



UNIVERSIDADE
ESTADUAL DE LONDRINA

ALEX SANDRO DE CASTILHO

**CARACTERIZAÇÃO DE SISTEMAS DE REGRAS
NEBULOSAS EM RETICULADO COMPLETO E UMA
APLICAÇÃO NO PROBLEMA DE TRANSMISSÃO DE VOZ
SOBRE IP**

Londrina
2009

ALEX SANDRO DE CASTILHO

**CARACTERIZAÇÃO DE SISTEMAS DE REGRAS
NEBULOSAS EM RETICULADO COMPLETO E UMA
APLICAÇÃO NO PROBLEMA DE TRANSMISSÃO DE VOZ
SOBRE IP**

Dissertação apresentada ao Programa de Pós-Graduação, em Matemática da Aplicada e Computacional da Universidade Estadual de Londrina, como requisito parcial para a obtenção do título de Mestre em Matemática Aplicada e Computacional.

Orientador: Prof. Dr. Marcos Eduardo Valle

Londrina
2009

**Catálogo na publicação elaborada pela Divisão de Processos Técnicos da
Biblioteca Central da Universidade Estadual de Londrina.**

Dados Internacionais de Catalogação-na-Publicação (CIP)

C352c Castilho, Alex Sandro de.

Caracterização de sistemas de regras nebulosas em reticulado completo e uma aplicação no problema de transmissão de voz sobre IP / Alex Sandro de Castilho. – Londrina, 2009.
ix, 57 f. : il.

Orientador: Marcos Eduardo Valle.

Dissertação (Mestrado em Matemática Aplicada e Computacional) –
Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de
Pós-Graduação em Matemática Aplicada e Computacional, 2009.

Inclui bibliografia.

1. Conjuntos difusos – Teses. 2. Teoria dos conjuntos – Teses. 3. Teoria dos reticulados – Teses. 4. Redes neurais – Teses. I. Valle, Marcos Eduardo. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Matemática Aplicada e Computacional. III. Título.

CDU 510.22

ALEX SANDRO DE CASTILHO

**CARACTERIZAÇÃO DE SISTEMAS DE REGRAS
NEBULOSAS EM RETICULADO COMPLETO E UMA
APLICAÇÃO NO PROBLEMA DE TRANSMISSÃO DE VOZ
SOBRE IP**

BANCA EXAMINADORA

Prof. Dr. Marcos Eduardo Valle
Departamento de Matemática – UEL

Profa. Dra. Maria José Castanho
Departamento de Matemática – UNICENTRO

Prof. Dr. Robinson Hoto
Departamento de Matemática – UEL

Londrina, 14 de janeiro de 2009.

*À meu avô,
Gabriel Acedo Neto,
1931-2008.*

AGRADECIMENTO

Aos meus pais Valter Josée de Castilho e Zulmira Acedo que sempre me diziam que eu podia fazer qualquer coisa, apesar das minhas limitações. Isto me inspirou a continuidade nos estudos até chegar a este mestrado.

A Deus, pois sem ele, nada disto seria possível.

A todos os meus amigos, professores e funcionários da UEL, que de alguma forma contribuíram para a elaboração deste trabalho. Em especial aos professores e colegas de turma do PGMAC.

A disponibilidade irrestrita, sua paciência comigo, sua forma exigente, crítica e criativa de expor suas idéias, creio que foram fundamentais para a conclusão desse trabalho.

A Marcos Eduardo Valle, sou grato pela orientação.

À CAPES, pelo apoio financeiro a esta dissertação de mestrado.

A todos sou profundamente grato e dedico esta dissertação.

CASTILHO, Alex Sandro de. **Caracterização de sistemas de regras nebulosas em reticulado completo e uma aplicação no problema de transmissão de voz sobre IP.** 2009. 63f. Dissertação (Mestrado em Matemática Aplicada e Computacional) – Universidade Estadual de Londrina, Londrina, 2009.

RESUMO

Conjuntos nebulosos são usados para descrever conceitos vagos ou imprecisos. Sistemas de regras nebulosas são uma poderosa ferramenta matemática para modelar fenômenos usando uma linguagem natural. Essa dissertação está dividida em duas partes. A primeira parte refere-se a questões teóricas relativas a conjuntos nebulosos, sistemas baseados em regras nebulosas e teoria dos reticulados. Estabelecemos aqui uma analogia entre espaços vetoriais e famílias de conjuntos nebulosos, transformações lineares e sistemas de regras nebulosas. Essa analogia segue identificando as operações aritméticas usuais de soma e multiplicação com as operações correspondentes de um reticulado. Concluímos a primeira parte com um teorema que caracteriza sistemas de regras nebulosas que comutam com as operações do reticulado. A segunda parte apresenta uma aplicação de voz sobre IP (VoIP). Especificamente, elaboramos um sistema de regras nebulosas que avalia a qualidade na transmissão de um serviço VoIP considerando apenas a taxa de pacotes pedidos e o atraso durante a transmissão. Nosso modelo pode ser usado para efeitos de monitoramento e/ou controle do serviço de modo que cumpra alguns requisitos técnicos ou comerciais, tais como acordo de nível de serviço.

Palavras-chaves: Reticulados completos. Conjuntos nebulosos. Sistema de regras. Serviço VoIP.

CASTILHO, Alex Sandro de. **Caracterização de sistemas de regras nebulosas em reticulado completo e uma aplicação no problema de transmissão de voz sobre IP.** 2009. 63f. Dissertação (Mestrado em Matemática Aplicada e Computacional) – Universidade Estadual de Londrina, Londrina, 2009.

ABSTRACT

Fuzzy sets can be used to describe vague or uncertain concepts. Rule-based systems constitute a powerful tool to model phenomena by means of a natural language. This work is divided in two parts. The first refers to theoretical issues concerning fuzzy sets, rule-based systems, and lattice theory. We establish here analogies between vector spaces and classes of fuzzy sets, linear transforms and rule-based systems. The analogies follow by replacing the usual arithmetic operations of addition and multiplication by corresponding lattice operations. We conclude the first part by demonstrating a theorem that completely characterizes rule-based systems that commute with lattice operations. The second part provides an application on voice over IP (VoIP). Specifically, we develop a rule-based system that evaluates the quality of a VoIP service by considering only package loss and network delay. Our model can be used for monitoring and/or control purposes in order to meet some technical or commercial requirement such as service level agreements.

Keywords: Complete lattices. Fuzzy sets. Rule-based system. Service VoIP.

LISTA DE FIGURAS

Figura 3.1 – Função f que representa o conceito de pessoas altas por meio do conjunto clássico X	22
Figura 3.2 – Função x que representa o conceito de pessoa alta por meio de uma função exponencial.....	23
Figura 3.3 – Funções de pertinência que representam as classes de pessoas de estatura média x^e e pessoas altas x^a , onde as linhas tracejada e a linha contínua representam x^e e x^a , respectivamente	25
Figura 3.4 – Função de pertinência que representa a intersecção entre as classes de pessoas de estatura média x^e e pessoas altas x^a	26
Figura 3.5 – Função de pertinência que representa a união entre as classes de pessoas de estatura média x^e e pessoas altas x^a	26
Figura 4.1 – Conjuntos nebulosos usados para modelar a taxa de pacotes perdidos. As linhas tracejadas indicam o grau de compatibilidade de uma perda 0,6 com os conceitos perda aceitável e perda inaceitável.....	37
Figura 4.2 – Conjuntos nebulosos usados para modelar o <i>jitter</i> de atraso da rede. As linhas tracejadas indicam o grau de compatibilidade de um atraso de 0,2 com os conceitos atraso imperceptível e atraso aceitável.....	38
Figura 4.3 – Conjuntos nebulosos usados para modelar a qualidade do serviço VoIP	39
Figura 4.4 – Função de pertinência do conjunto nebuloso y deduzido pelo SRNs da Tabela 4.1 quando uma perda de 0,6 e um atraso de 0,2 são apresentados. A linha tracejada indica o centro de área	40
Figura B.1 – Modelo não-linear de um neurônio	56
Figura B.2 – Representação gráfica de um perceptron de múltiplas camadas.....	59
Figura B.3 – Ilustração da regra de parada antecipada utilizando o Erro Médio Quadrado (EMQ).....	61

LISTA DE TABELAS

Tabela 3.1 – Relação de aproximação entre os elementos dos universos de discurso U e V.....	27
Tabela 4.1 – Sistema de regras usado para avaliar a qualidade do serviço VoIP	39
Tabela B.1 – Valores de dados adquiridos através de uma variação do sistema QVoIP [14], os dados estão normalizados no intervalo [0; 1].....	63

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO	10
CAPÍTULO 2 – TEORIA DOS RETICULADOS E ANALOGIAS COM ESPAÇOS VETORIAIS	14
2.1 RETICULADOS.....	14
2.2 ANALOGIAS COM ESPAÇOS VETORIAIS.....	17
CAPÍTULO 3 – TEORIA DOS CONJUNTOS NEBULOSOS	21
3.1 CONJUNTOS NEBULOSOS E CONJUNTOS CLÁSSICOS	21
3.2 OPERAÇÕES COM CONJUNTOS NEBULOSOS	24
3.3 RELAÇÃO NEBULOSA	26
3.4 SISTEMAS DE REGRAS NEBULOSAS	28
3.5 REGRA COMPOSICIONAL DE INFERÊNCIA E OPERADORES SUP- \wedge	29
3.6 SRNS COM MÚLTIPLAS ENTRADAS E MÚLTIPLAS SAÍDAS	34
CAPÍTULO 4 – APLICAÇÃO: AVALIAÇÃO DA QUALIDADE DE UM SERVIÇO DE TRANSMISSÃO DE VOZ SOBRE IP	36
4.1 MODELAGEM DO PROBLEMA	37
4.2 EXPERIMENTOS COMPUTACIONAIS.....	41
CAPÍTULO 5 – CONCLUSÃO	43
REFERÊNCIAS	45
APÊNDICES	51
Apêndice A – Espaços Vetoriais e Operadores Lineares	52
Apêndice B – Redes Neurais Artificiais.....	54

CAPÍTULO 1

INTRODUÇÃO

Por volta dos anos 1960, o matemático Loft Zadeh introduziu o conceito de *conjunto nebuloso* (fuzzy set) no intuito de representar matematicamente conceitos vagos empregados no nosso cotidiano, como os conceitos de “pessoa jovem”, “temperatura confortável” e “erro pequeno” [8]. Surgiu, então, a teoria dos conjuntos nebulosos, onde cada conjunto nebuloso representa um conceito vago ou impreciso. Na teoria clássica, os conjuntos são de tal forma que um dado elemento de um universo de discurso pertence ou não pertence ao conjunto referido. Na teoria dos conjuntos nebulosos existe um grau de pertinência de cada elemento do universo de discurso à um conjunto determinado. Desse modo, podemos dizer que a teoria dos conjuntos nebulosos generaliza os conceitos da teoria clássica de conjuntos.

Lembramos que, a teoria dos conjuntos nebulosos não é uma teoria nebulosa ou vaga [42]. Em outras palavras, embora usada para descrever conceitos que não são claramente definidos, a teoria dos conjuntos nebulosos possui uma rigorosa base matemática [12, 27]. Especificamente, essa teoria pode ser muito bem conduzida numa estrutura algébrica chamada reticulado [40], as definições, operações e propriedades de um reticulado serão vistas com detalhes no Capítulo 2.

Atualmente, vários pesquisadores utilizam a teoria dos conjuntos nebulosos para resolver problemas em diversas áreas da ciência, como por exemplo problemas de dinâmica populacional [4, 22, 23], diagnóstico [11], controle [8, 28], otimização [1] e previsão de séries temporais [51, 53]. Uma das formas possíveis de se trabalhar com esses problemas é utilizar sistemas de regras nebulosas (SRNs). Um SRNs utiliza métodos de inferência para produzir saídas para cada entrada nebulosa [4, 42, 27]. O método de inferência utilizado nessa dissertação será o método de Mamdani. Optamos por este método, pois possui um baixo custo operacional, efetua apenas operações de máximos e mínimos e é de fácil compreensão. Além disso, o método de Mamdani efetua poucas operações de ponto flutuante (*flops*), o que nos facilitará a avaliação da qualidade de um serviço de transmissão de voz sobre IP (VoIP).

A tecnologia VoIP surgiu como uma alternativa para os sistemas de telecomunicações convencionais. Junto com essa tecnologia veio a necessidade de se

melhorar a qualidade na transmissão VoIP. No intuito de melhorar essa transmissão, foram criados métodos para se avaliar a qualidade do serviço VoIP.

Atualmente há duas maneiras de avaliar a qualidade do serviço de uma rede: utilizando métodos subjetivos ou métodos objetivos. Os métodos subjetivos surgiram para simular a opinião humana. O MOS (Mean Opinion Score) pode ser considerado um modelo subjetivo. O MOS é a média de notas atribuídas por avaliadores para a transmissão de voz realizada em uma chamada VoIP [14]. No intuito de simular esses avaliadores, podemos usar redes neurais artificiais e/ou conjuntos nebulosos para avaliar uma transmissão VoIP.

Por outro lado, os métodos objetivos são assim chamados, pois, não necessitam que várias pessoas avaliem o áudio transmitido durante uma chamada IP. Um exemplo de método objetivo é o Modelo E [31], que é um modelo que mede diretamente transmissões degenerativas e depois prediz que qualidade de voz irá resultar dessas transmissões. Os métodos objetivos são baseados em equações matemáticas que tentam modelar o comportamento da rede, utilizando parâmetros que são utilizados durante a transmissão do áudio até chegar ao usuário final [14].

Nesta dissertação vamos utilizar um modelo baseado em um SRNs que avalia a qualidade do serviço VoIP, nosso modelo pode ser considerado um método subjetivo, pois foi construído a partir do nosso conhecimento de como funciona o serviço VoIP.

OBJETIVOS

Nesta dissertação temos dois objetivos. Primeiramente, estabeleceremos uma analogia entre espaços vetoriais e transformações lineares com famílias de conjuntos nebulosos e sistemas de regras nebulosas. Essa analogia segue identificando as duas operações de um reticulado com as operações usuais de soma e multiplicação. Além disso, delinearemos o conceito de espaço reticulado. Um espaço reticulado é um conceito semelhante à noção de espaço vetorial, mas definido usando as operações de junção e reunião. Apontaremos definições análogas a noção de operador linear e observaremos que os principais métodos de inferência, incluindo o *método de Mamdani*, resultam em operadores “reticulado-lineares” [4, 27, 42]. Sobretudo, apresentaremos um teorema que caracteriza operadores “reticulado-lineares”. Em outras palavras, demonstraremos um teorema análogo ao clássico resultado que afirma que todo operador linear definido sobre espaços finitos corresponde a uma matriz, e vice-versa.

Além disso, utilizaremos um sistema de regras nebulosas para avaliar a qualidade de um serviço de transmissão de voz sobre IP (VoIP) com o objetivo de oferecer uma nova proposta de solução para esse problema. Nosso modelo fornece uma nota para a qualidade do serviço. Essa nota poderá ser usada para o monitoramento e/ou controle. É importante observar que o modelo utilizado pode ser considerado como um método subjetivo, pois os conjuntos nebulosos foram definidos a partir do nosso conhecimento do problema VoIP.

Salientamos que, pelo nosso conhecimento, esta dissertação é um trabalho inédito na literatura atual, pois nenhuma pesquisa feita até então utilizou um sistema de regras nebulosas que avalia a qualidade de uma transmissão de voz sobre IP usando apenas como variáveis de entrada o jitter de atraso e a perda de pacotes. Este fato reforça a nossa afirmação que esta dissertação oferece uma solução alternativa e de fácil compreensão para o problema de avaliar a qualidade na transmissão de voz sobre IP.

ORGANIZAÇÃO DO TRABALHO

Essa dissertação está organizada da seguinte forma. No Capítulo 2 introduziremos as definições e propriedades da teoria dos reticulados. Estabeleceremos também nesse capítulo uma analogia entre reticulados completos e espaços vetoriais.

No Capítulo 3 introduziremos o conceito de conjuntos nebulosos e suas operações. Deduziremos o método de inferência de Mamdani utilizando um sistema de regras nebulosas. Neste capítulo, mostraremos que esse método gera um conjunto nebuloso que pode ser representado por um operador sobre um reticulado.

Finalmente, no Capítulo 4, estudaremos a qualidade de um serviço de voz sobre IP usando um sistema de regras nebulosas. Compararemos os resultados do SRNs com os resultados obtidos por uma rede neural artificial. Terminaremos essa dissertação com uma conclusão sobre os resultados obtidos.

CAPÍTULO 2

TEORIA DOS RETICULADOS E ANALOGIAS COM ESPAÇOS VETORIAIS

A teoria dos reticulados¹ tem suas origens nos estudos de Dedekind sobre álgebra booleana, no final do século XIX [5]. Muitos outros matemáticos eminentes contribuíram para o desenvolvimento dessa teoria, incluindo Ore, Von Neumann, Tarski e, principalmente, Birkhoff [5, 32, 47]. O termo “reticulado”, embora, não seja muito conhecido pela comunidade brasileira de matemática aplicada, é uma teoria baseada apenas em conceitos matemáticos elementares. Sobretudo, a teoria dos reticulados possui um papel importante em diversas aplicações modernas. Por exemplo, reticulados completos formam a base para muitas técnicas de processamento e análise de imagens [3, 18, 35] e inteligência computacional, incluindo redes neurais artificiais e a teoria dos conjuntos nebulosos [12, 25, 24, 40].

2.1 RETICULADOS

A teoria dos reticulados tem como base uma relação de ordem \leq , com propriedades bem definidas que nos leva a definição de um conjunto parcialmente ordenado.

Definição 2.1.1 Um conjunto parcialmente ordenado é um conjunto não vazio P munido de uma relação de ordem \leq que satisfaz as seguintes propriedades para todo $x, y, z \in P$:

P1. Reflexiva: $x \leq x$;

P2. Anti-simétrica: $x \leq y$ e $y \leq x$ implica $x = y$;

P3. Transitividade: $x \leq y$ e $y \leq z$ implica $x \leq z$.

Dizemos que um conjunto P é totalmente ordenado (ou forma uma cadeia) se

P4. $x \leq y$ ou $y \leq x$, para todo $x, y \in P$.

¹ Tradução do termo inglês “lattice theory”.

Uma cota superior de um subconjunto X de um conjunto parcialmente ordenado P é um elemento $s \in P$, tal que $x \leq s$ para todo $x \in X$. O supremo de X é a menor das cotas superiores de X e será denotado por $\bigvee X$. De forma análoga, uma cota inferior de X é um elemento $r \in P$, tal que $r \leq x$ para todo $x \in X$. O ínfimo de X é a maior das cotas inferiores X e será denotado por $\bigwedge X$.

É importante observar que se $\bigvee X$ existe, ele é único. De fato, se s e s' são supremos de X , então $s \leq s'$ e $s' \leq s$, portanto por $P \mathcal{O}$ temos $s = s'$.

Definição 2.1.2 Um reticulado L é um conjunto parcialmente ordenado onde todo subconjunto finito possui supremo e ínfimo [5].

Nesse trabalho denotamos um reticulado por \mathbf{L} ou \mathbf{M} . Temos a seguinte definição quando um conjunto possui subconjuntos infinitos com supremo e ínfimo.

Definição 2.1.3 Um reticulado é dito completo se todo subconjunto, finito ou infinito, possui supremo e ínfimo.

O intervalo fechado $[0,1]$ é um exemplo de reticulado completo. O conjunto dos números reais \mathbb{R} representa um reticulado, mas não é um reticulado completo. De fato, se tomarmos o conjunto dos números naturais $\mathbb{N} \subseteq \mathbb{R}$, temos que \mathbb{N} é um subconjunto infinito de \mathbb{R} que não possui supremo. Portanto, todo reticulado completo é também um reticulado, mas nem todo reticulado é completo.

Em um reticulado as operações de supremo e ínfimo definem duas operações binárias, chamadas junção e reunião², e denotadas respectivamente pelos símbolos \vee e \wedge . Especificamente, definimos essas operações através das seguintes equações, para todo $x, y \in L$:

$$x \wedge y = \bigwedge \{x, y\} \quad \text{e} \quad x \vee y = \bigvee \{x, y\}. \quad (2.1)$$

Exemplo 1 Dado um conjunto X , o conjunto das partes de X , denotado por $\mathcal{P}(X)$, é um reticulado completo com as operações de reunião e junção dadas pela interseção e união de conjuntos, respectivamente. O ínfimo é o conjunto vazio \emptyset e o supremo é o próprio conjunto X , i.e., $\bigwedge \mathcal{P}(X) = \emptyset$ e $\bigvee \mathcal{P}(X) = X$.

² Tradução dos termos "join" e "meet", respectivamente [5].

Exemplo 2 Como visto anteriormente, o intervalo fechado $[0; 1]$ é um reticulado completo com as operações de junção e reunião definidas como o máximo e o mínimo, respectivamente. O supremo de $[0; 1]$ é 1 e o ínfimo é 0.

Note que as operações de junção e reunião satisfazem as seguintes propriedades para todo $x, y, z \in \mathbf{L}$:

$$1. \text{ Comutatividade: } x \wedge y = y \wedge x \text{ e } x \vee y = y \vee x. \quad (2.2)$$

$$2. \text{ Associatividade: } x \wedge (y \wedge z) = (x \wedge y) \wedge z \text{ e } x \vee (y \vee z) = (x \vee y) \vee z. \quad (2.3)$$

$$3. \text{ Absorção: } x \wedge (x \vee y) = x \vee (x \wedge y) = x. \quad (2.4)$$

$$4. \text{ Idempotência: } x \wedge x = x \text{ e } x \vee x = x. \quad (2.5)$$

A Proposição 2.1.1 a seguir apresenta um importante resultado sobre reticulados, pois mostra que todo conjunto munido de duas operações binárias que satisfaça as propriedades (2.2)-(2.5) é um reticulado. A demonstração dessa proposição pode ser vista em [5].

Proposição 2.1.1 *Um conjunto munido de duas operações binárias é um reticulado se, e somente se, as operações sobre esse conjunto satisfazem as propriedades (2.2)-(2.5).*

Na teoria dos reticulados existe um importante conceito afirmando que todo reticulado \mathbf{L} possui um reticulado dual \mathbf{L}' . O reticulado \mathbf{L}' é definido pela relação de ordem inversa \leq' sobre os mesmos elementos de \mathbf{L} , i.e., dados $x, y \in \mathbf{L}$ e $x \leq y$, a relação de ordem \leq' em \mathbf{L} entre x e y é dada por $y \leq' x$. As operações de junção e reunião do reticulado dual \mathbf{L}' correspondem às operações de reunião e junção do reticulado \mathbf{L} . Além disso, toda expressão em \mathbf{L} corresponde à uma sentença dual obtida substituindo operações de junção por reunião, e vice-versa. Este conceito é chamado princípio de dualidade [5, 18].

Princípio de Dualidade. *Se \mathbf{L} é um reticulado, então existe um reticulado dual, tal que, para toda definição, propriedade, operação, etc., referente a \mathbf{L} existe uma correspondência dual referente a \mathbf{L}' .*

Em muitos reticulados, além dos itens (2.2)-(2.5) serem satisfeitos, temos também a propriedade de distributividade sobre as operações de junção (\vee) e reunião (\wedge) como segue.

Definição 2.1.4 *Um reticulado \mathbf{L} é distributivo se as seguintes equações valem para todo $\alpha, x, y \in \mathbf{L}$:*

$$\alpha \wedge (x \vee y) = (\alpha \wedge x) \vee (\alpha \wedge y) \quad e \quad \alpha \vee (x \wedge y) = (\alpha \vee y) \wedge (\alpha \vee x). \quad (2.6)$$

Um reticulado completo \mathbf{L} é infinitamente distributivo se as seguintes equações valem para todo $\alpha \in \mathbf{L}$ e $X \subseteq \mathbf{L}$ [5, 18]:

$$\alpha \wedge \left(\bigvee X \right) = \bigvee_{x \in X} (\alpha \wedge x) \quad e \quad \alpha \vee \left(\bigwedge X \right) = \bigwedge_{x \in X} (\alpha \vee x). \quad (2.7)$$

Exemplo 3 O intervalo fechado $[0,1]$ é um exemplo de reticulado completo infinitamente distributivo. De fato, para todo $\alpha \in [0,1]$ e $X \subseteq [0,1]$ as duas eqs. em (2.7) são satisfeitas.

2.2 ANALOGIAS COM ESPAÇOS VETORIAIS

Como os reticulados são estruturas algébricas bem definidas, podemos delinear uma analogia entre reticulados e espaços vetoriais. Note que as duas eqs. em (2.6) são análogas a distributividade do produto com a soma, isto é, $\alpha(x + y) = \alpha x + \alpha y$, para todo $\alpha \in \mathbb{K}$ e $x, y \in \mathcal{V}$, onde \mathbb{K} é um corpo e \mathcal{V} é um espaço vetorial (cf. Apêndice A). Por outro lado, as duas eqs. em (2.7) podem ser identificadas com a equação $\alpha \left(\int_X dx \right) = \int_X \alpha dx$. Desse modo, podemos delinear uma analogia entre as operações de junção e reunião de reticulados com as operações de adição e multiplicação dos números reais. Tendo em vista o princípio de dualidade, podemos identificar sem perda de generalidade “ \wedge ” com “ \times ”, “ \vee ” com “ $+$ ” e, no caso de reticulados completos, $\int_X dx$ com $\bigvee X$.

Dessa analogia podemos estabelecer uma estrutura algébrica semelhante ao conceito de espaço vetorial, mas definido sobre um reticulado³ que será chamado espaço reticulado sup- \wedge .

Definição 2.2.1 Um espaço reticulado sup- \wedge é um conjunto não vazio \mathcal{L} definido sobre um reticulado distributivo \mathbb{L} onde a operação junção associa a cada par $x, y \in \mathcal{L}$ um elemento $x \vee y \in \mathcal{L}$ e a operação reunião associa a cada $x \in \mathcal{L}$ e $\alpha \in \mathbb{L}$ um elemento $\alpha \wedge x \in \mathcal{L}$.

Exemplo 4 Considere o reticulado completo $\mathbf{L} = [0,1]$. O hipercubo $\mathbf{L}^n = [0,1]^n$ é um exemplo de espaço reticulado análogo à \mathbb{R}^n . Por exemplo, dado um escalar $\alpha \in \mathbb{L}$ e vetores

³ Nesse trabalho, usaremos fontes em negrito para denotar elementos de um espaço vetorial e fontes em itálico para denotar escalares.

coluna $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{L}^n$ e $\mathbf{y} = [y_1, \dots, y_n]^T \in \mathbb{L}^n$, podemos definir $\mathbf{x} \vee \mathbf{y} = [x_1 \vee y_1, \dots, x_n \vee y_n]^T \in \mathbb{L}^n$ e $\alpha \wedge \mathbf{x} = [\alpha \wedge x_1, \dots, \alpha \wedge x_n]^T \in \mathbb{L}^n$. Neste caso, $\mathbf{x} \vee \mathbf{y}$ é simplesmente o máximo elemento-a-elemento de \mathbf{x} e \mathbf{y} . Por outro lado, o vetor $\alpha \wedge \mathbf{x}$ é obtido efetuando um corte de nível α no vetor \mathbf{x} .

Exemplo 5 Dado um conjunto U e o reticulado completo \mathbb{L} , o conjunto de todas as funções $\mathbf{x} : U \rightarrow \mathbb{L}$, denotado por $\mathcal{F}(U, \mathbb{L})$, é um exemplo de espaço reticulado análogo ao espaço das funções contínuas definidas sobre um conjunto compacto [29]. Por exemplo, dados $\alpha \in \mathbb{L}$ e funções $\mathbf{x}, \mathbf{y} \in \mathcal{F}(U, \mathbb{L})$, definimos $\mathbf{x} \vee \mathbf{y} \in \mathcal{F}(U, \mathbb{L})$ e $\alpha \wedge \mathbf{x} \in \mathcal{F}(U, \mathbb{L})$ como segue para todo $u \in U$:

$$(\mathbf{x} \vee \mathbf{y})(u) = \mathbf{x}(u) \vee \mathbf{y}(u) \quad \text{e} \quad (\alpha \wedge \mathbf{x})(u) = \alpha \wedge \mathbf{x}(u). \quad (2.8)$$

Observação 2.2.1 Dualmente, podemos definir o conceito de espaço reticulado *inf- \vee* . Note também que o conjunto de todos os elementos de \mathbb{L} forma um reticulado completo com as operações de supremo e ínfimo empregados nos espaços *sup- \wedge* e *inf- \vee* , respectivamente. Em particular, vale o seguinte lema.

Lema 2.2.1 Se \mathbb{L} é um reticulado, então o conjunto das funções $\mathcal{F}(U, \mathbb{L})$ é um reticulado. Além disso, se \mathbb{L} é um reticulado completo, então $\mathcal{F}(U, \mathbb{L})$ também é um reticulado completo.

Demonstração: Parte 1. Considere um reticulado \mathbb{L} . Mostraremos que $\mathcal{F}(U, \mathbb{L})$ é também um reticulado. Para isso, $\mathcal{F}(U, \mathbb{L})$ deve satisfazer as propriedades (2.2)-(2.5). Sejam $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{F}(U, \mathbb{L})$, temos que

1. Comutatividade:

$$(\mathbf{x} \wedge \mathbf{y})(u) = \mathbf{x}(u) \wedge \mathbf{y}(u) = \mathbf{y}(u) \wedge \mathbf{x}(u) = (\mathbf{y} \wedge \mathbf{x})(u), \quad \forall u \in U.$$

$$(\mathbf{x} \vee \mathbf{y})(u) = \mathbf{x}(u) \vee \mathbf{y}(u) = \mathbf{y}(u) \vee \mathbf{x}(u) = (\mathbf{y} \vee \mathbf{x})(u), \quad \forall u \in U.$$

2. Associatividade:

$$\begin{aligned} (\mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z}))(u) &= \mathbf{x}(u) \wedge (\mathbf{y} \wedge \mathbf{z})(u) \\ &= \mathbf{x}(u) \wedge \mathbf{y}(u) \wedge \mathbf{z}(u) \\ &= (\mathbf{x} \wedge \mathbf{y})(u) \wedge \mathbf{z}(u) \\ &= ((\mathbf{x} \wedge \mathbf{y}) \wedge \mathbf{z})(u), \quad \forall u \in U. \end{aligned}$$

$$\begin{aligned}
(\mathbf{x} \vee (\mathbf{y} \vee \mathbf{z}))(u) &= \mathbf{x}(u) \vee (\mathbf{y} \vee \mathbf{z})(u) \\
&= \mathbf{x}(u) \vee \mathbf{y}(u) \vee \mathbf{z}(u) \\
&= (\mathbf{x} \vee \mathbf{y})(u) \vee \mathbf{z}(u) \\
&= ((\mathbf{x} \vee \mathbf{y}) \vee \mathbf{z})(u), \quad \forall u \in U.
\end{aligned}$$

3. Absorção: Para todo $\mathbf{x}(u), \mathbf{y}(u) \in \mathbb{L}$, temos por hipótese que \mathbb{L} satisfaz a propriedade de absorção. De fato,

$$\begin{aligned}
(\mathbf{x} \wedge (\mathbf{x} \vee \mathbf{y}))(u) &= \mathbf{x}(u) \wedge (\mathbf{x} \vee \mathbf{y})(u) \\
&= \mathbf{x}(u) \wedge (\mathbf{x}(u) \vee \mathbf{y}(u)) \\
&= \mathbf{x}(u), \quad \forall u \in U,
\end{aligned}$$

$$\begin{aligned}
(\mathbf{x} \vee (\mathbf{x} \wedge \mathbf{y}))(u) &= \mathbf{x}(u) \vee (\mathbf{x} \wedge \mathbf{y})(u) \\
&= \mathbf{x}(u) \vee (\mathbf{x}(u) \wedge \mathbf{y}(u)) \\
&= \mathbf{x}(u), \quad \forall u \in U.
\end{aligned}$$

4. Idempotência:

$$(\mathbf{x} \wedge \mathbf{x})(u) = \mathbf{x}(u) \wedge \mathbf{x}(u) = \mathbf{x}(u), \quad \forall u \in U.$$

$$(\mathbf{x} \vee \mathbf{x})(u) = \mathbf{x}(u) \vee \mathbf{x}(u) = \mathbf{x}(u), \quad \forall u \in U.$$

Portanto, $\mathcal{F}(U, \mathbb{L})$ é um reticulado.

Parte 2. Agora, vamos mostrar que $\mathcal{F}(U, \mathbb{L})$ é um reticulado completo quando \mathbb{L} for um reticulado completo. Seja X um subconjunto, finito ou infinito, de $\mathcal{F}(U, \mathbb{L})$. Note que $\mathbf{x}(u) \in \mathbb{L}$ para todo $u \in U$, logo o conjunto $Y(u) = \{\mathbf{x}(u), \mathbf{x} \in X\}$ é um subconjunto de \mathbb{L} . Como \mathbb{L} é um reticulado completo $Y(u)$ possui um supremo $\mathbf{y}^1(u)$ e um ínfimo $\mathbf{y}^2(u)$. Assim,

$$\left(\bigvee X\right)(u) = \bigvee_{\mathbf{x} \in X} \mathbf{x}(u) = \mathbf{y}^1(u), \quad (2.9)$$

$$\left(\bigwedge X\right)(u) = \bigwedge_{\mathbf{x} \in X} \mathbf{x}(u) = \mathbf{y}^2(u). \quad (2.10)$$

Desse modo, o supremo de X é \mathbf{y}^1 e o ínfimo é \mathbf{y}^2 , i. e., $\bigvee X = \mathbf{y}^1$ e $\bigwedge X = \mathbf{y}^2$. Portanto $\mathcal{F}(U, \mathbb{L})$ é um reticulado completo.

Finalmente, podemos importar também da álgebra linear um conceito semelhante à noção de operador linear, mas sobre espaços reticulados.

Definição 2.2.2 *Sejam \mathcal{L} e \mathcal{M} espaços reticulados sobre \mathbb{L} . Dizemos que $T : \mathcal{L} \rightarrow \mathcal{M}$ é um operador $\text{sup-}\wedge$ se a seguinte equação valer para todo $\mathbf{x}, \mathbf{y} \in \mathcal{L}$ e $\alpha, \beta \in \mathbb{L}$:*

$$T[(\alpha \wedge \mathbf{x}) \vee (\beta \wedge \mathbf{y})] = [\alpha \wedge T(\mathbf{x})] \vee [\beta \wedge T(\mathbf{y})]. \quad (2.11)$$

Se \mathbb{L} for um reticulado completo infinitamente distributivo, então um operador $\text{sup-}\wedge$ deve satisfazer a seguinte equação para todo $\mathbf{X} \subseteq \mathcal{L}$ e $\alpha \in \mathbb{L}$:

$$T\left(\alpha \wedge \left(\bigvee \mathbf{X}\right)\right) = \alpha \wedge \left[\bigvee_{\mathbf{x} \in \mathbf{X}} T(\mathbf{x})\right]. \quad (2.12)$$

Exemplo 6 *Sejam U e V conjuntos quaisquer. O operador $T_R : \mathcal{F}(U, \mathbb{L}) \rightarrow \mathcal{F}(V, \mathbb{L})$ definido como segue onde $R : U \times V \rightarrow \mathbb{L}$ é um operador $\text{sup-}\wedge$.*

$$\mathbf{y}(v) = T_R(\mathbf{x})(v) = \bigvee_{u \in U} \left(R(u, v) \wedge \mathbf{x}(u) \right), \quad \forall v \in V. \quad (2.13)$$

De fato, seja $\mathbf{X} \subseteq \mathcal{F}(U, \mathbb{L})$ e $\alpha \in \mathbb{L}$. Assim, temos

$$\begin{aligned} T_R\left(\alpha \wedge \left(\bigvee \mathbf{X}\right)\right)(v) &= \bigvee_{u \in U} \left(R(u, v) \wedge \left(\alpha \wedge \left(\bigvee \mathbf{X}\right)\right)(u) \right) \\ &= \alpha \wedge \left[\bigvee_{u \in U} \left(R(u, v) \wedge \left(\bigvee \mathbf{X}\right)(u) \right) \right] \\ &= \alpha \wedge \left[\bigvee_{u \in U} \left(\bigvee_{\mathbf{x} \in \mathbf{X}} (R(u, v) \wedge \mathbf{x}(u)) \right) \right] \\ &= \alpha \wedge \left[\bigvee_{\mathbf{x} \in \mathbf{X}} \left(\bigvee_{u \in U} (R(u, v) \wedge \mathbf{x}(u)) \right) \right] \\ &= \alpha \wedge \left[\bigvee_{\mathbf{x} \in \mathbf{X}} T_R(\mathbf{x})(v) \right], \end{aligned}$$

para todo $v \in V$. Portanto, T_R é um operador $\text{sup-}\wedge$ pois satisfaz a eq. (2.12).

Nos próximos capítulos mostraremos que os principais sistemas baseados em regras nebulosas representam operadores $\text{sup-}\wedge$. Gostaríamos de observar, entretanto, que podemos definir de forma análoga o conceito de operador $\text{inf-}\vee$ substituindo “ \vee ” por “ \wedge ”, e vice-versa, nas eqs. (2.12) e (2.11).

CAPÍTULO 3

TEORIA DOS CONJUNTOS NEBULOSOS

A *Teoria dos Conjuntos Nebulosos*⁴ (ou Conjuntos Fuzzy) foi introduzida em 1965 pelo matemático Loft Zadeh no intuito de representar matematicamente conceitos vagos empregados frequentemente na linguagem cotidiana [54, 56]. Conjuntos nebulosos foram usados com sucesso para resolver diversos problemas com aplicações científicas e/ou na área de engenharia [4, 28, 30, 42]. Sobretudo, a teoria dos conjuntos nebulosos oferece uma ferramenta eficiente para a manipulação de termos expressos em palavras [57]. Nesse trabalho, usaremos conjuntos nebulosos para modelar e manipular conceitos vagos como a noção de “serviço ruim”, “perda imperceptível” e “atraso inaceitável” [8].

Neste capítulo, apresentamos os conceitos básicos da teoria dos conjuntos nebulosos e sistemas de regras nebulosas. Mostraremos também, que os principais métodos de inferência representam operadores “reticulado-lineares” e um teorema que caracteriza operadores “reticulado-lineares”.

3.1 CONJUNTOS NEBULOSOS E CONJUNTOS CLÁSSICOS

A teoria dos conjuntos nebulosos é utilizada para representar matematicamente conceitos vagos como, por exemplo, o conceito de “pessoa alta” ou “pessoa jovem”. Esses conceitos, apesar de utilizados com frequência no nosso cotidiano, são imprecisos no seguinte sentido: considere a sentença “João é alto”. Duas abordagens podem ser empregadas para descrever matematicamente o conceito de pessoa alta. A primeira, usando a teoria clássica de conjuntos, estabelece um valor para que uma pessoa possa ser considerada “alta”. Por exemplo, podemos supor que uma pessoa é considerada alta a partir de 1,80 metros. Neste caso, uma pessoa com 1,79 metros não é considerada alta. Podemos então nos perguntar se esse conceito está sendo modelado corretamente.

⁴ Tradução do termo inglês “*fuzzy sets*”

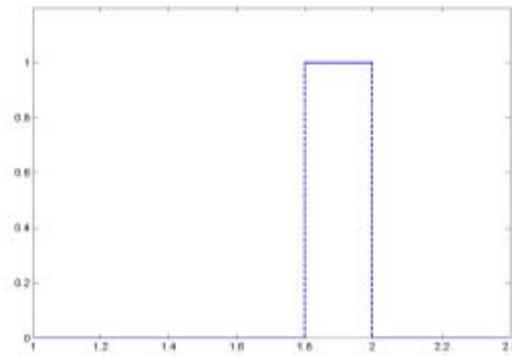


Figura 3.1 – Função f que representa o conceito de pessoas altas por meio do conjunto clássico X .

A segunda abordagem é baseada na teoria de conjuntos nebulosos. Aqui atribuímos a cada pessoa um grau de intensidade que representa o quanto a pessoa pertence à classe de pessoas altas. Esta segunda abordagem é a que será tratada neste capítulo.

Quando trabalhamos com conjuntos clássicos (ou crisp) as pessoas são consideradas altas ou não, i.e., as pessoas pertencem a classe de pessoas altas ou não. Por exemplo, seja U o conjunto de todas as alturas possíveis. O conjunto U é chamado universo de discurso. O conjunto das “pessoas altas” pode ser representado por um subconjunto $X \subset U$. Vamos assumir que $X = \{x \in U \mid 1,80 \leq u \leq 2,00\}$. Note que uma pessoa com altura x é alta se e somente se $x \in X$. Além disso, observe que o conjunto X também pode ser representado pela função $\mathbf{x} : U \rightarrow \{0,1\}$ definida por

$$\mathbf{x}(u) = \begin{cases} 1 & \text{se } 1,80 \leq u \leq 2,00 \\ 0 & \text{caso contrário,} \end{cases} \quad (3.1)$$

para todo $u \in U$. A Figura 3.1 representa o gráfico de \mathbf{x} . Note que $u \in X$ se, e somente se, $\mathbf{x}(u) = 1$. Caso contrário, u não pertence a X .

Por outro lado, o conceito de pessoa alta pode ser representado por um conjunto nebuloso como segue [27, 4]. Seja $\mathbf{x} : U \rightarrow [0,1]$ a função que associa a cada altura u um valor no intervalo $[0,1]$.

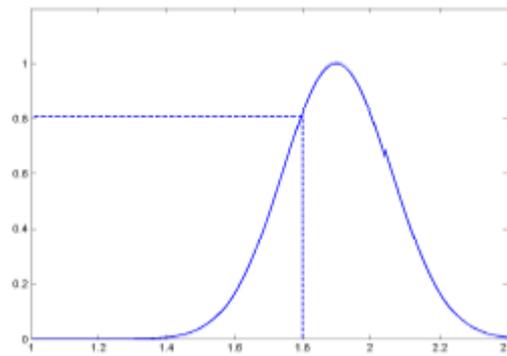


Figura 3.2 – Função x que representa o conceito de pessoa alta por meio de uma função exponencial.

A função x pode ser definida, por exemplo, como

$$x(u) = \exp[-20(u - 1,90)^2], \quad \forall u \in U. \quad (3.2)$$

O gráfico de x está apresentado na Figura 3.2. Note que x assume o valor 0,81 para $u = 1,80$. Isso significa que uma pessoa com 1,80 m de altura é 81% compatível com o conceito de pessoa alta. De um modo similar, uma pessoa com 1,79 m de altura é 78% compatível com a noção de pessoa alta. Portanto, pessoas com alturas próximas foram classificadas de forma semelhante ao conceito de pessoa alta. Esse simples fato nos permite descrever problemas complexos usando poucos conceitos ou conjuntos nebulosos [42, 57].

Definição 3.1.1 Um conjunto nebuloso, definido sobre um universo de discurso U , é caracterizado por uma função de pertinência que atribui um valor no intervalo $[0,1]$ para cada elemento $u \in U$, isto é,

$$x : U \rightarrow [0,1]. \quad (3.3)$$

Em outras palavras, o valor $x(u)$ representa o grau de compatibilidade de $u \in U$ com o conceito representado por x , i.e., $x(u)$ representa o quanto u é x .

Denotaremos o conjunto de todos os conjuntos nebulosos em U por $\mathcal{F}(U, [0,1])$ ou para facilitar a notação, por $\mathcal{F}(U)$. Note que $\mathcal{F}(U)$ é um caso particular de $\mathcal{F}(U, \mathbf{L})$ no Exemplo 5 do Capítulo 2.

3.2 OPERAÇÕES COM CONJUNTOS NEBULOSOS

Muitas operações com conjuntos nebulosos são efetuadas usando o conceito de conjunção nebulosa e disjunção nebulosa. Conjunções e disjunções nebulosas generalizam os conceitos clássicos de conjunção e disjunção, isto é, os operadores “e” e “ou” da lógica booleana [27, 42, 54].

Definição 3.2.1 *Uma conjunção nebulosa é uma operação binária $C_{\mathcal{F}} : [0, 1] \times [0, 1] \rightarrow [0, 1]$ crescente em ambos os argumentos que satisfaz as equações*

1. $C_{\mathcal{F}}(1, 1) = 1$, $C_{\mathcal{F}}(1, 0) = 0$, $C_{\mathcal{F}}(0, 1) = 0$ e $C_{\mathcal{F}}(0, 0) = 0$ (condição de fronteira);

2. Se $x \leq x'$ e $y \leq y'$, então $C_{\mathcal{F}}(x, y) \leq C_{\mathcal{F}}(x', y')$ (monotonicidade),

Para todo $x, x', y, y' \in [0, 1]$ [50, 53].

Definição 3.2.2 *Uma disjunção nebulosa é uma operação binária $D_{\mathcal{F}} : [0, 1] \times [0, 1] \rightarrow [0, 1]$ crescente em ambos os argumentos que satisfaz as condições*

1. $D_{\mathcal{F}}(0, 0) = 0$, $D_{\mathcal{F}}(1, 0) = 1$, $D_{\mathcal{F}}(0, 1) = 1$ e $D_{\mathcal{F}}(1, 1) = 1$ (condição de fronteira);

2. Se $x \leq x'$ e $y \leq y'$, então $D_{\mathcal{F}}(x, y) \leq D_{\mathcal{F}}(x', y')$ (monotonicidade),

para todo $x, y, x', y' \in [0, 1]$.

As operações de máximo e mínimo são exemplos de disjunção e conjunção nebulosas, respectivamente. Outros exemplos de disjunção e conjunção podem ser encontrados em [4, 40]. É importante observar, entretanto que o máximo e o mínimo são os operadores lógicos mais empregados na literatura. De fato, muitos livros definem apenas esses dois operadores na lógica nebulosa.

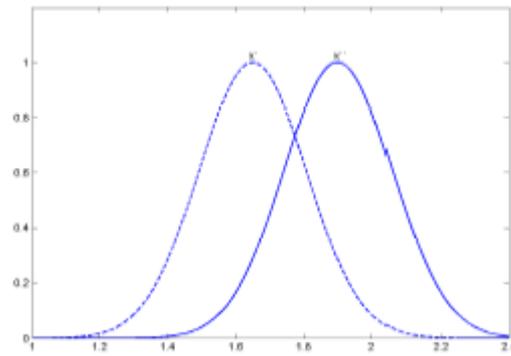


Figura 3.3 – Funções de pertinência que representam as classes de pessoas de estatura média x' e pessoas altas x'' , onde as linhas tracejada e a linha contínua representam x' e x'' , respectivamente.

Disjunção e conjunção nebulosas podem ser usadas para definir a união e a interseção de conjuntos nebulosos. Explicitamente, a união de conjuntos nebulosos $x, y \in \mathcal{F}(U)$ é definida como

$$(x \cup y)(u) = D_{\mathcal{F}}(x(u), y(u)), \quad \forall u \in U. \quad (3.4)$$

Analogamente, a interseção dos conjuntos nebulosos $x, y \in \mathcal{F}(U)$ é definida como

$$(x \cap y)(u) = C_{\mathcal{F}}(x(u), y(u)), \quad \forall u \in U. \quad (3.5)$$

Exemplo 7 Considere os conceitos de “pessoas de estatura média” e “pessoa alta”. Estes dois conceitos podem ser representados pelos conjuntos nebulosos x' e x'' apresentados na Figura 3.3. A união e a interseção nebulosas de x' e x'' podem ser calculados usando o mínimo e o máximo. O resultado é apresentado nas Figuras 3.4 e 3.5, respectivamente.

Nesse trabalho empregaremos apenas o mínimo e o máximo como conjunção e disjunção nebulosa. Antes de prosseguir, entretanto, gostaríamos de apresentar algumas observações sobre essas duas operações. Do ponto de vista matemático, a classe de todos os conjuntos nebulosos, representada pelo conjunto $\mathcal{F}(U)$, com as operações de máximo e mínimo, aplicadas pontualmente, representa um reticulado completo. De fato, todo subconjunto finito ou infinito de $\mathcal{F}(U)$ possui supremo e ínfimo, como visto no Capítulo 2.

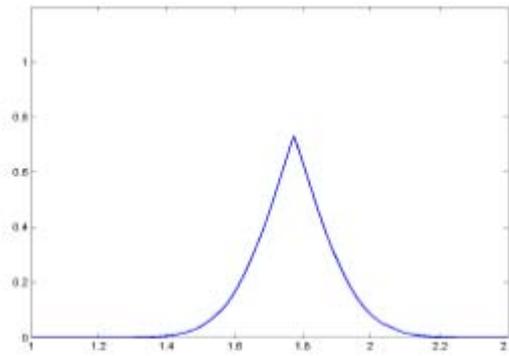


Figura 3.4 – Função de pertinência que representa a intersecção entre as classes de pessoas de estatura média x^e e pessoas altas x^h .

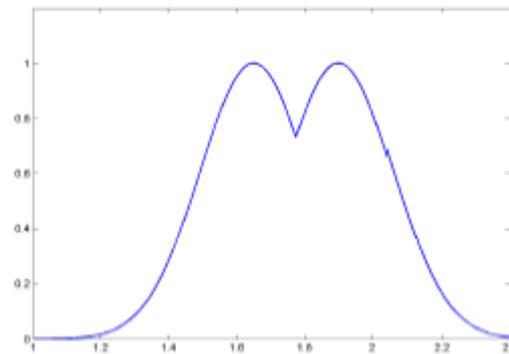


Figura 3.5 – Função de pertinência que representa a união entre as classes de pessoas de estatura média x^e e pessoas altas x^h .

Do ponto de vista computacional, as operações de máximo e mínimo não geram erros de arredondamento e podem ser efetuadas de forma rápida e eficiente [45].

3.3 RELAÇÃO NEBULOSA

Antes de falarmos de relações nebulosas, vamos fazer uma pequena introdução sobre o que são relações no sentido clássico. Relações estabelecem associações, comparações, semelhanças, identidades, etc. entre objetos ou elementos de conjuntos não necessariamente iguais. Em termos matemáticos, uma relação clássica corresponde a um conjunto de pares ordenados (u, v) onde $u \in U$ e $v \in V$ [42].

Sejam U e V dois universos de discursos, vamos definir o produto cartesiano de U e V por $U \times V$, que é o conjunto de todos os pares (u, v) tais que $u \in U$ e $v \in V$. Uma relação R , definida sobre o cartesiano $U \times V$ também pode ser vista como uma aplicação

$$R: U \times V \rightarrow \{0, 1\}, \quad (3.6)$$

onde

$$R(u, v) = \begin{cases} 1 & \text{se } (u, v) \in R \\ 0 & \text{se } (u, v) \notin R \end{cases}. \quad (3.7)$$

Relações nebulosas generalizam o conceito clássico de relações, pois admitem a noção de associação parcial. Vejamos a seguinte definição.

Definição 3.3.1 *Dados os universos de discurso U e V . Uma relação nebulosa R é um subconjunto do produto cartesiano $U \times V$ [42]. Equivalentemente, uma relação nebulosa sobre $U \times V$ é uma aplicação*

$$R: U \times V \rightarrow [0, 1]. \quad (3.8)$$

Dizemos que dois elementos u e v estão totalmente relacionados quando $R(u, v) = 1$. Por outro lado, se $R(u, v) = 0$, então os elementos não estão relacionados. Finalmente, valores entre $0 < R(u, v) < 1$ representam um grau de relacionamento parcial.

Observação 3.3.1 Uma relação clássica indica se existe ou não um relacionamento entre dois elementos. Uma relação nebulosa, além de dizer se existe ou não um relacionamento entre dois elementos, indica também qual o grau desse relacionamento.

Tabela 3.1 – Relação de aproximação entre os elementos dos universos de discurso U e V .

R	1,60	1,90	2,00
1,50	0,9	0,6	0,5
1,60	1	0,7	0,6
1,80	0,8	0,9	0,8

Observe ainda, que relações nebulosas são conjuntos nebulosos formados sobre o produto cartesiano dos universos de discursos U e V . As operações básicas sobre

relações nebulosas são a união e a interseção, que conceitualmente seguem as operações correspondentes sobre conjuntos nebulosos [42]. A família de todas as relações nebulosas sobre $U \times V$ é denotada por $\mathcal{F}(U \times V)$.

Exemplo 8 *Por simplicidade considere os universos de discurso $U = \{1, 60; 1, 90; 1, 95\}$ e $V = \{1, 50; 1, 60; 1, 80\}$. Definimos uma relação R de proximidade, onde R representa o conceito “aproximado”, tal que R verifica o quanto os elementos de U e V estão próximos. Esta relação pode ser representada pela Tabela 3.1.*

Note que a relação nebulosa $R(u, v)$, onde $u \in U$ e $v \in V$, recebe valores com grau de pertinência no intervalo $[0, 1]$.

3.4 SISTEMAS DE REGRAS NEBULOSAS

Uma regra nebulosa é qualquer sentença da forma *se-então* onde os antecedentes e/ou consequentes são conjuntos nebulosos [4, 27, 40, 42]. Várias regras nebulosas formam um *Sistema de Regras Nebulosas* (SRNs). Matematicamente, podemos representar um SRNs como segue onde $x^\xi \in \mathcal{F}(U)$ e $y^\xi \in \mathcal{F}(V)$ são conjuntos nebulosos para todo $\xi = 1, \dots, k$.

$$R^\xi : \text{Se } u \text{ é } x^\xi \text{ então } v \text{ é } y^\xi, \quad \text{para todo } \xi = 1, \dots, k. \quad (3.9)$$

SRNs são usados para modelar problemas que podem ser descritos usando linguagem natural. Aplicações de SRNs inclui avaliação da qualidade de um serviço de transmissão de voz sobre IP, modelagem da evolução de soropositivos para HIV numa determinada população e diagnóstico de câncer de próstata [4, 11, 8, 22, 23]. Apresentamos abaixo um exemplo simples de SRNs [42].

Exemplo 9 *Vamos supor que uma sala possua um aparelho de ar condicionado. Podemos representar a regulação automática do aparelho em relação à temperatura ambiente através de um SRN com regras do tipo*

$$\text{Se a temperatura é alta, então a potência é máxima,} \quad (3.10)$$

onde os universos de discurso U e V representam a temperatura ambiente e a potência do aparelho de ar-condicionado, respectivamente. Na regra (3.10), o conjunto nebuloso do

antecedente representa o conceito de “temperatura alta” enquanto que o conjunto nebuloso do conseqüente corresponde à noção “potência máxima”.

É importante observar que regras nebulosas podem ser formuladas de forma simples usando nossa concepção de como o sistema (ar-condicionado) deve funcionar. O SRNs é usado posteriormente para inferir a potência apropriada para uma dada temperatura ambiente.

Dado um SRNs e um conjunto nebuloso $\mathbf{x} \in \mathcal{F}(U)$ (e.g. a temperatura ambiente), a forma mais usada para deduzir um conjunto nebuloso $\mathbf{y} \in \mathcal{F}(V)$ (e.g. a potência do ar-condicionado) consiste em efetuar os seguintes passos [4, 27]:

1. Calcular o maior grau de interseção de \mathbf{x} com os antecedentes \mathbf{x}^ξ das regras nebulosas R^ξ . Em termos matemáticos, calculamos:

$$r^\xi = \bigvee_{u \in U} (\mathbf{x}^\xi(u) \wedge \mathbf{x}(u)), \quad \forall \xi = 1, \dots, k. \quad (3.11)$$

O valor r^ξ mede o grau de compatibilidade dos conjuntos nebulosos \mathbf{x} e \mathbf{x}^ξ .

2. Definir o conjunto nebuloso \mathbf{y} como a união dos conseqüentes \mathbf{y}^ξ truncados nos respectivos valores r^ξ , para $\xi = 1, \dots, k$. Especificamente, definimos

$$\mathbf{y}(v) = \bigvee_{\xi=1}^k (r^\xi \wedge \mathbf{y}^\xi(v)), \quad \forall v \in V. \quad (3.12)$$

A estratégia descrita pelos itens 1 e 2 acima é conhecida como *método de inferência de Mamdani* [4, 27, 42].

3.5 REGRA COMPOSICIONAL DE INFERÊNCIA E OPERADORES SUP- \wedge

O método de Mamdani representa um caso particular de *Regra Composicional de Inferência* (RCI), um conceito introduzido nos anos 1970 por Zadeh [55]. Uma RCI define o conjunto nebuloso deduzido \mathbf{y} através de uma composição do conjunto

nebuloso $\mathbf{x} \in \mathcal{F}(U)$ com uma relação nebulosa $R \in \mathcal{F}(U \times V)$ que captura o comportamento do SRNs. Formalmente, $\mathbf{y} \in \mathcal{F}(V)$ é dado pela equação

$$\mathbf{y}(v) = T_R(\mathbf{x})(v) = \bigvee_{u \in U} (R(u, v) \wedge \mathbf{x}(u)), \quad \forall v \in V. \quad (3.13)$$

A relação nebulosa $R \in \mathcal{F}(U \times V)$ pode ser definida de diversas formas [27, 42]. Por exemplo, no método de Mamdani definimos $R \in \mathcal{F}(U \times V)$ como segue:

$$R(u, v) = \bigvee_{\xi=1}^k (\mathbf{x}^\xi(u) \wedge \mathbf{y}^\xi(v)), \quad \text{para todo } u \in U \text{ e } v \in V. \quad (3.14)$$

Observe que o método de inferência de Mamdani apresenta como saída um conjunto nebuloso \mathbf{y} em V . Podemos utilizar um método de defuzzificação para determinar um valor $v^* \in V$ que representa o conjunto nebuloso \mathbf{y} [42]. Neste trabalho, adotamos o centro de área como método de defuzzificação. O centro de área define v^* através da seguinte equação onde $\{v_1, \dots, v_n\}$ é uma partição do universo de discurso V :

$$v^* = \frac{\sum_{j=1}^n \mathbf{y}(v_j) v_j}{\sum_{j=1}^n \mathbf{y}(v_j)}. \quad (3.15)$$

No Exemplo 9, o método de defuzzificação dado pela eq. (3.15) fornece um valor indicando a potência que o aparelho de ar-condicionado deve usar para uma dada temperatura ambiente.

Observação 3.5.1 *Um SRNs com o método de inferência de Mamdani e o centro de área como técnica de defuzzificação (e.g. Exemplo 9) representa um aproximador universal, i.e., esse processo pode ser usado para aproximar uniformemente qualquer função contínua definida em um compacto [2, 9].*

Note também que a eq. (3.13) define um operador $T_R : \mathcal{F}(U) \rightarrow \mathcal{F}(V)$ análogo ao conceito de operador integral, mas definido sobre um espaço reticulado. Lembramos que um operador integral $T_K : \mathcal{C}(U) \rightarrow \mathcal{C}(V)$ é dado pela seguinte equação onde $\mathcal{C}(U)$ e $\mathcal{C}(V)$ denotam o conjunto de todas as funções reais contínuas sobre conjuntos compactos U e V respectivamente [29]:

$$\mathbf{y}(v) = T_K(\mathbf{x})(v) = \int_U K(u, v) \mathbf{x}(u) du, \quad \forall v \in V. \quad (3.16)$$

Operadores integrais são exemplos de operadores lineares sobre o espaço das funções reais contínuas (cf. Apêndice A). Portanto, podemos nos perguntar se T_R representa um operador reticulado. O seguinte teorema mostra que T_R dado pela eq. (3.13) é um operador sup- \wedge . Em outras palavras, o método de inferência de Mamdani ou, de modo mais geral, a regra composicional de inferência, representa um operador sup- \wedge . O Teorema 3.5.1 mostra também que a recíproca é válida, i.e., todo operador sup- \wedge é descrito pela eq. (3.13).

Teorema 3.5.1 *Sejam U e V universos de discurso. Um operador $T_R : \mathcal{F}(U) \rightarrow \mathcal{F}(V)$ é um operador sup- \wedge se, e somente se, T_R é dado pela eq. (3.13) para alguma relação nebulosa $R \in \mathcal{F}(U \times V)$.*

Demonstração: Primeiramente vamos supor que T_R seja um operador sup- \wedge .

Seja $\mathbf{I}_u : U \rightarrow \{0, 1\}$ a função indicadora definida como segue: $\mathbf{I}_u(\mu) = 1$ se $\mu = u$ e $\mathbf{I}_u(\mu) = 0$ caso contrário. Note que podemos representar $\mathbf{x} \in \mathcal{F}(U)$ como segue usando funções indicadoras:

$$\mathbf{x} = \bigvee_{u \in U} (\mathbf{x}(u) \wedge \mathbf{I}_u).$$

Como T_R é um operador sup- \wedge , temos

$$\begin{aligned} T_R(\mathbf{x})(v) &= T_R \left(\bigvee_{u \in U} (\mathbf{x}(u) \wedge \mathbf{I}_u) \right) (v) \\ &= \bigvee_{u \in U} T_R (\mathbf{x}(u) \wedge \mathbf{I}_u) (v) \\ &= \bigvee_{u \in U} (\mathbf{x}(u) \wedge (T_R(\mathbf{I}_u)(v))), \end{aligned}$$

para todo $v \in V$. Logo T_R satisfaz a eq. (3.13) onde a relação $R \in \mathcal{F}(U \times V)$ é dada por $R(u, v) = T_R(\mathbf{I}_u)(v)$ para todo $u \in U$ e $v \in V$.

Por outro lado, suponha que T_R seja dado pela eq. (3.13), mostraremos que T_R é um operador sup- \wedge . Para isso, seja $\mathbf{X} \subseteq \mathcal{F}(U)$ e $\alpha \in [0, 1]$. Assim, temos

$$\begin{aligned}
T_R(\alpha \wedge (\bigvee \mathbf{X}))(v) &= \bigvee_{u \in U} (R(u, v) \wedge (\alpha \wedge (\bigvee \mathbf{X}))(u)) \\
&= \alpha \wedge \left[\bigvee_{u \in U} (R(u, v) \wedge (\bigvee \mathbf{X})(u)) \right] \\
&= \alpha \wedge \left[\bigvee_{u \in U} \left(\bigvee_{x \in \mathbf{X}} (R(u, v) \wedge x(u)) \right) \right] \\
&= \alpha \wedge \left[\bigvee_{x \in \mathbf{X}} \left(\bigvee_{u \in U} (R(u, v) \wedge x(u)) \right) \right] \\
&= \alpha \wedge \left[\bigvee_{x \in \mathbf{X}} T_R(x)(v) \right],
\end{aligned}$$

para todo $v \in V$. Portanto, T_R é um operador sup- \wedge pois satisfaz a eq. (2.11).

Gostaríamos de observar que muitos livros e artigos sobre a teoria dos conjuntos nebulosos e sobre lógica nebulosa apresentam a eq. (3.13) como a forma geral de se representar um SRNs. O argumento usado frequentemente é o seguinte: Como SRNs capturam relações entre as variáveis $u \in U$ e $v \in V$, esses podem então ser representados de forma geral usando uma relação nebulosa. Todavia, o Teorema 3.5.1 mostra que a eq. (3.13) descreve apenas um subconjunto de todos os operadores entre as famílias de conjuntos nebulosos $\mathcal{F}(U)$ e $\mathcal{F}(V)$, i.e., a classe dos operadores sup- \wedge .

Exemplo 10 *O método de inferência Soma- \wedge introduzido por Kosko constitui um exemplo de operador entre conjuntos nebulosos que não pode ser descrito pela eq. (3.13) [28, 42].*

Vejamos, dado um SRNs (cf. Exemplo 9) e um conjunto nebuloso x (e.g. a temperatura ambiente). O método de inferência soma- \wedge deduz um conjunto nebuloso y (e.g. a potência do ar-condicionado) da seguinte forma:

1. Para cada ξ calculamos a relação entre o conjunto $y^\xi(v)$ e o grau de intersecção de x e x^ξ como segue

$$y_s^\xi(v) = \bigvee_{u \in U} \left((x(u) \wedge x^\xi(u)) \wedge y^\xi(v) \right), \quad (3.17)$$

onde $v \in V$.

2. Deduzimos o conjunto nebuloso $y(v)$ como a soma de todos $y_s^\xi(v)$ com um peso w^ξ , isto é,

$$y(v) = \sum_{\xi=1}^k w^\xi y_s^\xi(v), \quad \forall v \in V, \quad (3.18)$$

onde w^ξ reflete a relevância de cada regra R^ξ .

Note que, o método soma- \wedge processa paralelamente cada regra, e como o conjunto nebuloso inferido y é a soma da contribuição individual de cada regra R^k , podemos obter um valor para $y(v)$ maior que 1 para algum $v \in V$. Para evitar esse problema, geralmente adotamos w^k , de forma que $\sum_{k=1}^k w^k = 1$. Se todas as regras R^k tem a mesma contribuição para o sistema, então definimos $w^k = 1/k$ [28].

Esse tipo de método é chamado de método de inferência aditiva, pois utilizamos a soma no lugar do supremo para deduzir o conjunto nebuloso y [42]. Este método de inferência não pode ser representado pelo operador da eq. (3.13), pois a soma não comuta com o supremo e o mínimo. Vejamos um contra-exemplo:

Vamos supor que $U = \{u\}$ e $V = \{v\}$ g sejam dois universos de discurso com um único elemento, e sejam $x^1(u) = 0,25$, $x^2(u) = 0,75$, $y^1(v) = 0,30$ e $y^2(v) = 0,60$ conjuntos nebulosos em U e V . Construímos o SRNs com duas regras R^1 e R^2 , como segue

$$R^1 : \text{ Se } u \text{ é } x^1, \text{ então } v \text{ é } y^1, \quad (3.19)$$

$$R^2 : \text{ Se } u \text{ é } x^2, \text{ então } v \text{ é } y^2, \quad (3.20)$$

Mostraremos que a seguinte propriedade é violada se adotarmos o método de inferência de Kosko.

$$T(\alpha \wedge (\bigvee X)) = \alpha \wedge \left[\bigvee_{x \in X} T(x) \right]. \quad (3.21)$$

Para isso, considere o conjunto nebuloso $x(u) = 0,40$ e a constante $\alpha = 0,35$. Dado um valor de entrada $\alpha \wedge x$ aplicado ao operador de Kosko temos

$$\begin{aligned} \sum_{k=1}^2 w^k \left[([\alpha \wedge x(u)] \wedge x^k(u)) \wedge y^k(v) \right] &= \sum_{k=1}^2 \frac{1}{2} \left[([\alpha \wedge x(u)] \wedge x^k(u)) \wedge y^k(v) \right] \\ &= \frac{1}{2} \left(([0,35 \wedge 0,40] \wedge 0,25) \wedge 0,30 \right) \\ &+ \frac{1}{2} \left(([0,35 \wedge 0,40] \wedge 0,75) \wedge 0,60 \right) \\ &= 0,125 + 0,175 = 0,300. \end{aligned}$$

Por outro lado, temos que

$$\begin{aligned}
\alpha \wedge \left(\sum_{\xi=1}^2 w^{\xi} \left((x(u) \wedge x^{\xi}(u)) \wedge y^{\xi}(v) \right) \right) &= \alpha \wedge \left(\sum_{\xi=1}^2 \frac{1}{2} \left((x(u) \wedge x^{\xi}(u)) \wedge y^{\xi}(v) \right) \right) \\
&= 0,35 \wedge \left[\frac{1}{2} \left((0,40 \wedge 0,25) \wedge 0,30 \right) \right. \\
&\quad \left. + \frac{1}{2} \left((0,40 \wedge 0,75) \wedge 0,60 \right) \right] \\
&= 0,35 \wedge (0,125 + 0,20) \\
&= 0,35 \wedge 0,325 = 0,325.
\end{aligned}$$

Portanto,

$$\sum_{\xi=1}^2 w^{\xi} \left(([\alpha \wedge x(u)] \wedge x^{\xi}(u)) \wedge y^{\xi}(v) \right) \neq \alpha \wedge \left(\sum_{\xi=1}^2 w^{\xi} \left((x(u) \wedge x^{\xi}(u)) \wedge y^{\xi}(v) \right) \right),$$

o que contraria a definição de operador sup- \wedge .

3.6 SRNs COM MÚTIPLAS ENTRADAS E MÚTIPLAS SAÍDAS

Em muitos problemas utilizamos SRNs onde cada regra possui múltiplas entradas e múltiplas saídas. Vamos representar o número de entradas por n e o número de saídas por m . O caso mais simples de regras nebulosas acontece quando temos $n = m = 1$, visto em (3.9). Um outro caso muito comum é quando $n = 2$ e $m = 1$, que será utilizado para avaliar a qualidade do serviço VoIP, como será visto no Capítulo 4. Nesta seção mostraremos que um SRNs com múltiplas entradas e múltiplas saídas pode ser representado por um sistema que possui uma única entrada e uma única saída [41, 42].

Um SRNs com múltiplas entradas e múltiplas saídas pode ser representado da seguinte forma

$$R^{\xi} : \text{ Se } u_1 \text{ é } x_1^{\xi} \text{ e } u_2 \text{ é } x_2^{\xi} \text{ e } \dots \text{ e } u_n \text{ é } x_n^{\xi}, \text{ então } v_1 \text{ é } y_1^{\xi} \text{ e } v_2 \text{ é } y_2^{\xi} \text{ e } \dots \text{ e } v_m \text{ é } y_m^{\xi}, \quad (3.22)$$

onde $\xi = 1, \dots, k$, u_i e v_j são as variáveis pertencentes aos universos de discursos U_i e V_j , respectivamente, x_i^{ξ} e y_j^{ξ} são conjuntos nebulosos definidos sobre os universos U_i e V_j para todo $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, m$. Note, entretanto, que podemos definir conjuntos nebulosos x^{ξ} e y^{ξ} , para $\xi = 1, \dots, k$, que capturam as informações dos conjuntos nebulosos do antecedente e do consequente sobre os produtos cartesianos $U = U_1 \times U_2 \times \dots \times U_n$ e

$V = V_1 \times V_2 \times \cdots \times V_m$, respectivamente. Os conjuntos nebulosos x^ξ e y^ξ são definidos como segue, para todo $u_i \in U_i$ e $v_i \in V_i$,

$$x^\xi(u) = \bigwedge_{i=1}^n x_i^\xi(u_i), \quad \text{e} \quad y^\xi(v) = \bigwedge_{j=1}^m y_j^\xi(v_j). \quad (3.23)$$

Assim, podemos converter uma regra nebulosa com múltiplas entradas e múltiplas saídas em uma regra da forma

$$R^\xi : \text{ Se } u \text{ é } x^\xi, \text{ então } v \text{ é } y^\xi, \quad (3.24)$$

onde $u = (u_1, u_2, \dots, u_n)$ e $v = (v_1, v_2, \dots, v_m)$ são as variáveis.

É importante observar que os conjuntos x^ξ e y^ξ representam o produto cartesiano $x_1^\xi \times x_2^\xi \times \cdots \times x_n^\xi$ e $y_1^\xi \times y_2^\xi \times \cdots \times y_m^\xi$, respectivamente.

CAPÍTULO 4

APLICAÇÃO: AVALIAÇÃO DA QUALIDADE DE UM SERVIÇO DE TRANSMISSÃO DE VOZ SOBRE IP

Atualmente, os serviços de telecomunicações não estão mais restritos apenas aos sistemas de telefonia convencionais. Hoje, temos também a opção de usar a internet como meio de comunicação em tempo real. Visto que a internet emprega um protocolo especial chamado IP¹, denominamos VoIP² qualquer tecnologia que usa a internet para a comunicação de voz [37].

É importante lembrar, entretanto, que a internet foi projetada para o tráfego de dados genéricos, e não necessariamente para a transmissão de voz em tempo real. Conseqüentemente, a qualidade dos serviços VoIP pode ser inferior aos serviços convencionais de telefonia, que foram projetados especificamente para a transmissão de voz [34]. Nesse contexto, surge a necessidade de medir a qualidade de um serviço VoIP [49]. Essa medida pode ser usada para o monitoramento e/ou controle do serviço de modo que alguns requisitos sejam satisfeitos. Nesse trabalho usaremos um sistema de regras nebulosas para estimar a qualidade de um serviço (QoS)³ VoIP.

Muitos fatores podem influenciar a qualidade de um serviço VoIP, incluindo a *taxa de pacotes perdidos* (ou simplesmente a *perda de pacotes*), *jitter*⁴ de *atraso* (ou simplesmente o *atraso*) e o *codec* empregado para a codificação e decodificação do sinal de voz. Nesse trabalho, iremos estimar a qualidade do serviço VoIP considerando apenas a perda de pacotes e o atraso. O codec utilizado durante todos os experimentos foi o *codec G. 711* [7].

Existe na literatura basicamente dois tipos de métodos para avaliar a qualidade de um serviço VoIP: *métodos objetivos* e *métodos subjetivos* [14, 49, 39]. Os métodos objetivos avaliam a qualidade do serviço VoIP utilizando dados da rede e algumas informações do sistema como o codec empregado. O *modelo-E* é um exemplo de método objetivo [21].

¹ do inglês *internet protocol*.

² do inglês *Voice over IP*.

³ do inglês, *quality of service*

⁴ *Jitter* é o valor absoluto da diferença entre o atraso de dois pacotes recebidos consecutivamente em um mesmo stream [14].

Em contraste, os métodos subjetivos simulam a opinião dos usuários sobre a qualidade do serviço VoIP [20]. É importante observar que a opinião dos usuários pode variar dependendo da língua empregada e dos serviços de telefonia que os usuários estão acostumados a utilizar [39]. O método que utilizaremos nesse trabalho para analisar a qualidade do sistema VoIP pode ser classificado como um método subjetivo pois foi projetado a partir da nossa concepção de como deve ser a qualidade do serviço VoIP, conhecidos a perda e o atraso da rede.

Finalmente, gostaríamos de lembrar que a estimativa da QoS deve ser efetuada em tempo real na ordem de milissegundos. Portanto, nesse trabalho buscamos um algoritmo que efetue poucas operações de ponto flutuante (ops) [6]. Optamos assim por um método que efetua, na grande maioria, operações de máximos e mínimos, i.e., as operações de junção e reunião em reticulados completos $[0; 1]$ e $\mathcal{F}(U)$.

4.1 MODELAGEM DO PROBLEMA

Nesse trabalho utilizaremos sistemas de regras nebulosas (SRNs) para estimar a qualidade de um serviço VoIP, onde cada regra nebulosa foi construída através do nosso conhecimento de como funciona um sistema de transmissão de voz sobre IP. Os conjuntos nebulosos usados no SRNs representam a taxa de pacotes perdidos, o atraso na transmissão e a qualidade do serviço.

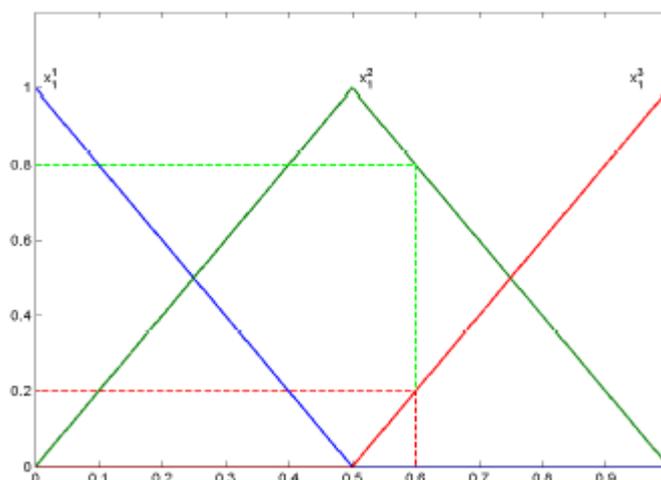


Figura 4.1 – Conjuntos nebulosos usados para modelar a taxa de pacotes perdidos. As linhas tracejadas indicam o grau de compatibilidade de uma perda 0,6 com os conceitos perda aceitável e perda inaceitável.

A Figura 4.1 apresenta os três conjuntos nebulosos que foram usados para descrever a taxa de pacotes perdidos, ou simplesmente a perda de pacotes. Esses conjuntos nebulosos representam os conceitos de *perda imperceptível*, *perda aceitável* e *perda inaceitável* de pacotes e foram denotados por x_1^1 , x_1^2 e x_1^3 , respectivamente. Assim, uma perda de 0,6 dos pacotes possui compatibilidade 0 (0%) com o conceito perda imperceptível, 0,8 (80%) com o conceito perda aceitável e 0,2 (20%) com o conceito perda inaceitável. De um modo similar, a Figura 4.2 apresenta conjuntos nebulosos x_2^1 , x_2^2 e x_2^3 que representam os conceitos de *atraso imperceptível*, *aceitável* e *inaceitável*, respectivamente. Desse modo, um atraso de 0,2 é 0,6 (60%) compatível com o conceito atraso imperceptível, 0,4 (40%) compatível com o conceito atraso aceitável e nada compatível com o conceito atraso inaceitável. Finalmente, a Figura 4.3 apresenta os conjuntos nebulosos que descrevem a qualidade do serviço VoIP. Especificamente, essa figura apresenta os conjuntos nebulosos y^1, y^2, \dots, y^5 que representam os conceitos de serviço *péssimo*, *ruim*, *regular*, *bom* e *excelente*, respectivamente. Considerando os conjuntos nebulosos definidos nas Figuras 4.1, 4.2 e 4.3, construímos regras nebulosas para um SRNs. A seguinte sentença é um exemplo de regra nebulosa:

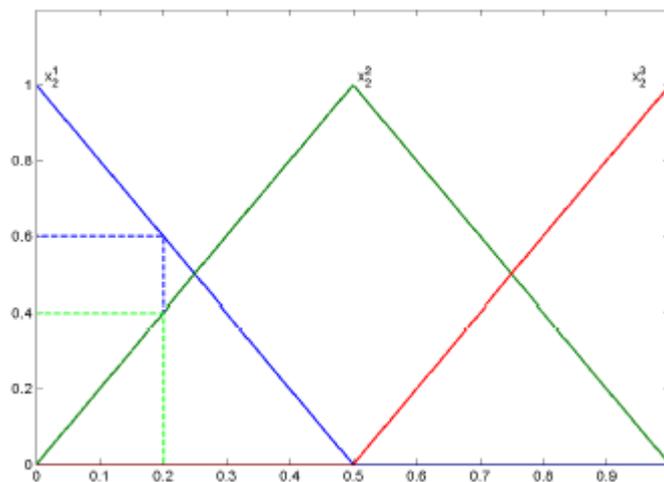


Figura 4.2 – Conjuntos nebulosos usados para modelar o *jitter* de atraso da rede. As linhas tracejadas indicam o grau de compatibilidade de um atraso de 0,2 com os conceitos atraso imperceptível e atraso aceitável.

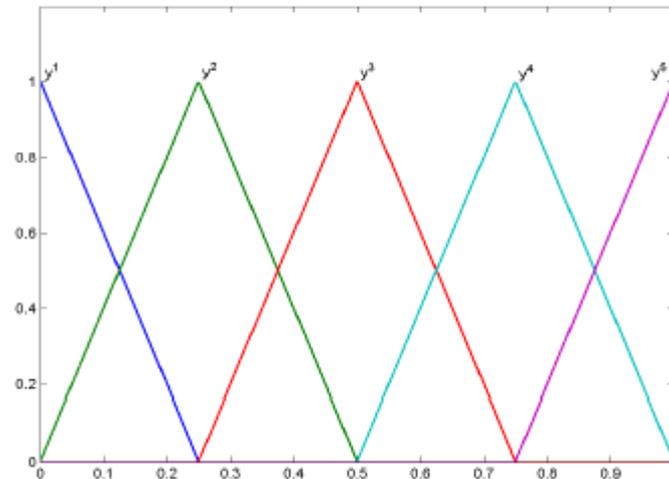


Figura 4.3 – Conjuntos nebulosos usados para modelar a qualidade do serviço VoIP.

Tabela 4.1 – Sistema de regras usado para avaliar a qualidade do serviço VoIP.

SRNs	x_2^1	x_2^2	x_2^3
x_1^1	y^5	y^4	y^3
x_1^2	y^4	y^3	y^2
x_1^3	y^3	y^2	y^1

Se a *perda é imperceptível* e o *atraso é imperceptível*, então o *serviço é ótimo*. (4.1) Ressaltamos que elaboramos um SRNs para avaliar a qualidade do serviço VoIP usando nossa percepção sobre a taxa de pacotes perdidos e o *jitter* de atraso da rede.

É importante observar que um SRNs com dois antecedentes e um consequente pode ser expresso de forma compacta usando uma tabela onde os conjuntos nebulosos do antecedente formam as linhas e colunas, e os elementos no interior da tabela representam os consequentes. Por exemplo, nesse trabalho consideramos o SRNs apresentado na Tabela 4.1. Aqui, as linhas e as colunas correspondem aos conjuntos nebulosos definidos nas Figuras 4.1 e 4.2, respectivamente. Os conjuntos nebulosos no interior da tabela correspondem aos conjuntos nebulosos apresentados na Figura 4.3. Note que a interação entre a primeira linha e a primeira coluna, i.e., a interação de x_1^1 e x_2^1 , resulta no conjunto nebuloso y^5 . Em outras palavras, temos a regra apresentada em (4.1).

Como vimos nas Seções 3.4 e 3.6 do Capítulo 3, podemos representar matematicamente um SRNs como segue onde x_1^η , x_2^μ e y^ν são conjuntos nebulosos em U_1 , U_2 e V , respectivamente:

$$\text{Se } u_1 \text{ é } x_1^\eta \text{ e } u_2 \text{ é } x_2^\mu, \text{ então } v \text{ é } y^\nu, \quad (4.2)$$

onde $\eta, \mu = 1, 2, 3$ e $\nu \in \{1, \dots, 5\}$ depende de η e μ , conforme a Tabela 4.1.

Dados elementos $u_1 \in U_1$ e $u_2 \in U_2$, podemos deduzir (ou inferir) um conjunto nebuloso y em V utilizando o método de inferência de Mamdani, onde o conjunto nebuloso y é definido por:

$$y(v) = \bigvee_{\eta, \mu=1}^3 (r^\nu \wedge y^\nu(v)) = \bigvee_{\eta, \mu=1}^3 (x_1^\eta(u_1) \wedge x_2^\mu(u_2) \wedge y^\nu(v)), \quad \forall v \in V, \quad (4.3)$$

onde o valor $r^\nu = x_1^\eta(u_1) \wedge x_2^\mu(u_2)$ avalia a veracidade da sentença “ u_1 é x_1^η e u_2 é x_2^μ ”, para cada $\eta, \mu = 1, 2, 3$.

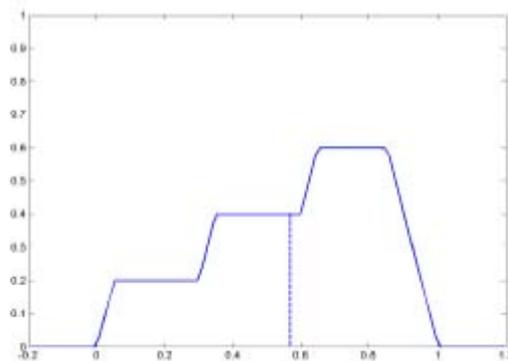


Figura 4.4 – Função de pertinência do conjunto nebuloso y deduzido pelo SRNs da Tabela 4.1 quando uma perda de 0,6 e um atraso de 0,2 são apresentados. A linha tracejada indica o centro de área.

Por exemplo, considere um valor de perda igual a 0,6 e um atraso de 0,2. Introduzindo esses dados no SRNs apresentado na Tabela 4.1 com os conjuntos nebulosos das Figuras 4.1, 4.2 e 4.3, obtemos o conjunto nebuloso y representado na Figura 4.4.

Finalmente, como o método de inferência de Mamdani apresenta como saída um conjunto nebuloso y em V , utilizamos o método de defuzzificação de centro de área (cf. eq. (3.15) do Capítulo 3) para determinar um valor $v^* \in V$ que representa o conjunto nebuloso y [42]. Por exemplo, o centro de área do conjunto nebuloso y da Figura 4.4 é 0,57.

Esse valor indica a nota dada pelo SRNs para a qualidade do serviço VoIP para uma taxa de 0,6 pacotes perdidos e um *jitter* de atraso de 0,2.

4.2 EXPERIMENTOS COMPUTACIONAIS

Para avaliar o desempenho do modelo baseado no SRNs da Tabela 4.1, consideramos os dados adquiridos por Ferrari [14]. Esses dados contém 47 amostras que foram coletadas usando uma variação do sistema QVoIP (cf. Tabela B.1 no Apêndice B), que emprega o codec G. 711. Para cada amostra, temos a taxa de pacotes perdidos, o *jitter* de atraso e uma nota atribuída para a qualidade do serviço⁵. As notas foram obtidas calculando a média das notas fornecidas por 7 pessoas. É importante observar que todas as variáveis, i.e., perda, atraso e nota, foram confinadas ao intervalo [0,1], por meio de uma transformação linear.

O *Erro Quadrático Médio* (EQM) e *Erro Percentual Médio Normalizado* (EPMN) obtido considerando as 47 amostras foram 0:008 e 10%, respectivamente. Lembre-se que essas medidas são calculadas através das eqs. em (4.4) a seguir, onde $N = 47$ refere-se ao número de amostras, os valores q_1, \dots, q_N denotam a média das notas fornecidas pelas 7 pessoas e $\hat{q}_1, \dots, \hat{q}_N$ correspondem aos valores fornecidos pelo SRNs

$$EQM = \frac{1}{N} \sum_{i=1}^N (q_i - \hat{q}_i)^2 \quad \text{e} \quad EPMN = \frac{100}{N} \times \left(\sum_{i=1}^N \frac{|q_i - \hat{q}_i|}{q_i} \right). \quad (4.4)$$

Realizamos também experimentos considerando uma rede neural artificial (cf. Apêndice B). Especificamente, utilizamos o Perceptrons de Múltiplas Camadas com uma única camada oculta com 5 neurônios artificiais e treinado com o Algoritmo de Retropropagação tradicional [16, 15, 43]. É importante observar que esse é um dos modelos de rede neural artificial mais empregado na literatura. Sobretudo, vários pesquisadores utilizaram essa rede neural artificial para avaliar a qualidade de um serviço VoIP [14, 39, 49].

Para avaliar o desempenho da rede neural artificial, dividimos aleatoriamente as 47 amostras em 3 conjuntos [16, 43]. Precisamente, 15 amostras foram usadas para treinamento (i.e., para determinar os valores das conexões sinápticas da rede

⁵ Especificamente, as notas para a QoS correspondem à uma normalização do *Mean Opinion Score* (MOS). [20]

neural), 10 amostras foram usadas para validação (i.e., foram usadas como critério de parada do algoritmo de retropropagação) e 22 amostras foram usadas para teste (para avaliar realmente a performance da rede neural). O EQM e o EPMN fornecidos pela rede neural após 100 simulações foram 0:003 e 4%, respectivamente. Note que esses valores são menores que os erros fornecidos pelo SRNs. Entretanto, algumas observações podem ser feitas com respeito às redes neurais artificiais e o SRNs.

Redes neurais artificiais são modelos que “aprendem” com dados. Portanto, uma rede neural terá uma boa performance somente se uma grande quantidade de dados estiverem disponíveis [15, 43]. No nosso caso, temos apenas 47 amostras que foram divididas em 3 conjuntos. Logo, temos poucos dados. Além disso, quando trabalhamos com redes neurais artificiais não conseguimos extrair muitas informações sobre esses modelos [42]. Por esse motivo, muitos pesquisadores afirmam que redes neurais artificiais representam “caixas-pretas”.

Em contraste, SRNs são projetados partindo da nossa concepção de como deve ser a qualidade do serviço VoIP conhecidos a taxa de pacotes perdidos e o *jitter* de atraso da rede. Portanto, não precisam de um grande banco de dados, como é o nosso caso. Sobretudo, SRNs são modelos transparentes e de fácil compreensão.

Finalmente, gostaríamos de observar que o SRNs pode ser melhorado alterando os conjuntos nebulosos definidos nas Figuras 4.1, 4.2 e 4.3. De fato, fizemos experimentos alterando os conjuntos nebulosos da seguinte forma: para os conjuntos nebulosos que representam a perda, definimos as seguintes funções de pertinência

$$\begin{aligned}x_1^1(u) &= \left(0 \vee [(u + 0,8)/0,8 \wedge (0,8 - u)/0,8]\right), \\x_1^2(u) &= \left(0 \vee [u/0,8 \wedge (1 - u)/(1 - 0,8)]\right), \\x_1^3(u) &= \left(0 \vee [(u - 0,8)/(1 - 0,8) \wedge (1,6 - u)/(1,6 - 0,8)]\right),\end{aligned}$$

para todo $u \in [0, 1]$.

Esta mudança gerou um EQM de 0:0012 e um EPMN de 3%, i.e., obtivemos resultados melhores que o modelo original e a rede neural artificial. Todavia, esses resultados melhores foram obtidos examinando os dados que temos disponíveis, o que torna os valores questionáveis pois não sabemos como o SRNs se comportará quando apresentarmos outros dados. Mesmo assim, esses experimentos confirmam que os SRNs oferecem uma boa opção para estimar a qualidade de um serviço VoIP.

CAPÍTULO 5

CONCLUSÃO

Nesta dissertação de mestrado discutimos sobre sistemas de regras nebulosas (SRNs) e suas aplicações. É importante ressaltar que SRNs é uma ferramenta eficaz para modelar problemas que envolvem conceitos nebulosos ou imprecisos, com pouco custo operacional e de fácil compreensão. Observamos também que um SRNs com o método de inferência de Mamdani e o centro de área como técnica de defuzzificação representa um aproximador universal, i.e., tal processo de inferência pode ser usado para aproximar uniformemente qualquer função contínua definida sobre um compacto [2, 9].

Dividimos essa dissertação em duas partes, na primeira tratamos de alguns resultados teóricos como o Teorema 3.5.1 da Seção 3.5. Já na segunda, utilizamos um SRNs para avaliar a qualidade do serviço VoIP.

Nos Capítulos 2 e 3, foi estabelecida uma analogia entre as operações de junção e reunião de reticulados com as operações usuais de soma e multiplicação. Além disso, delineamos conceitos de espaço reticulado e operadores \sup - \wedge , que são análogos às noções de espaço vetorial e operadores lineares. Todavia, lembramos que uma analogia deve ser sempre vista com cuidado pois resultados válidos para espaço vetoriais podem não valer para espaços reticulados.

Observamos também que o operador que representa o conjunto consequente do método de Mamdani, visto na eq. (3.13) da Seção 3.5, pode ser apresentado como a forma geral de se representar um SRNs. O argumento usado frequentemente é o seguinte: Como SRNs capturam relações entre as variáveis $u \in U$ e $v \in V$, esses podem então ser representados de forma geral usando uma relação nebulosa. Todavia, o Teorema 3.5.1 mostra que a eq. (3.13) descreve apenas um subconjunto de todos os operadores entre as famílias de conjuntos nebulosos $\mathcal{F}(U)$ e $\mathcal{F}(V)$, i.e., a classe dos operadores \sup - \wedge . Mostramos que o método de inferência introduzido por Kosko constitui um exemplo de operador entre conjuntos nebulosos que não pode ser descrito pela eq. (3.13).

No Capítulo 4, falamos sobre a tecnologia de transmissão de voz sobre IP (VoIP), que surgiu como uma alternativa para os sistemas de telecomunicações

convencionais. Junto com essa tecnologia veio a necessidade de se melhorar a qualidade na transmissão VoIP. No intuito de melhorar essa transmissão, foram criados métodos para se avaliar a qualidade do serviço VoIP.

Nessa dissertação foi construído um modelo baseado em SRNs para avaliar a qualidade da transmissão de voz sobre IP. Nosso modelo pode ser considerado como um método subjetivo, pois foi construído a partir do nosso conhecimento de como funciona o serviço VoIP.

O SRNs foi construído com duas variáveis de entrada e uma única variável de saída. O modelo baseado no SRNs fornece uma nota para a qualidade do serviço considerando apenas a taxa de pacotes perdidos e o atraso durante a transmissão.

Para testar nosso modelo, além de utilizarmos o método de inferência de Mamdani aplicado a um SRNs, utilizamos também uma rede neural artificial (RNA) com múltiplas camadas. Como forma de comparação dos métodos foram utilizadas 47 amostras adquiridas por Ferrari, onde cada amostra contém a taxa de pacotes perdidos, o atraso na transmissão (jitter) e uma nota atribuída para a qualidade do serviço.

Os resultados obtidos tiveram um *Erro Quadrático Médio* (EQM) de 0,008 e 0,003, e um *Erro Percentual Médio Normalizado* (EPMN) de 10% e 4% para o SRNs e a RNA, respectivamente. Note que o erro do SRNs é maior que o da RNA. A vantagem de se trabalhar com sistemas de regras nebulosas é a seguinte: SRNs são projetados partindo da nossa concepção de como deve ser a qualidade do serviço VoIP conhecidos a taxa de pacotes perdidos e o *jitter* de atraso da rede. Portanto, não precisam de um grande banco de dados, como é o nosso caso. Sobretudo, SRNs são modelos transparentes, de baixo custo operacional e de fácil compreensão. Observamos que, o SRNs utilizado pode ser melhorado, alterando os conjuntos nebulosos que representam os antecedentes e os consequentes do SRNs, como visto na Seção 4.

Em estudos futuros pretendemos coletar mais dados sobre a qualidade do serviço VoIP. Esses dados serão usados para ajustar (ou refinar) os conjuntos nebulosos que definem os antecedentes e os consequentes do SRNs.

Nessa dissertação não apresentamos uma definição formal de espaço reticulado. Pretendemos no futuro desenvolver uma teoria bem estruturada sobre espaços reticulados como tratam alguns autores em assuntos similares [5, 10, 35, 44]. Pretendemos também, ao invés de definimos um espaço reticulado $\text{sup-}\wedge$, definir um espaço reticulado $\text{sup-}T$, onde T é uma T -norma [5, 35]. Um outro estudo possível é avaliar se existe um operador sobre espaços reticulados que melhor represente um SRNs em um certo sentido.

REFERÊNCIAS

- [1] A. CHERRI, D. J., and SILVA, I. Uma abordagem fuzzy para o problema de corte de estoque unidimensional com sobras de material aproveitáveis. In: *Anais do XXXIX Simpósio Brasileiro de Pesquisa Operacional* (Fortaleza, Brasil, 2007).
- [2] BANDO, F. Sistemas fuzzy e aproximação universal. Master's thesis, Universidade Estadual de Campinas - Instituto de Matemática, Estatística e Computação Científica, 2002.
- [3] BANON, G., BARRERA, J., and BRAGA-NETO, U., Eds. *Mathematical Morphology and its Applications to Signal and Image Processing*. INPE, São José dos Campos, 2007. Proceedings of the 8th International Symposium on Mathematical Morphology.
- [4] BARROS, L., and BASSANEZI, R. *Tópicos de Lógica Fuzzy e Biomatemática*. Editora do Instituto de Matemática, Estatística e Computação Científica (IMECCUnicamp), Campinas, 2006.
- [5] BIRKHOFF, G. *Lattice Theory*, 3 ed. American Mathematical Society, Providence, 1993.
- [6] BURDEN, R. L., and FAIRES, J. D. *Numerical Analysis*. 2004.
- [7] C., H. W. *VoIP Service Quality: Measuring and Evaluating Packet Switched Voice*, 1.ed. ed. New York: McGraw Hill, 2003.
- [8] CaStilho, A., and VALLE, M. Avaliação da qualidade de um serviço de transmissão de voz sobre IP usando sistemas de regras nebulosas. In *XII Encontro Regional de Matemática Aplicada e Computacional (XII ERMAC)* (Foz do Iguaçu, Brazil, August 11-13 2008). Available at: <http://www.uel.br/pessoal/valle>.
- [9] CASTRO, J. L., and DELGADO, M. Fuzzy systems with defuzzification are universal approximators. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* (1996), 26(1):149{152.
- [10] CUNINGHAME-GREEN, R. *Minimax Algebra: Lecture Notes in Economics and Mathematical Systems 166*. Springer-Verlag, New York, 1979.

- [11] DE PAULA CASTANHO, M. J., de BARROS, L. C., YAMAKAMI, A., and VENDITE, L. L. Fuzzy expert system: An example in prostate cancer. *Applied Mathematics and Computation* 202, 1 (2008), 78{85.
- [12] DUBOIS, D., and PRADE, H. *Fuzzy sets and systems: theory and applications*. Academic Press, New York, 1980.
- [13] ENGELBRECHT, A. P. *Computational Inteligence - An Introduction*, 2^a ed. Wiley, South Africa, 2007.
- [14] FERRARI, H. V. Análise e classificação de transmissões de voz sobre IP. Tech. rep., Trabalho de conclusão de curso de graduação. Universidade Estadual de Londrina, 2007.
- [15] HASSOUN, M. H. *Fundamentals of Artificial Neural Networks*. MIT Press, Cambridge, MA, 1995.
- [16] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Upper Saddle River, NJ, 1999.
- [17] HEBB, D. *The Organization of Behavior*. John Wiley & Sons, New York, 1949.
- [18] HEIJMANS, H. *Morphological Image Operators*. Academic Press, New York, NY, 1994.
- [19] HOFFMAN, K., and KUNZE, R. *Linear Algebra*. Prentice-Hall, Englewood Cliffs, New Jersey, 1971.
- [20] INTERNATIONAL TELECOMMUNICATION UNION. *Methods for Subjective Determination of Transmission Quality*, August 1996. ITU-T Recommendation P.800. Series P: 53 Telephone Transmission Quality; Methods for Objective and Subjective Assessment of Quality.
- [21] INTERNATIONAL TELECOMMUNICATION UNION. *The E-model, a computational model for use in transmission planning*, March 2005. ITU-T Recommendation G.107. Series G: Transmission Systems and Media, Digital Systems and Networks.
- [22] JAFELICE, R. M., de BARROS, L. C., BASSANEZI, R. C., and GOMIDE, F. A. C. Fuzzy modeling in symptomatic hiv virus infected population. *Bulletin of Mathematical Biology* 66, 6 (November 2004), 1597{1620.

- [23] JAFELICE, R. M., de BARROS, L. C., BASSANEZI, R. C., and GOMIDE, F. A. C. Methodology to determine the evolution of asymptomatic HIV population using fuzzy set theory. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 13, 1 (2005), 39{58.
- [24] KABURLASOS, V., and PAPADAKIS, S. Fuzzy lattice neurocomputing (n) models. *Neural Networks* 13, 10 (Dec. 2000), 1145{1169.
- [25] KABURLASOS, V., and RITTER, G., Eds. *Computational Intelligence Based on Lattice Theory*. Springer-Verlag, Heidelberg, Germany, 2007.
- [26] KASABOV, N. K. *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*. MIT Press, Cambridge, MA, USA, 1996.
- [27] KLIR, G. J., and YUAN, B. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, Upper Saddle River, NY, 1995.
- [28] KOSKO, B. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, Englewood Cliffs, NJ, 1992.
- [29] KREYSZIG, E. *Introductory Functional Analysis with Applications*. J. Wiley, New York, 1989.
- [30] LEE, K. H. *First Course on Fuzzy Theory and Applications*. Springer, Berlin, Germany, 2005.
- [31] LUSTOSA, L. C. G. Arquitetura de monitoração de qualidade de chamadas telefônicas ip. Master's thesis, Instituto de Matemática / Núcleo de Computação Eletrônica da Universidade Federal do Rio de Janeiro, Rio de Janeiro - RJ, 2005.
- [32] MacLANE, S. Garrett Birkhoff and the survey of modern algebra. *Notices of the American Mathematical Society* 44, 11 (December 1997), 1438{1439. Available at: <http://www.ams.org/notices/199711/comm-maclane.pdf>.
- [33] MAGALHÃES, M., BALLINI, R., GONÇALVES, R., and GOMIDE, F. Predictive fuzzy clustering model for natural streamflow forecasting. In *Proceedings of the IEEE International Conference on Fuzzy Systems* (Budapest, Hungary, 2004), pp. 390-394.

- [34] MANOUSOS, M., APOSTOLACOS, S., GRAMMATIKAKIS, I., MEXIS, D., KAGKLIS, D., and SYKAS, E. Voice-quality monitoring and control for VoIP. *IEEE Internet Computing* 9, 4 (2005), 35{42.
- [35] MARAGOS, P. Lattice image processing: A unification of morphological and fuzzy algebraic systems. *Journal of Mathematical Imaging and Vision* 22, 2-3 (2005), 333{353.
- [36] MCCULLOCH, W., and PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5 (1943), 115{133.
- [37] MINISTÉRIO DAS COMUNICAÇÕES. ANATEL - *Agência Nacional de Telecomunicações*, Governo Brasileiro. Página Eletrônica da Agência Nacional de Telecomunicações. Disponível no endereço: <http://www.anatel.gov.br/>. Acessado em 30 de Junho de 2008.
- [38] MIYASAKI, R. Um estudo das ligações entre memórias associativas fuzzy implicativas e equações relacionais fuzzy com aplicações. Master's thesis, Universidade Estadual de Campinas - UNICAMP, Campinas - SP, Brasil, 2008.
- [39] MOHAMED, S., C.-P. F., and AFIFI, H. Integrating networks measurements and speech quality subjective scores for control purposes. *Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2001)*. volume 2 (2001), 641{649.
- [40] NGUYEN, H. T., and WALKER, E. A. *A First Course in Fuzzy Logic*, 2 ed. Chapman & Hall/CRC, Boca Raton, 2000.
- [41] PEDRYCZ, W., and GOMIDE, F. *An Introduction to Fuzzy Sets: Analysis and Design*. MIT Press, Cambridge, Massachusetts, 1998.
- [42] PEDRYCZ, W., and GOMIDE, F. *Fuzzy Systems Engineering: Toward Human-Centric Computing*. Wiley-IEEE Press, New York, 2007.
- [43] RIPLEY, B. D. *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, 1996.
- [44] RITTER, G. X., and GADER, P. Fixed points of lattice transforms and lattice associative memories. In *Advances in Imaging and Electron Physics*, P. Hawkes, Ed., vol. 144. Academic Press, New York, NY, 2006.

- [45] RITTER, G. X., and WILSON, J. N. *Handbook of Computer Vision Algorithms in Image Algebra*, 2 ed. CRC Press, Boca Raton, 2001.
- [46] ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65 (1958), 386{408.
- [47] ROTA, G. The many lives of lattice theory. *Notices of the American Mathematical Society*.
- [48] RUMELHART, D., HINTON, G., and WILLIAMS, R. *Learning internal representation by error propagation*. Parallel Distributed Processing. MIT Press, Cambridge, MA, 1988.
- [49] SUN, L., and IFEACHOR, E. Voice quality prediction models and their application in VoIP networks. *IEEE Transactions on Multimedia* 8, 4 (2006), 809{820.
- [50] SUSSNER, P., and VALLE, M. E. Classification of fuzzy mathematical morphologies based on concepts of inclusion measure and duality. *Journal of Mathematical Imaging and Vision* (2007). Accepted for publication.
- [51] SUSSNER, P., and VALLE, M. E. Fuzzy associative memories and their relationship to mathematical morphology. In *Handbook of Granular Computing*, W. Pedrycz, A. Skowron, and V. Kreinovich, Eds. John Wiley and Sons, Inc., New York, 2008, ch. 33, pp. 733{754. Invited contribution, to appear.
- [52] TATIBANA, C. Y., and Kaetsu, D. Y. Uma introdução a redes neurais artificiais. Home page: <http://www.din.uem.br/ia/neurais/>.
- [53] VALLE, M. E. *Fundamentos e Aplicações de Memórias Associativas Morfológicas Nebulosas*. PhD thesis, Universidade Estadual de Campinas (UNICAMP), Campinas, Brasil, Fevereiro 2007.
- [54] ZADEH, L. A. Fuzzy sets. *Information and Control* 8, 3 (1965), 338{353.
- [55] ZADEH, L. A. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics SMC-3, 1* (January 1973), 28{44.

[56] ZADEH, L. A. Fuzzy sets and information granularity. In *Advances in Fuzzy Set Theory and Applications*, M. M. Gupta, R. K. Ragade, and R. R. Yager, Eds. North Holland, Amsterdam, 1979, pp. 3{18.

[57] ZADEH, L. A. Fuzzy logic = computing with words. *Fuzzy Sets and Systems* 4(2) (1996), 103{111.

APÊNDICES

APÊNDICE A
Espaços Vetoriais e Operadores Lineares

Apêndice A – Espaços Vetoriais e Operadores Lineares

Um *espaço vetorial* \mathcal{V} definido sobre um conjunto de escalares \mathbf{E} é um conjunto não vazio munido de duas operações algébricas, a adição de vetores e a multiplicação por escalar [19]. Exemplos de espaços vetoriais incluem \mathbb{R}^n e o conjunto $\mathcal{C}(U)$ de todas as funções reais contínuas definidas sobre um subconjunto compacto $U \subseteq \mathbb{R}$ [29].

Definição A.0.1 *Sejam \mathcal{V} e \mathcal{W} espaços vetoriais definidos sobre \mathbf{E} , um operador $T : \mathcal{V} \rightarrow \mathcal{W}$ é linear se e somente se a seguinte equação vale para todo $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ e $\alpha, \beta \in \mathbf{E}$:*

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y}). \quad (1.1)$$

Note que a definição de operador linear é analoga a definição de operador sup- \wedge em reticulados (cf. Seção 2.2).

Operadores integrais são exemplos de operadores lineares sobre o espaço das funções reais contínuas. Especificamente, operadores $T : \mathcal{C}(U) \rightarrow \mathcal{C}(V)$ dados pela seguinte equação, onde $U, V \subseteq \mathbb{R}$ são subconjuntos compactos, representam operadores lineares:

$$T(\mathbf{x}) = \int_U K(u, v)\mathbf{x}(u)du. \quad (1.2)$$

Aqui, $K : U \times V \rightarrow \mathbf{E}$ é o *núcleo* do operador linear T . Observe ainda, que a eq. (1.2) é analoga a eq. (2.12) da Seção 2.2. Lembre-se também que temos um operador linear $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ se e somente se existe uma matriz $A \in \mathbb{R}^{m \times n}$ tal que $T_A(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x} \in \mathbb{R}^n$. Em outras palavras, um operador linear entre \mathbb{R}^n e \mathbb{R}^m é completamente caracterizado por uma matriz $A \in \mathbb{R}^{m \times n}$. Esse resultado serviu como motivação para o Teorema 3.5.1 apresentado no Capítulo 3.

APÊNDICE B
Redes Neurais Artificiais

Apêndice B – Redes Neurais Artificiais

A partir do momento em que as máquinas começaram evoluir, um grande desejo do homem tem sido a criação de uma máquina que possa operar independentemente do controle humano. Uma máquina cuja independência seja desenvolvida de acordo com seu próprio aprendizado e que tenha a capacidade de interagir com ambientes incertos (desconhecidos por ela), uma máquina que possa ser chamada de autônoma, inteligente ou cognitiva. O cérebro humano é considerado a mais fascinante máquina processadora de informações, sendo composto por aproximadamente 10 bilhões de neurônios. Os neurônios estão conectados uns aos outros através de sinapses, e juntos formam uma grande rede, chamada *rede neural*. Esta grande rede proporciona uma fabulosa capacidade de processamento e armazenamento de informação [52].

As *Redes Neurais Artificiais* (RNAs) são modelos matemáticos inspirados no cérebro humano por sua habilidade de adquirir e armazenar conhecimento necessários para realizar uma dada tarefa [53]. Devido a motivação biológica, os elementos básicos de processamento de uma rede neural são chamados neurônios. Assim como o cérebro humano é capaz de transmitir e/ou armazenar informações através de impulsos nervosos entre os neurônios, as RNAs possuem conexões inter-neurais, representadas por pesos sinápticos, com função de armazenar e processar o conhecimento [33].

As primeiras informações mencionadas sobre redes neurais artificiais datam de 1943, em artigos de McCulloch e Pitts, em que sugeriam a construção de uma máquina baseada ou inspirada no cérebro humano [36]. Muitos outros artigos e livros surgiram desde então. Em 1949 Donald Hebb escreveu um livro intitulado “The Organization of Behavior” (A Organização do Comportamento) [17] que perseguia a idéia de que o condicionamento psicológico clássico está presente em qualquer parte dos animais pelo fato de que esta é uma propriedade de neurônios individuais. Suas idéias não eram completamente novas, mas Hebb foi o primeiro a propor uma lei de aprendizagem específica para as sinapses dos neurônios. Este primeiro e corajoso passo serviu de inspiração para que muitos outros pesquisadores perseguissem a mesma idéia [52]. Embora muito tenha sido estudado e publicado durante as décadas subsequentes, estes estudos serviram mais como base para desenvolvimento posterior que para o próprio desenvolvimento.

A primeira máquina a obter sucesso como redes neurais (Mark I Perceptron) surgiu em 1958, criado por Frank Rosenblatt. Devido a profundidade de seus estudos, suas contribuições técnicas e sua maneira moderna de pensar, muitos o vêem como o fundador da neuro computação na forma em que a temos hoje. Seu interesse inicial para a criação do Perceptron era o reconhecimento de padrões [46].

Em 1974, Werbos lançou bases para o algoritmo de retropropagação (backpropagation) que posteriormente foi estudado por Rumelhart, Hinton e Williams em 1986 e emergiu como um algoritmo de aprendizado para o treinamento de perceptrons de multiplas camadas [38].

B.1 Redes Neurais Artificiais

As redes neurais artificiais são modelos matemáticos semelhantes as estruturas neurais biológicas e que têm capacidade de armazenar conhecimento, aplicar esse conhecimento armazenado para resolver problemas e adquirir novos conhecimentos através de aprendizado e generalização. As RNAs se assemelham ao cérebro humano em dois aspectos:

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

Além da semelhança com o cérebro, as RNAs possuem características bastante atraentes [33]:

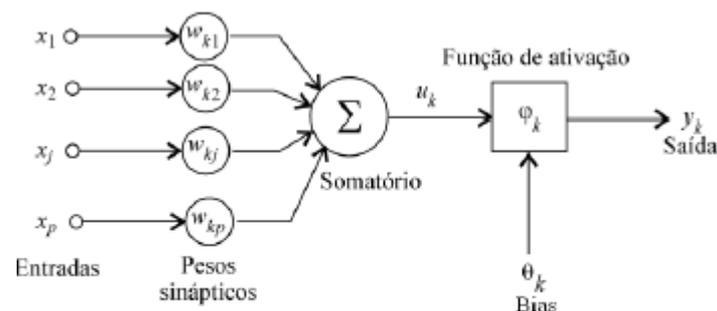


Figura B.1: Modelo não-linear de um neurônio.

- Não requerem *a priori* informações sobre o modelo, pois aprendem a partir de exemplos e de experiência. Para que uma RNA possa aprender e garantir um bom desempenho no resultado, o conjunto de dados disponíveis deve ser suficientemente grande. Nesse sentido, estes modelos podem ser tratados como um método não-linear.
- Têm a capacidade de generalizar. Após a fase de aprendizagem, quando já possuem algum conhecimento sobre as características do conjunto de dados, as RNAs conseguem inferir na população restante mesmo se os dados amostrais contêm ruído nas informações.
- Algumas RNAs são aproximadores universais, i.e., podem aproximar uma função contínua, definida sob um domínio compacto, com a precisão desejada. Por exemplo, os perceptrons de múltiplas camadas.

O funcionamento e o desempenho de uma RNA depende diretamente da escolha dos componentes adotados em sua estrutura e dos métodos adotados em seu processo de aprendizagem.

A estrutura de uma RNA é formada por diversas unidades de processamento chamadas *neurônios artificiais*. Essas unidades processam suas entradas através da soma ponderada e de uma *função de ativação*. Esquemáticamente, cada neurônio pode ser representado como na Figura B.1, onde $\varphi_k(\cdot)$ representa a função de ativação de um neurônio.

Em um modelo de neurônio, podem ser identificados três elementos básicos [16, 38]:

- Um conjunto de sinais de entrada e um conjunto de pesos sinápticos ou conexões sinápticas w_{kj} , onde o índice k corresponde ao número do neurônio e j ao estímulo de entrada. Quanto maior o peso w_{kj} , maior será a contribuição da entrada j para o somatório;
- Um somatório dos sinais de entrada, ponderados pelos seus respectivos pesos sinápticos do neurônio, resultando no valor u_k ;
- Uma função de ativação φ , que visa restringir ou limitar a amplitude da saída de um neurônio a um determinado intervalo. Aqui o intervalo referido é o intervalo $[0,1]$. Além disso, a função de ativação introduz uma não-linearidade no modelo.

Observe que o modelo neural apresentado na Figura B.1 inclui um parâmetro externo, denominado *bias*. O *bias* tem o papel de aumentar ou diminuir a entrada da função de ativação, dependendo se ele é positivo ou negativo, respectivamente [16].

Matematicamente, um neurônio k pode ser escrito pelo seguinte par de equações:

$$u_k = \sum_{j=1}^m w_{kj}x_j \quad (2.1)$$

e

$$y_k = \varphi(u_k + \theta_k), \quad (2.2)$$

onde os x_i são os sinais de entrada; w_{ki} são os pesos sinápticos do neurônio k ; u_k é a saída do combinador linear devido aos sinais de entrada; θ_k é o *bias*; $\varphi(\cdot)$ é a função de ativação e y_k é o sinal de saída do neurônio.

A função de ativação $\varphi(\cdot)$ representa a saída de um neurônio em termos de v_k , onde $v_k = u_k + \theta_k$. A função de ativação utilizada neste trabalho foi uma *função sigmoideal* definida como segue

$$\varphi(v_k) = \frac{1}{1 + e^{-av_k}}, \quad (2.3)$$

onde a é o ponto de inflexão da função $\varphi(\cdot)$. Outros exemplos de funções de ativação podem ser vistos em [13, 16].

A organização dos neurônios de uma RNA geralmente é feita em camadas, sendo a primeira, denominada camada de entrada e a última, camada de saída. O número de camadas intermediárias e o número de neurônios em cada uma delas são geralmente escolhidos empiricamente de acordo com a necessidade de se processar as entradas da rede, i.e., os números de camadas e neurônios aumentam caso o conjunto de dados exija um processamento mais refinado para capturar características não lineares [33]. É importante observar, entretanto, que uma única camada oculta é suficiente para resolver uma grande variedade de problemas [16, 26].

Os neurônios recebem as informações das entradas e as propagam através dos pesos w_i , $i = 1, \dots, m$, procedimento que simula o processamento sináptico. Os pesos sinápticos são parâmetros adaptáveis ajustados via processo de treinamento [16]. Esse processo pode ser classificado, principalmente, como treinamento supervisionado, treinamento não supervisionado ou treinamento por reforço. Neste trabalho, o processo de treinamento utilizado foi o *treinamento supervisionado*, descrito com mais detalhes a seguir.

B.2 Perceptrons de Múltiplas Camadas (PMCs)

O *perceptron de múltiplas camadas* (PMCs) com o algoritmo de retropropagação (back-propagation) representa uma importante classe das redes neurais artificiais. Este tipo de rede consiste de um conjunto de neurônios conectados, seguindo uma formação em camadas. As unidades que compõem a camada de entrada têm como objetivo difundir o sinal inicial sem nenhuma modificação para a segunda camada. O sinal de entrada é processado ao longo da rede e, após percorrer todas as camadas ocultas, produz uma resposta gerada pelos neurônios da camada de saída [16, 33], como visto na Figura B.2.

Os PMCs são utilizados com sucesso para resolver diversos problemas, através de seu treinamento de forma supervisionada utilizando o *algoritmo de retropropagação* [16].

O treinamento supervisionado caracteriza-se pela disponibilidade de conhecimento ou informação sobre um sistema na forma de padrões *entrada-saída* [16]. Aplicando um algoritmo de treinamento supervisionado, uma RNA adquire conhecimento ou informação relevante sobre um problema de forma analoga aquela utilizada pelo ser humano e outros animais, ou seja, a partir de exemplos e de experiência.

Um processo supervisionado segue basicamente o seguinte esquema. Considerando um conjunto de treinamento composto por padrões *entrada-saída* desejada; para uma dada entrada, o objetivo é ajustar os parâmetros da rede através da comparação entre a

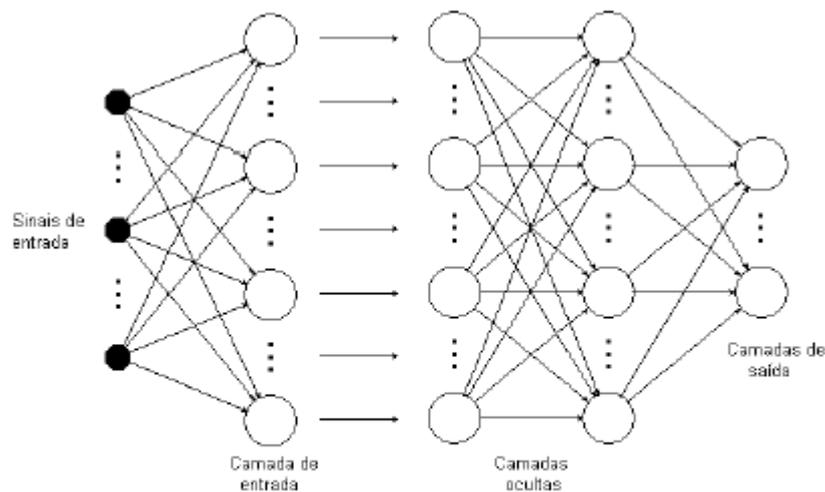


Figura B.2: Representação grafica de um perceptron de multiplas camadas.

resposta da rede e a saída desejada correspondente. O ajuste é feito com base no sinal de erro, calculado pela diferença entre a saída desejada (conjunto de treinamento) e a saída da rede.

O algoritmo de treinamento mais difundido e utilizado neste trabalho é o de retropropagação (backpropagation) [48]. Esse método consiste em duas fases. Na primeira, conhecida como fase de propagação direta (forward), as entradas são apresentadas e propagadas através da rede, camada a camada, calculando a saída de cada neurônio. Durante essa fase os pesos são fixos e a saída calculada é comparada com a saída desejada, resultando em um erro para cada unidade de saída.

Na segunda fase, o erro calculado é propagado da camada de saída para a camada de entrada, fase de propagação reversa (backward), e os pesos são ajustados de acordo com a regra de correção do erro, originando o termo retropropagação do erro. A segunda fase pode ser formulada como um problema de otimização não linear irrestrito, tendo os pesos sinápticos como variáveis de decisão e objetivo de minimizar o somatório dos erros quadráticos de todos os neurônios de saída [33].

A extração do conhecimento é realizada a partir de um conjunto de n padrões entrada-saída, chamado conjunto de treinamento, $\{(x^1, d^1), \dots, (x^l, d^l), \dots, (x^n, d^n)\}$, onde $x^l = (x_1^l, \dots, x_i^l, \dots, x_N^l)$ e $d^l = (d_1^l, \dots, d_j^l, \dots, d_M^l)$ representam os sinais de entrada e as saídas desejadas, respectivamente. Na fase de propagação direta (forward), um sinal de entrada x^l é propagado camada a camada e a respectiva saída $y^l = (y_1^l, \dots, y_j^l, \dots, y_M^l)$ é comparada a saída desejada d^l . Assim, para cada neurônio k da camada de saída, na t -ésima apresentação de um padrão (x^l, d^l) , define-se um erro associado.

$$e_k^l(t) = d_k^l(t) - y_k^l(t). \quad (2.4)$$

A soma dos erros quadráticos é obtida para todos os neurônios da camada de saída e é dada por

$$e^l(t) = \frac{1}{2} \sum_{k=1}^M (e_k^l(t))^2, \quad (2.5)$$

onde M é o número de saídas.

Considerando o número total n de pares entrada-saída contidos no conjunto de treinamento, o erro quadrático médio é obtido pela soma dos erros correspondentes a cada padrão, normalizado com relação ao número de padrões n :

$$E(t) = \frac{1}{n} \sum_{l=1}^n e^l(t). \quad (2.6)$$

O objetivo do processo de aprendizagem é ajustar os parâmetros livres (pesos) da rede para minimizar E . Para a derivação do algoritmo de retropropagação, é considerado um método de treinamento em que os pesos são ajustados padrão-à-padrão, ou seja, o ajuste dos pesos é realizado à partir do erro calculado para cada entrada apresentada a rede, conforme a Eq. 2.5.

O processo de aprendizagem termina quando uma das seguintes condições é satisfeita:

- A rede alcança o índice de desempenho estabelecido, i.e., $E(t) \leq \delta$, onde δ é suficientemente pequeno;
- A aprendizagem termina caso o número de iterações t ultrapasse um valor máximo especificado;

No entanto, quando utilizamos um PMCs treinado com o algoritmo de retropropagação, o erro médio quadrado decresce conforme o número de iterações durante o treinamento, i.e., ele começa com um valor grande, decresce rapidamente e então continua diminuindo

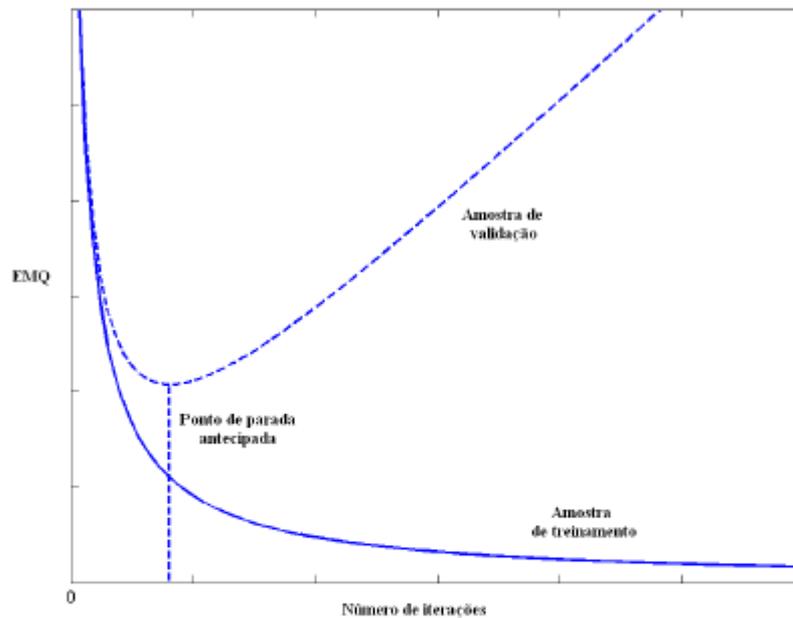


Figura B.3: Ilustração da regra de parada antecipada utilizando o Erro Médio Quadrado (EMQ).

lentamente conforme os pesos sinápticos se aproximam de um mínimo local de E . Por outro lado, é muito difícil estimar o melhor momento para encerrar o treinamento, pois é possível que a rede seja excessivamente ajustada aos dados de treinamento [16].

Podemos então identificar o início do excesso de treinamento dividindo os dados de treinamento em dois subconjuntos, um de estimação e outro de validação. O subconjunto de exemplos de estimação é usado para treinar a rede da maneira usual, exceto por uma pequena modificação: a sessão de treinamento é interrompida periodicamente após um determinado número de iterações t , e a rede é testada com o subconjunto de validação. Este processo é conhecido como o *método de treinamento com parada antecipada*.

A Figura B.3 mostra duas curvas de erro de aprendizagem, uma relativa ao erro sobre o subconjunto de estimação e a outra relativa ao subconjunto de validação. Note que, o modelo gera um erro maior para o subconjunto de validação do que o erro gerado para o subconjunto de estimação, sobre o qual a rede foi baseada. Assim, o ponto de mínimo na curva de aprendizagem de validação é escolhido como um critério sensato para encerrar o treinamento, pois o que a rede aprende após esse ponto é essencialmente o ruído contido nos dados de treinamento [16].

B.3 Aplicação do PMCs para Avaliar a Qualidade de Um Serviço VoIP

Neste trabalho realizamos experimentos com um PMCs com o algoritmo de retropropagação como processo de aprendizagem. O algoritmo utilizado para as simulações está disponível no seguinte endereço eletrônico:

<http://www.uel.br/pessoal/valle/MRLNN/matlabcodes.htm>.

Primeiramente, o conjunto de dados disponíveis foi dividido aleatoriamente em um conjunto de treinamento e em um conjunto de teste. Particionamos o conjunto de treinamento em dois subconjuntos distintos:

1. Um subconjunto usado durante a aprendizagem da rede para a estimação dos parâmetros do modelo, denominado *subconjunto de estimação*. Nos experimentos, utilizamos 15 amostras de variações do sistema QVoIP (cf. Tabela B.1);
2. Um subconjunto usado para avaliar a evolução do desempenho do modelo, ou seja, para a validação da rede, denominado *subconjunto de validação* com 10 amostras. Este subconjunto corresponde a uma pequena parte do conjunto de treinamento, variando de 10% a 20% do seu tamanho.

A seguir, a partir do subconjunto de treinamento, foram estimados os parâmetros da rede e utilizando esses parâmetros no subconjunto de validação, verificou-se o desempenho de várias topologias da rede. A melhor topologia foi escolhida de acordo com erros apresentados durante a validação. Especificamente, a topologia escolhida foi uma rede com uma única camada oculta com 5 neurônios artificiais. Lembrando que, quanto menores

os erros, melhor é o desempenho da rede testada.

Após encontrada a melhor arquitetura e os pesos sinápticos da rede, iniciou-se a fase de teste. Aqui, as 22 amostras restantes, foram utilizadas para verificar a capacidade de generalização da rede, i.e., para avaliar realmente a performance da rede neural.

Tabela B.1: Valores de dados adquiridos através de uma variação do sistema QVoIP [14], os dados estão normalizados no intervalo $[0, 1]$.

Atraso	Perda	Estimativa do MoS	Atraso	Perda	Estimativa do MoS
0,9075	0,1600	0,03939	0,8958	0,2600	0,06465
0,8071	0,5000	0,05081	0,8594	0,4000	0,05086
0,5447	1,0000	0,05734	0,9002	0,2600	0,01772
0,9042	0,2000	0,03452	0,8952	0,3200	0,01113
0,8984	0,2400	0,05668	0,8971	0,3000	0,01114
0,8782	0,3200	0,19318	0,8985	0,2800	0,01619
0,8753	0,3200	0,22600	0,9012	0,2600	0,00712
0,8720	0,3600	0,22056	0,8976	0,3000	0,00605
0,8755	0,3600	0,18218	0,8953	0,3200	0,01015
0,5794	0,9200	0,13463	0,8979	0,2800	0,02234
0,6728	0,7400	0,14315	0,9001	0,2600	0,01921
0,8732	0,3600	0,20690	0,8988	0,2800	0,01297
0,6909	0,7000	0,17473	0,8994	0,2600	0,02603
0,8841	0,3000	0,15010	0,5580	0,9800	0,02730
0,8802	0,3400	0,15185	0,8967	0,2800	0,03502
0,8855	0,2800	0,15547	0,6628	0,7800	0,02687
0,5412	0,9800	0,20837	0,8971	0,2800	0,03112
0,8760	0,3600	0,17725	0,8951	0,3200	0,01199
0,8852	0,2800	0,15858	0,8939	0,3200	0,02503
0,8745	0,3800	0,00203	0,6383	0,8200	0,06439
0,8994	0,2800	0,00610	0,8740	0,3800	0,00692
0,8425	0,4400	0,00808	0,8998	0,2800	0,00201
0,8973	0,3000	0,00916	0,9014	0,1600	0,10505
0,8953	0,3200	0,01015			