



ESTEVÃO ESMI LAUREANO

Θ -FAMs: MEMÓRIAS ASSOCIATIVAS FUZZY BASEADAS EM
FUNÇÕES Θ

CAMPINAS
2014



UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística
e Computação Científica

ESTEVÃO ESMI LAUREANO

Θ -FAMs: MEMÓRIAS ASSOCIATIVAS FUZZY BASEADAS EM
FUNÇÕES Θ

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em matemática aplicada.

Orientador: Peter Sussner

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO ESTEVÃO ESMI LAUREANO, E ORIENTADA PELO PROF. DR. PETER SUSSNER.

Assinatura do Orientador

CAMPINAS
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Maria Fabiana Bezerra Muller - CRB 8/6162

Es53f Esmi, Estevão, 1982-
'*Theta*'-FAMs : memórias associativas fuzzy baseadas em funções-'*theta*' /
Estevão Esmi Laureano. – Campinas, SP : [s.n.], 2014.

Orientador: Peter Sussner.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de
Matemática, Estatística e Computação Científica.

1. Memória associativa. 2. Conjuntos difusos. 3. Redes neurais
(Computação). 4. Classificação. 5. Visão de robô. 6. Sistemas de reconhecimento
de padrões. I. Sussner, Peter, 1961-. II. Universidade Estadual de Campinas.
Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: '*Theta*'-FAMs : fuzzy associative memories based on functions-'*theta*'

Palavras-chave em inglês:

Associative memory

Fuzzy sets

Neural networks (Computer science)

Classification

Robot vision

Pattern recognition systems

Área de concentração: Matemática Aplicada

Titulação: Doutor em Matemática Aplicada

Banca examinadora:

Peter Sussner [Orientador]

Roberto Andreani

Laécio Carvalho de Barros

Sandra Aparecida Sandri

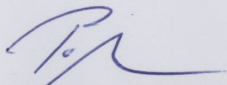
Benjamín René Callejas Bedregal

Data de defesa: 30-04-2014

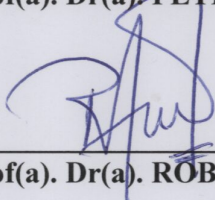
Programa de Pós-Graduação: Matemática Aplicada

Tese de Doutorado defendida em 30 de abril de 2014 e aprovada

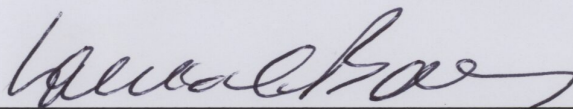
Pela Banca Examinadora composta pelos Profs. Drs.



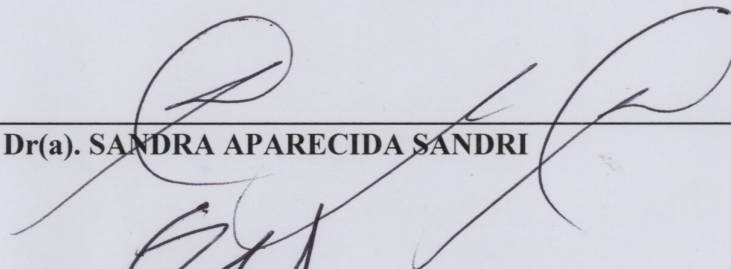
Prof(a). Dr(a). PETER SUSSNER



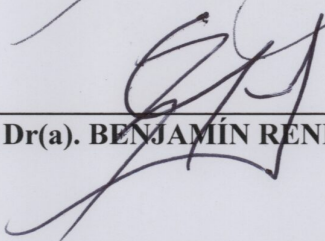
Prof(a). Dr(a). ROBERTO ANDREANI



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



Prof(a). Dr(a). SANDRA APARECIDA SANDRI



Prof(a). Dr(a). BENJAMÍN RENÉ CALLEJAS BEDREGAL

Abstract

Most *fuzzy associative memories* in the literature correspond to neural networks with a single layer of weights that distributively contains the information about the associations to be stored. The main applications of these types of associative memory can be found in fuzzy rule-base systems. In contrast, we present in this thesis the class of Θ -fuzzy associative memories (Θ -FAMs) that represent fuzzy neural networks with two layers. Particular cases of Θ -FAMs, called (dual) S-FAMs and E-FAMs, are based on fuzzy subsethood and equivalence measures. We provide theoretical results concerning the storage capability and error correction capability of Θ -FAMs. Furthermore, we introduce a general training algorithm for Θ -FAM that is guaranteed to converge in a finite numbers of iterations. We also proposed another alternative training algorithm for a certain type of E-FAM that not only adjusts the parameters of the corresponding network but also automatically determines its topology. We compare the classification rates produced by Θ -FAMs with that ones of some well-known classifiers in several benchmark classification problems that are available on the internet. Finally, we successful apply Θ -FAM approach to a problem of vision-based self-localization in mobile robotics.

Keywords: Fuzzy associative memory, fuzzy (parametrized) subsethood fuzzy, fuzzy (parametrized) equivalence fuzzy, selection of fundamental memories, classification, robotics, vision-based self-localization.

Resumo

Muitas das *memórias associativas fuzzy* (FAMs) da literatura correspondem a redes neurais com uma única camada de pesos que armazenam de forma distributiva as informações das associações desejadas. As principais aplicações deste tipo de memórias associativas são encontradas em sistemas baseados em regras *fuzzy*. Nesta tese introduzimos a classe de memórias associativas *fuzzy- Θ* (Θ -FAMs) que, em contraste com estes outros modelos, representam redes neurais fuzzy com duas camadas. Caso particulares de Θ -FAMs, denominadas S-FAMs (duais) e E-FAMs, são baseadas em medidas de *subsethood* e equivalência fuzzy. Resultados gerais sobre a capacidade de armazenamento e a capacidade de correção de erro das Θ -FAMs também foram providenciados. Adicionalmente, introduzimos um algoritmo geral de treinamento para Θ -FAM cuja convergência é sempre garantida. Apresentamos ainda um algoritmo alternativo para treinamento de uma certa classe de E-FAMs que além de ajustar os seus parâmetros também determina automaticamente a topologia da rede. Finalmente, comparamos as taxas de classificação produzidas pelas Θ -FAMs com alguns classificadores bem conhecidos em diversos problemas de classificação disponíveis na

internet. Além disso, aplicamos com sucesso as Θ -FAMs em um problema de auto-localização de robô móvel baseado em visão.

Palavras-chave: Memória associativa fuzzy, medida de subethood (parametrizada) fuzzy, medida de equivalência (parametrizada) fuzzy, seleção de memórias fundamentais, classificação, robótica, auto-localização baseada em visão.

Sumário

| | |
|--|-----------|
| Dedicatória | ix |
| Agradecimentos | x |
| 1 Introdução | 1 |
| 2 Conceitos Matemáticos | 4 |
| 2.1 Alguns Conceitos Básicos e Notações | 4 |
| 2.2 Medidas de <i>Subsethood</i> e Equivalência <i>Fuzzy</i> | 6 |
| 3 Memória Associativa <i>Fuzzy</i> - Θ | 16 |
| 3.1 Alguns Resultados Teóricos Referentes as Θ -FAMs | 19 |
| 3.2 Algoritmo de Treinamento Geral para Θ -FAMs | 19 |
| 3.3 Algoritmo de Treinamento para TE-FAM do Tipo $\mathcal{E}_{W,v}$ | 27 |
| 3.3.1 Seleção de Memórias Fundamentais | 28 |
| 3.3.2 Seleção de Parâmetros para TE-FAM $\mathcal{E}_{W,v}$ | 32 |
| 4 Resultados Experimentais | 34 |
| 4.1 Alguns de Problemas de Classificação de Referência | 36 |
| 4.2 Um Problema de Auto-localização de Robô Baseado em Visão | 39 |
| 5 Conclusão | 44 |
| Referências | 46 |
| A Alguns Resultados Experimentais Usando Θ-FAMs | 55 |
| A.1 Alguns Problemas de Classificação de Referência | 56 |
| A.2 Problema de Auto-localização de Robô Baseado em Visão | 65 |
| Índice Remissivo | 69 |

Dedicado a minha esposa Damaris e aos meus pais Brasil e Tânia.

Agradecimentos

Este trabalho foi apoiado pela FAPESP sob o processo no. 2009/16284-2.

Eu gostaria de agradecer a minha esposa, meus pais e meus amigos por me encorajarem e me apoiarem neste últimos quatro anos. Em especial, não posso deixar de mencionar Damaris, Tânia, Brasil, Renata, Willian, Gustavo, Marcos, etc. que estiveram sempre presente me agraciando com suas amizades.

Agradeço a todos os professores e pesquisadores que diretamente ou indiretamente me ensinaram e me forneceram o conhecimento necessário para concluir esta tese de doutorado. Em particular, agradeço ao Prof. Dr. Peter Sussner por me orientar neste projeto.

Lista de Ilustrações

| | | |
|-----|---|----|
| 3.1 | Topologia de uma Θ -FAM. | 17 |
| 3.2 | Topologia de uma TE-FAM. | 28 |
| 3.3 | Uma aplicação do Algoritmo 3.2 em \mathcal{T}^* | 30 |
| 4.1 | Posições de referência para o problema de auto-localização. | 41 |

Lista de Tabelas

| | | |
|------|--|----|
| 2.1 | Alguns exemplos de t-normas e implicações <i>fuzzy</i> . | 6 |
| 3.1 | Resultados dos Passos 4-6 do Algoritmo 3.2 | 30 |
| 3.2 | Resultados dos Passos 9-13 do Algoritmo 3.2. | 31 |
| 4.1 | Descrição dos conjuntos de dados. | 36 |
| 4.2 | Taxa de classificação em 15 problemas | 38 |
| 4.3 | Tempo de treinamento das TE-FAMs | 39 |
| 4.4 | Número de imagens por percurso e posição. | 40 |
| 4.5 | Resultados para o problema de auto-localização de robô | 43 |
| A.1 | Descrição dos conjuntos de dados. | 56 |
| A.2 | Resultados produzidos pelas Θ -FAMs para Appendicitis. | 57 |
| A.3 | Resultados produzidos pelas Θ -FAMs para Cleveland. | 58 |
| A.4 | Resultados produzidos pelas Θ -FAMs para Crx. | 58 |
| A.5 | Resultados produzidos pelas Θ -FAMs para Ecoli. | 59 |
| A.6 | Resultados produzidos pelas Θ -FAMs para Glass. | 59 |
| A.7 | Resultados produzidos pelas Θ -FAMs para Heart. | 60 |
| A.8 | Resultados produzidos pelas Θ -FAMs para Iris. | 60 |
| A.9 | Resultados produzidos pelas Θ -FAMs para Monk. | 61 |
| A.10 | Resultados produzidos pelas Θ -FAMs para Movementlibras | 61 |
| A.11 | Resultados produzidos pelas Θ -FAMs para Pima. | 62 |
| A.12 | Resultados produzidos pelas Θ -FAMs para Spectfheart. | 62 |
| A.13 | Resultados produzidos pelas Θ -FAMs para Sonar. | 63 |
| A.14 | Resultados produzidos pelas Θ -FAMs para Vowel. | 63 |
| A.15 | Resultados produzidos pelas Θ -FAMs para Wdbc. | 64 |
| A.16 | Resultados produzidos pelas Θ -FAMs para Wine. | 64 |
| A.17 | Número de imagens por percurso e posição. | 65 |
| A.18 | Os resultados produzidos pelas S-FAMs usando \mathcal{M}_e . | 66 |
| A.19 | Os resultados produzidos pelas S-FAMs duais usando \mathcal{M}_e . | 66 |
| A.20 | Os resultados produzidos pelas E-FAMs usando \mathcal{M}_e . | 67 |
| A.21 | Os resultados produzidos pelas S-FAMs usando \mathcal{M}_o . | 67 |
| A.22 | Os resultados produzidos pelas S-FAMs duais usando \mathcal{M}_o . | 68 |
| A.23 | Os resultados produzidos pelas E-FAMs usando \mathcal{M}_o . | 68 |

Lista de Abreviaturas e Siglas

AM Memória associativa.

FAM Memória associativa *fuzzy*.

FNN Rede neural *fuzzy*.

ANN Rede neural artificial.

MNN Rede neural morfológica.

MM Morfológia matemática.

FMM Morfológia matemática *fuzzy*.

FMAM Memórias associativa morfológica *fuzzy*.

Θ -FAM Memórias associativa *fuzzy*- Θ .

S-FAM Memórias associativa *fuzzy* de *subsethood*.

TS-FAM Memórias associativa *fuzzy* de *subsethood* ajustável.

E-FAM Memórias associativa *fuzzy* de equivalência.

TE-FAM Memórias associativa *fuzzy* de equivalência ajustável.

Lista de Símbolos

$\mathbb{R}_{\pm\infty}$ Representa a conjunto dos reais estendidos.

$\mathcal{P}(X)$ Representa a coleção dos subconjuntos de X .

$\mathcal{F}(X)$ Representa a coleção dos conjuntos *fuzzy* em um universo arbitrário X .

\mathbb{L}^X Representa a coleção dos conjuntos *L-fuzzy* em um universo arbitrário X .

\circ_T Representa a composição sup-T.

\mathcal{M} Representa o conjunto de memórias associativas.

p Denota o número de memórias fundamentais (*i.e.* $|\mathcal{M}|$).

\mathcal{T} Representa o conjunto de treinamento.

k Denota o número de padrões de treinamento (*i.e.* $|\mathcal{T}|$).

\mathcal{O}_v Representa uma Θ -FAM.

\mathcal{E} Representa uma E-FAM.

\mathcal{S} Representa uma S-FAM.

\mathcal{D} Representa uma S-FAM dual.

S_K Denota a medida de *subsethood* de Kosko.

S_W Denota a medida de *subsethood* de Willmot.

Lista de Algoritmos

| | | |
|-----|--|----|
| 3.1 | Algoritmo Geral de Treinamento para Θ -FAM | 24 |
| 3.2 | Extração de Memórias Fundamentais do Conjunto de Treinamento | 29 |

Capítulo 1

Introdução

Memórias associativas fuzzy (FAMs) pertencem à classe de *redes neurais fuzzy* (FNNs). Uma FNN é uma *rede neural artificial* (ANN) cujas entradas e/ou os pesos das conexões correspondem a valores *fuzzy* [12]. Muitas memórias associativas *fuzzy* bem conhecidas da literatura já foram classificadas como *memórias associativas morfológicas fuzzy* (FMAMs) [107], incluindo as FAMs max-min e max-product de Kosko [64], a FAM generalizada de Chung e Lee [17], a FAM de Junbo [47], a FAM max-min com limiar de Liu [70], a memória associativa bidirecional lógica *fuzzy* de Belohlávek [9] e as memórias associativas implicativas *fuzzy* [104]. Todos estes modelos de FAMs representam redes neurais *fuzzy* totalmente conectadas sem camadas ocultas [12, 102].

Uma classe mais geral de ANN que inclui os modelos de FMAMs é chamada de *redes neurais morfológicas* (MNNs). Em um sentido algébrico de reticulados, [6, 87, 91], uma MNN é definida como uma ANN que executa uma das operações elementares da *morfologia matemática* (MM) em cada neurônio ou nó de processamento, eventualmente seguida da aplicação de um função de ativação [101]. As operações elementares da MM consistem das operações de erosão, dilatação, anti-erosão e anti-dilatação [6].

MNNs e redes neurais híbridas do tipo morfológico/posto/linear [81] foram aplicadas com sucesso em uma variedade de problemas, tais como classificação [24, 25, 27, 96, 97], reconhecimento automático de alvo [32, 42, 55, 56, 111], predição [4, 3, 99, 102, 105], reconhecimento de caracteres [81], visão computacional [84, 85], e análise de imagens hiperespectrais [37, 36, 86].

Muitos outros modelos de ANN da literatura são explícita ou implicitamente baseados em idéias que surgem em análise e processamento de imagens usando “elementos estruturantes” (SEs) [74, 90]. Por exemplo, temos as redes neurais de regularização e pesos-compartilhados morfológicos [32, 57, 58], modelos de neurocomputação em reticulados *fuzzy* [48, 49, 51, 52, 82]. Nesta tese, apresentaremos as FAMs *subsethood*, *subsethood* dual, e equivalência (ponderadas) [24, 27, 98], ou simplesmente S-FAM, S-FAM dual, and E-FAM (ponderadas), que são FNN de duas camadas baseadas nos conceitos de medidas de *subsethood* e equivalência *fuzzy* que também encontram inspiração na morfologia matemática *fuzzy* (FMM) [93]. Relembre que, em um sentido geométrico, podemos definir operadores elementares da MM em termos de medidas de inclusão (ou *subsethood*) e intersecção [79, 101]. Especificamente, o operador elementar de erosão *fuzzy* é definido em termos do grau de inclusão (ou grau de *subsethood*) de objetos (*i.e.*, versões transladadas do chamado elemento estruturante) em outro objeto (dado pela imagem em questão) [94]. Adicionalmente,

certas medidas de equivalência empregadas em E-FAMs podem ser relacionadas com transformadas hit-or-miss *fuzzy*. Portanto, as S-FAMs (duais) e as E-FAMs (ponderadas) podem ser vistas como redes neurais morfológicas *fuzzy*, mais especificamente como *memórias associativas morfológicas fuzzy* [107].

Subsethood é um conceito da teoria de conjuntos *fuzzy* que visa avaliar o quanto um conjunto *fuzzy* está contido em outro. Kitainik [59] e Sinha e Dougherty [92] foram alguns dos primeiros pesquisadores a apresentar uma definição axiomática para a medida de *subsethood* (ou inclusão). Subsequentemente, Young criticou os nove axiomas que uma medida de *subsethood* deveria satisfazer, segundo Sinha e Dougherty, e propôs uma definição mais simples e intuitiva consistindo de apenas três axiomas [114]. Tanto a definição axiomática de Sinha e Dougherty quanto a de Young serviram como base para muitas outras formulações e caracterizações de medidas de *subsethood*, em particular a definição axiomática introduzida por Fan *et al.* [29] que será adotada nesta tese. As chamadas medidas de **-subsethood* representam uma ampla classe de medidas de *subsethood* que satisfazem os axiomas de Fan *et al.*. Além da ligeira modificação da definição de Young, Fan *et al.* mostraram como combinar duas medidas de *subsethood* para produzir uma nova medida desta classe. Assim como tanto outros autores, tais como Bandler e Kohout [5], Willmott [110] e Zhang e Zhang [117], Fan *et al.* também providenciaram fórmulas para gerar medidas de *subsethood* a partir de implicações *fuzzy*. Zhang e Zhang definiram o que eles chamaram de medida de *S*-inclusão substituindo o terceiro axioma de Young por um par de axiomas de Sinha & Dougherty. Entretanto, tal modificação produz condições mais restritivas do que o terceiro axioma de Young [117].

Bustince *et al.* estudaram extensivamente tanto a noção de medida de *subsethood* quanto a de **-subsethood* e introduziram uma nova noção chamada medida de DI-*subsethood* [15] que cumpre todos os axiomas de Young. Neste mesmo artigo, eles observaram que a definição de medida de **-subsethood* generaliza a definição de medida de *subsethood* segundo Young. Nesta tese, adotaremos a definição de Fan *et al.* que é mais geral do que todas as definições mencionadas anteriormente. Em particular, a medida de **-subsethood* pertence à classe de medidas de *subsethood* proposta por Fan *et al.* [29].

Kaburlasos *et al.* apresentaram uma abordagem às “medidas de inclusão *fuzzy*” em reticulados completos [10, 49], que inclui a classe de conjuntos *fuzzy* como um caso especial. Estas medidas de inclusão *fuzzy* foram usadas para derivar vários modelos de *fuzzy lattice reasoning* (FLR) que têm sido aplicados com sucesso em várias tarefas de classificação e agrupamento (ou *clustering*) [49, 50, 54].

Extensões do conceito de medida de inclusão para estruturas mais gerais podem também ser encontradas na teoria de conjuntos \mathbb{L} -*fuzzy* [33], ou, mais especificamente, na morfologia matemática \mathbb{L} -*fuzzy* [100], onde erosões e dilatações \mathbb{L} -*fuzzy* podem ser definidas em termos de medidas de inclusão e intersecção \mathbb{L} -*fuzzy*. No caso especial onde $\mathbb{L} = [0, 1]$, obtemos abordagens à morfologia matemática *fuzzy* [101]. Certos tipos de medidas de inclusão *fuzzy*, amplamente utilizadas na FMM [79, 101], satisfazem a propriedade de herança, isto é, sua restrição para a classe de conjuntos *crisp* (ou clássicos) coincide com a inclusão clássica.

Para certas tarefas em classificação de imagens e visão computacional [95, 98], a propriedade de herança não é desejada. Portanto, nesta tese, desenvolveremos FAMs baseadas em outros tipos de medidas de *subsethood* [98] tal como a de Kosko [63] que satisfaz tanto os axiomas de Fan *et*

al. quanto os axiomas de Young [114]. De fato, Young encontrou inspiração no trabalho de Kosko para derivar sua definição axiomática de medida de *subsethood*. Note que as medidas de *subsethood* segundo Fan *et al.* não necessariamente satisfazem a propriedade de herança. Além de medidas de *subsethood*, também desenvolveremos FAMs baseadas no conceito de medidas de equivalência ou similaridade, que é um outro conceito importante na teoria de conjuntos *fuzzy* estritamente relacionado com a noção de *subsethood* [16, 116]. Os modelos resultantes foram aplicados com sucesso em problemas de auto-localização de robô, reconstrução de imagens e identificação automática de locutor [27, 98].

Esta tese está organizada como se segue. Iniciamos por apresentar os conceitos matemáticos básicos utilizados no decorrer deste texto. Neste mesmo capítulo, dedicamos um seção inteira para discutir medidas de *subsethood* e equivalência *fuzzy*, propondo novas fórmulas gerais para se obter tais medidas. As principais contribuições originais deste trabalho estão enunciados em forma de teoremas.

No Capítulo 3, introduzimos uma família de modelos FAMs chamada Θ -FAMs que correspondem às FNNs de duas camadas. Este modelo é determinado por funções Θ^ξ que são aplicadas no ξ -ésimo neurônio oculto. Medidas de *subsethood* e equivalência podem ser empregadas para produzir funções Θ^ξ e, neste caso, obtemos os modelos S-FAMs (duais) e E-FAMs mencionados acima [24, 27, 98]. Estes modelos são equipados com pesos v_ξ na camada oculta que ponderam a contribuição do ξ -ésimo neurônio oculto. Na Seção 3.2, introduzimos um algoritmo específico para treinar os pesos v_ξ de um modelo de Θ -FAM arbitrário. Provamos que este algoritmo converge em número finito de iterações e, sob algumas condições fracas, atinge um mínimo local da função objetiva. Além desse algoritmo, também propomos um algoritmo alternativo para treinar os pesos de certos modelos de E-FAMs chamado *Algoritmo TE*. Este último algoritmo, além de ajustar os parâmetros da rede, também determina automaticamente, em um primeiro estágio, a topologia da rede através de um algoritmo de seleção de memórias fundamentais discutido na Subseção 3.3.2.

O Capítulo 4 refere-se à aplicação dos modelos Θ -FAMs em vários problemas de classificação de referência disponíveis na internet [2] e um problema de auto-localização baseado em visão da literatura [108, 109]. Finalmente, terminamos apresentando algumas conclusões e considerações finais.

Como produto direto desta tese, confeccionamos dois artigos de revistas que compilam a maioria dos principais resultados aqui enunciados [25, 24]. Além desses dois artigos de revistas, também submetemos outros dois artigos de congresso internacional com árbitro divulgando resultados teóricos parciais e aplicações variadas dos modelos Θ -FAMs [26, 27]. Todos estes trabalhos tiveram a colaboração de pesquisadores renomados na comunidade científica da área de sistemas *fuzzy*.

Capítulo 2

Conceitos Matemáticos

Este capítulo é organizado como se segue. Começamos por apresentar os conceitos básicos já bem estabelecidos na literatura que serão utilizados no decorrer do texto. Na seção subsequente serão introduzidas as noções de medidas de *subsethood* e equivalência baseadas em diversos trabalhos [15, 16, 29, 30, 31].

2.1 Alguns Conceitos Básicos e Notações

Um par (P, \leq) consistindo de um conjunto não vazio P e uma relação binária reflexiva, anti-simétrica e transitiva “ \leq ” é dito ser um *conjunto parcialmente ordenado* ou *poset* [40]. Se a relação de ordem \leq claramente surge do contexto, então, nos referimos ao poset em questão simplesmente usando o símbolo P ao invés de (P, \leq) . Se $X \subseteq P$, então, um elemento $l \in P$ é dito ser um *limitante inferior* de X se $l \leq x$ para todo $x \in X$. Similarmente, $u \in P$ é dito ser um *limitante superior* de X se $x \leq u$ para todo $x \in X$. O *ínfimo* de $X \subseteq P$, denotado pelo símbolo $\bigwedge X$, é definido como o maior limitante inferior de X . Similarmente, o *supremo* de $X \subseteq P$, denotado pelo símbolo $\bigvee X$, é definido como o menor limitante superior de X . Se $X = \{x, y\}$, então, alternativamente, escrevemos $x \vee y$ e $x \wedge y$ ao invés de $\bigvee X$ e $\bigwedge X$, respectivamente.

Um conjunto parcialmente ordenado \mathbb{L} é um *reticulado* se todo subconjunto finito não vazio de \mathbb{L} possui um ínfimo e um supremo em \mathbb{L} [10]. Se, adicionalmente, tivermos $\bigwedge \mathbb{L} \in \mathbb{L}$ e $\bigvee \mathbb{L} \in \mathbb{L}$, então, o reticulado \mathbb{L} é dito ser *limitado*. Neste caso, os símbolos $0_{\mathbb{L}}$ e $1_{\mathbb{L}}$ denotam, respectivamente, $\bigwedge \mathbb{L}$ e $\bigvee \mathbb{L}$. Um reticulado \mathbb{L} é *completo* se todo subconjunto de \mathbb{L} possui um ínfimo e um supremo em \mathbb{L} . Em particular, todo reticulado completo é limitado [10].

Um reticulado (\mathbb{L}, \leq) é dito uma *corrente* se tivermos $x \leq y$ ou $y \leq x$ para todo $x, y \in \mathbb{L}$. Neste caso, nos referimos a \leq como uma *ordem total*. Se \mathbb{L} é adicionalmente um reticulado completo, então, dizemos que \mathbb{L} é uma *corrente completa*. O intervalo unitário $[0, 1]$ e o conjunto dos números reais estendidos $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{+\infty, -\infty\}$ com a ordem total convencional representam exemplos de correntes completas. Assim, o supremo de $\{r, s\} \subset [0, 1]$, é s se e somente se $r \leq s$. Similarmente, $r \vee s = r$ se e somente se $s \leq r$. A estrutura de reticulado de $[0, 1]$ é central na teoria de conjuntos *fuzzy* uma vez que muitas das suas regras tais como inclusão, composições relacionais *fuzzy*, etc. dependem tanto da relação de ordem quanto dos operadores ínfimo e supremo de $[0, 1]$.

Em 1965, Lotfi A. Zadeh estendeu a noção clássica de um conjunto, também referido como um conjunto *crisp*, introduzindo conjuntos *fuzzy* [115]. Um conjunto *fuzzy* A consiste de um conjunto X , chamado de universo, junto com uma função de pertinência $\mu_A : X \rightarrow [0, 1]$ que produz o grau de pertinência $\mu_A(x) \in [0, 1]$ para cada $x \in X$. O símbolo $\mathcal{F}(X)$ denota a coleção de todos os conjuntos *fuzzy* do universo X .

Matematicamente falando, um conjunto *fuzzy* A em um universo X pode ser identificado com sua função de pertinência μ_A e pode ser visto simplesmente como uma função do universo X para o intervalo unitário $[0, 1]$. Se X é finito, digamos $X = \{x_1, \dots, x_n\}$, então, $\mathcal{F}(X)$ pode ser identificado com $[0, 1]^n$ via a bijeção que mapeia $A \in \mathcal{F}(X)$ para o vetor $(A(x_1), \dots, A(x_n))^t \in [0, 1]^n$. Aqui o símbolo sobrescrito “ t ” denota o operador transposto da álgebra matricial. A noção de conjunto *crisp* surge se A (ou, mais precisamente, μ_A) apenas adota valores em $\{0, 1\}$. Mais precisamente, cada $Y \subseteq X$ corresponde ao conjunto *fuzzy* de X tal que o grau de pertinência de $x \in X$ é igual a 1 se e somente se $x \in Y$. Em particular, X e \emptyset denotam os conjuntos *fuzzy* dados, respectivamente, por $\mu_X(x) = 1$ e $\mu_\emptyset(x) = 0$ para todo $x \in X$. Aqui, utilizaremos o símbolo $\mathcal{P}(X)$ para denotar o conjunto dos subconjuntos *crisp* de X .

Uma vez que o conceito clássico de uma relação é dado por um subconjunto de $X \times Y$, onde X e Y são universos arbitrários e, portanto, pode ser visto como uma função $X \times Y \rightarrow \{0, 1\}$, podemos estender este conceito clássico para o caso *fuzzy*. Assim, uma relação *fuzzy* é dada por uma função $R : X \times Y \rightarrow [0, 1]$, onde $R(x, y)$ pode ser interpretado como o grau de relacionamento entre x e y [20]. Em outras palavras, uma relação *fuzzy* em $X \times Y$ nada mais é do que um conjunto *fuzzy* em $X \times Y$ [7].

Podemos produzir uma família de operações entre conjuntos *fuzzy* a partir de operações simples conhecidas como conectores *fuzzy* que nada mais são do que extensões dos operadores lógicos booleanos para o conjunto $[0, 1]$ [61, 80]. Por exemplo, uma conjunção *fuzzy* C é uma função $[0, 1] \times [0, 1] \rightarrow [0, 1]$ crescente em ambos os argumentos tal que $C(0, 0) = C(1, 0) = C(0, 1) = 0$ e $C(1, 1) = 1$. Além disso, se uma conjunção *fuzzy* T é comutativa, associativa e satisfaz $T(1, a) = a$ para todo $a \in [0, 1]$, então, T é dita ser uma t-norma. Definições de disjunção *fuzzy* e t-conorma são dadas de maneira similar. Uma implicação *fuzzy* é uma função $I : [0, 1] \times [0, 1] \rightarrow [0, 1]$ que é decrescente no primeiro argumento, crescente no segundo argumento e que satisfaz $I(0, 0) = I(0, 1) = I(1, 1) = 1$ e $I(1, 0) = 0$. Podemos utilizar t-normas para produzir certos tipos de implicações *fuzzy* conhecidas como *residuais*. Especificamente, seja T uma t-norma, o operador $I_T : [0, 1] \times [0, 1] \rightarrow [0, 1]$ dado por $I_T(a, b) = \bigvee \{c \in [0, 1] \mid T(a, c) \leq b\}$ para todo $a, b \in [0, 1]$ é a implicação residual *fuzzy* (ou R-implicação) com respeito a T . A Tabela 2.1 exibe alguns exemplos de t-normas e suas correspondentes implicações residuais *fuzzy* amplamente encontradas na literatura:

Duas relações *fuzzy* $R \in \mathcal{F}(X \times Y)$ e $S \in \mathcal{F}(Y \times Z)$, onde X, Y e Z são universos arbitrários, podem ser combinadas usando uma composição relacional *fuzzy* que produz uma relação *fuzzy* no universo $X \times Z$. Nesta tese, iremos apenas empregar composições relacionais *fuzzy* do tipo sup- T definida na Equação (2.1.2) abaixo. Formalmente, se T é uma t-norma, então, a composição sup- T de $R \in \mathcal{F}(X \times Y)$ e $S \in \mathcal{F}(Y \times Z)$ produz a seguinte relação *fuzzy* $R \circ_T S \in \mathcal{F}(X \times Z)$:

$$R \circ_T S(x, z) = \bigvee_{y \in Y} T(R(x, y), S(y, z)) \quad \forall (x, z) \in X \times Z. \quad (2.1.1)$$

Tabela 2.1: Alguns exemplos de t-normas e implicações *fuzzy*.

| t-norma | | implicação <i>fuzzy</i> | |
|--------------|----------------------------------|-------------------------|--|
| Mínimo: | $T_M(x, y) = x \wedge y$ | Gödel: | $I_M(x, y) = \begin{cases} 1, & x \leq y, \\ y, & x > y \end{cases}$ |
| Produto: | $T_P(x, y) = xy$ | Goguen: | $I_P(x, y) = \begin{cases} 1, & x \leq y, \\ y/x, & x > y \end{cases}$ |
| Lukasiewicz: | $T_L(x, y) = 0 \vee (x + y - 1)$ | Lukasiewicz: | $I_L(x, y) = 1 \wedge (y - x + 1)$. |

Similarmente, a composição sup- T de $R \in \mathcal{F}(X \times Y)$ e $S \in \mathcal{F}(Y)$ produz o seguinte conjunto *fuzzy* $R \circ_T S \in \mathcal{F}(X)$:

$$R \circ_T S(x) = \bigvee_{y \in Y} T(R(x, y), S(y)) \quad \forall x \in X. \quad (2.1.2)$$

Em [33], Goguen generalizou as noções de conjuntos *crisp* e *fuzzy* definindo um conjunto \mathbb{L} -*fuzzy* como uma função $X \rightarrow \mathbb{L}$, onde \mathbb{L} é usualmente exigido ser um conjunto parcialmente ordenado, um reticulado, ou um reticulado completo. A seguir, iremos rever rapidamente estes conceitos.

Se \mathbb{L}_i são reticulados, reticulados limitados ou reticulados completos para todo $i = 1, \dots, n$, então, $\mathbb{L} = \mathbb{L}_1 \times \dots \times \mathbb{L}_n$ também representa, respectivamente, um reticulado, um reticulado limitado ou um reticulado completo e é chamado um produto de reticulado (de $\mathbb{L}_1, \dots, \mathbb{L}_n$). Se $\mathbb{L}_i = \mathbb{M}$ para todo $i = 1, \dots, n$, então, o símbolo \mathbb{M}^n denota $\mathbb{L}_1 \times \dots \times \mathbb{L}_n$. A fim de simplificar nossa notação, denotaremos $(\mathbb{R}_{\pm\infty})^n$ usando o símbolo $\mathbb{R}_{\pm\infty}^n$.

\mathbb{L}^X denota a classe de conjuntos \mathbb{L} -*fuzzy* em um universo arbitrário X , *i.e.*, a classe das funções $X \rightarrow \mathbb{L}$. Se \mathbb{L} representa um poset, então, a ordem parcial \leq em \mathbb{L} induz uma ordem parcial em \mathbb{L}^X que é denotada usando o símbolo \subseteq e é definida como se segue para todo $f, g \in \mathbb{L}^X$:

$$f \subseteq g \Leftrightarrow f(x) \leq g(x) \quad \forall x \in X. \quad (2.1.3)$$

A estrutura de reticulado em \mathbb{L} induz a mesma estrutura de reticulado em \mathbb{L}^X . Por exemplo, se \mathbb{L} constitui um reticulado limitado ou completo, então, \mathbb{L}^X também constitui, respectivamente, um reticulado limitado ou completo [10]. Considere, por exemplo, o reticulado completo $[0, 1]$. A classe de conjuntos *fuzzy* no universo X , $\mathcal{F}(X)$, é dada pelo reticulado completo $[0, 1]^X$ e os operadores de ínfimo e supremo em $\mathcal{F}(X)$ coincidem, respectivamente, com as operações usuais de intersecção e união entre conjuntos *fuzzy*. Especificamente, sejam $A, B \in \mathcal{F}(X)$, a intersecção e a união de A e B são os conjuntos *fuzzy* $A \cap B \in \mathcal{F}(X)$ e $A \cup B \in \mathcal{F}(X)$ dados, respectivamente, por

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \quad \text{e} \quad \mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} \quad \forall x \in X.$$

2.2 Medidas de *Subsethood* e Equivalência *Fuzzy*

Neste trabalho faremos uso de conjunções e implicações *fuzzy* para produzir medidas de *subsethood* e equivalência *fuzzy* [5, 16, 15, 29, 30, 31, 112, 116].

O conceito de *subsethood* refere-se ao grau com que um conjunto *fuzzy* está contido em outro conjunto *fuzzy*. A fim de ilustrar esta questão, considere os seguintes subconjuntos de \mathbb{R} :

$$A = \{x > 0\}, B = \{x \geq 0\} \text{ e } C = \{x < 0\}. \quad (2.2.1)$$

Apesar de $B \not\subseteq A$ e $C \not\subseteq A$, parece razoável afirmar que B está contido em A com um grau maior do que C está em A . O conceito de *subsethood* emerge com o objetivo de mensurar este tipo de relação entre dois conjuntos *fuzzy*. Existem várias definições axiomáticas para medidas de *subsethood* na literatura. Neste trabalho, adotaremos a definição proposta por Fan *et al.* [29] pelo fato dela incluir todas as outras medidas de *subsethood* mencionadas na Introdução. Além disso, esta definição é restrita o suficiente para gerar modelos Θ -FAM tal como explicado no capítulo subsequente.

Definição 2.2.1 (*Subsethood*). Uma função $S : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$, onde $X \neq \emptyset$, é dita ser uma *medida de subsethood* se S satisfaz as seguintes propriedades para todo $A, B, C \in \mathcal{F}(X)$:

(P1) $S(A, B) = 1$ se $A \subseteq B$;

(P2) $S(X, \emptyset) = 0$;

(P3) Se $A \subseteq B \subseteq C$, então, $S(C, A) \leq S(B, A)$ e $S(C, A) \leq S(C, B)$.

Alguns autores referem-se à noção de medida de *subsethood* como medida de *inclusão fuzzy*. Para fazer uma distinção clara entre estes dois termos técnicos, preferimos reservar o último termo para funções $Inc_{\mathcal{F}} : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$ que satisfazem a propriedade de herança, *i.e.*, que estendem a medida de inclusão crisp: $Inc_{\mathcal{F}}(A, B) = 1$ se $A \subseteq B$, caso contrário, $Inc_{\mathcal{F}}(A, B) = 0$, para todo $A, B \in \mathcal{P}(X)$ [5, 79, 101]. O seguinte resultado [5, 29] implica que podemos usar implicações *fuzzy* para produzir operadores que são ao mesmo tempo medidas de *subsethood* e inclusão *fuzzy*:

Proposição 2.2.2 ([5, 29]). Seja I uma implicação *fuzzy* que satisfaz $I(a, b) = 1$ para todo $a \leq b$. O operador $Inc_{\mathcal{F}} : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$ dado por

$$Inc_{\mathcal{F}}(A, B) = \inf_{x \in X} I(\mu_A(x), \mu_B(x)), \quad \forall A, B \in \mathcal{F}(X), \quad (2.2.2)$$

é tanto uma medida de inclusão *fuzzy* quanto uma medida de *subsethood*.

Vale a pena notar que toda R-implicação derivada de uma t-norma T satisfaz a propriedade $I(a, b) = 1$ para todo $a \leq b$ uma vez que $T(1, a) = a \leq b$ [13]. Além disso, uma medida de *subsethood* pode não ser uma medida de inclusão e vice-versa. Nesta tese, focaremos em medidas de *subsethood*. Usualmente, medidas de *subsethood* são construídas em termos de uma função de agregação apropriada atuando sob implicações (veja a Proposição 5 de [15]). Aqui, adotaremos um ponto de vista diferente, invertendo o papel da implicação e da função de agregação dada pela função v abaixo, cuja definição coincide com a definição de Klir e Folger para uma função de agregação em um universo finito [60].

Teorema 2.2.3. Seja I uma implicação *fuzzy* que satisfaz $I(a, b) = 1$ para todo $a \leq b$ e seja $v : \mathcal{F}(X) \rightarrow [0, 1]$ uma função crescente tal que

$$v(\emptyset) = 0 \text{ e } v(X) = 1. \quad (2.2.3)$$

Suponham que $H, F : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow \mathcal{F}(X)$ são operadores que satisfazem as seguintes condições:

- (i) $H(X, \emptyset) = X$ e $F(X, \emptyset) = \emptyset$,
- (ii) $H(A, B) \subseteq F(A, B)$ se $A \subseteq B$,
- (iii) $H(C, A) \supseteq (H(B, A) \cup H(C, B))$ e $F(C, A) \subseteq (F(B, A) \cap F(C, B))$ para todo $A, B, C \in \mathcal{F}(X)$ tal que $A \subseteq B \subseteq C$.

O seguinte operador $S : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$ representa uma medida de *subsethood* em $\mathcal{F}(X)$:

$$S(A, B) = I(v(H(A, B)), v(F(A, B))), \forall A, B \in \mathcal{F}(X). \quad (2.2.4)$$

Demonstração. Seja $A, B, C \in \mathcal{F}(X)$. Se $A \subseteq B$, então, temos que $H(A, B) \subseteq F(A, B)$ e consequentemente $v(H(A, B)) \leq v(F(A, B))$. A última afirmação implica que

$$S(A, B) = I(v(H(A, B)), v(F(A, B))) = 1.$$

Além disso, temos que $S(X, \emptyset) = 0$:

$$S(X, \emptyset) = I(v(H(X, \emptyset)), v(F(X, \emptyset))) = I(v(X), v(\emptyset)) = I(1, 0) = 0.$$

Finalmente, a Propriedade (iii) implica que $H(C, A) \supseteq H(B, A)$ e $F(C, A) \subseteq F(B, A)$ se $A \subseteq B \subseteq C$. Desde que v é crescente e I é decrescente no primeiro argumento e crescente no segundo argumento, obtemos as seguintes desigualdades:

$$S(C, A) = I(v(H(C, A)), v(F(C, A))) \quad (2.2.5)$$

$$\leq I(v(H(B, A)), v(F(C, A))) \quad (2.2.6)$$

$$\leq I(v(H(B, A)), v(F(B, A))) = S(B, A). \quad (2.2.7)$$

A desigualdade $S(C, A) \leq S(B, A)$ pode ser demonstrada de maneira análoga. \square

Corolário 2.2.4. Seja I uma medida de implicação *fuzzy* que satisfaz $I(a, b) = 1$ para todo $a \leq b$ e seja $v : \mathcal{F}(X) \rightarrow [0, 1]$ uma função crescente tal que

$$v(\emptyset) = 0 \text{ e } v(X) = 1. \quad (2.2.8)$$

Os seguintes operadores S^\cap e S^\cup representam medidas de *subsethood*:

$$(a) \quad S^\cap(A, B) = I(v(A), v(A \cap B)) \quad (2.2.9)$$

$$(b) \quad S^\cup(A, B) = I(v(A \cup B), v(B)) \quad (2.2.10)$$

Demonstração. A Equação (2.2.9) pode ser obtida aplicando o Teorema 2.2.3 com F e H dadas por $F(A, B) = A \cap B$ e $H(A, B) = A$ para todo $A, B \in \mathcal{F}(X)$. De fato, estas funções satisfazem as seguintes propriedades:

- (i) $H(X, \emptyset) = X$ e $F(X, \emptyset) = X \cap \emptyset = \emptyset$;
- (ii) Se $A \subseteq B$, então, $H(A, B) = A \subseteq A = A \cap B = F(A, B)$;
- (iii) Se $A \subseteq B \subseteq C$, então, $H(C, A) = C \supseteq C = (B \cup C) = (H(B, A) \cup H(C, B))$ e $F(C, A) = A \cap B = A \subseteq A = ((B \cap A) \cap (C \cap B)) = (F(B, A) \cap F(C, B))$.

Similarmente, uma aplicação do Teorema 2.2.3 com as funções \hat{F} e \hat{H} dadas por $\hat{F}(A, B) = B$ e $\hat{H}(A, B) = A \cup B$ para todo $A, B \in \mathcal{F}(X)$ produz a Equação (2.2.10). \square

Seja X um universo finito, digamos $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\}$. Considere a implicação de Goguen e a função $v_M : \mathcal{F}(X) \rightarrow [0, 1]$ crescente dada por $v_M(A) = \frac{\sum_{i=1}^k \mu_A(\mathbf{x}^i)}{k}$. Neste caso, as Equações (2.2.9) e (2.2.10) correspondem às medidas de *subsethood* de Kosko [63] e Willmot [110] que denotaremos respectivamente pelos símbolos S_K e S_W . Uma infinidade de medidas de *subsethood* surge se considerarmos a implicação de Goguen e a seguinte função $v_\rho : \mathcal{F}(X) \rightarrow [0, 1]$, onde $\rho \in (0, +\infty)$, nas Equações (2.2.9) e (2.2.10):

$$v_\rho(A) = \sum_{i=1}^k \frac{1 - \cos(\pi[\mu_A(\mathbf{x}^i)]^\rho)}{2k}, \quad \forall A \in \mathcal{F}(X). \quad (2.2.11)$$

Denotaremos as medidas de *subsethood* resultantes usando os símbolos S_ρ^\cap e S_ρ^\cup , respectivamente.

Note que tanto a medida de Kosko quanto S_ρ^\cap não representam medidas de DI-*subsethood* em $[0, 1]^n$ para $n > 1$ [15, 29], isto é, medidas de *subsethood* decrescentes no primeiro argumento e crescentes no segundo argumento. Especificamente, para $n = 2$, $A = (1, 0)$, $B = (1, 0.5)$, e $C = (0, 1)$, temos que

$$S_\rho^\cap(A, C) = 0 < S_\rho^\cap(B, C) = \frac{1 - \cos(0.5^\rho \pi)}{3 - \cos(0.5^\rho \pi)}.$$

Assim, S_ρ^\cap não é decrescente no primeiro argumento e portanto não produz uma medida de DI-*subsethood*.

Para um universo finito X , uma outra interessante classe de medidas de *subsethood* surgem do Corolário 2.2.4 usando *funções de sobreposições*. Lembre que uma função de sobreposição é uma função de agregação $G_O : [0, 1]^2 \rightarrow [0, 1]$ contínua e simétrica tal que $G_O(x, y) = 0$ se e somente se $xy = 0$ e $G_O(x, y) = 1$ se e somente se $xy = 1$ [14]. Temos que $G_O(x, y) = \min(x, y)$ ou $G_O(x, y) = (xy)^\iota$ com $\iota > 0$ são exemplos de funções de sobreposições. Note que se definirmos

$$v_O(A) = \sum_{i=1}^k \frac{G_O(\mu_A(x_i), 1)}{k}, \quad \forall A \in \mathcal{F}(X), \quad (2.2.12)$$

então, $v_O : \mathcal{F}(X) \rightarrow [0, 1]$ é crescente e satisfaz a Equação (2.2.8). Em particular, $v_O = v_M$ se $G_O(x, y) = \min(x, y)$.

Seja X um conjunto arbitrário não vazio. Para $\alpha \in [0, 1]$ e $A \in \mathcal{F}(X)$, usaremos o símbolo $[A]^\alpha$ para denotar o α -nível do conjunto *fuzzy* A , *i.e.*, $[A]^\alpha = \{\mathbf{x} \in X : \mu_A(\mathbf{x}) \geq \alpha\} \subseteq X$. Seja $\rho : \mathcal{P}(X) \rightarrow [0, 1]$ uma função crescente tal que $\rho(\emptyset) = 0$ e $\rho(X) = 1$. O operador $v_S : \mathcal{F}(X) \rightarrow [0, 1]$ dado por

$$v_S(A) = \sup_{\alpha \in [0, 1]} \alpha \wedge \rho([A]^\alpha), \forall A \in \mathcal{F}(X) \quad (2.2.13)$$

satisfaz os requerimentos do Corolário 2.2.4. De fato, desde que $[X]^\alpha = X$ e $[\emptyset]^\alpha = \emptyset$ para todo $\alpha \in (0, 1]$, temos que $v_S(X) = 1$ e $v_S(\emptyset) = 0$. Adicionalmente, se $A \subseteq B$ para $A, B \in \mathcal{F}(X)$, então, temos que $v_S(A) \leq v_S(B)$, pois ρ é crescente e $[A]^\alpha \subseteq [B]^\alpha$ para todo $\alpha \in [0, 1]$. Pelo Corolário 2.2.4, podemos combinar v_S com certas implicações *fuzzy* para produzir medidas de *subsethood* em $\mathcal{F}(X)$. Note que o operador v_S corresponde a integral de Sugeno [77] sob X de A com respeito à medida ρ que é definida na σ -álgebra $\mathcal{P}(X)$.

Enquanto medidas de *subsethood* e inclusão *fuzzy* constituem abordagens à *fuzzificação* da relação de inclusão, medidas de equivalência (ou similaridade) *fuzzy* e igualdade de índices podem ser vistas como abordagens à *fuzzificação* da igualdade [16, 21, 61, 112, 116].

Estendendo a definição de Fodor e Roubens de equivalência como um operador binário em $[0, 1]$ [30, 31], Bustince *et al.* definiram a noção de medida de equivalência como uma função $\mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$ como se segue [16].

Definição 2.2.5. Uma medidade de equivalência em $\mathcal{F}(X)$ é uma função $E : \mathcal{F}(X)^2 \rightarrow [0, 1]$ que satisfaz as seguintes condições:

- E1) $E(A, B) = E(B, A)$ para todo $A, B \in \mathcal{F}(X)$;
- E2) $E(\emptyset, X) = 0$;
- E3) $E(A, A) = 1$ para todo $A \in \mathcal{F}(X)$;
- E4) $E(A, C) \leq E(A, B)$ e $E(A, C) \leq E(B, C)$ se $A \subseteq B \subseteq C$.

Na literatura da teoria de conjuntos *fuzzy*, podemos encontrar dois outros conceitos muito relacionados com o de equivalência (*fuzzy*) chamados de medida de similaridade *fuzzy* e índice de igualdade [16, 22, 28, 61, 112, 116, 117]. De fato, estes termos técnicos ou são usados intecaladamente ou suas definições diferem ligeiramente da Definição 2.2.5 (muitas dessas diferenças foram comentadas em [16]). Em nossa opinião, a principal propriedade que distingue os conceitos de similaridade do de equivalência é a propriedade de que a similaridade entre um conjunto *crisp* com seu complemento deve ser zero [112, 117]. Esta propriedade é essencial para definir medidas de entropia ou *fuzziness* em termos de medida de similaridade uma vez que é razoável assumir que o grau de entropia de um conjunto *crisp* seja igual a zero [16, 112, 116, 117]. Medidas de equivalência e similaridade *fuzzy* foram extensivamente estudadas nos últimos anos e muitos autores propuseram fórmulas para produzir medidas de equivalência na classe dos conjuntos *fuzzy* de um universo X [16, 28, 112, 117]. Em [88], um mapeamento $E : \mathbb{R}^2 \rightarrow [0, 1]$ satisfazendo as propriedades E1, E2 e E3 da Definição 2.2.5 foi chamado de uma *relação fuzzy compatível com a ordem* (OCFR) e a propriedade E4 foi chamada de “compatibilidade com ordem total”.

Dependendo do tipo de aplicação, propriedades adicionais podem ser desejadas. A próxima definição coincide com a definição de Zeng e Li para uma medida de similaridade. Já que existem várias definições diferentes de medidas de similaridade na literatura [16, 22, 61, 112, 116], preferimos falar de uma medida de *equivalência forte*. Note que a Definição 2.2.6 surge da Definição 2.2.5 pela imposição de uma condição adicional.

Definição 2.2.6. Uma medida de equivalência E é dita ser *forte* se $E(A, B) = 1 \Leftrightarrow A = B$.

A medida de equivalência forte definida no Exemplo 2.2.7 foi empregada para recuperação de consistência em sistemas de base de regras *fuzzy* baseados em implicações [88] e em sistemas de raciocínio baseados em casos (CBR) *fuzzy* [73].

Exemplo 2.2.7. Para cada $\lambda \in (0, 1]$, a função $E_\lambda : [0, 1] \times [0, 1] \rightarrow [0, 1]$ dada por

$$E_\lambda(x, y) = \max\left(0, 1 - \frac{|x - y|}{\lambda}\right) \quad (2.2.14)$$

produz uma medida de equivalência forte em $[0, 1]$.

Uma infinidade de medidas de equivalência em $\mathcal{F}(X)$ podem ser geradas a partir de medidas de *subsethood* ou inclusão *fuzzy* [16, 24, 27, 116, 117]. O conceito de medida de equivalência também pode ser conectado aos conceitos de medidas de entropia e distância [28, 112, 117]. O próximo teorema é uma contribuição original desta tese e estabelece uma fórmula geral para se produzir medidas de equivalência baseadas em medidas de *subsethood* e certas conjunções *fuzzy*.

Teorema 2.2.8. Sejam C_o e S uma conjunção *fuzzy* comutativa e uma medida de *subsethood*, respectivamente. Suponha que $H, F : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow \mathcal{F}(X)$ são operadores que satisfazem as seguintes propriedades para todo $A, B, C \in \mathcal{F}(X)$:

- (i) $H(X, \emptyset) = X$ e $F(X, \emptyset) = \emptyset$,
- (ii) $H(A, B) \subseteq F(A, B)$ e $F(B, A) \subseteq H(B, A)$ se $A \subseteq B$,
- (iii) $H(C, A) \supseteq H(C, B) \cup H(B, A)$ e $F(C, A) \subseteq F(C, B) \cap F(B, A)$ se $A \subseteq B \subseteq C$.

O operador $E^* : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$, dado por

$$E^*(A, B) = C_o(S(H(A, B), F(A, B)), S(H(B, A), F(B, A))), \quad \forall A, B \in \mathcal{F}(X), \quad (2.2.15)$$

representa uma medida de equivalência.

Demonstração. Como C_o é comutativo, temos que E^* é também comutativo. A propriedade (i) e as definições de S e C_o implicam que $E^*(X, \emptyset) = C_o(S(X, \emptyset), S(\emptyset, X)) = C_o(0, 1) = 0$. Para todo $A \in \mathcal{F}(X)$, temos que $H(A, A) = F(A, A)$ uma vez que $A \subseteq A$. Usando esta última observação e as propriedades de S e C_o , obtemos que $E^*(A, A) = C_o(1, 1) = 1$.

Seja $A, B \in \mathcal{F}(X)$ tais que $A \subseteq B$, a propriedade (ii) garante que $H(A, B) \subseteq F(A, B)$ e portanto $S(H(A, B), F(A, B)) = 1$. Além disso, se $A \subseteq B \subseteq C$ para $A, B, C \in \mathcal{F}(X)$, então, as propriedades (ii) e (iii) de H e F implicam que $F(C, A) \subseteq F(C, B) \subseteq H(C, B) \subseteq H(C, A)$. Estas

observações junto com os fatos de que C_o é crescente em ambos os argumentos e que S satisfaz o axioma P3 da Definição 2.2.1 produzem as seguintes desigualdades:

$$\begin{aligned} E^*(C, A) &= C_o(S(H(C, A), F(C, A)), 1) \\ &\leq C_o(S(H(C, A), F(C, B)), 1) \\ &\leq C_o(S(H(C, B), F(C, B)), 1) \\ &= E^*(C, B). \end{aligned}$$

Logo, $E^*(C, A) \leq E^*(C, B)$.

Similarmente, as hipóteses (ii) e (iii) implicam que $F(C, A) \subseteq F(B, A) \subseteq H(B, A) \subseteq H(C, A)$ para $A \subseteq B \subseteq C$. Novamente, podemos usar as propriedades de C e S para obter

$$\begin{aligned} E^*(C, A) &= C_o(S(H(C, A), F(C, A)), 1) \\ &\leq C_o(S(H(C, A), F(B, A)), 1) \\ &\leq C_o(S(H(B, A), F(B, A)), 1) \\ &= E^*(B, A). \end{aligned}$$

Portanto, $E^*(C, A) \leq E^*(B, A)$. □

Dois tipos de expressões algébricas para produzir medidas de equivalência que são usualmente encontradas na literatura surgem como um caso particular do Teorema 2.2.8. Em particular, medidas de equivalência são definidas baseada nas seguintes expressões

$$T(S(A, B), S(B, A)), \forall A, B \in \mathcal{F}(X), \quad (2.2.16)$$

e

$$S(A \cup B, A \cap B), \forall A, B \in \mathcal{F}(X), \quad (2.2.17)$$

onde S e T denotam, respectivamente, uma certa medida de *subsethood* e uma t-norma [16, 116, 117]. O próximo corolário garante que de fato as expressões 2.2.16 e 2.2.17 obedecem os axiomas da Definição 2.2.5.

Corolário 2.2.9. Sejam C_o e S uma conjunção *fuzzy* comutativa e uma medida de *subsethood*, respectivamente. Os operadores $E^S, \bar{E}^S : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$ dados por

$$E^S(A, B) = C_o(S(A, B), S(B, A)) \forall A, B \in \mathcal{F}(X), \quad (2.2.18)$$

e

$$\bar{E}^S(A, B) = \begin{cases} C_o(1, S(A \cup B, A \cap B)) & \text{se } A \subseteq B \text{ ou } B \subseteq A, \\ C_o(S(A \cup B, A \cap B), S(A \cup B, A \cap B)) & \text{se } A \not\subseteq B \text{ e } B \not\subseteq A, \end{cases} \quad \forall A, B \in \mathcal{F}(X), \quad (2.2.19)$$

representam medidas de equivalência.

Demonstração. A Equação (2.2.18) pode ser obtida aplicando-se o Teorema 2.2.8 com F e H dados por $F(A, B) = B$ e $H(A, B) = A$ para todo $A, B \in \mathcal{F}(X)$. Note que de fato estas funções satisfazem as seguintes propriedades:

(i) $H(X, \emptyset) = X$ e $F(X, \emptyset) = \emptyset$;

(ii) se $A \subseteq B$, então, $H(A, B) = A \subseteq B = F(A, B)$ e $F(B, A) = A \subseteq B = H(B, A)$;

(iii) se $A \subseteq B \subseteq C$, então,

$$H(C, A) = C \supseteq C = C \cup B = H(C, B) \cup H(B, A)$$

e

$$F(C, A) = A \subseteq A = B \cap A = F(C, B) \cap F(B, A).$$

Uma aplicação do Teorema 2.2.8 com as funções \bar{F} e \bar{H} definidas como

$$\bar{F}(A, B) = \begin{cases} A \cup B & \text{se } A \subseteq B, \\ A \cap B & \text{se } A \not\subseteq B. \end{cases} \quad \text{e} \quad \bar{H}(A, B) = \begin{cases} A \cap B & \text{se } A \subseteq B, \\ A \cup B & \text{se } A \not\subseteq B. \end{cases}$$

para todo $A, B \in \mathcal{F}(X)$ produz a Equação (2.2.19). Além disso, \bar{F} e \bar{H} cumprem os requisitos (i), (ii) e (iii):

(i) $\bar{H}(X, \emptyset) = X \cup \emptyset = X$ e $\bar{F}(X, \emptyset) = \emptyset \cap X = \emptyset$;

(ii) se $A \subseteq B$, então,

$$\bar{H}(A, B) = A \cap B = A \subseteq B = A \cup B = \bar{F}(A, B)$$

e

$$\bar{F}(B, A) = A \cap B = A \subseteq B = A \cup B = \bar{H}(B, A);$$

(iii) se $A \subseteq B \subseteq C$, então,

$$\bar{H}(C, A) = C \cup A = C \supseteq C = (C \cup B) \cup (B \cup A) = \bar{H}(C, B) \cup \bar{H}(B, A)$$

e

$$\bar{F}(C, A) = C \cap A = A \subseteq A = (C \cap B) \cap (B \cap A) = \bar{F}(C, B) \cap \bar{F}(B, A).$$

Dados $A, B \in \mathcal{F}(X)$, a propriedade (ii) implica que $S(\bar{H}(A, B), \bar{F}(A, B)) = 1$ se $A \subseteq B$. Similarmemente, se $B \subseteq A$, então, $S(\bar{H}(B, A), \bar{F}(B, A)) = 1$. Assim, podemos provar o primeiro caso da Equação (2.2.19) usando a comutatividade de C_o . Finalmente, temos que $\bar{H}(A, B) = \bar{H}(B, A) = A \cup B$ e $\bar{F}(A, B) = \bar{F}(B, A) = A \cap B$ se $A \not\subseteq B$ e $B \not\subseteq A$. \square

Os Teoremas 1 e 2 apresentados por Zeng *et al.* em [116] surgem como casos especiais da primeira parte do Corolário 2.2.9 se consideramos as t-normas do produto e do mínimo. Além disso, o Teorema 5.1 de [117] difere do Corolário 2.2.9 uma vez que os autores empregam a t-norma do mínimo e as definições de medidas de *subsethood* de Young [114] e de similaridade de Liu [112], respectivamente. Note que, de acordo com o Teorema 3.3 do artigo de Zhang e Zhang [117], a medida de *subsethood* de Willmot (S_W) cumpre todos os axiomas propostos por Young mas o operador $\bar{E}^W(A, B) = \min\{S_W(A, B), S_W(B, A)\}$ não satisfaz $\bar{E}^W(A, A^c) = 0$ para todo $A \in \mathcal{P}(X)$. Portanto, \bar{E}^W não é uma medida de similaridade segundo Liu contradizendo assim o

Teorema 5.1 deste mesmo artigo. Finalmente, a Equação (2.2.17), que também é encontrada em [16, 117], surge da Equação (2.2.19) do Corolário 2.2.9 se C_o corresponde à t-norma do mínimo.

O próximo teorema nos fornece uma fórmula para produzir medidas de equivalência em $\mathcal{F}(X)$ baseada em uma função de agregação v e em uma medida de equivalência simples em $[0, 1]$.

Teorema 2.2.10. Seja X um universo arbitrário e seja $v : \mathcal{F}(X) \rightarrow [0, 1]$ uma função crescente tal que

$$v(\emptyset) = 0 \text{ e } v(X) = 1. \quad (2.2.20)$$

Se $\varepsilon_x : [0, 1]^2 \rightarrow [0, 1]$ denota uma medida de equivalência em $[0, 1]$ para todo $x \in X$, então, o seguinte operador $E_v : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$ é uma medida de equivalência em $\mathcal{F}(X)$:

$$E_v(A, B) = v(\varepsilon_X(A, B)) \forall A, B \in \mathcal{F}(X), \quad (2.2.21)$$

onde $\varepsilon_X(A, B)$ denota o conjunto *fuzzy* de $\mathcal{F}(X)$ cuja função de pertinência é definida como $\varepsilon_X(A, B)(x) = \varepsilon_x(A(x), B(x))$ para todo $x \in X$.

Além disso, se para cada $x \in X$ a função ε_x for uma medida de equivalência *forte* e se v satisfaz $v(A) = 1 \Leftrightarrow A(x) = 1 \forall x \in X$, então, E_v é uma medida de equivalência *forte*.

Demonstração. Vejamos que de fato E_v satisfaz os axiomas da Definição 2.2.5. Para $A, B, C \in \mathcal{F}(X)$,

- temos que $\varepsilon_X(A, B) = \varepsilon_X(B, A)$, pois ε_x é comutativo para todo $x \in X$. Logo, $E_v(A, B) = E_v(B, A)$;
- $E_v(\emptyset, X) = 0$, pois $\varepsilon_X(\emptyset, X)(x) = \varepsilon_x(\mu_\emptyset(x), \mu_X(x)) = \varepsilon_x(0, 1) = 0$ para todo $x \in X$;
- $E_v(A, A) = 1$, pois $\varepsilon_X(A, A)(x) = \varepsilon_x(A(x), A(x)) = 1$ para todo $x \in X$;
- temos que $\varepsilon_x(A(x), C(x)) \leq \varepsilon_x(A(x), B(x))$, $\forall x \in X$ se $A \subseteq B \subseteq C$. Como v é uma função crescente, segue que $E_v(A, C) \leq E_v(A, B)$. Analogamente, podemos mostrar que $E_v(A, C) \leq E_v(B, C)$.

□

Note que, se X é um universo finito, digamos $X = \{x_1, \dots, x_n\}$, então, a *média ponderada* $M_{\mathbf{w}} : [0, 1]^n \rightarrow [0, 1]$ dada por

$$M_{\mathbf{w}}(A_1, \dots, A_n) = \sum_{i=1}^n w_i A_i, \quad \forall A \in [0, 1]^n, \quad (2.2.22)$$

onde $\mathbf{w} \in [0, 1]^n$ tal que $\sum_{j=1}^n w_j = 1$, pode fazer o papel da função de agregação v no Teorema 2.2.10. O Exemplo 2.2.11 exhibe algumas aplicações do Teorema 2.2.10 com $v = M_{\mathbf{w}}$.

Exemplo 2.2.11. Seja E_i uma medida de equivalência em $[0, 1]$ para $i = 1, \dots, n$ e seja $\mathbf{w} \in [0, 1]^n$ tal que $\sum_{i=1}^n w_i = 1$. Uma medida de equivalência $E_{\mathbf{w}} : [0, 1]^n \times [0, 1]^n \rightarrow [0, 1]$ pode ser definida como se segue:

$$E_{\mathbf{w}}(A, B) = \sum_{i=1}^n w_i E_i(A_i, B_i) \forall A, B \in [0, 1]^n. \quad (2.2.23)$$

Em particular, usando (2.2.7) obtemos a seguinte medida de equivalência em $[0, 1]^n$:

$$E_{\lambda, \mathbf{w}}(A, B) = \sum_{i=1}^n w_i E_{\lambda_i}(A_i, B_i) \quad \forall A, B \in [0, 1]^n. \quad (2.2.24)$$

where $\lambda = (\lambda_1, \dots, \lambda_n)^t \in (0, 1]^n$.

Capítulo 3

Memória Associativa *Fuzzy* - Θ

O propósito de uma memória associativa (AM) é armazenar um conjunto finito de associações, isto é, pares ordenados $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) \in X \times Y \mid \xi = 1, \dots, p\}$, também conhecida como conjunto de memórias fundamentais, tal que um padrão de saída desejado \mathbf{y} pode ser recuperado (ou recordado) sob a apresentação de uma versão eventualmente incompleta ou ruidosa de um padrão de entrada \mathbf{x} [41, 62]. Formalmente, uma AM descreve um mapeamento $\Phi : X \rightarrow Y$ que visa associar \mathbf{x}^ξ com \mathbf{y}^ξ para todo $\xi = 1, \dots, p$ e que adicionalmente seja equipado com algum tipo de *tolerância a ruídos* ou *capacidade de correção de erro* (de acordo com a aplicação). Idealmente, para cada $\xi = 1, \dots, p$, temos que Φ satisfaz $\Phi(\mathbf{x}^\xi) = \mathbf{y}^\xi$ e adicionalmente $\Phi(\tilde{x}^\xi) = y^\xi$ se \tilde{x}^ξ é uma versão ruidosa ou corrompida da entrada x^ξ [71]. Na prática, muitos modelos de AM nem sempre são capazes de armazenar todas as memórias fundamentais e/ou produzem uma capacidade de correção de erro insatisfatória ou inexistente, conseguindo recuperar aproximadamente versões corrompidas e ruidosas de \mathbf{x}^ξ , *i.e.*, $\Phi(\tilde{\mathbf{x}}^\xi) \simeq \mathbf{y}^\xi$. Se Φ representa uma rede neural *fuzzy* [12], isto é, uma rede neural artificial cujas as entradas ou pesos são valores *fuzzy*, então, nos referimos à respectiva AM como uma memória associativa *fuzzy* (FAM) [102].

Os modelos de Θ -FAM que serão discutidos neste capítulo surgiram das FAMs de *subsethood* e equivalência (ou similaridade) [27], em particular, da FAM de *subsethood* de Kosko (KS-FAM) [98]. Generalizando estes modelos, introduzimos a classe de Θ -FAMs que envolve tipos mais gerais de funções do que medidas de *subsethood* e equivalência. Em contraste com os primeiros modelos, as Θ -FAMs incluem pesos com valores reais que podem ser ajustados utilizando um algoritmo de treinamento [25, 24].

Definição 3.0.12. Seja $\mathcal{M} = \{(A^\xi, B^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y) \mid \xi \in \mathcal{P}\}$, onde X e Y são universos arbitrários e \mathcal{P} é um conjunto de índices finito, digamos $\mathcal{P} = \{1, \dots, p\}$. Dados operadores $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ satisfazendo $\Theta^\xi(A^\xi) = 1 \forall \xi \in \mathcal{P}$, um vetor de pesos $\mathbf{v} \in \mathbb{R}^p$, uma t-norma arbitrária T e uma função $F : \mathbb{R}^p \rightarrow [0, 1]^p$, o seguinte mapeamento $\mathcal{O}_{\mathbf{v}} : \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$ é dito ser uma memória associativa *fuzzy*- Θ , ou, simplesmente, uma Θ -FAM:

$$\mathcal{O}_{\mathbf{v}}(A) = R \circ_T F(v_1 \Theta^1(A), \dots, v_p \Theta^p(A)), \quad (3.0.1)$$

onde R denota a relação *fuzzy* em $Y \times \mathcal{P}$ dada por $R(y, \xi) = B^\xi(y)$ para todo $y \in Y$ e $\xi \in \mathcal{P}$.

A partir de agora, iremos focar apenas em modelos Θ -FAMs tais que a função $F : \mathbb{R}^p \rightarrow [0, 1]^p$ na Definição 3.0.12 é dada para todo $\mathbf{a} \in \mathbb{R}^p$ como se segue

$$F(a_1, \dots, a_p)_i = \begin{cases} 1 & \text{se } a_i = \bigvee_{j=1}^p a_j, \\ 0 & \text{caso contrário,} \end{cases} \text{ para } i = 1, \dots, p. \quad (3.0.2)$$

Para universos finitos $X = \{x_1, \dots, x_n\}$ e $Y = \{y_1, \dots, y_m\}$, a Figura 3.1 mostra a topologia de uma Θ -FAM com a t-norma do produto.

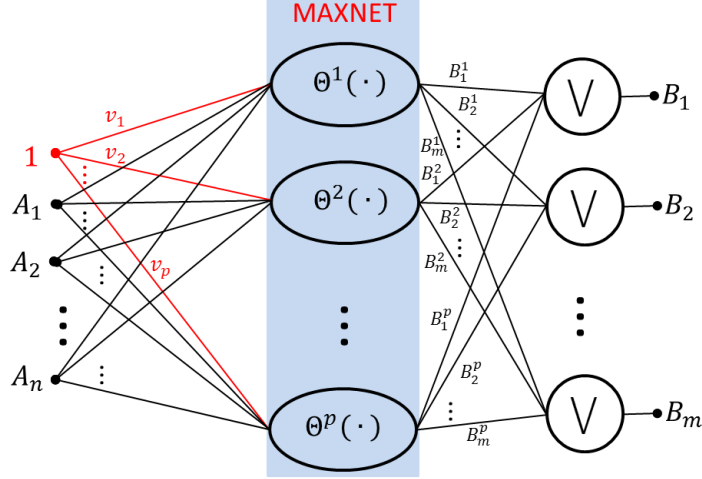


Figura 3.1: Topologia de uma Θ -FAM com universos finitos $X = \{x_1, \dots, x_n\}$ e $Y = \{y_1, \dots, y_m\}$.

Note que a Θ -FAM ilustrada na Figura 3.1 corresponde a uma rede neural com p e m neurônios nas camadas oculta e de saída, respectivamente, produzindo um mapeamento $[0, 1]^n \rightarrow [0, 1]^m$. Dado um padrão de entrada $A = (A_1, A_2, \dots, A_n)^t$, o j -ésimo neurônio escondido calcula $v_j \Theta^j(A)$ seguido por uma função de ativação binária que produz 1 se e somente se $v_j \Theta^j(A) \geq v_i \Theta^i(A)$ para todo $i \neq j$. Logo, a camada oculta pode ser implementada em termos de uma rede MAXNET [67]. Para todo $j = 1, \dots, p$, o j -ésimo neurônio oculto se conecta com cada um dos m neurônios de saída com pesos fixos. Em particular, o i -ésimo peso é dado por B_i^j , *i.e.*, a i -ésima componente de $B^j \in [0, 1]^m$. Cada neurônio de saída calcula o máximo das suas entradas. Enquanto a maioria dos modelos de memórias associativas *fuzzy* representam redes neurais *fuzzy* totalmente conectadas sem camadas ocultas [102], as Θ -FAMs possuem uma camada oculta competitiva como mostrado na Figura 3.1. Neste contexto, as Θ -FAMs são similares às bem conhecidas redes de Hamming [66] ou memória associativa de Hamming e suas extensões [44, 76].

As funções $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ podem ser geradas a partir de medidas de *subsethood* ou equivalência. Note que, para qualquer medida de *subsethood* S e para qualquer medida de equivalência E , temos que $S(A^\xi, A^\xi) = 1$ e $E(A^\xi, A^\xi) = 1$. Portanto, $S(A^\xi, \cdot)$, $S(\cdot, A^\xi)$ e $E(\cdot, A^\xi)$ podem fazer o papel de Θ^ξ na Definição 3.0.12. Esta observação dá origem às seguintes definições:

Definição 3.0.13. Considere uma Θ -FAM dada pela Fórmula (3.0.1) da Definição 3.0.12.

1. Se cada função $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ for dada por $\Theta^\xi(\cdot) = S^\xi(A^\xi, \cdot)$ para alguma medida de *subsethood* S^ξ , então, a correspondente Θ -FAM é chamada uma *memória associativa fuzzy de subsethood* (*S-FAM*). Note que uma S-FAM gera a seguinte saída $\mathcal{S}(A)$ para uma entrada $A \in \mathcal{F}(X)$:

$$\mathcal{S}(A) = R \circ_T F(v_1 S^1(A^1, A), \dots, v_p S^p(A^p, A)). \quad (3.0.3)$$

Falamos de uma *memória associativa fuzzy de subsethood ajustável*, ou, simplesmente, *TS-FAM*, se cada S^ξ é igual a uma medida de *subsethood* parametrizada;

2. Se cada função $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ for dada por $\Theta^\xi(\cdot) = S^\xi(\cdot, A^\xi)$ para alguma medida de *subsethood* S^ξ , então, a correspondente Θ -FAM é chamada uma *memória associativa fuzzy de subsethood dual* (*S-FAM dual*). Note que uma dual S-FAM gera a seguinte saída $\mathcal{D}(A)$ para uma entrada $A \in \mathcal{F}(X)$:

$$\mathcal{D}(A) = R \circ_T F(v_1 S^1(A, A^1), \dots, v_p S^p(A, A^p)). \quad (3.0.4)$$

Falamos de uma *memória associativa fuzzy de subsethood ajustável dual*, ou, simplesmente, *TS-FAM dual*, se cada S^ξ é igual a uma medida de *subsethood* parametrizada;

3. Se cada função $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ for dada por $\Theta^\xi(\cdot) = E^\xi(\cdot, A^\xi)$ para alguma medida de equivalência E^ξ , então, a correspondente Θ -FAM é chamada uma *memória associativa fuzzy de equivalência* (*E-FAM*). Note que uma E-FAM gera a seguinte saída $\mathcal{E}(A)$ para uma entrada $A \in \mathcal{F}(X)$:

$$\mathcal{E}(A) = R \circ_T F(v_1 E^1(A, A^1), \dots, v_p E^p(A, A^p)). \quad (3.0.5)$$

Falamos de uma *memória associativa fuzzy de equivalência ajustável*, ou, simplesmente, *TE-FAM*, se cada E^ξ é igual a uma medida de equivalência parametrizada.

Na morfologia matemática *fuzzy*, o operador elementar de erosão *fuzzy* é definido em termos do grau de inclusão ou *subsethood* de objetos (*i.e.*, versões transladadas de um objeto dito elemento estruturante) em outro objeto [94]. Adicionalmente, as medidas de equivalência definidas na primeira parte do Corolário 2.2.9 podem ser relacionadas com transformadas *hit-or-miss fuzzy*. Portanto, S-FAMs (duais) e E-FAMs podem ser vistas como redes neurais morfológicas *fuzzy*, mais especificamente, como memórias associativas morfológicas *fuzzy* [107]. Relembre que redes neurais morfológicas executam um operação da morfologia matemática *fuzzy* em cada neurônio seguido, eventualmente, pela aplicação de uma função de ativação [102]. Redes neurais morfológicas (*fuzzy*) incluem os modelos de *raciocínio em reticulados (fuzzy)* de Kaburlasos *et al.* que empregam um certo tipo de “medida de inclusão” que pode ser vista como uma ordem parcial *fuzzy* em um reticulado [49, 53, 54].

3.1 Alguns Resultados Teóricos Referentes as Θ -FAMs

Os Teoremas 3.1.1 e 3.1.2 descrevem, respectivamente, as condições suficientes para que uma Θ -FAM seja capaz de armazenar todas as memórias fundamentais e exibir uma certa tolerância a ruídos.

Teorema 3.1.1. Sejam os operadores $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ tais que $\Theta^\xi(A^\xi) = 1$ para $\xi = 1, \dots, p$ e um vetor $\mathbf{v} \in \mathbb{R}^p$. Seja $\mathcal{O}_{\mathbf{v}}$ o símbolo que denota a correspondente Θ -FAM segundo a Definição 3.0.12 com alguma t-norma. Se $\{(A^\xi, B^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y) \mid \xi = 1, \dots, p\}$ satisfaz $v_\gamma \Theta^\gamma(A^\xi) < v_\xi$ para todo $\gamma \neq \xi$, então, temos que

$$\mathcal{O}_{\mathbf{v}}(A^\xi) = B^\xi \quad \forall \xi = 1, \dots, p. \quad (3.1.1)$$

Demonstração. Seja $\xi \in \{1, \dots, p\}$. Por hipótese, temos que $\Theta^\xi(A^\xi) = 1$. Se $v_\gamma \Theta^\gamma(A^\xi) < v_\xi = v_\xi \Theta^\xi(A^\xi)$, então, $F(v_1 \Theta^1(A^\xi), \dots, v_p \Theta^p(A^\xi))_i \in \{0, 1\}$ é igual a 1 se e somente se $i = \xi$, segundo a Equação (3.0.2). Pelas definições de t-norma e da composição sup-T, concluímos que $\mathcal{O}_{\mathbf{v}}(A^\xi) = B^\xi$. \square

Teorema 3.1.2. Sejam $\mathbf{v} \in \mathbb{R}^p$ e $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ funções contínuas (com respeito a uma métrica d) que satisfazem $\Theta^\xi(A^\xi) = 1$ para $\xi = 1, \dots, p$. Além disso, seja $\mathcal{O}_{\mathbf{v}}$ o símbolo que denota a correspondente Θ -FAM segundo a Definição 3.0.12 com alguma t-norma.

Se $v_\gamma \Theta^\gamma(A^\xi) < v_\xi$ para $\gamma = 1, \dots, p$ e $\gamma \neq \xi$, então, existe $\delta^\xi > 0$ tal que $\mathcal{O}_{\mathbf{v}}(A) = B^\xi$ se $d(A^\xi, A) < \delta^\xi$.

Demonstração. Seja $\beta^\xi = v_\xi - \left(\bigvee_{\gamma \neq \xi} v_\gamma \Theta^\gamma(A^\xi) \right) > 0$ para $\xi = 1, \dots, p$. Primeiramente, note que os operadores $v_\gamma \Theta^\gamma(\cdot)$ são contínuos em A^ξ para todo $\gamma = 1, \dots, p$ e em particular para $\gamma \neq \xi$. Portanto, para cada $\gamma \neq \xi$ existe $\delta_\gamma^\xi > 0$ tal que $v_\gamma \Theta^\gamma(A) < (v_\xi - 0.5\beta^\xi)$ se $d(A, A^\xi) \leq \delta_\gamma^\xi$. Similarmente, existe $\delta_\xi^\xi > 0$ tal que $v_\xi \Theta^\xi(A) > (v_\xi - 0.5\beta^\xi)$ se $d(A, A^\xi) \leq \delta_\xi^\xi$. Definindo $\delta^\xi = \bigwedge_{\gamma=1, \dots, p} \delta_\gamma^\xi$, temos que $v_\xi \Theta^\xi(A) > v_\gamma \Theta^\gamma(A)$ se $d(A, A^\xi) \leq \delta^\xi$. Assim, pela Equação (3.0.2), para $A \in \mathcal{F}(X)$ tal que $d(A, A^\xi) \leq \delta^\xi$ temos que $F(v_1 \Theta^1(A), \dots, v_p \Theta^p(A))_i \in \{0, 1\}$ é igual a 1 se e somente se $i = \xi$, $i = 1, \dots, p$. Portanto, pela definição de t-norma e da composição sup-T, obtemos que $\mathcal{O}_{\mathbf{v}}(A) = B^\xi$ se $d(A, A^\xi) \leq \delta^\xi$. \square

3.2 Algoritmo de Treinamento Geral para Θ -FAMs

Suponha que nos seja fornecido um conjunto de memórias fundamentais $\mathcal{M} = \{(A^\xi, B^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y) \mid \xi = 1, \dots, p\}$, um subconjunto \mathbf{V} de \mathbb{R}^p não vazio, fechado e convexo e funções $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$ satisfazendo $\Theta^\xi(A^\xi) = 1$ para $\xi = 1, \dots, p$. Nesta seção vamos considerar o problema de determinar um vetor apropriado $\mathbf{v} \in \mathbf{V}$ que minimiza algum tipo de função de erro dependente de um conjunto de treinamento \mathcal{T} . Informalmente falando, dado um conjunto de associações $\mathcal{T} = \{(\bar{A}^\xi, \bar{B}^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y) : \xi = 1, \dots, k\}$, a fase de treinamento deve produzir um vetor de pesos $\mathbf{v} \in \mathbf{V}$ tal que $\mathcal{O}_{\mathbf{v}}(\bar{A}^\xi) \approx \bar{B}^\xi$ para $\xi = 1, \dots, k$ e $\mathcal{O}_{\mathbf{v}}$ dado pela Equação (3.0.1). Para

isto, uma função objetiva tal como a seguinte função de erro quadrático poderia ser minimizada sujeita a $\mathbf{v} \in \mathbf{V}$:

$$MSE(\mathbf{v}) = \sum_{\xi=1}^k \|\mathcal{O}_{\mathbf{v}}(\bar{A}^{\xi}) - \bar{B}^{\xi}\|^2. \quad (3.2.1)$$

Note que, estritamente falando, não impomos nenhuma restrição sobre a relação entre os conjuntos \mathcal{M} e \mathcal{T} . Entretanto, o objetivo de atingir perfeita recordação, *i.e.*, $\mathcal{O}_{\mathbf{v}}(A^{\xi}) = B^{\xi}$ para $\xi = 1, \dots, p$, sugere escolher $\mathcal{M} \subseteq \mathcal{T}$. Neste contexto, lembre que o conjunto de treinamento \mathcal{T} determina a função erro a ser minimizada, enquanto, o conjunto de memórias fundamentais \mathcal{M} determina a topologia da Θ -FAM em questão. Em particular, as funções Θ^{ξ} dos neurônios ocultos ilustrados na Figura 3.1 satisfazem $\Theta^{\xi}(A^{\xi}) = 1$.

Como mencionado acima, uma possível estratégia para determinar $\mathbf{v} \in \mathbf{V}$ consiste em minimizar a função objetiva $MSE(\mathbf{v})$ na Equação (3.2.1). Note que a função objetiva $MSE(\mathbf{v})$ não é diferenciável com respeito a \mathbf{v} , pois o operador $\mathcal{O}_{\mathbf{v}}$ é dado pela composição da função F (*cf.* Eq. (3.0.2)) e do operador sup-T que não são diferenciáveis com respeito a \mathbf{v} . Uma solução aproximada para este problema pode ser obtida utilizando-se de métodos de otimização não linear sem derivadas tal como algoritmos genéticos. Entretanto, a aplicação de um algoritmo genético para problemas tendo um grande número de variáveis, como os considerados nesta tese, muitas vezes se tornam impraticáveis devido ao seu alto custo computacional [89].

Em uma tentativa de desenvolver uma alternativa factível para determinar $\mathbf{v} \in \mathbf{V}$ na fase de treinamento, desenvolvemos um algoritmo de treinamento específico para Θ -FAMs. Primeiramente, introduziremos algumas notações. Dado $B \in \mathcal{F}(Y)$, definimos os seguintes conjuntos de índices:

$$T(B) = \{\gamma \in \mathcal{P} : B^{\gamma} = B\} \text{ e } T^c(B) = \mathcal{P} \setminus T(B), \quad (3.2.2)$$

$$\bar{T}(B) = \{\gamma \in \mathcal{K} : \bar{B}^{\gamma} = B\} \text{ e } \bar{T}^c(B) = \mathcal{K} \setminus \bar{T}(B), \quad (3.2.3)$$

onde $\mathcal{P} = \{1, \dots, p\}$ e $\mathcal{K} = \{1, \dots, k\}$. Para cada $A \in \mathcal{F}(X)$, associamos o vetor binário $F(v_1\Theta^1(A), \dots, v_p\Theta^p(A))$ com o subconjunto $I_{\mathbf{v}}(A)$ de \mathcal{P} dado por

$$I_{\mathbf{v}}(A) = \{\gamma \in \mathcal{P} \mid F(v_1\Theta^1(A), \dots, v_p\Theta^p(A))_{\gamma} = 1\}. \quad (3.2.4)$$

Note que, pela definição de t-norma e da composição sup-T, a Equação (3.0.1) na Definição 3.0.12 pode ser reescrita como segue:

$$\mathcal{O}_{\mathbf{v}}(A) = R \circ_T F(v_1\Theta^1(A), \dots, v_p\Theta^p(A)) = \bigcup_{\gamma \in I_{\mathbf{v}}(A)} B^{\gamma}. \quad (3.2.5)$$

Além disso, temos que o conjunto $I_{\mathbf{v}}(A)$ é não vazio uma vez que existe um índice $j \in \mathcal{P}$ tal que $v_j\Theta^j(A)$ maximiza o subconjunto finito $\{v_{\xi}\Theta^{\xi}(A)\}_{\xi=1}^p$ de \mathbb{R} .

Além disso, seja $D : \mathbb{R} \rightarrow \{0, 1\}$ a função degrau dada por

$$D(a) = \begin{cases} 1, & \text{se } a \geq 0 \\ 0, & \text{se } a < 0 \end{cases} \quad \forall a \in \mathbb{R}. \quad (3.2.6)$$

Seja $\mathcal{O}_{\mathbf{v}}$ o símbolo que denota uma Θ -FAM baseada em uma família de mapeamentos $\Theta^\xi : \mathcal{F}(X) \rightarrow [0, 1]$, um vetor de pesos $\mathbf{v} \in \mathbf{V}$ e uma t-norma qualquer. Obtemos erro de treinamento 0 se e somente se $\mathcal{O}_{\mathbf{v}}(\bar{A}^\xi) = \bar{B}^\xi$ para todo $\xi = 1, \dots, k$. Considere um índice arbitrário $\xi \in \{1, \dots, k\}$. Note que, pela Equação (3.2.5), $\mathcal{O}_{\mathbf{v}}(\bar{A}^\xi) = \bar{B}^\xi$ se $I_{\mathbf{w}}(\bar{A}^\xi) \subseteq T(\bar{B}^\xi)$ ou, equivalentemente, se para cada $\gamma \in T^c(\bar{B}^\xi)$ existe $j \in T(\bar{B}^\xi)$ tal que $v_\gamma \Theta^\gamma(\bar{A}^\xi) < v_j \Theta^j(\bar{A}^\xi)$. Para um dado índice $\gamma \in T^c(\bar{B}^\xi)$, a função $P_\gamma^\xi : \mathbf{V} \rightarrow \{0, 1\}$ definida pela seguinte equação avalia se $v_\gamma \Theta^\gamma(\bar{A}^\xi) \geq v_j \Theta^j(\bar{A}^\xi)$ para todo $j \in T(\bar{B}^\xi)$:

$$P_\gamma^\xi(\mathbf{v}) = \prod_{j \in T(\bar{B}^\xi)} D(v_\gamma \Theta^\gamma(\bar{A}^\xi) - v_j \Theta^j(\bar{A}^\xi)). \quad (3.2.7)$$

Especificamente, dado $\gamma \in T^c(\bar{B}^\xi)$, temos que existe $j \in T(\bar{B}^\xi)$ tal que $v_\gamma \Theta^\gamma(\bar{A}^\xi) < v_j \Theta^j(\bar{A}^\xi)$ se e somente se $P_\gamma^\xi(\mathbf{v}) = 0$. Caso contrário, temos que $P_\gamma^\xi(\mathbf{v}) = 1$. Além disso, note que $\mathcal{O}_{\mathbf{v}}(\bar{A}^\xi) = \bar{B}^\xi$ se $\sum_{\gamma \in T^c(\bar{B}^\xi)} P_\gamma^\xi(\mathbf{v})$ é igual a 0. Portanto, propomos minimizar a seguinte função objetiva $f : \mathbf{V} \rightarrow \mathbb{Z}$ ao invés de $MSE : \mathbf{V} \rightarrow \mathbb{R}$:

$$f(\mathbf{v}) = \sum_{\xi=1}^k \left(\sum_{\gamma \in T^c(\bar{B}^\xi)} P_\gamma^\xi(\mathbf{v}) \right). \quad (3.2.8)$$

Logo, estamos interessado em encontrar uma solução do seguinte problema de otimização não linear:

$$\min_{\mathbf{v} \in \mathbf{V}} f(\mathbf{v}) \quad (3.2.9)$$

Uma vez que $P_\gamma^\xi(\mathbf{v}) \in \{0, 1\}$, temos que $0 \leq f(\mathbf{v}) \leq \sum_{\xi=1}^k |T^c(\bar{B}^\xi)| \leq kp$ para todo $\mathbf{v} \in \mathbf{V}$. Assim, f pode ser considerada uma função $\mathbf{V} \rightarrow \{0, \dots, kp\}$ o que implica que o problema de otimização da Equação (3.2.9) está bem definido e possui ao menos uma solução global.

Nossa estratégia para resolver o problema de otimização da Equação (3.2.9) consiste em procurar direções de descidas entre os vetores canônicos \mathbf{e}^j de \mathbb{R}^p . Dado $\mathbf{v} \in \mathbf{V}$, vamos analisar o efeito da j -ésima componente de \mathbf{v} no valor de f . Para cada $j = 1, \dots, p$, temos que $f(\mathbf{v})$ pode ser reescrita como se segue:

$$f(\mathbf{v}) = \sum_{\xi \in \bar{T}(B^j)} \sum_{\gamma \in T^c(B^j)} P_\gamma^\xi(\mathbf{v}) + \sum_{\xi \in \bar{T}^c(B^j)} \sum_{\gamma \in T^c(\bar{B}^\xi)} P_\gamma^\xi(\mathbf{v}). \quad (3.2.10)$$

Vamos olhar um pouco mais a fundo nas expressões do lado esquerdo e do lado direito do símbolo “+” na Equação (3.2.10). Pelas definições de \bar{T} e T , temos que $\xi \in \bar{T}(B^j)$ se e somente se $j \in T(\bar{B}^\xi)$, *i.e.*, $B^j = \bar{B}^\xi$, o que implica que, para todo γ (em particular para $\gamma \in T^c(\bar{B}^\xi)$), o conjunto de índices considerado no produto da Equação (3.2.7) inclui j . Em outras palavras, $P_\gamma^\xi(\mathbf{v})$ depende da j -ésima componente de \mathbf{v} se $\xi \in \bar{T}(B^j)$. Por outro lado, se $\xi \in \bar{T}^c(B^j)$, então, $j \in T^c(\bar{B}^\xi)$ e, para $\gamma \in T^c(\bar{B}^\xi) \setminus \{j\}$, $P_\gamma^\xi(\mathbf{v})$ não depende da j -ésima componente de \mathbf{v} , pois, j não pertence ao conjunto de índices em que ocorre o produto da Equação (3.2.7). Portanto,

$$f(\mathbf{v}) = \sum_{\xi \in \bar{T}(B^j)} \sum_{\gamma \in T^c(B^j)} P_\gamma^\xi(\mathbf{v}) + \sum_{\xi \in \bar{T}^c(B^j)} P_j^\xi(\mathbf{v}) + C_j(\mathbf{v}), \quad (3.2.11)$$

onde $C_j(\mathbf{v}) = \sum_{\xi \in \bar{T}^c(B^j)} \left(\sum_{\gamma \in T^c(\bar{B}^\xi) \setminus \{j\}} P_\gamma^\xi(\mathbf{v}) \right)$. Uma vez que o termo $C_j(\mathbf{v})$ não depende do índice j , uma modificação de v_j apenas afetará o valor de

$$R_j(\mathbf{v}) = f(\mathbf{v}) - C_j(\mathbf{v}) = \sum_{\xi \in \bar{T}(B^j)} \sum_{\gamma \in T^c(B^j)} P_\gamma^\xi(\mathbf{v}) + \sum_{\xi \in \bar{T}^c(B^j)} P_j^\xi(\mathbf{v}) =$$

$$\sum_{\xi \in \bar{T}(B^j)} \sum_{\gamma \in T^c(B^j)} a_{\gamma,j,\xi} D(v_\gamma \Theta^\gamma(\bar{A}^\xi) - v_j \Theta^j(\bar{A}^\xi)) + \sum_{\xi \in \bar{T}^c(B^j)} P_j^\xi(\mathbf{v}),$$

onde o coeficiente $a_{\gamma,j,\xi} \in \{0, 1\}$ é dado por

$$a_{\gamma,j,\xi} = \prod_{i \in T(B^j) \setminus \{j\}} D(v_\gamma \Theta^\gamma(\bar{A}^\xi) - v_i \Theta^i(\bar{A}^\xi)). \quad (3.2.12)$$

Considere $\mathbf{v} \in \mathbf{V}$ arbitrário. Nosso objetivo é encontrar um valor real v_j^* tal que o vetor $\mathbf{v} + (v_j^* - v_j)\mathbf{e}^j$ está contido em \mathbf{V} e minimiza o valor de R_j . Para isto, seja V_j o conjunto $\{\alpha \in \mathbb{R} : (\mathbf{v} + (\alpha - v_j)\mathbf{e}^j) \in \mathbf{V}\}$, onde \mathbf{e}^j denota a j -ésimo vetor canônico de \mathbb{R}^p , definimos as seguintes funções $g_j, h_j : V_j \rightarrow \mathbb{Z}$:

$$g_j(v) = \sum_{\xi \in \bar{T}(B^j)} \sum_{\gamma \in T^c(B^j)} a_{\gamma,j,\xi} D(v_\gamma \Theta^\gamma(\bar{A}^\xi) - v \Theta^j(\bar{A}^\xi)),$$

$$h_j(v) = \sum_{\xi \in \bar{T}^c(B^j)} \prod_{i \in T(\bar{B}^\xi)} D(v \Theta^j(\bar{A}^\xi) - v_i \Theta^i(\bar{A}^\xi)). \quad (3.2.13)$$

No que se segue, desenvolveremos um método de otimização que, sob algumas condições fracas, produz um mínimo local do problema na Equação (3.2.9). Nossa abordagem é baseada em resolver problemas simples da forma

$$\min_{w \in V_j} f_j(w), \quad (3.2.14)$$

onde f_j denota a soma das funções g_j e h_j .

Tal como f , a função f_j também possui valores em $\{0, \dots, kp\}$ e, portanto, é garantido existir uma solução global para o problema de minimização na Equação (3.2.14). Desde de que $f_j(v_j) = R_j(\mathbf{v})$ e $f_j = g_j + h_j$ por definição, temos que $f_j(v^*) \leq R_j(\mathbf{v})$ para qualquer solução global v^* da Equação (3.2.14). Note que a seguinte afirmação é verdadeira para todo ξ tal que $\Theta^j(\bar{A}^\xi) \neq 0$:

$$D(v_\gamma \Theta^\gamma(\bar{A}^\xi) - v \Theta^j(\bar{A}^\xi)) = 0 \Leftrightarrow \frac{v_\gamma \Theta^\gamma(\bar{A}^\xi)}{\Theta^j(\bar{A}^\xi)} < v. \quad (3.2.15)$$

A função g_j pode ser escrita como

$$g_j(v) = d_j + \sum_{\substack{\xi \in \bar{T}(B^j), \\ \Theta^j(\bar{A}^\xi) \neq 0}} \left\{ \sum_{\substack{\gamma \in T^c(B^j), \\ v \leq \beta_{\gamma,j,\xi}}} a_{\gamma,j,\xi} \right\}, \quad (3.2.16)$$

onde

$$\beta_{\gamma,j,\xi} = \frac{v_\gamma \Theta^\gamma(\bar{A}^\xi)}{\Theta^j(\bar{A}^\xi)} \quad \text{e} \quad d_j = \sum_{\substack{\xi \in \bar{T}(B^j), \\ \Theta^j(\bar{A}^\xi) = 0}} \sum_{\gamma \in T^c(B^j)} a_{\gamma,j,\xi} D(v_\gamma \Theta^\gamma(\bar{A}^\xi)). \quad (3.2.17)$$

Dado que temos um número finito de coeficientes $\beta_{\gamma,j,\xi}$, a Equação (3.2.16) revela que a função g_j é constante por partes, decrescente e com um número finito de descontinuidades. Adicionalmente, a seguinte afirmação é válida para todo ξ tal que $\Theta^j(\bar{A}^\xi) \neq 0$:

$$D(v \Theta^j(\bar{A}^\xi) - v_i \Theta^i(\bar{A}^\xi)) = 0 \Leftrightarrow v < \frac{v_i \Theta^i(\bar{A}^\xi)}{\Theta^j(\bar{A}^\xi)} \quad (3.2.18)$$

o que implica que

$$\prod_{i \in T(\bar{B}^\xi)} D(v\Theta^j(\bar{A}^\xi) - v_i\Theta^i(\bar{A}^\xi)) = 0 \Leftrightarrow v < \bigvee_{i \in T(\bar{B}^\xi)} \frac{v_i\Theta^i(\bar{A}^\xi)}{\Theta^j(\bar{A}^\xi)}. \quad (3.2.19)$$

A função h_j pode ser expressa por

$$h_j(v) = e_j + \sum_{\substack{\xi \in \bar{T}^c(B^j), \\ \Theta^j(\bar{A}^\xi) \neq 0, \\ v \geq \zeta_{j,\xi}}} 1, \quad (3.2.20)$$

onde

$$\zeta_{j,\xi} = \bigvee_{i \in T(\bar{B}^\xi)} \frac{v_i\Theta^i(\bar{A}^\xi)}{\Theta^j(\bar{A}^\xi)} \quad \text{e} \quad e_j = \sum_{\substack{\xi \in \bar{T}^c(B^j), \\ \Theta^j(\bar{A}^\xi) = 0}} \left\{ \prod_{i \in T(\bar{B}^\xi)} D(-v_i\Theta^i(\bar{A}^\xi)) \right\}. \quad (3.2.21)$$

A Equação (3.2.20) implica que h_j é crescente, constante por partes com um número finito de descontinuidades, pois, o conjunto de coeficientes $\zeta_{j,\xi}$ é finito.

Para sumarizar, temos que a função objetiva $f_j = g_j + h_j$ é limitada, não negativa e contante por partes com um número finito de descontinuidades. Relembre ainda que o problema na Equação (3.2.14) possui ao menos uma solução global e muitas vezes um número infinito de soluções globais. A resolução deste problema de otimização requer a computação dos coeficientes $\beta_{\gamma,j,\xi}, \zeta_{j,\xi} \in V_j$ que incluem os pontos de descontinuidades de f_j em V_j .

O seguinte algoritmo de treinamento para Θ -FAMs determina uma aproximação para um minimizador global $\mathbf{v}^* \in \mathbf{V}$ de f .

Suponha que $\mathcal{M} = \{(A^\xi, B^\xi) : \xi = 1, \dots, p\}$ e $\mathcal{T} = \{(\bar{A}^\xi, \bar{B}^\xi) : \xi = 1, \dots, k\}$ são, respectivamente, os conjuntos de memórias fundamentais e treinamento. Considere uma Θ -FAM baseada em funções Θ^ξ , onde $\xi = 1, \dots, p$, que satisfazem $\Theta^\xi(A^\xi) = 1$. Na n -ésima iteração, o Algoritmo 3.1 produz um vetor $\mathbf{v}^{(n)} = (v_1^{(n)}, \dots, v_p^{(n)})^t \in \mathbf{V}$. Seja $f_j^{(n)} = g_j + h_j$, onde as funções g_j e h_j são, respectivamente, dadas pelas Equações (3.2.16) e (3.2.20) com $v_\gamma = v_\gamma^{(n)}$ e $v_i = v_i^{(n)}$ nas Equações (3.2.12), (3.2.17) e (3.2.21).

Teorema 3.2.1. No Algoritmo 3.1, as desigualdades $f(\mathbf{v}^{(n-1)}) > f(\mathbf{v}^{(n)})$ são verdadeiras para todas as iterações $n > 0$ até que o critério de parada $J \neq \emptyset$ seja atingido.

Demonstração. Suponha que $j^* \in J$ tenha sido escolhido no Passo 6) do Algoritmo 3.1. Após a execução do Passo 8), temos que $f_{j^*}^{(n-1)}(v_{j^*}^{(n-1)}) > f_{j^*}^{(n-1)}(v_{j^*}^{(n)})$. Note que $C_{j^*}(\mathbf{v}^{(n-1)}) = C_{j^*}(\mathbf{v}^{(n)})$ and $f_{j^*}^{(n-1)} = f_{j^*}^{(n)}$, pois, $v_i^{(n-1)} = v_i^{(n)}$ para todo $i \neq j^*$. Pelas Equações (3.2.7) e (3.2.11), temos

Algoritmo 3.1 Algoritmo Geral de Treinamento para Θ -FAM

- 1: Dado $\mathbf{v}^{(0)} \in \mathbf{V}$.
- 2: $n \leftarrow 0$;
- 3: Para $j = 1, \dots, p$, compute uma solução global v_j^* de

$$\min_{v \in V_j} f_j^{(0)}(v);$$

- 4: Defina $J = \{j : f_j^{(0)}(v_j^*) < f_j^{(0)}(v_j^{(0)})\}$;
- 5: **while** $J \neq \emptyset$ **do**
- 6: Escolha $j \in J$;
- 7: Atualize o vetor \mathbf{v} :
$$v_i^{(n+1)} \leftarrow \begin{cases} v_j^* & \text{se } i = j \\ v_i^{(n)} & \text{se } i \neq j \end{cases}, \text{ para } i = 1, \dots, p;$$
- 8: $n \leftarrow n + 1$;
- 9: Para $j = 1, \dots, p$, compute uma solução global v_j^* de

$$\min_{v \in V_j} f_j^{(n)}(v);$$

- 10: Defina $J = \{j : f_j^{(n)}(v_j^*) < f_j^{(n)}(v_j^{(n)})\}$;
 - 11: **end while**
 - 12: Retorne $\mathbf{v} \leftarrow \mathbf{v}^{(n)}$.
-

que

$$\begin{aligned} f(\mathbf{v}^{(n-1)}) &= R_{j^*}(\mathbf{v}^{(n-1)}) + C_{j^*}(\mathbf{v}^{(n-1)}) \\ &= f_{j^*}^{(n-1)}(v_{j^*}^{(n-1)}) + C_{j^*}(\mathbf{v}^{(n-1)}) \\ &> f_{j^*}^{(n-1)}(v_{j^*}^{(n)}) + C_{j^*}(\mathbf{v}^{(n)}) \\ &= f_{j^*}^{(n)}(v_{j^*}^{(n)}) + C_{j^*}(\mathbf{v}^{(n)}) \\ &= R_{j^*}(\mathbf{v}^{(n)}) + C_{j^*}(\mathbf{v}^{(n)}) \\ &= f(\mathbf{v}^{(n)}). \end{aligned}$$

□

O Teorema 3.2.1 implica que o Algoritmo 3.1 converge em um número finito de iterações, pois, $0 \leq f(\mathbf{v}^{(n)}) \leq f(\mathbf{v}^{(n-1)}) - 1$ para $n > 0$. Assim, obtemos os seguinte corolário.

Corolário 3.2.2. O Algoritmo 3.1 converge em no máximo $f(\mathbf{v}^{(0)})$ iterações.

Demonstração. Pelo Teorema 3.2.1, na iteração n do algoritmo temos que $0 \leq f(\mathbf{v}^{(n)}) \leq f(\mathbf{v}^{(n-1)}) - 1$ para $n > 0$. Por indução, mostra-se que $0 \leq f(\mathbf{v}^{(n)}) \leq f(\mathbf{v}^{(0)}) - n$. Portanto, o Algoritmo 3.1 converge no máximo $f(\mathbf{v}^{(0)})$ iterações. □

A seguir, mostraremos que o Algoritmo 3.1 atinge um mínimo local de f em \mathbf{V} sob algumas condições. Dado $\mathbf{v} = (v_1, \dots, v_p)^t \in \mathbf{V}$, definimos G_j e H_j como sendo, respectivamente, os conjuntos dos pontos de descontinuidades das funções g_j e h_j e $F_j = G_j \cup H_j$. Para todo $j = 1, \dots, p$, temos:

$$G_j = \{\beta_{\gamma,j,\xi} : \xi \in \bar{T}(B^j) \text{ t.q. } \Theta^j(\bar{A}^\xi) \neq 0, \gamma \in T^c(B^j)\}, \quad (3.2.22)$$

$$H_j = \{\zeta_{j,\xi} : \xi \in \bar{T}(B^j) \text{ t.q. } \Theta^j(\bar{A}^\xi) \neq 0\}. \quad (3.2.23)$$

Primeiramente, iremos providenciar um método para minimizar $f_j(v)$ sujeito a $v \in V_j$ nas linhas 3. e 9. do Algoritmo 3.1. Para isto, empregaremos o lema a seguir que estabelece que as funções f_j são de fato constantes por partes. No que se segue, todos os ínfimos e supremos são tomados em $\mathbb{R}_{\pm\infty}$.

Para cada $j \in \{1, \dots, p\}$, definimos $l_j = \wedge V_j \in \mathbb{R}_{\pm\infty}$, $u_j = \vee V_j \in \mathbb{R}_{\pm\infty}$ e $S_j = (F_j \cap V_j) \cup \{l_j, u_j\} \subseteq \mathbb{R}_{\pm\infty}$. Como V_j é um conjunto fechado e não vazio, concluímos que $(l_j, u_j) \subseteq V_j$.

Lema 3.2.3. Para cada $j = 1, \dots, p$, a função f_j é constante por partes. Além disso, cada $\alpha \in F_j \cap V_j$ é um ponto de máximo local de f_j .

Demonstração. Sejam $j \in \{1, \dots, p\}$ e $\alpha \in V_j$. Considere $\beta = \vee\{s \in S_j : s < \alpha\} \vee l_j$ e $\zeta = \wedge\{s \in S_j : s > \alpha\} \wedge u_j$. Temos que $l_j \leq \beta \leq \zeta \leq u_j$ e $\beta = \zeta \Leftrightarrow l_j = u_j$ pois S_j é finito. Se $l_j = u_j$, então, $|V_j| = 1$ e, neste caso, a afirmação do Lema 3.2.3 segue imediatamente. Agora, vamos assumir que $l_j < u_j$.

Se $\alpha \notin S_j$, então, $\alpha \in (\beta, \zeta)$ e $(\beta, \zeta) \cap S_j = \emptyset$ pela definições de β e ζ . Como o intervalo (β, ζ) não possui tanto pontos de descontinuidades $\beta_{\gamma,j,\xi}$ de g_j quanto pontos de descontinuidades $\zeta_{j,\xi}$ de h_j , as Equações (3.2.16) e (3.2.20) revelam que $g_j(x) = g_j(\alpha)$ e $h_j(x) = h_j(\alpha)$ para todo $x \in (\beta, \zeta)$. Assim, $f_j(x) = f_j(\alpha)$ para todo $x \in (\beta, \zeta)$ e, portanto, f_j é constante por partes.

Se $\alpha \in F_j$, então, $(\beta, \zeta) \cap F_j = \{\alpha\}$. Pela Equação (3.2.16), temos que $g_j(x) = g_j(\alpha)$ se $x \in (\beta, \alpha)$ e $g_j(x) \leq g_j(\alpha)$ se $x \in (\alpha, \zeta)$. Similarmente, a Equação (3.2.20) implica que $h_j(x) < h_j(\alpha)$ se $x \in (\beta, \alpha)$ e $h_j(x) = h_j(\alpha)$ se $x \in (\alpha, \zeta)$. Portanto, como $f_j(x) = g_j(x) + h_j(x)$, concluímos que $f_j(\alpha) \geq f_j(x)$ para todo $x \in (\beta, \zeta)$. Logo, α é um máximo local de f_j . \square

Como a função f_j é constante por partes, é suficiente avaliar f_j em um número finitos de elementos de V_j para encontrar um mínimo global de f_j . De fato, basta computar f_j para um único elemento de cada conjunto onde f_j é constante. Assim, a idéia é substituir a restrição V_j por um conjunto finito U_j dado como se segue:

1. Atribua $U_j = \emptyset$;
2. Para cada $\alpha \in \{l_j\} \cup (F_j \cap V_j)$, faça:
 - (a) Defina $\zeta_\alpha = \wedge\{s \in S_j : s > \alpha\} \wedge u_j$;
 - (b) Se $\alpha = \zeta_\alpha$, então, atribua $t_\alpha = \alpha$, caso contrário, escolha $t_\alpha \in (\alpha, \zeta_\alpha)$ (e.g. $t_\alpha = 0.5 * (\alpha + \zeta_\alpha)$);
 - (c) $U_j := U_j \cup \{t_\alpha\}$.

Portanto, uma solução global de f_j , sujeito a V_j , é dada por

$$v_j^* = \arg \min_{t \in U_j} f_j(t).$$

Um problema com esta abordagem diz respeito a questões numéricas. Mais especificamente, calcular o conjunto U_j introduz problemas numéricos se o tamanho do intervalo (α, ζ_α) é menor do que duas vezes o *epsilon* máquina. Para evitar problemas numéricos desse tipo, excluimos estes intervalos do cálculo de U_j . No caso extremo onde essa estratégia produz $U_j = \emptyset$, desconsideramos a direção \mathbf{e}^j , isto é, removemos o índice j do conjunto J .

Finalmente, seja $\mathbf{v} \in \mathbf{V}$ o resultado do Algoritmo 3.1. Se \mathbf{v} satisfaz $v_j \notin F_j$ para todo $j = 1, \dots, p$, então, \mathbf{v} é um mínimo local de f em \mathbf{V} . Este fato segue da propriedade da função f estabelecida no Teorema 3.2.5, cuja demonstração utiliza o resulta do próximo lema.

Lema 3.2.4. Seja $\Lambda : \mathbb{R}^2 \rightarrow \{0, 1\}$ a função dada por

$$\Lambda(\mathbf{x}) = D(x_1\alpha_1 - x_2\alpha_2), \forall \mathbf{x} \in \mathbb{R}^2, \quad (3.2.24)$$

onde $\alpha_1, \alpha_2 \in [0, 1]$. Dado $\mathbf{x} \in \mathbb{R}^2$ tal que $x_1\alpha_1 \neq x_2\alpha_2$ para $\alpha_1 + \alpha_2 > 0$. Existe $\delta > 0$ tal que $\Lambda(\mathbf{y}) = \Lambda(\mathbf{x})$ para todo $\mathbf{y} = (y_1, y_2)^t \in \mathbb{R}^2$ satisfazendo $\|\mathbf{y} - \mathbf{x}\|_\infty < \delta$.

Demonstração. Por um lado, se $\alpha_1 = \alpha_2 = 0$, então, para todo $\delta > 0$, temos que $\Lambda(\mathbf{y}) = \Lambda(\mathbf{x})$ para todo $\mathbf{y} \in \mathbb{R}^2$. Por outro lado, se $\alpha_1 \neq 0$ ou $\alpha_2 \neq 0$, então, temos que $(\alpha_1 + \alpha_2) > 0$ e $\alpha_1 x_1 \neq \alpha_2 x_2$ pela hipótese acima. Sejam $\delta = \frac{|x_1\alpha_1 - x_2\alpha_2|}{2(\alpha_1 + \alpha_2)} > 0$ e $\mathbf{y} = (y_1, y_2)^t \in \mathbb{R}^2$ tal que $\|\mathbf{y} - \mathbf{x}\|_\infty < \delta$. Se $x_1\alpha_1 - x_2\alpha_2 > 0$, então, temos que

$$\begin{aligned} y_1\alpha_1 - y_2\alpha_2 &> (x_1 - \delta)\alpha_1 - (x_2 + \delta)\alpha_2 \\ &= (x_1\alpha_1 - x_2\alpha_2) - (\alpha_1 + \alpha_2)\delta \\ &= (x_1\alpha_1 - x_2\alpha_2) - (x_1\alpha_1 - x_2\alpha_2)/2 \\ &= (x_1\alpha_1 - x_2\alpha_2)/2 > 0. \end{aligned}$$

Se $x_1\alpha_1 - x_2\alpha_2 < 0$, então,

$$\begin{aligned} y_1\alpha_1 - y_2\alpha_2 &< (x_1 + \delta)\alpha_1 - (x_2 - \delta)\alpha_2 \\ &= (x_1\alpha_1 - x_2\alpha_2) + (\alpha_1 + \alpha_2)\delta \\ &= (x_1\alpha_1 - x_2\alpha_2) - (x_1\alpha_1 - x_2\alpha_2)/2 \\ &= (x_1\alpha_1 - x_2\alpha_2)/2 < 0. \end{aligned}$$

Logo, existe $\delta > 0$ tal que $\text{sign}(y_1\alpha_1 - y_2\alpha_2) = \text{sign}(x_1\alpha_1 - x_2\alpha_2)$ para todo $\mathbf{y} \in \mathbb{R}^2$ satisfazendo $\|\mathbf{y} - \mathbf{x}\|_\infty < \delta$. Portanto, a afirmação do lema segue imediatamente. \square

Teorema 3.2.5. Seja $\mathbf{v} \in \mathbf{V}$ tal que $v_j \notin F_j$ para todo $j = 1, \dots, p$. Para qualquer norma $\|\cdot\|$ em \mathbb{R}^p , existe $\delta > 0$ tal que $f(\mathbf{z}) = f(\mathbf{v})$ para todo $\mathbf{z} \in \mathbf{V}$ satisfazendo $\|\mathbf{z} - \mathbf{v}\| < \delta$.

Demonstração. Seja $\alpha_\gamma^\xi = \Theta^\gamma(\bar{A}^\xi)$ para todo $\xi = 1, \dots, k$ e $\gamma = 1, \dots, p$. Note que a função P_γ^ξ pode ser reescrita como

$$P_\gamma^\xi(\mathbf{v}) = \prod_{j \in T(\bar{B}^\xi)} \Lambda_{\gamma,j}^\xi(v_\gamma, v_j),$$

onde $\Lambda_{\gamma,j}^\xi(\mathbf{x}) = D(x_1 \alpha_\gamma^\xi - x_2 \alpha_j^\xi)$ para todo $\mathbf{x} \in \mathbb{R}^2$. Se $\mathbf{v} \in \mathbf{V}$ tal que $v_j \notin F_j$ para todo $j = 1, \dots, p$, então, pelo Lema 3.2.4, para cada função $\Lambda_{\gamma,j}^\xi$ existe $\delta_{\gamma,j}^\xi > 0$ tal que $\Lambda_{\gamma,j}^\xi$ é constante em $\{(z_1, z_2)^t \in \mathbb{R}^2 : |z_1 - v_\gamma| \vee |z_2 - v_j| < \delta_{\gamma,j}^\xi\}$. Assim, se $\delta_\gamma^\xi = \bigwedge_{j \in T(\bar{B}^\xi)} \delta_{\gamma,j}^\xi$, então, $P_\gamma^\xi(\mathbf{z})$ é constante para todo $\mathbf{z} \in \mathbf{V}$ tal que $\|\mathbf{z} - \mathbf{v}\|_\infty < \delta_\gamma^\xi$. Portanto, $f(\mathbf{z}) = f(\mathbf{v})$ para $\mathbf{z} \in \mathbf{V}$ tal que $\|\mathbf{z} - \mathbf{v}\|_\infty < \delta$, onde $\delta = \bigwedge_{\xi=1}^k (\bigwedge_{\gamma \in T(\bar{B}^\xi)} \delta_\gamma^\xi)$.

Finalmente, uma vez que qualquer norma $\|\cdot\|$ em \mathbb{R}^p é equivalente a norma do máximo $\|\cdot\|_\infty$, existe $c > 0$ tal que $\|\mathbf{x}\| \leq c \|\mathbf{x}\|_\infty$ para todo $\mathbf{x} \in \mathbb{R}^p$. Logo, concluímos que $f(\mathbf{z}) = f(\mathbf{v})$ para $\mathbf{z} \in \mathbf{V}$ tal que $\|\mathbf{z} - \mathbf{v}\| < \bar{\delta}$, onde $\bar{\delta} = c\delta > 0$. \square

3.3 Algoritmo de Treinamento para TE-FAM do Tipo $\mathcal{E}_{W,\mathbf{v}}$

O número de memórias fundamentais determina a topologia da Θ -FAM, mais especificamente, o número de neurônios ocultos (veja Figura 3.1). O algoritmo de treinamento apresentado na seção anterior requer que o conjunto de memórias fundamentais seja pré-conhecido e apenas lida com a seleção do vetor de pesos \mathbf{v} . Nesta seção, apresentaremos um algoritmo de treinamento alternativo para TE-FAMs baseadas em medidas de equivalência dadas pela Equação (2.2.23) que inclui a seleção automática das memórias fundamentais.

Até o resto desta seção nos concentraremos em TE-FAMs baseadas em $E_{\mathbf{w}^\xi}$ para algum $\mathbf{w}^\xi \in [0, 1]^n$ que satisfaz $\sum_{i=1}^n w_i^\xi = 1$. Fazendo $\mathbf{w}^1, \dots, \mathbf{w}^p \in [0, 1]^n$ como vetores colunas de uma matriz $W = [\mathbf{w}^1, \dots, \mathbf{w}^p] \in [0, 1]^{n \times p}$, denotaremos a correspondente TE-FAM usando o símbolo $\mathcal{E}_{W,\mathbf{v}}$. Dado um padrão de entrada $A \in \mathcal{F}(X)$, a TE-FAM $\mathcal{E}_{W,\mathbf{v}}$ gera a seguinte saída $\mathcal{E}_{W,\mathbf{v}}(A) \in \mathcal{F}(Y)$:

$$\mathcal{E}_{W,\mathbf{v}}(A) = B \circ_T F(v_1 E_{\mathbf{w}^1}(A, A^1), \dots, v_p E_{\mathbf{w}^p}(A, A^p)). \quad (3.3.1)$$

Relembre que as medidas de equivalência $E_{\mathbf{w}^\xi}$ são definidas em $[0, 1]^n$, para algum $n > 0$, segundo a Equação (2.2.23). Deste modo, incorporamos um outro conjunto de pesos w_i^ξ , que permite ajustar a relevância do i -ésimo atributo (ou componente) do ξ -ésimo neurônio oculto em termos de $w_i^\xi E_i(\cdot, A_i^\xi)$, enquanto que, como mencionamos anteriormente, os pesos v_ξ , $\xi = 1, \dots, p$ são responsáveis por ajustar a contribuição do ξ -ésimo neurônio oculto. Seja Y um universo finito, a Figura 3.2 exibe a topologia de uma TE-FAM do tipo $\mathcal{E}_{W,\mathbf{v}}$ com a t-norma do produto.

Aqui, introduziremos um algoritmo supervisionado específico para TE-FAMs do tipo $\mathcal{E}_{W,\mathbf{v}}$ chamado de *Algoritmo TE* que pode ser dividido em dois estágios [25]:

1. Extração do conjunto de memórias fundamentais $\mathcal{M} = \{(A^1, B^1), \dots, (A^p, B^p)\}$ do conjunto de treinamento \mathcal{T} ;
2. Otimização dos vetores de pesos \mathbf{w}^ξ e normalização do vetor de pesos \mathbf{v} .

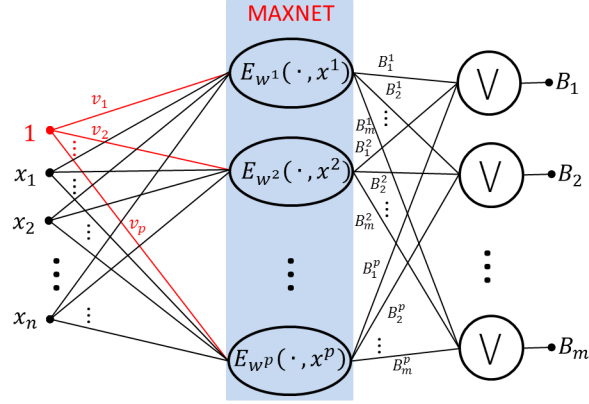


Figura 3.2: Topologia de uma TE-FAM do tipo $\mathcal{E}_{W,v}$.

O Estágio 1 pode ser executado para qualquer E-FAM e determina exatamente a topologia dessa rede neural *fuzzy* de duas camadas uma vez que, inicialmente, p e A^1, \dots, A^p são desconhecidos. O Estágio 1 visa gerar um conjunto de memórias fundamentais \mathcal{M} com $|\mathcal{M}| \leq |\mathcal{T}|$ para facilitar a otimização dos vetores \mathbf{w}^ξ (se existirem) e/ou \mathbf{v} baseado no conjunto de memórias fundamentais resultante $\mathcal{M} \subseteq \mathcal{T}$. A seguir, providenciaremos alguns detalhes sobre os dois estágios individuais.

3.3.1 Seleção de Memórias Fundamentais

Suponha que nos seja fornecido um conjunto de treinamento $\mathcal{T} = \{(\bar{A}^\xi, \bar{B}^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y), \xi = 1, \dots, k\}$. Seja E^ξ uma medida de equivalência em $\mathcal{F}(X)$ para $\xi = 1, \dots, k$. Ao invés de apresentar todos os elementos de \mathcal{T} como memórias fundamentais à E-FAM baseada em E^ξ , onde $\xi = 1, \dots, k$, iremos selecionar um subconjunto \mathcal{M} de \mathcal{T} como um conjunto de memórias fundamentais de uma E-FAM. Esse processo de extração de memórias fundamentais do conjunto de treinamento \mathcal{T} baseia-se em avaliar $E^\gamma(\bar{A}^\xi, \bar{A}^\gamma)$ para algum $\xi \neq \gamma$ a fim de assegurar que \mathcal{M} produz em algum sentido uma boa representação de \mathcal{T} . A seguir, daremos algumas explicações detalhadas dessa abordagem.

Seja $e_{\xi,\gamma} = E^\gamma(\bar{A}^\xi, \bar{A}^\gamma)$ para $\xi, \gamma = 1, \dots, k$. Se $(\bar{A}^\xi, \bar{B}^\xi), (\bar{A}^\gamma, \bar{B}^\gamma) \in \mathcal{T}$ e $\bar{B}^\xi = \bar{B}^\gamma$, então, $e_{\xi,\gamma}$ pode ser visto como um potencial, medido como um número em $[0, 1]$, de \bar{A}^γ substituir \bar{A}^ξ . Para cada $(\bar{A}^\xi, \bar{B}^\xi)$, definimos \bar{e}^ξ como a equivalência ou similaridade máxima de \bar{A}^ξ e qualquer \bar{A}^γ tal que $\bar{B}^\gamma \neq \bar{B}^\xi$. Formalmente, temos que:

$$\bar{e}_\xi = \bigvee_{\gamma: \bar{B}^\gamma \neq \bar{B}^\xi} e_{\xi,\gamma}. \quad (3.3.2)$$

Para todo $\rho \in [0, 1)$ e $\xi \in \{1, \dots, k\}$, definimos \mathcal{R}_ρ^ξ como o conjunto de índices γ tal que \bar{A}^γ pode substituir \bar{A}^ξ com um potencial maior do que ρ e

$$\bar{e}_\gamma \leq \bar{e}_\xi < e_{\xi,\gamma}. \quad (3.3.3)$$

Formalmente, temos

$$\mathcal{R}_\rho^\xi = \{\gamma \in \{1, \dots, k\} \mid \bar{B}^\gamma = \bar{B}^\xi, \bar{e}_\gamma \leq \bar{e}_\xi < e_{\xi, \gamma}, e_\rho < e_{\xi, \gamma}\}. \quad (3.3.4)$$

Se \mathcal{C}^ξ denota o conjunto de \bar{A}^γ tal que $\bar{B}^\gamma = \bar{B}^\xi$ para cada $\xi = 1, \dots, k$, então, $\bar{e}_\gamma \leq \bar{e}_\xi$ expressa a seguinte situação: a equivalência máxima de \bar{A}^γ e qualquer $\bar{A}^\zeta \notin \mathcal{C}^\xi$ é no máximo tão grande quanto a equivalência máxima de \bar{A}^ξ para qualquer $\bar{A}^\zeta \notin \mathcal{C}^\xi$. A condição $\bar{e}_\xi < e_{\xi, \gamma}$ indica que a equivalência máxima de \bar{A}^ξ para qualquer $\bar{A}^\zeta \notin \mathcal{C}^\xi$ é menor que o grau de equivalência de \bar{A}^ξ e \bar{A}^γ .

Note que a cardinalidade de \mathcal{R}_ρ^ξ é decrescente com respeito a ρ e que \mathcal{R}_ρ^ξ contém ξ se $\bar{e}_\xi < 1$, pois $e_{\xi, \xi} = 1$. Em particular, temos que $\xi \in \mathcal{R}_\rho^\xi$ se E^γ é uma medida de equivalência forte e $\mathcal{C}^\xi \cap \mathcal{C}^\gamma = \emptyset$ para todo $B^\xi \neq B^\gamma$.

O Algoritmo 3.2 abaixo descreve como gerar um conjunto de memórias fundamentais $\mathcal{M} \subseteq \mathcal{T}$ para uma E-FAM. Uma breve olhada no Algoritmo 3.2 revela que sua complexidade no pior caso é de $O(k^2)$.

Algoritmo 3.2 Extração de Memórias Fundamentais do Conjunto de Treinamento

Require: Dado $\mathcal{T} = \{(\bar{A}^\xi, \bar{B}^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y), \xi = 1, \dots, k\}$;

- 1: Calcule $e_{\xi, \gamma}$ para $\xi, \gamma = 1, \dots, k$;
 - 2: Calcule \bar{e}_ξ para $\xi = 1, \dots, k$;
 - 3: Defina $\rho \in [0, 1)$;
 - 4: Calcule \mathcal{R}_ρ^ξ para $\xi = 1, \dots, k$;
 - 5: Defina $\mathcal{S}^\gamma = \{\xi \in \{1, \dots, k\} \mid \gamma \in \mathcal{R}_\rho^\xi\}$ para $\gamma = 1, \dots, k$;
 - 6: Calcule $s_\gamma = \sum_{\xi \in \mathcal{S}^\gamma} e_{\xi, \gamma}$ para $\gamma = 1, \dots, k$;
 - 7: Atribua $\mathcal{M} \leftarrow \emptyset$;
 - 8: **for** $\xi = 1, \dots, k$ **do**
 - 9: **for all** $\gamma \in \mathcal{R}_\rho^\xi$ **do**
 - 10: Calcule $v_{\xi, \gamma} \leftarrow e_{\xi, \gamma} s_\gamma$;
 - 11: **end for**
 - 12: Atribua $m_\xi \leftarrow \bigvee_{\zeta \in \mathcal{R}_\rho^\xi} v_{\xi, \zeta}$;
 - 13: Para cada $\gamma \in \mathcal{R}_\rho^\xi$ t.q. $v_{\xi, \gamma} = m_\xi$ atualize $\mathcal{M} \leftarrow \mathcal{M} \cup \{(\bar{A}^\gamma, \bar{B}^\gamma)\}$;
 - 14: **end for**
 - 15: **return** \mathcal{M} .
-

Exemplo 3.3.1. Neste exemplo vamos ilustrar o funcionamento do Algoritmo 3.2 por meio de um pequeno conjunto de treinamento artificial $\mathcal{T}^* \subset [0, 1]^2 \times [0, 1]$ mostrado na Figura 3.3. Especificamente, temos que

$$\mathcal{T}^* = \left\{ \begin{array}{l} (\bar{A}^1 = (0.66, 0.13)^t, 0), (\bar{A}^2 = (0.93, 0.33)^t, 0), \\ (\bar{A}^3 = (0.77, 0.37)^t, 0), (\bar{A}^4 = (0.47, 0.33)^t, 0), \\ (\bar{A}^5 = (0.27, 0.87)^t, 1), (\bar{A}^6 = (0.17, 0.80)^t, 1) \end{array} \right\}.$$

Seja E^ξ uma medida de equivalência em $[0, 1]^2$ dada por $E_{\lambda, \mathbf{w}}$ com $\lambda_1 = \lambda_2 = 1$ e $\mathbf{w} = (0.5, 0.5)$ para $\xi = 1, \dots, 6$. Ou seja, $E^\xi(A, B) = E_{(1,1), (0.5, 0.5)}(A, B) = 0.5(E_1(A_1, B_1) + E_1(A_2, B_2))$ para todo $\xi = 1, \dots, 6$ e $A, B \in [0, 1]^2$.

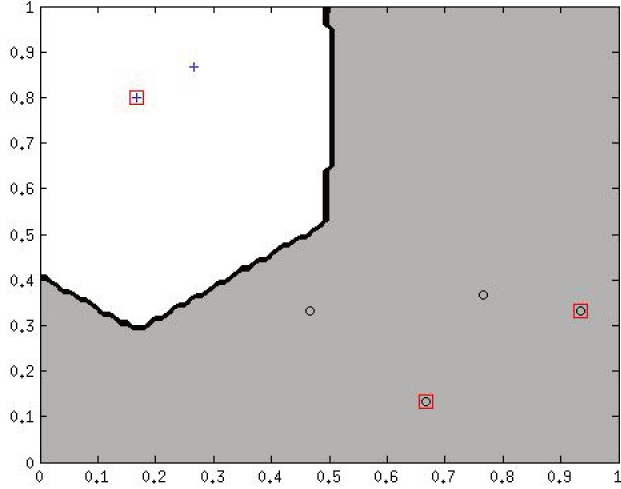


Figura 3.3: Os símbolos \circ e $+$ marcam as localizações dos padrões pertencentes às classes “0” e “1”, respectivamente. O Algoritmo 3.2 produz o conjunto \mathcal{M} consistindo dos três padrões marcados com quadrados vermelhos. A TE-FAM $E_{W,\nu}$ treinada utilizando o Algorithm TE com \mathcal{M} produz a superfície de decisão dividida pela linha preta.

Sejam $\mathbf{E} \in [0, 1]^{6 \times 6}$ e $\bar{\mathbf{e}} \in [0, 1]^6$, respectivamente, a matriz e o vetor cujas entradas $e_{i,j}$ e \bar{e}_i são computadas no Passos 1 e 2 do Algoritmo 3.2:

$$\mathbf{E} = \begin{pmatrix} 1.00 & 0.76 & 0.82 & 0.80 & 0.43 & 0.42 \\ 0.76 & 1.00 & 0.90 & 0.77 & 0.40 & 0.38 \\ 0.82 & 0.90 & 1.00 & 0.83 & 0.50 & 0.48 \\ 0.80 & 0.77 & 0.83 & 1.00 & 0.63 & 0.61 \\ 0.43 & 0.40 & 0.50 & 0.63 & 1.00 & 0.91 \\ 0.42 & 0.38 & 0.48 & 0.61 & 0.91 & 1.00 \end{pmatrix} \quad \text{e} \quad \bar{\mathbf{e}} = \begin{pmatrix} 0.43 \\ 0.40 \\ 0.50 \\ 0.63 \\ 0.63 \\ 0.61 \end{pmatrix}.$$

As Tabelas 3.1 e 3.2 mostram, respectivamente, os valores computados nos Passos 4–6 e 9–13 do Algoritmo 3.2, com $\rho = 0$. Logo, o Algoritmo 3.2 produziu $\mathcal{M} = \{(\bar{A}^1, 0), (\bar{A}^2, 0), (\bar{A}^6, 1)\}$.

Tabela 3.1: Resultados dos Passos 4, 5, e 6 do Algoritmo 3.2.

| ξ | 1 | 2 | 3 | 4 | 5 | 6 |
|------------------------|-----------|--------------|-----------|--------------|--------|--------|
| \mathcal{R}_ρ^ξ | {1, 2} | {2} | {1, 2, 3} | {1, 2, 3, 4} | {5, 6} | {6} |
| \mathcal{S}^ξ | {1, 3, 4} | {1, 2, 3, 4} | {3, 4} | {4} | {5} | {5, 6} |
| s_ξ | 2.63 | 3.43 | 1.83 | 1.00 | 1.00 | 1.91 |

Dado um conjunto de associações $\mathcal{T} = \{(\bar{A}^\xi, \bar{B}^\xi) : \xi = 1, \dots, k\}$ e uma E-FAM \mathcal{E} , o próximo teorema mostra que podemos utilizar o conjunto $\mathcal{M} \subseteq \mathcal{T}$ produzido pelo Algoritmo 3.2 para obter perfeita recordação de \bar{B}^ξ ao apresentar \bar{A}^ξ para todo $\xi = 1, \dots, k$. Mais precisamente, temos que:

Tabela 3.2: Resultados dos Passos 9-13 do Algoritmo 3.2.

| ξ | $\{v_{\xi,\gamma} \mid \gamma \in \mathcal{R}_\rho^\xi\}$ | m_ξ | Atualize \mathcal{M} |
|-------|--|---------|--|
| 1 | $\{\mathbf{v}_{1,1} = \mathbf{2.63}, v_{1,2} = 2.62\}$ | 2.63 | $\mathcal{M} \leftarrow \mathcal{M} \cup \{(A^1, 0)\}$ |
| 2 | $\{\mathbf{v}_{2,2} = \mathbf{3.43}\}$ | 3.43 | $\mathcal{M} \leftarrow \mathcal{M} \cup \{(\bar{A}^2, 0)\}$ |
| 3 | $\{v_{3,1} = 2.17, \mathbf{v}_{3,2} = \mathbf{3.09}, v_{3,3} = 1.83\}$ | 3.09 | $\mathcal{M} \leftarrow \mathcal{M} \cup \{(\bar{A}^2, 0)\}$ |
| 4 | $\{v_{4,1} = 2.11, \mathbf{v}_{4,2} = \mathbf{2.64}, v_{4,3} = 1.52, v_{4,4} = 1.00\}$ | 2.64 | $\mathcal{M} \leftarrow \mathcal{M} \cup \{(\bar{A}^2, 0)\}$ |
| 5 | $\{v_{5,5} = 1.00, \mathbf{v}_{5,6} = \mathbf{1.75}\}$ | 1.75 | $\mathcal{M} \leftarrow \mathcal{M} \cup \{(\bar{A}^6, 1)\}$ |
| 6 | $\{\mathbf{v}_{6,6} = \mathbf{1.91}\}$ | 1.91 | $\mathcal{M} \leftarrow \mathcal{M} \cup \{(\bar{A}^6, 1)\}$ |

Teorema 3.3.2. Seja $\mathcal{T} = \{(\bar{A}^\xi, \bar{B}^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y), \xi = 1, \dots, k\}$, onde X e Y são universos arbitrários. Suponha que E^γ seja uma medida de equivalência em $\mathcal{F}(X)$ tal que $E^\gamma(\bar{A}^\xi, \bar{A}^\gamma) < 1$ para todo $\bar{B}^\xi \neq \bar{B}^\gamma$, onde $\xi, \gamma = 1, \dots, k$. Se o Algoritmo 3.2 gera $\mathcal{M} = \{(A^\xi, B^\xi), \xi = 1, \dots, p\} \subseteq \mathcal{T}$, então, a correspondente E-FAM \mathcal{E} (cf. Equação (3.0.5)) com vetor de pesos $\mathbf{v} = (1, \dots, 1) \in \mathbb{R}^p$ recupera perfeitamente \bar{B}^ξ ao apresentar o padrão \bar{A}^ξ para todo $\xi = 1, \dots, k$, i.e.,

$$\mathcal{E}(\bar{A}^\xi) = \bar{B}^\xi \quad \forall \xi = 1, \dots, k. \quad (3.3.5)$$

Demonstração. Como $e_{\xi,\gamma} = E^\gamma(\bar{A}^\xi, \bar{A}^\gamma) < 1$ para todo $\xi, \gamma = 1, \dots, k$ tal que $\bar{B}^\xi \neq \bar{B}^\gamma$, temos que $\bar{e}_\xi < 1$ para todo $\xi = 1, \dots, k$ (cf. Equação (3.3.2)). Assim, temos que $\xi \in \mathcal{R}_\rho^\xi$ para $\xi = 1, \dots, k$ pois $\max\{\bar{e}_\xi, \rho\} < 1 = E^\xi(\bar{A}^\xi, \bar{A}^\xi) = e_{\xi,\xi}$ e $\bar{e}_\xi \leq \bar{e}_\xi$. Logo $\gamma \in \mathcal{S}^\gamma$ para todo $\gamma = 1, \dots, k$. Portanto, \mathcal{R}_ρ^ξ e \mathcal{S}^γ são não vazios para todo $\xi, \gamma = 1, \dots, k$ o que implica que os valores s_γ e $v_{\xi,\zeta}$ estão bem definidos para todo $\gamma, \xi = 1, \dots, k$ e $\zeta \in \mathcal{R}_\rho^\xi$.

Além disso, como $1 \leq |\mathcal{R}_\rho^\xi| \leq k$ temos que existe um índice ζ em \mathcal{R}_ρ^ξ tal que $m_\xi = v_{\xi,\zeta}$. Uma breve olhada no Passo 13 do Algoritmo 3.2 revela que $(\bar{A}^\zeta, \bar{B}^\zeta) \in \mathcal{M}$. Pela definição de \mathcal{R}_ρ^ξ , temos que $\bar{B}^\zeta = \bar{B}^\xi$ e $e_{\xi,\zeta} > \bar{e}_\xi$. Portanto, para qualquer associação $(A^\gamma, B^\gamma) \in \mathcal{M}$ tal que

$$\Theta^\gamma(\bar{A}^\xi) = E^\gamma(\bar{A}^\xi, A^\gamma) = \bigvee_{\kappa:(A^\kappa, B^\kappa) \in \mathcal{M}} \Theta^\kappa(\bar{A}^\xi) = E^\gamma(\bar{A}^\xi, A^\gamma) = \bigvee_{\kappa:(A^\kappa, B^\kappa) \in \mathcal{M}} E^\kappa(\bar{A}^\xi, A^\kappa), \quad (3.3.6)$$

temos que $B^\gamma = \bar{B}^\xi$, pois,

$$\bigvee_{\kappa:(A^\kappa, B^\kappa) \in \mathcal{M}} E^\kappa(\bar{A}^\xi, A^\kappa) \geq e_{\xi,\zeta} > \bar{e}_\xi. \quad (3.3.7)$$

Pelas definições de \mathcal{E} em termos da F dada na Equação (3.0.2), \mathbf{v} , e da composição sup-T (veja Equação (3.0.5)), temos que $\mathcal{E}(\bar{A}^\xi) = \bar{B}^\xi$. \square

Corolário 3.3.3. Seja $\mathcal{T} = \{(\bar{A}^\xi, \bar{B}^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y), \xi = 1, \dots, k\}$ tal que $\bar{A}^\xi \neq \bar{A}^\gamma$ para todo $\xi \neq \gamma$. Suponha que E^ξ é uma medida de equivalência forte em $\mathcal{F}(X)$ para todo $\xi = 1, \dots, k$. Se o Algoritmo 3.2 retorna $\mathcal{M} \subseteq \mathcal{T}$, então, a respectiva E-FAM \mathcal{E} com vetor de pesos $\mathbf{v} = (1, \dots, 1) \in \mathbb{R}^{|\mathcal{M}|}$ recupera perfeitamente \bar{B}^ξ ao apresentar o padrão \bar{A}^ξ para todo $\xi = 1, \dots, k$, i.e.,

$$\mathcal{E}(\bar{A}^\xi) = \bar{B}^\xi \quad \forall \xi = 1, \dots, k. \quad (3.3.8)$$

Demonstração. Como E^ξ é uma medida de equivalência forte e $\bar{A}^\xi \neq \bar{A}^\gamma$ para $\xi \neq \gamma$, temos que $E^\gamma(\bar{A}^\xi, \bar{A}^\gamma) < 1$ para todo $\xi \neq \gamma$. Em particular, $E^\gamma(\bar{A}^\xi, \bar{A}^\gamma) < 1$ para todo $\xi, \gamma = 1, \dots, k$ tal que $\bar{B}^\xi \neq \bar{B}^\gamma$. Logo, o Teorema 3.3.2 pode ser acionado para provar a afirmação do corolário. \square

3.3.2 Seleção de Parâmetros para TE-FAM $\mathcal{E}_{W,\mathbf{v}}$

Nesta subseção, vamos considerar o problema de ajustar os parâmetros de uma TE-FAM da forma $\mathcal{E}_{W,\mathbf{v}}$ conforme ilustrada na Figura 3.2. Para isto, suponha que já seja conhecido o conjunto de memórias fundamentais $\mathcal{M} = \{(A^\xi, B^\xi) \in [0, 1]^n \times \mathcal{F}(Y) \mid \xi = 1, \dots, p\}$, onde $n > 0$. Relembre que este modelo depende da escolha de vetores $\mathbf{w}^1, \dots, \mathbf{w}^p \in [0, 1]^n$, compilados em uma matriz $W = [\mathbf{w}^1, \dots, \mathbf{w}^p] \in [0, 1]^{n \times p}$, e $\mathbf{v} \in \mathbb{R}^p$. Como já mencionado anteriormente, a competição que ocorre na camada oculta pode ser expressa em termos da aplicação da função F dada na Equação (3.0.2). Infelizmente, tanto a função F quanto a composição sup-T não são diferenciáveis, dificultando assim a otimização dos parâmetros W e \mathbf{v} na Equação (3.3.1). Investigaremos a seguir como superar esses problemas.

Note primeiramente que, pela definição de F , temos que o conjunto de pontos fixo de F é $\{0, 1\}^p \setminus (0, \dots, 0)$. Além disso, se $\tilde{\mathbf{z}}^\xi = (v_1 E_{\mathbf{w}^1}(A^\xi, A^1), \dots, v_p E_{\mathbf{w}^p}(A^\xi, A^p))$ é um ponto fixo de F tal que $\tilde{z}_i^\xi = 0$ para todo $B^\xi \neq B^i$, então, obviamente temos que $\mathcal{E}_{W,\mathbf{v}}(A^\xi) = B^\xi$. Uma vez que essa última condição é suficiente para que a memória $\mathcal{E}_{W,\mathbf{v}}$ recorde perfeitamente B^ξ sob a apresentação de A^ξ , podemos formular uma função objetiva que visa encontrar W e \mathbf{v} tais que $\tilde{\mathbf{z}}^\xi$ é um ponto fixo onde $\tilde{z}_i^\xi = 0$ se e somente se $B^\xi \neq B^i$. Note que a condição $\tilde{z}_i^\xi = 1$ para $B^\xi = B^i$ pode ser vista como uma tentativa de distribuir as informações das memórias fundamentais com o mesmo consequente em $\mathcal{F}(Y)$ em mais de um neurônio oculto. Baseado nessas observações, propomos minimizar a seguinte função objetiva:

$$\sum_{\xi=1}^p \left(\frac{1}{d_\xi} \sum_{\gamma: B^\gamma = B^\xi} (1 - v_\xi E_{\mathbf{w}^\xi}(A^\gamma, A^\xi))^2 + \left(1 - \frac{1}{d_\xi}\right) \sum_{\gamma: B^\gamma \neq B^\xi} (v_\xi E_{\mathbf{w}^\xi}(A^\gamma, A^\xi))^2 \right), \quad (3.3.9)$$

onde $d_\xi = |\{\kappa : B^\kappa = B^\xi\}|$. Definindo $u_i^\xi = v_\xi w_i^\xi$ para $\xi = 1, \dots, p$ e $i = 1, \dots, n$, podemos expressar $v_\xi E_{\mathbf{w}^\xi}(A, A^\xi) = \sum_{i=1}^n v_\xi w_i^\xi E_i(A_i, A_i^\xi)$ em termos de um produto interno dos vetores $\mathbf{c}^\xi(A) = (E_1(A_1, A_1^\xi), \dots, E_n(A_n, A_n^\xi))^t$ e $\mathbf{u}^\xi = (u_1^\xi, \dots, u_n^\xi)^t$, *i.e.*, $v_\xi E_{\mathbf{w}^\xi}(A, A^\xi) = (\mathbf{c}^\xi(A))^t \mathbf{u}^\xi$. Além disso, seja $\mathbf{z}^\xi \in \{0, 1\}^p$ tal que $z_i^\xi = 0$ se e somente se $B^\xi \neq B^i$ para todo $\xi = 1, \dots, p$ e seja $D^\xi = \text{diag}(\tilde{d}_1^\xi, \dots, \tilde{d}_p^\xi) \in [0, 1]^{p \times p}$ onde:

$$\tilde{d}_i^\xi = \begin{cases} \frac{1}{d_\xi} & \text{se } B^i = B^\xi, \\ 1 - \frac{1}{d_\xi} & \text{caso contrário,} \end{cases} \quad \forall \xi = 1, \dots, p, i = 1, \dots, p. \quad (3.3.10)$$

Se $C^\xi = [\mathbf{c}^\xi(A^1), \dots, \mathbf{c}^\xi(A^p)]^t \in [0, 1]^{p \times n}$, então, a função objetiva acima pode ser reescrita como

$$\sum_{\xi=1}^p (\mathbf{z}^\xi - C^\xi \mathbf{u}^\xi)^t D^\xi (\mathbf{z}^\xi - C^\xi \mathbf{u}^\xi) \quad (3.3.11)$$

Aqui, seguiremos o seguinte procedimento para ajustar os parâmetros W e \mathbf{v} :

1. Encontrar $\mathbf{u}^* \in \mathbb{R}^{pn}$ que resolva o seguinte problema de otimização:

$$\begin{aligned} \min \frac{1}{2} (\mathbf{z} - \mathbf{C}\mathbf{u})^t \mathbf{D} (\mathbf{z} - \mathbf{C}\mathbf{u}) \\ \text{sujeito a } \mathbf{u} \geq 0. \end{aligned} \quad (3.3.12)$$

onde $\mathbf{z} \in \mathbb{R}^{p^2}$, $\mathbf{C} \in [0, 1]^{p^2 \times pn}$ e $\mathbf{D} \in \mathbb{R}^{p^2 \times p^2}$ são dados por

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}^1 \\ \vdots \\ \mathbf{z}^p \end{pmatrix}, \mathbf{C} = \begin{pmatrix} C^1 & 0 & \dots & 0 \\ 0 & C^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & C^p \end{pmatrix} \text{ e } \mathbf{D} = \begin{pmatrix} D^1 & 0 & \dots & 0 \\ 0 & D^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D^p \end{pmatrix}.$$

2. Para $\xi = 1, \dots, p$, defina o vetor $\mathbf{u}^\xi \in \mathbb{R}^n$ como $u_i^\xi = u_{p(1-\xi)+i}^*$ para todo $i = 1, \dots, n$.

3. Atribua $v_\xi = \|\mathbf{u}^\xi\|_1$ e $w_i^\xi = \frac{u_i^\xi}{v_\xi}$ para todo $\xi = 1, \dots, p$ e $i = 1, \dots, n$.

Encerraremos essa seção com alguns comentários referente ao Passo 1 do algoritmo acima, *i.e.*, a minimização da função objetiva $f(\mathbf{u}) = \frac{1}{2}(\mathbf{z} - \mathbf{C}\mathbf{u})^t \mathbf{D}(\mathbf{z} - \mathbf{C}\mathbf{u})$ sujeita à $u_i \geq 0$ para $i = 1, \dots, pn$. Primeiramente, note que a matriz Hessiana $\nabla^2 f(\mathbf{u}) = \mathbf{C}^t \mathbf{D} \mathbf{C}$ é semi-definida positiva, pois, \mathbf{D} também é semi-definida positiva. Portanto, estamos lidando com um problema de otimização convexo ou, mais precisamente, com um problema de programação quadrática [11]. Tal como acontece para qualquer problema de otimização convexo, todos os pontos de mínimos locais são também ótimos globais. Em outras palavras, todo ponto de mínimo local é também um ponto de mínimo global. Por fim, as restrições de desigualdades $u_i \geq 0$, para $i = 1, \dots, np$, são lineares e, portanto, as condições de Karush-Kuhn-Tucker podem ser usadas como critério suficiente de otimalidade [8].

Capítulo 4

Resultados Experimentais

Neste capítulo descreveremos vários experimentos que realizamos a fim de comparar a acurácia de certos modelos de Θ -FAMs com as de outros classificadores da literatura. Especificamente, consideramos os modelos de S-FAMs, S-FAMs duais e E-FAMs (*cf.* Definição 3.0.13). Relembre que todos estes modelos de Θ -FAMs são equipados com pesos ajustáveis v_1, \dots, v_p , como ilustrado na Figura 3.1, que podem ser ajustados usando o Algoritmo 3.1. Um outro algoritmo alternativo chamado Algoritmo TE, que foi apresentado na Seção 3.3, pode ser usado para treinar TE-FAMs do tipo $\mathcal{E}_{W, \mathbf{v}}$. Além de determinar os pesos v_1, \dots, v_p , o Algoritmo TE também determina automaticamente a topologia da TE-FAM, selecionando o conjunto de memórias fundamentais e ajustando os parâmetros \mathbf{w}^ξ nas funções de agregações da camada oculta ilustrada na Figura 3.2. Aqui, todas as TE-FAMs do tipo $\mathcal{E}_{W, \mathbf{v}}$ utilizadas em nossos experimentos foram treinadas usando o Algoritmo TE.

Em nossas simulações, empregamos S-FAMs (duais) baseadas nas medidas de *subsethood* de Kosko [63] e de Willmott [110] e também nas medidas de *subsethood* S_ϱ^\cap e S_ϱ^\cup (*i.e.*, Eqs. (2.2.9) e (2.2.10) com a implicação de Goguen e v_ϱ dado pela Eq. (2.2.11)) para $\varrho \in P = \{0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3, 4\}$. Adicionalmente, construímos medidas de equivalência E^S para E-FAMs seguindo a receita da Equação (2.2.18) com a t-norma do mínimo. As Θ -FAMs resultantes foram treinadas usando o Algoritmo 3.1 com $\mathbf{V} = [0, 1]^p$, onde p é o número de memórias fundamentais. Para cada Θ -FAM, utilizamos como vetor inicial $\mathbf{v}^{(0)} \in \mathbf{V}$ onde cada componente $v_j^{(0)}$ minimiza a função $f_j(v) = g_j(v) + h_j(v)$ sujeito a $v \in [0, 1]$, para $j = 1, \dots, p$. Aqui, as funções g_j e h_j são, respectivamente, dadas pelas Equações (3.2.16) e (3.2.20) com $v_\gamma = 1$ e $v_i = 1$ nas Equações (3.2.12), (3.2.17) e (3.2.21).

Além disso, também aplicamos TE-FAMs $\mathcal{E}_{W, \mathbf{v}}$ baseadas em medidas de equivalência $E_{\lambda, \mathbf{w}}$ dadas como na Equação (2.2.24). Esses modelos de TE-FAMs foram treinados usando o Algoritmo TE que foi detalhado na Seção 3.3. Relembre que o Algoritmo TE possui dois estágios: 1) extração de memórias fundamentais \mathcal{M} e 2) seleção de parâmetros ótimos. No Estágio 1, que corresponde à extração de \mathcal{M} via o Algoritmo 3.2, definimos o parâmetro ρ como a média dos valores $\bar{e}^1, \dots, \bar{e}^k$ que são obtidos no passo anterior (*i.e.*, Passo 2). Note que o parâmetro $\rho = \frac{1}{k} \sum_{\xi=1}^k \bar{e}^\xi$ atua como um limite inferior para a seleção de potenciais candidatos para substituir um certo elemento A^γ em termos de medidas de equivalência. Essa escolha de ρ parece razoável uma vez que, em problemas de classificação, cada valor \bar{e}^ξ denota o valor máximo de $E^\gamma(A^\xi, A^\gamma)$ para A^γ pertencente a uma

classe distinta daquela de A^ξ e que A^ξ pode ser substituída apenas por elementos da mesma classe. No Estágio 2, resolvemos o problema de otimização da Equação (3.3.12) usando o método *Trust-Region-Reflective Optimization* [18, 19] com o vetor inicial $(\frac{1}{n}, \dots, \frac{1}{n})^t \in [0, 1]^n$.

Como qualquer modelo de FAM lida apenas com classes de conjuntos *fuzzy*, faz-se necessário aplicar algum pré-processamento para converter os dados originais em conjuntos *fuzzy*. Cada conjunto de dados analisado neste capítulo consiste de instâncias ou vetores cujas entradas representam atributos numéricos ou categóricos. Formalmente, cada conjunto de dados considerado neste trabalho é dado por um subconjunto finito de $X_1 \times \dots \times X_n$, onde cada X_i denota um conjunto de valores numéricos, *i.e.*, $X_i = [t_{min}^i, t_{max}^i] \subset \mathbb{R}$, ou um conjunto de valores categóricos, *i.e.*, $X_i = \{l_1, \dots, l_c\}$ para algum $c > 0$.

Associamos cada elemento $\mathbf{x} \in X_1 \times \dots \times X_n$ com um conjunto *fuzzy* $\eta(\mathbf{x}) = (\eta_1(x_1)^t, \dots, \eta_n(x_n)^t)$ que é dado pela concatenação de $\eta_j(x_j)$ para $j = 1, \dots, n$. Seja $i \in \{1, \dots, n\}$. Por um lado, se X_i é da forma $[t_{min}^i, t_{max}^i] \subset \mathbb{R}$, então,

$$\eta_i(x) = \frac{(x - t_{min}^i)}{(t_{max}^i - t_{min}^i)} \in [0, 1]$$

para todo $x \in X_i$. Por outro lado, se X_i é da forma $\{l_1, \dots, l_c\}$, $c > 0$, então, $\eta_i(l_j) = B^j \in \{0, 1\}^c \subseteq [0, 1]^c$ onde

$$\mu_{B^j}(i) = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{caso contrário.} \end{cases} \quad (4.0.1)$$

Note que, em problemas de classificação, o rótulo da classe é dado por um atributo categórico. Logo, identificamos cada classe com um conjunto *fuzzy* segundo a Equação (4.0.1).

Em [98], propusemos o uso adicional de uma função de normalização Ψ como uma tentativa de reduzir os efeitos da variação de luminosidade em padrões de imagens. Especificamente, definimos Ψ como uma função $[0, 1]^n \rightarrow [0, 1]^n$ dada componente a componente por

$$\Psi(\mathbf{a})_i = \begin{cases} a_i & \text{se } a_i = 0 \text{ ou } a_i = 1 \\ \frac{a_i - m_{\mathbf{a}} + 1.5}{3} & \text{caso contrário} \end{cases}, \quad \forall i = 1, \dots, n \text{ e } \mathbf{a} \in [0, 1]^n, \quad (4.0.2)$$

onde $m_{\mathbf{a}}$ denota a média de $\mathbf{a} \in [0, 1]^n$ sob a restrição do domínio $I = \{i : 0 < a_i < 1\}$, isto é,

$$m_{\mathbf{a}} = \frac{\sum_{i \in I} a_i}{|I|}.$$

Uma vez que o uso de Ψ favoreceu à aplicação bem sucedida de uma S-FAM baseada na medida de *subsethood* S_K de Kosko em um problema de auto-localização de robô baseado em visão [98], consideramos também a possibilidade do uso da função de normalização Ψ após a aplicação da função n em nossas simulações. Neste caso, obtemos a seguinte sequência:

$$\mathbf{x} \xrightarrow{\eta} \eta(\mathbf{x}) \xrightarrow{\Psi} \Psi(\eta(\mathbf{x})). \quad (4.0.3)$$

4.1 Alguns de Problemas de Classificação de Referência

Nesta seção aplicamos os modelos Θ -FAMs em vários problemas de classificação que estão disponíveis na internet no repositório *Knowledge Extraction Based on Evolutionary Learning (KEEL)* [2]. Especificamente, consideramos os seguintes conjuntos de dados: Iris, Monks, Appendicitis, Ecoli, Pima, Glass, Wine, Heart, Cleveland, Vowel, Crx, Wdbc, Spectfheart, Sonar e Movementlibras. A Tabela 4.1 exibe as informações sobre números de instâncias, atributos numéricos e categóricos e classes para cada conjunto de dados considerado.

Tabela 4.1: Descrição dos conjuntos de dados.

| | Instâncias | Atributos Categóricos | Atributos Numéricos | Classes |
|----------------|------------|--------------------------|------------------------|---------|
| Appendicitis | 106 | 0 | 7 | 2 |
| Cleveland | 297 | 0 | 13 | 5 |
| Crx | 653 | 9 | 6 | 2 |
| Ecoli | 336 | 0 | 7 | 8 |
| Glass | 214 | 0 | 9 | 7 |
| Heart | 270 | 0 | 13 | 2 |
| Iris | 150 | 0 | 4 | 3 |
| Monks | 432 | 0 | 6 | 2 |
| Movementlibras | 360 | 0 | 90 | 15 |
| Pima | 768 | 0 | 8 | 2 |
| Sonar | 208 | 0 | 60 | 2 |
| Spectfheart | 267 | 0 | 44 | 2 |
| Vowel | 990 | 0 | 13 | 11 |
| Wdbc | 569 | 0 | 30 | 2 |
| Wine | 178 | 0 | 13 | 3 |

Comparamos os resultados de classificação produzidos pelas Θ -FAMs com os produzidos por vários classificadores competitivos que foram discutidos em um artigo recente de Alcalá-Fdez *et al.* [1]. São estes: árvore de decisão C4.5 [83], algoritmo de aprendizagem estrutural em ambiente vago (2SLAVE) [34], classificação baseada em associações (CBA) [68], uma versão melhorada do método CBA (CBA2) [69], classificação baseada em regras de associação múltipla (CMAR) [65], classificação baseada em regras de associação preditiva (CPAR) [113], algoritmo genético híbrido *fuzzy* baseada em algoritmo de aprendizagem de máquina (FH-GBML) [45], algoritmo genético de estado equilibrado para extração de regras de classificação *fuzzy* de dados (SGERD) [72], método de classificação baseado em regras de associação *fuzzy* para problemas com alta dimensão (FARC-HD) [1]. A fim de assegurar uma comparação justa com nossos modelos de Θ -FAMs, seguimos a abordagem sugerida no artigo mencionado anteriormente [1] calculando o desempenho de classificação para cada Θ -FAM sob um particionamento dos dados em 10 *folds* (*i.e.*, subconjuntos disjuntos de mesmo tamanho). Aqui, utilizamos o mesmo particionamento que em [2, 1] e conduzimos os experimentos como se segue para cada conjunto de dados. Cada *fold* serviu como um conjunto de

teste para calcular a taxa de classificação, enquanto, os *fold*s restantes, cujo tamanho consiste de 9/10 de todo conjunto de dados, foram utilizados com conjunto de treinamento. Sumarizando, calculamos a taxa de classificação média para cada Θ -FAM utilizando o método de validação cruzada com 10 *fold*s, onde foram realizados três simulações para cada *fold*.

Neste contexto, lembre que o uso do Algoritmo 3.1 na fase de treinamento de uma Θ -FAM requer a especificação de um conjunto de memórias fundamentais \mathcal{M} além do conjunto de treinamento \mathcal{T} . Para as Θ -FAMs que foram treinadas usando o Algoritmo 3.1 procedemos da seguinte maneira. Particionamos arbitrariamente \mathcal{T} , *i.e.*, a união dos *fold*s restantes, em duas metades. Primeiro, uma metade foi associada com \mathcal{M} e uma Θ -FAM, cujos neurônios ocultos correspondem a \mathcal{M} , foi treinada usando \mathcal{T} . Se $|\mathcal{T}| = 2p$, então, este passo produz p pesos conectados com p neurônios escondidos determinados por \mathcal{M} (*cf.* Figura 3.1). Segundo, tomamos a outra metade como \mathcal{M} e treinamos uma outra Θ -FAM com os correspondentes neurônios ocultos usando todo o conjunto de treinamento \mathcal{T} , produzindo assim p pesos que se conectam aos p neurônios ocultos. Finalmente, construímos ainda uma outra Θ -FAM com $2p$ neurônios ocultos correspondendo a \mathcal{T} . Para esta Θ -FAM resultante, simplesmente tomamos a união dos pesos determinados no primeiro e segundo passos. Esta estratégia produz pesos para uma Θ -FAM com $2p$ neurônios ocultos, visando assegurar um desempenho de generalização adequado.

Para cada modelo de Θ -FAM treinado utilizando o Algoritmo 3.1, realizamos o procedimento descrito acima três vezes para cada um dos dez *fold*s e calculamos a média da taxa de classificação sob os 10 conjuntos de treinamento. Além disso, para estes modelos, conduzimos nossas simulações considerando o uso ou não da normalização Ψ após a aplicação de η . Então, extraímos o conjunto de Θ -FAMs (entre o conjunto de S-FAMs, S-FAMs duais e E-FAMs baseadas em S_K , S_W , S_ρ^\cap ou S_ρ^\cup para $\rho \in P$) que produziram o menor erro de classificação médio com a menor variância durante a fase de treinamento. A taxa de classificação média para todas as Θ -FAMs treinadas usando o Algoritmo 3.1 podem ser encontradas no Apêndice A desta tese e também na internet [23].

As colunas “ Θ -FAM” e “TE-FAM” da Tabela 4.2 exibe a taxa de classificação média na fase de teste, respectivamente, para o conjunto de Θ -FAMs selecionadas entre aquelas treinadas com o Algoritmo 3.1 e para as TE-FAMs $\mathcal{E}_{W,v}$ que foram treinadas usando o Algoritmo TE.

Em primeiro lugar, a Tabela 4.2 revela que os modelos de Θ -FAMs com os pesos v_1, \dots, v_p ajustados usando o Algoritmo 3.1 produziram um desempenho geral de classificação muito satisfatório, exibindo a maior precisão de classificação, com média de 82.60%. Apenas as taxas de classificação para os conjuntos de dados Appendicitis e Pima ficaram ligeiramente atrás das atingidas pelos outros classificadores. Os resultados restantes exibidos por estes modelos de Θ -FAMs podem ser vistos como altamente competitivos. De fato, estes modelos de Θ -FAMs atingiram o melhor desempenho de classificação em cinco dos quinze conjuntos de dados considerados, exibindo uma média de erros de classificação excepcionalmente baixa em comparação aos demais modelos para os conjuntos de dados Movementlibras e Vowel.

Além disso, note que os modelos TE-FAMs, que também são casos particulares de Θ -FAMs segundo a Definição 3.0.13, tiveram um desempenho relativamente bom, exibindo os melhores resultados de classificação média em cinco conjuntos de dados (Cleveland, Iris, Pima, Sonar e Spectfheart), além de produzir a segunda melhor classificação média geral de 82.25%, o que corresponde a uma ligeira melhora do resultado produzido pelo classificador FARC-HD.

A Tabela 4.2 mostra que ambas abordagens para treinar certos modelos de Θ -FAMs produziram

Tabela 4.2: Taxa de classificação obtida pelos modelos Θ -FAMs em comparação com aquelas apresentadas em [1].

| | 2SLAVE | FH-GBML | SGERD | CBA | CBA2 | CMAR | CPAR | C4.5 | FARC-HD | Θ -FAM | TE-FAM |
|----------------|--------------|---------|-------|------------|-------------|-------------|-------------|------------|-------------|---------------|--------------|
| Appendicitis | 82.91 | 86 | 84.48 | 89.6 | 89.6 | 89.7 | 87.8 | 83.3 | 84.2 | 81.18 | 84.09 |
| Cleveland | 48.82 | 53.51 | 51.59 | 56.9 | 54.9 | 53.9 | 54.9 | 54.5 | 55.2 | 51.17 | 61.01 |
| Crx | 74.06 | 86.6 | 85.03 | 83.6 | 85 | 85 | 87.3 | 85.3 | 86 | 82.02 | 86.66 |
| Ecoli | 84.53 | 69.38 | 74.05 | 78 | 77.1 | 77.7 | 76.2 | 79.5 | 82.2 | 76.78 | 80.97 |
| Glass | 58.05 | 57.99 | 58.49 | 70.8 | 71.3 | 70.3 | 68.9 | 67.4 | 70.2 | 70.49 | 63.85 |
| Heart | 71.36 | 75.93 | 73.21 | 83 | 81.5 | 82.2 | 80.7 | 78.5 | 84.4 | 78.15 | 82.59 |
| Iris | 94.44 | 94 | 94.89 | 93.3 | 93.3 | 94 | 96 | 96 | 96 | 96 | 96 |
| Monks | 97.26 | 98.18 | 80.65 | 100 | 100 | 100 | 100 | 100 | 99.8 | 98.63 | 97.73 |
| Movementlibras | 67.04 | 68.89 | 68.09 | 36.1 | 7.2 | 39.2 | 63.6 | 69.4 | 76.7 | 84.72 | 69.17 |
| Pima | 73.71 | 75.26 | 73.37 | 72.7 | 72.5 | 75.1 | 74.5 | 74 | 75.7 | 67.44 | 76.18 |
| Sonar | 71.42 | 68.24 | 71.9 | 75.4 | 77.9 | 78.8 | 75 | 70.5 | 80.2 | 80.69 | 83.62 |
| Spectfheart | 79.17 | 72.36 | 78.16 | 79.8 | 79.8 | 79.4 | 78.3 | 76.5 | 79.8 | 81.3 | 81.67 |
| Vowel | 71.11 | 67.07 | 65.83 | 63.6 | 74.9 | 60.4 | 63.00 | 81.5 | 71.8 | 97.07 | 80 |
| Wdbc | 92.33 | 92.26 | 90.68 | 94.7 | 95.1 | 94.9 | 95.1 | 95.2 | 95.3 | 96.14 | 93.67 |
| Wine | 89.47 | 92.61 | 91.88 | 93.8 | 93.8 | 96.7 | 95.6 | 93.3 | 94.3 | 97.24 | 96.63 |
| Média | 77.05 | 77.22 | 76.15 | 78.09 | 76.93 | 78.49 | 79.79 | 80.33 | 82.12 | 82.60 | 82.25 |

resultados similares em termos de classificação média geral. De fato, os resultados obtidos pelas TE-FAMs do tipo $\mathcal{E}_{W,v}$ superaram aqueles obtidos pelos outros tipos de Θ -FAMs em oito dos quinze conjuntos de dados: Appendicitis, Cleveland, Crx, Ecoli, Heart, Pima, Sonar e Spectfheart. Em contraste, as taxas de classificação produzidas pelo modelo TE-FAM para seis dos quinze conjuntos de dados, principalmente os conjuntos de dados Movementlibras e Vowel, foram inferiores das produzidas pelas Θ -FAMs treinadas com o Algoritmo 3.1. Note que os conjuntos de dados Movementlibras e Vowel possuem o maiores números de classes entre os 15 conjuntos de dados. Este fato sugere que a tarefa de classificação dos dados do Movementlibras e do Vowel é bastante complicada e, portanto, a tarefa de extração de regras para sistemas de classificação baseados em regras se torna também difícil. Esta última observação é corroborada pelos erros relativamente altos que foram produzidos pela maioria dos modelos listados nas Tabelas IV e X de [1] nos conjuntos de treinamento do Movementlibras e Vowel. Neste contexto, note que o primeiro estágio do Algoritmo TE que lida com a extração de memórias fundamentais do conjunto de treinamento pode ser visto como um procedimento de extração de regras, onde cada regra corresponde a uma memória fundamental, justificando assim também o baixo desempenho dos modelos TE-FAMs. Para estes conjuntos de dados, os outros modelos de Θ -FAMs, que foram treinados com o Algoritmo 3.1, produziram as maiores taxas de classificação tanto na fase de teste quanto na fase de treinamento, atingindo na última 98.95% e 99.99% de acurácia, respectivamente (veja as Tabelas A.10 e A.14). A segunda melhor taxa média de classificação nos dados de treinamento do Movementlibras e Vowel foi produzida pela árvore de decisão C4.5, porém, sua capacidade de generalização as vezes pode ser afetada pela forma da sua fronteira de decisão, composta por hiperplanos [78]. Finalmente, note que o modelo FARC-HD não apenas produziu os menores erros de treinamento e teste para o conjunto de dados Movementlibras mas também produziu o terceiro melhor desempenho geral, atingindo as maiores taxas de classificação nos conjuntos de dados Heart e Iris [1].

Todos os experimentos referentes aos modelos Θ -FAMs foram implementados utilizando MATLAB[®]. A Tabela 4.3 exibe o tempo médio de CPU em cada estágio da fase de treinamento das

TE-FAMs do tipo $\mathcal{E}_{W,v}$ para cada conjunto de dados (usando um computador com CPU Intel[®] Core[™] i5-3330 com velocidade de processamento de 3.00 GHz). Note que o método de validação cruzada com 10 *fold*s se mostrou útil para selecionar uma Θ -FAM, com os pesos v_1, \dots, v_p ajustados pelo Algoritmo 3.1, apropriada para cada conjunto de dados. Apesar do modelo Θ -FAM treinado com o algoritmo geral apresentado na Seção 3.2 ter produzido a maior taxa de classificação geral de 82.60% nos quinze conjuntos de dados, esta abordagem de treinamento envolve um enorme custo computacional em comparação à Tabela 4.3. Mesmo para um escolha fixa de funções Θ^ξ , a fase de treinamento requer uma grande quantidade de tempo de processamento. Por exemplo, a fase de treinamento de uma Θ -FAM baseada na medida de *subsethood* de Kosko [64] requer um tempo de CPU de aproximadamente 1 hora apenas para o conjunto de dados Vowel.

Tabela 4.3: O valor médio da fração $|\mathcal{M}|/|\mathcal{T}|$ e os tempos de CPU para treinamento.

| | Instâncias | Média % $ \mathcal{M} / \mathcal{T} $ | Estágio 1 (seg.) | Estágio 2 (seg.) | Total (seg.) |
|----------------|------------|---------------------------------------|------------------|------------------|--------------|
| Appendicitis | 106 | 62.16 | 0.01 | 0.12 | 0.12 |
| Cleveland | 297 | 78.71 | 0.03 | 0.29 | 0.32 |
| Crx | 653 | 63.42 | 0.48 | 4.04 | 4.51 |
| Ecoli | 336 | 65.15 | 0.02 | 0.22 | 0.24 |
| Glass | 214 | 68.07 | 0.01 | 0.17 | 0.19 |
| Heart | 270 | 68.23 | 0.02 | 0.22 | 0.24 |
| Iris | 150 | 38.89 | 0.01 | 0.09 | 0.10 |
| Monks | 432 | 87.22 | 0.03 | 0.24 | 0.27 |
| Movementlibras | 360 | 64.88 | 0.31 | 8.27 | 8.58 |
| Pima | 768 | 73.77 | 0.14 | 0.50 | 0.64 |
| Sonar | 208 | 72.44 | 0.09 | 0.98 | 1.06 |
| Spectfheart | 267 | 88.43 | 0.10 | 1.45 | 1.55 |
| Vowel | 990 | 35.54 | 0.35 | 0.44 | 0.80 |
| Wdbc | 569 | 46.83 | 0.25 | 1.18 | 1.43 |
| Wine | 178 | 45.38 | 0.01 | 0.15 | 0.16 |

Apesar de Alcalá-Fdez *et al.* também reportarem os tempos de CPU requeridos pelos demais modelos listados na Tabela 4.2 para cada um dos quinze conjuntos de dados [1], estes tempos de CPU não podem ser comparados com aqueles listados na Tabela 4.3, pois, as simulações deles foram realizadas utilizando um ferramenta de software Java (GPLv3) *open source* chamada *KEEL* [2].

4.2 Um Problema de Auto-localização de Robô Baseado em Visão

Recentemente, aplicamos com sucesso o modelo FAM de *subsethood* de Kosko (KS-FAM), que corresponde a um modelo S-FAM baseado na medida de *subsethood* de Kosko com o vetor de pesos $\mathbf{w} = (1, \dots, 1)^t$, em um problema de identificação de localização (ou posições de referências) de um robô móvel baseado nas imagens coletadas durante sua operação [98]. Villaverde *et al.*

aplicaram várias abordagens a mapas topológicos baseados em aparência neste problema utilizando exatamente o mesmo conjunto de dados [109, 108] e desenvolveram algoritmos de mapeamentos *online* e *offline*.

Cada posição do mapa correspondente a uma certa localização é associado com um conjunto de imagens que foram capturadas pela câmera do robô durante a fase de treinamento. A tarefa de auto-localização é encarregada de confrontar as imagens armazenadas no mapa com aquelas adquiridas pelo robô enquanto ele está se movendo. Uma das possíveis abordagens à construção de um mapa baseado em aparência para auto-localização de robô baseado em visão consiste em armazenar pares imagem-posição em termos de um memória associativa tal como uma Θ -FAM. Aqui, empregamos esta metodologia em conjunto com Θ -FAMs e os resultados experimentais são descritos abaixo.

O conjunto de dados em questão compreende 2265 imagens de tamanho 61×79 que foram associadas com certas posições de referências em um prédio. Como o número total de posições de referências é 11, este problema de auto-localização baseado em visão pode ser visto como um problema de classificação com 11 classes. As imagens foram capturadas pela câmera óptica *on-board* de um robô móvel ao longo de seis percursos de um caminho específico em um prédio tal como ilustrado na Figura 4.1. A Tabela 4.4 mostra o número de imagens para cada percurso e posição no prédio.

Tabela 4.4: Número de imagens por percurso e posição.

| <i>Posição:</i> | 1 ^a | 2 ^a | 3 ^a | 4 ^a | 5 ^a | 6 ^a | 7 ^a | 8 ^a | 9 ^a | 10 ^a | 11 ^a | Total |
|-------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|-------|
| Percurso 1 | 32 | 15 | 37 | 30 | 29 | 45 | 21 | 49 | 36 | 36 | 32 | 362 |
| Percurso 2 | 30 | 9 | 33 | 26 | 27 | 42 | 34 | 58 | 27 | 32 | 39 | 357 |
| Percurso 3 | 19 | 27 | 38 | 36 | 35 | 47 | 38 | 59 | 36 | 35 | 34 | 404 |
| Percurso 4 | 33 | 22 | 30 | 40 | 30 | 45 | 33 | 49 | 35 | 26 | 42 | 385 |
| Percurso 5 | 26 | 22 | 39 | 38 | 36 | 46 | 39 | 41 | 27 | 34 | 35 | 383 |
| Percurso 6 | 32 | 21 | 35 | 33 | 32 | 39 | 36 | 50 | 30 | 32 | 34 | 374 |

Uma aplicação de um modelo Θ -FAM ao presente problema requer primeiro um pré-processamento das imagens em tons de cinza e a associação das imagens obtidas a um rótulo de classe. Primeiramente, convertamos as imagens em vetores $\mathbf{x}^\xi \in \mathbb{R}^{4819}$ através do método *row scan*. Sujeitando \mathbf{x}^ξ ao esquema dado na Equação (4.0.3), obtivemos os conjuntos *fuzzy* $\tilde{A}^\xi \in [0, 1]^{4819}$. Cada informação de posição, que correspondente a um rótulo de classe no conjunto $\{1, \dots, 11\}$, foi convertido em um conjunto *fuzzy* $\tilde{B}^\xi \in [0, 1]^{11}$ como descrito na Equação (4.0.1).

Aqui, consideramos o conjunto de Θ -FAMs consistindo de S-FAMs, S-FAMs duais e E-FAMs baseadas, respectivamente, em medidas de *subsethood* S e medidas de equivalência E^S , onde $S = S_K, S_W, S_\rho^\cap$ ou S_ρ^\cup para $\rho \in P$. Além disso, especificamos o conjuntos de memórias fundamentais, treinamento, validação e teste. Atribuimos o conjunto de associações $\{(\bar{A}^\xi, \bar{B}^\xi) : \xi = 1, \dots, 362\}$ correspondentes ao Percurso 1 como sendo o conjunto de treinamento $\mathcal{T} = \cup_{i=1}^{11} \mathcal{T}_i$, onde \mathcal{T}_i representa o conjunto de associações correspondente a i -ésima posição de referência. Os elementos de

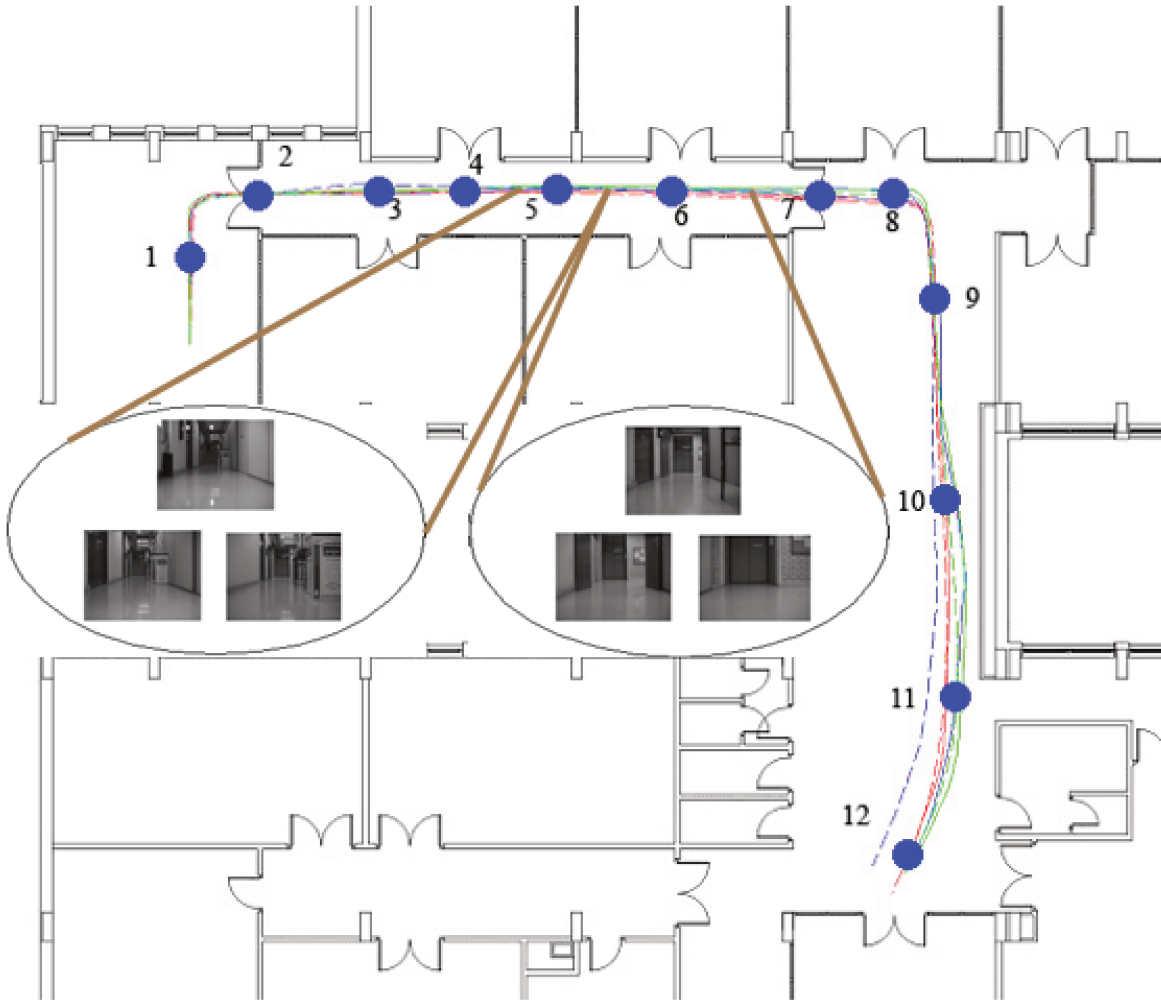


Figura 4.1: Posições de referência para regiões do particionamento. (1) Laboratório. (2) Porta do Laboratório. (3) Porta do Laboratório. (4) Porta do Laboratório. (5) Impressoras. (6) Portas do Laboratório. (7) Porta do corredor. (8) Porta do Laboratório. (9) Gabinete do roteador. (10) Sala de café. (11) Cruzamento. (12) Salão.

\mathcal{T}_i são numerados de acordo com a sequência com que as imagens foram obtidas. Cada \mathcal{T}_i foi dividido em duas metades: uma consistindo das associação com numeração par e a outra consistindo das associação com numeração ímpar. Dois experimentos independentes foram executados com cada uma das metades representando o conjunto de memórias fundamentais \mathcal{M} e o desempenho de classificação foi avaliado usando o conjunto de validação. Aqui, adotamos como conjunto de validação as associações correspondentes ao Percurso 2. Os Percursos de 3 a 6 foram empregados para teste.

Uma aplicação do Algoritmo 3.1 resulta em um vetor de pesos \mathbf{v} em $[0, 1]^{181}$ para cada uma das 60 Θ -FAMs e cada escolha de \mathcal{M} , tal como explicado acima. Assim, cada Θ -FAM resultante provê uma representação ponderada das funções Θ^ξ para $\xi = 1, \dots, 181$ segundo a importância da contribuição de cada Θ^ξ . Os modelos de $S_{1.5}^\cap$ -FAM dual e $S_{1.5}^\cup$ -FAM dual, isto é, os modelos de Θ -

FAM baseados, respectivamente, em $S_{1.5}^{\cap}(\cdot, A^{\xi})$ e $S_{1.5}^{\cup}(\cdot, A^{\xi})$ para $\xi = 1, \dots, 181$, exibiram a maior precisão de classificação no conjunto de validação com o valor de de 90.19% e também obiveram uma precisão de 100% de classificação no conjunto de treinamento. As taxas de classificação para todas as Θ -FAMs testadas podem ser encontradas na internet [23] e também nas Tabelas A.18 - A.23 do Apêndice A. A Tabela 4.5 compara o desempenho médio de classificação produzido pelas $S_{1.5}^{\cap}$ -FAM dual e $S_{1.5}^{\cup}$ -FAM dual com os resultados produzidos por outros modelos da literatura [98, 108, 109].

Uma das abordagens discutidas em [108] é baseada *lattice independent component analysis* (LICA) [38, 35]. LICA se refere à extração de vetores *lattice independent* para decomposição linear. Esta abordagem produz um vetor de atributos que são usados para classificação e comparação de imagens, ambas para construção *offline* de mapa [108] e simultânea localização e mapeamento (SLAM) [109]. Aqui, o processo SLAM implica que as posições de referências são descobertas e a transformação LICA é atualizada durante o vôo usando a caracterização dos pontos fixos das memórias auto-associativas morfológicas [103, 109]. Além disso, duas variantes do procedimento de *análise de componentes independentes* (ICA), a *Mean Field* ICA (MF-ICA) [43] e Mogeldey-Schuster ICA (MS-ICA) [75], foram aplicadas em [108]. ICA procura por regressores estatisticamente independentes e a correspondente matriz de mistura.

Os resultados de classificação obtidos usando as abordagens ICA e LICA dependem, respectivamente, do número de componentes independentes e das escolhas de um parâmetro de tolerância a ruído α e de um *endmember* inicial para o algoritmo *Endmember Induction Heuristic Algorithm* (EIHA) [39]. Portanto, optamos por selecionar os parâmetros que produziram os melhores resultados no conjunto de validação, isto é, nos dados correspondentes ao Percurso 2. Como anteriormente, utilizamos os dados do Percurso 1 como conjunto de treinamento e os Percursos de 3 a 6 foram usados na fase de teste. Aqui, nós nos baseamos nos resultados de classificação previamente publicados referentes às abordagens ICA (MF-ICA and MS-ICA), com números de componentes independentes variando entre 5, 10, \dots , 25, e referente à LICA com $\alpha = 7$ e 8 e com 5 diferentes *endmembers* iniciais para cada α [108]. Os melhores desempenhos nos dados de validação dos métodos baseados em MF-ICA e MS-ICA foram em ambos os casos obtidos usando 25 componentes independentes. A abordagem LICA produziu a maior precisão de classificação nos dados do Percurso 2 usando o primeiro dos cinco *endmembers* inicializados e $\alpha = 7$. Os resultados de classificação para a abordagem baseada em SLAM foram extraídos de [109].

A Tabela 4.5 revela que o modelo Θ -FAM produziu uma melhora significativa em comparação às demais abordagens para auto-localização de robô móvel baseado em visão que foram previamente aplicadas a este problema pela Universidade do País Basco, San Sebastián, Espanha.

Tabela 4.5: Comparação dos resultados produzidos pelas Θ -FAMs selecionadas com os melhores resultados obtidos para cada Percurso usando LICA [108], MF-ICA [43] e MS-ICA [75], e SLAM [109].

| | Percurso 3 | Percurso 4 | Percurso 5 | Percurso 6 | Média |
|----------------------------|-------------|-------------|-------------|-------------|-------------|
| $S_{1.5}^{\cap}$ -FAM dual | 0.78 | 0.72 | 0.78 | 0.77 | 0.76 |
| $S_{1.5}^{\cup}$ -FAM dual | 0.79 | 0.72 | 0.79 | 0.77 | 0.77 |
| LICA | 0.75 | 0.66 | 0.73 | 0.75 | 0.72 |
| MF-ICA | 0.62 | 0.54 | 0.65 | 0.53 | 0.58 |
| MS-ICA | 0.69 | 0.62 | 0.74 | 0.69 | 0.68 |
| SLAM | 0.76 | 0.6 | 0.69 | 0.64 | 0.67 |

Capítulo 5

Conclusão

Nesta tese apresentamos uma classe de memórias associativa *fuzzy* chamada Θ -FAM que é projetada para armazenar um conjunto de memórias fundamentais $(A^\xi, B^\xi) \in \mathcal{F}(X) \times \mathcal{F}(Y)$, para $\xi = 1, \dots, p$, e para exibir um desempenho de generalização adequado na apresentação de um novo padrão $A \in \mathcal{F}(X)$. Se $(A^\xi, B^\xi) \in [0, 1]^n \times [0, 1]^m$, então, a Θ -FAM pode ser vista como uma rede neural *fuzzy* [12] com p neurônios ocultos competitivos com funções de agregação Θ^ξ para $\xi = 1, \dots, p$. A contribuição do ξ -ésimo neurônio oculto pode ser ajustada em termos de um fator de ponderação v_ξ antes de ocorrer uma competição entre os neurônios ocultos. O peso de conexão do ξ -ésimo neurônio oculto com i -ésimo neurônio de saída é igual a B_i^ξ , *i.e.*, o i -ésimo componente do conjunto *fuzzy* B^ξ . Devido à natureza competitiva dos neurônios ocultos em termos da função F dada pela Equação (3.0.2), apenas aqueles com valores máximos são ativados.

Adicionalmente, providenciamos condições suficientes para se obter a recordação perfeita das memórias fundamentais. Nós também provamos que sob certas condições, dependendo das funções Θ^ξ e do vetor de pesos $\mathbf{v} = (v_1, \dots, v_p)^t$, o raio de atração das memórias fundamentais é positiva. Nesta tese, focamos em funções Θ^ξ dadas por medidas de *subsethood* ou medidas de equivalência. Note que outras escolhas de Θ^ξ também seriam possíveis [46, 106]. Todavia, uma caracterização geral de Θ^ξ , tal como propusemos neste trabalho, é importante pelo menos por dois motivos. Primeiro, esta caracterização assegura os requisitos mínimos para se obter os resultados dos Teoremas 3.1.1 e 3.1.2 referentes às capacidades de armazenamento e correção de erros das Θ -FAMs. Segundo, permite estender a abordagem Θ -FAM e inclusive os Teoremas 3.1.1 e 3.1.2 para domínios mais gerais, onde os conceitos de medidas de *subsethood* ou equivalência não estão necessariamente definidos.

Se as funções Θ^ξ são escolhidas como sendo $S(\cdot, A^\xi)$, $S(A^\xi, \cdot)$ ou $E(\cdot, A^\xi)$, onde S e E são, respectivamente, dados por medidas de *subsethood* e equivalência, então, falamos de S-FAM, S-FAM dual e E-FAM, respectivamente. Em particular, medidas de *subsethood* e equivalências parametrizadas tais como aquelas construídas nesta tese dão origem às TS-FAMs, TS-FAMs duais e TE-FAMs. Os pesos v_i de uma Θ -FAM podem ser determinados a partir de um algoritmo de treinamento geral para Θ -FAM introduzido na Seção 3.2. Sob algumas condições, é garantido que este algoritmo converge em um número finito de iterações para um mínimo local.

Além disso, apresentamos um algoritmo de treinamento alternativo para certos tipos de TE-FAMs chamado Algoritmo TE, que otimiza tanto as contribuições dos atributos dos dados ajus-

tando os parâmetros das medidas de equivalência quanto os pesos correspondentes à contribuição dos neurônios ocultos. O Algoritmo TE primeiro constrói o conjunto de memórias fundamentais \mathcal{M} como um subconjunto de um dado conjunto de treinamento \mathcal{T} antes de ajustar os parâmetros da rede. A construção de $\mathcal{M} \subseteq \mathcal{T}$ pode ser executada para qualquer E-FAM. Neste contexto, lembre que o número de neurônios ocultos é igual a $|\mathcal{M}|$ em uma Θ -FAM. Provamos que uma E-FAM construída utilizando-se o conjunto de memórias fundamentais \mathcal{M} , resultante da aplicação do Algoritmo TE, com pesos v_ξ constantes iguais a 1 na camada oculta produz erro zero no conjunto de treinamento se as medidas de equivalência empregadas nos neurônios ocultos satisfazem algumas condições fracas. Após construir $\mathcal{M} \subseteq \mathcal{T}$, mostramos como ajustar os parâmetros de uma TE-FAM baseada em medidas de equivalência da forma $E_{\mathbf{w}}$ (veja Eq. (2.2.23)), através da resolução de um problema de otimização quadrático no segundo estágio do Algoritmo TE.

A parte principal desta tese termina com aplicações de Θ -FAMs em dois conjuntos de problemas: 1) problemas de classificação de referência retirados da base de dados KEEL e 2) um problema em robótica de auto-localização baseado em visão. Para isto, selecionamos um modelo Θ -FAM entre uma família finita de Θ -FAMs treinadas com o Algoritmo 3.1 através da avaliação de seus desempenhos nos dados de treinamento ou, se existir, nos dados de validação. Como apontamos anteriormente, nos dois conjuntos de problemas, estes modelos selecionados produziram os melhores resultados em termo de acurácia em comparação aos resultados produzidos por outros classificadores bem conhecidos e outras abordagens à auto-localização baseada em visão da literatura. Adicionalmente, também aplicamos TE-FAMs baseadas em medidas de equivalência parametrizadas da forma $E_{\lambda, \mathbf{w}}$ (veja Equation (2.2.24)) em conjunto com Algoritmo TE para o primeiro conjunto de problemas que estão disponíveis na internet [2]. Os modelos TE-FAMs exibiram o segundo melhor resultado com uma ligeira melhora na taxa de classificação média em comparação a vários classificadores baseados em regras *fuzzy* competitivas, que foram aplicados recentemente a estes mesmos problemas por Alcalá-Fdez *et al.* [1]. Por fim, também verificamos que as TE-FAMs treinadas usando o Algoritmo TE exigiu um tempo de CPU relativamente menor em comparação aos modelos de Θ -FAMs treinados usando o Algoritmo 3.1.

No futuro, pretendemos generalizar a abordagem Θ -FAM usando outros tipos de função de ativação na camada oculta a fim de lidar com aplicações de regressão e agrupamento. Repare que uma modificação deste tipo exige o desenvolvimento de novos algoritmos de treinamento para Θ -FAMs, pois tanto o Algoritmo 3.1 quanto o Algoritmo TE, aqui discutidos, levam em conta apenas a função de ativação F dada pela Equação (3.0.2). Além de novos algoritmos de treinamento, novas versões dos Teoremas 3.1.1 e 3.1.2 também precisam ser reformuladas considerando outras escolhas de funções de ativação.

Referências

- [1] J. Alcalá-Fdez, R. Alcalá, and F. Herrera. “A Fuzzy Association Rule-Based Classification Model for High-Dimensional Problems With Genetic Rule Selection and Lateral Tuning”. In: *Fuzzy Systems, IEEE Transactions on* 19.5 (2011), pp. 857–872.
- [2] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, and S. García. “KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework.” In: *Multiple-Valued Logic and Soft Computing* 17.2-3 (2011), pp. 255–287.
- [3] R. A. Araújo, F. Madeiro, T. A. F. Ferreira, R. P. Sousa, and L. F. C. Pessoa. “Improved Evolutionary Hybrid Method for Designing Morphological Operators”. In: *Proceedings of the IEEE International Conference on Image Processing*. Oct. 2006, pp. 2417–2420. DOI: 10.1109/ICIP.2006.312950.
- [4] R. A. Araújo, F. Madeiro, R. P. Sousa, L. F. C. Pessoa, and T. A. E. Ferreira. “An Evolutionary Morphological Approach for Financial Time Series Forecasting”. In: *Proceedings of the IEEE Congress on Evolutionary Computation*. 2006, pp. 2467–2474. DOI: 10.1109/CEC.2006.1688615.
- [5] W. Bandler and L. Kohout. “Fuzzy power sets and fuzzy implication operators”. In: *Fuzzy Sets and Systems* 4.1 (July 1980), pp. 13–30.
- [6] G. J. F. Banon and J. Barrera. “Decomposition of Mappings between Complete Lattices by Mathematical Morphology, Part 1. General Lattices”. In: *Signal Processing* 30.3 (Feb. 1993), pp. 299–327.
- [7] L.C. Barros and R.C. Bassanezi. *Tópicos de lógica fuzzy e biomatemática*. 2rd. Campinas, SP, Brazil: G. P. Silveira, 2010.
- [8] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Wiley, 2006.
- [9] R. Belohlávek. “Fuzzy logical bidirectional associative memory”. In: *Information Sciences* 128.1-2 (2000), pp. 91–103.
- [10] G. Birkhoff. *Lattice Theory*. 3rd. Providence: American Mathematical Society, 1993.
- [11] S. Boyd and L. Vandenberghe. *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [12] J. J. Buckley and Y. Hayashi. “Fuzzy neural networks: A survey”. In: *Fuzzy Sets and Systems* 66 (1994), pp. 1–13.

- [13] H. Bustince, P. Burillo, and F. Soria. “Automorphisms, negations and implication operators”. In: *Fuzzy Sets and Systems* 134.2 (2003), pp. 209–229.
- [14] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, and R. Orduna. “Overlap functions”. In: *Nonlinear Analysis: Theory, Methods & Applications* 72 (2010), pp. 1488–1499.
- [15] H. Bustince, V. Mohedano, E. Barrenechea, and M. Pagola. “Definition and construction of fuzzy DI-subsethood measures”. In: *Information Sciences* 176.21 (2006), pp. 3190–3231.
- [16] H. Bustince, M. Pagola, and E. Barrenechea. “Construction of fuzzy indices from fuzzy DI-subsethood measures: Application to the global comparison of images”. In: *Information Sciences* 177.3 (2007), pp. 906–929.
- [17] F. Chung and T. Lee. “On Fuzzy Associative Memory with Multiple-Rule Storage Capacity”. In: *IEEE Transactions on Fuzzy Systems* 4.3 (Aug. 1996), pp. 375–384.
- [18] T. F. Coleman and Y. Li. *An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds*. Tech. rep. Ithaca, NY, USA, 1993.
- [19] T. F. Coleman and Y. Li. “On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds”. In: *Mathematical Programming* 67.1-3 (1994), pp. 189–224.
- [20] A. Di Nola, S. Sessa, W. Pedrycz, and E. Sanchez. *Fuzzy Relation Equations and Their Applications to Knowledge Engineering*. Norwell, MA: Kluwer Academic Publishers, 1989.
- [21] D. Dubois, W. Ostasiewicz, and H. Prade. “Fuzzy Sets: History and Basic Notions”. In: *Fundamentals of Fuzzy Sets*. Ed. by Didier Dubois and Henri Prade. Vol. 7. The Handbooks of Fuzzy Sets Series. Springer US, 2000, pp. 21–124.
- [22] D. Dubois and H. Prade. “A unifying view of comparison indices in a fuzzy set-theoretic framework”. In: *Recent Developments in Fuzzy Set and Possibility Theory*. Ed. by R. R. Yager. Pergamon Press, 1982, pp. 3–13.
- [23] E. Esmi and P. Sussner. *Some Experimental Results on the Use of Θ -FAMs*. University of Campinas, Dept. of Applied Math., Mathematical Imaging and Computational Intelligence Group. <http://www.milab.ime.unicamp.br/ieeeThetaFam.pdf>. 2011.
- [24] E. Esmi, P. Sussner, H. Bustince, and J. Fernández. “ Θ -Fuzzy Associative Memories”. In: *IEEE Transactions on Fuzzy Systems* (2014). accepted for publication, to appear.
- [25] E. Esmi, P. Sussner, and S. Sandri. “Tunable Equivalence Fuzzy Associative Memories”. In: (2014). submitted for publication.
- [26] E. Esmi, P. Sussner, and M. Valle. “Introdução as memórias associativas *fuzzy* - Θ ”. In: *Proceedings of the 2nd Brazilian Congress on Fuzzy Systems*. 2012, pp. 352–366.
- [27] E. Esmi, P. Sussner, M. E. Valle, F. Sakuray, and L. C. Barros. “Fuzzy Associative Memories Based on Subsethood and Similarity Measures with Applications to Speaker Identification”. In: *Hybrid Artificial Intelligent Systems*. Ed. by Emilio Corchado, Václav Snásel, Ajith Abraham, Michal Wozniak, Manuel Graña, and Sung-Bae Cho. Vol. 7209. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2012, pp. 479–490.

- [28] J. Fan and W. Xie. “Some notes on similarity measure and proximity measure”. In: *Fuzzy Sets and Systems* 101.3 (1999), pp. 403–412.
- [29] J. Fan, W. Xie, and J. Pei. “Subsethood measure: new definitions”. In: *Fuzzy Sets and Systems* 106.2 (Sept. 1999), pp. 201–209.
- [30] J. Fodor and R. R. Yager. “Fuzzy Set-Theoretic Operators and Quantifiers”. In: *Fundamentals of Fuzzy Sets*. Ed. by Didier Dubois and Henri Prade. Vol. 7. The Handbooks of Fuzzy Sets Series. Springer US, 2000, pp. 125–193.
- [31] J.C. Fodor and R. Roubens. *Fuzzy Preference Modelling and Multicriteria Decision Support*. Fundamental Theories of Physics. Springer, 1994.
- [32] P. D. Gader, M. A. Khabou, and A. Koldobsky. “Morphological Regularization Neural Networks”. In: *Pattern Recognition, Special Issue on Mathematical Morphology and Its Applications* 33.6 (June 2000), pp. 935–945.
- [33] J. Goguen. “L-fuzzy Sets”. In: *Journal of Mathematical Analysis and Applications* 18.1 (Apr. 1967), pp. 145–174.
- [34] A. Gonzalez and R. Perez. “Selection of relevant features in a fuzzy genetic learning algorithm”. In: *Trans. Sys. Man Cyber. Part B* 31.3 (June 2001), pp. 417–425.
- [35] M. Graña, D. Chyzhyk, M. García-Sebastián, and C. Hernández. “Lattice independent component analysis for functional magnetic resonance imaging”. In: *Information Sciences* 181.10 (2011), pp. 1910–1928.
- [36] M. Graña, J. Gallego, F. J. Torrealdea, and A. D’Anjou. “On the application of associative morphological memories to hyperspectral image analysis”. In: *Lecture Notes in Computer Science* 2687 (2003), pp. 567–574.
- [37] M. Graña, B. Raducanu, P. Sussner, and G. X. Ritter. “On endmember detection in hyperspectral images with Morphological Associative Memories”. In: *Lecture Notes in Artificial Intelligence - Proceedings of IBERAMIA 2002*. 2002, pp. 526–535.
- [38] M. Graña, A. M. Savio, M. García-Sebastián, and E. Fernandez. “A lattice computing approach for on-line fMRI analysis”. In: *Image and Vision Computing* 28.7 (2010), pp. 1155–1161.
- [39] M. Graña, I. Villaverde, J. O. Maldonado, and C. Hernandez. “Two lattice computing approaches for the unsupervised segmentation of hyperspectral images”. In: *Neurocomputing* 72 (2009), pp. 2111–2120.
- [40] G. A. Grätzer. *Lattice Theory: First Concepts and Distributive Lattices*. San Francisco, CA: W. H. Freeman, 1971.
- [41] M. H. Hassoun, ed. *Associative Neural Memories: Theory and Implementation*. Oxford, U. K.: Oxford University Press, 1993.
- [42] A. K. Hocaoglu and P. D. Gader. “Domain learning using Choquet integral-based morphological shared weight neural networks”. In: *Image and Vision Computing* 21.7 (2003), pp. 663–673.

- [43] P. A. D. F. R. Højen-Sørensen, O. Winther, and L. K. Hansen. “Mean-field approaches to independent component analysis”. In: *Neural Comput.* 14.4 (Apr. 2002), pp. 889–918.
- [44] N. Ikeda, P. Watta, M. Artiklar, and M. H. Hassoun. “A two-level Hamming network for high performance associative memory”. In: *Neural Networks* 14.9 (2001), pp. 1189–1200.
- [45] H. Ishibuchi, T. Yamamoto, and T. Nakashima. “Hybridization of fuzzy GBML approaches for pattern classification problems”. In: *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 35.2 (2005), pp. 359–365.
- [46] H. Izakian, W. Pedrycz, and I. Jamal. “Clustering Spatiotemporal Data: An Augmented Fuzzy C-Means”. In: *Fuzzy Systems, IEEE Transactions on* 21.5 (Oct. 2013), pp. 855–868.
- [47] F. Junbo, J. Fan, and S. Yan. “A learning rule for fuzzy associative memories”. In: *Proceedings of the IEEE International Joint Conference on Neural Networks*. Vol. 7. June 1994, pp. 4273–4277.
- [48] V. G. Kaburlasos. “FINs: Lattice theoretic tools for improving prediction of sugar production from populations of measurements”. In: *IEEE Transactions on Systems Man and Cybernetics, Part B - Cybernetics* 34.2 (2004), pp. 1017–1030.
- [49] V. G. Kaburlasos, I. N. Athanasiadis, and P. A. Mitkas. “Fuzzy lattice reasoning (FLR) classifier and its application for ambient ozone estimation”. In: *International Journal of Approximate Reasoning* 45.1 (May 2007), pp. 152–188.
- [50] V. G. Kaburlasos, L. Moussiades, and A. Vakali. “Fuzzy lattice reasoning (FLR) type neural computation for weighted graph partitioning”. In: *Neurocomput.* 72 (June 2009), pp. 2121–2133.
- [51] V. G. Kaburlasos and S. E. Papadakis. “A granular extension of the fuzzy-ARTMAP (FAM) neural classifier based on fuzzy lattice reasoning (FLR)”. In: *Neurocomputing* 72.10–12 (2009), pp. 2067–2078.
- [52] V. G. Kaburlasos and V. Petridis. “Fuzzy lattice neurocomputing (FLN) models”. In: *Neural Networks* 13.10 (2000), pp. 1145–1170.
- [53] V.G. Kaburlasos and A. Kehagias. “Fuzzy Inference System (FIS) Extensions Based on Lattice Theory”. In: *Fuzzy Systems, IEEE Transactions on* (2013). to appear, available online.
- [54] V.G. Kaburlasos, S.E. Papadakis, and G.A. Papakostas. “Lattice Computing Extension of the FAM Neural Classifier for Human Facial Expression Recognition”. In: *Neural Networks and Learning Systems, IEEE Transactions on* 24.10 (2013), pp. 1526–1538.
- [55] M. A. Khabou and P. D. Gader. “Automatic target detection using entropy optimized shared-weight neural networks”. In: *IEEE Transactions on Neural Networks* 11.1 (Jan. 2000), pp. 186–193.
- [56] M. A. Khabou, P. D. Gader, and J. M. Keller. “LADAR Target Detection Using Morphological Shared-Weight Neural Networks”. In: *Machine Vision and Applications* 11.6 (May 2000), pp. 300–305.

- [57] M. A. Khabou, P. D. Gader, and J. M. Keller. “LADAR Target Detection Using Morphological Shared-Weight Neural Networks”. In: *Machine Vision and Applications* 11.6 (May 2000), pp. 300–305.
- [58] M. A. Khabou, P. D. Gader, and H. Shi. “Entropy optimized morphological shared-weight neural networks”. In: *Optical Engineering* 38.2 (1999), pp. 263–273.
- [59] L. Kitainik. “Fuzzy inclusions and fuzzy dichotomous decision procedures”. In: *Optimization models using fuzzy sets and possibility theory*. Ed. by J. Kacprzyk and S. A. Orlovski. Dordrecht: D. Reidel, 1987.
- [60] G. J. Klir and T. A. Folger. *Fuzzy Sets, Uncertainty, and Information*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1987.
- [61] G. J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic; Theory and Applications*. Upper Saddle River, N. Y.: Prentice Hall, 1995. ISBN: 0-13-101171-5.
- [62] T. Kohonen. *Self-Organization and Associative Memory*. Springer Verlag, 1984.
- [63] B. Kosko. “Fuzziness vs. probability”. In: *Int. J. General Systems* 17 (1990), pp. 211–240.
- [64] B. Kosko. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Englewood Cliffs, N. J.: Prentice Hall, 1992.
- [65] W. Li, J. Han, and J. Pei. “CMAR: accurate and efficient classification based on multiple class-association rules”. In: *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*. 2001, pp. 369–376.
- [66] R. P. Lippmann. “An Introduction to Computing with Neural Nets”. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-4* (1987), pp. 4–22.
- [67] R. P. Lippmann and E. A. Martin. “Neural Classifiers useful for speech recognition”. In: *1st International Conference on Neural Networks IEEE* (1987), pp. 400–417.
- [68] B. Liu, W. Hsu, and Y. Ma. “Integrating Classification and Association Rule Mining”. In: *Proceedings of the International Conference on Knowledge Discovery and Data Mining (SIGKDD)*. New York, USA, 1998, pp. 80–86.
- [69] B. Liu, Y. Ma, and C.-K. Wong. “Classification Using Association Rules: Weaknesses and Enhancements”. In: *Data Mining for Scientific and Engineering Applications*. Ed. by C. Kamath R. Grossman and V. Kumar. Kluwer Academic Publishers, 2001, pp. 591–601.
- [70] P. Liu. “The Fuzzy Associative Memory of Max-Min Fuzzy Neural Networks with Threshold”. In: *Fuzzy Sets and Systems* 107 (1999), pp. 147–157.
- [71] M. H. Hassoun and P. B. Watta. “Associative Memory Networks”. In: *Handbook of Neural Computation*. Oxford, U. K.: Oxford University Press, 1997.
- [72] E.G. Mansoori, M.J. Zolghadri, and S.D. Katebi. “SGERD: A Steady-State Genetic Algorithm for Extracting Fuzzy Classification Rules From Data”. In: *Fuzzy Systems, IEEE Transactions on* 16.4 (2008), pp. 1061–1071.

- [73] F. T. Martins-Bedé, L. Godo, S. Sandri, L. V. Dutra, C. C. Freitas, O. S. Carvalho, R. J. Guimarães, and R. S. Amaral. “Classification of Schistosomiasis Prevalence Using Fuzzy Case-Based Reasoning”. In: *Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part I: Bio-Inspired Systems: Computational and Ambient Intelligence*. IWANN '09. Salamanca, Spain: Springer-Verlag, 2009, pp. 1053–1060.
- [74] G. Matheron. *Random Sets and Integral Geometry*. New York: Wiley, 1975.
- [75] L. Molgedey and H. G. Schuster. “Separation of a mixture of independent signals using time delayed correlations”. In: *Phys. Rev. Lett.* 72 (23 June 1994), pp. 3634–3637.
- [76] X. Mu, P. Watta, and M.H. Hassoun. “A Weighted Voting Model of Associative Memory”. In: *Neural Networks, IEEE Transactions on* 18.3 (2007), pp. 756–777.
- [77] T. Murofushi and M. Sugeno. “An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure”. In: *Fuzzy Sets and Systems* 29.2 (1989), pp. 201–227.
- [78] S. Murthy. “Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey”. In: *Data Mining and Knowledge Discovery* 2.4 (1998), pp. 345–389.
- [79] M. Nachtegael and E. E. Kerre. “Connections between binary, gray-scale and fuzzy mathematical morphologies”. In: *Fuzzy Sets and Systems* 124.1 (Nov. 2001), pp. 73–85.
- [80] W. Pedrycz and F. Gomide. *Fuzzy Systems Engineering: Towards Human-Centric Computing*. New York: Wiley, IEEE Press, 2007.
- [81] L. F. C. Pessoa and P. Maragos. “Neural networks with hybrid morphological/rank/linear nodes: a unifying framework with applications to handwritten character recognition”. In: *Pattern Recognition* 33 (2000), pp. 945–960.
- [82] V. Petridis and V. G. Kaburlasos. “Fuzzy lattice neural network (FLNN): a hybrid model for learning”. In: *IEEE Transactions on Neural Networks* 9.5 (Sept. 1998), pp. 877–890.
- [83] J. R. Quinlan. *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.
- [84] B. Raducanu, M. Graña, and X. F. Albizuri. “Morphological Scale Spaces and Associative Morphological Memories: Results on Robustness and Practical Applications”. In: *Journal of Mathematical Imaging and Vision* 19.2 (2003), pp. 113–131.
- [85] B. Raducanu, M. Graña, and P. Sussner. “Morphological Neural Networks for Vision Based Self-Localization”. In: *Proceedings of the International Conference on Robotics and Automation*. Seoul, South Korea, May 2001.
- [86] G. X. Ritter, G. Urcid, and M. S. Schmalz. “Autonomous single-pass endmember approximation using lattice auto-associative memories”. In: *Neurocomputing* 72.10–12 (2009), pp. 2101–2110.
- [87] C. Ronse. “Why Mathematical Morphology Needs Complete Lattices”. In: *Signal Processing* 21.2 (Oct. 1990), pp. 129–154.

- [88] S. Sandri and F.T. Martins-Bedê. “A method for deriving order compatible fuzzy relations from convex fuzzy partitions”. In: *Fuzzy Sets and Systems* (2014). In Press.
- [89] B. Schölkopf and A.J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Adaptive Computation and Machine Learning. Mit Press, 2002.
- [90] J. Serra. *Image Analysis and Mathematical Morphology*. London: Academic Press, 1982.
- [91] J. Serra. *Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances*. New York: Academic Press, 1988.
- [92] D. Sinha and E. R. Dougherty. “Fuzzification of set inclusion: theory and applications”. In: *Fuzzy Sets and Systems* 55.1 (Apr. 1993), pp. 15–42.
- [93] D. Sinha and E. R. Dougherty. “Fuzzy Mathematical Morphology”. In: *J. Vis. Comun. Image Represent.* 3.3 (Sept. 1992), pp. 286–302. ISSN: 1047-3203.
- [94] D. Sinha, P. Sinha, E. R. Dougherty, and S. Batman. “Design and analysis of fuzzy morphological algorithms for image processing”. In: *IEEE Transactions on Fuzzy Systems* 5.4 (Nov. 1997), pp. 570–583.
- [95] P. Sussner. “Generalizing Operations of Binary Morphological Autoassociative Memories using Fuzzy Set Theory”. In: *Journal of Mathematical Imaging and Vision* 9.2 (Sept. 2003). Special Issue on Morphological Neural Networks, pp. 81–93.
- [96] P. Sussner and E. Esmi. “Constructive morphological neural networks: some theoretical aspects and experimental results in classification”. In: *Constructive Neural Networks*. Ed. by J. Kacprzyk. Studies in Computational Intelligence. Heidelberg, Germany: Springer Verlag, 2009. URL: <http://www.springer.com/engineering/book/978-3-642-04511-0>.
- [97] P. Sussner and E. Esmi. “Introduction to Morphological Perceptrons with Competitive Learning”. In: *Proceedings of the International Joint Conference on Neural Networks*. Atlanta, GA, 2009, pp. 3024–3031.
- [98] P. Sussner, E. L. Esmi, I. Villaverde, and M. Graña. “The Kosko Subsethood Fuzzy Associative Memory (KS-FAM): Mathematical Background and Applications in Computer Vision”. In: *Journal of Mathematical Imaging and Vision* 42 (2 2012), pp. 134–149.
- [99] P. Sussner, R. Miyasaki, and M. E. Valle. “An Introduction to Parameterized IFAM Models with Applications in Prediction”. In: *Proceedings of IFSA-EUSFLAT 2009*. Lisbon, Portugal, 2009, pp. 3024–3031.
- [100] P. Sussner, M. Nachtegael, T. Mélangé, G. Deschrijver, E. Esmi, and E. Kerre. “Interval-Valued and Intuitionistic Fuzzy Mathematical Morphologies as Special Cases of L-Fuzzy Mathematical Morphology”. In: *Journal of Mathematical Imaging and Vision* 43 (1 2012), pp. 50–71.
- [101] P. Sussner and M. E. Valle. “Classification of Fuzzy Mathematical Morphologies Based on Concepts of Inclusion Measure and Duality”. In: *Journal of Mathematical Imaging and Vision* 32.2 (2008), pp. 139–159.

- [102] P. Sussner and M. E. Valle. “Fuzzy Associative Memories and Their Relationship to Mathematical Morphology”. In: *Handbook of Granular Computing*. Ed. by W. Pedrycz, A. Skowron, and V. Kreinovich. New York: John Wiley and Sons, Inc., 2008. Chap. 33.
- [103] P. Sussner and M. E. Valle. “Grayscale Morphological Associative Memories”. In: *IEEE Transactions on Neural Networks* 17.3 (May 2006), pp. 559–570.
- [104] P. Sussner and M. E. Valle. “Implicative Fuzzy Associative Memories”. In: *IEEE Transactions on Fuzzy Systems* 14.6 (2006), pp. 793–807.
- [105] P. Sussner and M. E. Valle. “Morphological and Certain Fuzzy Morphological Associative Memories for Classification and Prediction”. In: *Computational Intelligence Based on Lattice Theory*. Ed. by V. G. Kaburlassos and G. X. Ritter. Vol. 67. Heidelberg, Germany: Springer Verlag, 2007, pp. 149–173.
- [106] X.-Q. Tang and P. Zhu. “Hierarchical Clustering Problems and Analysis of Fuzzy Proximity Relation on Granular Space”. In: *Fuzzy Systems, IEEE Transactions on* 21.5 (Oct. 2013), pp. 814–824.
- [107] M. E. Valle and P. Sussner. “A General Framework for Fuzzy Morphological Associative Memories”. In: *Fuzzy Sets and Systems* 159.7 (2008), pp. 747–768.
- [108] I. Villaverde, B. Fernandez-Gauna, and E. Zulueta. “Lattice Independent Component Analysis for Mobile Robot Localization”. In: *Hybrid Artificial Intelligence Systems*. Ed. by E. Corchado, M. Graña, and A. Manhaes. Vol. 6077. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, pp. 335–342.
- [109] I. Villaverde, M. Graña, and A. d’Anjou. “Morphological neural networks and vision based simultaneous localization and mapping”. In: *Integrated Computer-Aided Engineering* 14 (4 Dec. 2007), pp. 355–363.
- [110] R. Willmott. “On the transitivity of containment and equivalence in fuzzy power set theory”. In: *Journal of Mathematical Analysis and Applications* 120.1 (1986), pp. 384–396.
- [111] Y. Won, P. D. Gader, and P. Coffield. “Morphological Shared-Weight Networks with Applications to Automatic Target Recognition”. In: *IEEE Transactions on Neural Networks* 8.5 (Sept. 1997), pp. 1195–1203.
- [112] Liu X. “Entropy, distance measure and similarity measure of fuzzy sets and their relations”. In: *Fuzzy Sets Syst.* 52 (3 Dec. 1992), pp. 305–318.
- [113] X. Yin and J. Han. “CPAR: Classification based on Predictive Association Rules”. In: *Proceedings of 3rd SIAM International Conference on Data Mining (SDM)*. San Francisco, CA, USA, 2003, pp. 331–335.
- [114] V. R. Young. “Fuzzy subsethood”. In: *Fuzzy Sets and Systems* 77.3 (Feb. 1996), pp. 371–384.
- [115] L. A. Zadeh. “Fuzzy Sets”. In: *Information and Control* 8.3 (1965), pp. 338–353.
- [116] W. Zeng and H. Li. “Inclusion measures, similarity measures, and the fuzziness of fuzzy sets and their relations”. In: *International Journal of Intelligent Systems* 21.6 (2006), pp. 639–653.

- [117] H.-Y. Zhang and W.-X. Zhang. “Hybrid monotonic inclusion measure and its use in measuring similarity and distance between fuzzy sets”. In: *Fuzzy Sets Syst.* 160.1 (Jan. 2009), pp. 107–118.

Apêndice A

Alguns Resultados Experimentais Usando \ominus -FAMs

Aqui, apresentaremos os resultados de classificação obtidos pelos modelos \ominus -FAMs em alguns problemas de classificação de referência [2] em um problema de auto-localização baseado visão em robótica [98, 108, 109].

Seja X um conjunto finito e seja $\mathcal{F}(X)$ a classe dos conjuntos *fuzzy* em X . Em nossas simulações, empregamos S-FAMs (duais) baseadas tanto em medidas de *subsethood* de Kosko e Willmott [63, 110] quanto em medidas de *subsethood* S_ϱ^\cap e S_ϱ^\cup dadas por

$$(a) \quad S_\varrho^\cap(A, B) = I_P(v_\varrho(A), v_\varrho(A \cap B)) \quad \forall A, B \in \mathcal{F}(X), \quad (A.0.1)$$

$$(b) \quad S_\varrho^\cup(A, B) = I_P(v_\varrho(A \cup B), v_\varrho(B)) \quad \forall A, B \in \mathcal{F}(X). \quad (A.0.2)$$

Aqui, o símbolo I_P denota a implicação fuzzy de Goguen e v_ϱ é uma função $\mathcal{F}(X) \rightarrow [0, 1]$ definida como

$$v_\varrho(C) = \sum_{i=1}^k \frac{1 - \cos(\pi[\mu_C(x_i)]^\varrho)}{k}, \quad \forall C \in \mathcal{F}(X) \quad (A.0.3)$$

para $\varrho \in P = \{0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3, 4\}$. Além disso, empregamos E-FAMs baseadas em medidas de equivalência dadas por $E^S(A, B) = \min\{S(A, B), S(B, A)\}$ para todo $A, B \in \mathcal{F}(X)$, onde S corresponde a uma das medidas de *subsethood* mencionadas acima.

Cada conjunto de dados é composto por intâncias ou vetores cujas entradas representam atributos numéricos ou categóricos. Formalmente, cada conjunto de dados considerado neste trabalho é dado por um subconjunto finito de $X_1 \times \dots \times X_n$, onde cada X_i denota ou um conjunto de valores numéricos, *i.e.*, $X_i = [t_{min}^i, t_{max}^i] \subset \mathbb{R}$, ou um conjunto de valores categóricos, *i.e.*, $X_i = \{l_1, \dots, l_c\}$ para algum $c > 0$. Cada elemento $\mathbf{x} \in X_1 \times \dots \times X_n$ foi associado com um conjunto *fuzzy* $\eta(\mathbf{x}) = (\eta_1(x_1)^t, \dots, \eta_n(x_n)^t)^t$ que é dado pela concatenação de $\eta_j(x_j)$ para $j = 1, \dots, n$. Seja $i \in \{1, \dots, n\}$. Por um lado, se X_i é da forma $[t_{min}^i, t_{max}^i] \subset \mathbb{R}$, então,

$$\eta_i(x) = \frac{(x - t_{min}^i)}{(t_{max}^i - t_{min}^i)} \in [0, 1]$$

para todo $x \in X_i$. Por outro lado, se X_i é da forma $\{l_1, \dots, l_c\}$, $c > 0$, então, $\eta_i(l_j) = B^j \in \{0, 1\}^c \subseteq [0, 1]^c$ onde

$$\mu_{B^j}(i) = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{caso contrário.} \end{cases} \quad (\text{A.0.4})$$

Note que, em problemas de classificação, o rótulo de classe é dado por um atributo categórico. Logo, identificamos cada classe com um conjunto *fuzzy* de acordo com a Equação (A.0.4). Adicionalmente, também consideramos a aplicação da função de normalização $\Psi : [0, 1]^n \rightarrow [0, 1]^n$ definida abaixo para o conjunto *fuzzy* $A \in [0, 1]^n$:

$$\Psi(A)_i = \begin{cases} A_i & \text{se } A_i = 0 \text{ ou } A_i = 1, \\ \frac{A_i - m_A + 1.5}{3} & \text{caso contrário,} \end{cases} \quad \forall i = 1, \dots, n. \quad (\text{A.0.5})$$

Aqui, m_A denota a média das pertinências de A sob a restrição do domínio $I = \{i : 0 < A_i < 1\}$, *i.e.*, $m_A = \sum_{i \in I} A_i / |I|$.

A.1 Alguns Problemas de Classificação de Referência

Aplicamos os modelos S-FAM, S-FAM dual e E-FAM em vários problemas de classificação de referência que estão disponíveis na internet [2]. Especificamente, consideramos os seguintes quinze conjuntos de dados: Iris, Monks, Appendicitis, Ecoli, Pima, Glass, Wine, Heart, Cleveland, Vowel, Crx, Wdbc, Spectfheart, Sonar e Movimentlibras. A Tabela A.1 provê algumas informações de cada conjunto de dados referente à quantidade de instâncias, atributos categóricos, atributos numéricos e classes.

Tabela A.1: Descrição dos conjuntos de dados.

| | Instâncias | Atributos Categóricos | Atributos Numéricos | Classes |
|----------------|------------|--------------------------|------------------------|---------|
| Appendicitis | 106 | 0 | 7 | 2 |
| Cleveland | 297 | 0 | 13 | 5 |
| Crx | 653 | 9 | 6 | 2 |
| Ecoli | 336 | 0 | 7 | 8 |
| Glass | 214 | 0 | 9 | 7 |
| Heart | 270 | 0 | 13 | 2 |
| Iris | 150 | 0 | 4 | 3 |
| Monks | 432 | 0 | 6 | 2 |
| Movementlibras | 360 | 0 | 90 | 15 |
| Pima | 768 | 0 | 8 | 2 |
| Sonar | 208 | 0 | 60 | 2 |
| Spectfheart | 267 | 0 | 44 | 2 |
| Vowel | 990 | 0 | 13 | 11 |
| Wdbc | 569 | 0 | 30 | 2 |
| Wine | 178 | 0 | 13 | 3 |

Para cada conjunto de dados, avaliamos o desempenho de classificação para cada Θ -FAM utilizando o método de validação cruzada com 10 *fold*s. Utilizamos o mesmo particionamento de dados utilizado em [1, 2]. As Tabelas A.2-A.16 exibem a precisão de classificação média para cada modelo Θ -FAM considerado nos quinze conjuntos de dados. Os valores em negrito indicam os melhores resultados na fase de treinamento.

Tabela A.2: Resultados produzidos pelas Θ -FAMs para Appendicitis.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|-------------------|-----------------|-------|------------|-------|--------------|--------------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 94.55 | 82.09 | 95.60 | 78.36 | 97.48 | 79.27 | 96.23 | 67.82 | 95.90 | 59.45 | 96.75 | 67.00 |
| S_W | 91.19 | 77.45 | 94.55 | 81.00 | 98.64 | 81.27 | 97.38 | 69.55 | 95.92 | 66.91 | 96.64 | 68.82 |
| $S_{0.25}^{\cap}$ | 94.56 | 81.09 | 95.08 | 78.27 | 98.74 | 82.27 | 96.43 | 71.91 | 96.01 | 61.55 | 6.12 | 69.73 |
| $S_{0.25}^{\cup}$ | 91.92 | 75.64 | 95.29 | 80.00 | 98.43 | 83.18 | 97.07 | 68.82 | 97.38 | 57.64 | 97.28 | 67.91 |
| $S_{0.5}^{\cap}$ | 93.50 | 78.18 | 93.81 | 76.18 | 97.06 | 77.55 | 95.38 | 68.73 | 97.27 | 67.82 | 96.75 | 68.73 |
| $S_{0.5}^{\cup}$ | 91.72 | 70.82 | 94.33 | 83.91 | 98.12 | 81.18 | 96.33 | 71.45 | 97.27 | 64.36 | 97.48 | 66.82 |
| $S_{0.75}^{\cap}$ | 94.23 | 79.27 | 95.60 | 81.91 | 97.69 | 79.27 | 96.64 | 64.91 | 96.02 | 66.82 | 97.06 | 73.27 |
| $S_{0.75}^{\cup}$ | 87.43 | 72.73 | 93.30 | 75.91 | 98.74 | <i>81.18</i> | 94.65 | 69.91 | 94.65 | 65.09 | 95.90 | 59.55 |
| $S_{1.0}^{\cap}$ | 92.24 | 79.64 | 93.50 | 81.00 | 97.80 | 82.27 | 95.50 | 67.82 | 94.66 | 60.27 | 97.28 | 70.73 |
| $S_{1.0}^{\cup}$ | 92.88 | 82.27 | 93.92 | 82.00 | 98.54 | 81.27 | 95.91 | 69.00 | 95.29 | 57.64 | 97.70 | 64.18 |
| $S_{1.5}^{\cap}$ | 90.68 | 83.18 | 93.81 | 78.36 | 98.53 | 75.45 | 95.90 | 68.82 | 98.22 | 68.91 | 97.80 | 66.00 |
| $S_{1.5}^{\cup}$ | 87.32 | 75.45 | 91.72 | 76.27 | 98.74 | 76.55 | 96.12 | 66.09 | 94.65 | 55.73 | 96.96 | 66.00 |
| $S_{2.0}^{\cap}$ | 91.93 | 79.09 | 93.92 | 79.00 | 98.43 | 70.64 | 96.12 | 67.82 | 95.91 | 68.82 | 96.75 | 67.09 |
| $S_{2.0}^{\cup}$ | 90.05 | 76.36 | 92.66 | 74.55 | 98.43 | 76.45 | 96.22 | 63.27 | 93.72 | 60.27 | 96.54 | 63.27 |
| $S_{2.5}^{\cap}$ | 91.30 | 81.18 | 94.45 | 75.18 | 97.90 | 74.55 | 95.60 | 61.55 | 97.59 | 59.82 | 97.17 | 69.00 |
| $S_{2.5}^{\cup}$ | 90.37 | 82.18 | 92.24 | 66.00 | 97.90 | 74.64 | 96.43 | 67.00 | 95.17 | 63.55 | 96.12 | 63.00 |
| $S_{3.0}^{\cap}$ | 92.03 | 78.55 | 95.07 | 79.45 | 97.59 | 72.64 | 96.75 | 60.09 | 96.44 | 66.36 | 97.60 | 63.09 |
| $S_{3.0}^{\cup}$ | 89.21 | 81.18 | 90.66 | 67.00 | 98.11 | 78.27 | 95.07 | 67.73 | 93.93 | 70.82 | 96.13 | 66.09 |
| $S_{4.0}^{\cap}$ | 94.13 | 79.18 | 95.39 | 82.00 | 97.28 | 76.64 | 93.91 | 67.00 | 97.90 | 65.18 | 96.12 | 61.45 |
| $S_{4.0}^{\cup}$ | 92.76 | 77.36 | 91.40 | 67.18 | 97.80 | 77.36 | 95.18 | 62.45 | 91.83 | 64.18 | 96.75 | 65.91 |

Tabela A.3: Resultados produzios pelas Θ -FAMs para Cleveland.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------------|--------------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 85.33 | 52.86 | 87.84 | 48.90 | 92.85 | 49.79 | 84.74 | 52.45 | 85.41 | 46.22 | 91.02 | 55.30 |
| S_W | 84.10 | 49.82 | 89.04 | 46.92 | 93.90 | 51.58 | 84.96 | 53.61 | 84.67 | 47.56 | 90.80 | 50.21 |
| $S_{0.25}^n$ | 85.60 | 53.51 | 84.55 | 45.92 | 91.17 | 51.24 | 87.32 | 55.31 | 85.37 | 47.25 | 88.63 | 48.79 |
| $S_{0.25}^u$ | 84.22 | 51.88 | 84.85 | 46.53 | 92.59 | 51.91 | 85.30 | 52.54 | 84.85 | 48.58 | 91.25 | 49.32 |
| $S_{0.5}^n$ | 85.67 | 53.55 | 87.73 | 44.84 | 93.53 | 52.20 | 87.35 | 54.20 | 85.11 | 49.62 | 90.68 | 50.13 |
| $S_{0.5}^u$ | 85.45 | 50.47 | 87.51 | 49.52 | 94.24 | 52.56 | 84.10 | 56.66 | 84.85 | 45.64 | 89.71 | 49.86 |
| $S_{0.75}^n$ | 87.36 | 51.88 | 89.15 | 50.58 | 93.64 | 47.84 | 87.46 | 57.62 | 85.45 | 49.55 | 91.32 | 45.87 |
| $S_{0.75}^u$ | 85.41 | 53.66 | 89.34 | 46.49 | 94.57 | 49.83 | 84.48 | 51.88 | 85.79 | 45.23 | 91.92 | 48.25 |
| $S_{1.0}^n$ | 85.97 | 59.30 | 88.81 | 47.25 | 93.19 | 51.09 | 87.43 | 54.28 | 86.34 | 48.89 | 91.77 | 48.49 |
| $S_{1.0}^u$ | 85.78 | 55.59 | 89.75 | 49.61 | 94.58 | <i>51.17</i> | 86.50 | 56.59 | 84.89 | 45.19 | 92.33 | 49.20 |
| $S_{1.5}^n$ | 85.98 | 51.88 | 88.63 | 46.27 | 92.93 | 49.47 | 86.53 | 51.92 | 86.53 | 49.54 | 90.87 | 51.52 |
| $S_{1.5}^u$ | 83.99 | 55.91 | 90.09 | 51.30 | 93.01 | 48.46 | 86.05 | 52.85 | 86.83 | 47.52 | 91.36 | 52.89 |
| $S_{2.0}^n$ | 83.70 | 47.82 | 87.77 | 49.21 | 91.51 | 51.94 | 85.78 | 54.64 | 86.90 | 49.57 | 90.42 | 50.61 |
| $S_{2.0}^u$ | 83.95 | 52.16 | 89.75 | 50.39 | 91.47 | 47.87 | 83.58 | 52.94 | 87.84 | 48.23 | 90.83 | 50.91 |
| $S_{2.5}^n$ | 84.70 | 51.48 | 88.03 | 45.88 | 90.42 | 53.50 | 84.85 | 50.88 | 86.49 | 49.92 | 90.38 | 50.90 |
| $S_{2.5}^u$ | 82.46 | 48.82 | 88.48 | 47.14 | 91.28 | 52.55 | 85.33 | 52.23 | 88.62 | 47.49 | 91.06 | 50.78 |
| $S_{3.0}^n$ | 83.06 | 54.21 | 87.99 | 48.88 | 89.90 | 51.52 | 86.46 | 52.52 | 87.28 | 47.21 | 90.54 | 48.52 |
| $S_{3.0}^u$ | 82.98 | 50.96 | 88.10 | 48.53 | 90.94 | 51.17 | 84.44 | 49.94 | 87.02 | 46.90 | 90.65 | 49.88 |
| $S_{4.0}^n$ | 83.95 | 52.92 | 86.65 | 53.06 | 89.15 | 51.21 | 83.95 | 51.21 | 87.92 | 48.95 | 89.82 | 51.21 |
| $S_{4.0}^u$ | 82.52 | 47.52 | 86.98 | 51.33 | 90.20 | 49.14 | 84.73 | 55.22 | 86.34 | 49.55 | 90.61 | 49.53 |

Tabela A.4: Resultados produzios pelas Θ -FAMs para Crx.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------|-------|-----------------------|-------|--------------|--------------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 98.11 | 82.16 | 98.16 | 80.82 | 98.42 | 81.90 | 97.94 | 79.44 | 98.06 | 78.07 | 97.86 | 77.77 |
| S_W | 98.21 | 79.93 | 98.47 | 81.77 | 98.71 | 81.77 | 98.18 | 77.96 | 97.98 | 78.25 | 97.96 | 78.56 |
| $S_{0.25}^n$ | 98.09 | 79.93 | 98.08 | 78.58 | 98.20 | 79.90 | 98.20 | 78.71 | 98.21 | 79.42 | 98.33 | 78.69 |
| $S_{0.25}^u$ | 98.13 | 81.47 | 98.16 | 79.61 | 98.26 | 78.99 | 98.33 | 79.49 | 98.71 | <i>82.02</i> | 98.43 | 78.56 |
| $S_{0.5}^n$ | 98.21 | 81.43 | 98.06 | 80.08 | 98.33 | 80.68 | 98.21 | 79.65 | 98.06 | 79.46 | 98.18 | 79.16 |
| $S_{0.5}^u$ | 98.06 | 81.24 | 98.23 | 80.28 | 98.20 | 79.92 | 98.52 | 80.07 | 98.52 | 79.31 | 98.32 | 80.55 |
| $S_{0.75}^n$ | 98.08 | 80.83 | 98.30 | 80.50 | 98.37 | 79.61 | 98.09 | 79.46 | 98.45 | 79.13 | 98.23 | 78.67 |
| $S_{0.75}^u$ | 98.16 | 79.89 | 98.38 | 81.56 | 98.09 | 80.03 | 98.62 | 79.78 | 98.25 | 79.74 | 98.44 | 78.42 |
| $S_{1.0}^n$ | 97.81 | 80.54 | 98.21 | 80.57 | 98.06 | 79.75 | 98.26 | 78.55 | 98.25 | 79.32 | 97.99 | 79.18 |
| $S_{1.0}^u$ | 98.26 | 79.60 | 98.35 | 81.90 | 98.54 | 80.11 | 98.47 | 80.39 | 98.23 | 80.98 | 98.03 | 79.92 |
| $S_{1.5}^n$ | 98.06 | 80.24 | 98.13 | 79.79 | 98.54 | 79.80 | 98.28 | 80.23 | 98.38 | 81.90 | 98.40 | 79.30 |
| $S_{1.5}^u$ | 98.03 | 81.40 | 98.04 | 80.37 | 98.33 | 80.22 | 98.09 | 80.28 | 98.33 | 79.61 | 98.08 | 79.02 |
| $S_{2.0}^n$ | 97.84 | 80.56 | 97.82 | 81.66 | 98.49 | 80.71 | 97.74 | 80.11 | 97.99 | 80.35 | 98.06 | 80.10 |
| $S_{2.0}^u$ | 97.94 | 78.85 | 98.14 | 82.41 | 98.35 | 80.87 | 98.11 | 79.18 | 98.32 | 80.38 | 98.25 | 77.04 |
| $S_{2.5}^n$ | 97.70 | 80.04 | 97.48 | 80.09 | 98.23 | 81.72 | 98.06 | 81.32 | 98.16 | 78.20 | 98.54 | 80.25 |
| $S_{2.5}^u$ | 98.14 | 81.12 | 97.89 | 80.40 | 98.33 | 81.73 | 97.94 | 79.62 | 98.50 | 81.10 | 98.47 | 79.75 |
| $S_{3.0}^n$ | 97.94 | 81.74 | 97.62 | 79.92 | 98.01 | 80.95 | 98.30 | 80.28 | 98.45 | 79.94 | 98.37 | 80.24 |
| $S_{3.0}^u$ | 98.25 | 81.56 | 98.18 | 80.85 | 98.38 | 80.23 | 98.40 | 79.89 | 98.55 | 80.69 | 98.57 | 79.89 |
| $S_{4.0}^n$ | 97.65 | 81.91 | 97.11 | 80.88 | 97.79 | 81.33 | 97.92 | 80.24 | 97.92 | 80.07 | 98.18 | 78.57 |
| $S_{4.0}^u$ | 97.96 | 80.68 | 97.82 | 79.63 | 98.28 | 82.22 | 98.13 | 78.09 | 98.37 | 80.99 | 98.30 | 80.19 |

Tabela A.5: Resultados produzios pelas Θ -FAMs para Ecoli.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------------|-------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 91.07 | 71.75 | 89.85 | 71.77 | 95.90 | 76.48 | 93.92 | 67.01 | 92.16 | 68.18 | 93.09 | 67.59 |
| S_W | 92.23 | 71.76 | 89.95 | 69.39 | 96.43 | 76.78 | 93.82 | 67.59 | 91.30 | 67.89 | 93.02 | 67.60 |
| $S_{0.25}^n$ | 91.96 | 67.60 | 88.16 | 71.52 | 93.98 | 69.66 | 93.82 | 64.88 | 90.77 | 64.96 | 93.55 | 67.56 |
| $S_{0.25}^u$ | 92.86 | 69.03 | 89.32 | 73.25 | 94.41 | 67.61 | 93.88 | 66.35 | 90.71 | 65.46 | 93.78 | 69.03 |
| $S_{0.5}^n$ | 91.47 | 67.87 | 90.24 | 70.04 | 95.04 | 74.06 | 94.15 | 68.19 | 91.70 | 66.41 | 93.22 | 69.92 |
| $S_{0.5}^u$ | 91.37 | 72.04 | 88.23 | 70.54 | 94.98 | 70.26 | 94.21 | 65.26 | 91.57 | 62.52 | 93.85 | 68.22 |
| $S_{0.75}^n$ | 91.11 | 71.45 | 90.11 | 71.19 | 95.77 | 74.44 | 93.88 | 63.99 | 91.24 | 63.10 | 93.49 | 65.82 |
| $S_{0.75}^u$ | 91.93 | 73.56 | 89.68 | 68.79 | 95.37 | 74.17 | 93.85 | 65.50 | 91.07 | 64.60 | 94.08 | 71.11 |
| $S_{1.0}^n$ | 92.13 | 73.89 | 90.77 | 70.59 | 95.44 | 76.52 | 94.18 | 64.59 | 92.13 | 64.60 | 93.65 | 64.86 |
| $S_{1.0}^u$ | 92.06 | 72.01 | 89.48 | 69.42 | 96.30 | 76.19 | 94.11 | 64.00 | 91.01 | 64.62 | 94.51 | 68.46 |
| $S_{1.5}^n$ | 91.34 | 73.27 | 89.48 | 73.86 | 96.13 | 75.30 | 93.91 | 66.71 | 92.39 | 67.90 | 94.31 | 67.31 |
| $S_{1.5}^u$ | 89.95 | 72.64 | 90.01 | 74.75 | 95.47 | 76.80 | 93.85 | 66.10 | 90.61 | 66.69 | 93.82 | 69.69 |
| $S_{2.0}^n$ | 88.69 | 71.12 | 90.81 | 74.19 | 95.14 | 75.91 | 92.69 | 64.87 | 90.41 | 66.08 | 93.29 | 67.87 |
| $S_{2.0}^u$ | 89.05 | 69.70 | 90.28 | 75.00 | 95.87 | 74.76 | 94.18 | 66.96 | 91.20 | 65.20 | 93.85 | 70.60 |
| $S_{2.5}^n$ | 88.16 | 72.04 | 91.50 | 78.32 | 95.17 | 77.09 | 92.20 | 68.20 | 90.61 | 63.10 | 92.79 | 68.46 |
| $S_{2.5}^u$ | 87.83 | 70.29 | 92.00 | 75.06 | 95.50 | 78.57 | 94.12 | 64.90 | 89.65 | 64.26 | 93.39 | 68.14 |
| $S_{3.0}^n$ | 85.95 | 72.04 | 91.70 | 78.58 | 94.94 | 78.30 | 92.50 | 61.61 | 90.35 | 61.60 | 93.29 | 63.10 |
| $S_{3.0}^u$ | 88.89 | 72.91 | 91.27 | 76.52 | 95.17 | 77.67 | 93.19 | 67.30 | 89.95 | 66.10 | 93.55 | 65.47 |
| $S_{4.0}^n$ | 87.57 | 70.54 | 92.00 | 75.05 | 95.44 | 78.26 | 92.46 | 63.09 | 91.27 | 65.45 | 93.42 | 66.36 |
| $S_{4.0}^u$ | 87.60 | 69.68 | 90.84 | 75.63 | 94.78 | 76.52 | 93.42 | 67.90 | 89.62 | 61.94 | 93.52 | 68.48 |

Tabela A.6: Resultados produzios pelas Θ -FAMs para Glass.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------------|-------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 95.69 | 72.07 | 95.12 | 70.65 | 95.32 | 69.75 | 96.26 | 73.71 | 88.70 | 61.92 | 92.83 | 69.91 |
| S_W | 96.67 | 71.04 | 94.96 | 72.01 | 96.21 | 71.19 | 96.05 | 69.64 | 83.69 | 54.58 | 93.51 | 64.35 |
| $S_{0.25}^n$ | 95.28 | 70.70 | 91.64 | 64.54 | 93.92 | 66.08 | 95.38 | 74.68 | 87.55 | 63.27 | 92.79 | 67.39 |
| $S_{0.25}^u$ | 96.16 | 67.34 | 91.17 | 67.21 | 93.46 | 69.92 | 95.54 | 71.39 | 84.07 | 59.73 | 92.42 | 67.33 |
| $S_{0.5}^n$ | 95.28 | 72.62 | 93.77 | 65.92 | 94.76 | 67.76 | 95.43 | 73.14 | 88.90 | 60.81 | 93.56 | 66.41 |
| $S_{0.5}^u$ | 95.49 | 71.06 | 93.30 | 69.46 | 94.97 | 67.05 | 96.06 | 70.76 | 84.11 | 60.37 | 92.11 | 66.27 |
| $S_{0.75}^n$ | 96.21 | 70.56 | 95.74 | 69.27 | 96.16 | 70.75 | 94.91 | 72.10 | 86.60 | 57.41 | 93.62 | 63.90 |
| $S_{0.75}^u$ | 95.69 | 70.72 | 95.43 | 70.53 | 96.52 | 70.58 | 96.52 | 70.15 | 84.28 | 59.96 | 92.84 | 65.58 |
| $S_{1.0}^n$ | 94.54 | 68.83 | 95.53 | 70.37 | 96.21 | 69.25 | 96.11 | 74.07 | 86.35 | 58.75 | 92.63 | 70.47 |
| $S_{1.0}^u$ | 95.74 | 71.28 | 95.48 | 71.79 | 97.15 | 70.49 | 96.11 | 71.18 | 84.48 | 56.42 | 93.97 | 70.25 |
| $S_{1.5}^n$ | 94.77 | 65.82 | 94.55 | 68.19 | 95.59 | 67.30 | 95.29 | 68.21 | 87.08 | 65.66 | 94.08 | 71.81 |
| $S_{1.5}^u$ | 94.76 | 64.15 | 94.60 | 70.98 | 95.54 | 69.10 | 96.26 | 69.58 | 84.63 | 59.08 | 93.61 | 67.23 |
| $S_{2.0}^n$ | 93.41 | 69.42 | 92.26 | 68.10 | 93.77 | 63.90 | 96.01 | 72.23 | 88.63 | 61.48 | 94.80 | 68.44 |
| $S_{2.0}^u$ | 95.89 | 65.89 | 90.91 | 63.70 | 93.71 | 62.44 | 97.10 | 67.32 | 85.77 | 61.45 | 95.96 | 68.72 |
| $S_{2.5}^n$ | 93.82 | 65.23 | 91.33 | 61.98 | 92.63 | 60.00 | 95.43 | 67.79 | 89.73 | 60.18 | 94.76 | 68.49 |
| $S_{2.5}^u$ | 94.34 | 66.16 | 87.07 | 60.05 | 92.16 | 60.77 | 96.63 | 67.12 | 85.94 | 58.06 | 94.86 | 66.77 |
| $S_{3.0}^n$ | 94.76 | 62.15 | 89.88 | 58.84 | 93.16 | 55.15 | 95.43 | 70.57 | 88.49 | 58.04 | 95.58 | 68.13 |
| $S_{3.0}^u$ | 94.92 | 62.56 | 86.50 | 57.57 | 90.86 | 56.60 | 96.83 | 65.68 | 82.62 | 56.48 | 93.48 | 64.12 |
| $S_{4.0}^n$ | 92.78 | 59.84 | 88.42 | 53.13 | 89.66 | 49.34 | 95.01 | 63.80 | 88.53 | 61.36 | 95.43 | 67.18 |
| $S_{4.0}^u$ | 93.98 | 58.03 | 83.91 | 49.63 | 88.95 | 53.84 | 96.88 | 67.03 | 80.44 | 47.73 | 93.71 | 61.74 |

Tabela A.7: Resultados produzios pelas Θ -FAMs para Heart.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|--------------|--------------|--------|-------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 96.71 | 71.11 | 98.52 | 76.67 | 97.90 | 78.52 | 96.54 | 69.26 | 98.23 | 76.67 | 98.35 | 76.30 |
| S_W | 96.83 | 74.44 | 98.93 | <i>78.15</i> | 98.48 | 77.04 | 96.63 | 70.37 | 97.78 | 72.96 | 97.94 | 77.41 |
| $S_{0.25}^n$ | 97.57 | 74.44 | 97.94 | 70.37 | 98.11 | 80.37 | 96.21 | 69.63 | 97.94 | 74.44 | 97.78 | 75.93 |
| $S_{0.25}^u$ | 97.16 | 70.37 | 97.61 | 74.44 | 97.86 | 78.52 | 96.01 | 69.26 | 97.49 | 71.85 | 97.78 | 74.81 |
| $S_{0.5}^n$ | 96.79 | 71.48 | 98.31 | 76.30 | 98.31 | 78.89 | 95.93 | 70.00 | 98.11 | 74.44 | 97.78 | 75.56 |
| $S_{0.5}^u$ | 98.07 | 73.33 | 98.07 | 73.70 | 98.23 | 81.11 | 97.04 | 70.74 | 97.94 | 74.07 | 97.70 | 74.81 |
| $S_{0.75}^n$ | 96.95 | 72.22 | 98.48 | 75.56 | 98.15 | 79.63 | 95.80 | 69.26 | 97.94 | 75.19 | 97.61 | 77.04 |
| $S_{0.75}^u$ | 97.70 | 74.44 | 98.72 | 75.19 | 98.35 | 82.59 | 96.58 | 68.89 | 97.24 | 75.19 | 98.11 | 76.67 |
| $S_{1.0}^n$ | 97.33 | 71.48 | 98.89 | 75.19 | 98.40 | 79.26 | 96.05 | 71.11 | 98.35 | 75.56 | 98.23 | 75.93 |
| $S_{1.0}^u$ | 97.74 | 72.59 | 98.40 | 77.04 | 98.27 | 74.81 | 97.04 | 71.85 | 97.74 | 70.74 | 97.98 | 77.41 |
| $S_{1.5}^n$ | 97.04 | 72.22 | 98.31 | 76.30 | 97.12 | 73.70 | 95.64 | 70.00 | 97.90 | 76.67 | 97.86 | 75.93 |
| $S_{1.5}^u$ | 97.57 | 69.63 | 98.52 | 74.44 | 98.19 | 76.67 | 96.67 | 74.07 | 97.98 | 67.41 | 98.02 | 78.15 |
| $S_{2.0}^n$ | 97.33 | 71.48 | 98.40 | 71.85 | 98.19 | 75.56 | 96.30 | 69.26 | 97.94 | 74.07 | 98.19 | 76.67 |
| $S_{2.0}^u$ | 97.57 | 71.11 | 98.48 | 71.85 | 98.60 | 77.41 | 96.67 | 69.26 | 97.94 | 72.96 | 98.02 | 76.30 |
| $S_{2.5}^n$ | 97.57 | 71.85 | 97.65 | 74.07 | 97.98 | 74.07 | 95.84 | 71.11 | 98.02 | 73.33 | 98.31 | 76.30 |
| $S_{2.5}^u$ | 97.08 | 72.59 | 98.52 | 72.59 | 97.90 | 75.93 | 95.56 | 68.52 | 97.82 | 72.96 | 97.74 | 75.19 |
| $S_{3.0}^n$ | 98.27 | 72.96 | 97.53 | 74.44 | 97.41 | 75.56 | 95.60 | 67.04 | 97.94 | 70.74 | 98.19 | 75.56 |
| $S_{3.0}^u$ | 97.61 | 75.93 | 98.07 | 75.93 | 96.71 | 73.70 | 96.54 | 67.04 | 97.78 | 70.74 | 97.74 | 75.56 |
| $S_{4.0}^n$ | 97.24 | 74.44 | 97.94 | 75.19 | 97.37 | 76.30 | 95.10 | 70.37 | 97.41 | 71.11 | 97.86 | 74.44 |
| $S_{4.0}^u$ | 97.78 | 68.52 | 97.86 | 74.07 | 97.70 | 72.59 | 95.93 | 66.67 | 97.61 | 71.85 | 98.15 | 72.59 |

Tabela A.8: Resultados produzios pelas Θ -FAMs para Iris.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|---------------|--------------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 98.74 | 92.67 | 99.26 | 96.67 | 99.78 | 93.33 | 99.19 | 85.33 | 98.37 | 83.33 | 99.48 | 86.67 |
| S_W | 98.52 | 95.33 | 99.11 | 95.33 | 99.48 | 94.67 | 98.74 | 85.33 | 98.89 | 84.67 | 99.04 | 88.00 |
| $S_{0.25}^n$ | 98.30 | 94.00 | 98.96 | 92.67 | 99.41 | 91.33 | 98.59 | 83.33 | 98.07 | 81.33 | 99.04 | 86.67 |
| $S_{0.25}^u$ | 98.52 | 92.00 | 98.67 | 92.00 | 99.33 | 92.67 | 98.96 | 83.33 | 98.74 | 80.67 | 99.04 | 85.33 |
| $S_{0.5}^n$ | 98.52 | 92.00 | 99.04 | 94.67 | 99.85 | 92.00 | 98.44 | 82.67 | 98.44 | 80.00 | 99.04 | 87.33 |
| $S_{0.5}^u$ | 98.67 | 92.67 | 98.52 | 94.67 | 99.63 | 92.00 | 98.67 | 81.33 | 98.74 | 83.33 | 99.48 | 87.33 |
| $S_{0.75}^n$ | 98.81 | 91.33 | 99.04 | 93.33 | 99.70 | 94.67 | 98.96 | 84.00 | 98.89 | 84.67 | 99.48 | 87.33 |
| $S_{0.75}^u$ | 99.11 | 92.00 | 99.19 | 94.00 | 99.33 | 92.67 | 98.44 | 83.33 | 98.52 | 83.33 | 98.74 | 86.67 |
| $S_{1.0}^n$ | 98.81 | 93.33 | 99.19 | 94.67 | 99.78 | 94.00 | 99.11 | 84.67 | 98.74 | 85.33 | 98.96 | 86.00 |
| $S_{1.0}^u$ | 98.37 | 96.00 | 99.41 | 94.67 | 99.70 | 94.00 | 99.04 | 85.33 | 98.59 | 86.00 | 99.63 | 85.33 |
| $S_{1.5}^n$ | 98.96 | 92.67 | 98.89 | 95.33 | 99.85 | 95.33 | 99.48 | 84.67 | 98.89 | 84.67 | 99.63 | 87.33 |
| $S_{1.5}^u$ | 98.81 | 93.33 | 99.41 | 96.67 | 99.85 | 95.33 | 98.52 | 81.33 | 98.96 | 87.33 | 99.41 | 88.00 |
| $S_{2.0}^n$ | 99.04 | 94.00 | 99.48 | 96.00 | 100.00 | <i>96.00</i> | 98.96 | 86.00 | 98.96 | 85.33 | 99.26 | 88.67 |
| $S_{2.0}^u$ | 98.89 | 94.00 | 99.41 | 94.00 | 99.85 | 95.33 | 99.41 | 88.00 | 99.19 | 90.00 | 99.11 | 88.67 |
| $S_{2.5}^n$ | 98.74 | 94.00 | 99.41 | 95.33 | 99.56 | 96.00 | 98.74 | 82.67 | 98.89 | 86.00 | 98.74 | 88.67 |
| $S_{2.5}^u$ | 98.81 | 95.33 | 99.33 | 96.67 | 99.63 | 94.67 | 99.41 | 85.33 | 99.19 | 87.33 | 99.19 | 86.67 |
| $S_{3.0}^n$ | 99.33 | 94.00 | 99.48 | 96.00 | 99.70 | 95.33 | 98.74 | 83.33 | 98.96 | 86.67 | 98.67 | 89.33 |
| $S_{3.0}^u$ | 98.74 | 92.00 | 99.63 | 95.33 | 99.48 | 96.00 | 99.48 | 82.67 | 99.26 | 84.67 | 99.41 | 88.67 |
| $S_{4.0}^n$ | 99.19 | 96.00 | 99.63 | 96.00 | 99.04 | 96.00 | 98.96 | 85.33 | 99.19 | 88.00 | 99.33 | 88.67 |
| $S_{4.0}^u$ | 98.96 | 94.00 | 99.26 | 94.00 | 99.85 | 96.00 | 99.04 | 83.33 | 99.41 | 89.33 | 99.41 | 88.67 |

Tabela A.9: Resultados produzios pelas Θ -FAMs para Monk.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|--------------|------------|-------|--------|-------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 99.30 | 95.40 | 98.79 | 97.71 | 98.33 | 65.97 | 98.10 | 94.44 | 97.58 | 95.14 | 97.74 | 83.17 |
| S_W | 98.94 | 96.54 | 98.18 | 97.49 | 98.89 | 62.41 | 98.17 | 94.94 | 97.69 | 94.60 | 97.63 | 67.00 |
| $S_{0.25}^n$ | 99.02 | 95.61 | 97.74 | 96.58 | 95.83 | 52.21 | 98.12 | 94.00 | 97.50 | 96.58 | 97.30 | 75.61 |
| $S_{0.25}^U$ | 99.05 | 95.41 | 97.33 | 95.80 | 95.37 | 42.91 | 98.04 | 95.37 | 97.56 | 95.06 | 96.66 | 68.74 |
| $S_{0.5}^n$ | 99.38 | 96.36 | 97.58 | 96.12 | 95.94 | 50.09 | 98.12 | 93.12 | 97.40 | 95.14 | 97.12 | 73.40 |
| $S_{0.5}^U$ | 98.90 | 95.87 | 97.28 | 95.83 | 95.11 | 44.62 | 98.20 | 93.76 | 97.68 | 95.14 | 96.81 | 67.33 |
| $S_{0.75}^n$ | 99.18 | 94.42 | 98.07 | 95.61 | 95.83 | 52.31 | 98.07 | 93.53 | 97.63 | 95.40 | 97.04 | 74.90 |
| $S_{0.75}^U$ | 98.54 | 95.30 | 97.69 | 96.13 | 94.96 | 44.72 | 97.79 | 95.16 | 97.53 | 94.60 | 96.81 | 68.74 |
| $S_{1.0}^n$ | 98.84 | 94.72 | 97.51 | 96.29 | 97.56 | 64.73 | 98.10 | 93.99 | 97.63 | 94.65 | 97.81 | 82.45 |
| $S_{1.0}^U$ | 99.36 | 97.72 | 97.22 | 95.37 | 97.76 | 59.78 | 98.20 | 94.92 | 97.74 | 95.83 | 97.48 | 67.84 |
| $S_{1.5}^n$ | 99.82 | 98.10 | 98.87 | 97.91 | 99.56 | 93.84 | 98.22 | 95.37 | 97.92 | 94.81 | 97.84 | 93.05 |
| $S_{1.5}^U$ | 99.87 | 98.18 | 98.28 | 97.27 | 99.56 | 94.04 | 98.23 | 95.18 | 97.99 | 95.56 | 98.10 | 92.17 |
| $S_{2.0}^n$ | 99.82 | 98.86 | 99.25 | 97.50 | 99.64 | 94.96 | 98.07 | 96.54 | 98.02 | 95.14 | 97.61 | 92.37 |
| $S_{2.0}^U$ | 99.92 | 98.63 | 97.58 | 96.81 | 99.31 | 94.99 | 98.25 | 95.59 | 97.87 | 95.82 | 97.97 | 93.32 |
| $S_{2.5}^n$ | 99.92 | 98.41 | 98.89 | 98.14 | 99.64 | 93.12 | 98.28 | 95.82 | 97.53 | 96.28 | 97.69 | 91.37 |
| $S_{2.5}^U$ | 99.72 | 99.32 | 97.58 | 97.27 | 99.59 | 93.82 | 98.30 | 95.13 | 97.89 | 96.28 | 98.02 | 92.14 |
| $S_{3.0}^n$ | 99.79 | 99.06 | 99.33 | 98.10 | 99.31 | 94.23 | 98.23 | 96.29 | 97.89 | 95.29 | 97.87 | 91.60 |
| $S_{3.0}^U$ | 99.54 | 97.72 | 98.40 | 96.97 | 99.36 | 94.27 | 98.17 | 96.06 | 97.92 | 94.88 | 98.04 | 92.17 |
| $S_{4.0}^n$ | 99.79 | 97.46 | 99.25 | 98.64 | 99.20 | 94.95 | 98.17 | 96.50 | 97.94 | 96.25 | 97.87 | 92.52 |
| $S_{4.0}^U$ | 99.79 | 98.63 | 98.95 | 98.40 | 99.25 | 93.35 | 98.25 | 95.14 | 97.84 | 95.60 | 98.05 | 93.77 |

Tabela A.10: Resultados produzios pelas Θ -FAMs para Movementlibras .

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------|-------|-----------------------|-------|--------------|--------------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 94.91 | 69.44 | 93.89 | 69.44 | 97.84 | 76.11 | 98.61 | 81.94 | 98.30 | 80.00 | 98.46 | 80.28 |
| S_W | 95.71 | 65.56 | 93.43 | 69.17 | 98.15 | 75.28 | 98.18 | 79.17 | 98.52 | 80.83 | 98.06 | 80.83 |
| $S_{0.25}^n$ | 95.83 | 67.50 | 85.28 | 54.72 | 95.59 | 68.89 | 98.89 | 81.94 | 98.21 | 78.61 | 98.52 | 80.56 |
| $S_{0.25}^U$ | 94.72 | 67.50 | 85.99 | 55.56 | 95.93 | 67.50 | 98.49 | 79.44 | 97.87 | 78.61 | 98.49 | 79.72 |
| $S_{0.5}^n$ | 95.77 | 71.67 | 89.57 | 61.94 | 97.47 | 70.56 | 98.09 | 80.56 | 97.90 | 78.33 | 98.12 | 77.78 |
| $S_{0.5}^U$ | 95.09 | 65.83 | 89.26 | 62.22 | 97.25 | 72.50 | 98.36 | 80.83 | 97.96 | 80.00 | 98.12 | 79.72 |
| $S_{0.75}^n$ | 94.69 | 71.67 | 92.01 | 69.17 | 97.44 | 75.00 | 98.80 | 79.44 | 98.36 | 78.33 | 98.36 | 80.28 |
| $S_{0.75}^U$ | 95.80 | 73.06 | 91.67 | 66.39 | 98.33 | 74.72 | 98.52 | 81.94 | 97.93 | 80.28 | 98.24 | 83.06 |
| $S_{1.0}^n$ | 96.54 | 75.56 | 94.54 | 70.83 | 98.27 | 76.67 | 98.73 | 81.94 | 98.67 | 78.33 | 98.64 | 80.00 |
| $S_{1.0}^U$ | 94.63 | 63.89 | 93.33 | 68.33 | 98.02 | 74.44 | 98.49 | 78.33 | 98.24 | 80.56 | 98.15 | 81.11 |
| $S_{1.5}^n$ | 92.04 | 65.56 | 94.54 | 67.78 | 98.33 | 76.94 | 98.21 | 79.44 | 98.80 | 80.56 | 98.67 | 81.11 |
| $S_{1.5}^U$ | 96.64 | 68.61 | 95.93 | 69.17 | 98.27 | 75.00 | 98.27 | 80.28 | 98.33 | 80.83 | 97.99 | 82.22 |
| $S_{2.0}^n$ | 95.31 | 69.17 | 95.43 | 70.28 | 98.09 | 80.00 | 97.69 | 79.17 | 98.30 | 83.33 | 98.24 | 79.72 |
| $S_{2.0}^U$ | 96.17 | 68.06 | 95.65 | 71.11 | 97.87 | 75.28 | 98.02 | 79.17 | 98.95 | <i>84.72</i> | 98.61 | 80.83 |
| $S_{2.5}^n$ | 93.12 | 63.61 | 94.07 | 66.39 | 96.98 | 76.11 | 97.93 | 78.89 | 98.33 | 80.28 | 98.21 | 79.72 |
| $S_{2.5}^U$ | 94.63 | 63.89 | 95.59 | 67.78 | 97.07 | 69.44 | 98.02 | 77.78 | 98.58 | 80.56 | 98.49 | 80.28 |
| $S_{3.0}^n$ | 94.17 | 64.72 | 94.66 | 70.00 | 96.48 | 72.78 | 97.04 | 78.89 | 97.35 | 80.00 | 97.75 | 79.44 |
| $S_{3.0}^U$ | 94.60 | 61.67 | 94.57 | 62.78 | 96.48 | 66.11 | 97.47 | 78.06 | 97.75 | 80.00 | 97.93 | 75.56 |
| $S_{4.0}^n$ | 93.86 | 60.28 | 93.46 | 65.56 | 95.68 | 70.28 | 97.84 | 74.72 | 98.27 | 79.72 | 97.87 | 78.06 |
| $S_{4.0}^U$ | 94.72 | 55.28 | 92.28 | 57.22 | 95.37 | 62.22 | 96.85 | 76.11 | 96.91 | 74.17 | 96.82 | 75.00 |

Tabela A.11: Resultados produzios pelas Θ -FAMs para Pima.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------------|--------------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 95.04 | 66.40 | 94.04 | 66.79 | 97.35 | 66.68 | 96.69 | 60.53 | 96.53 | 60.83 | 97.25 | 64.05 |
| S_W | 95.10 | 67.95 | 92.98 | 63.68 | 97.63 | 65.75 | 97.06 | 60.54 | 96.61 | 61.46 | 97.35 | 61.05 |
| $S_{0.25}^n$ | 94.85 | 65.61 | 94.52 | 65.62 | 97.24 | 63.16 | 96.43 | 61.46 | 96.51 | 60.44 | 97.38 | 61.99 |
| $S_{0.25}^u$ | 94.92 | 67.05 | 94.11 | 65.76 | 97.45 | 61.35 | 96.99 | 58.84 | 96.14 | 62.90 | 97.66 | 61.06 |
| $S_{0.5}^n$ | 94.73 | 65.48 | 93.58 | 65.88 | 97.71 | 67.33 | 96.72 | 60.68 | 96.63 | 61.34 | 97.16 | 61.84 |
| $S_{0.5}^u$ | 95.34 | 67.71 | 93.58 | 65.22 | 97.53 | 63.94 | 96.90 | 59.11 | 96.61 | 61.87 | 97.31 | 62.49 |
| $S_{0.75}^n$ | 94.57 | 66.79 | 94.20 | 63.67 | 97.57 | 66.82 | 96.31 | 59.12 | 96.35 | 58.62 | 97.50 | 60.80 |
| $S_{0.75}^u$ | 95.07 | 69.12 | 93.82 | 62.63 | 97.51 | 66.42 | 97.50 | 61.06 | 96.50 | 60.04 | 97.41 | 60.03 |
| $S_{1.0}^n$ | 94.47 | 64.05 | 94.17 | 64.70 | 97.82 | 67.06 | 96.59 | 60.54 | 96.22 | 60.17 | 97.29 | 61.58 |
| $S_{1.0}^u$ | 95.18 | 67.95 | 93.65 | 66.53 | 97.74 | 67.20 | 96.93 | 61.84 | 96.46 | 60.83 | 97.21 | 62.63 |
| $S_{1.5}^n$ | 93.94 | 68.11 | 94.40 | 66.14 | 97.86 | <i>67.44</i> | 96.85 | 60.15 | 96.37 | 63.03 | 97.57 | 61.33 |
| $S_{1.5}^u$ | 94.88 | 62.11 | 93.92 | 65.37 | 97.32 | 69.28 | 96.80 | 61.97 | 96.44 | 63.69 | 97.45 | 63.54 |
| $S_{2.0}^n$ | 94.72 | 64.97 | 93.68 | 65.88 | 97.16 | 67.18 | 97.08 | 60.95 | 96.17 | 62.38 | 97.40 | 61.59 |
| $S_{2.0}^u$ | 94.01 | 62.49 | 93.19 | 65.23 | 97.19 | 66.16 | 97.54 | 60.80 | 96.59 | 62.26 | 97.60 | 60.93 |
| $S_{2.5}^n$ | 93.61 | 65.51 | 93.74 | 64.30 | 96.98 | 65.77 | 96.43 | 63.55 | 96.31 | 64.22 | 97.27 | 61.59 |
| $S_{2.5}^u$ | 93.55 | 64.70 | 93.48 | 66.53 | 97.19 | 63.41 | 97.03 | 61.07 | 95.96 | 61.60 | 96.95 | 61.31 |
| $S_{3.0}^n$ | 92.92 | 68.08 | 94.05 | 66.40 | 96.98 | 65.36 | 97.16 | 61.86 | 96.04 | 62.14 | 97.14 | 61.20 |
| $S_{3.0}^u$ | 93.08 | 66.01 | 93.76 | 66.01 | 96.86 | 65.76 | 97.05 | 63.28 | 95.38 | 61.10 | 97.14 | 61.06 |
| $S_{4.0}^n$ | 92.98 | 64.20 | 93.59 | 67.57 | 96.61 | 65.75 | 97.06 | 60.02 | 95.72 | 61.86 | 96.72 | 66.15 |
| $S_{4.0}^u$ | 92.56 | 64.44 | 92.27 | 65.89 | 96.92 | 65.11 | 97.15 | 60.42 | 95.20 | 63.16 | 97.11 | 63.92 |

Tabela A.12: Resultados produzios pelas Θ -FAMs para Spectfheart.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|--------------|------------|-------|--------|-------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 99.54 | 76.38 | 98.54 | 67.05 | 98.50 | 78.30 | 98.59 | 69.17 | 98.29 | 69.20 | 98.71 | 70.70 |
| S_W | 99.29 | 79.03 | 98.79 | 68.95 | 98.38 | 76.50 | 98.67 | 70.37 | 98.54 | 70.38 | 98.67 | 69.64 |
| $S_{0.25}^n$ | 99.58 | 76.79 | 99.00 | 65.88 | 98.50 | 76.07 | 98.59 | 71.87 | 98.67 | 70.38 | 98.50 | 69.60 |
| $S_{0.25}^u$ | 99.63 | 74.90 | 98.63 | 68.53 | 99.13 | 74.96 | 98.75 | 73.43 | 98.63 | 66.62 | 98.33 | 70.04 |
| $S_{0.5}^n$ | 99.13 | 79.76 | 98.84 | 65.57 | 98.38 | 73.45 | 98.38 | 71.89 | 98.42 | 66.65 | 98.59 | 69.60 |
| $S_{0.5}^u$ | 99.33 | 77.89 | 98.71 | 66.31 | 98.75 | 72.28 | 98.38 | 70.41 | 98.54 | 67.01 | 98.38 | 69.26 |
| $S_{0.75}^n$ | 99.67 | 81.64 | 98.84 | 65.16 | 98.13 | 73.85 | 98.75 | 70.04 | 98.34 | 71.95 | 98.67 | 70.85 |
| $S_{0.75}^u$ | 99.67 | 79.83 | 98.88 | 69.33 | 98.50 | 73.08 | 98.83 | 74.13 | 98.50 | 66.57 | 98.96 | 71.88 |
| $S_{1.0}^n$ | 99.46 | 77.15 | 99.17 | 67.39 | 98.46 | 73.48 | 99.04 | 71.51 | 98.54 | 73.77 | 98.84 | 74.19 |
| $S_{1.0}^u$ | 99.33 | 77.14 | 98.71 | 67.42 | 98.92 | 73.06 | 98.34 | 70.80 | 98.63 | 71.21 | 98.29 | 70.81 |
| $S_{1.5}^n$ | 99.63 | 76.78 | 99.25 | 65.16 | 98.50 | 74.97 | 98.46 | 71.11 | 98.21 | 70.10 | 98.38 | 71.50 |
| $S_{1.5}^u$ | 99.50 | 79.40 | 99.04 | 68.52 | 98.63 | 74.96 | 98.67 | 68.90 | 98.29 | 75.27 | 98.63 | 73.35 |
| $S_{2.0}^n$ | 99.42 | 78.66 | 98.88 | 65.87 | 98.29 | 78.62 | 98.58 | 68.87 | 97.92 | 74.93 | 98.38 | 70.75 |
| $S_{2.0}^u$ | 99.58 | 77.19 | 98.54 | 70.43 | 98.50 | 76.45 | 98.84 | 70.00 | 98.88 | 71.92 | 98.25 | 72.99 |
| $S_{2.5}^n$ | 99.50 | 79.44 | 98.50 | 69.29 | 98.46 | 74.27 | 98.63 | 68.50 | 98.67 | 73.43 | 98.54 | 73.77 |
| $S_{2.5}^u$ | 99.63 | 82.04 | 99.00 | 69.64 | 98.75 | 75.30 | 98.83 | 69.20 | 98.13 | 70.75 | 98.42 | 75.24 |
| $S_{3.0}^n$ | 99.25 | 77.92 | 99.04 | 66.98 | 98.38 | 75.71 | 98.92 | 68.49 | 98.38 | 72.69 | 98.67 | 76.75 |
| $S_{3.0}^u$ | 99.71 | 78.23 | 99.17 | 67.79 | 98.54 | 74.20 | 98.79 | 70.01 | 99.21 | 73.85 | 98.21 | 75.63 |
| $S_{4.0}^n$ | 99.29 | 77.91 | 98.79 | 66.67 | 98.38 | 76.88 | 99.08 | 67.07 | 98.92 | 75.66 | 98.92 | 74.13 |
| $S_{4.0}^u$ | 99.75 | <i>81.30</i> | 98.75 | 67.44 | 98.25 | 78.72 | 98.79 | 67.05 | 99.75 | 73.12 | 98.25 | 74.52 |

Tabela A.13: Resultados produzios pelas Θ -FAMs para Sonar.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------|-------|-----------------------|-------|--------------|--------------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 99.47 | 74.45 | 99.63 | 80.74 | 99.73 | 84.52 | 99.52 | 84.57 | 99.89 | <i>80.69</i> | 99.52 | 80.69 |
| S_W | 99.36 | 76.00 | 99.20 | 75.52 | 99.68 | 82.14 | 99.47 | 79.26 | 99.63 | 81.69 | 99.63 | 79.26 |
| $S_{0.25}^n$ | 99.63 | 69.21 | 98.40 | 70.24 | 99.25 | 79.31 | 99.52 | 83.07 | 99.68 | 79.76 | 99.47 | 81.21 |
| $S_{0.25}^U$ | 99.63 | 68.24 | 98.98 | 66.83 | 99.36 | 81.17 | 99.25 | 82.12 | 99.79 | 80.71 | 99.57 | 84.10 |
| $S_{0.5}^n$ | 99.41 | 77.40 | 98.98 | 71.17 | 99.47 | 83.60 | 99.68 | 86.50 | 99.57 | 80.19 | 99.63 | 83.67 |
| $S_{0.5}^U$ | 99.68 | 71.74 | 98.24 | 67.79 | 99.57 | 84.14 | 99.47 | 83.19 | 99.63 | 83.07 | 99.57 | 79.24 |
| $S_{0.75}^n$ | 99.68 | 78.43 | 99.68 | 79.33 | 99.68 | 81.19 | 99.68 | 84.12 | 99.36 | 81.19 | 99.47 | 80.69 |
| $S_{0.75}^U$ | 99.52 | 74.00 | 99.14 | 74.45 | 99.68 | 79.74 | 99.63 | 85.55 | 99.25 | 81.14 | 99.41 | 82.64 |
| $S_{1.0}^n$ | 99.63 | 76.50 | 99.41 | 83.21 | 99.57 | 84.10 | 99.63 | 86.00 | 99.79 | 82.17 | 99.57 | 83.57 |
| $S_{1.0}^U$ | 99.52 | 72.21 | 99.20 | 77.86 | 99.52 | 78.29 | 99.52 | 84.62 | 99.57 | 84.55 | 99.41 | 80.71 |
| $S_{1.5}^n$ | 99.31 | 70.19 | 99.79 | 81.21 | 98.99 | 79.24 | 99.68 | 85.10 | 99.57 | 85.00 | 99.63 | 87.00 |
| $S_{1.5}^U$ | 99.14 | 71.71 | 99.79 | 80.26 | 99.47 | 77.26 | 99.20 | 83.17 | 99.52 | 83.07 | 99.52 | 82.19 |
| $S_{2.0}^n$ | 99.47 | 69.76 | 99.47 | 76.88 | 99.31 | 78.24 | 99.57 | 80.81 | 99.68 | 80.24 | 99.15 | 79.31 |
| $S_{2.0}^U$ | 99.20 | 65.98 | 99.25 | 75.57 | 99.47 | 76.40 | 99.36 | 79.76 | 99.63 | 82.21 | 99.15 | 82.14 |
| $S_{2.5}^n$ | 98.77 | 73.07 | 98.45 | 77.90 | 98.61 | 76.83 | 98.66 | 81.79 | 99.04 | 78.26 | 98.82 | 76.43 |
| $S_{2.5}^U$ | 99.20 | 65.95 | 98.13 | 73.10 | 98.83 | 72.67 | 98.98 | 77.93 | 99.36 | 81.31 | 98.93 | 74.98 |
| $S_{3.0}^n$ | 98.83 | 73.57 | 98.56 | 75.14 | 98.50 | 75.90 | 99.04 | 77.43 | 99.09 | 78.83 | 99.31 | 74.52 |
| $S_{3.0}^U$ | 98.61 | 63.50 | 97.81 | 67.74 | 98.50 | 75.02 | 99.31 | 77.90 | 98.98 | 78.83 | 98.82 | 71.14 |
| $S_{4.0}^n$ | 98.40 | 68.29 | 97.49 | 67.83 | 98.98 | 76.50 | 99.09 | 75.98 | 98.77 | 73.55 | 98.61 | 74.57 |
| $S_{4.0}^U$ | 96.26 | 51.98 | 97.97 | 58.26 | 98.45 | 67.90 | 98.77 | 73.64 | 98.98 | 67.74 | 98.66 | 68.21 |

Tabela A.14: Resultados produzios pelas Θ -FAMs para Vowel.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------|-------|-----------------------|-------|--------------|--------------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 99.90 | 96.67 | 99.83 | 96.67 | 99.88 | 97.98 | 99.89 | 97.37 | 99.91 | 96.77 | 99.93 | 97.07 |
| S_W | 99.81 | 95.96 | 99.78 | 95.66 | 99.83 | 97.37 | 99.87 | 97.98 | 99.96 | 98.28 | 99.82 | 98.18 |
| $S_{0.25}^n$ | 99.75 | 94.24 | 99.44 | 89.19 | 99.73 | 95.15 | 99.76 | 98.18 | 99.87 | 96.16 | 99.85 | 97.47 |
| $S_{0.25}^U$ | 99.72 | 94.04 | 99.33 | 89.39 | 99.71 | 95.15 | 99.90 | 97.37 | 99.96 | 97.68 | 99.87 | 98.38 |
| $S_{0.5}^n$ | 99.85 | 96.77 | 99.73 | 93.84 | 99.78 | 97.68 | 99.91 | 98.48 | 99.84 | 97.37 | 99.91 | 98.89 |
| $S_{0.5}^U$ | 99.90 | 95.76 | 99.66 | 93.54 | 99.88 | 98.28 | 99.91 | 97.88 | 99.99 | <i>97.07</i> | 99.92 | 98.59 |
| $S_{0.75}^n$ | 99.92 | 97.27 | 99.80 | 95.66 | 99.88 | 97.47 | 99.90 | 98.18 | 99.81 | 97.58 | 99.84 | 98.18 |
| $S_{0.75}^U$ | 99.93 | 96.06 | 99.82 | 94.55 | 99.91 | 97.47 | 99.94 | 97.88 | 99.92 | 97.17 | 99.87 | 97.78 |
| $S_{1.0}^n$ | 99.81 | 97.07 | 99.88 | 96.87 | 99.93 | 97.37 | 99.84 | 98.59 | 99.90 | 97.58 | 99.85 | 98.18 |
| $S_{1.0}^U$ | 99.88 | 96.57 | 99.76 | 95.25 | 99.84 | 97.07 | 99.89 | 97.17 | 99.91 | 97.58 | 99.85 | 97.37 |
| $S_{1.5}^n$ | 99.96 | 94.65 | 99.84 | 95.66 | 99.91 | 97.88 | 99.89 | 98.28 | 99.93 | 97.98 | 99.84 | 98.69 |
| $S_{1.5}^U$ | 99.82 | 94.04 | 99.79 | 92.93 | 99.90 | 96.67 | 99.91 | 97.58 | 99.90 | 97.37 | 99.91 | 98.79 |
| $S_{2.0}^n$ | 99.64 | 92.02 | 99.63 | 91.72 | 99.81 | 97.27 | 99.82 | 96.77 | 99.88 | 97.17 | 99.78 | 97.98 |
| $S_{2.0}^U$ | 99.54 | 88.59 | 99.64 | 88.08 | 99.89 | 96.57 | 99.84 | 96.57 | 99.84 | 96.06 | 99.94 | 97.58 |
| $S_{2.5}^n$ | 99.27 | 87.58 | 99.39 | 88.08 | 99.90 | 96.57 | 99.85 | 97.47 | 99.96 | 97.47 | 99.91 | 98.08 |
| $S_{2.5}^U$ | 99.21 | 84.65 | 99.16 | 86.57 | 99.80 | 95.96 | 99.90 | 95.86 | 99.73 | 93.13 | 99.92 | 97.07 |
| $S_{3.0}^n$ | 98.35 | 81.62 | 99.15 | 84.34 | 99.72 | 94.44 | 99.81 | 95.05 | 99.93 | 95.35 | 99.87 | 97.07 |
| $S_{3.0}^U$ | 98.27 | 76.77 | 98.71 | 78.69 | 99.57 | 92.63 | 99.72 | 93.94 | 99.74 | 88.08 | 99.81 | 96.16 |
| $S_{4.0}^n$ | 95.71 | 73.23 | 97.83 | 77.78 | 99.06 | 89.39 | 99.62 | 92.63 | 99.79 | 93.33 | 99.85 | 95.56 |
| $S_{4.0}^U$ | 96.77 | 70.81 | 96.86 | 68.08 | 99.38 | 88.59 | 99.71 | 92.53 | 99.21 | 71.82 | 99.85 | 95.35 |

Tabela A.15: Resultados produzios pelas Θ -FAMs para Wdbc.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|-------|------------|-------|--------------|--------------|-----------------------|-------|------------|-------|--------|-------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 99.45 | 91.91 | 99.73 | 95.96 | 99.61 | 95.78 | 99.63 | 92.27 | 99.71 | 92.97 | 99.69 | 93.67 |
| S_W | 99.63 | 93.85 | 99.77 | 94.03 | 99.57 | 94.90 | 99.67 | 92.27 | 99.71 | 92.97 | 99.67 | 93.49 |
| $S_{0.25}^n$ | 99.43 | 92.79 | 99.71 | 94.55 | 99.73 | 94.72 | 99.79 | 93.14 | 99.71 | 92.96 | 99.71 | 93.14 |
| $S_{0.25}^u$ | 99.47 | 91.57 | 99.69 | 92.96 | 99.47 | 94.55 | 99.73 | 94.37 | 99.73 | 92.26 | 99.75 | 92.61 |
| $S_{0.5}^n$ | 99.41 | 91.74 | 99.63 | 94.37 | 99.63 | 95.60 | 99.73 | 93.49 | 99.69 | 92.44 | 99.69 | 92.96 |
| $S_{0.5}^u$ | 99.38 | 92.44 | 99.73 | 94.73 | 99.79 | 95.25 | 99.67 | 92.79 | 99.57 | 93.32 | 99.61 | 92.97 |
| $S_{0.75}^n$ | 99.61 | 92.27 | 99.69 | 93.67 | 99.73 | 96.31 | 99.51 | 92.61 | 99.79 | 92.97 | 99.65 | 94.02 |
| $S_{0.75}^u$ | 99.47 | 92.79 | 99.69 | 93.84 | 99.86 | 95.78 | 99.53 | 92.79 | 99.57 | 92.61 | 99.59 | 93.49 |
| $S_{1.0}^n$ | 99.38 | 91.38 | 99.84 | 93.67 | 99.77 | 95.61 | 99.57 | 92.97 | 99.79 | 93.15 | 99.69 | 93.67 |
| $S_{1.0}^u$ | 99.39 | 92.62 | 99.69 | 94.38 | 99.88 | <i>96.14</i> | 99.59 | 93.15 | 99.79 | 93.15 | 99.69 | 93.67 |
| $S_{1.5}^n$ | 99.39 | 92.79 | 99.67 | 92.62 | 99.69 | 95.08 | 99.41 | 91.38 | 99.67 | 93.14 | 99.49 | 92.96 |
| $S_{1.5}^u$ | 99.53 | 90.34 | 99.57 | 93.32 | 99.59 | 94.91 | 99.49 | 92.78 | 99.67 | 92.44 | 99.59 | 92.61 |
| $S_{2.0}^n$ | 99.34 | 92.09 | 99.55 | 91.74 | 99.79 | 94.19 | 99.41 | 91.21 | 99.69 | 91.74 | 99.57 | 93.31 |
| $S_{2.0}^u$ | 99.22 | 90.49 | 99.63 | 91.04 | 99.75 | 93.85 | 99.49 | 91.20 | 99.71 | 90.86 | 99.65 | 92.61 |
| $S_{2.5}^n$ | 99.08 | 90.85 | 99.55 | 92.10 | 99.82 | 93.49 | 99.53 | 91.56 | 99.73 | 91.74 | 99.55 | 91.91 |
| $S_{2.5}^u$ | 99.38 | 92.09 | 99.51 | 91.75 | 99.77 | 94.55 | 99.49 | 90.68 | 99.71 | 89.98 | 99.57 | 91.74 |
| $S_{3.0}^n$ | 99.30 | 90.68 | 99.67 | 92.27 | 99.67 | 91.73 | 99.55 | 90.33 | 99.73 | 91.39 | 99.63 | 91.38 |
| $S_{3.0}^u$ | 99.16 | 89.45 | 99.45 | 90.86 | 99.49 | 90.15 | 99.38 | 91.56 | 99.67 | 89.10 | 99.59 | 90.32 |
| $S_{4.0}^n$ | 99.18 | 90.33 | 99.20 | 90.69 | 99.49 | 91.03 | 99.30 | 89.80 | 99.69 | 90.86 | 99.45 | 91.02 |
| $S_{4.0}^u$ | 99.30 | 88.21 | 99.20 | 90.15 | 99.18 | 88.74 | 99.38 | 89.98 | 98.77 | 84.35 | 99.59 | 90.85 |

Tabela A.16: Resultados produzios pelas Θ -FAMs para Wine.

| Subsethood | $n(\mathbf{x})$ | | | | | | $\Psi(n(\mathbf{x}))$ | | | | | |
|--------------|-----------------|--------------|------------|-------|---------------|--------------|-----------------------|--------------|---------------|--------------|---------------|--------------|
| | S-FAM | | S-FAM dual | | E-FAM | | S-FAM | | S-FAM dual | | E-FAM | |
| | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste | Trein. | Teste |
| S_K | 99.88 | 94.31 | 98.94 | 90.98 | 99.94 | 92.68 | 100.00 | <i>97.22</i> | 100.00 | <i>97.22</i> | 100.00 | <i>97.78</i> |
| S_W | 99.94 | 96.08 | 99.31 | 90.42 | 100.00 | <i>94.97</i> | 99.94 | 96.11 | 99.94 | 95.56 | 99.94 | 96.11 |
| $S_{0.25}^n$ | 99.94 | 95.56 | 99.75 | 91.54 | 100.00 | <i>94.93</i> | 100.00 | <i>97.22</i> | 99.88 | 96.63 | 100.00 | <i>94.97</i> |
| $S_{0.25}^u$ | 100.00 | <i>96.08</i> | 99.19 | 92.16 | 100.00 | <i>94.90</i> | 100.00 | <i>97.22</i> | 100.00 | <i>96.63</i> | 100.00 | <i>96.63</i> |
| $S_{0.5}^n$ | 100.00 | <i>95.56</i> | 99.13 | 91.54 | 100.00 | <i>94.97</i> | 100.00 | <i>97.78</i> | 100.00 | <i>97.78</i> | 100.00 | <i>96.08</i> |
| $S_{0.5}^u$ | 99.88 | 95.56 | 99.69 | 90.39 | 100.00 | <i>93.82</i> | 100.00 | <i>97.78</i> | 100.00 | <i>97.78</i> | 99.94 | 97.22 |
| $S_{0.75}^n$ | 99.81 | 95.56 | 98.88 | 89.87 | 99.75 | 93.30 | 100.00 | <i>98.33</i> | 100.00 | <i>97.75</i> | 100.00 | <i>97.75</i> |
| $S_{0.75}^u$ | 99.88 | 95.52 | 98.69 | 88.10 | 99.81 | 94.38 | 100.00 | <i>97.78</i> | 99.88 | 96.08 | 99.94 | 96.63 |
| $S_{1.0}^n$ | 99.81 | 93.82 | 99.00 | 90.95 | 99.88 | 93.27 | 100.00 | <i>97.22</i> | 100.00 | <i>97.22</i> | 100.00 | <i>97.78</i> |
| $S_{1.0}^u$ | 99.94 | 92.71 | 98.69 | 88.14 | 99.94 | 91.60 | 99.94 | 97.78 | 99.88 | 97.22 | 99.94 | 97.78 |
| $S_{1.5}^n$ | 99.81 | 93.27 | 99.06 | 89.31 | 100.00 | <i>90.98</i> | 99.94 | 96.08 | 99.94 | 96.67 | 99.81 | 94.97 |
| $S_{1.5}^u$ | 99.56 | 93.30 | 99.00 | 91.01 | 99.62 | 90.49 | 99.94 | 96.63 | 99.94 | 96.11 | 99.94 | 94.93 |
| $S_{2.0}^n$ | 99.75 | 89.93 | 98.94 | 90.98 | 99.69 | 86.54 | 99.81 | 93.86 | 100.00 | <i>96.08</i> | 99.94 | 92.68 |
| $S_{2.0}^u$ | 99.38 | 87.12 | 98.94 | 88.69 | 99.63 | 86.47 | 99.88 | 96.11 | 99.94 | 94.97 | 99.81 | 94.35 |
| $S_{2.5}^n$ | 99.19 | 83.17 | 99.06 | 86.63 | 99.06 | 84.31 | 99.94 | 94.41 | 99.94 | 93.82 | 99.88 | 93.79 |
| $S_{2.5}^u$ | 99.63 | 83.20 | 98.81 | 84.22 | 99.19 | 82.55 | 99.75 | 94.41 | 99.75 | 89.25 | 99.62 | 93.79 |
| $S_{3.0}^n$ | 98.50 | 84.90 | 98.88 | 82.03 | 98.81 | 81.47 | 99.63 | 92.71 | 99.94 | 94.38 | 99.75 | 91.01 |
| $S_{3.0}^u$ | 98.75 | 79.84 | 98.88 | 82.61 | 98.31 | 81.44 | 99.94 | 92.68 | 99.81 | 89.41 | 99.81 | 89.84 |
| $S_{4.0}^n$ | 98.88 | 77.55 | 98.07 | 82.71 | 98.63 | 77.55 | 99.63 | 92.19 | 99.81 | 88.79 | 99.81 | 89.31 |
| $S_{4.0}^u$ | 98.19 | 78.14 | 97.25 | 78.56 | 97.63 | 66.27 | 99.81 | 93.27 | 98.63 | 73.99 | 99.56 | 87.61 |

A.2 Problema de Auto-localização de Robô Baseado em Visão

Nesta seção, apresentaremos os resultados obtidos pelos modelos Θ -FAMs em um problema de identificação de localização (ou posições de referência) de um robô móvel baseado em um conjunto de imagens adquiridas durante sua operação [98, 108, 109].

O conjunto de dados é composto de 61×79 imagens de 11 posições de referências diferentes capturadas em 6 percursos. A Tabela A.17 mostra o número de imagens para cada percurso e posição de referência.

Tabela A.17: Número de imagens por percurso e posição.

| <i>Percurso</i> | 1^a | 2^a | 3^a | 4^a | 5^a | 6^a | 7^a | 8^a | 9^a | 10^a | 11^a | Total |
|-------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|-----------------------|-----------------------|--------------|
| Percurso 1 | 32 | 15 | 37 | 30 | 29 | 45 | 21 | 49 | 36 | 36 | 32 | 362 |
| Percurso 2 | 30 | 9 | 33 | 26 | 27 | 42 | 34 | 58 | 27 | 32 | 39 | 357 |
| Percurso 3 | 19 | 27 | 38 | 36 | 35 | 47 | 38 | 59 | 36 | 35 | 34 | 404 |
| Percurso 4 | 33 | 22 | 30 | 40 | 30 | 45 | 33 | 49 | 35 | 26 | 42 | 385 |
| Percurso 5 | 26 | 22 | 39 | 38 | 36 | 46 | 39 | 41 | 27 | 34 | 35 | 383 |
| Percurso 6 | 32 | 21 | 35 | 33 | 32 | 39 | 36 | 50 | 30 | 32 | 34 | 374 |

Cada imagem foi convertida em um vetor em \mathbb{R}^{4819} via o método *row scan* e, então, cada vetor foi associado a um conjunto *fuzzy* pela aplicação do operador $\Psi(n(\cdot))$ tal como descrito acima. O Percurso 1 e o Percurso 2 foram usados como conjuntos de treinamento e validação, respectivamente. Os outros percursos foram usados como conjunto de teste. Em nossos experimentos, consideramos dois conjuntos de memórias fundamentais diferentes de acordo com a sequência com que as imagens foram obtidas: um formado pelas imagens com numeração pares \mathcal{M}_e e outro formado pelas imagens com numeração ímpares (\mathcal{M}_o). As Tabelas A.18, A.19 e A.20 exibem, respectivamente, os resultados obtidos pelos modelos S-FAM, S-FAM dual e E-FAM usando o conjunto \mathcal{M}_e como memória fundamental. As Tabelas A.21, A.22 e A.23 mostram, respectivamente, os resultados obtidos pelos modelos S-FAM, S-FAM dual e E-FAM usando o conjunto \mathcal{M}_o como memória fundamental. Os valores em negrito indicam os melhores resultados no conjunto de validação (correspondente ao Percurso 2).

Tabela A.18: Os resultados produzidos pelas S-FAMs usando \mathcal{M}_e .

| Subsethood | Percurso 1 | Percurso 2 | Percurso 3 | Percurso 4 | Percurso 5 | Percurso 6 | Média |
|--------------|------------|------------|------------|------------|------------|------------|-------|
| S_K | 99.72 | 87.1148 | 76.24 | 72.21 | 75.98 | 71.39 | 73.95 |
| S_W | 99.72 | 87.1148 | 76.24 | 72.21 | 75.98 | 71.39 | 73.95 |
| $S_{0.25}^n$ | 99.72 | 88.7955 | 79.70 | 73.77 | 80.94 | 78.88 | 78.32 |
| $S_{0.25}^u$ | 99.72 | 88.7955 | 79.70 | 73.77 | 80.94 | 78.88 | 78.32 |
| $S_{0.5}^n$ | 99.72 | 87.6751 | 78.22 | 72.73 | 79.63 | 78.34 | 77.23 |
| $S_{0.5}^u$ | 99.72 | 87.9552 | 78.71 | 73.51 | 79.90 | 78.34 | 77.61 |
| $S_{0.75}^n$ | 99.72 | 87.9552 | 77.48 | 71.95 | 78.85 | 76.74 | 76.25 |
| $S_{0.75}^u$ | 99.72 | 87.9552 | 77.48 | 72.21 | 78.85 | 76.74 | 76.32 |
| $S_{1.0}^n$ | 99.72 | 87.1148 | 76.49 | 71.95 | 77.28 | 72.99 | 74.68 |
| $S_{1.0}^u$ | 99.72 | 87.1148 | 76.73 | 72.21 | 77.28 | 72.99 | 74.80 |
| $S_{1.5}^n$ | 99.72 | 84.8739 | 73.27 | 68.31 | 72.32 | 60.43 | 68.58 |
| $S_{1.5}^u$ | 99.72 | 84.3137 | 72.77 | 66.75 | 71.02 | 60.16 | 67.68 |
| $S_{2.0}^n$ | 99.72 | 80.6723 | 68.07 | 64.94 | 66.06 | 54.28 | 63.33 |
| $S_{2.0}^u$ | 99.72 | 76.1905 | 67.57 | 62.86 | 65.54 | 52.67 | 62.16 |
| $S_{2.5}^n$ | 99.72 | 74.5098 | 65.10 | 57.14 | 62.14 | 49.73 | 58.53 |
| $S_{2.5}^u$ | 99.72 | 66.9468 | 54.95 | 54.03 | 57.18 | 47.86 | 53.50 |
| $S_{3.0}^n$ | 99.72 | 67.2269 | 53.71 | 54.03 | 57.44 | 47.06 | 53.06 |
| $S_{3.0}^u$ | 99.72 | 53.5014 | 40.84 | 45.19 | 51.44 | 43.05 | 45.13 |
| $S_{4.0}^n$ | 99.72 | 52.9412 | 40.35 | 42.34 | 51.17 | 41.98 | 43.96 |
| $S_{4.0}^u$ | 99.45 | 46.2185 | 37.38 | 36.62 | 46.21 | 44.39 | 41.15 |

Tabela A.19: Os resultados produzidos pelas S-FAMs duais usando \mathcal{M}_e .

| Subsethood | Percurso 1 | Percurso 2 | Percurso 3 | Percurso 4 | Percurso 5 | Percurso 6 | Média |
|--------------|------------|----------------|--------------|--------------|--------------|--------------|--------------|
| S_K | 99.72 | 87.1148 | 76.24 | 72.21 | 75.98 | 71.39 | 73.95 |
| S_W | 99.72 | 87.1148 | 76.24 | 72.21 | 75.98 | 71.39 | 73.95 |
| $S_{0.25}^n$ | 99.72 | 82.0728 | 71.29 | 66.23 | 70.50 | 62.30 | 67.58 |
| $S_{0.25}^u$ | 99.72 | 82.0728 | 71.29 | 65.97 | 70.50 | 62.30 | 67.51 |
| $S_{0.5}^n$ | 99.72 | 84.0336 | 73.02 | 67.27 | 71.54 | 64.17 | 69.00 |
| $S_{0.5}^u$ | 99.72 | 83.7535 | 73.02 | 67.01 | 71.54 | 63.90 | 68.87 |
| $S_{0.75}^n$ | 99.72 | 85.7143 | 74.75 | 71.17 | 73.89 | 68.72 | 72.13 |
| $S_{0.75}^u$ | 99.72 | 85.7143 | 74.50 | 70.91 | 73.63 | 67.91 | 71.74 |
| $S_{1.0}^n$ | 99.72 | 87.1148 | 76.24 | 72.21 | 76.24 | 71.66 | 74.09 |
| $S_{1.0}^u$ | 99.72 | 87.1148 | 76.24 | 72.21 | 76.24 | 71.39 | 74.02 |
| $S_{1.5}^n$ | 99.72 | 88.5154 | 78.71 | 71.95 | 80.16 | 75.13 | 76.49 |
| $S_{1.5}^u$ | 99.72 | 88.2353 | 79.46 | 72.21 | 80.16 | 75.40 | 76.81 |
| $S_{2.0}^n$ | 99.72 | 89.0756 | 81.68 | 73.25 | 81.20 | 76.74 | 78.22 |
| $S_{2.0}^u$ | 99.72 | 88.5154 | 81.44 | 74.03 | 82.51 | 78.07 | 79.01 |
| $S_{2.5}^n$ | 99.72 | 90.1961 | <i>81.68</i> | <i>74.29</i> | <i>82.25</i> | <i>77.81</i> | <i>79.01</i> |
| $S_{2.5}^u$ | 99.72 | 89.6359 | 81.68 | 75.58 | 83.81 | 80.48 | 80.39 |
| $S_{3.0}^n$ | 99.72 | 90.1961 | <i>82.67</i> | <i>76.10</i> | <i>83.03</i> | <i>78.88</i> | <i>80.17</i> |
| $S_{3.0}^u$ | 99.72 | 89.9160 | 83.66 | 77.40 | 86.16 | 83.69 | 82.73 |
| $S_{4.0}^n$ | 99.72 | 89.9160 | 84.41 | 77.14 | 86.42 | 82.35 | 82.58 |
| $S_{4.0}^u$ | 99.72 | 89.0756 | 86.39 | 77.92 | 86.68 | 89.04 | 85.01 |

Tabela A.20: Os resultados produzidos pelas E-FAMs usando \mathcal{M}_e .

| Subsethood | Percurso 1 | Percurso 2 | Percurso 3 | Percurso 4 | Percurso 5 | Percurso 6 | Média |
|-------------------|------------|------------|------------|------------|------------|------------|-------|
| S_K | 99.72 | 87.1148 | 76.24 | 72.21 | 75.98 | 71.39 | 73.95 |
| S_W | 99.72 | 87.1148 | 76.24 | 72.21 | 75.98 | 71.39 | 73.95 |
| $S_{0.25}^{\cap}$ | 99.72 | 83.4734 | 71.29 | 66.23 | 72.06 | 62.57 | 68.04 |
| $S_{0.25}^{\cup}$ | 99.72 | 83.4734 | 71.29 | 65.97 | 72.06 | 62.57 | 67.97 |
| $S_{0.5}^{\cap}$ | 99.72 | 84.8739 | 73.02 | 69.09 | 72.85 | 64.71 | 69.92 |
| $S_{0.5}^{\cup}$ | 99.72 | 84.8739 | 73.02 | 68.57 | 72.85 | 64.71 | 69.79 |
| $S_{0.75}^{\cap}$ | 99.72 | 85.9944 | 75.50 | 71.43 | 74.41 | 69.25 | 72.65 |
| $S_{0.75}^{\cup}$ | 99.72 | 85.9944 | 75.25 | 71.69 | 74.41 | 68.72 | 72.52 |
| $S_{1.0}^{\cap}$ | 99.72 | 87.1148 | 76.49 | 72.21 | 76.24 | 71.66 | 74.15 |
| $S_{1.0}^{\cup}$ | 99.72 | 87.1148 | 76.73 | 72.21 | 76.24 | 71.93 | 74.28 |
| $S_{1.5}^{\cap}$ | 99.72 | 85.4342 | 74.01 | 68.57 | 73.11 | 61.50 | 69.30 |
| $S_{1.5}^{\cup}$ | 99.72 | 84.8739 | 73.27 | 67.01 | 72.58 | 60.70 | 68.39 |
| $S_{2.0}^{\cap}$ | 99.72 | 83.4734 | 70.05 | 65.45 | 68.41 | 55.08 | 64.75 |
| $S_{2.0}^{\cup}$ | 99.72 | 82.0728 | 68.32 | 65.45 | 67.10 | 53.48 | 63.59 |
| $S_{2.5}^{\cap}$ | 99.72 | 82.0728 | 66.83 | 63.64 | 65.01 | 49.73 | 61.30 |
| $S_{2.5}^{\cup}$ | 99.72 | 78.4314 | 62.13 | 58.18 | 61.88 | 47.86 | 57.51 |
| $S_{3.0}^{\cap}$ | 99.72 | 79.2717 | 62.13 | 59.22 | 61.36 | 47.06 | 57.44 |
| $S_{3.0}^{\cup}$ | 99.72 | 72.2689 | 54.70 | 52.99 | 56.40 | 43.85 | 51.98 |
| $S_{4.0}^{\cap}$ | 99.72 | 73.6695 | 55.69 | 52.99 | 57.18 | 43.05 | 52.23 |
| $S_{4.0}^{\cup}$ | 99.72 | 62.1849 | 37.62 | 36.88 | 46.48 | 39.30 | 40.07 |

Tabela A.21: Os resultados produzidos pelas S-FAMs usando \mathcal{M}_o .

| Subsethood | Percurso 1 | Percurso 2 | Percurso 3 | Percurso 4 | Percurso 5 | Percurso 6 | Média |
|-------------------|------------|------------|------------|------------|------------|------------|-------|
| S_K | 100.00 | 88.7955 | 74.75 | 71.43 | 75.46 | 72.73 | 73.59 |
| S_W | 100.00 | 88.7955 | 74.75 | 71.43 | 75.46 | 72.73 | 73.59 |
| $S_{0.25}^{\cap}$ | 100.00 | 89.6359 | 78.22 | 71.43 | 80.94 | 83.69 | 78.57 |
| $S_{0.25}^{\cup}$ | 100.00 | 89.6359 | 78.22 | 71.43 | 80.94 | 83.96 | 78.64 |
| $S_{0.5}^{\cap}$ | 100.00 | 89.3557 | 77.72 | 71.43 | 79.63 | 82.62 | 77.85 |
| $S_{0.5}^{\cup}$ | 100.00 | 89.3557 | 77.97 | 71.43 | 79.63 | 82.62 | 77.91 |
| $S_{0.75}^{\cap}$ | 100.00 | 89.3557 | 75.74 | 71.95 | 78.59 | 79.41 | 76.42 |
| $S_{0.75}^{\cup}$ | 100.00 | 89.3557 | 75.99 | 71.95 | 78.59 | 79.68 | 76.55 |
| $S_{1.0}^{\cap}$ | 100.00 | 89.0756 | 75.25 | 71.69 | 76.24 | 74.60 | 74.44 |
| $S_{1.0}^{\cup}$ | 100.00 | 89.0756 | 75.25 | 71.69 | 76.24 | 74.60 | 74.44 |
| $S_{1.5}^{\cap}$ | 100.00 | 86.2745 | 71.29 | 70.39 | 72.85 | 63.90 | 69.61 |
| $S_{1.5}^{\cup}$ | 100.00 | 85.4342 | 70.79 | 70.39 | 72.06 | 63.10 | 69.09 |
| $S_{2.0}^{\cap}$ | 100.00 | 79.8319 | 69.31 | 64.94 | 66.58 | 54.81 | 63.91 |
| $S_{2.0}^{\cup}$ | 100.00 | 76.7507 | 68.07 | 62.08 | 63.19 | 51.07 | 61.10 |
| $S_{2.5}^{\cap}$ | 100.00 | 71.4286 | 66.34 | 58.44 | 60.31 | 48.40 | 58.37 |
| $S_{2.5}^{\cup}$ | 100.00 | 65.8263 | 58.42 | 52.21 | 54.31 | 46.52 | 52.86 |
| $S_{3.0}^{\cap}$ | 100.00 | 66.3866 | 58.17 | 52.47 | 54.31 | 45.45 | 52.60 |
| $S_{3.0}^{\cup}$ | 100.00 | 52.3810 | 41.09 | 43.38 | 49.35 | 40.64 | 43.61 |
| $S_{4.0}^{\cap}$ | 100.00 | 52.6611 | 40.84 | 42.34 | 47.78 | 40.11 | 42.77 |
| $S_{4.0}^{\cup}$ | 99.72 | 31.6527 | 27.48 | 26.75 | 37.08 | 32.35 | 30.91 |

Tabela A.22: Os resultados produzidos pelas S-FAMs duais usando \mathcal{M}_o .

| Subsethood | Percurso 1 | Percurso 2 | Percurso 3 | Percurso 4 | Percurso 5 | Percurso 6 | Média |
|--------------|------------|----------------|------------|------------|------------|------------|-------|
| S_K | 100.00 | 88.7955 | 74.75 | 71.43 | 75.46 | 72.73 | 73.59 |
| S_W | 100.00 | 88.7955 | 74.75 | 71.43 | 75.46 | 72.73 | 73.59 |
| $S_{0.25}^n$ | 100.00 | 83.1933 | 69.31 | 64.42 | 72.32 | 64.71 | 67.69 |
| $S_{0.25}^u$ | 100.00 | 82.9132 | 69.06 | 64.16 | 72.32 | 64.44 | 67.49 |
| $S_{0.5}^n$ | 100.00 | 86.5546 | 71.04 | 70.13 | 73.63 | 66.31 | 70.28 |
| $S_{0.5}^u$ | 100.00 | 86.5546 | 71.04 | 69.87 | 73.37 | 65.78 | 70.01 |
| $S_{0.75}^n$ | 100.00 | 87.9552 | 72.28 | 71.17 | 74.67 | 70.32 | 72.11 |
| $S_{0.75}^u$ | 100.00 | 87.6751 | 72.03 | 71.17 | 74.41 | 70.32 | 71.98 |
| $S_{1.0}^n$ | 100.00 | 89.0756 | 75.00 | 71.43 | 75.46 | 72.99 | 73.72 |
| $S_{1.0}^u$ | 100.00 | 89.0756 | 74.75 | 71.43 | 75.46 | 72.99 | 73.66 |
| $S_{1.5}^n$ | 100.00 | 90.1961 | 77.97 | 72.21 | 78.07 | 76.74 | 76.25 |
| $S_{1.5}^u$ | 100.00 | 90.1961 | 78.71 | 72.21 | 79.37 | 76.74 | 76.76 |
| $S_{2.0}^n$ | 100.00 | 89.9160 | 81.19 | 72.99 | 79.37 | 77.27 | 77.71 |
| $S_{2.0}^u$ | 100.00 | 89.6359 | 81.19 | 72.47 | 80.16 | 79.95 | 78.44 |
| $S_{2.5}^n$ | 100.00 | 89.6359 | 81.44 | 72.99 | 80.16 | 79.14 | 78.43 |
| $S_{2.5}^u$ | 100.00 | 89.6359 | 81.68 | 74.29 | 80.16 | 84.22 | 80.09 |
| $S_{3.0}^n$ | 100.00 | 89.3557 | 81.44 | 74.29 | 80.16 | 81.02 | 79.22 |
| $S_{3.0}^u$ | 100.00 | 87.9552 | 82.92 | 75.58 | 83.03 | 83.96 | 81.37 |
| $S_{4.0}^n$ | 100.00 | 87.9552 | 83.42 | 76.62 | 82.77 | 82.35 | 81.29 |
| $S_{4.0}^u$ | 100.00 | 83.4734 | 84.16 | 76.88 | 79.11 | 81.55 | 80.43 |

Tabela A.23: Os resultados produzidos pelas E-FAMs usando \mathcal{M}_o .

| Subsethood | Percurso 1 | Percurso 2 | Percurso 3 | Percurso 4 | Percurso 5 | Percurso 6 | Média |
|--------------|------------|------------|------------|------------|------------|------------|-------|
| S_K | 100.00 | 88.7955 | 74.75 | 71.43 | 75.46 | 72.73 | 73.59 |
| S_W | 100.00 | 88.7955 | 74.75 | 71.43 | 75.46 | 72.73 | 73.59 |
| $S_{0.25}^n$ | 100.00 | 84.5938 | 69.31 | 68.05 | 72.85 | 65.51 | 68.93 |
| $S_{0.25}^u$ | 100.00 | 84.5938 | 69.31 | 67.53 | 72.85 | 65.24 | 68.73 |
| $S_{0.5}^n$ | 100.00 | 86.8347 | 71.29 | 70.39 | 73.89 | 67.65 | 70.80 |
| $S_{0.5}^u$ | 100.00 | 86.8347 | 71.29 | 70.39 | 73.63 | 66.84 | 70.54 |
| $S_{0.75}^n$ | 100.00 | 88.2353 | 72.77 | 71.43 | 74.93 | 70.86 | 72.50 |
| $S_{0.75}^u$ | 100.00 | 88.2353 | 72.52 | 71.43 | 74.93 | 70.86 | 72.44 |
| $S_{1.0}^n$ | 100.00 | 89.3557 | 75.25 | 71.69 | 75.98 | 73.53 | 74.11 |
| $S_{1.0}^u$ | 100.00 | 89.3557 | 75.25 | 71.69 | 75.98 | 73.53 | 74.11 |
| $S_{1.5}^n$ | 100.00 | 86.5546 | 71.29 | 71.43 | 72.85 | 64.71 | 70.07 |
| $S_{1.5}^u$ | 100.00 | 85.9944 | 71.04 | 71.17 | 72.58 | 63.64 | 69.61 |
| $S_{2.0}^n$ | 100.00 | 83.4734 | 69.80 | 68.31 | 69.71 | 56.15 | 65.99 |
| $S_{2.0}^u$ | 100.00 | 81.5126 | 68.81 | 65.97 | 67.36 | 52.41 | 63.64 |
| $S_{2.5}^n$ | 100.00 | 80.9524 | 67.33 | 66.23 | 63.45 | 48.93 | 61.48 |
| $S_{2.5}^u$ | 100.00 | 77.3109 | 65.59 | 58.70 | 60.84 | 47.06 | 58.05 |
| $S_{3.0}^n$ | 100.00 | 77.3109 | 65.84 | 59.74 | 60.84 | 45.99 | 58.10 |
| $S_{3.0}^u$ | 100.00 | 69.4678 | 54.46 | 53.51 | 55.09 | 41.71 | 51.19 |
| $S_{4.0}^n$ | 100.00 | 72.2689 | 57.67 | 55.58 | 55.87 | 41.44 | 52.64 |
| $S_{4.0}^u$ | 100.00 | 59.1036 | 40.35 | 41.82 | 45.69 | 38.50 | 41.59 |

Índice Remissivo

- \mathbb{L} -fuzzy, 6
- Θ -FAM, 16
 - capacidade de armazenamento, 19
 - E-FAM, 18
 - TE-FAM, 18
 - S-FAM, 18
 - TS-FAM, 18
 - S-FAM dual, 18
 - TS-FAM dual, 18
 - tolerância a ruídos, 19
- subsethood*, 7
 - de Kosko, 9
 - de Willmot, 9
- agregação, 7
- conjunção *fuzzy*, 5
 - t-norma, 5
- conjunto *fuzzy*, 5
 - intersecção, 6
 - união, 6
- equivalência, 10
 - forte, 11
- implicação *fuzzy*, 5
 - Gödel, 6
 - Goguen, 6
 - Lukasiewicz, 6
 - residual, 5
- inclusão *fuzzy*, 7
- memória associativa, 16
 - fuzzy*, 16
- memória fundamental, 16
 - seleção, 29
- poset, 4
- problema
 - de auto-localização de robô, 40
 - de referência, 36
- relação *fuzzy*, 5
- reticulado, 4
 - completo, 4
 - limitado, 4
 - totalmente ordenado, 4
- sup- T , 5
- t-norma
 - Lukasiewicz, 6
 - mínimo, 6
 - produto, 6
- treinamento
 - convergência, 24
 - de Θ -FAMs, 23
 - de TE-FAMs $\mathcal{E}_{W,v}$, 27