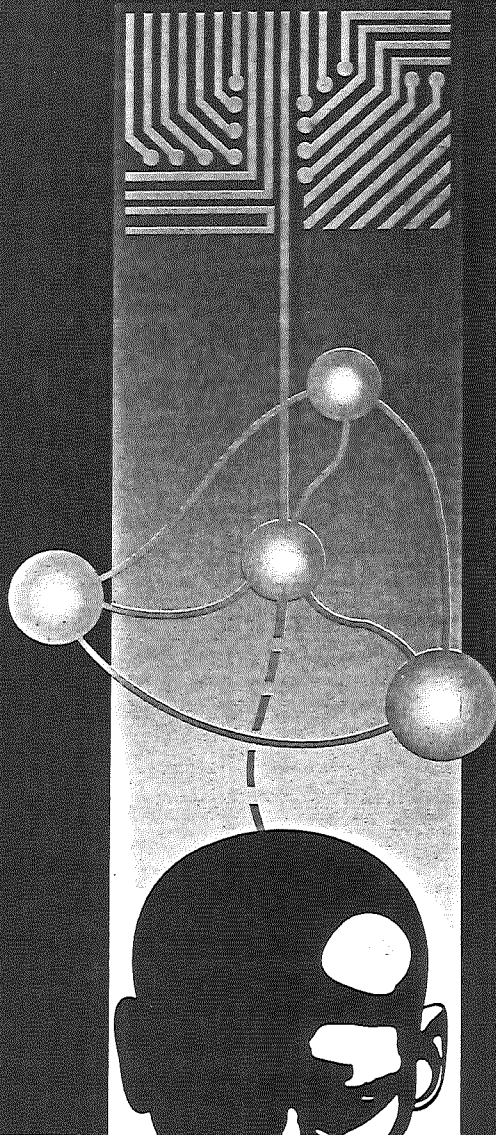


3.30.00.00-7



PROGRAMA
FRONTEIRA DO CONHECIMENTO

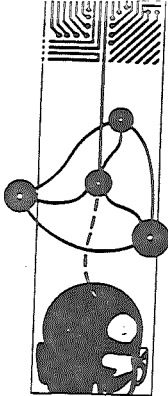
ANAIS DO
**II CONGRESSO BRASILEIRO
DE REDES NEURAIAS**
III ESCOLA DE REDES NEURAIAS



COPEL



31. APLICAÇÕES DE REDES NEURAIS ARTIFICIAIS À PREVISÃO DE CARGA E À DETECÇÃO DE ALARMES EM SISTEMAS DE POTÊNCIA
→ ANTONIO M. DE A. MASSOLA & ANTONIO S. R. MUNGIOLI
32. μ FILE - UMA MÁQUINA DE INFERÊNCIA FUZZY EM TEMPO REAL
→ GLAUCO A. DE P. CAURIN & SJUR J. VESTLI
33. CLASSIFICAÇÃO DE PADRÕES E EXTRAÇÃO DE PARÂMETROS DISCRIMINANTES UTILIZANDO REDES NEURAIS
→ LEE LUAN LING & MIGUEL GUSTAVO LIZÁRRAGA ESPINOSA
34. REDES NEURAIS ARTIFICIAIS E INFORMAÇÕES DE EXITAÇÃO NO RECONHECIMENTO AUTOMÁTICO DO LOCUTOR
→ ANDRÉ BORDIN MAGNI & EUVALDO F. CABRAL JR.
35. A RELATIONSHIP BETWEEN RECEPTIVE-FIELD OVERLAP IN EQUILIBRIUM SOMATOTOPIC MAPS AND LATERAL SYNAPTIC STRENGTH
→ ANTONIO C. R. DA SILVA FILHO
36. UMA OTIMIZAÇÃO VETORIAL EM REDES NEURAIS UTILIZANDO O ALGORITMO HÍBRIDO NEURO-GENÉTICO "SHAKE"
→ SILVIO S. PACHECO & ANTONIO g. THOMÉ
37. UMA ARQUITETURA PARCIALMENTE RECURSIVA APLICADA AO DIAGNÓSTICO DE TRANSFORMADORES
→ MARCELLO B. DE MARTINO & FRANCISCO DE ASSIS FILHO
38. REDES NEURAIS ARTIFICIAIS PARA RECONSTRUÇÃO DE IMAGENS TOMOGRÁFICAS
→ RENATO G. DE S. RODRIGUES; ANTONIO C. R. DA SILVA FILHO & CARLOS A. PELÁ
39. ARTIFICIAL NEURAL NETS APPLIED TO THE IDENTIFICATION OF NON LINEAR DYNAMICAL SYSTEMS
→ PEDRO S. P. LIMA & EUVALDO F. CABRAL JR.
40. IDENTIFYING RELEVANT FEATURES FOR A REAL-TIME NEURAL DISCRIMINATING PROCESS UNDER HIGH-EVENT RATE CONDITIONS
→ J. M. SEIXAS; L. P. CABOLA; C. D. P. POPPE & I. DELPINO
41. APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS PARA PREVISÃO DE CARGA NO PERÍODO DE DEMANDA DE PONTA
→ W. S. LIMA & T. OHISHI
42. IDENTIFICAÇÃO DE COMANDOS FRASAS UTILIZANDO REDES NEURAIS ARTIFICIAIS
→ DANIEL N. MÜLLER & PHILIPPE O. A. NAVAU
43. PREDIÇÃO DE VARIÁVEIS DE ESTADO DE UM PROCESSO FERMENTATIVO UTILIZANDO REDES NEURAIS
→ V. A. OLIVEIRA; L. F. MANESCO; J. B. VARGAS & A. M. F. SILVA
44. OPTIMAL LOCAL LEARNING IN A MULTILAYER MACHINE
→ MAURO COPELLI; OSAME KINOCHI & NESTOR CATICHA
45. PARALLEL IMPLEMENTATION OF RAM-BASED NEURAL NETWORKS
→ ANDRÉ C. P. DE LEON F. DE CARVALHO
46. VALIDAÇÃO DA RESPOSTA FINAL DE SISTEMAS BASEADOS EM ADAPTIVE LOGIC NETWORKS
→ WAGNER DA S. LIMA & WEBER MARTINS
47. ADAPTAÇÕES DO MODELO DE MEMÓRIA ASSOCIATIVA BIDIRECIONAL (BAM) PARA TAREFA HETEROASSOCIATIVA COM NÍVEIS VARIÁVEIS DE DIFICULDADE
→ ALUIZIO F. R. ARAÚJO



II CONGRESSO BRASILEIRO DE REDES NEURAIS III ESCOLA DE REDES NEURAIS

CURITIBA, 29/OUT A 01/NOV/1995

Adaptações no Modelo de Memória Associativa Bidirecional (BAM) para Tarefa Heteroassociativa com Níveis Variáveis de Dificuldade

Aluizio F. R. Araújo

Departamento de Engenharia Elétrica - USP
C.P. 359 - CEP 13.560-970 - São Carlos - SP
e-mail: aluizioa@peterpan.sel.eesc.sc.usp.br



Resumo: Este trabalho modela o desempenho variável da memória humana na memorização de pares de listas de palavras do tipo estímulo-resposta (S-R), cujos pares tenham diferentes níveis de dificuldade de associação. O algoritmo chamado Memória Associativa Bidirecional (BAM) é candidato a modelo pois lida com tarefas heteroassociativas, interferência e é sempre convergente. Contudo, a desejada melhora de performance na fase de testes demanda modificações no modelo original. Foram propostas e implementadas quatro maneiras de melhorar BAM para modelar o fenômeno em questão: percepção de dados com ruído, armazenamento de dados com ruído, “desaprendizagem” das respostas erradas e facilitação de acesso a dados na memória através de implantação de mecanismo que varia individualmente cada “threshold” seguido de estágio autoassociativo. As duas últimas opções se mostraram mais adequadas.

1. Introdução

Algoritmos de redes neurais são amplamente usados em modelamento de memória associativa pois possuem várias características pertinentes à memória humana (Kohonen, 1989).

O problema tratado neste artigo inspirou-se em teorias que discutem como fatores cognitivos e fatores afetivos influenciam o armazenamento de dados na memória humana e seu posterior acesso (Zajonc, 1980). Este artigo enfoca apenas fatores cognitivos. Em particular, deseja-se modelar o efeito da dificuldade de associação de palavras na performance da memória humana.

Muitas das conclusões das teorias citadas são oriundas de experimentos cuja tarefa consistiu na associação de pares de palavras. Alguns destes resultados são usados para testar a validade dos modelos propostos (Eysenck e Eysenck, 1985).

Desta forma, estuda-se a adaptação do modelo chamado Memória Associativa Bidirecional (BAM), proposto por Kosko (1988), para execução de uma tarefa heteroassociativa cuja

performance é função de seu nível variável de dificuldade. Este modelo deve ainda realizar aprendizagem “batch” e “on-line” de curto prazo e deve aprender com poucas ou apenas uma única apresentação dos padrões de treinamento.

A partir de hipóteses neuro-cognitivas de armazenamento e recuperação de dados na memória humana, propõe-se quatro variações do modelo original de BAM. O principal objetivo deste trabalho é analisar como tais versões de BAM atendem aos requisitos da tarefa.

Na seção 2, este artigo descreveu os testes psicológicos que serviram para avaliar a validade dos modelos propostos. Em seguida, apresentou-se o modelo BAM original. Nas seções 4 e 5 discutiu-se as hipóteses de alteração de BAM e como tais mudanças foram implementadas. As simulações dos modelos modificados foram apresentadas na seção 6 e seus resultados foram discutidos na seção 7.

2. Tarefa Associativa: Experimentos Psicológicos e Representação para Redes Neurais

Este trabalho visa modelar adição de novos dados à memória semântica (Baddeley, 1990) considerando como este procedimento possa ser facilitado ou prejudicado. Eysenck e Eysenck (1985) discutiram efeitos de fatores cognitivos e afetivos no desempenho de memória associativa humana. Eles destacaram o trabalho de Spence e Spence (1966), que ilustrou que o desempenho médio de sujeitos aprendendo associações de pares fáceis (não-competitivas) e difíceis (competitivas) oscila significativamente.

Nos experimentos originais, os sujeitos foram solicitados a associar pares de palavras do tipo estímulo-resposta (S_i-R_i , $i = 1, \dots, p$, onde p é o número de pares). Mais tarde eles deveriam responder que palavra era o par de um estímulo S_i .

Os pesquisadores estabeleceram o grau de dificuldade da tarefa a partir do nível de associação entre os dois membros de cada par, da similaridade dos estímulos entre si e das respostas entre si. Foram conduzidos dois experimentos onde a competição foi menor no primeiro que no segundo.

A diferença de dificuldade da tarefa foi obtida através da construção de duas listas do tipo S-R. A primeira, relativa à tarefa fácil, tinha associação S_i-R_i inicial alta, associação S_i-R_j baixa e baixa sinonimidade entre S_i-S_j e R_i-R_j . Na segunda lista, da tarefa difícil, S_i-R_i e S_i-R_j eram muito e pouco associados respectivamente, S_i-S_j podia ter alta ou baixa sinonimidade enquanto que R_i-R_j tinha baixo nível de sinonimidade.

A primeira lista foi construída selecionando-se 15 pares de palavras S-R não-competitivos. A segunda lista tinha 12 pares de palavras e foi formada a partir de 4 pares de palavras escolhidos na primeira. Para cada estímulo original (S_i) foram achados dois sinônimos (S_i' e S_i''). Cada novo estímulo foi emparelhado a uma resposta de baixa associatividade (R_i' e R_i'') criando uma tendência inicial dos sujeitos associarem novos estímulos à resposta do estímulo original (Figura 1).

Em ambos experimentos os pares de palavras foram apresentados aos indivíduos apenas por três vezes. Os resultados experimentais médios¹ sugeriam que: (i) o aprendizado inicial foi em torno de 50% (20%); (ii) o aprendizado ao final das 5 primeiras tentativas ficou em torno de 80%

(40%); (iii) os sujeitos responderam com mais de 90% de acerto após 11 (22) tentativas.

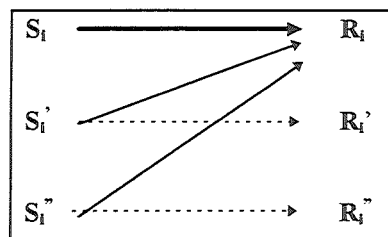


Figura 1: Associação entre pares competitivos de palavras. O par S_i-R_i é fortemente associado, enquanto que $S_i'-R_i'$ e $S_i''-R_i''$ são pares fracamente associados. Todos os estímulos são sinônimos, logo todos eles estão inicialmente associados à resposta original.

Para se construir um modelo de redes neurais, as palavras foram representadas por vetores binários no espaço $\mathcal{R}^{60}$. Dois vetores tinham nível máximo de diferença se a distância de Hamming deles era de 30 bits². A similaridade entre os vetores aumentava à medida que esta distância diminuía. Portanto, na tarefa fácil todos os vetores (S_i e R_i) distavam em torno de 30 bits uns dos outros. Na tarefa difícil os estímulos sinônimos (S_i , S_i' e S_i'') tinham suas distâncias encurtadas, enquanto que as distâncias dos estímulos originais (S_i) entre si e as respostas (R_i , R_i' e R_i'') entre si ficaram em torno de 30 bits.

O modelo "delta rule" de redes neurais foi empregado com sucesso para resolução do problema de associação clássica (Rescola, 1988). Contudo, quando o problema passa a ser associação de pares de palavras com níveis de dificuldade distintos, os resultados dos testes descritos em Araújo (1994) sugeriram que "delta rule" é inadequado. A nova perspectiva passou então a ser o emprego de modelos de memória na linha da rede de Hopfield (1982). Dentre tais opções o modelo chamado Memória Associativa Bidirecional (BAM) tem características adequadas para sua aplicação neste tipo de problema. Este modelo lida com tarefas heteroassociativas, interferência e é sempre convergente.

¹Os resultados das tarefas fácil e difícil estão fora e dentro dos parênteses respectivamente.

² Em BAM as operações de resgate de um dado vetor e seu bit-complementar são equivalentes.

3. Descrição de BAM

BAM modela memória de conteúdo endereçável, heteroassociativa que lida com pares de vetores. BAM compõe-se de duas camadas total e bidirecionalmente conectadas (Figura 2).

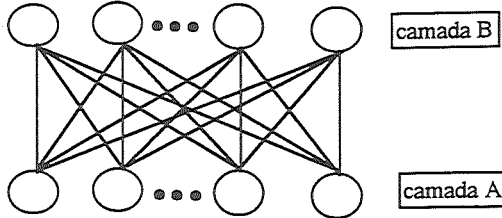


Figura 2: Diagrama esquemático do modelo BAM original.

Na entrada, a camada A tem m unidades e a camada de saída, a camada B, tem n unidades.

Cada padrão de entrada é representado por:

$$A_k = \{a_1, a_2, \dots, a_m\}, \quad k = 1, \dots, p \quad (1);$$

Cada padrão de saída é representado por:

$$B_k = \{b_1, b_2, \dots, b_n\}, \quad k = 1, \dots, p \quad (2);$$

onde p : é o número de pares de padrões.

As duas camadas estão acopladas através de uma matriz de pesos, a matriz correlação (W), de dimensão $m \times n$. Cada elemento desta matriz (w_{ij}) representa o valor da conexão entre o elemento i da camada A e a unidade j da camada B. O valor de w_{ij} é positivo se a conexão é excitatória e negativo se ela é inibitória. A matriz de correlação é construída como se segue:

$$W = \sum_{k=1}^p X_k^T Y_k \quad (3)$$

onde X_k e Y_k são as versões bipolares dos vetores A_k e B_k respectivamente.

O treinamento resume-se a formação da matriz correlação a partir de uma apresentação única dos p pares de padrões (A_k, B_k).

A informação flui no modelo de acordo com a regra de propagação abaixo:

$$\text{Camada B: } J_j = \sum_{i=1}^m w_{ij} a_i \quad (4.a)$$

$$\text{Camada A: } I_i = \sum_{j=1}^n w_{ij} b_j \quad (4.b)$$

A representação é binária e segue a regra de ativação abaixo:

Para uma unidade b_j na camada B:

$$b_j(t) = \begin{cases} 1, & \text{se } J_j > T_j \\ 0, & \text{se } J_j < T_j \\ b_j(t-1), & \text{se } J_j = T_j \end{cases} \quad (5.a)$$

Para uma unidade a_i na camada A:

$$a_i(t) = \begin{cases} 1, & \text{se } I_i > S_i \\ 0, & \text{se } I_i < S_i \\ a_i(t-1), & \text{se } I_i = S_i \end{cases} \quad (5.b)$$

Os valores T_j e S_i são "thresholds".

Na fase de testes, são apresentadas versões com ou sem ruído (α_{ok}) dos vetores de entrada (A_k) para se obter o par de vetores anteriormente "ensinado" à rede. Esta apresentação dispara a seguinte sequência de estados na rede:

$$\begin{aligned} \alpha_{ok} &\rightarrow W \rightarrow \beta_{ok} \\ \alpha_{1k} &\leftarrow W^T \leftarrow \beta_{ok} \\ \alpha_{1k} &\rightarrow W \rightarrow \beta_{1k} \\ &\quad M \\ \alpha_{hk} &\rightarrow W \rightarrow \beta_{hk} \\ \alpha_{hk} &\leftarrow W^T \leftarrow \beta_{hk} \end{aligned} \quad (6)$$

Inicialmente a camada A assume o estado de ativação do padrão de entrada, α_{ok} , este é propagado para a camada B através da matriz correlação, gerando o estado β_{ok} na camada B. β_{ok} é propagado na direção da camada A através da transposta da matriz correlação, W^T , produzindo α_{1k} . Esta dinâmica dura até o par de vetores (α_{hk}, β_{hk}) passar a se repetir, convergindo o processo. Idealmente, o par (α_{hk}, β_{hk}) é o par desejado (A_k, B_k). Isto nem sempre acontece embora a rede sempre convirja (Kosko, 1992). Logo a rede pode convergir para pares de atratores espúrios que aparecem devido à modificação das bacias de atração. Nelas são formados novos mínimos locais, diferentes daqueles gerados pelos pares treinados, a partir da apresentação de novos padrões à rede.

A falta de métodos analíticos gerais adequados para se estudar estas bacias de atração implicam na proposição de procedimentos heurísticos para redução do aparecimento destes atratores. Esta diminuição acarreta melhoria no desempenho de BAM que deve ser considerado no modelamento da tarefa cognitiva atual.

4. Hipóteses de Aprendizagem

Para modelar a melhora de performance dos indivíduos durante a fase de respostas nos experimentos originais Araújo (1994) testou adaptações da matriz correlação de BAM através da Regra de Hebb e Regra Delta. Os resultados dos testes sugeriram que tal sistemática não funciona para BAM. Em seguida, foi proposto a modificação da matriz de correlação a partir de outras hipóteses cognitivas ou neuro-cognitivas.

Deste modo, o desempenho da rede seria melhorado devido à percepção ou codificação de padrões com ruído, “desaprendizagem” de pares erradamente associados ou a troca do processo puramente heteroassociativo por outro com uma fase heteroassociativa que simula hiperpolarização ou despolarização da célula seguida por uma fase autoassociativa. As quatro hipóteses são:

- Sujeitos perceberiam versões com ruídos dos pares de padrões. Os vários ruídos aleatórios em diferentes versões dos padrões deveriam cancelar uns aos outros com o armazenamento das várias entradas.

- Os pares de padrões seriam percebidos de maneira perfeita pelos sujeitos, mas ruído seria adicionado na memorização dos padrões devido ao conhecimento prévio dos sujeitos. A influência das perturbações deveria desaparecer durante o estágio de testes.

- Sujeitos “desaprenderiam” associações equivocadas a partir de realimentações do meio. Tal mecanismo foi originalmente proposto para a rede de Hopfield (Hopfield e outros, 1983).

- Sujeitos aprenderiam devido a mecanismo que varia valor do “threshold” de cada unidade na direção da diminuição do erro de BAM, modelo biased-IBAM (Araújo, 1994). Esta variação permitiu um melhor resultado inicial de BAM, aperfeiçoado por um estágio autoassociativo.

5. Mudanças no Algoritmo BAM Original

As alternativas acima foram implementadas como é mostrado a seguir.

Na primeira hipótese, o ruído é adquirido durante a percepção, gerando a nova matriz correlação (W) abaixo:

$$W = \sum_{k=1}^P (X_k^T + N_X^T)(Y_k + N_Y) \quad (7)$$

onde N_X e N_Y são ruídos.

No estágio de teste, o produto externo de cada par com ruído foi somado a matriz correlação, após a recuperação de cada par, como a seguir:

$$\Delta W_k = (X_k^T + N_X^T)(Y_k + N_Y) \quad (8)$$

Na segunda hipótese, o ruído é adicionado durante a “memorização” dos padrões de treinamento da seguinte maneira:

$$W = \sum_{k=1}^P (X_k^T Y_k) + N_{XY} \quad (9)$$

onde N_{XY} é o ruído adicionado.

Na fase de teste, a matriz correlação variou depois de cada coleta dos pares como a seguir:

$$\Delta W_k = (X_k^T Y_k) \quad (10)$$

Na terceira hipótese os pares recuperados incorretamente da memória são “esquecidos” pela matriz correlação (3) como a seguir:

$$\Delta W_k = -(1/(2 \times p)) \times (S_k^T R_k) \quad (11)$$

onde S_k e R_k são atratores espúrios e $1/(2 \times p)$ é a taxa de aprendizagem.

Na quarta hipótese ocorre dois tipos de mudanças. A função de propagação passa a ser:

$$J_j = (1-k_1) \left(\sum_{i=1}^m w_{ij} a_i(t) + V_j(t) \right) + k_1 \left(\sum_{l=1}^n w_{lj}^B b_l(t-1) \right) \quad (12.a)$$

$$I_i = (1-k_1) \left(\sum_{j=1}^n w_{ij} b_j(t) + V_i(t) \right) + k_1 \left(\sum_{l=1}^m w_{il}^A a_l(t-1) \right) \quad (12.b)$$

onde o inteiro k_1 é zero durante o estágio heteroassociativo e “um” durante o autoassociativo, $V_j(t) = \alpha(t) y_j(t-1)$; $V_i(t) = \alpha(t) x_i(t-1)$; e $\alpha(t)$ é uma função crescente no tempo.

As matrizes de correlação para a fase autoassociativa (W_A e W_B) são:

$$W_A = \sum_{k=1}^P X_k^T X_k \quad (13.a)$$

$$W_B = \sum_{k=1}^P Y_k^T Y_k \quad (13.b).$$

Os resultados obtidos com estas modificações são discutidos na próxima seção.

6. Simulações

Esta seção relata os principais resultados dos experimentos realizados para testar os modelos propostos³. Empregou-se uma rede com 60 unidades em cada camada. Para as três primeiras hipóteses, 18 conjuntos distintos de 12 pares de vetores de diferentes graus de dificuldade foram usados. Na quarta opção, a mais promissora, 23 conjuntos de 12 pares foram empregados. O nível de ruído escolhido, 5% em todos os casos, foi compatível com a performance dos sujeitos nos experimentos originais.

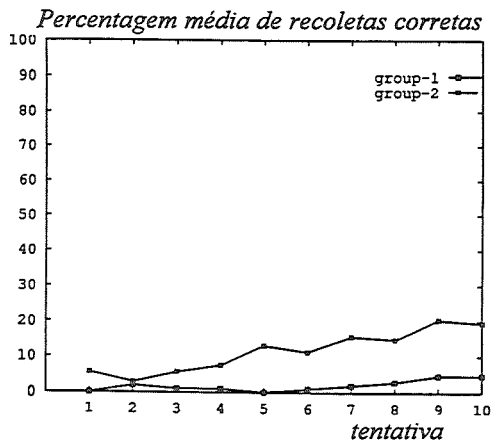


Figura 3: Pares corretamente recuperados no caso percepção com ruído.

Na primeira hipótese (Figura 3) o grupo-1 (0-23 bits de diferença) tem performance inferior à do grupo-2 (23-27 bits distintos)

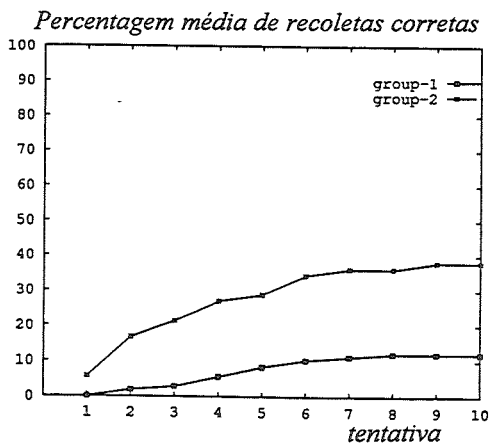


Figura 4: Pares corretamente coletados no caso armazenamento com ruído.

Nos experimentos relativo à segunda hipótese (Figura 4) os resultados obtidos mostraram progresso relativo ao caso anterior.

Os resultados da desaprendizagem de atratores espúrios são plotados na Figura 5.

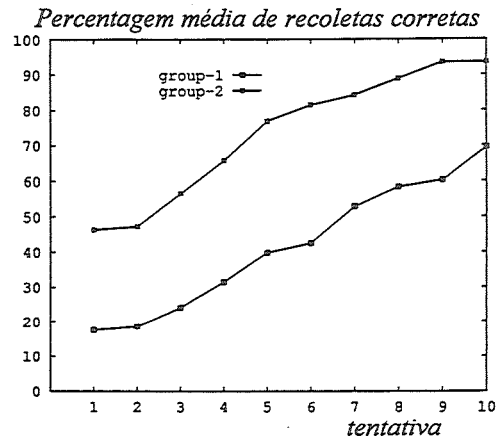


Figura 5: Pares corretamente recuperados no caso desaprendizagem de atratores espúrios.

A Figura 6 esboça os resultados do modelo decorrente da quarta hipótese. Quatro níveis de dificuldade foram considerados: leve, moderado, difícil e muito difícil que tinham, respectivamente, 24-27, 19-23, 15-18, 1-14 bits diferentes.

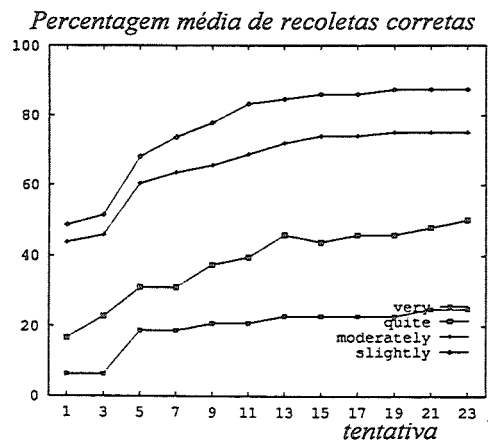


Figura 6: Pares corretamente recuperados no modelo biased-IBAM.

Em todos os casos os conjuntos de pares foram apresentados só uma vez.

Os resultados sugerem que houve melhora gradual na capacidade dos modelos de reproduzir o efeito da dificuldade na tarefa heteroassociativa. A próxima seção discute estes resultados.

³ Para resultados complementares consulte Araújo (1994).

7. Discussão

Nesta seção compara-se os resultados obtidos com aqueles desejados, baseados no efeito da dificuldade de associação de palavras para seres humanos.

Os resultados plotados na Figura 3 sugerem que a performance da rede cresce a cada tentativa, melhorando em cerca de 10% após 10 tentativas. A partir deste ponto seu crescimento é marginal. Os aprendizados inicial e final são baixos e não há crescimento rápido de desempenho nas tentativas iniciais.

Os resultados expostos na Figura 4 sugerem que o desempenho da rede cresce a cada tentativa, melhorando em cerca de 22% após 10 tentativas. A partir deste ponto seu crescimento é marginal. Os aprendizados inicial e final são baixos mas existe um crescimento mais rápido de desempenho nas tentativas iniciais que nas finais.

A Figura 5 mostra resultados sugerindo que a recuperação de pares corretos cresce, passo a passo, quase 54% após 10 tentativas. A partir deste ponto, o crescimento rápido inicial se torna mais lento. O nível de aprendizagem inicial é alto, mais alto do que o desejado.

Os resultados da Figura 6 sugerem que o modelo apresenta melhor aprendizagem em tentativas iniciais que nas finais, a performance melhora a cada tentativa em todos os níveis de dificuldade, o desempenho piora gradativamente para cada tentativa à medida que aumenta a dificuldade, a aprendizagem depois da vigésima primeira tentativa é apenas marginal, a aprendizagem final não é perfeita e o passo de aprendizagem é semelhante ao dos sujeitos.

Os resultados discutidos ilustram que na primeira estratégia a única característica de desempenho desejada é seu crescimento a cada nova tentativa. A segunda hipótese melhora o aprendizado final e cresce a taxa de aprendizagem inicial. A terceira hipótese mostra um nível significativo de aprendizagem inicial e final e bom aumento de performance que é mais rápido inicialmente. Contudo esta hipótese ainda mostrou uma taxa de aprendizagem e um número inicial de recuperação correta mais altos que o desejável. Estes dois problemas podem ser resolvidos misturando-se a segunda e terceira hipóteses e diminuindo-se o valor da taxa de aprendizagem no algoritmo. A quarta hipótese atende a maior parte dos requisitos levantados anteriormente. Ela mostra um desempenho final

um pouco inferior ao desejado que pode ser melhorado aumentando-se a dimensão da rede.

Os resultados das simulações sugerem que os dois últimos modelos são mais apropriados para replicar o efeito da dificuldade de associação de palavras para os humanos.

Referências

- Araújo, A. F. R. (1994). Memory, Emotions, and Neural Networks: Associative Learning and Memory Recall Influenced by Affective Evaluation and Task Difficulty. *PhD Thesis*. School of Cognitive and Computing Sciences, University of Sussex, Brighton, UK.
- Baddeley, A. D. (1990). *Human Memory: Theory and Practice*. Lawrence Erlbaum, Hove, UK.
- Eysenck, M. J. and Eysenck, M. W. (1985). *Personality and Individual Differences. A Natural Science Approach*. Plenum Press, New York, USA.
- Hopfield, J. J. (1982). Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554-2558.
- Hopfield, J. J., Feinstein, D. I., and Palmer, R. G. (1983). Unlearning has a Stabilizing Effect in Collective Memories. *Nature*, 304.
- Kohonen, T. (1989). *Self-Organization and Associative Memory*. Springer-Verlag, Berlin, Germany.
- Kosko, B. (1988). Bidirectional Associative Memories. In *IEEE Transactions on Systems Man and Cybernetics*, 18, pp 49-60.
- Kosko, B. (1992). *Neural Networks and Fuzzy Systems. A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, New Jersey, USA.
- Rescola, R. A (1988). Pavlovian Conditioning: It's not what you think it is. *American Psychologist*, 43.
- Spence, J. T. and Spence, K. W. (1966). Anxiety and Behaviour. In Spielberger, C. D. (Ed.), *The Motivational Components of Manifest Anxiety: Drive and Drive Stimuli*, pp 291-326. Academic Press, London.
- Zajonc, R. B. (1980). Feeling and Thinking: Preferences Need no Inferences. *American Psychologist*, 35 (2).