

IFAMs - Memórias Associativas Baseadas no Aprendizado Nebuloso Implicativo

Marcos Eduardo Valle

Instituto de Mat., Est. e Comp. Científica
Univ. Estadual de Campinas
Campinas, SP 13083-859
E-mail: mevalle@ime.unicamp.br

Peter Sussner

Instituto de Mat., Est. e Comp. Científica
Univ. Estadual de Campinas
Campinas, SP 13083-859
E-mail: sussner@ime.unicamp.br

Resumo—Memórias associativas são modelos que descrevem o fenômeno biológico que permite o armazenamento de padrões e a recordação destes após a apresentação de uma entrada ruidosa ou incompleta. Neste artigo apresentamos o aprendizado nebuloso implicativo e discutimos as memórias associativas nebulosas implicativas (IFAMs). No aprendizado nebuloso implicativo, os pesos sinápticos são determinados computando o mínimo das implicações das ativações pré e pós-sinápticas. Uma IFAM é descrita por uma rede neural nebulosa totalmente conexa com neurônios de Pedrycz como unidades básicas de processamento e treinada usando o aprendizado nebuloso implicativo. Apresentamos uma série de resultados comparando as IFAMs com outros modelos encontrados na literatura e uma aplicação das IFAMs em um problema de previsão.

I. INTRODUÇÃO

Uma memória associativa é um sistema de entrada e saída que permite o armazenamento e a recuperação de informação. Além disso, uma memória associativa é capaz de recuperar correntemente uma saída mesmo após a apresentação de uma entrada incompleta ou ruidosa. Estas características fazem das memórias associativas uma área de pesquisa com aplicações em diferentes ramos da ciência tais como processamento de imagens, reconhecimento de padrões e otimização [1], [2], [3]. Os estudos em memórias associativas iniciaram nos anos 50 e vários modelos foram apresentados deste então [4], [5], [6], [7], [8], [9], [10]. O modelo de memória associativa mais famoso é a rede de Hopfield usada para padrões binários [11], [12]. O modelo de Hopfield e os modelos mencionados acima representam redes neurais semi-lineares tradicionais onde em cada camada calculamos um produto matriz-vetor seguido de uma função de ativação possivelmente não-linear. Em contraste com os modelos de memória associativa semi-lineares, as memórias associativas nebulosas e morfológicas utilizam um produto matriz-vetor não-linear e geralmente nenhuma função de ativação é adicionada [13], [14], [15].

O primeiro modelo de memória associativa nebulosa (FAM, Fuzzy Associative Memory) foi apresentado por Kosko em 1992 [15]. Esta é uma rede neural de camada única descrita em termos das composições max-min e max-prod. A matriz dos pesos sinápticos da FAM de Kosko é obtida através da regra de Hebb nebulosa. O modelo de Kosko apresenta uma capacidade de armazenamento baixa devido à interferência cruzada entre os padrões armazenados. Chung e Lee generali-

zaram a FAM de Kosko para uma composição max-t qualquer e mostraram que uma recordação perfeita é possível se o conjunto das memórias fundamentais for normal e ortogonal no sentido da composição max-t [16]. As *memórias associativas nebulosas implicativas* (IFAMs, Implicative Fuzzy Associative Memories) discutidas aqui não encontram estas restrições [17]. A grande diferença entre uma IFAM e uma FAM Generalizada de Chung-Lee (GFAM, Generalized FAM) está na regra de aprendizado que foi introduzida em [18], [19] para o caso particular da composição max-min. Em vista da implicação nebulosa usada neste aprendizado, iremos nos referir a ele como *aprendizado nebuloso implicativo*. Junbo *et. al.* aplicaram o aprendizado nebuloso implicativo para o caso particular da composição max-min obtendo a chamada Max-min FAM [19]. Liu aperfeiçoou posteriormente a Max-min FAM de Junbo *et. al.* introduzindo um vetor limiar (threshold) no modelo [20]. Wang e Lu introduziram recentemente as chamadas memórias associativas nebulosas morfológicas [21], um modelo inspirado nas memórias associativas morfológicas [13], [14].

Neste artigo generalizamos o aprendizado nebuloso implicativo para incluir qualquer composição max-t e introduzimos as IFAMs como sendo redes neurais nebulosas com vetor limiar treinadas usando o aprendizado nebuloso implicativo. Verificamos que o modelo de Liu pode ser visto como um caso particular das IFAMs. Apresentamos comparações entre as IFAMs e outros modelos através de exemplos encontrados na literatura [20], [22]. Mostramos também que as memórias associativas morfológicas com vetor limiar podem ser vistas como um caso particular das IFAMs se as memórias fundamentais forem conjuntos nebulosos e se a matriz dos pesos sinápticos estiver restrita ao hiper-cubo $[0, 1]^{m \times n}$.

Este artigo está organizado em sete seções. Na próxima seção apresentamos os conceitos básicos de redes neurais nebulosas. Na seção III introduzimos o aprendizado nebuloso implicativo. Na seção IV introduzimos as IFAMs e apresentamos um exemplo comparando a capacidade absoluta de armazenamento das IFAMs com outros modelos encontrados na literatura. Na seção V apresentamos uma aplicação das IFAMs em um problema de previsão. Na seção VI apresentamos a relação entre as IFAMs e as memórias associativas morfológicas. O artigo termina com a conclusão na seção VII.

II. REDES NEURAIS NEBULOSAS

As memórias associativas nebulosas implicativas são redes neurais nebulosas com neurônios lógicos de Pedrycz com limiar [23], [24]. Se x_1, x_2, \dots, x_n são os sinais de entrada, a saída é calculada através das equações

$$y_i = \bigvee_{j=1}^n (w_{ij} \text{t} x_j) \vee \theta_i, \quad \text{for } i = 1, 2, \dots, m. \quad (1)$$

ou

$$y_i = \bigwedge_{j=1}^n (m_{ij} \text{s} x_j) \wedge \vartheta_i, \quad \text{for } i = 1, 2, \dots, m. \quad (2)$$

onde w_{ij} e m_{ij} são os pesos sinápticos e, θ_i e ϑ_i representam um limiar (threshold), para $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$. Os símbolos t e s representam normas (t-norma) e co-normas triangulares (s-norma), respectivamente. Note que θ_i e ϑ_i podem ser interpretados como pesos sinápticos conectados a uma entrada constante x_0 . Neste caso, $\theta_i = w_{i0} \text{t} x_0$, onde $x_0 = 1$ na equação 1, e $\vartheta_i = m_{i0} \text{s} x_0$, onde $x_0 = 0$ na equação 2.

O neurônio descrito pela equação 1 é conhecido como OU-neurônio e o modelo destrito pela equação 2 é chamado E-neurônio. As equações 1 e 2 descrevem redes neurais nebulosas alimentadas adiante de camada única. Esta será a única arquitetura de rede neural considerada neste artigo. Sendo as normas e co-normas triangulares funções não-lineares, não adicionaremos funções de ativação nos nossos modelos.

As equações 1 e 2 podem ser escritas numa forma matricial. Se $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in [0, 1]^n$ é o vetor de entrada e se $W = (w_{ij}) \in [0, 1]^{m \times n}$ e $M = (m_{ij}) \in [0, 1]^{m \times n}$ são as matrizes dos pesos sinápticos então o vetor de saída $\mathbf{y} = [y_1, y_2, \dots, y_m]^T \in [0, 1]^m$ é dado por

$$\mathbf{y} = (W \circ \mathbf{x}) \vee \boldsymbol{\theta}, \quad (3)$$

e

$$\mathbf{y} = (M \bullet \mathbf{x}) \wedge \boldsymbol{\vartheta}, \quad (4)$$

onde \circ e \bullet representam as composições max-t e min-s [24]. Lembre-se que a composição max-t de duas matrizes A e B de dimensões apropriadas é $C = A \circ B$ onde

$$c_{ij} = \bigvee_{k=1}^p (a_{ik} \text{t} b_{kj}). \quad (5)$$

Analogamente, a composição min-s de A e B é $D = A \bullet B$ onde

$$d_{ij} = \bigwedge_{k=1}^p (a_{ik} \text{s} b_{kj}). \quad (6)$$

Existe uma relação de dualidade entre as equações 3 e 4. Seja $\bar{R} = (\bar{r}_{ij}) \in [0, 1]^{m \times n}$ a matrix da negação de $R = (r_{ij}) \in [0, 1]^{m \times n}$ dada por

$$\bar{r}_{ij} = 1 - r_{ij}, \quad (7)$$

para todo i, j . Se A e B são matrizes nebulosas, então

$$\overline{A \bullet B} = \bar{A} \circ \bar{B}, \quad (8)$$

A equação 8 é válida somente se a t-norma e a s-norma nas composições \circ e \bullet forem também operações duais, isto é,

$$x \text{t} y = 1 - (1 - x) \text{t} (1 - y) = \overline{(\bar{x} \text{t} \bar{y})}, \quad (9)$$

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

Exemplo 1. O mínimo (\wedge) é uma t-norma e sua s-norma dual é o máximo (\vee). A s-norma dual do produto é a soma probabilística dada por

$$x \text{s} y = x + y - xy, \quad (10)$$

para $x, y \in [0, 1]$. A t-norma de Lukasiewicz (ou produto limitado) e sua s-norma dual, conhecida como s-norma de Lukasiewicz (ou soma limitada), são definidas como:

$$x \text{t} y = 0 \vee (x + y - 1), \quad (11)$$

$$x \text{s} y = 1 \wedge (x + y). \quad (12)$$

Neste trabalho nos concentramos nas redes neurais nebulosas descritas pela equação 3. Resultados para as redes neurais nebulosas descritas pela equação 4 podem ser obtidos através da relação de dualidade fornecida pela equação 8.

III. APRENDIZADO NEBULOSO IMPLICATIVO

O postulado de aprendizagem de Hebb afirma que a variação do peso de uma conexão sináptica depende de ambas ativações pré-sináptica (entrada) e pós-sináptica (saída) [25]. Em [26], Haykin apresenta a seguinte formulação matemática para a aprendizagem hebbiana. Considere um peso sináptico w_{ij} com ativação pré-sináptica x_j e ativação pós-sináptica y_i . O valor do pesos sináptico w_{ij} após o aprendizado de Hebb é expresso na forma geral através da equação

$$w_{ij} = F(x_j, y_i), \quad (13)$$

onde F é uma função binária. Haykin afirma que todas as formas da equação 13 são classificadas como hebbianas.

No aprendizado nebuloso implicativo consideramos uma implicação nebulosa como função das ativações pré e pós-sinápticas. Se x_j é a entrada e y_i é a saída, então a conexão sináptica w_{ij} é dada pela equação

$$w_{ij} = (x_j \Rightarrow y_i). \quad (14)$$

Lembre-se que uma *implicação nebulosa* é uma função binária $\Rightarrow: [0, 1]^2 \rightarrow [0, 1]$ satisfazendo as propriedades [27]:

- 1) Monotonicidade no segundo argumento,
- 2) Dominância da falsidade,
- 3) Neutralidade da verdade.

O aprendizado nebuloso implicativo pode ser interpretado do seguinte modo: a variação de um peso sináptico é o grau de veracidade da afirmação "Se estímulo, então resposta".

A forma tradicional do aprendizado de Hebb considera o produto das ativações x_j e y_i . Neste caso, temos uma natureza correlativa nas sinapses hebbianas. O aprendizado nebuloso implicativo dado pela equação 14 não possui uma natureza correlativa. Um aprendizado com natureza correlativa pode ser obtido substituindo a implicação por uma bi-implicação na equação 14. Neste caso, a conexão sináptica

seria computada através da equação $w_{ij} = (x_j \Leftrightarrow y_i)$ e o valor w_{ij} representaria o grau de veracidade da afirmação “Estímulo se e somente se resposta”. Esta segunda proposta não será considerada neste artigo porque ela pode restringir muito o valor das conexões sinápticas produzindo valores próximos de zero. Além disso, não teremos o teorema 1 para as memórias associativas nebulosas treinadas com este aprendizado alternativo.

Neste artigo nos concentraremos nas chamadas *R-implicações* associadas com a t-norma utilizada no modelo neural nebuloso descrito pela equação 3. Uma R-implicação é uma aplicação binária $\Rightarrow_R: [0, 1]^2 \rightarrow [0, 1]$ definida através da equação

$$(x \Rightarrow_R y) = \sup\{z \in [0, 1] : x t z \leq y\}, \quad (15)$$

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

Exemplo 2. A R-implicação associada ao mínimo é a implicação de Gödel dada por

$$x \Rightarrow_R y = \begin{cases} 1, & x \leq y \\ y, & x > y \end{cases}. \quad (16)$$

A R-implicação associada ao produto é a implicação de Goguen dada por

$$x \Rightarrow_R y = 1 \wedge \left(\frac{y}{x}\right), \quad (17)$$

e a R-implicação associada a t-norma de Lukasiewicz é dada por

$$x \Rightarrow_R y = 1 \wedge (1 - x + y). \quad (18)$$

Suponha que queremos armazenar um conjunto de associações $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, \dots, p\}$, onde $\mathbf{x}^\xi \in [0, 1]^n$ e $\mathbf{y}^\xi \in [0, 1]^m$, numa rede neural nebulosa descrita pela equação 3 usando o aprendizado nebuloso implicativo. Queremos que os pares de entrada e saída $(\mathbf{x}^1, \mathbf{y}^1)$ e $(\mathbf{x}^2, \mathbf{y}^2)$ e ... e $(\mathbf{x}^p, \mathbf{y}^p)$ sejam todos armazenados. Lembrando que a operação de mínimo corresponde à conjunção lógica “e”, a matriz dos pesos sinápticos $W = (w_{ij}) \in [0, 1]^{m \times n}$ será o mínimo das matrizes dos pesos sinápticos obtida para cada par $(\mathbf{x}^\xi, \mathbf{y}^\xi)$, isto é,

$$w_{ij} = \bigwedge_{\xi=1}^p (x_j^\xi \Rightarrow y_i^\xi). \quad (19)$$

Em outras palavras, a matriz dos pesos sinápticos W é sintetizada através da regra

$$W = Y \circledast_R X^T, \quad (20)$$

onde o produto $C = A \circledast_R B \in [0, 1]^{m \times n}$ de duas matrizes nebulosas $A \in [0, 1]^{m \times k}$ e $B \in [0, 1]^{k \times n}$ e calculado através da equação

$$c_{ij} = \bigwedge_{l=1}^k (b_{lj} \Rightarrow a_{il}). \quad (21)$$

A regra definida pela equação 20 será chamada *aprendizado nebuloso implicativo*. Quando uma R-implicação for usada, teremos um *aprendizado nebuloso R-implicativo* e usaremos

o símbolo \circledast_R para lembrar que uma R-implicação é utilizada na equação 21.

A matriz dos pesos sinápticos obtida através do aprendizado nebuloso R-implicativo é a solução do problema

$$\max \{W \in [0, 1]^{m \times n} : W \circ X \leq Y\}. \quad (22)$$

Logo, se existe uma matriz $W_0 \in [0, 1]^{m \times n}$ tal que $W_0 \circ X = Y$ então $W = Y \circledast_R X^T$ obtida através do aprendizado nebuloso R-implicativo é tal que $W \circ X = Y$ e $W_0 \leq W$ [28].

O aprendizado nebuloso implicativo será usado para treinar uma rede neural nebulosa descrita pela equação 3. Uma rede descrita pela equação 4 pode ser treinada aplicando o aprendizado nebuloso implicativo em conjunto com a relação de dualidade dada pela equação 8.

IV. MEMÓRIAS ASSOCIATIVAS NEBULOSAS IMPLICATIVAS

Uma memória associativa é um sistema de entrada e saída capaz de armazenar diferentes pares de padrões (\mathbf{x}, \mathbf{y}) [1]. O problema de formular uma memória associativa pode ser interpretado matematicamente como: Dado um conjunto finito de memórias fundamentais, $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, \dots, p\}$, determine uma aplicação G tal que $G(\mathbf{x}^\xi) = \mathbf{y}^\xi$ para todo $\xi = 1, \dots, p$. Além disso, a aplicação G deve apresentar tolerância com respeito a ruído, isto é, $G(\tilde{\mathbf{x}}^\xi) = \mathbf{y}^\xi$ para qualquer versão ruidosa $\tilde{\mathbf{x}}^\xi$ de \mathbf{x}^ξ .

Nas memórias associativas neurais nebulosas, a aplicação G é descrita por uma rede neural nebulosa e o processo usado para determinar G torna-se uma regra de aprendizado. Se o aprendizado nebuloso R-implicativo é usado para sintetizar a matriz dos pesos sinápticos, obtemos uma *Memória Associativa Nebulosa Implicativa* (IFAM, Implicative Fuzzy Associative Memory). Note que existem infinitas IFAMs induzidas pelas normas triangulares e R-implicações. Uma IFAM particular é especificada pela R-implicação utilizada no aprendizado nebuloso R-implicativo. Lembre-se que a R-implicação considerada está diretamente ligada a t-norma usada na composição max-t através da equação 15.

Os padrões de entrada e saída de uma IFAM são conjuntos nebulosos. Assim, os vetores de entrada $\mathbf{x}^\xi \in [0, 1]^n$ e os vetores de saída $\mathbf{y}^\xi \in [0, 1]^m$, para $\xi = 1, \dots, p$, são interpretados como conjuntos nebulosos definidos sobre um universo de discurso finito com m e n elementos, respectivamente.

Resumindo, uma IFAM é uma rede neural nebulosa descrita pela equação

$$\mathbf{y} = (W \circ \mathbf{x}) \vee \boldsymbol{\theta}, \quad (23)$$

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

$$W = Y \circledast_R X^T \quad \text{e} \quad \boldsymbol{\theta} = \bigwedge_{\xi=1}^p \mathbf{y}^\xi. \quad (24)$$

Aqui, a implicação usada no produto $Y \circledast_R X^T$ é a R-implicação associada a t-norma da composição max-t da equação 23. Como consequência do aprendizado nebuloso R-implicativo temos o seguinte teorema:

Teorema 1. Se existe W_0 e θ_0 tal que $(W_0 \circ \mathbf{x}^\xi) \vee \theta_0 = \mathbf{y}^\xi$ para todo $\xi = 1, 2, \dots, p$ então $W_0 \leq W = Y \circledast_R X^T$, $\theta_0 \leq \theta = \bigwedge_{\xi=1}^p \mathbf{y}^\xi$ e $\mathbf{y}^\xi = (W \circ \mathbf{x}^\xi) \vee \theta$ para todo $\xi = 1, \dots, p$.

Exemplo 3. Um exemplo apresentado por Liu revela que a capacidade absoluta de armazenamento da Max-min FAM com limiar supera a capacidade absoluta de armazenamento da FAM de Junbo *et. al.*, da FAM de Kosko e da FAM Generalizada de Chung-Lee [20], [19], [15], [16].

Considere o conjunto de memórias fundamentais $(\mathbf{x}^\xi, \mathbf{y}^\xi)$, para $\xi = 1, \dots, 8$, apresentadas na tabela I. Armazenamos

ξ	\mathbf{x}^ξ	\mathbf{y}^ξ
1	$[0.5, 0.5, 0.4, 0.4, 0.3]^T$	$[0.5, 0.6, 0.3]^T$
2	$[0.1, 0.3, 0.3, 0.4, 0.4]^T$	$[0.5, 0.6, 0.4]^T$
3	$[0.8, 0.4, 0.6, 0.7, 0.4]^T$	$[0.6, 0.8, 0.4]^T$
4	$[0.3, 0.4, 0.4, 0.3, 0.4]^T$	$[0.5, 0.6, 0.4]^T$
5	$[0.6, 0.4, 0.7, 0.7, 0.5]^T$	$[0.7, 0.7, 0.5]^T$
6	$[0.1, 0.1, 0.2, 0.2, 0.1]^T$	$[0.5, 0.6, 0.3]^T$
7	$[0.7, 0.2, 0.4, 0.3, 0.2]^T$	$[0.5, 0.7, 0.3]^T$
8	$[0.8, 0.4, 0.3, 0.4, 0.2]^T$	$[0.5, 0.8, 0.3]^T$

TABELA I

PADRÕES DE ENTRADA E SAÍDA UTILIZADOS NO EXEMPLO 3

estas 8 associações nos modelos de memória associativa nebulosa mencionados anteriormente. A Max-min FAM com limiar recordou corretamente todos os padrões. A FAM de Junbo recordou corretamente somente os pares $(\mathbf{x}^3, \mathbf{y}^3)$, $(\mathbf{x}^5, \mathbf{y}^5)$, $(\mathbf{x}^7, \mathbf{y}^7)$ e $(\mathbf{x}^8, \mathbf{y}^8)$. A FAM de Kosko é capaz de recordar somente o par $(\mathbf{x}^5, \mathbf{y}^5)$ e a FAM Generalizada de Chung-Lee usando a t-norma de Lukasiewicz é incapaz de armazenar e recordar corretamente este conjunto de memórias fundamentais. Armazenamos também este conjunto com 8 memórias fundamentais na memória associativa de Wang e Lu [21]. Este modelo foi capaz de recordar os pares $(\mathbf{x}^1, \mathbf{y}^1)$, $(\mathbf{x}^3, \mathbf{y}^3)$, $(\mathbf{x}^5, \mathbf{y}^5)$ e $(\mathbf{x}^8, \mathbf{y}^8)$.

Construímos a IFAM de Lukasiewicz usando a equação 24 e encontramos a matriz de pesos sinápticos

$$W_{\text{Lukasiewicz}} = \begin{bmatrix} 0.7 & 1.0 & 1.0 & 0.9 & 1.0 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ 0.5 & 1.0 & 0.8 & 0.7 & 1.0 \end{bmatrix}, \quad (25)$$

e o vetor de limiar

$$\theta = [0.5 \quad 0.6 \quad 0.3]^T. \quad (26)$$

A fase de recordação da IFAM de Lukasiewicz é realizada usando a t-norma de Lukasiewicz na composição max-t. Verificamos que $W \circ \mathbf{x}^\xi = \mathbf{y}^\xi$ para todo $\xi = 1, \dots, 8$. Analogamente para as IFAMs de Goguen e Gödel, onde as matrizes dos pesos sinápticos são

$$W_{\text{Goguen}} = \begin{bmatrix} 0.6250 & 1.0000 & 1.0000 & 0.8571 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.3750 & 0.6000 & 0.6667 & 0.5714 & 1.0000 \end{bmatrix}, \quad (27)$$

e

$$W_{\text{Gödel}} = \begin{bmatrix} 0.5 & 1.0 & 1.0 & 0.6 & 1.0 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ 0.3 & 0.3 & 0.3 & 0.3 & 1.0 \end{bmatrix}, \quad (28)$$

respectivamente. O vetor de limiar da IFAM de Goguen e da IFAM de Gödel coincidem e é dado pela equação 26,

pois este não depende da t-norma da composição max-t. A IFAM de Goguen utiliza a composição max-prod e a IFAM de Gödel utiliza a composição max-min na fase de recordação. Resumindo, apresentamos na tabela II a capacidade absoluta de armazenamento dos modelos apresentados neste exemplo. A FAM Generalizada de Chung-Lee com a t-norma de Lukasiewicz será referida como GFAM de Lukasiewicz.

Modelo	Capacidade Absoluta de Armazenamento
Max-min FAM com limiar	8
IFAM de Lukasiewicz	8
IFAM de Gödel	8
IFAM de Goguen	8
FAM de Junbo	4
M.A. de Wang e Lu	4
FAM de Kosko	1
GFAM de Lukasiewicz	0

TABELA II

CAPACIDADE ABSOLUTA DE ARMAZENAMENTO DOS MODELOS APRESENTADOS NO EXEMPLO 3

Note que a capacidade absoluta de armazenamento da Max-min FAM com limiar de Liu e a capacidade de armazenamento das IFAMs de Lukasiewicz, Goguen e Gödel coincidiram neste exemplo. De fato, com base no teorema 1, podemos mostrar que o desempenho da IFAM de Gödel é no mínimo tão bom quanto o desempenho da Max-min FAM com limiar. Neste sentido, podemos ver as IFAMs como uma extensão da Max-min FAM com limiar para qualquer composição max-t. Na próxima seção apresentamos um exemplo onde a IFAM de Lukasiewicz apresenta um resultado melhor que o modelo de Liu.

V. APLICAÇÃO DAS IFAMs COMO SISTEMA DE REGRAS NEBULOSAS

As memórias associativas nebulosas podem representar sistemas de regras nebulosas onde armazenamos um conjunto de regras da forma SE-ENTÃO. Nesta seção, apresentamos uma aplicação das IFAMs em um modelo de previsão. Precisamente, consideremos o problema apresentado em [22] que consiste na previsão de mão-de-obra requerida nas indústrias metalúrgicas do estado Bengal do Oeste na Índia. Inicialmente, definimos 5 variáveis linguísticas A_i , $i = 1, \dots, 5$ e um conjunto de sentenças nebulosas como “Se a mão-de-obra requerida no ano n é A_i , então a mão-de-obra requerida no $n+1$ é A_j ”. Assim, obtemos um conjunto de pares de entrada e saída que serão armazenados numa memória associativa nebulosa. Se W é a matriz dos pesos sinápticos e θ é o vetor limiar obtido após o processo de aprendizagem, então a estimativa da mão-de-obra requerida no ano $n+1$ é dada por:

$$A_{n+1} = (W \circ A_n) \vee \theta, \quad (29)$$

onde A_n é a mão-de-obra requerida no ano n e \circ é a composição max-t.

Usando a FAM de Kosko, Choudhury *et. al.* encontraram um erro médio de 2.669% enquanto que os métodos estatísticos

ARIMA1 e ARIMA2 apresentaram um erro médio de 9.79% e 5.48%, respectivamente. Seguindo o mesmo procedimento usando a GFAM de Lukasiewicz de Chung e Lee, a Max-min FAM com limiar de Liu, a memória associativa de Wang e Lu, a IFAM de Gödel, a IFAM de Goguen e a IFAM de Lukasiewicz, encontramos os erros médios apresentados na tabela III. A FAM de Kosko e a GFAM de Lukasiewicz apresentaram os mesmos erros. A memória associativa de Wang e Lu e a IFAM de Goguen também apresentaram os mesmos resultados. Finalmente, a Max-min FAM com limiar e a IFAM de Gödel também apresentaram os mesmos resultados.

Método	Erro médio
IFAM de Lukasiewicz	2.29%
FAM de Kosko	2.67%
GFAM de Lukasiewicz	2.67%
IFAM de Gödel	2.73%
Max-min FAM com limiar	2.73%
IFAM de Goguen	2.99%
M.A. de Wang e Lu	2.99%
ARIMA2	5.48%
ARIMA1	9.79%

TABELA III

ERRO MÉDIO NA PREVISÃO DE MÃO-DE-OBRA.

Na figura 1 apresentamos a mão-de-obra requerida nos anos 1984 até 1995. O valor atual é comparado com as previsões obtidas pelos métodos citados anteriormente.

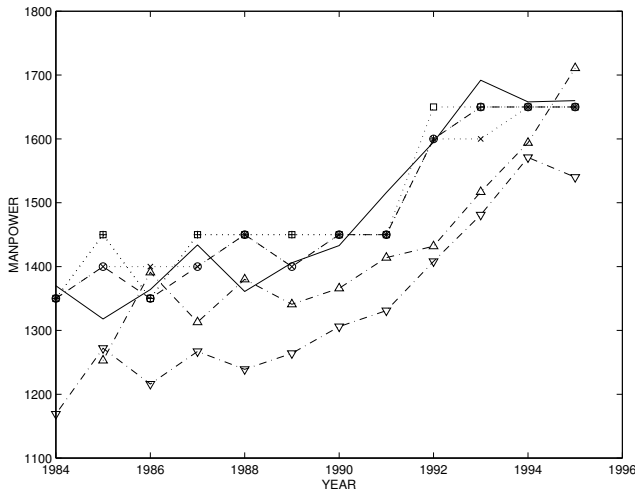


Fig. 1. Comparação da previsão de mão-de-obra estimada por diferentes modelos de memória associativa nebulosa. A linha contínua representa o valor atual da mão-de-obra. A linha tracejada marcada com 'o' representa a IFAM de Lukasiewicz, a linha pontilhada com 'x' corresponde a FAM de Kosko e a GFAM de Lukasiewicz, a linha pontilhada com '+' corresponde a Max-min FAM com limiar e a IFAM de Gödel e a linha pontilhada com '□' corresponde a memória associativa de Wang e Lu e a IFAM de Goguen. As linhas marcadas com '△' e '▽' representam os métodos ARIMA1 e ARIMA2, respectivamente.

A IFAM de Lukasiewicz apresentou o melhor resultado para este problema de previsão superando a FAM de Kosko, a GFAM de Lukasiewicz, a Max-min FAM com limiar e

o método estatístico ARIMA [22]. Este exemplo revela a aplicação e utilidade das IFAMs como sistemas de regras nebulosas.

VI. IFAMS E AS MEMÓRIAS ASSOCIATIVAS MORFOLÓGICAS

Existe uma relação entre a IFAM de Lukasiewicz e as memórias associativas morfológicas [13], [14]. As memórias associativas morfológicas são redes neurais descritas pela equação

$$\mathbf{y} = W_{XY} \boxtimes \mathbf{x}, \quad (30)$$

onde a matriz dos pesos sinápticos é dada por

$$W_{XY} = Y \boxtimes (-X^T). \quad (31)$$

Aqui, para matrizes $A \in \mathbb{R}^{m \times p}$ e $B \in \mathbb{R}^{k \times p}$, o produto $C = A \boxtimes B \in \mathbb{R}^{m \times n}$ e $D = A \boxtimes B \in \mathbb{R}^{m \times n}$ são definidos como

$$c_{ij} = \bigvee_{l=1}^p (a_{il} + b_{lj}) \quad \text{e} \quad d_{ij} = \bigwedge_{l=1}^p (a_{il} + b_{lj}), \quad (32)$$

para todo $i = 1, \dots, m$ e $j = 1, \dots, n$.

Podemos mostrar que se $X = [\mathbf{x}^1, \dots, \mathbf{x}^p] \in [0, 1]^{n \times p}$ e $Y = [\mathbf{y}^1, \dots, \mathbf{y}^p] \in [0, 1]^{m \times p}$ são as matrizes com as memórias fundamentais, então a IFAM de Lukasiewicz $W = Y \circledast_R X^T$ e a memória associativa morfológica $W_{XY} = Y \boxtimes (-X)^T$ estão relacionadas através das equações:

$$W = (W_{XY} \wedge 0) + 1, \quad (33)$$

$$(W \circ \mathbf{x}) \vee \theta = [(W_{XY} \wedge 0) \boxtimes \mathbf{x}] \vee \theta, \quad (34)$$

onde $\mathbf{x} \in [0, 1]^n$ é o padrão de entrada da memória associativa e $\theta = \bigwedge_{\xi=1}^p \mathbf{y}^\xi$ é o vetor de limiar.

Note que a fase de armazenamento e a fase de recordação da IFAM de Lukasiewicz podem ser descritas em termos da fase correspondente da memória associativa morfológica W_{XY} . Além disso, ambos modelos produzem a mesma saída \mathbf{y} se esta estiver dentro do hiper-cubo $[0, 1]^m$. Neste sentido, podemos ver as IFAMs como uma generalização das memórias associativas morfológicas. Uma aplicação da memória associativa morfológica W_{XY} para o problema de previsão apresentado na seção V produz os mesmos resultados obtidos pela IFAM de Lukasiewicz.

Existe também uma relação entre a IFAM de Goguen e a memória associativa de Wang e Lu, conhecida como *Memória Associativa Morfológica Nebulosa* [21] apesar do fato que não foi estabelecida nenhuma relação entre este modelo e a morfologia matemática nebulosa.

Na memória associativa de Wang e Lu, definimos

$$W_{XY}^{prod} = Y \circledast X^*, \quad (35)$$

onde a i, j -ésima componente de X^* é $1/x_{ji}$ e definimos $1/0 = \infty$ e $1/\infty = 0$. A recordação de um padrão é dada através da equação

$$\mathbf{y} = W_{XY}^{prod} \circledast \mathbf{x}. \quad (36)$$

As operações \otimes e \odot são definidas para matrizes A, B com elementos em $\mathbb{R}_{\geq 0}^{\infty} = \{r \in \mathbb{R} \cup +\infty : r \geq 0\}$ de modo que $C = A \odot B$ e $D = A \otimes B$ são matrizes dadas por

$$c_{ij} = \bigvee_{l=1}^k (a_{il} \cdot b_{lj}) \quad \text{e} \quad d_{ij} = \bigwedge_{l=1}^k (a_{il} \cdot' b_{lj}), \quad (37)$$

para todo índice $i = 1, \dots, m$ e $j = 1, \dots, n$. As operações \cdot e \cdot' diferem do produto usual somente no seguinte sentido

$$0 \cdot \infty = \infty \cdot 0 = 0, \quad (38)$$

$$0 \cdot' \infty = \infty \cdot' 0 = \infty. \quad (39)$$

Note que a fase de recordação da IFAM de Goguen e da memória associativa de Wang e Lu são idênticas exceto pelo vetor limiar usado na IFAM. A diferença entre estes dois modelos está na fase de armazenamento onde podemos verificar que a matriz dos pesos sinápticos W_{GG} da IFAM de Goguen está em $[0, 1]^{m \times n}$ enquanto que W_{XY}^{prod} está em $(\mathbb{R}_{\geq 0}^{\infty})^{m \times n}$. Como consequência deste último fato, um padrão recordado pela a memória associativa de Wang e Lu pode não ser nebuloso.

VII. CONCLUSÃO

Neste artigo apresentamos as memórias associativas nebulosas implicativas (IFAMs) que são redes neurais nebulosas treinadas usando o aprendizado nebuloso R-implicativo. Uma IFAM particular é definida pela R-implicação utilizada no aprendizado nebuloso R-implicativo, por exemplo, IFAM de Lukasiewicz, IFAM de Gödel e IFAM de Goguen.

Comparações entre as memórias associativas nebulosas implicativas e modelos de memória associativa nebulosa encontrados na literatura foram realizados através de um exemplo extraído de [20] e um problema de previsão apresentado em [22]. No primeiro experimento, o conjunto das memórias fundamentais foi perfeitamente armazenado nas IFAMs e na FAM de Liu. Os demais modelos não apresentaram uma recordação perfeita dos padrões originais devido a interferência cruzada. No segundo experimento, relacionado ao problema de predição de mão-de-obra em uma indústria metalúrgica, a IFAM de Lukasiewicz apresentou um resultado superior aos modelos de Kosko, Liu e Chung-Lee, bem como os métodos estatísticos ARIMA1 e ARIMA2. Este experimento revelou uma aplicação das IFAMs como sistema de regras nebulosas e indicou que as IFAMs podem ser usadas em problemas de aproximação de funções. Temos que observar porém que falta ainda conduzir pesquisas mais profundas sobre a utilização das IFAMs como aproximadores de funções. Finalmente, mostramos que a IFAM de Lukasiewicz e as memórias associativas morfológicas estão fortemente relacionadas se as memórias fundamentais forem conjuntos nebulosos. Uma relação semelhante foi mostrada para IFAM de Goguen e a memória associativa nebulosa morfológica de Wang e Lu.

AGRADECIMENTOS

Os autores gostariam de agradecer ao Prof. Fernando Gomide pelo incentivo e atenção. Este trabalho teve apoio do CNPq, processo no. 142196/03-7.

REFERÊNCIAS

- [1] M. H. Hassoun, "Dynamic associative neural memories," in *Associative Neural Memories: Theory and Implementation*, M. H. Hassoun, Ed. Oxford, U.K.: Oxford University Press, 1993.
- [2] M. M. Nelson and W. Illingworth, *A Practical Guide to Neural Nets*. Addison-Wesley Publishing Company, 1994.
- [3] B. Raducanu, M. Graña, and X. F. Albizuri, "Morphological scale spaces and associative morphological memories: Results on robustness and practical applications," *Journal of Mathematical Imaging and Vision*, vol. 19, no. 2, pp. 113–131, 2003.
- [4] K. Steinbruch, "Die lernmatrix," *Kybernetick*, vol. 1, pp. 36–45, 1961.
- [5] J. Anderson, "A simple neural network generating interactive memory," *Mathematical Biosciences*, vol. 14, pp. 197–220, 1972.
- [6] K. Nakano, "Associatron: A model of associative memory," *IEEE Trans. on Systems, Man, Cybernetics*, vol. SMC-2, pp. 380–388, 1972.
- [7] T. Kohonen, "Correlation matrix memory," *IEEE Transactions on Computers*, vol. C-21, pp. 353–359, 1972.
- [8] —, *Self-Organization and Associative Memory*. Springer Verlag, 1984.
- [9] B. Kosko, "Bidirectional associative memories," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, pp. 49–60, 1988.
- [10] T. Chiueh and R. Goodman, "Recurrent correlation associative memories," *IEEE Trans. on Neural Networks*, vol. 2, pp. 275–284, Feb. 1991.
- [11] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences*, vol. 79, pp. 2554–2558, Apr. 1982.
- [12] J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proceedings of the National Academy of Sciences*, vol. 81, pp. 3088–3092, May 1984.
- [13] G. X. Ritter, P. Sussner, and J. L. D. de Leon, "Morphological associative memories," *IEEE Transactions on Neural Networks*, vol. 9, no. 2, pp. 281–293, 1998.
- [14] G. X. Ritter, J. L. D. de Leon, and P. Sussner, "Morphological bidirectional associative memories," *Neural Networks*, vol. 6, no. 12, pp. 851–867, 1999.
- [15] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Englewood Cliffs, N.J.: Prentice Hall, 1992.
- [16] F. Chung and T. Lee, "On fuzzy associative memory with multiple-rule storage capacity," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 3, pp. 375–384, August 1996.
- [17] M. Valle, P. Sussner, and F. Gomide, "Introduction to implicative fuzzy associative memories," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, Hungary, July 2004, pp. 925 – 931.
- [18] Q. Cheng and Z.-T. Fan, "The stability problem for fuzzy bidirectional associative memories," *Fuzzy Sets and Systems*, vol. 132, no. 1, pp. 83–90, November 2002.
- [19] F. Junbo, J. Fan, and S. Yan, "A learning rule for fuzzy associative memories," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 7, June 1994, pp. 4273 – 4277.
- [20] P. Liu, "The fuzzy associative memory of max-min fuzzy neural networks with threshold," *Fuzzy Sets and Systems*, vol. 107, pp. 147–157, 1999.
- [21] S. Wang and H. Lu, "On new fuzzy morphological associative memories," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 3, pp. 316–323, 2004.
- [22] J. Choudhury, B. Sarkar, and S. Mukherjee, "Forecasting of engineering manpower through fuzzy associative memory neural network with arima: a comparative study," *Neurocomputing*, vol. 47, pp. 241–257, 2002.
- [23] W. Pedrycz, "Neurocomputations in relational systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 289–296, 1991.
- [24] —, "Fuzzy neural networks and neurocomputations," *Fuzzy Sets and Systems*, vol. 56, no. 1, pp. 1–28, 1993.
- [25] D. Hebb, *The Organization of Behavior*. New York: John Wiley & Sons, 1949.
- [26] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ: Prentice Hall, 1999.
- [27] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy Sets: Analysis and Design*. MIT Press Complex Adaptive Systems, 1998.
- [28] J. Drewniak, "Fuzzy relation equations and inequalities," *Fuzzy Sets and Systems*, vol. 14, no. 3, pp. 237–247, December 1984.