a seguinte desigualdade é satisfeita [69]:

$$A \leq W_{XY} \leq M_{XY} \leq B \quad \text{e} \quad W_{XY} \boxtimes X = Y = M_{XY} \boxtimes X \quad (4.28)$$

Portanto, a memória morfológica $W_{XY} = Y \boxtimes X^*$ representa o supremo das matrizes A tal que $A \boxtimes X \leq Y$ e $M_{XY} = Y \boxtimes X^*$ representa o ínfimo das matrizes B tal que $Y \leq B \boxtimes X$.

Corolário 4.1.2. Para $X \in \mathbb{R}^{n \times k}$, os conjuntos dos pontos fixos $\mathcal{F}(W_{XX})$ e $\mathcal{F}(M_{XX})$ coincidem. Se \mathcal{F} denota esses conjuntos, então \mathcal{F} é dada pela seguinte expressão:

$$\bigvee_{i=1}^n \bigwedge_{\xi=1}^k (a_i^\xi + \mathbf{x}^\xi), \text{ onde } a_i^\xi \in \mathbb{R}. \quad (4.29)$$

Além disso, dado um padrão de entrada arbitrário $\mathbf{x} \in \mathbb{R}^n$, tem-se

$$W_{XX} \boxtimes \mathbf{x} = \hat{\mathbf{x}} \quad e \quad M_{XX} \boxtimes \mathbf{x} = \check{\mathbf{x}}, \quad (4.30)$$

onde $\hat{\mathbf{x}}$ é o supremo de \mathbf{x} em \mathcal{F} e $\check{\mathbf{x}}$ é o ínfimo de \mathbf{x} em \mathcal{F} .



Figura 4.2: A primeira linha representa versões incompletas das imagens originais e a segunda linha corresponde aos padrões de saída da memória morfológica W_{XX}

Imagens incompletas das imagens originais (Figura 4.1) foram perfeitamente recuperadas por W_{XX} e o resultado pode ser visto na Figura 4.2. Note que o conjunto dos pontos fixos de W_{XX} contém as memórias fundamentais \mathbf{x}^ξ , para $\xi = 1, \dots, k$. Portanto, pode-se armazenar quantos padrões forem desejados numa MAM no caso autoassociativo. E mais, o Corolário 4.1.2 revela que os modelos de MAMs autoassociativas W_{XX} possuem convergência em um única iteração, i.e, o padrão recordado pela MAM $\mathbf{y} = W_{XX} \boxtimes \mathbf{x}$ é um ponto fixo de W_{XX} para todo $\mathbf{x} \in \mathbb{R}^n$ [55, 69].

4.2 Memórias Associativas Morfológicas Fuzzy

4.2.1 Introdução

As *memórias associativas morfológicas fuzzy* (FMAMs - Fuzzy Morphological Associative Memories) são um desenvolvimento muito recente na área de sistemas neuro-fuzzy, estabelecidas por Valle e Sussner [71, 80]. De modo geral, entende-se como uma FMAM um modelo de memória associativa descrita por uma rede neuro-fuzzy cujos neurônios executam uma das operações morfológicas de dilatação, erosão, anti-dilatação ou anti-erosão. A matemática básica para as FMAMs é encontrada na morfologia matemática fuzzy que confia no fato do conjunto $[0, 1]^{\mathbf{X}}$ representar um reticulado completo para qualquer universo \mathbf{X} [46, 56].

As *memórias associativas fuzzy* (FAMs - Fuzzy Associative Memories) foram introduzidas por Kosko na década de 90 [37, 36] como redes neurais fuzzy progressivas de camada única que armazena regras fuzzy em termos de composições max-min ou max-produto para sintetizar a matriz dos pesos sinápticos W . Os modelos de FAMs de Kosko apresentam baixa capacidade de armazenamento e posteriormente outros pesquisadores, como Chung e Lee, Junbo, Liu desenvolveram outras versões de memórias associativas fuzzy capazes de armazenar múltiplos pares de associações fuzzy [6, 10, 13, 32, 38]. Muitos modelos de FAMs existentes na literatura pertencem à classe das FMAMs [80], incluindo as *memórias associativas fuzzy implicativas* (IFAMs - Implicative Fuzzy Associative Memories) introduzidas por Sussner e Valle [72], e podem ser vistas como uma extensão das memórias associativas morfológicas para o caso fuzzy.

4.2.2 Neurônios Max-C, Min-D e Min-J

Os tipos de neurônios mais utilizados nos modelos de memórias associativas fuzzy podem ser formulados em termos dos produtos matriciais fuzzy max- C , min- D e min- J , definidas por 2.30 e 2.31. Considere o conjunto fuzzy de entrada $\mathbf{x} = [x_1, \dots, x_n]^t \in [0, 1]^n$, a saída fuzzy $y \in [0, 1]$ e $w_j \in [0, 1]$ os pesos sinápticos que formam o vetor $\mathbf{w} = [w_1, \dots, w_n]^t$. Defini-se o *neurônio max-C* da seguinte forma [72]:

$$y = \left[\bigvee_{j=1}^n C(w_j, x_j) \right] = (\mathbf{w}^t \circ \mathbf{x}) \quad (4.31)$$

De forma similar, defini-se o *neurônios min-D* e o *neurônios min-J*, através da Equação 4.32:

$$y = \left[\bigwedge_{j=1}^n D(m_j, x_j) \right] = (\mathbf{m}^t \bullet \mathbf{x}) \quad y = \left[\bigwedge_{j=1}^n J(x_j, m_j) \right] = (\mathbf{m}^t \otimes \mathbf{x}) \quad (4.32)$$

onde $\mathbf{m} = [m_1, \dots, m_n]^t \in [0, 1]^n$.

Para particulares escolhas de conjunções fuzzy, obtêm-se particulares neurônios max- C . Por isso, indica-se com um subscrito qual conjunção fuzzy adotada. Por exemplo, max- C_M denota o neurônio que é baseado na conjunção fuzzy do mínimo. Pela mesma razão, uma similar notação é aplicada aos neurônios min- D e min- J . Particularmente, esta dissertação de mestrado tem interesse nos neurônios max- T que ocorrem nos modelos de IFAMs, onde T é uma t-norma contínua e o símbolo \otimes_T denota o produto min- J_T , onde J_T é uma implicação fuzzy reversa que forma uma adjunção com T . Como uma t-norma contínua representa uma dilatação em $[0,1]$, tem-se que as IFAMs pertencem à classe das FMAMs.

Se a conjunção fuzzy C e a disjunção fuzzy D são operadores duais com respeito à alguma negação fuzzy N , então os neurônios max- C e min- D são *duais com respeito a negação fuzzy* N [72]. Em outras palavras, considere a operação $\delta(\mathbf{x}) = (\mathbf{w}^T \circ \mathbf{x})$ realizada pelo neurônio max- C e $\epsilon(\mathbf{x}) = (\mathbf{m}^T \bullet \mathbf{x})$ a operação realizada pelo neurônio min- D , para todo $\mathbf{x} \in [0, 1]^n$. Conseguem-se a seguinte igualdade $\epsilon(\mathbf{x}) = N(\delta(N(\mathbf{x})))$, $\forall \mathbf{x} \in [0, 1]^n$, onde $m_j = N(w_j)$. Por exemplo, o neurônio max- C_M e o neurônio min- D_M são duais com a respeito a negação fuzzy usual N_S .

Caso a conjunção fuzzy $C(a, \cdot) : [0, 1] \rightarrow [0, 1]$ represente uma dilatação para todo $a \in [0, 1]$, então fale-se em *neurônio morfológico max- C* . Exemplos de neurônios morfológicos

$\max-C$ são dados pelos neurônios $\max-C_M$, $\max-C_P$ e $\max-C_L$. Da mesma forma, se a disjunção fuzzy $D(a, \cdot) : [0, 1] \rightarrow [0, 1]$ represente uma erosão para todo $a \in [0, 1]$, então classifica-se esses neurônios como *neurônio morfológico min-D* [20, 72, 80].

4.2.3 FMAMs Max-C

Diversos modelos de FAMs são vistos como redes neurais progressivas de camada única, compostas de neurônios morfológicos $\max-C$ [80]. Esses modelos são descritos segundo a equação:

$$\mathbf{y} = \mathcal{W}(\mathbf{x}) = W \circ \mathbf{x} \quad (4.33)$$

onde $W \in [0, 1]^{m \times n}$ representa a matriz dos pesos sinápticos e $\mathbf{x} \in [0, 1]^n$ e $\mathbf{y} \in [0, 1]^m$ são os padrões de entrada e saída fuzzy, respectivamente.

Note que a Equação 4.33 descreve uma FMAM se, e somente se, a correspondente conjunção fuzzy corresponde uma dilatação. Neste caso, a função associativa \mathcal{W} representa uma dilatação de $[0, 1]^n$ em $[0, 1]^m$ e um modelo de AM dado pela Equação 4.33 pertence a classe das FMAMs $\max-C$ [80].

4.2.4 FMAMs Min-D, a Negação das FMAMs max-C

Os modelos de FAMs discutidos nesta seção são descritos pela Equação 4.34, onde $M \in [0, 1]^{m \times n}$ representa a matriz dos pesos sinápticos e $\mathbf{x} \in [0, 1]^n$ e $\mathbf{y} \in [0, 1]^m$ são os padrões de entrada e saída fuzzy, respectivamente.

$$\mathbf{y} = \mathcal{M}(\mathbf{x}) = M \bullet \mathbf{x} \quad (4.34)$$

Note que a Equação 4.34 descreve uma FMAM se e somente se a correspondente disjunção fuzzy corresponde à uma erosão. Neste caso, a função associativa \mathcal{M} representa uma erosão de $[0, 1]^n$ em $[0, 1]^m$. Na verdade, duas relações de dualidade da morfologia matemática

podem ser usadas para formular novos modelos de FMAMs. Em particular, seja N uma negação fuzzy arbitrária e $N(\mathbf{x})$ a negação fuzzy dos elementos de um vetor \mathbf{x} . Define-se a *negação de uma FMAM max-C* \mathcal{W} como o modelo de AM \mathcal{M} que corresponde a negação de \mathcal{W} com respeito a N , ou seja, a FAM \mathcal{M} dada pela Equação 4.35, onde \mathbf{x} e \mathbf{y} são o padrão de entrada e o padrão recordado, respectivamente:

$$\mathbf{y} = \mathcal{M}(\mathbf{x}) = \mathcal{M}^N(\mathbf{x}) = N(M \bullet N(\mathbf{x})) \quad (4.35)$$

O seguinte teorema revela que a negação de uma FMAM max-C representa uma FMAM min-D [80].

Teorema 4.2.1. *Considere N uma negação fuzzy. Seja a função ϕ_N definida por $\phi_N(\mathcal{W}) = \mathcal{W}^N$ para toda FAM max-C \mathcal{W} , então ϕ_N constitui uma bijeção entre o conjunto das FAMs max-C e o conjunto das FAMs min-D. Dada uma arbitrária FAM max-C \mathcal{W} com correspondente matriz dos pesos W , tem-se que a matriz dos pesos sinápticos da FAM min-D $\mathcal{M} = \mathcal{W}^N$ é simplesmente dada por $M = N(W)$. E mais, os operadores fuzzy C e D são duais com respeito a N . Em particular, se C é uma dilatação no segundo argumento, então D é uma erosão no segundo argumento. Portanto, restringindo ϕ_N ao conjunto das FMAMs max-C, tem-se uma bijeção entre o conjunto das FMAMs max-C e FMAMs min-D. Similarmente, se C é uma t -norma então D é uma s -norma e portanto, ϕ_N induz uma bijeção entre o conjunto das FAMs max-T e o conjunto das FAMs min-S.*

4.2.5 Aprendizado por Adjunção

Nesta seção, serão introduzidas estratégias de aprendizagem para as FMAMs max-C e min-D, formalizadas a partir das relações de dualidade por adjunção e por isso denominado *aprendizado por adjunção* ou *aprendizado fuzzy implicativo* introduzidos recentemente por Valle e Sussner [71, 80]. Considere o conjunto das memórias fundamentais $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, 2, \dots, k\}$. Por simplicidade, sejam $X \in [0, 1]^{n \times k}$ e $Y \in [0, 1]^{m \times k}$ as matrizes cujas as colunas são os vetores \mathbf{x}^ξ e \mathbf{y}^ξ , respectivamente. Seja a FMAM max-C dada pela Equação 4.33 e o operador $\mathcal{D}_X : [0, 1]^{m \times n} \rightarrow [0, 1]^{m \times k}$ definido da seguinte forma:

$$\mathcal{D}_X(W) = W \circ X \quad (4.36)$$

Se existe uma matriz $W \in [0, 1]^{m \times n}$ tal que $Y = \mathcal{D}_X(W)$ então a FMAM max- C produz a desejada saída após a apresentação de uma entrada não distorcida \mathbf{x}^ξ , i.e, a FMAM recorda perfeitamente as associações $(\mathbf{x}^\xi, \mathbf{y}^\xi)$ para cada $\xi = 1, \dots, k$. Note que, o operador \mathcal{D}_X representa uma dilatação para todo $X \in [0, 1]^{n \times k}$ se, e somente se, a conjunção fuzzy $C(\cdot, x)$ representa uma dilatação para todo $x \in [0, 1]$.

Se \mathcal{D}_X é uma dilatação então pela Proposição 2.1.1 existe uma única erosão fuzzy $\mathcal{E}_X : [0, 1]^{m \times k} \rightarrow [0, 1]^{m \times n}$ que forma uma adjunção com \mathcal{D}_X . A erosão fuzzy \mathcal{E}_X depende de $X \in [0, 1]^{n \times k}$ e produz uma matriz em $[0, 1]^{m \times n}$ para toda entrada $Y \in [0, 1]^{m \times k}$. A matriz dos pesos sinápticos para uma FMAM max- C é definida como se segue:

$$W = \mathcal{E}_X(Y) \quad (4.37)$$

O Teorema 4.2.2 descreve a principal propriedade de uma matriz dos pesos sinápticos W dada pela Equação 4.37. Em particular, esse teorema implica que W é a melhor aproximação inferior de Y em termos do produto max- C [79].

Teorema 4.2.2. *Sejam $X = [\mathbf{x}^1, \dots, \mathbf{x}^k] \in [0, 1]^{n \times k}$ e $Y = [\mathbf{y}^1, \dots, \mathbf{y}^k] \in [0, 1]^{m \times k}$. Considere a adjunção $(\mathcal{D}_X, \mathcal{E}_X)$, onde \mathcal{D}_X é dada pela Equação 4.36. A matriz dos pesos sinápticos $W = \mathcal{E}_X(Y)$ representa o supremo das matrizes $V \in [0, 1]^{m \times n}$ tal que $V \circ X \leq Y$, i.e., W satisfaz a seguinte equação:*

$$W = \bigvee \{V \in [0, 1]^{m \times n} : V \circ X \leq Y\} \quad (4.38)$$

Em particular, se existe uma matriz $V \in [0, 1]^{m \times n}$ tal que $V \circ X = Y$ então $V \leq W$ e $W \circ X = Y$.

Uma consequência do Teorema 4.2.2 envolvendo o caso autoassociativo é apresentado no Corolário 4.2.1 que garante a recordação perfeita da FMAM max- C após armazenar os padrões $\{\mathbf{x}^1, \dots, \mathbf{x}^k\}$. O produto max- C tem uma identidade à esquerda, se existe uma matriz fuzzy $I \in [0, 1]^{n \times n}$ tal que $I \circ X = X$ para todo $X \in [0, 1]^{n \times k}$ [79].

Corolário 4.2.1. *Considere a adjunção $(\mathcal{D}_X, \mathcal{E}_X)$, onde \mathcal{D}_X é dada pela Equação 4.36 e $X = [\mathbf{x}^1, \dots, \mathbf{x}^k] \in [0, 1]^{n \times k}$. Se o produto max- C tem uma identidade à esquerda $I \in [0, 1]^{n \times n}$ então a matriz dos pesos sinápticos $W = \mathcal{E}_X(X)$ é tal que $W \circ \mathbf{x}^\xi = \mathbf{x}^\xi$, para todo $\xi = 1, \dots, k$.*

Em particular, uma t-norma satisfaz $T(1, x) = x$ para todo $x \in [0, 1]$. Logo, pode-se armazenar quantos padrões forem desejados em uma FMAM max- T autoassociativa utilizando a regra de aprendizagem descrita pela Equação 4.37. O Teorema 4.2.3, mostra que a matriz dos pesos sinápticos W é sintetizada através do produto min- J [80].

Teorema 4.2.3. *Sejam $X = [\mathbf{x}^1, \dots, \mathbf{x}^k] \in [0, 1]^{n \times k}$ e $Y = [\mathbf{y}^1, \dots, \mathbf{y}^k] \in [0, 1]^{m \times k}$. Considere o operador \mathcal{D}_X dado pela Equação 4.36, baseado em uma conjunção fuzzy C que representa uma dilatação em ambos argumentos. A matriz dos pesos sinápticos $W = \mathcal{E}_X(Y)$ é dada pelo produto min- J , onde J é a implicação fuzzy reversa adjunta à C :*

$$W = Y \circledast X^t \quad (4.39)$$

A regra de aprendizagem para uma FMAM min- D pode ser formalizada a partir de uma FMAM max- C . De fato, considere uma negação fuzzy N arbitrária e suponha que a disjunção fuzzy D comuta com o operador de ínfimo no primeiro argumento, ou seja, $D(\cdot, x)$ é uma erosão para todo $x \in [0, 1]$. Pelo Teorema 4.2.1, existe uma FMAM max- C \mathcal{W} tal que \mathcal{W} e \mathcal{M} são duais com respeito a N . Essa FMAM max- C é baseada em uma conjunção fuzzy C que é dual a D com respeito a N . Na verdade, $C(\cdot, x)$ é uma dilatação para todo $x \in [0, 1]$ e o aprendizado por adjunção pode ser aplicado para armazenar um conjunto de associações. Em particular, dadas as matrizes $X = [\mathbf{x}^1, \dots, \mathbf{x}^k] \in [0, 1]^{n \times k}$ e $Y = [\mathbf{y}^1, \dots, \mathbf{y}^k] \in [0, 1]^{m \times k}$, define-se $W = N(Y) \circledast N(X)^t$ dado ao fato que $N(Y) = W \circ N(X)$ se e somente se $Y = M \bullet X$,

onde $M = N(W)$. Resumindo essas observações, a matriz dos pesos sinápticos $M \in [0, 1]^{m \times n}$ da FMAM min- D é dada segundo a equação:

$$M = N(W) = N(N(Y) \circledast N(X)^t) \quad (4.40)$$

O produto min- J é baseado na implicação fuzzy reversa que é adjunta à conjunção fuzzy C que é dual à D com respeito a N . O Teorema 4.2.4 revela que a matriz M dada pela Equação 4.40 é a melhor aproximação superior de Y no sentido do produto min- D [80].

Teorema 4.2.4. *A matriz dos pesos sinápticos M dada pela Equação 4.40 representa o ínfimo do conjunto das matrizes $U \in [0, 1]^{m \times n}$ tal que $U \circ X \geq Y$, i.e, M satisfaz a seguinte equação:*

$$M = \bigwedge \{U \in [0, 1]^{m \times n} : U \bullet X \geq Y\} \quad (4.41)$$

Em particular, se existe uma matriz $U \in [0, 1]^{m \times n}$ tal que $U \circ X = Y$ então $V \geq M$ e $M \bullet X = Y$.

Note que o Teorema 4.2.4 é dual ao Teorema 4.2.2. Assim, toda afirmação envolvendo a FMAM max- C com matriz dos pesos sinápticos dada pela Equação 4.39, induz uma correspondente afirmação dual à FMAM min- D com matriz dos pesos sinápticos dada pela Equação 4.41.

4.3 Memórias Associativas Fuzzy Implicativas

4.3.1 Introdução

As *memórias associativas fuzzy implicativas* (IFAMs - Implicative Fuzzy Associative Memories) foram introduzidas recentemente por Sussner e Valle [72] como modelos de memórias associativas descritas por redes neurais fuzzy progressivas de camada única, composta de

neurônios morfológicos max- T . A matriz dos pesos sinápticos é sintetizada através do *aprendizado por adjunção*, onde a implicação fuzzy reversa J_T é adjunta a uma t-norma contínua T . Os modelos de IFAMs pertencem à classe das FMAMs, uma vez que uma t-norma contínua representa uma dilatação em $[0,1]$. Neste contexto, também foram estabelecidos os modelos de *memórias associativas fuzzy implicativas duais*, compostas de neurônios morfológicos min- S que executam uma operação de erosão e portanto, as IFAMs duais também pertencem à classe das FMAMs. Na verdade, existe uma relação de dualidade entre a IFAM e a IFAM dual com respeito à negação fuzzy N . Portanto, toda afirmação envolvendo o modelo de IFAM produz uma afirmação dual para a IFAM dual. Aproveitando-se deste fato, serão focalizados os modelos de IFAMs e omitidos os resultados para o caso dual [72]. Por fim, serão apresentados os resultados mais importantes para os modelos de *memórias fuzzy implicativas autoassociativas* incluindo capacidade de armazenamento, convergência e tolerância a ruídos.

4.3.2 Mémória Associativa Fuzzy Implicativa

Seja J_T uma implicação fuzzy reversa que forma uma adjunção com uma t-norma contínua T . Dado uma padrão de entrada $\mathbf{x} \in [0, 1]^n$, a IFAM composta de neurônios max- T , produz a saída $\mathbf{y} \in [0, 1]^m$ segundo a equação:

$$\mathbf{y} = (W \circ_T \mathbf{x}) \vee \theta \quad \text{onde} \quad \theta = \bigwedge_{\xi=1}^k \mathbf{y}^\xi. \quad (4.42)$$

Os neurônios max- T da IFAM executam uma operação de dilatação, uma vez que a t-norma contínua $T(x, \cdot) : [0, 1] \rightarrow [0, 1]$ aplicada neste modelo de memória associativa fuzzy comuta com o operador de supremo, para todo $x \in [0, 1]$ [29]. Portanto, as IFAMs pertencem a classe das memórias associativas morfológicas fuzzy. A matriz dos pesos sinápticos W é sintetizada empregando o *aprendizado por adjunção*, definido no Teorema 4.2.3 [80]. Particulares escolhas de t-normas T e implicações fuzzy reversas J_T levam a particulares modelos de IFAM. As escolhas de T e J_T indicam o nome do modelo de IFAM. Por exemplo,

a IFAM de Lukasiewicz, adota-se a J_L associada a t-norma C_L . Tal modelo é dado pelas equações: $\mathbf{y} = (W \circ_L \mathbf{x}) \vee \theta$ e $W = Y \circledast_L X^T$.

Exemplo 4.3.1. *Seja $X \in [0, 1]^{5 \times 3}$ a matriz cujas colunas são formadas pelos padrões de entrada $\mathbf{x}^\xi \in [0, 1]^5$ e $Y \in [0, 1]^{3 \times 3}$ a matriz cujas colunas são formadas pelos padrões de saída $\mathbf{y}^\xi \in [0, 1]^3$ para $\xi = 1, 2, 3$:*

$$X = \begin{bmatrix} 0.5 & 0.1 & 0.8 \\ 0.5 & 0.3 & 0.4 \\ 0.4 & 0.3 & 0.6 \\ 0.4 & 0.4 & 0.7 \\ 0.3 & 0.3 & 0.4 \end{bmatrix} \quad e \quad Y = \begin{bmatrix} 0.5 & 0.5 & 0.6 \\ 0.6 & 0.6 & 0.8 \\ 0.3 & 0.4 & 0.4 \end{bmatrix} \quad (4.43)$$

Logo, o vetor bias θ e a matriz dos pesos sinápticos da IFAM W de Lukasiewicz são dados por:

$$W = Y \circledast X^t = \begin{bmatrix} 0.8 & 0.9 & 1 & 0.9 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0.6 & 0.8 & 0.8 & 0.7 & 1 \end{bmatrix} \quad e \quad \theta = [0.5 \ 0.6 \ 0.3]^t \quad (4.44)$$

Pode-se verificar que $\mathbf{y}^\xi = (W \circ_L \mathbf{x}^\xi) \vee \theta$, para $\xi = 1, 2, 3$:

$$(W \circ_L \mathbf{x}^1) \vee \theta = \left(\begin{bmatrix} 0.8 & 0.9 & 1 & 0.9 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0.6 & 0.8 & 0.8 & 0.7 & 1 \end{bmatrix} \circ_L \begin{bmatrix} 0.5 \\ 0.5 \\ 0.4 \\ 0.4 \\ 0.3 \end{bmatrix} \right) \vee \begin{bmatrix} 0.5 \\ 0.6 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.6 \\ 0.3 \end{bmatrix} = \mathbf{y}^1 \quad (4.45)$$

$$(W \circ_L \mathbf{x}^2) \vee \theta = \left(\left[\begin{array}{ccccc} 0.8 & 0.9 & 1 & 0.9 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0.6 & 0.8 & 0.8 & 0.7 & 1 \end{array} \right] \circ_L \left[\begin{array}{c} 0.1 \\ 0.3 \\ 0.3 \\ 0.4 \\ 0.4 \end{array} \right] \right) \vee \left[\begin{array}{c} 0.5 \\ 0.6 \\ 0.3 \end{array} \right] = \left[\begin{array}{c} 0.5 \\ 0.6 \\ 0.4 \end{array} \right] = \mathbf{y}^2 \quad (4.46)$$

$$(W \circ_L \mathbf{x}^3) \vee \theta = \left(\left[\begin{array}{ccccc} 0.8 & 0.9 & 1 & 0.9 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0.6 & 0.8 & 0.8 & 0.7 & 1 \end{array} \right] \circ_L \left[\begin{array}{c} 0.8 \\ 0.4 \\ 0.6 \\ 0.7 \\ 0.4 \end{array} \right] \right) \vee \left[\begin{array}{c} 0.5 \\ 0.6 \\ 0.3 \end{array} \right] = \left[\begin{array}{c} 0.6 \\ 0.8 \\ 0.4 \end{array} \right] = \mathbf{y}^3 \quad (4.47)$$

4.3.3 Mémória Associativa Fuzzy Implicativa Dual

Relembrando que o modelo de IFAM possui neurônios morfológicos max- T , onde T é uma t-norma contínua. O modelo de IFAM dual pode ser construída tomando neurônios duais com respeito a uma certa negação fuzzy N . A princípio, Valle e Sussner estabeleceram este modelo dual considerando somente a negação fuzzy padrão N_S [72]. Logo, a IFAM dual está associada a uma s-norma contínua S que é o operador dual a T com respeito a negação fuzzy padrão N_S . Portanto, a IFAM dual executa uma erosão em cada neurônio e pode ser descrita segundo a equação:

$$\mathbf{y} = (M \bullet_S \mathbf{x}) \wedge \vartheta \quad \text{e} \quad \vartheta = \bigvee_{\xi=1}^k \mathbf{y}^\xi. \quad (4.48)$$

onde a matriz dos pesos sinápticos $M \in [0, 1]^{m \times n}$ é dada pela Equação 4.40.

Toda afirmação envolvendo o modelo de IFAM produz uma correspondente afirmação dual envolvendo o modelo de IFAM dual. Portanto, na próxima seção serão discutidos apenas as

IFAMs autoassociativas e omitidos os resultados para o caso dual.

4.3.4 Memórias Implicativas Fuzzy Autoassociativas

Quando o conjunto das memórias fundamentais é da forma $\{(\mathbf{x}^\xi, \mathbf{x}^\xi) : \xi = 1, \dots, k\}$, a memória associativa fuzzy implicativa é denominada *memória implicativa fuzzy autoassociativa* (AFIM - Autoassociative Fuzzy Implicative Memory) [72]. Seja $X = [\mathbf{x}^1, \dots, \mathbf{x}^p] \in [0, 1]^{n \times p}$ a matriz cujas colunas são padrões originais. A matriz dos pesos sinápticos W e o vetor bias θ são definidos como:

$$W = X \circledast_T X^T \quad \text{e} \quad \theta = \bigwedge_{\xi=1}^k \mathbf{x}^\xi. \quad (4.49)$$

A AFIM pode ser convertida em um modelo dinâmico, tomando-se um padrão de entrada $\mathbf{x}(0) = \tilde{\mathbf{x}}$ e realiza-se a iteração:

$$\mathbf{x}(k+1) = (W \circ_T \mathbf{x}(k)) \vee \theta \quad \text{para } k = 0, 1, 2, \dots \quad (4.50)$$

Na verdade, tem-se que $\mathbf{x}(k+1) = \mathbf{x}(k)$ para todo $k \geq 1$ [72]. Dada a reflexividade de W ($I \leq W$) e a monotocidade da composição \max - T , o padrão de entrada é sempre menor ou igual ao padrão recordado, i.e, $\mathbf{x} \leq W \circ_T \mathbf{x}$ para todo $\mathbf{x} \in [0, 1]^n$. Por outro lado, como θ é dado pelo ínfimo dos padrões originais, concluí-se que todo padrão da forma $W \circ_T \mathbf{x}$ para algum $\mathbf{x} \in [0, 1]^n$ é um ponto fixo de W . Portanto, a AFIM exhibe *convergência em um único passo*. A proposição 4.3.1 caracteriza os pontos fixos e as saídas da AFIM [72]:

Proposição 4.3.1. *Considere o conjunto das memórias fundamentais $\{\mathbf{x}^1, \dots, \mathbf{x}^p\}$. Se $W = X \circledast_T X^t$ e $\theta = \bigwedge_{\xi=1}^p \mathbf{x}^\xi$ então para todo padrão de entrada $\mathbf{x} \in [0, 1]^n$, a saída $(W \circ_T \mathbf{x}) \vee \theta$ da AFIM é o supremo de \mathbf{x} no conjunto dos pontos fixos de W maior ou igual a θ , i.e, $(W \circ_T \mathbf{x}) \vee \theta$ é o menor ponto fixo \mathbf{y} de W tal que $\mathbf{y} \geq \mathbf{x}$ e $\mathbf{y} \geq \theta$. E mais, um padrão $\mathbf{y} \in [0, 1]^n$ é um ponto fixo de W se $\mathbf{y} = \mathbf{c}$ para algum vetor constante $\mathbf{c} = [c, c, \dots, c] \in [0, 1]^n$, ou se \mathbf{y} é da seguinte forma, para algum $L_l \subseteq \{1, \dots, p\}$ e para algum $k \in \mathbb{N}$*

$$\mathbf{y} = \bigvee_{l=1}^k \bigwedge_{\xi \in L_l} \mathbf{x}^\xi \quad (4.51)$$

A Proposição 4.3.1 revela que a região de atração dos padrão originais \mathbf{x}^ξ consiste somente dos padrões \mathbf{x} tal que $\mathbf{x} \leq \mathbf{x}^\xi$, i.e, \mathbf{x} é uma versão erodida do padrão \mathbf{x}^ξ para algum $\xi = 1, \dots, k$. Portanto, a AFIM exibe tolerância com respeito a padrões erosivos, além de possuir um grande número de pontos fixos, incluindo os padrões originais \mathbf{x}^ξ , onde $\xi = 1, \dots, p$ e muitas memórias espúrias [79]. Como não há restrições sobre o número de padrões, isto significa que pode-se armazenar o número de padrões desejados em uma AFIM.

Exemplo 4.3.2. *Considere a matriz*

$$X = \begin{bmatrix} 0.3 & 0.8 & 0.4 \\ 0.1 & 0.6 & 0.4 \\ 1 & 0.2 & 0.7 \end{bmatrix} \quad (4.52)$$

A matriz dos pesos sinápticos $W = X \circledast_L X^t$ de Lukasiewicz e o vetor bias $\theta = \bigwedge_{\xi=1}^3 \mathbf{x}^\xi$ são dados por:

$$W = \begin{bmatrix} 1 & 1 & 0.3 \\ 0.8 & 1 & 0.1 \\ 0.4 & 0.6 & 1 \end{bmatrix} \quad e \quad \theta = \begin{bmatrix} 0.3 & 0.1 & 0.2 \end{bmatrix}^t \quad (4.53)$$

Pode-se verificar que os padrões originais \mathbf{x}^ξ para $\xi = 1, 2, 3$ são pontos fixos de W :

$$(W \circledast_L \mathbf{x}^1) \vee \theta = \left(\begin{bmatrix} 1 & 1 & 0.3 \\ 0.8 & 1 & 0.1 \\ 0.4 & 0.6 & 1 \end{bmatrix} \circledast_L \begin{bmatrix} 0.3 \\ 0.1 \\ 1 \end{bmatrix} \right) \vee \begin{bmatrix} 0.3 \\ 0.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.1 \\ 1 \end{bmatrix} = \mathbf{x}^1 \quad (4.54)$$

$$(W \circ_L \mathbf{x}^2) \vee \theta = \left(\left(\begin{bmatrix} 1 & 1 & 0.3 \\ 0.8 & 1 & 0.1 \\ 0.4 & 0.6 & 1 \end{bmatrix} \circ_L \begin{bmatrix} 0.8 \\ 0.6 \\ 0.2 \end{bmatrix} \right) \vee \begin{bmatrix} 0.3 \\ 0.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.6 \\ 0.2 \end{bmatrix} = \mathbf{x}^2 \quad (4.55)$$

$$(W \circ_L \mathbf{x}^1) \vee \theta = \left(\left(\begin{bmatrix} 1 & 1 & 0.3 \\ 0.8 & 1 & 0.1 \\ 0.4 & 0.6 & 1 \end{bmatrix} \circ_L \begin{bmatrix} 0.4 \\ 0.4 \\ 0.7 \end{bmatrix} \right) \vee \begin{bmatrix} 0.3 \\ 0.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.4 \\ 0.7 \end{bmatrix} = \mathbf{x}^3 \quad (4.56)$$

Agora, dado um vetor constante arbitrário $\mathbf{c} \in [0, 1]^3$, tem-se que $\mathbf{c} = (W \circ_L \mathbf{c}) \vee \theta$:

$$(W \circ_L \mathbf{c}) \vee \theta = \left(\left(\begin{bmatrix} 1 & 1 & 0.3 \\ 0.8 & 1 & 0.1 \\ 0.4 & 0.6 & 1 \end{bmatrix} \circ_L \begin{bmatrix} 0.7 \\ 0.7 \\ 0.7 \end{bmatrix} \right) \vee \begin{bmatrix} 0.3 \\ 0.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.7 \\ 0.7 \end{bmatrix} = \mathbf{c} \quad (4.57)$$

Tomando um vetor arbitrário $\mathbf{x} \in [0, 1]^3$, por exemplo $\mathbf{x} = [0.9 \ 0.1 \ 0.5]^t$ e pela Proposição 4.3.1, tem-se que \mathbf{y} é o supremo de \mathbf{x} no conjunto dos pontos fixos de W maior do que θ :

$$\mathbf{y} = (W \circ_L \mathbf{x}) \vee \theta = \left(\left(\begin{bmatrix} 1 & 1 & 0.3 \\ 0.8 & 1 & 0.1 \\ 0.4 & 0.6 & 1 \end{bmatrix} \circ_L \begin{bmatrix} 0.9 \\ 0.1 \\ 0.5 \end{bmatrix} \right) \vee \begin{bmatrix} 0.3 \\ 0.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 0.7 \\ 0.5 \end{bmatrix} \geq \mathbf{x} \quad (4.58)$$

$$\mathbf{y} = \begin{bmatrix} 0.9 \\ 0.7 \\ 0.5 \end{bmatrix} \geq \begin{bmatrix} 0.3 \\ 0.1 \\ 0.2 \end{bmatrix} = \theta \quad (4.59)$$

e

$$(W \circ_L \mathbf{y}) \vee \theta = \left(\left(\begin{bmatrix} 1 & 1 & 0.3 \\ 0.8 & 1 & 0.1 \\ 0.4 & 0.6 & 1 \end{bmatrix} \circ_L \begin{bmatrix} 0.9 \\ 0.7 \\ 0.5 \end{bmatrix} \right) \vee \begin{bmatrix} 0.3 \\ 0.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 0.7 \\ 0.5 \end{bmatrix} = \mathbf{y} \quad (4.60)$$

Caso um vetor \mathbf{y} seja da forma $\mathbf{y} = \bigvee_{l=1}^p \bigwedge_{\xi \in L_l} x^\xi$ para algum $L_l \subseteq \{1, \dots, p\}$ e para algum $k \in \mathbb{N}$, então $\mathbf{y} = (W \circ_L \mathbf{y}) \vee \theta$. Tome os conjuntos de índices: $L_1 = \{1, 2\}$, $L_2 = \{2\}$ e $L_3 = \{1, 3\}$. Obtém-se,

$$\mathbf{y} = \bigvee_{l=1}^3 \bigwedge_{\xi \in L_l} x^\xi = \begin{bmatrix} 0,3 \\ 0,1 \\ 0,2 \end{bmatrix} \vee \begin{bmatrix} 0,8 \\ 0,6 \\ 0,2 \end{bmatrix} \vee \begin{bmatrix} 0,3 \\ 0,1 \\ 0,7 \end{bmatrix} = \begin{bmatrix} 0,8 \\ 0,6 \\ 0,7 \end{bmatrix} \quad (4.61)$$

e

$$(W \circ_L \mathbf{y}) \vee \theta = \left(\begin{bmatrix} 1 & 1 & 0,3 \\ 0,8 & 1 & 0,1 \\ 0,4 & 0,6 & 1 \end{bmatrix} \circ_L \begin{bmatrix} 0,8 \\ 0,6 \\ 0,7 \end{bmatrix} \right) \vee \begin{bmatrix} 0,3 \\ 0,1 \\ 0,2 \end{bmatrix} = \begin{bmatrix} 0,8 \\ 0,6 \\ 0,7 \end{bmatrix} = \mathbf{y} \quad (4.62)$$

Para concluir esta seção, as seguintes observações podem ser feitas em relação as memórias implicativas fuzzy autoassociativas [79]:

1. Capacidade de armazenagem ilimitada;
2. Os padrões originais são pontos fixos do modelo;
3. A saída permanece estável após repetidas aplicações da AFIM;
4. A memória exibe tolerância com respeito a padrões erosivos;
5. A AFIM não é adequada para padrões corrompidos com ruídos dilatativos;
6. A AFIM tem um grande número de memórias espúrias.

CAPÍTULO 5

Existência de Soluções nos Sistemas de FREs em Termos das IFAMs

5.1 Introdução

O objetivo deste capítulo é investigar a existência de soluções nos sistemas de Equações Relacionais Fuzzy (FREs - Fuzzy Relational Equations) em termos das Memórias Associativas Fuzzy Implicativas (IFAMs - Implicative Fuzzy Associative Memories). Acreditamos que um estudo envolvendo a questão da recordação perfeita através dos modelos de IFAMs poderá contribuir na existência de soluções para os sistemas de FREs.

5.2 Uma Comparação entre FREs e as Composições $\max-C$, $\min-D$ e $\min-J$

Recordando que os três principais produtos matriciais fuzzy usados para descrever diversos modelos FMAMs, combinam máximos e mínimos com as operações fuzzy de conjunção C , disjunção D e implicação reversa J . Comparando as equações relacionais fuzzy 2.34, 2.35 e 2.36 com as composições $\max-T$, $\min-S$ e $\min-J$ definidas nas Equações 2.30 e 2.31, tem-se as seguintes igualdades:

$$R \cdot \mathbf{x} = R^t \circ \mathbf{x} \quad R\Delta\mathbf{x} = R^t \bullet \mathbf{x} \quad R\phi\mathbf{x} = R^t \otimes \mathbf{x} \quad (5.1)$$

Dada a igualdade $\mathbf{y} = R \cdot \mathbf{x}^\xi = R^t \circ \mathbf{x}^\xi$ e como toda afirmação envolvendo o modelo de IFAM, produz uma correspondente afirmação dual envolvendo o modelo de IFAM dual. Nossos estudos serão direcionados as equações relacionais fuzzy do tipo $\mathbf{y} = R \cdot \mathbf{x}$ uma vez que pode-se inferir resultados similares para a FRE $\mathbf{y} = R\Delta\mathbf{x} = R^t \bullet \mathbf{x}$ dada a esta relação de dualidade. Portanto, deseja-se uma relação fuzzy \mathcal{R} que satisfaça o seguinte sistema de FRES para todo $\xi = 1, \dots, k$:

$$\mathbf{y}^\xi = R \cdot \mathbf{x}^\xi \quad (5.2)$$

Associando a Proposição 4.2.2 ao sistema de FRES descrito pela Equação 5.2, obtém-se o seguinte resultado: Se existe uma relação fuzzy R_0 tal que $\mathbf{y}^\xi = R_0 \cdot \mathbf{x}^\xi$ para todo $\xi = 1, \dots, k$, então $R_0 \leq R = (Y \otimes X^t)^t$ e $\mathbf{y}^\xi = R \circ \mathbf{x}^\xi$ para todo $\xi = 1, \dots, k$. Em outras palavras, tem-se o seguinte teorema [45]:

Teorema 5.2.1. *Se o sistema de FRES definido por 5.2 tem uma solução, então \mathcal{R} definida por $r_{ij} = \bigwedge_{\xi=1}^k J(\mathbf{y}_i^\xi, \mathbf{x}_j^\xi)$ representa a maior solução.*

Similarmente, pode-se aplicar o resultado dual dada pela Proposição 4.2.4 referente ao modelo de IFAM dual. Desta proposição, surge o seguinte teorema para as FRES do tipo $\mathbf{y} = \mathcal{R}\Delta\mathbf{x}$.

Teorema 5.2.2. *Se o sistema de FRES definido por $\mathbf{y}^\xi = \mathcal{R}\Delta\mathbf{x}^\xi$ para $\xi = 1, \dots, k$ tem uma solução, então \mathcal{R} definida por $r_{ij} = \bigvee_{\xi=1}^k J_S(\mathbf{y}_i^\xi, \mathbf{x}_j^\xi)$ representa a menor solução.*

Finalmente, tem-se o teorema referente as soluções do sistema de FRES $\mathbf{y}^\xi = \mathcal{R}\phi\mathbf{x}^\xi$:

Teorema 5.2.3. *Se o sistema de FRES $\mathbf{y}^\xi = \mathcal{R}\phi\mathbf{x}^\xi$, para $\xi = 1, \dots, k$ tem uma solução, então a menor solução é dada por $r_{ij} = \bigvee_{\xi=1}^k T(\mathbf{x}_i^\xi, \mathbf{y}_j^\xi)$.*

5.3 Relação entre FREs e IFAMs

A seguir, serão caracterizados os padrões recordados por uma IFAM . O Teorema 5.2.1 diz que a saída é menor ou igual ao padrão desejado \mathbf{y}^ξ quando apresentado um padrão não distorcido \mathbf{x}^ξ para algum $\xi \in \{1, \dots, k\}$ [72].

Teorema 5.3.1. *Se $\mathcal{W} : [0, 1]^n \rightarrow [0, 1]^m$ é uma IFAM então $\mathcal{W}(\mathbf{x}^\xi) \leq \mathbf{y}^\xi$ para todo $\xi = 1, \dots, k$*

O Teorema dual para um modelo de IFAM dual é dado a seguir:

Teorema 5.3.2. *Se $\mathcal{M} : [0, 1]^n \rightarrow [0, 1]^m$ é uma IFAM dual então $\mathcal{M}(\mathbf{x}^\xi) \geq \mathbf{y}^\xi$ para todo $\xi = 1, \dots, k$*

Considere que o par (T, J) forme uma adjunção. O seguinte teorema caracteriza algebricamente os padrões recordados por uma IFAM [79].

Teorema 5.3.3. *Sejam $X = [\mathbf{x}^1, \dots, \mathbf{x}^k] \in [0, 1]^{n \times k}$ e $Y = [\mathbf{y}^1, \dots, \mathbf{y}^k] \in [0, 1]^{m \times k}$. Para um padrão de entrada arbitrário $\mathbf{x} \in [0, 1]^n$, o padrão de saída $\mathbf{y} \in [0, 1]^m$ da IFAM pode ser escrito da seguinte forma:*

$$\mathbf{y} = \bigvee_{j=1}^n \bigwedge_{\xi=1}^k \mathcal{T}(\mathcal{J}(\mathbf{y}^\xi, x_j^\xi), x_j) \quad (5.3)$$

onde $[\mathcal{J}(\mathbf{y}, a)]_i = J(y_i, a)$ e $[\mathcal{T}(\mathbf{x}, b)]_i = T(x_i, b) \forall \mathbf{x}, \mathbf{y} \in [0, 1]^m$ e $a, b \in [0, 1]$.

Para os próximos teoremas e corolários, denota-se \mathcal{N} como o conjunto $\{1, \dots, n\}$ e \mathcal{K} como o conjunto $\{1, \dots, k\}$. O corolário abaixo descreve outro forma que os padrões de saída da IFAM podem assumir [79]:

Corolário 5.3.1. *Sejam $X \in [0, 1]^{n \times k}$ e $Y \in [0, 1]^{m \times k}$. Para um padrão de entrada arbitrário $\mathbf{x} \in [0, 1]^n$, o padrão de saída $\mathbf{y} \in [0, 1]^m$ da IFAM pode ser escrito algebricamente por:*

$$\mathbf{y} = \bigwedge_{F \in \mathcal{C}} \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{y}^{F(j)}, x_j^{F(j)}), x_j) \quad (5.4)$$

onde \mathcal{C} é o conjunto de todas as funções de $F : \mathcal{N} \rightarrow \mathcal{K}$.

O seguinte teorema caracteriza algebricamente os padrões recordados pela IFAM dual.

Teorema 5.3.4. *Sejam $X = [\mathbf{x}^1, \dots, \mathbf{x}^k] \in [0, 1]^{n \times k}$ e $Y = [\mathbf{y}^1, \dots, \mathbf{y}^k] \in [0, 1]^{m \times k}$. Para um padrão de entrada arbitrário $\mathbf{x} \in [0, 1]^n$, o padrão de saída $\mathbf{y} \in [0, 1]^m$ da IFAM dual pode ser escrito da seguinte forma:*

$$\mathbf{y} = \bigwedge_{j=1}^n \bigvee_{\xi=1}^k \mathcal{S}(\overline{\mathcal{J}}(\mathbf{y}^\xi, x_j^\xi), x_j) \quad (5.5)$$

onde $[\overline{\mathcal{J}}(\mathbf{y}, a)]_i = J_S(y_i, a)$ e $[\mathcal{S}(\mathbf{x}, b)]_i = S(x_i, b) \forall \mathbf{x}, \mathbf{y} \in [0, 1]^m$ e $a, b \in [0, 1]$.

O Corolário 5.3.2 descreve outra forma que os padrões de saída da IFAM dual podem assumir:

Corolário 5.3.2. *Sejam $X \in [0, 1]^{n \times k}$ e $Y \in [0, 1]^{m \times k}$. Para um padrão de entrada arbitrário $\mathbf{x} \in [0, 1]^n$, o padrão de saída $\mathbf{y} \in [0, 1]^m$ da IFAM dual pode ser escrito algebricamente por:*

$$\mathbf{y} = \bigvee_{F \in \mathcal{C}} \bigwedge_{j=1}^n \mathcal{S}(\overline{\mathcal{J}}(\mathbf{y}^{F(j)}, x_j^{F(j)}), x_j) \quad (5.6)$$

onde \mathcal{C} é o conjunto de todas as funções de $F : \mathcal{N} \rightarrow \mathcal{K}$.

O Teorema 5.3.3, revela que o padrão recordado \mathbf{y} pela IFAM representa uma polinômio reticulado em versões transformadas dos padrões $\mathbf{y}^1, \dots, \mathbf{y}^k$ [79]. O Teorema 5.3.3 juntamente com o Teorema 5.3.5 caracteriza completamente as saídas da IFAM e sua demonstração pode ser encontrado em [68].

Teorema 5.3.5. *Para $X \in [0, 1]^{n \times k}$, $Y \in [0, 1]^{m \times k}$, e seja $\mathbf{x} \in [0, 1]^n$ tal que $(X \otimes X^t) \circ \mathbf{x} = \mathbf{x}^\gamma$. Tem-se $(Y \otimes X^t) \circ \mathbf{x} = \mathbf{y}^\gamma$ se, e somente se, a seguinte implicação é válida $\forall F : \mathcal{N} \rightarrow \mathcal{K}$, $\forall a_j, b_j \in [0, 1]$ e $j = 1, \dots, n$.*

$$\mathbf{x}^\gamma \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{x}^{F(j)}, a_j), b_j) \Rightarrow \mathbf{y}^\gamma \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{y}^{F(j)}, a_j), b_j) \quad (5.7)$$

Um corolário imediato do Teorema 5.3.5 para um padrão de entrada arbitrário $\mathbf{x}^\gamma \in [0, 1]^n$ é dado por:

Corolário 5.3.3. *Para $X \in [0, 1]^{n \times k}$, $Y \in [0, 1]^{m \times k}$, e seja um padrão de entrada arbitrário $\mathbf{x}^\gamma \in [0, 1]^n$. Tem-se $(Y \circledast X^t) \circ \mathbf{x}^\gamma = \mathbf{y}^\gamma$ se, e somente se, a seguinte implicação é válida $\forall F : \mathcal{N} \rightarrow \mathcal{K}$, $\forall a_j, b_j \in [0, 1]$ e $j = 1, \dots, n$.*

$$\mathbf{x}^\gamma \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{x}^{F(j)}, a_j), b_j) \Rightarrow \mathbf{y}^\gamma \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{y}^{F(j)}, a_j), b_j) \quad (5.8)$$

O Corolário 5.3.3 revela um importante resultado, relacionando a recordação perfeita de uma IFAM com a existência de uma relação fuzzy \mathcal{R} que satisfaça o sistema de FRES do tipo $\mathbf{y}^\xi = W \cdot \mathbf{x}^\xi$.

Teorema 5.3.6. *O sistema de FRES $\mathbf{y}^\xi = \mathcal{R} \cdot \mathbf{x}^\xi$ para $\xi = 1, \dots, k$, tem solução se, e somente se, a seguinte implicação é satisfeita $\forall F : \mathcal{N} \rightarrow \mathcal{K}$, $\forall a_j, b_j \in [0, 1]$ e $j = 1, \dots, n$.*

$$\mathbf{x}^\xi \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{x}^{F(j)}, a_j), b_j) \Rightarrow \mathbf{y}^\xi \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{y}^{F(j)}, a_j^{F(j)}), b_j^{F(j)}) \quad (5.9)$$

E mais, $\mathcal{R} = (Y \circledast X^t)^t$ é a maior solução.

Agora, será enunciado os Teoremas e Corolários para a IFAM dual, relacionando a recordação perfeita deste modelo com a existência de soluções para o sistema de FRES do tipo $\mathbf{y}^\xi = \mathcal{R} \Delta \mathbf{x}^\xi$.

Teorema 5.3.7. *Para $X \in [0, 1]^{n \times k}$, $Y \in [0, 1]^{m \times k}$, e seja $\mathbf{x} \in [0, 1]^n$ tal que $(X \bar{\circledast} X^t) \bullet \mathbf{x} = \mathbf{x}^\gamma$. Temos $(Y \bar{\circledast} X^t) \bullet \mathbf{x} = \mathbf{y}^\gamma$ se, e somente se, a seguinte implicação é válida $\forall F : \mathcal{N} \rightarrow \mathcal{K}$, $\forall a_j, b_j \in [0, 1]$ e $j = 1, \dots, n$.*

$$\mathbf{x}^\gamma \geq \bigwedge_{j=1}^n \mathcal{S}(\bar{\mathcal{J}}(\mathbf{x}^{F(j)}, a_j), b_j) \Rightarrow \mathbf{y}^\gamma \geq \bigwedge_{j=1}^n \mathcal{S}(\bar{\mathcal{J}}(\mathbf{y}^{F(j)}, a_j), b_j) \quad (5.10)$$

Um corolário imediato do Teorema 5.3.7 para um padrão de entrada arbitrário $\mathbf{x}^\gamma \in [0, 1]^n$ é dado por:

Corolário 5.3.4. *Para $X \in [0, 1]^{n \times k}$, $Y \in [0, 1]^{m \times k}$, e seja um padrão de entrada arbitrário $\mathbf{x}^\gamma \in [0, 1]^n$. Tem-se $(Y \bar{\otimes} X^t) \bullet \mathbf{x}^\gamma = \mathbf{y}^\gamma$ se, e somente se, a seguinte implicação é válida $\forall F : \mathcal{N} \rightarrow \mathcal{K}$, $\forall a_j, b_j \in [0, 1]$ e $j = 1, \dots, n$.*

$$\mathbf{x}^\gamma \geq \bigwedge_{j=1}^n \mathcal{S}(\bar{\mathcal{J}}(\mathbf{x}^{F(j)}, a_j), b_j) \Rightarrow \mathbf{y}^\gamma \geq \bigwedge_{j=1}^n \mathcal{S}(\bar{\mathcal{J}}(\mathbf{y}^{F(j)}, a_j), b_j) \quad (5.11)$$

O Corolário 5.3.4 relaciona a recordação perfeita de uma IFAM dual com a existência de uma relação fuzzy \mathcal{R} que satisfaça o sistema de FRES $\mathbf{y}^\xi = \mathcal{R} \Delta \mathbf{x}^\xi$.

Teorema 5.3.8. *O sistema de FRES $\mathbf{y}^\xi = \mathcal{R} \Delta \mathbf{x}^\xi$ para $\xi = 1, \dots, k$ tem solução se, e somente se, a seguinte implicação é satisfeita $\forall F : \mathcal{N} \rightarrow \mathcal{K}$, $\forall a_j, b_j \in [0, 1]$ e $j = 1, \dots, n$.*

$$\mathbf{x}^\xi \geq \bigwedge_{j=1}^n \mathcal{S}(\bar{\mathcal{J}}(\mathbf{x}^{F(j)}, a_j), b_j) \Rightarrow \mathbf{y}^\xi \geq \bigwedge_{j=1}^n \mathcal{S}(\bar{\mathcal{J}}(\mathbf{y}^{F(j)}, a_j), b_j) \quad (5.12)$$

E mais, $\mathcal{R} = (Y \bar{\otimes} X^t)^t$ é a menor solução do sistema FRES.

Finalizamos este capítulo descrevendo os principais resultados obtidos, relacionando o recordação perfeita de uma IFAM com a existência de soluções para um sistema de FRES do tipo $\mathbf{y}^\xi = \mathcal{R} \cdot \mathbf{x}^\xi$. O mesmo vale para as IFAMs duais dado a relação de dualidade entre a IFAM e a IFAM dual. Se a seguinte implicação é válida $\forall F : \mathcal{N} \rightarrow \mathcal{K}$, $\forall a_j, b_j \in [0, 1]$ e $j = 1, \dots, n$, então

$$\mathbf{x}^\xi \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{x}^{F(j)}, a_j), b_j) \Rightarrow \mathbf{y}^\xi \leq \bigvee_{j=1}^n \mathcal{T}(\mathcal{J}(\mathbf{y}^{F(j)}, a_j), b_j) \quad \forall \xi = 1, \dots, k \quad (5.13)$$

$$\Leftrightarrow \exists \mathcal{R} \text{ tal que } \mathcal{R} \cdot \mathbf{x}^\xi = \mathbf{y}^\xi \quad \forall \xi = 1, \dots, k \quad (5.14)$$

$$\Leftrightarrow \exists \mathcal{R}^t \text{ tal que } \mathcal{R}^t \circ \mathbf{x}^\xi = \mathbf{y}^\xi \quad \forall \xi = 1, \dots, k \quad (5.15)$$

$$\Leftrightarrow \exists W \text{ tal que } W \circ \mathbf{x}^\xi = \mathbf{y}^\xi \quad \forall \xi = 1, \dots, k \quad (5.16)$$

$$\Leftrightarrow \exists (Y \circledast X^t)^t \circ \mathbf{x}^\xi = \mathbf{y}^\xi \quad \forall \xi = 1, \dots, k \quad (5.17)$$

CAPÍTULO 6

Determinação do Melhor Modelo de IFAM de Yager para uma Aplicação

6.1 Introdução

O fato de uma Memória Associativa Fuzzy Implicativa (IFAM - Implicative Fuzzy Associative Memory) ser caracterizada por uma certa t-norma contínua [72], surge o problema de escolher o melhor modelo de IFAM para uma determinada aplicação. Uma solução para esta questão é formular um problema de otimização no processo de recordação, com o objetivo de minimizar o erro entre o padrão desejado \mathbf{y} e o padrão recordado $\tilde{\mathbf{y}}$ pela IFAM. Isso é possível adotando-se a IFAM de Yager, uma vez que a t-norma parametrizada de Yager abrange outras t-normas [14]. Portanto, o problema de otimização é caracterizado por encontrar a t-norma de Yager que produza o menor erro no processo de recordação.

6.2 Formulação do Problema

Considere os padrões de entrada $\mathbf{x}^\xi \in [0, 1]^n$ e os padrões de saída $\mathbf{y}^\xi \in [0, 1]^m$ para $\xi = 1, 2, \dots, p$. Neste momento, interpretados como conjuntos fuzzy definidos sobre um universo de discurso finito com n e m elementos, respectivamente. Seja X , a matriz cujas colunas são formadas pelos padrões de entrada, ou seja, $X = [\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^p] \in [0, 1]^{n \times p}$ e a matriz Y com

colunas formadas pelos padrões de saída desejados, ou seja, $Y = [\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^p] \in [0, 1]^{m \times p}$.

Deseja-se determinar uma memória associativa fuzzy tal que:

$$\mathbf{y}^\xi = (W \circ_t \mathbf{x}^\xi) \vee \theta \quad (6.1)$$

onde a matriz dos pesos sinápticos W e o vetor limiar θ são dados por:

$$W = Y \otimes_Y X^t \quad \text{e} \quad \theta = \bigwedge_{\xi=1}^p \mathbf{y}^\xi \quad (6.2)$$

Neste capítulo, somente será adota a t-norma de Yager e por isso o subscrito “Y” no produto $Y \otimes_Y X^t$. Algumas vezes somente será mencionado IFAM em vez de IFAM de Yager já que não há perigo de confusão. Para a t-norma de Yager, será adotado o símbolo $T_Y^d = T(x, y)$ para todo $x, y \in [0, 1]$ e $d \in (0, +\infty)$. Para a R-implicação de Yager, será adotado o símbolo I_Y . Os elementos da matriz dos pesos W são descritos pela seguinte equação:

$$w_{ji} = \bigwedge_{\xi=1}^p I_Y(x_i^\xi, y_j^\xi) \quad \text{para } i = 1, \dots, n \text{ e } j = 1, \dots, m. \quad (6.3)$$

ou equivalentemente,

$$w_{ji} = \bigwedge_{\xi=1}^p \bigvee \left\{ z \in [0, 1] : T_Y^d(x_i^\xi, z) \leq y_j^\xi \right\} \quad (6.4)$$

Expandindo a t-norma de Yager T_Y^d , tem-se:

$$w_{ji} = \bigwedge_{\xi=1}^p \bigvee \left\{ z \in [0, 1] : 1 - \min \left\{ 1, \left\{ (1 - x_i^\xi)^d + (1 - z)^d \right\}^{1/d} \right\} \leq y_j^\xi \right\}, \quad d > 0 \quad (6.5)$$

Da mesma forma, escrevendo a função de pertinência da expressão 6.1, para $\xi = 1, \dots, p$ obtém-se:

$$y_j^\xi = \bigvee_{i=1}^n T_Y^d(w_{ji}, x_i^\xi) \vee \theta_j, \quad \text{para } j = 1, \dots, m \quad (6.6)$$

$$y_j^\xi = \bigvee_{i=1}^n \left(1 - \min \left\{ 1, \left\{ (1 - w_{ji})^d + (1 - x_i^\xi)^d \right\}^{1/d} \right\} \right) \vee \theta_j, \quad \text{para } d > 0 \quad (6.7)$$

Como o parâmetro d na t-norma de Yager é livre, e a IFAM W e os padrões recordados \tilde{y}^ξ dependem deste parâmetro, pode-se formular um problema de otimização para encontrar d que minimize a diferença entre os padrões desejados y^ξ e os padrões recordados \tilde{y}^ξ para $\xi = 1, \dots, p$. Assim, a partir deste momento, as seguintes notações para W e \tilde{y}^ξ serão dadas por W^d e $\tilde{y}^{\xi,d}$, respectivamente. O problema baseia-se em resolver 6.8:

$$\begin{cases} \text{minimize} & \left\| \tilde{Y}^d - Y \right\|_F \\ \text{sujeito a} & d \geq 0 \end{cases} \quad (6.8)$$

ou

$$\begin{cases} \text{minimize} & \left\| \left[\bigvee_{i=1}^n \left\{ 1 - \min \left\{ 1, \left\{ (1 - w_{ji}^d)^d + (1 - x_i^\xi)^d \right\}^{1/d} \right\} \right\} \vee \theta_j \right] - y_j^\xi \right\|_F \\ \text{sujeito a} & d \geq 0 \end{cases} \quad (6.9)$$

Mas Sussner e Valle [72], mostraram que

$$\tilde{\mathbf{y}}^\xi = (W \circ_T \mathbf{x}^\xi) \vee \theta \leq \mathbf{y}^\xi \quad \text{para } \xi = 1, \dots, p. \quad (6.10)$$

Logo, os padrões recordados $\tilde{\mathbf{y}}^\xi$ são limitados superiormente pelos padrões \mathbf{y}^ξ . Isto significa, que resolver o problema de otimização descrito por 6.8, é equivalente a resolver o problema de otimização descrito por 6.11:

$$\begin{cases} \text{maximize} & \left\| \tilde{Y}^d \right\|_2 \\ \text{sujeito a} & d \geq 0 \end{cases} \quad (6.11)$$

6.3 Aplicações de IFAMs de Yager em Séries Temporais

6.3.1 Previsão da Mão-de-obra em Indústrias Metalúrgicas

Considere o problema apresentado em [12] que consiste na previsão da mão-de-obra requerida nas indústrias metalúrgicas do estado Bengal do Oeste na Índia. As memórias associativas fuzzy podem apresentar sistemas de regras fuzzy onde armazenamos um conjunto de regras da forma SE-ENTÃO. Nesse problema em particular, com base nos valores passados, é definido um conjunto de regras nebulosas da forma "Se a mão-de-obra requerida no ano n é grande, então a mão-de-obra requerida no $n + 1$ é muito grande". Desse modo, obtém-se um conjunto de pares de entrada e saída, onde as entradas correspondem aos antecedentes e as saídas correspondem aos respectivos consequentes. Os pares de entrada e saída usados nesse exemplo são apresentados na Tabela 6.1.

ξ	\mathbf{x}^ξ	\mathbf{y}^ξ
1	$\begin{bmatrix} 1.0 & 0.5 & 0 & 0 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 1.0 & 0.5 & 0 & 0 & 0 \end{bmatrix}^T$
2	$\begin{bmatrix} 0.5 & 1.0 & 0.5 & 0 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 0.5 & 1.0 & 0.5 & 0 & 0 \end{bmatrix}^T$
3	$\begin{bmatrix} 0.5 & 1.0 & 0.5 & 0 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0.5 & 1.0 & 0.5 & 0 \end{bmatrix}^T$
4	$\begin{bmatrix} 0 & 0.5 & 1.0 & 0.5 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 0.5 & 1.0 & 0.5 & 0 & 0 \end{bmatrix}^T$
5	$\begin{bmatrix} 0 & 0.5 & 1.0 & 0.5 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0.5 & 1.0 & 0.5 & 0 \end{bmatrix}^T$
6	$\begin{bmatrix} 0 & 0.5 & 1.0 & 0.5 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0 & 0.5 & 1.0 & 0.5 \end{bmatrix}^T$
7	$\begin{bmatrix} 0 & 0 & 0.5 & 1.0 & 0.5 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0 & 0.5 & 1.0 & 0.5 \end{bmatrix}^T$
8	$\begin{bmatrix} 0 & 0 & 0.5 & 1.0 & 0.5 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0 & 0 & 0.5 & 1.0 \end{bmatrix}^T$
9	$\begin{bmatrix} 0 & 0 & 0 & 0.5 & 1.0 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0 & 0 & 0.5 & 1.0 \end{bmatrix}^T$

Tabela 6.1: Conjunto dos pares de entrada e saída usado na aplicação de previsão

A Tabela 6.2 apresenta a mão-de-obra requerida em cada ano e sua respectiva fuzzificação [12]:

Ano	Mão-de-obra	Fuzzificação
1983-84	1259	$\begin{bmatrix} 1.0 & 0.6 & 0 & 0 & 0 \end{bmatrix}^T$
1984-85	1370	$\begin{bmatrix} 0.4 & 1.0 & 0.6 & 0 & 0 \end{bmatrix}^T$
1985-86	1318	$\begin{bmatrix} 0.8 & 1.0 & 0.2 & 0 & 0 \end{bmatrix}^T$
1986-87	1364	$\begin{bmatrix} 0.4 & 1.0 & 0.6 & 0 & 0 \end{bmatrix}^T$
1987-88	1434	$\begin{bmatrix} 0 & 0.6 & 1.0 & 0.4 & 0 \end{bmatrix}^T$
1988-89	1361	$\begin{bmatrix} 0.4 & 1.0 & 0.6 & 0 & 0 \end{bmatrix}^T$
1989-90	1406	$\begin{bmatrix} 0 & 0.8 & 1.0 & 0.2 & 0 \end{bmatrix}^T$
1990-91	1433	$\begin{bmatrix} 0 & 0.6 & 1.0 & 0.4 & 0 \end{bmatrix}^T$
1991-92	1516	$\begin{bmatrix} 0 & 0 & 0.8 & 1.0 & 0.2 \end{bmatrix}^T$
1992-93	1595	$\begin{bmatrix} 0 & 0 & 0.2 & 1.0 & 0.8 \end{bmatrix}^T$
1993-94	1692	$\begin{bmatrix} 0 & 0 & 0 & 0.4 & 1.0 \end{bmatrix}^T$
1994-95	1658	$\begin{bmatrix} 0 & 0 & 0 & 0.4 & 1.0 \end{bmatrix}^T$
1995-96	1660	$\begin{bmatrix} 0 & 0 & 0 & 0.2 & 1.0 \end{bmatrix}^T$

Tabela 6.2: Mão-de-obra requerida e sua respectiva fuzzificação

Assim, a estimativa da mão-de-obra requerida no ano $n + 1$ é obtida “defuzzificando” o padrão recordado. Nesta aplicação, foram empregados o método do centro de massa e a média do máximo, através da rotina DEFUZZ da Toolbox de Conjuntos Fuzzy do MATLAB com a opção CENTROID e MOM, respectivamente. Logo, a mão-de-obra requerida no ano $n + 1$ é obtida pela sintaxe $defuzz(\tilde{a}_{n+1})$ onde \tilde{a}_{n+1} , dada pela Equação 6.12:

$$\tilde{a}_{n+1} = (W \circ_t a_n) \vee \theta \tag{6.12}$$

e a_n é o padrão que corresponde a mão-de-obra requerida no ano n .

SEÇÃO 6.3 APLICAÇÕES DE IFAMS DE YAGER EM SÉRIES TEMPORAIS 62

Denota-se a saída defuzzificada por s_ξ^d , ou seja, $s_\xi^d = defuzz(\tilde{y}^\xi)$ para $\xi = 1, \dots, p$, onde \tilde{y}^ξ é dado pela Equação 6.1. Se o vetor \mathbf{s}^d denota o vetor cujos elementos são a defuzzificação dos padrões \tilde{y}^ξ , i.e, $\mathbf{s}^d = [s_1^d, s_2^d, \dots, s_p^d]$ e \mathbf{s} o vetor cujos elementos são a mão-de-obra real requerida, então o erro percentual médio normalizado (EPMN) é dado pela Equação 6.13:

$$EPMN = \frac{100}{p} \times \left(\sum_{n=1}^N \frac{|s_\xi - s_\xi^d|}{s_\xi} \right) \quad (6.13)$$

A FAM max-min em [12] adotando como método de defuzzificação a média do máximo, gerou um EPMN de 2.669% superando os modelos estatísticos ARIMA1 e ARIMA2, como também a média aritmética (MA) e média geométrica (MG). Os resultados são apresentados na Tabela 6.3:

Método	MA	MG	ARIMA1	ARIMA2	FAM max-min
EPMN	5.65%	4.24%	9.79%	5.48%	2.669%

Tabela 6.3: EPMN produzidos pelos modelos da média aritmética, média geométrica, ARIMA1, ARIMA2 e FAM max-min para o problema de previsão de mão-de-obra.

Observando que a memória associativa fuzzy max-min gerou o menor EPMN, Sussner e Valle aplicaram este mesmo problema de previsão de mão-de-obra para outros modelos de FAMs [72]. Note que, vários modelos apresentaram os mesmos resultados. De fato, os valores estimados pela FAM max-min, FAM max-prod e GFAM de Lukasiewicz foram iguais. Além disso, as previsões fornecidas pelas IFAMs adjuntas de Gödel e Goguen, as IFAMs duais adjuntas de Gödel e Goguen, as FLBAMs e a IFAM dual de Lukasiewicz, e a IFAM de Gödel, FAM de Junbo e FAM max-min com limiar também coincidiram. Finalmente, a Figura 6.1 apresenta a mão-de-obra estimada nos anos 1984 até 1995. O valor real é comparado com os valores estimados por alguns dos métodos da Tabela 6.4. Observe que a IFAM de Lukasiewicz apresentou o melhor resultado para esse problema de previsão com EPMN de 2.29%, superando os demais modelos de FAMs. A implementação em MATLAB dessas FAMs para este problema pode ser encontrado em [77].

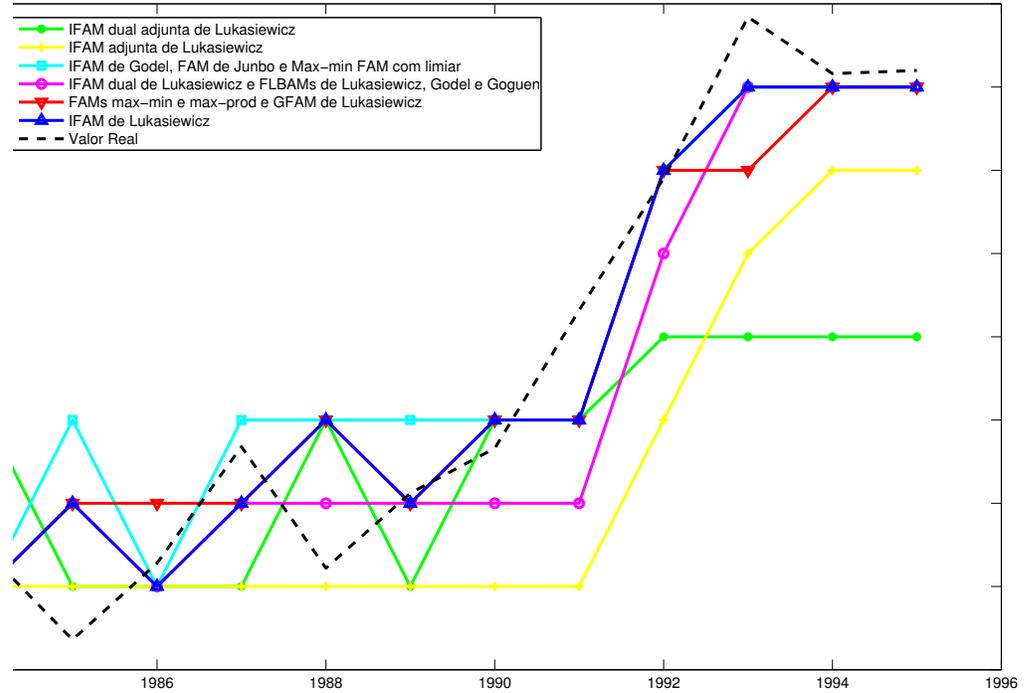


Figura 6.1: Comparação da previsão da mão-de-obra estimada por diferentes modelos de memória associativa fuzzy.

Observando que a IFAM de Lukasiewicz gerou o menor EPMN de 2.29%, isso nos motivou a formular um problema de otimização associado a IFAM de Yager. O objetivo é encontrar o parâmetro d na t-norma de Yager que minimize a diferença entre o mão-de-obra real s_ξ e a mão-de-obra estimada pela IFAM de Yager, defuzzificando o padrão recordado $s_\xi^d = defuzz(\tilde{y}^\xi)$. De fato, a IFAM de Yager *sempre* produzirá resultados iguais ou melhores que a IFAM de Lukasiewicz, pois basta tomar o parâmetro $d = 1$ na t-norma de Yager e obtém-se:

SEÇÃO 6.3 APLICAÇÕES DE IFAMS DE YAGER EM SÉRIES TEMPORAIS 64

Método	EPMN	Método	EPMN
FAM max-min:	2.67%	FLBAM de Lukasiewicz:	2.56%
fAM max-produto:	2.67%	FLBAM de Gödel:	2.56%
GFAM de Lukasiewicz:	2.67%	FLBAM de Goguen	2.56%
FAM Junbo:	2.73%	IFAM de Yager:	2.29%
IFAM de Lukasiewicz:	2.29%	IFAM dual de Lukasiewicz	2.56%
IFAM de Gödel:	2.73%	IFAM dual de Gödel:	8.08%
IFAM de Goguen:	2.99%	IFAM dual de Goguen:	8.08%
IFAM Adjunta de Lukasiewicz	4.74%	IFAM dual adjunta de Lukasiewicz	5.64%
IFAM Adjunta de Gödel:	6.89%	IFAM dual adjunta de Godel	7.26%
IFAM Adjunta de Goguen:	6.89%	IFAM dual adjunta de Goguen	7.26%

Tabela 6.4: Erros percentuais médios normalizados produzidos pelos modelos de FAMs na previsão de mão-de-obra.

$$\begin{aligned}
 T_Y^d(a, b) &= 1 - \min \left\{ 1, \left\{ (1-a)^d + (1-b)^d \right\}^{1/d} \right\} = 1 - \min \{ 1, \{(1-a) + (1-b)\} \} \\
 &= 1 - \min \{ 1, (2-a-b) \} = 1 - [1 - \max \{ 1-1, 1-(2-a-b) \}] \\
 &= \max \{ 0, a+b-1 \}
 \end{aligned} \tag{6.14}$$

ou seja, para $d = 1$ a t-norma de Yager é justamente a t-norma de Lukasiewicz. Então, o problema de otimização resume-se em:

$$\begin{cases} \text{minimize } \|\mathbf{s}^d - \mathbf{s}\|_2 \\ \text{sujeito a } d \geq 0 \end{cases} \tag{6.15}$$

Pelo gráfico da Figura 6.2, fica claro que neste caso, existe infinitos minimizadores no intervalo $(0, 1]$ para função $\|\mathbf{s}^d - \mathbf{s}\|_2$ utilizando-se como método de defuzzificação a média do máximo. Portanto, para qualquer $d \in (0, 1]$, o erro percentual médio normalizado gerado pela IFAM de Yager é também de 2.29%.

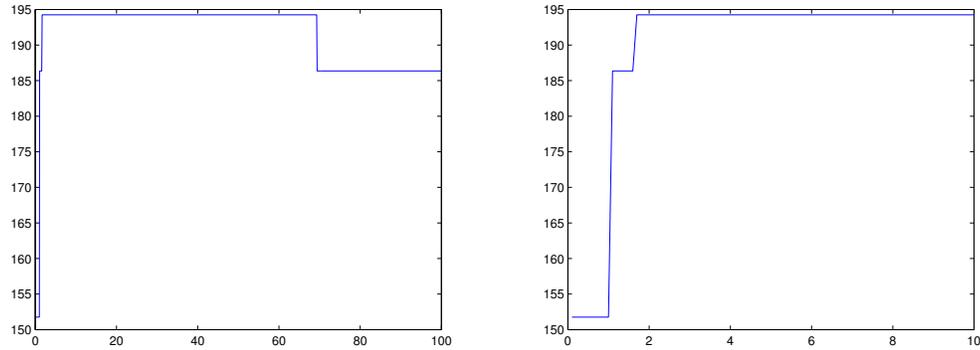


Figura 6.2: Gráfico da função $\|s^d - s\|_F$ adotando como método de defuzzificação a média do máximo. No gráfico da esquerda, tem-se $d \in (0, 100]$ e no gráfico da direita $d \in (0, 10]$

A implementação do algoritmo foi feita no software MATLAB, utilizando as funções da Toolbox de Otimização. Para resolver o problema descrito em 6.15, recorreu-se a função FMINBND que utiliza uma combinação de busca por secções áureas e interpolação parabólica [9, 23]. Sua sintaxe é: $[d_{min}, valor] = \text{FMINBND}(fun, x_1, x_2)$, onde fun representa a função objetiva, x_1 e x_2 representam os extremos do intervalo no qual procura-se pelo minimizador d_{min} e $valor$ retorna o valor da função objetiva em d_{min} .

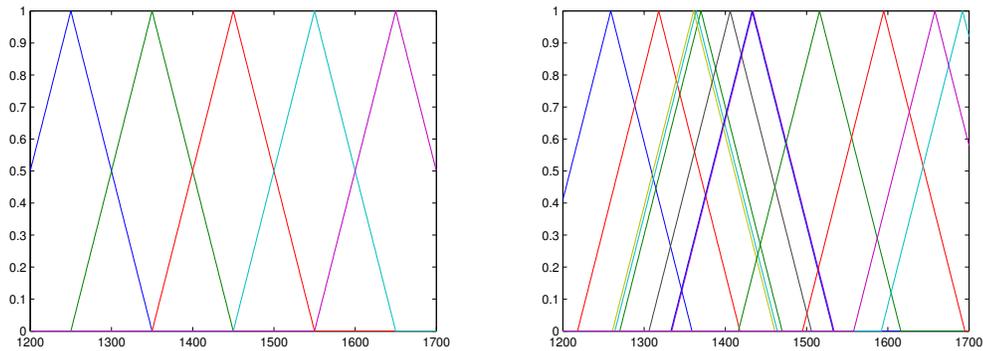


Figura 6.3: Conjuntos fuzzy triangulares da Tabela 6.1 e Tabela 6.2

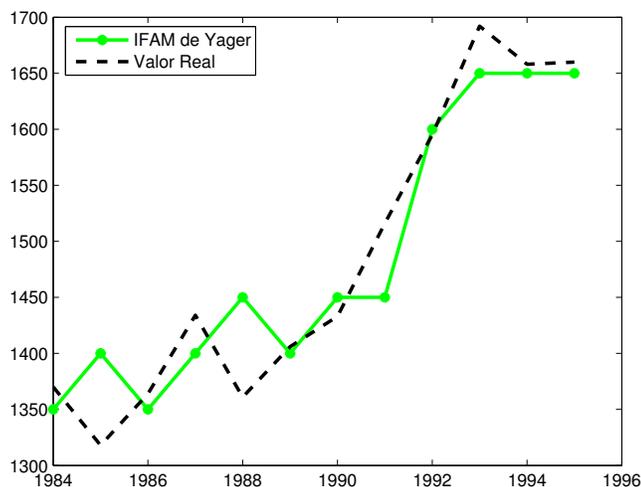


Figura 6.4: Previsão da mão-de-obra gerada pela IFAM de Yager adotando o método de defuzzificação a média do máximo e $d \in (0, 1]$

Esta aplicação de IFAM como sistemas de regras fuzzy usados para previsão de séries temporais apresentado por Choudhury *et al.* [12] apresenta alguns pontos questionáveis, como por exemplo as componentes das memórias fundamentais assumem apenas três valores: 0, 0.5 e 1. Também, cada conjunto fuzzy possui 5 elementos definidos no universo de discurso [1200,1700]. Além, disso o próprio conjunto de treinamento é usado para teste.

Outros testes foram realizados para esta mesma aplicação da previsão da mão-de-obra. Nesta etapa, utilizou-se o método da centróide como técnica de defuzzificação. Os conjuntos fuzzy da Tabela 6.1 e as respectivas fuzzificações da mão-de-obra real estão apresentados na Figura 6.4. A Tabela 6.6 representa a saída da função FMINBND para resolver o problema de otimização descrito pela Equação 6.15. O minimizador d encontrado foi igual a 6.36391 e a mão-de-obra estimada pela IFAM de Yager para este parâmetro é apresentado na Figura 6.5. Neste caso, a IFAM de Yager apresentou o melhor resultado para esse problema de previsão com EPMN de 2.43% e os demais erros gerados pelas diferentes FAMS são vistos na Tabela 6.5.

Método	EPMN	Método	EPMN
FAM max-min:	4.68%	FLBAM de Lukasiewicz:	2.89%
fAM max-produto:	3.49%	FLBAM de Gödel:	2.76%
GFAM de Lukasiewicz:	2.89%	FLBAM de Goguen	2.63%
FAM Junbo:	2.46%	IFAM de Yager	2.43%
IFAM de Lukasiewicz:	5.51%	IFAM dual de Lukasiewicz	6.76%
IFAM de Gödel:	2.46%	IFAM dual de Gödel:	7.93%
IFAM de Goguen:	2.56%	IFAM dual de Goguen	7.48%
IFAM Adjunta de Lukasiewicz	6.82%	IFAM dual adjunta de Lukasiewicz	7.26%
IFAM dual adjunta de Goguen	7.26 %	IFAM dual adjunta de Gödel	7.26 %

Tabela 6.5: EPMN na previsão de mão-de-obra, utilizando como método de defuzzificação a centróide e $d = 6.36391$ na t-norma de Yager

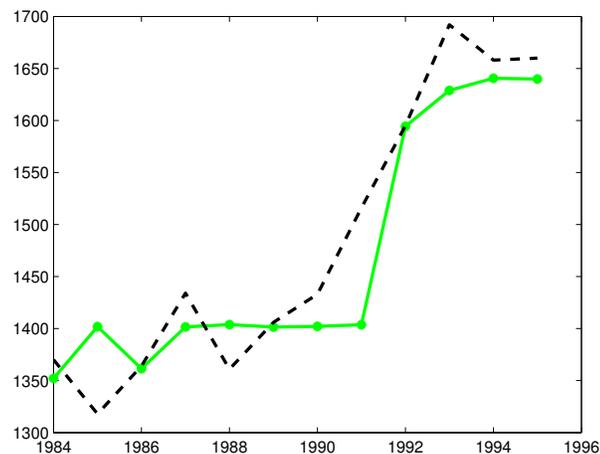


Figura 6.5: Previsão de mão-de-obra estimada pela IFAM de Yager adotando 6.36391, utilizando como método de defuzzificação a centróide. O EPMN obtido foi de 2.43%.

6.3.2 Previsão da Vazão Mensal de uma Usina Hidrelétrica

Esta seção apresenta uma aplicação da IFAM de Yager no problema e previsão da vazão mensal média da usina hidrelétrica de Furnas, localizada na região sudeste do Brasil. Esse

Func-count	d	$\ \mathbf{s}^d - \mathbf{s}\ _2$	Procedure
1	4.43769	173.086	initial
2	6.56231	168.833	golden
3	7.87539	168.955	golden
4	7.14266	168.886	parabolic
5	5.75078	168.96	golden
6	6.25233	168.832	golden
7	6.38372	168.828	parabolic
8	6.39662	168.828	parabolic
9	6.36503	168.828	parabolic
10	6.36152	168.828	parabolic
11	6.36561	168.828	parabolic
12	6.36422	168.828	parabolic
13	6.36319	168.828	golden
14	6.36418	168.828	parabolic
15	6.3638	168.828	golden
16	6.36401	168.828	parabolic
17	6.36397	168.828	parabolic
18	6.36391	168.828	golden

Tabela 6.6: Saída de cada iteração no processo de otimizar $\min \|\mathbf{s}^d - \mathbf{s}\|_2$, utilizando como método da centróide como técnica de defuzzificação.

problema foi discutido anteriormente em [39, 40], onde foi proposto uma técnica baseada em agrupamentos fuzzy, chamada FPM-PRP. Este problema de previsão foi revisto por Sussner e Valle aplicando a IFAM de Lukasiewicz [78, 70] cuja implementação pode ser encontrado em [77]. Apesar de não apresentar resultados tão bons quanto o FPM-PRP, superou os métodos estatísticos PARMA, a rede neural Perceptron de Múltiplas Camadas (MLP, Multi-layer Perceptron) e um certo modelo de rede neural fuzzy recorrente (NFN -

Neuro-Fuzzy Network).

O previsor baseado na IFAM de Yager associado ao problema de otimização para determinar o parâmetro d na t-norma de Yager, apresentou o menor Erro Quadrático Médio (EQM), superando todos os modelos apresentados anteriormente. Os erros produzidos pela IFAM de Yager e os demais modelos para a previsão da vazão média mensal de 1991 à 1998 são apresentados na Tabela 6.7. O Erro Quadrático Médio e o Erro Absoluto Médio (EAM) são calculados através da Equação 6.16, onde s_n e \hat{s}_n representam os valores reais e estimados, respectivamente. O EPMN é dado pela Equação 6.13. Para esta aplicação, tem-se $N = 96$.

$$EQM = \frac{1}{N} \sum_{n=1}^N (s_n - \hat{s}_n)^2 \quad \text{e} \quad EAM = \frac{1}{N} \sum_{n=1}^N |s_n - \hat{s}_n| \quad (6.16)$$

Método	EQM ($\times 10^5$)	EAM (m^3/s)	EPM(%)
IFAM de Lukasiewicz	1.27	229	24
IFAM de Yager	1.15	215.5	22
PARMA	1.85	280	28
MLP	1.82	271	30
NFN	1.73	234	20
FPM-PRP	1.20	200	18

Tabela 6.7: Erro Quadrático Médio (EQM), Erro Absoluto Médio (EAM) e Erro Percentual Médio (EPM) produzido pelos modelos de previsão.

As memórias fundamentais no modelo baseado na IFAM de Yager foi construído a partir dos dados históricos de vazão mensal de 1931 à 1990 e os dados dos históricos de 1991 à 1998 foram usados para teste. Conhecidas as amostras da série temporal s_ξ , para $\xi = 1, 2, \dots, q - 1$, deseja-se estimar o valor s_q , considerando-se um subconjunto de valores passados s_1, s_2, \dots, s_{q-1} . No entanto, é necessário transformar esses valores reais s_ξ em conjuntos fuzzy \mathbf{x}^ξ e \mathbf{y}^ξ , uma vez que o modelo baseado numa memória associativa é armazenar

SEÇÃO 6.3 APLICAÇÕES DE IFAMS DE YAGER EM SÉRIES TEMPORAIS 70

pares de associações $(\mathbf{x}^\xi, \mathbf{y}^\xi)$. Deste modo, dado um conjunto fuzzy \mathbf{x}^q que captura a informação da vazão média dos últimos meses, o padrão recordado \mathbf{y}^q pela IFAM representa o conjunto fuzzy que corresponde ao valor estimado para a vazão s_q . Este último pode ser obtido defuzzificando o padrão recordado \mathbf{y}^q com o auxílio do função DEFUZZ da Toolbox do MATLAB com a opção CENTROID que representa o método do centro de massa, ou seja, $s_q = defuzz(\mathbf{y}^q, centroid)$. É importante observar que os valores da série temporal foram normalizados antes de serem processados, isto é, os dados da vazão de cada mês possuem média 0 e variância 1.

O conjunto ds memórias fundamentais $\{(\mathbf{x}^\xi, \mathbf{y}^\xi), \xi = 1, \dots, k\}$ é armazenado implicitamente, isto é, não é necessário construir toda a matriz dos pesos sinápticos, mas sim a parte que será usada na fase de recordação. De fato, se o padrão de entrada é da forma $\mathbf{x}^q = [1, 0, \dots, 0] \in [0, 1]^n$, então o padrão recordado $\mathbf{y}^q = [y_1^q, \dots, y_m^q] \in [0, 1]^m$ é obtido pela Equação 6.17, para $j = 1, \dots, m$ e $d > 0$:

$$\begin{aligned} y_j^q &= (W \circ_Y \mathbf{x})_j = \bigvee_{i=1}^n T_Y^d(w_{ji}, x_i^\xi) \vee \theta_j = w_{j1} \vee \theta_j \\ &= \bigwedge_{\xi=1}^k I_Y(x_1^\xi, y_j^\xi) \vee \theta_j = \bigwedge_{\xi=1}^k \bigvee \left\{ z \in [0, 1] : T_Y^d(x_1^\xi, z) \leq y_j^\xi \right\} \vee \theta_j \\ &= \bigwedge_{\xi=1}^k \bigvee \left\{ z \in [0, 1] : 1 - \min \left\{ 1, \left\{ (1 - x_1^\xi)^d + (1 - z)^d \right\}^{1/d} \right\} \leq y_j^\xi \right\} \vee \theta_j \end{aligned} \quad (6.17)$$

Portanto, o padrão recordado é determinado com base somente nos valores x_1^ξ e nos padrões \mathbf{y}^ξ , para $\xi = 1, \dots, k$. Os próximos parágrafos descrevem como definir as memórias fundamentais para o problema de previsão. Primeiro defini-se os padrões \mathbf{p}^γ contendo os termos da série temporal necessários para estimar s_γ . Esses padrões são vetores da forma:

$$\mathbf{p}^\gamma = [s_{\gamma-h}, \dots, s_{\gamma-1}]^T \quad \text{para} \quad h \in \{1, 2, \dots, q-1\} \quad (6.18)$$

onde $\gamma > h$ tal que $\gamma = q, q-12, q-24, \dots$. O parâmetro h pode ser interpretado como o número de termos consecutivos da série temporal necessários para estimar s_γ . Por exemplo,

para $h = 2$ e s_γ correspondendo a vazão mensal do mês de Dezembro, então \mathbf{p}^γ é o vetor contendo a vazão de Outubro e Novembro. Como o parâmetro $\gamma = q, q - 12, \dots$ estamos informando para a memória associativa que as informações referentes as vazões de Janeiro à Setembro não possuem correlação com o mês de Dezembro. Nesta aplicação, adotou-se o parâmetro $h = 3$ por produzir os melhores resultados.

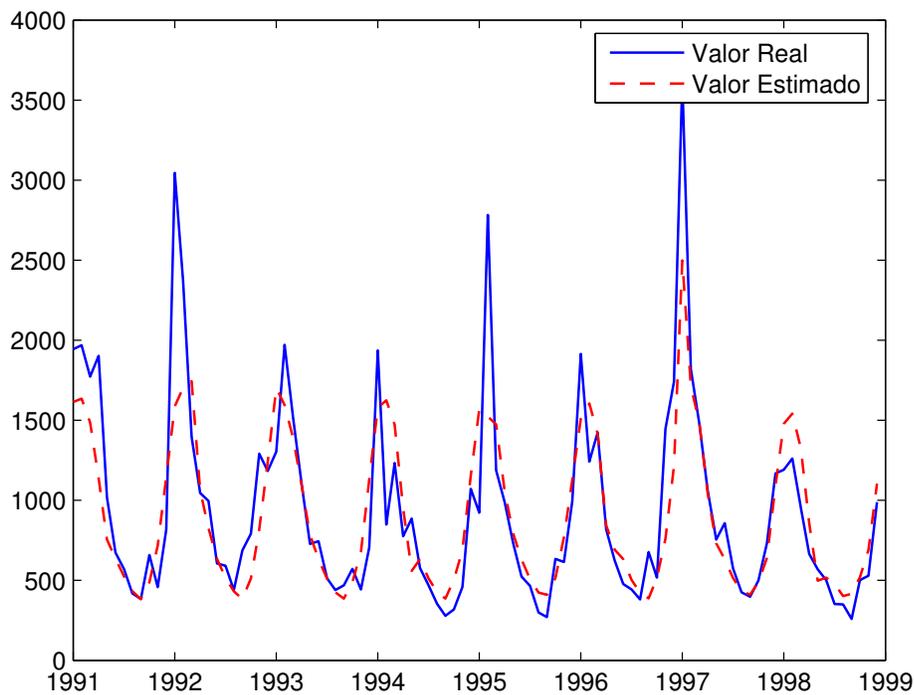


Figura 6.6: Previsão para a vazão média de Furnas de 1991 a 1998. A linha contínua corresponde aos valores reais e a linha tracejada corresponde aos valores estimados pela IFAM de Yager adotando $d = 2.10065$.

A fuzzificação de \mathbf{p}^ξ e s_ξ foram feitas através de funções de pertinência Gaussianas, produzindo conjuntos fuzzy $\mathbf{x}^\xi : U \rightarrow [0, 1]$ e $\mathbf{y}^\xi : V \rightarrow [0, 1]$, onde $U = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ e $V = \{v_1, \dots, v_m\}$, representam universos de discursos finitos e $\mathbf{u}_1 = \mathbf{p}^q$ e os v'_s são pontos

igualmente espaçados no intervalo $[-5,5]$ e $m = 500$. Note que o vetor $\mathbf{x}^q = [1, 0, \dots, 0]$ usado como padrão de entrada nessa aplicação é um conjunto clássico “crisp” que corresponde ao padrão \mathbf{p}^q .

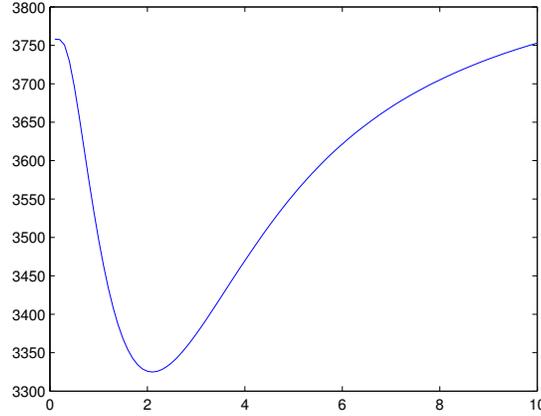


Figura 6.7: Gráfico da função $\|s^d - s\|_2$

Os conjuntos fuzzy \mathbf{x}^ξ e \mathbf{y}^ξ foram determinados utilizando o SUBTRACTIVE CLUSTERING METHOD (SCM), disponível na Toolbox de Conjuntos Fuzzy do MATLAB [11]. Esta função é utilizada para encontrar agrupamentos baseados na densidade dos dados. Sua sintaxe é dada por: $[C, d] = \text{subclust}(S, r)$, onde S contém os dados de 1931 a 1990 que serão agrupados e $r \in [0, 1]$ determina o raio de influência de um centro de grupo. Geralmente, boas escolhas para r são valores entre o intervalo 0.2 à 0.5. O melhor valor para r encontrado foi $r = 0.5$. Cada linha da matriz S foi definida concatenando \mathbf{p}^γ e s_γ , ou seja, $S(i, :) := [\mathbf{p}^\gamma, s_\gamma]$, para $s_\gamma = q - 12i$. Assim, cada linha da matriz C representa um centro. Para facilitar, denote c_x^ξ os primeiros h elementos e c_y^ξ o último elemento da linha ξ , isto é, $[c_x^\xi, c_y^\xi] = C(\xi, :)$ para $\xi = 1, \dots, k$. O vetor $\mathbf{d} = [d_1, \dots, d_{h+1}]^T$ é usado para determinar o espalhamento das funções de pertinência Gaussiana [81]:

$$\sigma_l^2 = \frac{1}{2d_l}, \quad \text{para } l = 1, 2, \dots, h + 1 \quad (6.19)$$

Finalmente os conjuntos fuzzy x_1^ξ e y^ξ são determinados através da Equação 6.20, para $\xi = 1, \dots, k$ e $i = 1, \dots, m$.

$$x_1^\xi = \exp \left[\frac{1}{2} \sum_{l=1}^h \left(\frac{(\mathbf{P}^q)_l - (c_{\mathbf{x}}^\xi)_l}{\sigma_l} \right)^2 \right] \quad \text{e} \quad y_i^\xi = \exp \left[\frac{1}{2} \sum_{l=1}^h \left(\frac{v_i - (c_{\mathbf{y}}^\xi)_l}{\sigma_{h+1}} \right)^2 \right] \quad (6.20)$$

A estimativa da vazão mensal s_q é obtida “defuzzificando” o padrão $\mathbf{y}^q = [y_1^q, \dots, y_m^q]^t$, onde y_1^q 's são determinados pela Equação 6.17. Se \mathbf{s}^q denota o vetor cujo os elementos são as estimativas para as previsões da vazão mensal de cada mês e \mathbf{s} o vetor com componentes com os valores reais da vazão mensal, então nosso problema de otimização passa a ser descrito pela Equação 6.21:

$$\begin{cases} \text{minimize } \|\mathbf{s}^q - \mathbf{s}\|_2 \\ \text{sujeito a } d \geq 0 \end{cases} \quad (6.21)$$

O parâmetro d da t-norma de Yager ótimo encontrado foi de $d = 2.10065$. O gráfico da função descrita pela Equação 6.21 pode ser visto na Figura 6.7. Finalmente, a Figura 6.6 mostra a previsão para a vazão média de Furnas de 1991 à 1998.

Func-count	d	$\ \mathbf{s}^q - \mathbf{s}\ _2$	Procedure
1	3.43776	3415.1	initial
2	5.56234	3595.32	golden
3	2.12469	3324.96	golden
4	1.31317	3405.27	golden
5	2.34615	3329.79	parabolic
6	2.14216	3325.06	parabolic
7	2.09934	3324.92	parabolic
8	2.10034	3324.92	parabolic
9	2.10062	3324.92	parabolic
10	2.10058	3324.92	parabolic
11	2.10065	3324.92	parabolic

Tabela 6.8: Saída de cada iteração no processo de otimizar $\min \|\mathbf{s}^q - \mathbf{s}\|_2$, utilizando como método de defuzzificação o centro de massa.

CAPÍTULO 7

Aplicação de IFAMs no Processo de Descompressão de Imagens

7.1 Introdução

Hirota e Pedrycz propuseram um método para compressão e descompressão (reconstrução) de imagens baseados em equações relacionais fuzzy [30]. Uma imagem em tons de cinza pode ser expressa como uma relação fuzzy $R \in [0, 1]^{n \times m}$ e dadas as famílias de conjuntos fuzzy $\mathbf{x}^\xi \in [0, 1]^n$, para $\xi = 1, \dots, p$ e $\mathbf{y}^\gamma \in [0, 1]^m$, para $\gamma = 1, \dots, q$, pode-se comprimir R em uma imagem $G \in [0, 1]^{p \times q}$, onde $p < n$ e $q < m$. Agora, dado a imagem $G \in [0, 1]^{p \times q}$ e os conjuntos fuzzy $\mathbf{x}^\xi \in [0, 1]^n$ e $\mathbf{y}^\gamma \in [0, 1]^m$, surge o problema inverso de reconstruir uma aproximação da imagem original $R \in [0, 1]^{n \times m}$. Uma solução dada por Hirota e Pedrycz foi encontrar a maior solução \hat{R} [30, 48]. O que pretende-se neste capítulo é relacionar o processo de reconstrução de imagens à teoria das memórias associativas fuzzy implicativas [72].

7.2 Processo de Compressão/Descompressão

Sejam $R \in [0, 1]^{n \times m}$ uma relação binária fuzzy e $T : [0, 1] \times [0, 1] \rightarrow [0, 1]$ uma t-norma contínua. Dadas as famílias de conjuntos fuzzy $\mathbf{x}^\xi \in [0, 1]^n$ para $\xi = 1, \dots, p$ e $\mathbf{y}^\gamma \in [0, 1]^m$

para $\gamma = 1, \dots, q$, o processo de compressão da relação fuzzy $R \in [0, 1]^{n \times m}$ (imagem) é definido pela seguinte equação:

$$G(\xi, \gamma) = (R^T \circ_T \mathbf{x}^\xi)^t \circ_T \mathbf{y}^\gamma \quad (7.1)$$

onde o símbolo \circ_T denota a composição max- T .

Para facilitar a compreensão das composições entre operadores fuzzy, será adotado neste capítulo uma nova notação. Especificamente, para a t-norma $T(a, b) = atb$ e para a implicação fuzzy $I(a, b) = a \Rightarrow b$ para todo $a, b \in [0, 1]$. Denotam-se a relação fuzzy $R = [r_{ij}]$ para $i = 1, \dots, n$ e $j = 1, \dots, m$, os conjuntos fuzzy $\mathbf{x}^\xi = [x_1^\xi, \dots, x_n^\xi]$ e $\mathbf{y}^\gamma = [y_1^\gamma, \dots, y_m^\gamma]$. Reescrevendo a Equação 7.1 na forma de função de pertinência:

$$G(\xi, \gamma) = \bigvee_{j=1}^m \left\{ \bigvee_{i=1}^n \{ r_{ij} t x_i^\xi \} t y_j^\gamma \right\} \quad \text{para } \xi = 1, \dots, p \text{ e } \gamma = 1, \dots, q \quad (7.2)$$

Portanto, a imagem $R \in [0, 1]^{n \times m}$ é comprimida na imagem $G \in [0, 1]^{p \times q}$ pela composição max- T de R , e os conjuntos fuzzy \mathbf{x}^ξ e \mathbf{y}^γ para $\xi = 1, \dots, p$ e $\gamma = 1, \dots, q$ com a condição de $p < n$ e $q < m$. Agora, considerando o problema inverso, ou seja, dados a família de conjuntos fuzzy \mathbf{x}^ξ , \mathbf{y}^γ e a imagem comprimida $G \in [0, 1]^{p \times q}$, deseja-se encontrar a reconstrução (aproximação) $\hat{R} \in [0, 1]^{n \times m}$ da imagem original $R \in [0, 1]^{n \times m}$. A solução dada por Hirota e Pedrycz foi encontrar maior solução da Equação 7.1 [30, 48], segundo a equação:

$$\hat{R} = \min_{\xi, \gamma} \{ (\mathbf{x}^\xi \times_t \mathbf{y}^\gamma) \Rightarrow_R G(\xi, \gamma) \} \quad (7.3)$$

O símbolos \times_t denota o produto t -cartesiano e \Rightarrow_R denota a R-implicação, definida na Equação 2.22, vista no capítulo referente a teoria dos conjuntos fuzzy. Estas equações são definidas a seguir:

$$(\mathbf{x}^\xi \times_t \mathbf{y}^\gamma)(i, j) = (\mathbf{x}_i^\xi t \mathbf{y}_j^\gamma) \quad \text{para } i = 1, \dots, n \text{ e } j = 1, \dots, m \quad (7.4)$$



Figura 7.1: Imagem original 256×256

$$(a \Rightarrow_R b) = \bigvee \{c \in [0, 1] : (atc) \leq b\}, \quad \forall a, b \in [0, 1] \quad (7.5)$$

Reescrevendo a Equação 7.3 na sua respectiva função de pertinência, obtém-se:

$$\hat{r}_{ij} = \min_{\xi, \gamma} \left\{ \left(\mathbf{x}_i^\xi t \mathbf{y}_j^\gamma \right) \Rightarrow_R G(\xi, \gamma) \right\} \quad \text{para } i = 1, \dots, n \text{ e } j = 1, \dots, m \quad (7.6)$$

Portanto, se a relação fuzzy $R \in [0, 1]^{n \times m}$ é solução da equação $(R^t \circ_T \mathbf{x}^\xi)^t \circ_T \mathbf{y}^\gamma = G(\xi, \gamma)$, então $R \leq \hat{R}$, e mais, \hat{R} também é solução.

Demonstração

Primeiro, mostra-se que a Equação 7.3 é a maior solução da Equação 7.1. Denotando a relação fuzzy $R \in [0, 1]^{n \times m}$ por $R = [r_{ij}]$ para $i = 1, \dots, n$ e $j = 1, \dots, m$, como também os conjuntos fuzzy $\mathbf{x}^\xi = [x_1^\xi, \dots, x_n^\xi]$ e $\mathbf{y}^\gamma = [y_1^\gamma, \dots, y_m^\gamma]$, para $\xi = 1, \dots, p$ e $\gamma = 1, \dots, q$. Pela Equação 7.2, tem-se para todo $\xi = 1, \dots, p$ e $\gamma = 1, \dots, q$:

$$\begin{aligned}
G(\xi, \gamma) &= \bigvee_{j=1}^m \bigvee_{i=1}^n \left\{ (r_{ij} t x_i^\xi) t y_j^\gamma \right\} = \bigvee_{j=1}^m \bigvee_{i=1}^n \left\{ r_{ij} t (x_i^\xi t y_j^\gamma) \right\} = \\
&= \bigvee_{j=1}^m \bigvee_{i=1}^n \left\{ (x_i^\xi t y_j^\gamma) t r_{ij} \right\} \geq (x_i^\xi t y_j^\gamma) t r_{ij}
\end{aligned} \tag{7.7}$$

Logo,

$$\begin{aligned}
(x_i^\xi t y_j^\gamma) t r_{ij} \leq G(\xi, \gamma) \quad \text{para } i = 1, \dots, n \text{ e } j = 1, \dots, m \\
\xi = 1, \dots, p \text{ e } \gamma = 1, \dots, q
\end{aligned} \tag{7.8}$$

Segue que, para $i = 1, \dots, n$ e $j = 1, \dots, m$

$$\begin{aligned}
r_{ij} &\in \left\{ w \in [0, 1] \mid (x_i^\xi t y_j^\gamma) t w \leq G(\xi, \gamma) \right\} \\
r_{ij} &\leq \bigvee \left\{ w \in [0, 1] \mid (x_i^\xi t y_j^\gamma) t w \leq G(\xi, \gamma) \right\} \\
r_{ij} &\leq (x_i^\xi t y_j^\gamma) \Rightarrow_R G(\xi, \gamma) \\
r_{ij} &\leq \min_{\xi, \gamma} \left\{ (x_i^\xi t y_j^\gamma) \Rightarrow_R G(\xi, \gamma) \right\} = \hat{r}_{ij}
\end{aligned} \tag{7.9}$$

Portanto, $R \leq \hat{R}$.

Falta mostrar que \hat{R} é solução da Equação 7.1. De um lado, tem-se para todo $i = 1, \dots, n$ e $j = 1, \dots, m$

$$r_{ij} \leq \hat{r}_{ij} \tag{7.10}$$

Pela monotonicidade da t-norma, tem-se que para $i = 1, \dots, n$ e $j = 1, \dots, m$

$$\begin{aligned}
 r_{ij}tx_i^\xi &\leq \hat{r}_{ij}tx_i^\xi \quad \text{para} \quad \xi = 1, \dots, p \\
 (r_{ij}tx_i^\xi)ty_j^\gamma &\leq (\hat{r}_{ij}tx_i^\xi)ty_j^\gamma \quad \text{para} \quad \xi = 1, \dots, p \text{ e } \gamma = 1, \dots, q
 \end{aligned} \tag{7.11}$$

$$\bigvee_{j=1}^m \bigvee_{i=1}^n (r_{ij}tx_i^\xi)ty_j^\gamma \leq \bigvee_{j=1}^m \bigvee_{i=1}^n (\hat{r}_{ij}tx_i^\xi)ty_j^\gamma \quad \text{para} \quad \xi = 1, \dots, p \text{ e } \gamma = 1, \dots, q$$

Ou seja,

$$G(\xi, \gamma) \leq (\hat{R}^T \circ_t \mathbf{x}^\xi)^T \circ_t \mathbf{y}^\gamma \tag{7.12}$$

Por outro lado, temos para $\xi = 1, \dots, p$ e $\gamma = 1, \dots, q$

$$\begin{aligned}
 \bigvee_{j=1}^m \bigvee_{i=1}^n (\hat{r}_{ij}tx_i^\xi)ty_j^\gamma &= \bigvee_{j=1}^m \bigvee_{i=1}^n \left\{ \left[\min_{\xi, \gamma} \left\{ (x_i^\xi ty_j^\gamma) \Rightarrow_R G(\xi, \gamma) \right\} \right] tx_i^\xi \right\} ty_j^\gamma \\
 &= \bigvee_{j=1}^m \bigvee_{i=1}^n \left\{ x_i^\xi t \left[\min_{\xi, \gamma} \left\{ (x_i^\xi ty_j^\gamma) \Rightarrow_R G(\xi, \gamma) \right\} \right] \right\} ty_j^\gamma \\
 &= \bigvee_{j=1}^m \bigvee_{i=1}^n y_j^\gamma t \left\{ x_i^\xi t \left[\min_{\xi, \gamma} \left\{ (x_i^\xi ty_j^\gamma) \Rightarrow_R G(\xi, \gamma) \right\} \right] \right\} \\
 &= \bigvee_{j=1}^m \bigvee_{i=1}^n (y_j^\gamma tx_i^\xi) t \left[\min_{\xi, \gamma} \left\{ (x_i^\xi ty_j^\gamma) \Rightarrow_R G(\xi, \gamma) \right\} \right] \\
 &= \bigvee_{j=1}^m \bigvee_{i=1}^n (x_i^\xi ty_j^\gamma) t \left[\min_{\xi, \gamma} \left\{ (x_i^\xi ty_j^\gamma) \Rightarrow_R G(\xi, \gamma) \right\} \right]
 \end{aligned} \tag{7.13}$$

Como $at(a \Rightarrow_R b) \leq b, \forall a, b \in [0, 1]$, tem-se

$$= \bigvee_{j=1}^m \bigvee_{i=1}^n (x_i^\xi t y_j^\gamma) t \left[\min_{\xi, \gamma} \left\{ (x_i^\xi t y_j^\gamma) \Rightarrow_R G(\xi, \gamma) \right\} \right] \leq G(\xi, \gamma) \quad (7.14)$$

Ou seja,

$$(\hat{R}^T \circ_t \mathbf{x}^\xi)^T \circ_t \mathbf{y}^\gamma \leq G(\xi, \gamma) \quad (7.15)$$

Exemplo 7.2.1. *Sejam os conjuntos fuzzy triangulares:*

$$\mathbf{x}^1 = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^T \quad \mathbf{x}^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T \quad (7.16)$$

$$\mathbf{y}^1 = \begin{bmatrix} 0 & 1 & 0.5 & 0 \end{bmatrix}^T \quad \mathbf{y}^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T$$

Esses conjuntos podem ser vistos na Figura 7.2 e a imagem $R \in [0, 1]^{4 \times 4}$ é dada por:

$$R = \begin{bmatrix} 0 & 0.9 & 0.5 & 0 \\ 0.3 & 0.7 & 0.6 & 0 \\ 0.3 & 0.4 & 0.2 & 0.3 \\ 0.1 & 0.4 & 1 & 0.7 \end{bmatrix} \quad (7.17)$$

Assumindo o operador mínimo como a t-norma contínua, calcula-se a imagem comprimida $G \in [0, 1]^{2 \times 2}$ de acordo com a Equação 7.1:

$$G(1, 1) = (R^T \circ_T \mathbf{x}^1)^t \circ_T \mathbf{y}^1 = \left(\left(\begin{bmatrix} 0 & 0.3 & 0.3 & 0.1 \\ 0.9 & 0.7 & 0.4 & 0.4 \\ 0.5 & 0.6 & 0.2 & 1 \\ 0 & 0 & 0.3 & 0.7 \end{bmatrix} \circ_T \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right)^t \circ_T \begin{bmatrix} 0 \\ 1 \\ 0.5 \\ 0 \end{bmatrix} \quad (7.18)$$

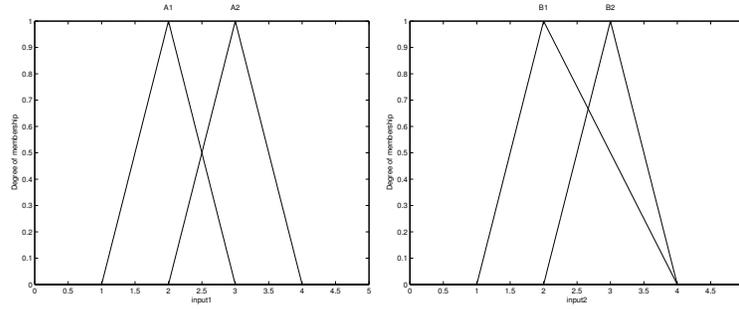


Figura 7.2: Conjuntos Fuzzy do Exemplo 7.2.1

$$= \begin{bmatrix} 0.3 & 0.7 & 0.6 & 0 \end{bmatrix} \circ_T \begin{bmatrix} 0 \\ 1 \\ 0.5 \\ 0 \end{bmatrix} = (0.3 \wedge 0) \vee (0.7 \wedge 1) \vee (0.6 \wedge 0.5) \vee (0 \wedge 0) = 0.7 \quad (7.19)$$

Da mesma forma, obtém-se:

$$G(1,2) = (R^t \circ_T \mathbf{x}^1)^t \circ_T \mathbf{y}^2 = 0.6 \quad (7.20)$$

$$G(2,1) = (R^t \circ_T \mathbf{x}^2)^t \circ_T \mathbf{y}^1 = 0.4 \quad (7.21)$$

$$G(2,2) = (R^t \circ_T \mathbf{x}^2)^t \circ_T \mathbf{y}^2 = 0.2 \quad (7.22)$$

Logo, a imagem comprimida G é dada por:

$$G = \begin{bmatrix} 0.7 & 0.6 \\ 0.4 & 0.2 \end{bmatrix} \quad (7.23)$$

Considere, agora, o problema inverso: dados os conjunto fuzzy \mathbf{x}^ξ e \mathbf{y}^γ para $\xi, \gamma = 1, 2$ e a imagem comprimida $G \in [0, 1]^{2 \times 2}$, deseja-se encontrar a descompressão (reconstrução) da imagem original R . Pela Equação 7.3, tem-se:

$$R_{11} = ((\mathbf{x}^1 \times_t \mathbf{y}^1) \Rightarrow_R G(1, 1)) = \left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \times_t \begin{bmatrix} 0 & 1 & 0.5 & 0 \end{bmatrix} \right) \Rightarrow_R 0.7 = \quad (7.24)$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow_R 0.7 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0.7 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (7.25)$$

Analogamente, obtém-se as seguintes soluções:

$$R_{12} = ((\mathbf{x}^1 \times_t \mathbf{y}^2) \Rightarrow_R G(1, 2)) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0.6 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (7.26)$$

$$R_{21} = ((\mathbf{x}^2 \times_t \mathbf{y}^1) \Rightarrow_R G(2, 1)) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0.4 & 0.4 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (7.27)$$

$$R_{22} = ((\mathbf{x}^2 \times_t \mathbf{y}^2) \Rightarrow_R G(2, 2)) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0.2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (7.28)$$

Portanto, a maior solução é dada pela seguinte equação:

$$\hat{R} = \min_{\xi, \gamma} \{(\mathbf{x}^\xi \times_t \mathbf{y}^\gamma) \Rightarrow_R G(\xi, \gamma)\} = R_{11} \wedge R_{12} \wedge R_{21} \wedge R_{22} \quad (7.29)$$

$$\hat{R} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0.7 & 0.6 & 1 \\ 1 & 0.4 & 0.2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (7.30)$$

7.3 Experimentos

Para as Figuras 7.3 e 7.4, foram adotados conjuntos fuzzy triangulares e a t-norma do mínimo. A Figura 7.2 mostra imagem original em tons de cinza de dimensões $n \times m = 256 \times 256$. A Figura 7.3, mostra a imagem original comprimida em $p \times q = 32 \times 32$ (razão de compressão = 0.0156) e $p \times q = 64 \times 64$ (razão de compressão = 0.0625), respectivamente. Já a Figura 7.4, representa as imagens reconstruídas a partir das imagens dadas pela Figura 7.3, respectivamente.

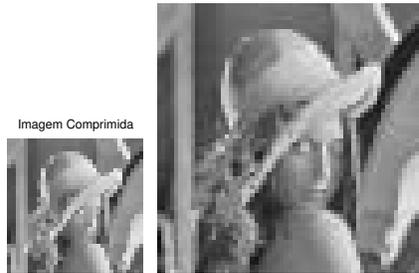


Figura 7.3: Imagens comprimidas com razão de compressão 0.0156 (32×32) e 0.0625 (64×64), respectivamente. Para esse processo, utilizou-se a t-norma contínua do mínimo e conjuntos fuzzy triangulares

Também realizou-se o processo de compressão/descompressão utilizando a família de conjuntos fuzzy $\{x^\xi : \xi = 1, 2, \dots, p\}$ e $\{y^\gamma : \gamma = 1, 2, \dots, q\}$ definidos a seguir:

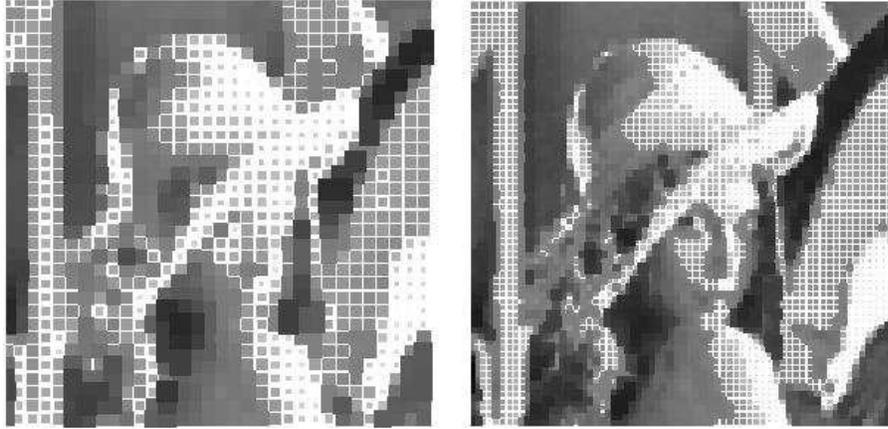


Figura 7.4: Imagens reconstruídas a partir da Figura 7.3, utilizando a t-norma contínua do mínimo e conjuntos fuzzy triangulares.

$$x_i^\xi = \exp \left(-Sh \left(\frac{\xi n}{p} - i \right)^2 \right) \quad \text{para } i = 1, 2, \dots, n \quad (7.31)$$

$$y_j^\gamma = \exp \left(-Sh \left(\frac{\xi m}{q} - j \right)^2 \right) \quad \text{para } j = 1, 2, \dots, m \quad (7.32)$$

onde $Sh > 0$ é um parâmetro tomado entre 0.01–0.05. Nesta etapa, será adotada a t-norma contínua de Yager, dada por:

$$a \underset{y}{t} b = 1 - \min \left\{ 1, \left\{ (1-a)^d + (1-b)^d \right\}^{1/d} \right\} \quad d \geq 1 \quad (7.33)$$

As Figuras 7.5 e 7.6, representam os resultados do processo de compressão e reconstrução a partir dos conjuntos fuzzy definidos nas Equações 7.31 e 7.32 e a t-norma de Yager, tomando $d = 1$. A Figura 7.5, mostra a imagem original comprimida em $p \times q = 32 \times 32$ (razão de compressão = 0.0156 e $Sh = 0.01$) e $p \times q = 64 \times 64$ (razão de compressão = 0.0625 e $Sh = 0.05$), respectivamente. O resultado do processo de descompressão da Figura 7.5 pode ser vista na Figura 7.6.



Figura 7.5: Imagens comprimidas com razão de compressão 0.0156 (32×32) e 0.0625 (64×64), respectivamente. Para esse processo, utilizou-se a t-norma contínua de Yager com o parâmetro $d = 1$ e os conjuntos fuzzy definidos nas Equações 7.31 e 7.32

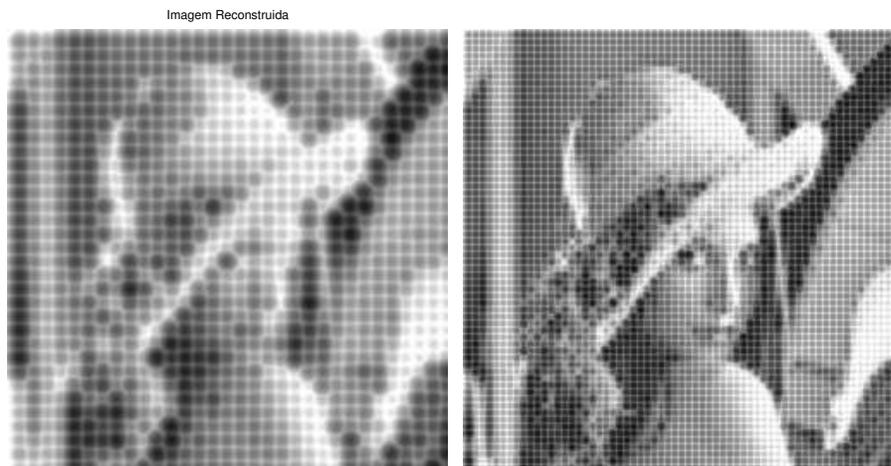


Figura 7.6: Imagens reconstruídas a partir da Figura 7.5, utilizando a t-norma contínua de Yager e os conjuntos fuzzy definidos nas Equações 7.31 e 7.32, com $Sh = 0.01$ para a razão de compressão igual a 0.0156 e $Sh = 0.05$ para a razão de compressão igual a 0.0625

7.4 O Processo de Reconstrução de Imagens e a Teoria das Memórias Associativas Fuzzy Implicativas

7.4.1 Considerações Iniciais

Nesta seção, será relacionado o problema de reconstrução de imagens através das equações relacionais fuzzy (FRE) à teoria das memórias associativas fuzzy implicativas (IFAMs). Pela

Equação 7.1, tem-se:

$$G(\xi, \gamma) = (R^t \circ_T \mathbf{x}^\xi)^t \circ_T \mathbf{y}^\gamma \quad (7.34)$$

Se \mathbf{g}^γ denota a γ -ésima coluna de $G \in [0, 1]^{p \times q}$, então a maior solução V^t da equação relacional fuzzy $\mathbf{g}^\gamma = V^t \circ_T \mathbf{y}^\gamma$ é dado por uma IFAM. Se \mathbf{v}^ξ denota a ξ -ésima coluna de $V \in [0, 1]^{m \times p}$, então a maior solução $R^T \in [0, 1]^{m \times n}$ da FRE $\mathbf{v}^\xi = R^t \circ_T \mathbf{x}^\xi$ também é dado por uma IFAM [72].

Recordando que uma IFAM é uma rede neuro-fuzzy descrita pelas seguintes equações [72]:

$$\mathbf{y} = (W \circ_T \mathbf{x}) \vee \theta \quad \text{onde} \quad W = Y \circledast_T X^t \quad \text{e} \quad \theta = \bigwedge_{\xi=1}^p \mathbf{y}^\xi \quad (7.35)$$

onde $X = [\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^p]$ e $Y = [\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^p]$ e a R-implicação fuzzy \Rightarrow_R é unicamente determinada pela equação [72]:

$$(a \Rightarrow_R b) = \bigvee \{c \in [0, 1] : (atc) \leq b\}, \quad \forall a, b \in [0, 1] \quad (7.36)$$

Neste contexto de compressão/reconstrução, assume-se $\theta = \mathbf{0}$, ou seja, o vetor nulo e conseqüentemente, a Equação 7.35 passa a ser regida apenas pelas equações:

$$\mathbf{y} = (W \circ_T \mathbf{x}) \quad \text{e} \quad W = Y \circledast_T X^t \quad (7.37)$$

7.4.2 Formulação do Problema

Dado a imagem comprimida $G \in [0, 1]^{p \times q}$ e os conjuntos fuzzy $\mathbf{x}^\xi \in [0, 1]^n$ para $\xi = 1, 2, \dots, p$ e $\mathbf{y}^\gamma \in [0, 1]^m$ para $\gamma = 1, 2, \dots, q$, deseja-se encontrar uma aproximação \hat{R} da imagem original $R \in [0, 1]^{n \times m}$ por meio de uma IFAM. Para cada $\gamma = 1, \dots, q$, deseja-se encontrar a solução $V^t \in [0, 1]^{p \times m}$ da equação abaixo:

$$\mathbf{g}^\gamma = V^t \circ_T \mathbf{y}^\gamma \quad \text{para} \quad \gamma = 1, \dots, q \quad (7.38)$$

onde \mathbf{g}^γ denota a γ -ésima coluna de G . A maior solução da Equação 7.38 é dada pela memória associativa fuzzy implicativa:

$$V^t = G \circledast_T Y^t \quad (7.39)$$

onde $Y = [\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^q] \in [0, 1]^{m \times q}$ e o símbolo \circledast_T denota o produto:

$$v_{\xi j} = \bigwedge_{\gamma=1}^q (y_j^\gamma \Rightarrow_R g_\xi^\gamma) \quad \text{para} \quad j = 1, \dots, m \quad \text{e} \quad \xi = 1, \dots, p \quad (7.40)$$

Obtido $V = [\mathbf{v}^1, \mathbf{v}^2, \dots, \mathbf{v}^p]$ da Equação 7.39, mais uma vez, recorre-se as IFAMs para encontrar a maior solução \hat{R}^t da equação abaixo:

$$\mathbf{v}^\xi = R^t \circ_T \mathbf{x}^\xi \quad \text{para} \quad \xi = 1, \dots, p \quad (7.41)$$

dada por:

$$\hat{R}^t = V \circledast_T X^t \quad (7.42)$$

onde $X = [\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^p] \in [0, 1]^{n \times p}$ ou na forma de função de pertinência:

$$\hat{r}_{ji} = \bigwedge_{\xi=1}^p (x_i^\xi \Rightarrow_R v_j^\xi) \quad \text{para} \quad i = 1, \dots, n \quad \text{e} \quad j = 1, \dots, m \quad (7.43)$$

Exemplo 7.4.1. *Sejam os conjuntos fuzzy triangulares, a t -norma do mínimo e a imagem comprimida G do Exemplo 7.2.1.*

$$\mathbf{x}^1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{x}^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{y}^1 = \begin{bmatrix} 0 \\ 1 \\ 0.5 \\ 0 \end{bmatrix} \quad \mathbf{y}^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad G = \begin{bmatrix} 0.7 & 0.6 \\ 0.4 & 0.2 \end{bmatrix} \quad (7.44)$$

Tomando $\gamma = 1$ e pela IFAM descrita pela Equação 7.40, tem-se:

$$\begin{aligned} v_{11}^1 &= y_1^1 \Rightarrow_R g_1^1 = \bigvee \{z \in [0, 1] \mid y_1^1 t z \leq g_1^1\} = \bigvee \{z \in [0, 1] \mid (0 \wedge z) \leq 0.7\} = 1 \\ v_{12}^1 &= y_2^1 \Rightarrow_R g_1^1 = \bigvee \{z \in [0, 1] \mid y_2^1 t z \leq g_1^1\} = \bigvee \{z \in [0, 1] \mid (1 \wedge z) \leq 0.7\} = 0.7 \\ v_{13}^1 &= y_3^1 \Rightarrow_R g_1^1 = \bigvee \{z \in [0, 1] \mid y_3^1 t z \leq g_1^1\} = \bigvee \{z \in [0, 1] \mid (0.5 \wedge z) \leq 0.7\} = 1 \\ v_{14}^1 &= y_4^1 \Rightarrow_R g_1^1 = \bigvee \{z \in [0, 1] \mid y_4^1 t z \leq g_1^1\} = \bigvee \{z \in [0, 1] \mid (0 \wedge z) \leq 0.7\} = 1 \end{aligned} \quad (7.45)$$

Seguindo este mesmo raciocínio, obtem-se: $v_{21}^1 = 1$, $v_{22}^1 = 0.4$, $v_{23}^1 = 0.4$ e $v_{24}^1 = 1$, assim:

$$(V^1)^t = \begin{bmatrix} 1 & 0.7 & 1 & 1 \\ 1 & 0.4 & 0.4 & 1 \end{bmatrix} \quad (7.46)$$

Similarmente, para $\gamma = 2$, obtém-se:

$$(V^2)^t = \begin{bmatrix} 1 & 1 & 0.6 & 1 \\ 1 & 1 & 0.2 & 1 \end{bmatrix} \quad (7.47)$$

Logo,

$$V^t = (V^1)^t \wedge (V^2)^t = \begin{bmatrix} 1 & 0.7 & 0.6 & 1 \\ 1 & 0.4 & 0.2 & 1 \end{bmatrix} \quad (7.48)$$

Obtido $V \in [0, 1]^{4 \times 2}$, de forma análoga, encontra-se \hat{R} pela IFAM descrita pela Equação 7.42. Tomando $\xi = 1$, tem-se:

$$\begin{aligned}
 \hat{r}_{11}^1 &= x_1^1 \Rightarrow_R v_1^1 = \bigvee \{w \in [0, 1] | x_1^1 t z \leq v_1^1\} = \bigvee \{w \in [0, 1] | (0 \wedge w) \leq 1\} = 1 \\
 \hat{r}_{12}^1 &= x_2^1 \Rightarrow_R v_1^1 = \bigvee \{w \in [0, 1] | x_2^1 t z \leq v_1^1\} = \bigvee \{w \in [0, 1] | (1 \wedge w) \leq 1\} = 1 \\
 \hat{r}_{13}^1 &= x_3^1 \Rightarrow_R v_1^1 = \bigvee \{w \in [0, 1] | x_3^1 t z \leq v_1^1\} = \bigvee \{w \in [0, 1] | (0 \wedge w) \leq 1\} = 1 \\
 \hat{r}_{14}^1 &= x_4^1 \Rightarrow_R v_1^1 = \bigvee \{w \in [0, 1] | x_4^1 t z \leq v_1^1\} = \bigvee \{w \in [0, 1] | (0 \wedge w) \leq 1\} = 1
 \end{aligned} \tag{7.49}$$

Seguindo este mesmo raciocínio, obtém-se:

$$\begin{aligned}
 \hat{r}_{21}^1 &= 1 & \hat{r}_{22}^1 &= 1 & \hat{r}_{23}^1 &= 1 & \hat{r}_{24}^1 &= 1 \\
 \hat{r}_{31}^1 &= 1 & \hat{r}_{32}^1 &= 0.7 & \hat{r}_{33}^1 &= 1 & \hat{r}_{34}^1 &= 1 \\
 \hat{r}_{41}^1 &= 1 & \hat{r}_{42}^1 &= 0.6 & \hat{r}_{43}^1 &= 1 & \hat{r}_{44}^1 &= 1
 \end{aligned} \tag{7.50}$$

$$(\hat{R}^1)^t = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0.7 & 1 & 1 \\ 1 & 0.6 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{7.51}$$

Similarmente, para $\xi = 2$, tem-se:

$$(\hat{R}^2)^t = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0.4 & 1 \\ 1 & 1 & 0.2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{7.52}$$

Daí,

$$\hat{R}^t = (\hat{R}^1)^t \wedge (\hat{R}^2)^t = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0.7 & 0.4 & 1 \\ 1 & 0.6 & 0.2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (7.53)$$

Portanto,

$$\hat{R} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0.7 & 0.6 & 1 \\ 1 & 0.4 & 0.2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (7.54)$$

Note que a reconstrução \hat{R} dada por 7.54 obtida por meio das IFAMs 7.39 e 7.42 é justamente a reconstrução 7.30 obtida por meio da Equação 7.3. Nobuhara, Pedrycz e Hirota apresentaram um método mais eficiente para calcular \hat{R} [48]:

$$\hat{r}_{ij} = \bigwedge_{\gamma=1}^q \left\{ y_j^\gamma \Rightarrow_R \left\{ \bigwedge_{\xi=1}^p [x_i^\xi \Rightarrow_R G(\xi, \gamma)] \right\} \right\} \quad (7.55)$$

Em [48], mostraram que a Equação 7.6 é equivalente a Equação 7.55 e iremos mostrar que o processo de reconstrução pelas IFAMs também é equivalente à essas duas equações. De fato, tomando o transposto em ambos os lados da Equação 7.39 e substituindo na Equação 7.42, tem-se:

$$\hat{R}^t = (G \otimes_T Y^t)^t \otimes_T X^t \quad \text{ou} \quad \hat{R} = \left[(G \otimes_T Y^t)^t \otimes_T X^t \right]^t \quad (7.56)$$

cuja função de pertinência revela:

$$\hat{r}_{ij} = \bigwedge_{\xi=1}^p \left\{ x_i^\xi \Rightarrow_R \left\{ \bigwedge_{\gamma=1}^q [y_j^\gamma \Rightarrow_R G(\xi, \gamma)] \right\} \right\} \quad (7.57)$$

e como

$$[a \Rightarrow_T (b \Rightarrow_T c)] = [b \Rightarrow_T (a \Rightarrow_T c)] \quad (7.58)$$

concluí-se da Equação 7.58 que a Equação 7.57 é equivalente a Equação 7.55 e a Equação 7.6.

Nobuhara, Pedrycz e Hirota mostraram em [48] que o processo de reconstrução é mais eficiente quando se emprega a Equação 7.55 ao invés da Equação 7.3. Considere L , T e I o tempo computacional para realizar as operações de mínimo, da t-norma e da R-implicação, respectivamente. Se NM e PQ denotam o tamanho da imagem original R e o tamanho da imagem comprimida G , respectivamente, então o tempo computacional estimado para efetuar a reconstrução dada pela Equação 7.3 é de $PQNM(L + T + I)$. No entanto, esse número reduz para $QN(P + M)(L + I)$ quando aplica-se a Equação 7.55 [48]. No nosso modelo proposto, também encontra-se a maior solução \hat{R} por meio das IFAMs com tempo computacional de $QN(P + M)(L + I)$.

7.5 Experimentos

Considere a segunda imagem comprimida $G \in [0, 1]^{64 \times 64}$ da Figura 7.5 com razão de compressão igual a 0.0625 e os conjuntos fuzzy das Equações 7.31 e 7.32 com o parâmetro $Sh = 0.05$. A imagem do lado esquerdo da Figura 7.7 representa a IFAM $V^t = G \circledast_T Y^t$ e a imagem do lado direito a maior aproximação da imagem original $R \in [0, 1]^{256 \times 256}$ por meio da IFAM $\hat{R}^t = V \circledast_T X^t$, adotando a R-implicação associada a t-norma contínua de Yager ($d = 1$).

O segundo experimento foi realizado empregando a t-norma contínua do mínimo e os conjuntos fuzzy triangulares. O resultado do processo de reconstrução da segunda imagem da Figura 7.3 utilizando as IFAMs cuja razão de compressão é 0.065 pode ser vista na Figura



Figura 7.7: Imagem reconstruída a partir da segunda imagem da Figura 7.5 utilizando as IFAMs V^t e \hat{R} .

7.8. Como se era de esperar, obteve-se a mesma imagem da Figura 7.4, no entanto de uma forma muito mais eficiente e portanto com um menor esforço computacional.

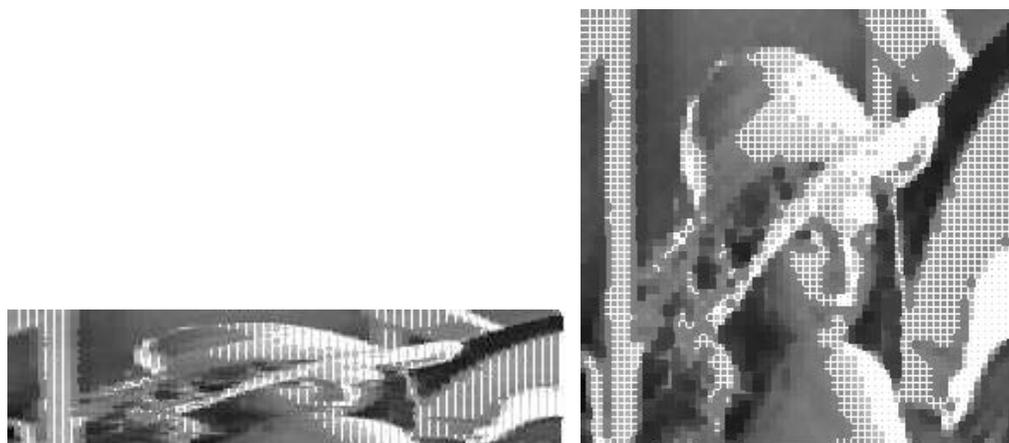


Figura 7.8: Imagem reconstruída a partir da segunda imagem da Figura 7.3 utilizando as IFAMs V^t e \hat{R} associadas a t-norma do mínimo e os conjuntos fuzzy triangulares.

7.6 Formulação do Problema de Otimização

Os melhores resultados obtidos no processo de compressão/descompressão por Nobuhara, Pedrycz e Hirota [48] foram alcançados através dos conjuntos fuzzy definidos nas Equações 7.31, 7.32, juntamente com a t-norma de Yager fixando o parâmetro $d = 1$. No entanto, levantamos a seguinte pergunta: “Qual é a melhor aproximação \hat{R} da imagem original R adotando a t-norma de Yager com o parâmetro d livre e uma razão de compressão fixa?”. Assim, gerou-se a Tabela 7.1, atribuindo outros valores à d e então calculou-se o EQM e a norma infinita entre a diferença de \hat{R} e R . Os parâmetros considerados foram $Sh = 0.05$ e a razão de compressão igual a 0.065, i.e, a imagem comprimida $G \in [0, 1]^{64 \times 64}$.

d	EQM	$\ \hat{R} - R\ _\infty$
1	0.1888	103.8646
1.5	0.1332	64.0980
2	0.1250	44.0746
2.5	0.1246	32.1365
3	0.1257	32.3635
3.5	0.1271	33.7226
4	0.1284	34.8631
4.5	0.1294	35.6268
5	0.1302	36.2466
6	0.1315	37.1636
10	0.1344	37.9977

Tabela 7.1: Tabela de erros para diferentes valores de d

A Tabela 7.1 revela que para outros valores de $d \neq 1$ encontrou-se menores erros na reconstrução. Isto nos motivou a formular um problema de otimização em busca do parâmetro d na t-norma de Yager que minimize a diferença entre \hat{R} e R . A Figura 7.9 apresenta as reconstruções para d igual a 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6 e 10. Observe que para valores



Figura 7.9: Reconstrução para diferentes valores de d

de d muito grandes, a reconstrução \hat{R} perde seus detalhes.

A solução do nosso problema de otimização é encontrar o parâmetro d na t-norma de Yager utilizada tanto no processo da compressão da imagem original $R \in [0, 1]^{n \times m}$ para a imagem comprimida $G \in [0, 1]^{p \times q}$ como também no processo de reconstrução da imagem $\hat{R} \in [0, 1]^{n \times m}$. Por isso, nesta seção será adotado uma nova notação, onde os símbolos com o sobrescrito d indica que os elementos das matrizes G e \hat{R} dependem do parâmetro d . Logo, expressa-se os elementos de $G^{[d]}$ por $g_{\xi\gamma}^d$ para $\xi = 1, \dots, p$ e $\gamma = 1, \dots, q$ e os elementos de $\hat{R}^{[d]}$ por \hat{r}_{ij}^d para $i = 1, \dots, n$ e $j = 1, \dots, m$.

Portanto, o problema de otimização concentra-se em encontrar o minimizador d que minimize a diferença entre $\hat{R}^{[d]}$ e R na t-norma de Frobenius, restrito a $d \geq 1$. Esta formulação é expressa na Equação 7.59:

$$\begin{cases} \text{minimize} & \left\| \hat{R}^{[d]} - R \right\|_F \\ \text{sujeito a} & d \geq 1 \end{cases} \quad (7.59)$$

ou

$$\begin{cases} \text{minimize} & \sqrt{\sum_{i=1}^n \sum_{j=1}^m (\hat{r}_{ij}^d - r_{ij})^2} \\ \text{sujeito a} & d \geq 1 \end{cases} \quad (7.60)$$

Relembrando que a t-norma de Yager e cada elemento \hat{r}_{ij}^d são expressos pelas equações:

$$at^d b = 1 - \min \left\{ 1, \left[(1-a)^d + (1-b)^d \right]^{1/d} \right\}, \quad d \geq 1 \quad (7.61)$$

$$\hat{r}_{ij}^d = \bigwedge_{\xi=1}^p \left\{ x_i^\xi \Rightarrow_R \left\{ \bigwedge_{\gamma=1}^q [y_j^\gamma \Rightarrow_R g_{\xi\gamma}^d] \right\} \right\} \quad (7.62)$$

ou equivalentemente,

$$\hat{r}_{ij}^d = \bigwedge_{\xi=1}^p \left\{ \bigvee \left\{ w_{ij}^{\xi,d} \in [0, 1] : x_i^\xi t^d w_{ij}^{\xi,d} \leq \left\{ \bigwedge_{\gamma=1}^q \left\{ \bigvee \left\{ z_j^{\gamma,d} \in [0, 1] : y_j^\gamma t^d z_j^{\gamma,d} \leq g_{\xi\gamma}^d \right\} \right\} \right\} \right\} \right\} \quad (7.63)$$

Finalmente, expandindo a t-norma de Yager, tem-se:

$$\begin{aligned} \hat{r}_{ij}^d &= \bigwedge_{\xi=1}^p \bigvee \left\{ w_{ij}^{\xi,d} \in [0, 1] : 1 - \min \left\{ 1, \left[(1-x_i^\xi)^d + (1-w_{ij}^{\xi,d})^d \right]^{1/d} \right\} \leq \right. \\ &\leq \left. \bigwedge_{\gamma=1}^q \bigvee \left\{ z_j^{\gamma,d} \in [0, 1] : 1 - \min \left\{ 1, \left[(1-y_j^\gamma)^d + (1-z_j^{\gamma,d})^d \right]^{1/d} \right\} \leq g_{\xi\gamma}^d \right\} \right\} \end{aligned} \quad (7.64)$$

Agora, separando a inequação:

$$\bigwedge_{\gamma=1}^q \left\{ \bigvee \left\{ z_j^{\gamma,d} \in [0, 1] : 1 - \min \left\{ 1, \left[(1 - y_j^\gamma)^d + (1 - z_j^{\gamma,d})^d \right]^{1/d} \right\} \leq g_{\xi\gamma}^d \right\} \right\} \quad (7.65)$$

Da Equação 7.64 cuja solução z_j^d é dada pela Equação 7.66:

$$z_j^d = \begin{cases} 1 - \left[(1 - g_{\xi\gamma}^d)^d - (1 - y_j^\gamma)^d \right]^{\frac{1}{d}} & \text{se } g_{\xi\gamma}^d \leq y_j^\gamma \\ 1 & \text{se } y_j^\gamma \leq g_{\xi\gamma}^d \end{cases} \quad (7.66)$$

Almejando a formulação do problema de otimização, reescreve-se z_j^d através da equação:

$$z_j^d = \alpha + \beta \left[1 - \left[\left((1 - g_{\xi\gamma}^d)^d - (1 - y_j^\gamma)^d \right)^{\frac{1}{d}} \right] \right] \quad (7.67)$$

onde $\alpha = \max \{0, \text{sign}(g_{\xi\gamma}^d - y_j^\gamma)\}$ e $\beta = 1 - \max \{0, \text{sign}(g_{\xi\gamma}^d - y_j^\gamma)\}$.

Definido z_j^d , a segunda inequação:

$$\bigwedge_{\xi=1}^p \left\{ \bigvee \left\{ w_{ij}^{\xi,d} \in [0, 1] : 1 - \min \left\{ 1, \left[(1 - x_i^\xi)^d + (1 - w_{ij}^{\xi,d})^d \right]^{1/d} \right\} \leq z_j^d \right\} \right\} \quad (7.68)$$

é expressa de forma similar:

$$\hat{r}_{ij}^d = w_{ij}^d = \phi + \varphi \left[1 - \left[\left((1 - z_j^d)^d - (1 - x_i^\xi)^d \right)^{\frac{1}{d}} \right] \right] \quad (7.69)$$

onde $\phi = \max \{0, \text{sign}(z_j^d - x_i^\xi)\}$ e $\varphi = 1 - \max \{0, \text{sign}(z_j^d - x_i^\xi)\}$.

Portanto, deseja-se resolver o problema de otimização restrito:

$$\begin{cases} \text{minimize} & \sqrt{\sum_{i=1}^n \sum_j^m \left(\phi + \varphi \left[1 - \left[\left((1 - z_j^d)^d - (1 - x_i^\xi)^d \right)^{\frac{1}{d}} \right] \right] - r_{ij} \right)^2} \\ \text{sujeito a} & d \geq 1 \end{cases} \quad (7.70)$$

onde z_j^d é dado pela Equação 7.67, $\varphi = 1 - \phi$ e $\phi = \max \{0, \text{sign}(z_j^d - x_i^\xi)\}$. O gráfico da Equação 7.70 pode ser visto na Figura 7.12 para razão de compressão igual a 0.065.

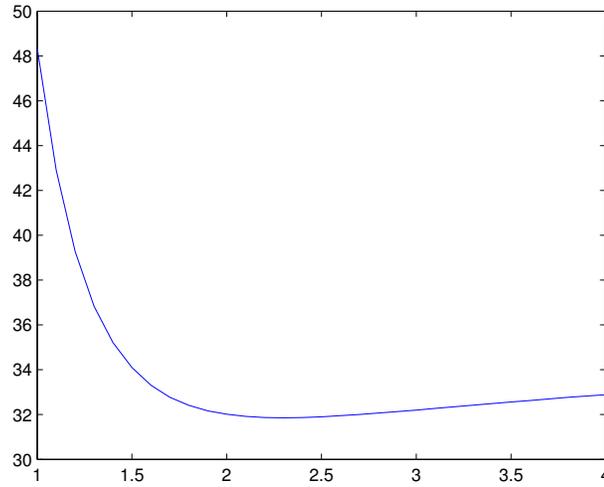


Figura 7.10: Gráfico da Equação 7.70

A implementação do algoritmo foi feita no software MATLAB, utilizando as funções da Toolbox de Otimização. Para resolver o problema descrito na Equação 7.70, recorreu-se a função FMINBND que utiliza uma combinação de busca por secções áureas e interpolação parabólica [9, 23]. Sua sintaxe é: $[d_{min}, valor] = fminbnd(fun, x_1, x_2, options)$, onde fun representa a função objetiva, x_1 e x_2 representam os extremos do intervalo no qual se procura pelo mínimo d_{min} e $valor$ retorna o valor da função objetiva em d_{min} . $Options$ permite estabelecer parâmetros para a função $fminbnd$ criada com a função $optimset$. No nosso caso, estabeleceu-se:

```
options = optimset('maxiter', 1000, 'tolfun', 1e - 4, 'tolx', 1e - 4, 'display', 'iter')
```

onde:

- maxiter: número máximo de iterações;
- tolfun: tolerância para o valor da função final;
- tolx: tolerância para a resposta final em d ;
- display: controla a quantidade de detalhes fornecidos enquanto a função está sendo executada;
- iter: mostra a iteração.



Figura 7.11: Tem-se na imagem da direita a reconstrução \hat{R} utilizando o parâmetro $d = 1$ na t-norma de Yager e a esquerda a reconstrução \hat{R} utilizando o parâmetro $d = 2.29902$ obtido no processo de otimização. Considerando a razão de compressão igual a 0.065 ($G \in [0, 1]^{64 \times 64}$)

Para a razão de compressão igual a 0.065, ou seja, para a imagem comprimida $G \in [0, 1]^{64 \times 64}$ obteve a saída à Tabela 7.2. O ótimo encontrado foi $d = 2.29902$ no intervalo $[1, 4]$ com valor $\|\hat{R} - R\|_F = 31.8514$. Já para a razão de compressão igual a 0.0156 ($G \in [0, 1]^{32 \times 32}$), o ótimo encontrado foi $d = 2.03135$ também no intervalo $[1, 4]$ com valor $\|\hat{R} - R\|_F = 51.8447$. As reconstruções para $d = 2.29902$ e $d = 2.03135$ podem ser vistas nas Figuras 7.11 e 7.12, respectivamente.

Func-count	d	$\ \hat{R} - R\ _F$	Procedure
1	2.1459	31.8894	initial
2	2.8541	32.1025	golden
3	1.7082	32.7347	golden
4	2.42275	31.8718	parabolic
5	2.32187	31.8525	parabolic
6	2.30622	31.8518	parabolic
7	2.29678	31.8515	parabolic
8	2.23915	31.8589	golden
9	2.26184	31.8549	parabolic
10	2.28343	31.8522	golden
11	2.29725	31.8515	parabolic
12	2.29893	31.8514	parabolic
13	2.30171	31.8515	golden
14	2.29921	31.8514	parabolic
15	2.29888	31.8514	parabolic
16	2.29899	31.8514	parabolic
17	2.29907	31.8514	golden
18	2.29902	31.8514	parabolic

Tabela 7.2: Saída de cada iteração para a razão de compressão igual a 0.065

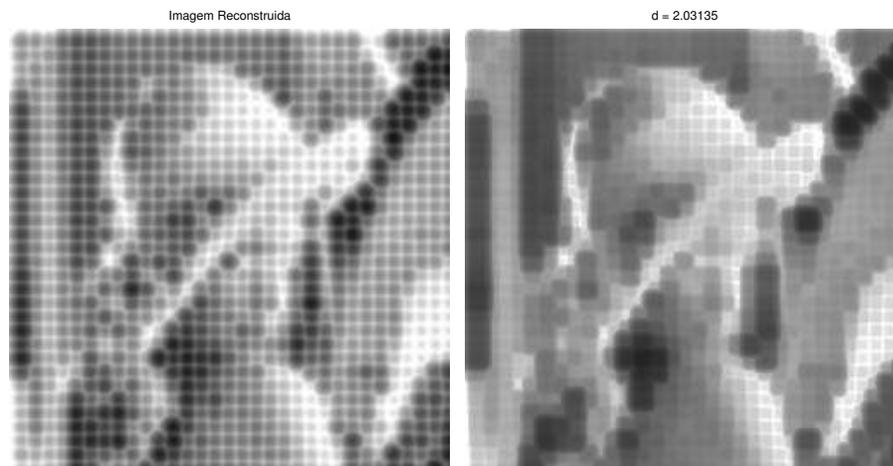


Figura 7.12: Tem-se na imagem da direita a reconstrução \hat{R} utilizando o parâmetro $d = 1$ na t-norma de Yager e a esquerda a reconstrução \hat{R} utilizando o parâmetro $d = 2.03135$ obtido no processo de otimização. Considerando a razão de compressão igual a 0.0156 ($G \in [0, 1]^{32 \times 32}$)

CAPÍTULO 8

Conclusão

Essa dissertação de mestrado discutiu as Memórias Associativas Fuzzy Implicativas (IFAMs - Implicative Fuzzy Associative Memories), uma classe das Memórias Associativas Morfológicas Fuzzy (FMAMs - Fuzzy Morphological Associative Memories), elaboradas para armazenar padrões fuzzy cujos neurônios realizam operações elementares da morfologia matemática de dilatação ou erosão. Foram estabelecidas relações envolvendo a existência de soluções para os sistemas de Equações Relacionais Fuzzy (FREs - Fuzzy Relational Equations) através da recordação perfeita das IFAMs. Também, determinou-se uma forma de escolher o melhor modelo de IFAM de Yager para uma certa aplicação, com base num algoritmo de otimização que determina a t -norma parametrizada de Yager que minimize o erro entre os padrões desejados e os padrões de saída. Para esse problema, aplicações de IFAMs de Yager foram conduzidas em predição de séries temporais. Finalmente, relacionou-se o processo de reconstrução de imagens através de FREs à teoria das IFAMs.

Precisamente, o capítulo 5 apresenta as igualdades entre as composições \max - T , \min - D e \min - I com os principais tipos de FREs. Caracterizamos os padrões de saída e os conjuntos dos pontos fixos das IFAMs. Por fim, estabelecemos condições suficientes para a existência de soluções nos sistemas de FREs em termos das recordações perfeitas das IFAMs e IFAMs duais.

No capítulo 6, apresentamos uma forma de escolher o melhor modelo de IFAM de Yager para uma certa aplicação. Como a t-norma parametrizada de Yager T_Y representa outras t-normas, implementamos um algoritmo de otimização que determina o parâmetro d na t-norma T_Y que minimize a diferença entre os padrões de recordados e os padrões desejados. Experimentos foram conduzidos em aplicações de séries temporais para estimar a mão-de-obra requerida em indústrias metalúrgicas do estado de Begal do Oeste da Índia e a vazão média mensal da usina hidrelétrica de Furnas. O previsor baseado na IFAM de Yager apresentou o menor erro quadrático médio, superando o modelo estatístico PARMA, a rede MLP, o modelo FPM-PRP e a IFAM de Lukasiewicz. O que falta é conduzir mais pesquisa sobre o processo de "fuzzificação" e "defuzzificação" empregados nas aplicações das IFAMs.

O capítulo 7 descreveu uma aplicação de IFAMs em problemas de reconstrução de imagens. Aproveitando a idéia do Capítulo 6, formulamos um problema de otimização utilizando a IFAM de Yager em busca da t-norma de Yager que retornasse a melhor aproximação da imagem original no processo de compressão e descompressão. Os resultados obtidos não foram tão bons quanto aos métodos clássicos DCT e VQ, mas superou o método proposto em [48] com um menor EQM. Uma boa direção para estudos futuros é investigar a distribuição dos conjuntos fuzzy \mathbf{x}^ξ e \mathbf{y}^γ para melhorar a qualidade da imagem descomprimida.

Referências Bibliográficas

- [1] ANDERSON, J. A. A simple neural network generating interactive memory. *Mathematical Biosciences* 14 (1972), 3197–220.
- [2] BANDLER, W., AND KOHOUT, L. Fuzzy power sets and fuzzy implication operators. *Fuzzy Sets and Systems*, 1 (July 1980), 13–30.
- [3] BANON, G. J. F., AND BARRERA, J. Decomposition of mappings between complete lattices by mathematical morphology: Part 1 - general lattices. *Signal Processing* 30 (February 1993), 299–327.
- [4] BARROS, L., AND BASSANEZI, R. C. *Tópicos de Lógica Fuzzy e Biomatemática*. Coleção IMECC - Textos Didáticos - N°5, 2006.
- [5] BIRKHOFF, G. *Lattice Theory*. 3 ed. American Mathematical Society, Providence, 1993.
- [6] BLANCO, A., DELGADO, M., AND REQUENA, I. Identification of fuzzy relational equations by fuzzy neural networks. *Fuzzy Sets and Systems* 71 (1995), 215–226.
- [7] BLOCH, I., AND MAITRE, H. Fuzzy mathematical morphologies - a comparative study. *Pattern Recognition*,.
- [8] BRAGA NETO, U., AND GOUTSIAS, J. Supremal multiscale signal analysis. *SIAM Journal of Mathematical Analysis* 36, 1 (2004), 94–120.

- [9] BRENT, R. P. *Algorithms for Minimization without Derivatives*. Prentice Hall, Englewood Cliffs, New Jersey, 1973.
- [10] CHENG, Q., AND FAN, Z. T. The stability problem for fuzzy bidirectional associative memories. *Fuzzy Sets and Systems* 132, 1 (November 2002), 83–90.
- [11] CHIU, S. Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems* 2, 3 (September 2004).
- [12] CHOUDHURY, J. P., SARKAR, B., AND MUKHERJEE, S. K. Forecasting of engineering manpower through fuzzy associative memory neural network with arima: a comparative study. *Neurocomputing* 47 (2002), 241–257.
- [13] CHUNG, F., AND LEE, T. On fuzzy associative memory with multiple-rule storage capacity. *IEEE Transactions on Fuzzy Systems* 4, 3 (August 1996), 375–384.
- [14] CIARAMELLA, A., TAGLIAFERRI, R., AND PEDRYCZ, W. The genetic development of ordinal sums. *Fuzzy Sets and Systems, 2004. In Press, Accept Manuscript, Available online*.
- [15] CUNINGHAME-GREEN, R. Gray-scale morphological associative memories. *Minimax Algebra: Lecture Notes in Economics and Mathematical Systems* 166. New York: Springer-Verlag (1979).
- [16] DAVIDSON, J. L., AND RITTER, G. X. Recursion and feedback in image algebra. In *SPIEs 19 AIPR Workshop on Image Understanding in the 90s, Proceedings of SPIE, Mclean* (October 1990).
- [17] DAVIDSON, J. L., AND RITTER, G. X. A theory of morphological neural networks. In *Digital Optical Computing II 1215 of Proceedings of SPIE* (1990), 378–388.
- [18] DE BAETS, B. Fuzzification of set inclusion: theory and applications. In *Uncertainty Analysis in Engineering and Science: Fuzzy Logic, Statistics, and Neural Network Ap-*

- proach*, B. M. Ayyub and M. M. Gupta, Eds. Kluwer Academic Publishers, Norwell (1997), 53–67.
- [19] DEDEKIND, R. Über Zerlegungen von Zahlen durch ihre grössten gemeinsamen Teiler. in Richard Dedekind. Gesammelte mathematische Werke. R. Fricke, E. Noether, and Ö. Ore, Eds. Chelsea Publishing Corporation, 1969 II (Appeared in 1987), 103–148.
- [20] DENG, T., AND HEIJMANS, H. Grey-scale morphology based on fuzzy logic. *Journal of Mathematical Imaging and Vision* 2, 16 (March 2002), 155–171.
- [21] DI NOLA, A., PEDRYCZ, W., SESSA, S., AND SANCHEZ, E. *Fuzzy relation equations and their applications to knowledge engineering: An overview*. Kluwer Academic Publishers, Dordrecht, Holland, 1989.
- [22] DI NOLA, A., PEDRYCZ, W., SESSA, S., AND SANCHEZ, E. Fuzzy relation equations theory as a basis of fuzzy modelling: An overview. *Fuzzy Sets and Systems*, 3 (April 1991), 415–429.
- [23] FORSYTHE, G., MALCOLM, M., AND MOLER, C. *Computer Methods for Mathematical Computations*. Prentice Hall, 1976.
- [24] FULLER, R. *Introduction to Neuro-Fuzzy Systems*. Springer-Verlag, New York, 2000.
- [25] HADWIGER, H. *Vorlesungen Über Inhalt, Oberfläche und Isoperimetrie*. Springer-Verlag, Berlin, 1957.
- [26] HASSUON, M. H. *Dynamic associative neural memories. In Associative Neural Memories: Theory and Implementation*. Oxford University Press, Oxford, U.K., 1993.
- [27] HAYKIN, S. *Princípios e práticas. 2º ed.* Bookman, Porto Alegre, 2001.
- [28] HEBB, D. *The Organization of Behavior*. John Wiley Sons, New York, 1949.
- [29] HEIJMANS, H. *Morphological Image Operators*. Academic Press, New York, 1994.

- [30] HIROTA, K., AND PEDRYCZ, W. Fuzzy relational compression. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics* 29, 3 (June 1999), 443–453.
- [31] HOPFIELD, J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* 79 (April 1982), 2554–2558.
- [32] JUNBO, F., FAN, J., AND YAN, S. A learning rule for fuzzy associative memories. *In Proceedings of the IEEE International Joint Conference on Neural Network 7* (June 1994), 4273–4277.
- [33] KLIR, G. J., AND YUAN, B. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, Upper Saddle River, N. Y., 1995.
- [34] KOHONEN, T. Correlation matrix memories. *IEEE Trans. Computers C-21* (1972), 353–359.
- [35] KOHONEN, T., AND RUOHONEN, M. Representation of associated data by matrix operators. *IEEE Trans. Computers C-22* (1973), 701–702.
- [36] KOSKO, B. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, Englewood Cliffs, N.J., 1992.
- [37] KOSKO, B. *Fuzzy Engineering*. Prentice Hall, Upper Saddle River, N.J., 1997.
- [38] LIU, P. The fuzzy associative memory of max-min fuzzy neural networks with threshold. *Fuzzy Set and Systems* 107 (1999), 147–157.
- [39] MAGALHÃES, M. Redes neurais, metodologias de agrupamento e combinação de previsores aplicados a previsão de vazões naturais. *Master's thesis, State University of Campinas* (2004).
- [40] MAGALHÃES, M., BALLINI, R., GONÇALVES, R., AND GOMIDE, F. Predictive fuzzy clustering model for natural streamflow forecasting. *In Proceedings of the IEEE International Conference on Fuzzy Systems (Budapest, Hungary, 2004)*, 390–394.

- [41] MARAGOS, P. Lattice image processing: A unification of morphological and fuzzy algebraic systems. *Journal of Mathematical Imaging and Vision* 22, 2-3 (May 2005), 333–353.
- [42] MATHERON, G. *Random Sets and Integral Geometry*. Wiley, New York, 1975.
- [43] MCCULLOCH, W., AND PITTS, W. A logical calculus of the ideas immanent em nervous activity. *Bolletín of Mathematical Biophysics* 5 (1943), 115–133.
- [44] MINKOWSKI, H. *Gesammelte Abhandlungen*. Teubner Verlag, Leipzig-Berlin, 1911.
- [45] MIYAKOSHI, M., AND SHIMBO, M. Lower solutions of systems of fuzzy equations. *Fuzzy Sets and Systems* 19(1) (May 1986), 37–46.
- [46] NACHTEGAEL, M., AND KERRE, E. Connections between binary, gray-scale and fuzzy mathematical morphologies. *Fuzzy Sets and Systems* 124, 1 (November 2001), 73–85.
- [47] NAKANO, K. Associatron: A model of associative memory. *IEEE Trans. System, Man, and Cybernetics SMC-2* (1972), 380–388.
- [48] NOBUHARA, H., PEDRYCZ, W., AND HIROTA, K. Fast solving method of fuzzy relational equation and its application to lossy image compression/reconstruction. *IEEE Transactions on Fuzzy Systems* 8, 3 (June 2000), 325–334.
- [49] PAO, Y. H. *Adaptive Pattern Recognition and Neural Networks*. Addison-Wesley, Reading, MA, 1989.
- [50] PEDRYCZ, W., AND GOMIDE, F. *An Introduction to Fuzzy Sets: Analysis and Design*. MIT Press, Complex Adaptive Systems Series, 1998.
- [51] PERFILIEVA, I. Fuzzy function as an approximate solution to a system of fuzzy relation equations. *Fuzzy Sets and Systems*, 3 (November 2004), 363–383.
- [52] PERFILIEVA, I., AND GOTTWALD, S. Solvability and approximate solvability of fuzzy relation equations. *International Journal of General Systems*, 4 (July 2003), 361–372.

- [53] RITTER, G. X., AND SUSSNER, P. An introduction to morphological neural networks. *In Proceedings of the 13th International Conference on Pattern Recognition (Vienna, Austria, 1996)*, 709–717.
- [54] RITTER, G. X., AND SUSSNER, P. Morphological neural networks. *In Intelligent Systems: A Semiotic Perspective; Proceedings of the 1996 International Multidisciplinary Conference (Gaithersburg, Maryland, 1997)*.
- [55] RITTER, G. X., SUSSNER, P., AND DE LEON, J. L. D. Morphological associative memories. *IEEE Transaction on Neural Networks* 9, 2 (March 1998), 281–293.
- [56] RONSE, C. Why mathematical morphology needs complete lattices. *Signal Processing* 21, 2 (October 1990), 129–154.
- [57] ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65 (1958), 386–408.
- [58] SANCHEZ, E. Resolution of composite fuzzy relation equations. *Inform and Contro* (1976), 38–47.
- [59] SERRA, J. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.
- [60] SERRA, J. *Image Analysis and Mathematical Morphology*. Volume 2: Theoretical Advances. Academic Press, New York, 1988.
- [61] SERRA, J. Image analysis and mathematical morphology. *Signal Processing* 2 (October 1990), 129–154.
- [62] SINHA, D., AND DOUGHERTY, R. Fuzzification of set inclusion: theory and applications. *Fuzzy Sets and Systems*, 1 (April 1993), 15–42.
- [63] SINHA, S., AND DOUGHERTY, E. R. A general axiomatic theory of intrinsically fuzzy mathematical morphologies. *IEEE Transactions on Fuzzy Systems* 3, 4 (November 1995), 389–403.

- [64] SOILLE, P. *Morphological Image Analysis*. Springer Verlag, Berlin, 1999.
- [65] STEINBUCH, K. Die Lernmatrix. *Kybernetik 1* (1961), 36–45.
- [66] STERNBEG, S. Grayscale morphology. *Computer Vision, Graphics and Image Processing 35* (1986), 333–355.
- [67] SUSSNER, P. Perceptrons. in *Encyclopedia for Electrical and Electronics Engineering, John Wiley and Sons 16* (1999), 44–59.
- [68] SUSSNER, P. Technical report. *Institute of Mathematics, Statistics and Scientific Computing, State University of Campinas* (January 2008).
- [69] SUSSNER, P., AND VALLE, M. Gray-scale morphological associative memories. *IEEE Transactions on Neural Networks 17*, 3 (May 2006), 559–570.
- [70] SUSSNER, P., AND VALLE, M. Recall of patterns using morphological and certain fuzzy morphological associative memories with applications in classification and prediction. *Proceedings of the IEEE World Congress on Computational Intelligence , Vancouver, Canada* (July 2006), 209–216.
- [71] SUSSNER, P., AND VALLE, M. E. *Fuzzy Associative Memories and Their Relationship to Mathematical Morphology*. Handbook of Granular Computing, Series: Studies in Computational Intelligence, Pedrycz W., Skowron A., and Kreinovich V. (Eds), Chapter 33, Wiley, September 2008.
- [72] SUSSNER, P., AND VALLE, M. E. Implicative fuzzy associative memories. *IEEE Transactions on Fuzzy Systems*, 6 (December 2006), 793–807.
- [73] SUSSNER, P., AND VALLE, M. E. Classification of fuzzy mathematical morphologies based on concepts of inclusion measure and duality. *Journal of Mathematical Imaging and Vision 32*, 2 (October 2008), 139–159.

- [74] SUSSNER, P., AND VALLE, M. E. A brief account of the relations between gray-scale mathematical morphologies. *In Proceedings of the Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)* (Natal, Brazil, October 2005), 79–86.
- [75] TAYLOR, W. Electrical simulation of some nervous system functional activities. *Information Theory* 3 (1956), 314–328.
- [76] TEROL-VILLALOBOS, I., AND MENDIOLA-SANTIBAÑES, J. Transformations with reconstruction criteria: Image segmentation and filtering. *Computational Imaging and Vision. Springer-Verlag, Dordrecht* 30 (2005), 75–84.
- [77] VALLE, M. E. *Matlab Source Code*. Disponível em <http://www.ime.unicamp.br/mevalle>.
- [78] VALLE, M. E. Fundamentos e aplicações de memórias associativas morfológicas nebulosas, dissertação de doutorado. *IMECC-Unicamp, Campinas-SP* (2007).
- [79] VALLE, M. E., AND SUSSNER, P. Storage and recall capabilities of fuzzy morphological associative memories with adjunction-based learning. *Submitted for publication, December 2007*.
- [80] VALLE, M. E., AND SUSSNER, P. A general framework for fuzzy morphological associative memories. *Fuzzy Sets and Systems* 159, 7 (April 2008), 747–768.
- [81] YAGER, R. R., AND FILEV, D. P. *Essentials of Fuzzy Modeling and Control*. John Wiley and Sons, New York, 1994.
- [82] ZADEH, L. A. Fuzzy sets. *Information and Control* 8, 3 (1965), 338–353.