

**Dinâmicas Emergentes na Família de Memórias
Associativas Bidirecionais Caóticas e sua
Habilidade para Saltar Passos**

Luciana Pavani de Paula Bueno

Tese apresentada à Escola de Engenharia de
São Carlos da Universidade de São Paulo,
como parte dos requisitos para obtenção do
título de Doutor em Engenharia Elétrica.

Orientador: Prof. Dr. Aluizio Fausto Ribeiro Araújo

São Carlos
2006

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Àquele que é poderoso para vos guardar de tropeços e para vos apresentar com exultação, imaculados diante da sua glória, ao único Deus, nosso Salvador, mediante Jesus Cristo, Senhor nosso, glória, majestade, império e soberania, antes de todas as eras, e agora, e por todos os séculos. Amém.

Judas 24-25

Agradecimentos

- Ao Prof. Dr. Aluizio Fausto Ribeiro Araújo, pelo tratamento amigo, pela dedicação, pelas sugestões e pelo incentivo durante toda a orientação deste trabalho;
- Aos professores e funcionários do Departamento de Engenharia Elétrica - EESC/USP, pelo conhecimento, atenção e apoio recebidos;
- À Profa. Dra. Maria Eunice Quilici Gonzales, pelo encorajamento para realização do trabalho de doutorado;
- Aos meus pais, João Manoel e Guiomar, pelo amor, apoio e auxílio incondicionais;
- À minha família, por todo o apoio;
- Ao Cleber Zanchettin, pelo carinho, incentivo e companheirismo;
- Aos amigos preciosos da 1ª IPI de São José do Rio Preto e da 1ª e 2ª IPI de Recife, pela confiança, apoio e carinho;
- Aos amigos do LASI, Guilherme, Aline, Karla, José Carlos, Arthur, e de outros laboratórios, Kátia, Patrícia, Renata, Ricardo, Augusto, João Bosco, Natache, Alessandra, Edmárcio, Delbem, Mário, e todos os outros que não foram aqui mencionados, pela convivência enriquecedora, pelo apoio e pelos momentos de diversão;
- Ao Centro de Informática (CIn) da Universidade Federal de Pernambuco, pelo apoio recebido durante o período em que lá estive realizando trabalhos com o Prof. Dr. Aluizio F. R. Araújo;
- Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo suporte financeiro fornecido para a realização deste trabalho.

Sumário

Sumário	iii
Lista de Figuras	vi
Lista de Tabelas	ix
Resumo	xi
Abstract	xiii
Capítulo 1	1
Introdução	1
1.1 Objetivo	3
1.2 Contribuições desta Tese	3
1.3 Organização Geral deste Documento	4
Capítulo 2	7
Caos: Ocorrência, Conceitos, Medidas, Controle e Presença em Redes Neurais Artificiais	7
2.1 Caos em Sistemas Naturais e Artificiais	9
2.2 Fenômenos Caóticos	13
2.2.1 Dimensões e Fractais	13
2.2.2 Atratores	16
2.2.3 Bifurcação e Rotas para o Caos	18
2.2.4 Medidas de Caos	20
2.3 Papel do Caos em Redes Neurais Artificiais	24

2.4	Controle de Caos	27
2.4.1	O Controle de Caos	27
2.4.2	Como Controlar o Caos	28
2.4.2.1	Controlando um Processo Caótico para Melhorar seu Desempenho (Meta 1)	28
2.4.2.2	Controlando um Processo Caótico para um Estado Desejado (Meta 2)	29
2.4.2.3	Controlando um Processo Caótico para Prevenir uma Região Indesejável no Espaço de Fase (Meta 3)	31
2.5	Resumo	32
Capítulo 3		33
Caos, Modelos Auto-associativos Caóticos e Fenômeno Saltar Passos: Fundamentos		
3.1	Caos em Modelos Auto-associativos	34
3.1.1	Modelo Neurônio Caótico	34
3.1.2	Aplicação do Modelo Neurônio Caótico em Memórias Auto-associativas	39
3.2	Controle de Caos para Modelos Auto-associativos	43
3.3	Saltar Passos Transforma um Noviço em um Especialista	46
3.3.1	Caracterização do Saltar Passos	46
3.3.2	Explicações do Saltar Passos	48
3.3.3	Discussões sobre o Saltar Passos	51
3.4	Resumo	54
Capítulo 4		56
Família de Modelos de Memória Associativa Bidirecional Caótica (Família C-BAM)		
<hr/>		
4.1	Modelos Originais de Memória Associativa Bidirecional	58
4.1.1	Modelo Memória Associativa Bidirecional Discreto (BAM)	58
4.1.2	Modelo Memória Associativa Bidirecional Discreto com Atraso	61
4.1.3	Modelo Memória Associativa Bidirecional Discreto Exponencial	62
4.2	Modelos de Memória Associativa Bidirecional Caótica (C-BAM)	63
4.2.1	Diferenças entre BAM e C-BAM	63
4.2.2	C-BAM	66
4.2.3	C-BAM com atraso	67
4.2.4	C-eBAM	68
4.3	Resumo	68
Capítulo 5		70
Experimentos com a Família de Modelos de Memória Associativa Bidirecional Caótica		
5.1	Comportamentos Dinâmicos da Família C-BAM	71
5.1.1	Planejamento de Experimentos	71
5.1.2	Comportamentos Emergentes na Família C-BAM	74
5.2	Expoentes de Lyapunov	82
5.2.1	Expoentes de Lyapunov na Família C-BAM	85
5.3	Recuperação de Memórias Não Acessíveis	91

5.4	Modelagem do Fenômeno Saltar Passos	95
5.4.1	Representação	96
5.4.2	Testes	99
5.5	Resumo	107
Capítulo 6		109
Controle de Caos na Família C-BAM		
6.1	Controle de Caos para a Recuperação de Memórias Inacessíveis	110
6.1.1	Controle de Caos para a Rede C-BAM	111
6.1.2	Controle de Caos para a Rede C-BAM com atraso	114
6.1.3	Controle de Caos para a Rede C-eBAM	116
6.1.4	Discussão da Estratégia de Controle de Caos Proposta para a Recuperação de Memórias Inacessíveis na Família C-BAM	118
6.2	Controle de Caos para o Fenômeno Saltar Passos	119
6.2.1	Método de Controle por Pinagem	122
6.2.2	Método de Controle Adaptativo	124
6.2.2.1	Modelos da Planta Controlada	125
6.2.2.2	Sistema Adaptativo Modelo-Referência (MRAS)	127
6.2.2.3	A Regra MIT	128
6.2.2.4	Controle de Caos da Rede C-BAM com o Esquema Adaptativo MRAS	130
6.2.2.5	Discussão da Estratégia de Controle de Caos Proposta para a Modelagem do Fenômeno Saltar Passos	134
6.3	Resumo	135
Capítulo 7		137
Conclusões e Trabalhos Futuros		
7.1	Desenvolvimento e Limitação da Família C-BAM	137
7.2	Resumo das Contribuições Científicas da Tese	142
7.3	Propostas para Novos Trabalhos	144
Referências Bibliográficas		147
Apêndice A		156
Experimento Fatorial		156
A.1.	Experimentos Fatoriais	156
A.2.	Aplicação de Experimentos Fatoriais	160

Lista de Figuras

Figura 2.1. Ilustração do experimento realizado por Shaw (1984).	10
Figura 2.2. Esquema da rede neural artificial apresentada por Wang (1991).	11
Figura 2.3. Diagrama de bifurcação do neurônio x da rede de Wang (1991), com a matriz de pesos da Equação 2.3.	12
Figura 2.4. Diagrama de bifurcação do neurônio y da rede de Wang (1991), com a matriz de pesos da Equação 2.3.	12
Figura 2.5. Ilustração do conjunto de Cantor.	14
Figura 2.6. Ilustração da aplicação da dimensão fractal para o conjunto de Cantor.	15
Figura 2.7. Diagrama de bifurcação mostrando as mudanças qualitativas na resposta de um sistema dinâmico devido à variação do parâmetro μ no mapa logístico $x_{k+1} = \mu x_k (1-x_k)$. O sistema converge para zero, para um valor constante, para valores periódicos depois de bifurcações até atingir o caos dependendo do valor de μ .	19
Figura 2.8. Ilustração da divergência de duas órbitas vizinhas. A dinâmica caótica leva duas curvas inicialmente próximas a distanciarem-se.	21
Figura 3.1. Modelo do neurônio caótico como um elemento da rede neural caótica.	37
Figura 3.2. Padrões armazenados utilizados nos testes com a rede neural auto-associativa contendo neurônios caóticos.	40
Figura 3.3. Exemplo de seqüência de padrões de saída da rede de memória auto-associativa caótica, com capacidade de refração relativamente fraca, em que $k_r=0.5$, $k_f=0.15$, $\rho=8.0$ e $a_i=2.0$ ($i=1,\dots,100$).	41
Figura 3.4. Exemplo de seqüência de padrões de saída da rede de memória auto-associativa caótica, com capacidade de refração alta, em que $k_r=0.8$, $k_f=0.2$, $\rho=8.0$ e $a_i=2.0$ ($i=1,\dots,100$). O padrão inicial selecionado é o padrão armazenado X^3 , da Figura 3.2.	42
Figura 4.1. Topologia do modelo BAM.	58
Figura 4.2. Representação do i -ésimo neurônio da rede BAM original.	64
Figura 4.3. Representação do i -ésimo neurônio da rede C-BAM.	64
Figura 4.4. Influência do tempo na força de refração da rede C-BAM.	65
Figura 4.5. Representação da neurodinâmica caótica da rede C-BAM. O ponto central representa um ponto fixo. O parâmetro k_f representa o efeito de estabilização devido às entradas de realimentação da rede. O parâmetro k_r representa o efeito de desestabilização devido à refração ou à inibição auto-recorrente.	65
Figura 4.6. Arquitetura C-BAM. A atividade na unidade de entrada i é determinada por $x_i(t)$, e a atividade na unidade de saída j é determinada por $y_j(t)$.	66
Figura 5.1. Pares de padrões armazenados (X^k, Y^k) , $k=1,\dots,4$, usados nos experimentos com a família C-BAM.	74
Figura 5.2. Distância Hamming (dH) para C-BAM, em que $k_r=0.85$, $k_f=0.2$, $\rho=8.0$ e $a=2.0$. Para o cálculo da distância Hamming entre o padrão de saída e o padrão armazenado ($dH(Y(t), Y^k)$) foi considerada a diferença entre os bits da seqüência numérica da saída da rede no tempo t ($Y(t)$) e o padrão Y^k armazenado ($k=1,\dots,4$). A cada bit diferente entre $Y(t)$ e Y^k (neurônio a neurônio da rede), foi adicionado o valor um na dH entre eles.	76

- Figura 5.3.** Distância Hamming para C-BAM com atraso, em que $k_r=0.8$, $k_f=0.15$, $\rho=8.0$, $a=2.0$ e $\sigma=5$. 76
- Figura 5.4.** Distância Hamming para C-eBAM, em que $k_r=0.9$, $k_f=0.2$, $\rho=10.0$, $a=2.0$ e $\beta=1.3$. 77
- Figura 5.5.** Seqüência de pares recuperados na C-BAM: $(\mathbf{X}^3, \mathbf{Y}^1)$ e $(\mathbf{X}^1, \mathbf{Y}^3)$. Período 2 ocorre do instante $t=7$ em diante (até $t=30$ nesta figura). Valores de parâmetros utilizados: $k_r=0.2$, $k_f=0.15$, $\rho=8.0$, e $a=2.0$. 78
- Figura 5.6.** Seqüência de pares recuperados na C-BAM. À esquerda, na legenda, é mostrada a seqüência binária de pares de padrões recuperados pela rede (retângulos pretos representam valor 1, enquanto os brancos representam valor 0). À direita da figura, cada par binário recuperado é transformado em seu número decimal equivalente, representado pelos pontos. Pode ser notada a ocorrência de comportamento não periódico dentro do intervalo $t=192$ a $t=211$. Valores de parâmetros utilizados: $k_r=0.8$, $k_f=0.15$, $\rho=8.0$, e $a=2.0$. 79
- Figura 5.7.** Recuperação dos pares associados treinados na C-BAM com atraso para os parâmetros $k_r=0.92$, $k_f=0.46$, $\rho=4.0$, $a=2.0$ e $\sigma=3$. Neste caso, para o cálculo da distância Hamming, $dH\{\mathbf{X}(t), \mathbf{Y}(t), (\mathbf{X}^k, \mathbf{Y}^k)\}$ foi considerada a diferença entre os bits da seqüência numérica binária do par associado encontrado no tempo t $(\mathbf{X}(t), \mathbf{Y}(t))$ e do par associado armazenado $(\mathbf{X}^k, \mathbf{Y}^k)$, $k=1, \dots, 4$. A cada bit diferente entre $(\mathbf{X}(t), \mathbf{Y}(t))$ e $(\mathbf{X}^k, \mathbf{Y}^k)$, foi adicionado o valor um na distância Hamming entre eles. 81
- Figura 5.8.** Limites entre comportamentos periódicos e caóticos de cada rede da família C-BAM, de acordo com o aumento do parâmetro k_r . Os outros parâmetros foram mantidos fixos em $k_f=0.3$, $\rho=9.0$, $a=2.0$, $\sigma=5$ (para C-BAM com atraso) e $\beta=1.3$ (para C-eBAM). As barras pretas indicam comportamento periódico, enquanto as barras horizontais indicam comportamento caótico. 82
- Figura 5.9.** Os pares de vetores $(\mathbf{X}^k, \mathbf{Y}^k)$, $k=1, \dots, 4$, usados nos experimentos com os modelos C-BAM. 93
- Figura 5.10.** Os quatros padrões utilizados na modelagem do saltar passos da equação algébrica $-\mathbf{X}-\mathbf{A}=\mathbf{B}$. 98
- Figura 5.11.** Os seis padrões armazenados utilizados na resolução da equação algébrica $-\mathbf{A}/\mathbf{B}-\mathbf{X}=-\mathbf{C}$ para mostrar a modelagem do fenômeno saltar passos. 99
- Figura 5.12.** Distância Hamming entre os pares recuperados e o par $(\mathbf{P}^1, \mathbf{P}^2)$ para $k_r=k_f=\rho=a=0$ e padrão inicial \mathbf{P}^1 . Período um (ponto fixo) ocorreu do instante $t=1$ em diante. 100
- Figura 5.13.** Distância Hamming entre os pares recuperados $(\mathbf{X}(t), \mathbf{Y}(t))$ e o pares associados armazenados $(\mathbf{P}^k, \mathbf{P}^{k+1})$, $k=1, \dots, 5$, para $k_r=0.995$, $k_f=0.25$, $\rho=8.0$ e $a=3.0$. Comportamento não periódico foi encontrado. 101
- Figura 5.14.** Seqüência de recuperação dos pares de padrões armazenados da Figura 5.11 para $k_r=0.9975$, $k_f=0.35$, $\rho=9.0$ e $a=4.0$. 102
- Figura 5.15.** Seqüência de recuperação dos pares de padrões armazenados para $k_r=0.9974$, $k_f=0.35$, $\rho=9.0$ e $a=4.0$. 103
- Figura 5.16.** Seqüência de recuperação dos pares de padrões armazenados para $k_r=0.9973$, $k_f=0.35$, $\rho=9.0$ e $a=4.0$. 104
- Figura 5.17.** Seqüência de recuperação dos pares de padrões armazenados para $k_r=0.9972$, $k_f=0.35$, $\rho=9.0$ e $a=4.0$. 104
- Figura 5.18.** Seqüência de recuperação dos pares de padrões armazenados para $k_r=0.998$, $k_f=0.7$, $\rho=5.0$ e $a=3.0$. 105
- Figura 5.19.** Seqüência de recuperação dos pares de padrões armazenados da Figura 5.10 para $k_r=0.999$, $k_f=0.7$, $\rho=5.0$ e $a=3.0$. 106
- Figura 6.1.** Esquema do método de controle aplicada na camada Y da rede C-BAM. 111
- Figura 6.2.** Distância Hamming entre a saída da camada Y $(\mathbf{Y}(t))$ da rede C-BAM e o padrão armazenado \mathbf{Y}^1 , no intervalo de tempo $[0, 50]$, para $k_r=k_f=\rho=a=0$ e $\mathbf{X}(t=0)=\mathbf{X}^3$. Pode ser aqui verificado que, a partir do padrão inicial \mathbf{X}^3 , a C-BAM sem caos não consegue recuperar o par associado \mathbf{Y}^3 . Ao invés de \mathbf{Y}^3 a rede recupera \mathbf{Y}^1 . 112
- Figura 6.3.** Distância Hamming entre a saída da camada Y $(\mathbf{Y}(t))$ da rede C-BAM e o padrão armazenado \mathbf{Y}^3 , de $t=0$ a 3000, para $k_r=0.95$, $k_f=0.25$, $\rho=10$, $a=4.0$ e $\mathbf{X}(t=0)=\mathbf{X}^3$. 113
- Figura 6.4.** Distância Hamming entre a saída da camada Y $(\mathbf{Y}(t))$ da rede C-BAM e o padrão armazenado \mathbf{Y}^3 , de $t=0$ a 50, para $k_r=0.95$, $k_f=0.25$, $\rho=10.0$, $a=4.0$, $K=20$ e $\mathbf{X}(t=0)=\mathbf{X}^3$. 113
- Figura 6.5.** Distância Hamming entre a saída da camada Y $(\mathbf{Y}(t))$ da rede C-BAM e o padrão armazenado \mathbf{Y}^1 , de $t=0$ a 3000, para $k_r=0.95$, $k_f=0.45$, $\rho=10.0$, $a=2.0$ e $\mathbf{X}(t=0)=\mathbf{X}^1$. 114
- Figura 6.6.** Distância Hamming entre a saída da camada Y $(\mathbf{Y}(t))$ da rede C-BAM e o padrão armazenado \mathbf{Y}^1 , de $t=0$ a 50, para $k_r=0.95$, $k_f=0.45$, $\rho=10.0$, $a=2.0$, $K=20$ e $\mathbf{X}(t=0)=\mathbf{X}^1$. 114
- Figura 6.7.** Distância Hamming entre a saída da camada Y $(\mathbf{Y}(t))$ da rede C-BAM com atraso e o padrão armazenado \mathbf{Y}^3 , de $t=0$ a 3000, para $k_r=0.96$, $k_f=0.45$, $\rho=9.0$, $a=2.0$, $\sigma=5$ e $\mathbf{X}(t=0)=\mathbf{X}^3$. 115

- Figura 6.8.** Distância Hamming entre a saída da camada Y ($Y(t)$) da rede C-BAM com atraso e o padrão armazenado Y^3 , de $t = 0$ a 50, para $k_r = 0.96$, $k_f = 0.45$, $\rho = 9.0$, $a = 2.0$, $\sigma = 5$, $K = 10$ e $X(t=0) = X^3$. _____ 116
- Figura 6.9.** Distância Hamming entre a saída da camada Y ($Y(t)$) da rede C-eBAM e o padrão armazenado Y^3 , de $t = 0$ a 3000, para $k_r = 0.9$, $k_f = 0.45$, $\rho = 8.0$, $a = 2.0$, $\beta = 1.3$ e $X(t=0) = X^3$. _____ 117
- Figura 6.10.** Distância Hamming entre a saída da camada Y ($Y(t)$) da rede C-eBAM e o padrão armazenado Y^3 , de $t = 0$ a 50, para $k_r = 0.9$, $k_f = 0.45$, $\rho = a = 0$, $\beta = 1.3$, $K = 10$ e $X(t=0) = X^3$. _____ 117
- Figura 6.11.** Seqüência de recuperação dos pares de padrões armazenados na C-BAM para o conjunto de padrões da Figura 5.10 e $k_r = 0.999$, $k_f = 0.7$, $\rho = 5.0$, $a = 3.0$. Pode ser notado que, após recuperar o par associado contendo o padrão final P^4 , no instante $t = 832$, a rede continua transitando por outros pares armazenados, sem se estabilizar no padrão desejado. _____ 121
- Figura 6.12.** Seqüência de recuperação dos pares de padrões armazenados na C-BAM para o conjunto de padrões da Figura 5.11 e $k_r = 0.9972$, $k_f = 0.35$, $\rho = 9.0$, $a = 4.0$. Pode ser notado que, após recuperar o par associado contendo o padrão final P^6 , no instante $t = 488$, a rede continua transitando por outros pares armazenados, sem se estabilizar no padrão desejado. _____ 121
- Figura 6.13.** Distância Hamming entre a saída da camada Y ($Y(t)$) da rede C-BAM e o padrão armazenado P^4 , de $t = 0$ a 50, para $k_r = 0.9$, $k_f = 0.7$, $\rho = 5$, $a = 3$, $I = 2$, $K = 3$ e padrão inicial P^1 . _____ 123
- Figura 6.14.** Diagrama de bloco de um sistema adaptativo. _____ 126
- Figura 6.15.** Diagrama de bloco de um sistema adaptativo modelo-referência (MRAS). _____ 127
- Figura 6.16.** Diagrama de bloco do MRAS com um ajuste paramétrico baseado na Regra MIT. _____ 129
- Figura 6.17.** Controlador adaptativo discreto com diagrama de bloco usando transformada-z. _____ 131
- Figura 6.18.** Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo (linha contínua) e a saída do modelo (linha pontilhada) é ilustrada para $k_f = 0.7$, $k_r = 0.9$, $\rho = 5$, $a = 3$, ganho de adaptação $\gamma = -2$, padrão inicial P^1 e $y_m = P^4$. _____ 132
- Figura 6.19.** Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo (linha contínua) e a saída do modelo (linha pontilhada) é ilustrada para $k_f = 0.7$, $k_r = 0.9$, $\rho = 5$, $a = 3$, ganho de adaptação $\gamma = -3$, padrão inicial P^1 e $y_m = P^4$. _____ 132
- Figura 6.20.** Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo e a saída do modelo é ilustrada para $k_f = 0.7$, $k_r = 0.9$, $\rho = 5$, $a = 3$, $\gamma = -4$, padrão inicial P^1 e $y_m = P^4$. Foi verificado que a dinâmica da rede não se modifica ao utilizar-se $\gamma > -4$. _____ 133
- Figura 6.21.** Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo e a saída do modelo é ilustrada para $k_f = 0.7$, $k_r = 0.9$, $\rho = 5$, $a = 3$, ganho de adaptação $\gamma = -1$, padrão inicial P^1 e $y_m = P^4$. _____ 133
- Figura 6.22.** Simulação do MRAS para a rede C-BAM. A distância Hamming entre a saída do processo e a saída do modelo é ilustrada para $k_f = 0.35$, $k_r = 0.9$, $\rho = 9$, $a = 4$, $y_m = P^6$ e ganho de adaptação $\gamma = -3$. _____ 134

Lista de Tabelas

Tabela 2.1. Relação entre comportamento de um sistema e os valores dos expoentes de Lyapunov. _____	22
Tabela 2.2. Resultado do cálculo do expoente de Lyapunov para diversos valores paramétricos do mapa logístico e seus respectivos comportamentos. _____	23
Tabela 2.3. Resumo da relação entre comportamento e valor dos expoentes de Lyapunov. _____	24
Tabela 3.1. Instantes de tempo de recuperação de padrões armazenados e seus inversos para apresentação de X^3 . _____	43
Tabela 3.2. Passos utilizados na resolução de uma equação algébrica. _____	47
Tabela 5.1. Resultados da tabela ANOVA (Montgomery e Runger, 2003) para o experimento fatorial com quatro fatores. _____	73
Tabela 5.2. Relação entre entradas da família C-BAM (X) e suas saídas (Y) para $k_r = k_f = \rho = a = 0$. _____	75
Tabela 5.3. Comportamento da C-BAM para k_r variando de 0 a 1 e $k_f = 0.15$, $\rho = 8.0$, $a = 2.0$. _____	78
Tabela 5.4. Crise na rede C-BAM com atraso, para k_f variando e $k_r = 0.65$, $\rho = 9.0$, $a = 2.0$ e $\sigma = 5$ fixos. _	80
Tabela 5.5. Derivadas parciais da matriz Jacobiana DF de F em $z(t)$. _____	88
Tabela 5.6. Expoentes de Lyapunov da rede C-BAM com os parâmetros $k_f = 0.2$, $k_r = 0.6$, $\rho = 8.0$ e $a = 2.0$. _____	89
Tabela 5.7. Expoentes de Lyapunov da rede C-BAM com os parâmetros $k_f = 0.3$, $k_r = 0.9$, $\rho = 10.0$ e $a = 2.0$. _____	90
Tabela 5.8. Expoentes de Lyapunov da rede C-BAM com atraso com os parâmetros $k_f = 0.1$, $k_r = 0.7$, $\rho = 8.0$, $a = 2.0$ e $\sigma = 3$. _____	90
Tabela 5.9. Expoentes de Lyapunov da rede C-eBAM com os parâmetros $k_f = 0.25$, $k_r = 0.7$, $\rho = 9.0$, $a = 3.0$ e $\beta = 1.3$. _____	91
Tabela 5.10. Relação entre as entradas (X) e as saídas (Y) das redes da família C-BAM para $k_r = k_f = \rho = a = 0$. _____	93
Tabela 5.11. Freqüência de recuperação dos padrões armazenados da Figura 5.10 na C-eBAM, para $k_r = 0.9$, $k_f = 0.45$, $\rho = 8.0$, $a = 2.0$ e $\beta = 1.3$. _____	94
Tabela 5.12. Freqüência de recuperação dos padrões armazenados da Figura 5.10 na C-BAM, para $k_r = 0.95$, $k_f = 0.45$, $\rho = 7.0$ e $a = 3.0$. _____	95
Tabela 5.13. Freqüência de recuperação dos padrões armazenados da Figura 5.10 na C-BAM com atraso, para $k_r = 0.9$, $k_f = 0.45$, $\rho = 8.0$, $a = 2.0$ e $\sigma = 5$. _____	95
Tabela 5.14. Passos utilizados na modelagem da resolução da equação algébrica exemplificada em Blessing e Anderson (1996) pela C-BAM. _____	97
Tabela 5.15. Passos utilizados na modelagem da resolução da equação algébrica $-A / B - X = -C$. _____	98
Tabela 5.16. Resumo com os exemplos de comportamentos interessantes alcançados nos experimentos com a família C-BAM. _____	107
Tabela 5.17. C-BAM para k_r variando e $k_f = 0.70$, $\rho = 5.0$ e $a = 3.0$. (P^1 a P^4); $k_f = 0.35$, $\rho = 9.0$ e $a = 4.0$ (P^1 a P^6). _____	108
Tabela 6.1. Relação entre as entradas e as saídas das redes da família C-BAM para $k_r = k_f = \rho = a = 0$ considerando-se os padrões algébricos da Figura 5.11. _____	120

Tabela 6.2. Relação entre as entradas e as saídas das redes da família C-BAM para $k_r = k_f = \rho = a = 0$ considerando-se os padrões algébricos da Figura 5.10. _____	120
Tabela 6.3. Correlação entre a distância de pinagem (I) e a intensidade mínima de controle (K) para C-BAM com os padrões da Figura 5.10 e $k_r=0.9$, $k_f=0.7$, $\rho=5$, $a=3$, padrão inicial \mathbf{P}^1 . _____	124
Tabela 6.4. Correlação entre a distância de pinagem (I) e a intensidade mínima de controle (K) para C-BAM com os padrões da Figura 5.11 e $k_r=0.9$, $k_f=0.35$, $\rho=9$, $a=4$, padrão inicial \mathbf{P}^1 . _____	124
Tabela A.1. Arranjo dos Dados para um Planejamento de Experimentos com Dois Fatores. _____	158
Tabela A.2. Método utilizado para converter um par associado da C-BAM (binário) em um número decimal, a fim de que ele seja utilizado como réplica no planejamento de experimentos. _____	161
Tabela A.3. Formatação do tamanho das réplicas de um experimento fatorial para o caso de período máximo 16. _____	162
Tabela A.4. Dados para um experimento fatorial com 2 fatores ($\rho = 2, 4$ e 8 ; $k_r = 0.3, 0.6$, e 0.7) realizado com a rede C-BAM. Os parâmetros k_f e a foram mantidos fixos. _____	163

Resumo

BUENO, L. P. P. (2006). **DINÂMICAS EMERGENTES NA FAMÍLIA DE MEMÓRIAS ASSOCIATIVAS BIDIRECIONAIS CAÓTICAS E SUA HABILIDADE PARA SALTAR PASSOS**. São Carlos, 2006, Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

Nesta tese, uma família de memórias associativas bidirecionais caóticas (família C-BAM) é proposta, implementada e testada com o objetivo de estender a relevância da presença e do estudo do fenômeno caótico a modelos de redes associativas. Na modelagem da família C-BAM, todos os neurônios da memória associativa bidirecional caótica (BAM), BAM com atraso e BAM exponencial (eBAM) foram substituídos por neurônios caóticos.

Cada parâmetro do neurônio caótico na família C-BAM tem sua influência estimada através do planejamento de experimentos, em diferentes dinâmicas. Com base no planejamento de experimentos, valores de parâmetros são selecionados a fim de ilustrar a emergência de comportamentos dinâmicos como bifurcação, caos determinístico e crise. A existência de dinâmicas caóticas é confirmada pelo cálculo dos expoentes de Lyapunov.

Experimentos empíricos mostraram que a dinâmica caótica modifica a acessibilidade à memória da família C-BAM. Ao invés de recuperar um único par, como a família BAM fazia, a versão caótica é capaz de gerar uma grande diversidade de padrões recuperados, envolvendo complexas transições entre os padrões armazenados, para algumas variações paramétricas. Tal comportamento permite à família C-BAM acessar padrões inacessíveis às redes BAMs originais.

Além disso, a nova acessibilidade à memória, na qual seqüências de recuperação (com diferentes tamanhos) compostas de padrões treinados e não treinados têm emergido, pode ser usada para modelar a habilidade de um indivíduo saltar passos na solução de uma tarefa. Esta tese seleciona a rede C-BAM para ilustrar que a seqüência de recuperação da rede pode modelar a habilidade de um noviço ou a habilidade de um especialista executar uma tarefa.

Embora a família C-BAM possa alcançar todos os padrões armazenados durante o comportamento caótico, ela não consegue convergir para um padrão específico. Duas estratégias de controle são propostas para permitir que as redes caóticas convirjam para a memória desejada: o método de controle por pinagem e um método de controle adaptativo. Conseqüentemente, os modelos C-BAM podem, de fato, realizar a hetero-associação de memórias antes inacessíveis, e a rede C-BAM pode estabilizar-se no estado final de uma tarefa, dado o primeiro estado.

Palavras-chave – Memórias associativas bidirecionais, dinâmicas caóticas, recuperação de memórias, fenômeno saltar passos, controle de caos.

Abstract

BUENO, L. P. P. (2006). **EMERGENT DYNAMICS IN FAMILY OF CHAOTIC BIDIRECTIONAL ASSOCIATIVE MEMORIES AND ITS ABILITY TO SKIP STEPS.** São Carlos, 2006, *Thesis (PhD)* – Escola de Engenharia de São Carlos, Universidade de São Paulo.

In this thesis, a family of bidirectional associative memories (C-BAM family) is proposed, implemented and tested to extend the study of chaotic phenomenon in associative models. In the C-BAM model, all the original neurons of bidirectional associative memory (BAM), BAM with delay and exponential BAM (eBAM) were substituted for chaotic neurons.

Based on the experimental design, values of C-BAM family parameters are set to illustrate the emergence of a diversity of dynamic behavior, such as bifurcation, deterministic chaos and crisis. The existence of the chaotic dynamics is confirmed by calculation of Lyapunov exponents.

Empiric experiments showed that the chaotic dynamics modifies the behavior of memory accessibility. Instead of recalling a single pair, as BAM did, its chaotic version yielded a wide diversity of recalled patterns, involving complex transitions via memorized patterns for some parametric variations. Hence, C-BAM family can access patterns that original BAM family cannot.

Moreover, the new way of memory accessibility, in which several recall sequences (with distinct sizes) composed of trained and nontrained patterns have emerged, can be used to model the ability of skipping steps by an individual in a task solution. This thesis selected C-BAM network to illustrate that the retrieval sequence can model the ability of a novice or the ability of an expert to execute a task. There are also illustrated cases in which a novice recall can be transformed into an expert recall through parametric variation.

Although C-BAM family can reach all stored patterns during the chaotic behavior, it can not converge towards a specific pattern, consequently a desired output is not produced. In this thesis, two control strategies are proposed in order to make the chaotic networks to converge towards the desired memory: the pinning control method and the adaptive control method. Consequently, the C-BAM models can effectively realize the correct heteroassociation to former non-accessible memories and the C-BAM network can quickly be stabilized in the final state of a task, given the first state.

Keywords – *Bidirectional associative memories, chaotic dynamics, retrieval of memories, step skipping phenomenon, chaos control.*

Capítulo 1

Introdução

Os fenômenos caóticos têm sido um dos tópicos de pesquisa de maior expansão nas últimas décadas. Embora a possibilidade de dinâmicas caóticas tenha sido demonstrada matematicamente, há pelo menos cem anos, por Hadamard e Poincaré (Aihara, 2002), o entendimento do caos determinístico e a conscientização de sua importância para uma ampla variedade de situações somente se intensificou dentro da última década.

Sistemas dinâmicos caóticos possuem muitas propriedades interessantes, tais como a não periodicidade, comportamentos complexos, dependência sensível às condições iniciais, estrutura fractal e imprevisibilidade a longo prazo. Além disso, o caos é um tópico interdisciplinar, estudado por especialistas de várias áreas e com muitas funções potencialmente úteis em cada uma delas. Há algum tempo, a complexidade e a aparente não previsibilidade do caos levaram a não estudá-lo. Contudo, mais recentemente, essa tendência tem se invertido, isto é, o entendimento e a utilização desse comportamento pseudo-aleatório revelaram sua importância nas ciências.

O caos tem sido encontrado, por exemplo, tanto experimentalmente quanto numericamente, em muitos circuitos elétricos e eletrônicos, tais como os circuitos de Chua (Chua et al., 1993). A implicação disso é que as dinâmicas caóticas de modelos matemáticos podem ser implementadas em sistemas do mundo real ao se usar circuitos elétricos e eletrônicos não lineares. Outro exemplo de existência de caos foi encontrado ao se examinar eletrofisiologicamente o comportamento de neurônios biológicos. Foi possível observar o fenômeno caótico em membranas nervosas tanto experimentalmente, com axônios de lula gigante (Aihara e Matsumoto, 1986), quanto numericamente, com as equações de Hodgkin-Huxley (Aihara e Matsumoto, 1983). De fato, as equações Hodgkin-Huxley podem descrever adequadamente as dinâmicas não lineares das membranas nervosas (Hodgkin-Huxley, 1952).

Em particular, modelos pioneiros de neurônios artificiais discretos no tempo, tais como os neurônios de McCulloch e Pitts (1943), e de modelos de neurônios analógicos como os de Rumelhart, Hinton e Williams (1986), utilizados na rede *back-propagation*, só apresentavam dinâmicas estáveis, normalmente periódicas. Esta ausência de comportamentos caóticos em modelos de redes neurais convencionais foi colocada por alguns autores (Ikeguchi, Hasegawa e Aihara, 1999) como uma crítica à modelagem artificial dos neurônios biológicos.

Mais recentemente, em especial a partir da década de 90, começaram a surgir na literatura relatos de pesquisas de redes neurais artificiais com dinâmicas caóticas. Em 1990, Aihara, Takabe e Toyoda propuseram um modelo, chamado modelo caótico de redes neurais, como uma extensão dos modelos de neurônios discretos e analógicos convencionais. Esta extensão foi realizada ao se considerar: (i) as propriedades fisiológicas de refração, (ii) uma curva estímulo-resposta contínua inerente a dinâmicas neuronais e (iii) somatórios espaço-temporais de entradas externas e de entradas realimentadas vindas de outros neurônios (Komuro e Aihara, 2001). Alguns artigos (Aihara, Takabe e Toyoda, 1990; Adachi e Aihara, 1997; Aihara, 2002) mostram que esse modelo de rede neural caótico gera um comportamento dinâmico de caos determinístico espaço-temporal no domínio de redes neurais artificiais.

Em 1997, Adachi e Aihara apresentaram uma aplicação da rede neural caótica para memórias auto-associativas e analisaram as características de recuperação da rede construída. Foi mostrado que a inclusão do neurônio caótico na rede de memória associativa convencional gera um comportamento curioso, no qual a rede recupera um padrão armazenado e escapa dele após alguns instantes de tempo. Em sua dinâmica, essa rede nunca se estabiliza em um ponto de equilíbrio, mas move-se não periodicamente ao redor dos padrões armazenados.

Nesta tese, é proposta uma extensão da rede neural caótica citada para uma família de memórias hetero-associativas, a chamada família de redes de memórias associativas bidirecionais caóticas (família C-BAM), ao se “caotizar” o neurônio nos seguintes modelos de redes hetero-associativas convencionais: BAM, BAM com atraso e BAM exponencial. Para este novo paradigma, foram analisadas as características de seu comportamento dinâmico e as propriedades de recuperação de seus padrões armazenados, sendo que os experimentos escolhidos para as análises sugeriram a emergência de diversos comportamentos dinâmicos.

A partir dos resultados observados experimentalmente, procurou-se formalizar a presença de caos determinístico nos modelos C-BAM, de modo a confirmar a existência de dinâmicas caóticas nesses modelos. Comprovada a existência de dinâmica caótica, buscou-se tratar da aplicabilidade de tal dinâmica. A princípio, o desafio foi o de mostrar a possibilidade de as redes C-BAM escaparem de atratores indesejáveis ocorridos em memórias dinâmicas do tipo de Hopfield e o de apresentar a habilidade de estas redes C-BAM gerarem uma nova forma de acessibilidade à

memória, capaz de lidar com problemas de recuperação de memórias até então não resolvidos pelas redes de memória associativas convencionais (problema da recuperação de memórias inacessíveis). A aplicação da dinâmica caótica foi também explorada em problemas envolvendo fenômenos psicológicos. Foi então considerado o possível impacto da dinâmica caótica das redes C-BAM com respeito à modelagem de fenômenos psicológicos, em particular com o fenômeno saltar passos.

Na última etapa deste trabalho foram propostos dois métodos de controle de caos para lidar com a rica dinâmica caótica da família C-BAM e permitir que as redes de memória associativas bidirecionais caóticas converjam para um padrão desejado. O método de controle por pinagem foi aplicado com o objetivo de permitir a convergência da família C-BAM para uma memória inacessível e para um padrão final, representando a solução da operação algébrica memorizada pela rede durante o fenômeno saltar passos. O método de controle adaptativo, implementado para a rede C-BAM, foi utilizado para que a rede se estabilizasse no padrão final na modelagem do saltar passos, dado o padrão inicial. Comparando-se com o método de controle por pinagem, o método de controle adaptativo diminui o número de parâmetros pré-determinados para a convergência da rede.

1.1 Objetivo

O objetivo desta tese é a proposição, a implementação, a análise e a utilização em ao menos um fenômeno cognitivo de uma família de redes de memória associativa bidirecional caótica (família C-BAM), através da inclusão do modelo neurônio caótico proposto por Adachi e Aihara (1997). A família C-BAM é testada e analisada com o objetivo de se buscar o entendimento do comportamento dinâmico emergente desta família de redes criadas. Pretende-se também mostrar a aplicabilidade das dinâmicas caóticas e a proposição de um método de controle de caos para lidar com as dinâmicas das redes de memória hetero-associativas caóticas, de modo a se destacar a relevância da presença e do estudo do fenômeno caótico em tais modelos de redes.

1.2 Contribuições desta Tese

Em especial, esta tese traz os seguintes avanços e contribuições para a ciência:

- (i) Criação de uma família de redes associativas bidirecionais caóticas, a família C-BAM. Este paradigma consiste em uma extensão das redes de memória auto-associativas propostas por Adachi e Aihara (1997) e, desse modo, pretende contribuir para um maior aprofundamento nos estudos das influências dos fenômenos caóticos em modelos de memórias associativas;

- (ii) Análise dos comportamentos dinâmicos emergentes apresentados pelas redes da família C-BAM, com atenção particular em fenômenos como o caos determinístico, bifurcação e crise;
- (iii) Demonstração da existência de caos nas redes da família C-BAM, através do cálculo dos expoentes de Lyapunov;
- (iv) Estudo da acessibilidade a todas as memórias treinadas e armazenadas pelas redes da família C-BAM;
- (v) Mostrar a aplicabilidade da dinâmica caótica no acesso a memórias não recuperáveis pelas redes BAM convencionais;
- (vi) Proposição da modelagem do fenômeno psicológico chamado saltar passos a partir da aplicabilidade da dinâmica caótica das redes C-BAM.
- (vii) Proposição de ao menos uma estratégia de controle de caos para lidar com a dinâmica caótica das redes C-BAM.

Parte desses tópicos geraram os seguintes artigos:

- Araújo, A.F.R, Bueno, L.P.P. e Campos, M.A. (2004). Emergent Behaviors in Bidirectional Associative Memory (BAM) Family with Chaotic Neurons. **Anais do 8th Brazilian Symposium on Neural Networks (SBRN 2004), São Luís – MA.**
- Bueno, L.P.P. e Araújo, A.F.R. (2005). Ability to Skip Steps Emerging from Chaotic Dynamics. **Proceedings of the International Joint Conference on Neural Networks (IJCNN 2005), Montreal - Canadá.**
- Bueno, L.P.P. e Araújo, A.F.R. (2006). Controlling Chaos in Chaotic Bidirectional Associative Memories. **Proceedings of the International Joint Conference on Neural Networks (IJCNN 2006), Vancouver - Canadá.**

1.3 Organização Geral deste Documento

Esta tese foi redigida de acordo com a seguinte estrutura: (i) apresentação do fenômeno caótico, (ii) busca de uma abordagem em redes neurais artificiais capaz de gerar dinâmicas caóticas; (iii) proposição, descrição e implementação da família C-BAM; (iv) análise dos comportamentos e propriedades emergentes desta família; (v) aplicação dos principais resultados

encontrados; (vi) proposição e implementação de métodos de controle de caos para a família C-BAM e (vii) avaliação dos principais resultados e sugestões de trabalhos futuros com memórias associativas caóticas. O conteúdo foi dividido em 7 capítulos, descritos brevemente a seguir:

- **CAPÍTULO 2: Caos: Ocorrência, Conceitos, Medidas, Controle e Presença em Redes Neurais Artificiais.**

Este capítulo tem como objetivo apresentar alguns princípios fundamentais de caos para a compreensão das análises do comportamento dinâmico caótico dentro do domínio das redes neurais artificiais. São apresentados os tópicos mais importantes da teoria do caos, visando sua posterior aplicação nas redes hetero-associativas caóticas a serem criadas. Estes tópicos incluem os conceitos de caos determinístico, atratores, bifurcação, crise, dimensões fractais e medidas de caos. São analisadas, no contexto de sistemas caóticos, as três metas de controle encontradas na literatura. O capítulo também oferece uma visão geral da ocorrência de caos em sistemas naturais e artificiais e do papel do caos em redes neurais artificiais, o que serviu de motivação para o uso de neurônios caóticos nas redes de memórias associativas bidirecionais.

- **CAPÍTULO 3: Caos, Modelos Auto-associativos e Fenômeno Saltar Passos: Fundamentos.**

Em sua primeira parte, este capítulo tem como objetivo descrever o modelo neurônio caótico (Aihara, Takabe e Toyoda, 1990), capaz de reproduzir o fenômeno caótico encontrado em experimentos com neurônios biológicos, e apresentar um modelo de memória auto-associativa construído com o modelo neurônio caótico. É também apresentado o método de controle por pinagem utilizado para controlar o modelo de memória auto-associativa caótica. Em sua última parte, este capítulo tem como objetivo apresentar os conceitos necessários para a compreensão do fenômeno psicológico chamado saltar passos (*step skipping*). Neste trabalho, o fenômeno saltar passos é modelado (Capítulo 5) por redes neurais artificiais de memória hetero-associativa bidirecionais com neurônios caóticos.

- **CAPÍTULO 4: Família de Modelos de Memória Associativa Bidirecional Caótica (Família C-BAM).**

Este capítulo tem como objetivo descrever a criação da família de modelos de memória hetero-associativa bidirecional caótica (família C-BAM). Inicialmente são descritos os modelos originais de memória hetero-associativa bidirecional selecionados: BAM, BAM com atraso e BAM exponencial (e-BAM). Em seguida, o capítulo introduz a diferença entre os neurônios utilizados nas redes BAM originais e os neurônios utilizados na criação das redes BAM caóticas. Por fim, o capítulo descreve cada rede da família C-BAM: C-BAM, C-BAM com atraso e a C-eBAM.

- **CAPÍTULO 5: Experimentos com a Família de Modelos de Memória Associativa Bidirecional Caótica.**

Este capítulo tem como objetivo apresentar os experimentos realizados com a família C-BAM e os resultados obtidos através deles. Primeiramente verificou-se, estatisticamente, a relevância de cada parâmetro dos neurônios da família C-BAM, através da aplicação do *planejamento fatorial de experimentos* (Montgomery, 1997). Por meio da variação dos diferentes parâmetros (k_r , k_f , ρ e a) da família C-BAM foi, então, verificada a presença de comportamentos dinâmicos emergentes na família C-BAM. Para a comprovação dos comportamentos caóticos nos modelos C-BAM sugeridos pelos resultados dos experimentos, foi mostrado o cálculo dos expoentes de Lyapunov. Nos experimentos realizados com a família C-BAM, foi também observado o processo e as propriedades de recuperação dos padrões armazenados nas redes. Por fim, foi possível mostrar como as C-BAMs são capazes de modelar o fenômeno saltar passos.

- **CAPÍTULO 6: Controle de Caos na Família C-BAM.**

Este capítulo tem como objetivo propor e implementar estratégias de controle de caos para o novo modo de acessibilidade a memórias armazenadas gerado pela família C-BAM. Dentro desse objetivo, o controle de caos foi proposto considerando-se dois casos distintos: (a) recuperação de memórias inacessíveis, através do método de controle por pinagem, e (b) o fenômeno saltar passos, através do método de controle por pinagem e do método de controle adaptativo. O esquema adaptativo proposto possui um controlador adaptativo do tipo modelo referência (MRAS), e o mecanismo utilizado para ajustar os parâmetros do MRAS foi a Regra MIT.

- **CAPÍTULO 7: Conclusões e Trabalhos Futuros.**

Este capítulo traz um resumo do que foi exposto ao longo deste documento, bem como diversas sugestões para continuação da pesquisa com os modelos de memórias associativas bidirecionais caóticas.

Um apêndice acompanha este documento. Tal apêndice contém a teoria do Planejamento Fatorial de Experimentos, utilizado na análise da relevância estatística de cada parâmetro presente nos neurônios caóticos da família C-BAM. Está também incluída uma ilustração detalhada da aplicação desta metodologia na rede C-BAM.

Capítulo 2

Caos: Ocorrência, Conceitos, Medidas, Controle e Presença em Redes Neurais Artificiais

O estudo moderno do caos, contando com uma estrutura de idéias articuladas, começou na década de 60, com a compreensão de que equações matemáticas muito simples podiam servir de modelo para uma série de sistemas complexos, com problemas muito difíceis de previsibilidade (Gleick, 1989). Desde então, novos e amplos debates sobre as relações de causa e efeito têm sido travados na tentativa de se entender um mundo sem lugar para o acaso, no qual o conceito deste, ligado aos conceitos de probabilidade e determinismo tornaram-se cada vez mais importantes na descrição do universo.

Antes do estudo de caos prevalecia o entendimento que um universo regido por leis simples de causa e efeito, descrito em termos da dinâmica clássica que privilegia trajetórias previsíveis, havia chegado ao fim. Com base em novos estudos, percebeu-se a presença de uma incrível e sutil ordem microscópica de relações, nas quais se presumia anteriormente que houvesse apenas a aleatoriedade. De fato, se fosse necessário resumir os acontecimentos aqui brevemente relatados, seria possível dizer que a barreira entre determinismo e acaso encontrava-se no chão. No entanto, é nesse ponto que se encontra a ‘nova ciência’, pois segundo Gleick (1989), ‘onde começa o caos, a ciência clássica termina’.

A descoberta que atemorizou e, ao mesmo tempo, fascinou os cientistas envolvidos nesta exploração científica era a de que sistemas que obedecem a leis imutáveis e precisas nem sempre atuam de formas previsíveis e regulares. Leis simples podem não produzir comportamentos simples, e, ainda, leis determinísticas podem produzir um comportamento que parece aleatório. Deste modo, a ordem pode gerar seu próprio tipo de caos, com sistemas simples e submetidos a leis

precisas evoluindo para resultados caóticos, enquanto que, em contrapartida, percebeu-se que fenômenos que parecem desestruturados e aleatórios, podem estar obedecendo a leis simples. Ainda como parte dessa percepção, foi verificado que sistemas aleatórios podem ser capturados por atratores bem definidos, passando a apresentar um comportamento regular.

Simplificadamente, pode-se dizer que esta capacidade de que equações simples possam gerar comportamento complexo, e, como será visto adiante, tão sensível à mensuração e parecendo aleatório, mas permitindo que se veja ordem e padrão onde só se tinha observado aleatoriedade, irregularidade e imprevisibilidade, é chamada de caos. Sendo assim, a ciência do caos estuda os fenômenos aparentemente irregulares e imprevisíveis que surgem na evolução temporal de muitos sistemas não-lineares, na busca de padrões escondidos e de leis simples que rejam os comportamentos complexos. A justificativa está em que, apesar da falta de regularidade, da não repetição de comportamentos passados, o sistema dinâmico caótico segue equações determinísticas, sendo que a fonte da irregularidade, para o caos, é parte da dinâmica intrínseca do sistema. Esse estudo tornou-se efetivamente plausível a partir da década de 60, devido ao suporte computacional disponível, que permitia capacidade gráfica e de processamento para encontrar padrões precisos em acontecimentos antes tidos como aleatórios.

Para lidar com esta nova ciência foi concebida, na década de 70, a teoria do caos, que é definida como a teoria matemática que trata de sistemas com comportamento imprevisível e aparentemente aleatório, embora sejam regidos por leis determinísticas. No cerne da teoria do caos está a ordem escondida em fatos que surgem, à primeira vista, desordenados.

Com as descobertas envolvendo caos, muitos cientistas entenderam que uma mudança de paradigma e um novo sentido para ciência seriam mais apropriados, e isso os conduziu para a ciência moderna, na qual a natureza é considerada complexa e não se limita a qualquer lei que transforme os fenômenos naturais em estáveis e determinados. Ao contrário de seu uso comum, portanto, o sentido matemático de caos não se caracteriza por desordem ou confusão, mas sim designa uma classe específica complexa de comportamento dinâmico. De acordo com Ott (2002), uma grande variedade de sistemas apresenta uma evolução complicada ao longo do tempo. Este comportamento complicado é conhecido como caos. Faure e Korn (2001) argumentam que não existe uma teoria dos fenômenos caóticos que seja completamente abrangente, mas sim uma junção de modelos teóricos, ferramentas matemáticas e técnicas experimentais.

O presente capítulo oferece uma visão geral da teoria do caos e das propriedades das trajetórias caóticas, caracterizando conceitos relacionados ao caos e procurando definir terminologias que serão utilizadas ao longo deste trabalho, tais como bifurcação e crise. Serão considerados apenas alguns fundamentos principais no estudo de caos, devido à amplitude do tema. O objetivo deste capítulo é apresentar os conceitos básicos necessários para a compreensão das

análises do comportamento caótico dentro do domínio das redes neurais artificiais que serão introduzidas nos capítulos a seguir.

A Seção 2.1 exemplifica a natureza multidisciplinar do caos e mostra que o comportamento caótico pode ser encontrado tanto em sistemas naturais quanto em sistemas artificiais. Na Seção 2.2 foram comentadas as idéias básicas dos principais termos relacionados ao caos, tais como atrator, bifurcação, crise, fractal, e expoentes de Lyapunov. A Seção 2.3 apresenta alguns dos papéis do caos em redes neurais artificiais, os quais serviram de motivação para o uso de neurônios caóticos nas redes de memórias associativas bidirecionais. Na Seção 2.4 são analisadas as três metas de controle que recebem atenção no contexto de sistemas caóticos, visando posterior proposição de métodos de controle de caos para a família de redes hetero-associativas caóticas. Um comentário final sobre o que foi desenvolvido neste Capítulo é introduzido na Seção 2.5.

2.1 Caos em Sistemas Naturais e Artificiais

Embora a existência das dinâmicas caóticas seja conhecida há muito tempo, há algumas décadas a sua importância para uma ampla variedade de áreas tornou-se estimada. De fato, o caos pode ser observado em muitos sistemas, tanto nos sistemas naturais quanto nos artificiais, de modo que alguns pesquisadores o consideram como um dos tópicos mais importantes dentro do estudo científico. Tal opinião provavelmente origina-se da natureza multidisciplinar do caos: há relatos de pesquisas afirmando a existência do caos na previsão do tempo, em cinéticas populacionais, em reações químicas, em lasers, em sistemas econômicos, em atividades elétricas do coração e do cérebro, em movimento de asteróides, e em muitos outros fenômenos (Williams, 1997). O objetivo desta Seção é apresentar alguns exemplos da ocorrência do comportamento caótico em sistemas naturais e artificiais.

Como um primeiro exemplo de movimento caótico será mostrado o experimento de Shaw (1984), ilustrado na Figura 2.1. Neste experimento, manteve-se um fluxo lento e permanente de gotas de água de uma torneira. Os instantes de tempo em que as gotas de água caíam da torneira eram registrados por um sensor. Desse modo, os dados do sistema consistiam de um conjunto de instantes de tempo discretos (t_1, t_2, \dots, t_n) nos quais as gotas eram observadas pelo sensor. A partir destes dados, foram montados intervalos de tempo entre as gotas: $\Delta t_n = t_{n+1} - t_n$. Quando a taxa de fluxo de água na torneira era suficientemente pequena, os intervalos de tempo Δt_n eram todos iguais. Com o aumento do fluxo, a seqüência de intervalos de tempo tornava-se periódica, com um pequeno intervalo Δt_a seguido por um intervalo maior Δt_b , gerando uma seqüência de intervalos de tempos na forma $\dots, \Delta t_a, \Delta t_b, \Delta t_a, \Delta t_b, \Delta t_a, \dots$. Esta seqüência é chamada de seqüência de período dois, uma vez que $\Delta t_n = \Delta t_{n+2}$. Conforme o fluxo aumentava, seqüências de períodos maiores eram

observadas até que, com um fluxo suficientemente grande, a seqüência $\Delta t_1, \Delta t_2, \Delta t_3, \dots$ não apresentava nenhuma regularidade aparente. Argumenta-se que esta seqüência irregular é devida à dinâmica caótica (Ott, 2002).

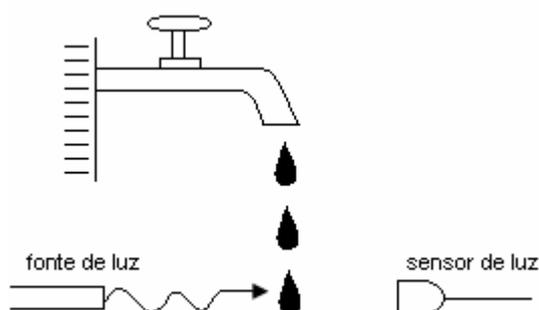


Figura 2.1. Ilustração do experimento realizado por Shaw (1984).

Há trabalhos interessantes apontando a presença de comportamento caótico em dinâmicas cerebrais. Um segundo exemplo mostra a presença de caos no estudo de atividades de neurônios biológicos. Freeman e colegas apresentam atividades caóticas relacionadas ao processo de aprendizagem no córtex olfatório de coelhos (Freeman, 1987, Eisenberg, Freeman e Burke, 1989, Yao e Freeman, 1990) e também no córtex visual (Freeman e Van Dijk, 1987). De acordo com estes estudos, a saída do bulbo olfatório é capaz de apresentar comportamento caótico. Na presença de um cheiro conhecido, o potencial olfativo mostra, primeiramente, um breve episódio de comportamento oscilatório de grande amplitude que, depois, converge para um comportamento quase-periódico. No caso em que o cheiro apresentado é desconhecido, a atividade neural torna-se caótica.

Babloyantz e colegas (1977, 1991, Destexhe, Sepulchre e Babloyantz, 1988) publicaram trabalhos apontando a presença de comportamentos caóticos em dinâmicas cerebrais. A presença de fenômenos caóticos foi detectada com o auxílio de eletroencefalogramas humanos realizados com relação a diferentes estados mentais. Foram observadas evidências de caos em estágios do sono, no ritmo alfa cortical, em pacientes apresentando a síndrome Creutzfeld-Jakob e em estados mentais caracterizados por grande atenção e concentração. Babloyantz e colegas argumentam que o caos aumenta a capacidade cerebral, enriquecendo suas respostas a estímulos externos. Basar e seus colegas (1980, 1988, 1989) também realizaram estudos importantes envolvendo a presença de comportamentos caóticos em atividades mentais relacionadas com o processo de memorização. Nestes estudos mostrou-se a presença de muitos atratores estranhos em sinais de eletroencefalogramas (EEG) realizados em experimentos mnemônicos. Basar considera a atividade cerebral base de processos cognitivos mnemônicos como sendo fundamentalmente caótica.

Um terceiro exemplo é incluído para ilustrar a existência de comportamento caótico em sistemas artificiais. Wang (1991) apresenta a prova analítica da existência de caos em uma rede neural artificial discreta no tempo que consiste de apenas dois neurônios, com função de ativação sigmoideal e nenhuma entrada externa ou atraso de tempo. Wang afirma que mesmo esta simples rede neural, com valores diferentes de conexões sinápticas, pode apresentar quase todos os tipos de comportamentos dinâmicos complexos encontrados na literatura para um sistema dinâmico unidimensional discreto no tempo. A rede neural de Wang é definida como uma rede de dois neurônios completamente conectada (Figura 2.2). Os estados dos dois neurônios são denotados como x e y , respectivamente, cujos valores estão no intervalo $I=[0,1]$, e o estado da rede é denotado como um vetor (x,y) no espaço de estado $I^2 = [0,1] \times [0,1]$. Os pesos da rede formam uma matriz 2×2 de valor real $\mathbf{W}=[w_{ij}]$. A atualização dos estados da rede em tempos discretos ($t = 0,1,2,\dots$) é feita de acordo com a equação:

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} \sigma_{\mu}(w_{11}x(t) + w_{12}y(t)) \\ \sigma_{\mu}(w_{21}x(t) + w_{22}y(t)) \end{pmatrix} \equiv \mathbf{F}_{\mu} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \quad (2.1)$$

em que:

$$\sigma_{\mu}(z) = \frac{1}{1 + e^{-\mu z}} \quad (2.2)$$

é a função de ativação de tipo sigmóide do neurônio, $\mu > 0$, e $\mathbf{F}_{\mu} = \sigma_{\mu} * \mathbf{W}$ é o mapa da rede com $\sigma_{\mu}(x,y) = (\sigma_{\mu}(x), \sigma_{\mu}(y))$.

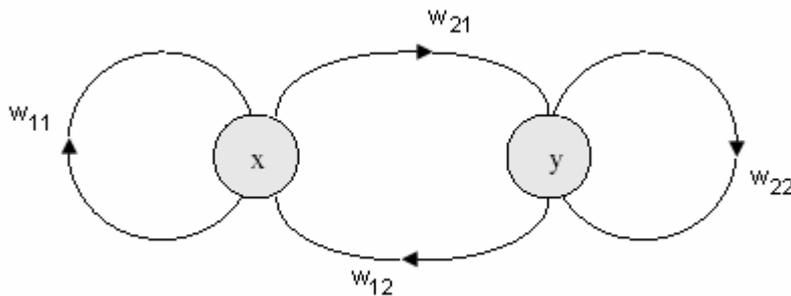


Figura 2.2. Esquema da rede neural artificial apresentada por Wang (1991).

No modelo de rede neural apresentado acima, o parâmetro μ possui um importante papel. Primeiramente porque ele controla a inclinação máxima da função $\sigma_{\mu}(z)$ e depois porque a mudança do valor de μ provoca modificações em todos pesos w_{ij} o que, conseqüentemente, afeta o comportamento dinâmico da rede. Para ilustrar o comportamento caótico da rede, foram realizadas simulações computacionais utilizando a seguinte matriz de pesos:

$$\mathbf{W} = \begin{bmatrix} -5 & 5 \\ -25 & 25 \end{bmatrix} \quad (2.3)$$

As Figuras 2.3 e 2.4 mostram os diagramas de bifurcação para os neurônios x e y , respectivamente, conforme o parâmetro μ varia de 0 a 6. Nestas figuras pode ser observada a ocorrência da rota de duplicação de período para o caos na rede apresentada por Wang (1991).

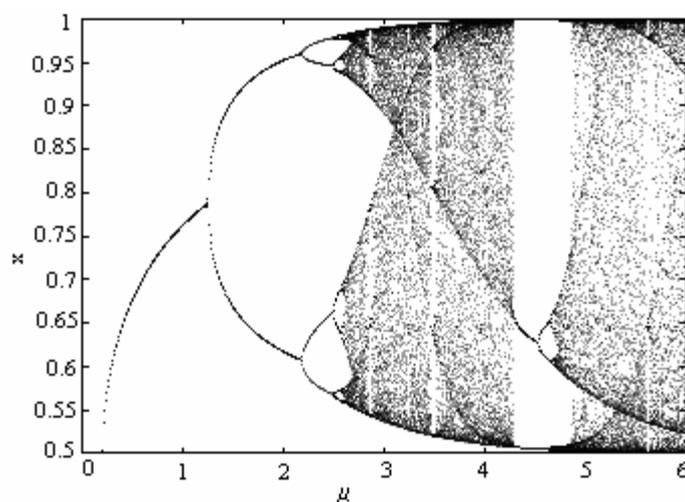


Figura 2.3. Diagrama de bifurcação do neurônio x da rede de Wang (1991), com a matriz de pesos da Equação 2.3.

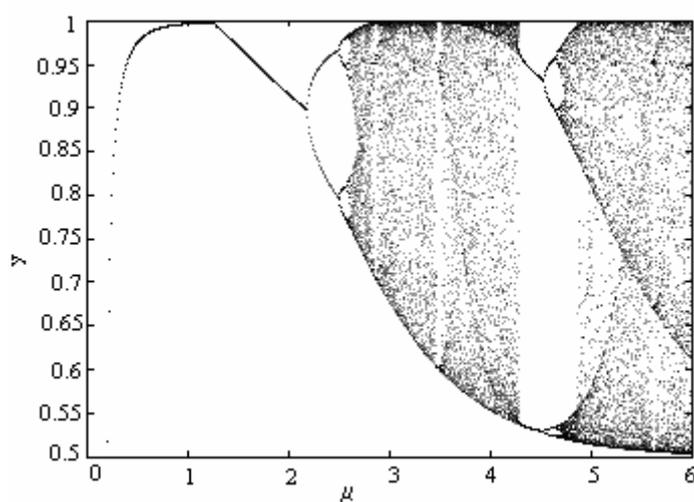


Figura 2.4. Diagrama de bifurcação do neurônio y da rede de Wang (1991), com a matriz de pesos da Equação 2.3.

Diante das evidências de caos observadas em diversos sistemas, Ford (1989) afirmou que “poucos observadores duvidam que o caos é encontrado por toda parte na natureza”. A aprovação de tal afirmação necessita de mais pesquisas. O fato é que o estudo de fenômenos caóticos continua a se desenvolver rapidamente e em muitas direções, e suas implicações continuam a crescer, o que torna o estudo do comportamento caótico um tópico interessante e relevante.

2.2 Fenômenos Caóticos

O caos, sendo uma nova ciência, gerou sua linguagem própria, contendo termos que podem expressar e ilustrar as características e a complexidade dessa descoberta. O objetivo desta seção é desenvolver a idéia básica de termos usados em caos e que serão utilizadas no contexto das redes neurais artificiais de memórias associativas bidirecionais caóticas. Mais detalhes sobre caos em sistemas dinâmicos podem ser encontradas em Ott (2002).

2.2.1 Dimensões e Fractais

Segundo Ott (2002), talvez a característica mais básica de um conjunto seja a sua *dimensão*. Um *espaço* pode ser definido como sendo uma coleção de entidades chamada *pontos*. Neste sentido, os pontos seriam quaisquer entidades subordinadas a algo superior chamado espaço. Por exemplo, pessoas são pontos no espaço das populações assim como telefones são pontos no espaço de uma rede de telecomunicações. Qualquer que seja o espaço imaginado, este apresenta sempre um número a ele associado, chamado de *dimensão*.

Faure e Korn (2001) argumentam que “uma perspectiva moderna generaliza a idéia de dimensão em termos de leis em escala”. De acordo com esta definição, ao considerar o espaço formado em um círculo de área πr^2 , tem-se que seus valores numéricos dependem da medida utilizada para r . Por exemplo, um círculo com raio de um metro apresentará área π quando r for medido em metros e área de $10^4 \pi$ quando r for medido em centímetros. Usar esta última definição de dimensão permite um modo de especificar uma nova classe de objetos geométricos chamada *fractais*. Foi o matemático Mandelbrot quem escolheu o termo *fractus*, do verbo *frangere* (fraturar, quebrar), que depois se tornou *fractal*, palavra que se enquadra bem com as inglesas *fracture* e *fraction*, e que, como será visto mais adiante, está relacionada com a dimensão apresentada pelos fractais.

Após a primordial característica do caos, que é a dependência às condições iniciais (Seção 2.2.4), pode-se considerar os fractais como sendo a segunda característica essencial do caos. Isso porque sistemas dinâmicos caóticos quase inevitavelmente resultam na formação de fractais, apesar de nem todos os fractais estarem associados ao caos. Também a análise de fractais é

freqüentemente útil em descrever a estrutura geométrica do atrator de um sistema dinâmico. De fato, os padrões do comportamento caótico estão relacionados com os observados na geometria fractal (Ott, 2002).

Basicamente, um fractal pode ser entendido como uma forma geométrica com uma propriedade essencial: a auto-similaridade. O objeto fractal possui uma estrutura similar em todas as extensões escalares. Diz-se que uma figura é auto-similar se pequenas porções dela assemelham-se com suas partes maiores, as quais, por sua vez, também assemelham-se ao objeto todo, de modo que se for cortada uma parte da figura e olhada como uma lente de aumento, será visto que ela se parece com o todo. Pode também se dizer que o todo reprisa a unidade mais elementar.

O exemplo possivelmente mais simples de fractal é uma forma auto-similar conhecida como conjunto de Cantor. A construção padrão desse tipo de fractal é feita considerando-se o conjunto $C = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$. A partir deste intervalo, devem ser divididos por três cada um deles e então removidos os intervalos internos resultantes dessa divisão. Esse processo é repetido diversas vezes de modo que, em cada estágio, é removida sempre a terça parte do meio de cada um dos intervalos que permaneceram no estágio anterior. Tal processo não é algo diferente do conhecido processo matemático chamado iteração de construções geométricas. A Figura 2.5 apresenta graficamente os três primeiros passos na construção do conjunto de Cantor, que são:

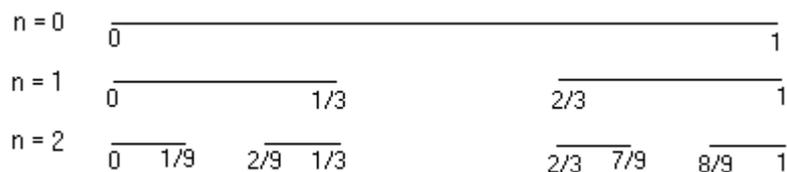


Figura 2.5. Ilustração do conjunto de Cantor.

Na Figura 2.5 fica evidente, no contexto do conjunto de Cantor, a auto-similaridade presente nos fractais: cada linha de dado estágio é similar à linha do estágio anterior na construção do conjunto. Além disto, será mostrado a seguir que o conjunto de Cantor tem uma dimensão não-inteira. Dimensões deste tipo (dimensões fracionárias) são chamadas *dimensões fractais*, de acordo com a definição dada por Mandelbrot (Ott, 2002).

Para a definição formal de dimensão fractal (Ferrara e Prado, 1995), considere um conjunto de pontos A num espaço de dimensão p . Estes pontos devem ser envolvidos por hiper-cubos (caixas) de lado ε . A dimensão fractal (ou dimensão de Hausdorff-Besicovitch) é definida como:

$$D = \lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)} \quad (2.4)$$

em que $N(\varepsilon)$ é o número mínimo de hiper-cubos de lado ε necessário para envolver todo o conjunto de pontos A . Esta proposta numérica para o cálculo da dimensão fractal é chamada método da contagem de caixas.

Para a aplicação desta definição ao conjunto de Cantor, pode ser observado na Figura 2.6 que $N(\varepsilon) \approx 2^n$, enquanto $\varepsilon \approx L/3^n$, conforme $\varepsilon \rightarrow \infty$.

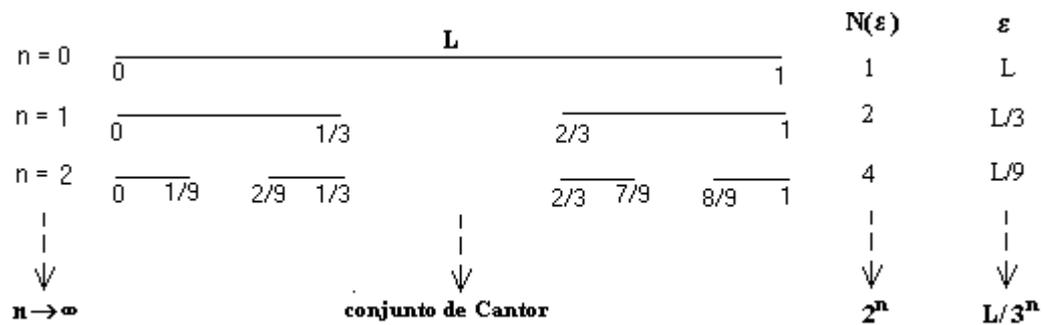


Figura 2.6. Ilustração da aplicação da dimensão fractal para o conjunto de Cantor.

Portanto, pode ser calculado que:

$$\begin{aligned}
 D &= \lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)} = \lim_{n \rightarrow \infty} \frac{\log 2^n}{\log(3^n/L)} = \left\langle \begin{array}{l} \text{para} \\ n \text{ grande} \end{array} \right\rangle = \frac{\log 2^n}{\log(3^n/L)} \\
 &= \frac{\log 2^n}{\log 3^n - \log L} = \frac{\log 2^n}{\log 3^n - \log 1} = \frac{\log 2^n}{\log 3^n} = \frac{n \log 2}{n \log 3} = \frac{\log 2}{\log 3} = 0.63 \quad (2.5)
 \end{aligned}$$

Encontrado o valor de 0.63 para a dimensão do conjunto de Cantor, pode-se confirmar que o conjunto de Cantor tem uma dimensão que não é um número inteiro, mas sim uma fração, o que indica que o conjunto é fractal.

Uma importante aplicação da análise da estrutura geométrica do sistema é possibilitar a classificação do nível de complexidade de um sistema através da análise da estrutura geométrica do atrator a ele associado. Portanto, a respeito do estudo de dimensão e fractais, vale enfatizar que ele fornece uma das importantes distinções entre tipos de dinâmicas de um sistema. Utilizando-se das definições aqui apresentadas, é possível afirmar que o cálculo da dimensão D fornece uma maneira relativamente simples de caracterizar se o sistema observado apresenta ou não comportamento caótico. O fator de diferenciação está no conjunto numérico em que D está contido (Ott, 2002). Em sistemas contínuos de comportamento caótico $D \in \mathfrak{R}$, enquanto que, em comportamentos não caóticos, a dimensão dos atratores é inteira ($D \in \mathcal{N}$).

Definições de fractal e dimensão foram incluídas nesta seção com o objetivo de fornecer subsídios para a definição de tipos de atratores que emergem na dinâmica das redes caóticas propostas nesta tese.

2.2.2 Atratores

De acordo com Williams (1997), o conceito de atrator é, possivelmente, o conceito mais fundamental na análise do caos, uma vez que caracteriza o comportamento e a variação no tempo dos sistemas. Segundo Ferrara e Prado (1995), um dos objetivos centrais no estudo dos atratores associados a sistemas dinâmicos é o de extrair informações a respeito da dinâmica característica do sistema.

Em sua essência, retendo-se a noção intuitiva, o termo atrator refere-se a um conjunto invariante S para o qual as órbitas que se coloquem em movimento em suas proximidades convergem para ele à medida que $t \rightarrow \infty$, como se ali existisse um campo gravitacional atraindo corpos em sua órbita. Na medida em que descrevem o comportamento de sistemas físicos para longos prazos, os atratores estão intimamente ligados à noção de estabilidade, isto é, a longo prazo, um sistema dinâmico geral estabiliza-se em um atrator. O conjunto de todas as possíveis condições iniciais que convergem para o mesmo atrator é chamado bacia de atração.

Esta seção comenta apenas os tipos de atratores diretamente ligados aos comportamentos encontrados na família de memória memórias associativas bidirecionais caóticas, proposta nesta tese.

O tipo mais simples de atrator, o *atrator ponto*, consiste em um único ponto fixo no espaço de fase. Ele representa os sistemas que tendem a convergir para um único estado com o passar do tempo. Uma vez que a condição de regime permanente é alcançada, o atrator do tipo ponto torna-se independente do tempo. Um exemplo de sistema com este tipo de atrator é o pêndulo, na situação em que ele se encontra parado por causa da fricção. Stewart (1991) afirma que os pontos fixos ou pontos de equilíbrio consistem de pontos singulares onde a evolução dos estados permanece estacionada.

Para uma definição formal de ponto fixo ou de equilíbrio, considere a equação de estado:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (2.6)$$

em que $\mathbf{x}(t) \in \mathfrak{R}^n$, $\mathbf{f}: \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ e $\mathbf{f}(\mathbf{x})$ não considera o tempo explicitamente. Os sistemas formados por equações diferenciais que são invariantes no tempo, tais como o sistema da Equação 2.6, são chamados *sistemas autônomos*. Além disso, a *solução* desta equação de estado será aqui denotada como sendo a trajetória $\phi_t(\mathbf{x}_0)$. Sendo assim, um *ponto de equilíbrio*, num sistema autônomo, é

definido como a solução constante da equação de estado dada acima, ou seja, $\phi_t(\mathbf{x}_{eq}) = \mathbf{x}_{eq}, \forall t, t \rightarrow \infty$, e $\mathbf{f}(\mathbf{x})=0$.

Outro tipo de atrator característico em sistemas dinâmicos é o *atrator periódico* (ou *ciclo limite*). Ele consiste de dois ou mais valores que se repetem na mesma ordem. Um ciclo limite descreve um sistema de periodicidade regular. A duração do ciclo ou do período define a periodicidade do atrator. Para a definição formal de ciclo limite, a equação de estado variante no tempo, usada para sistemas dinâmicos não autônomos é dada por:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (2.7)$$

Neste caso $\mathbf{f}(\mathbf{x})$ depende do tempo t . A *solução* desta equação de estado será aqui denotada como sendo a trajetória $\phi_t(\mathbf{x}_0, t_0)$. Caso exista um período $T > 0 \mid \mathbf{f}(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t+T), \forall \mathbf{x}, t$, a solução desse sistema é chamada solução periódica de período T , denotada por $\phi_t(\mathbf{x}^*, t_0) = \phi_{t+T}(\mathbf{x}^*, t_0)$. Em Mahla (1991) é encontrado que, no caso de um sistema autônomo, uma solução periódica isolada $\phi_t(\mathbf{x}^*)$ é chamada *ciclo limite*.

É uma característica das dinâmicas caóticas que o atrator resultante frequentemente apresente uma estrutura geométrica mais complexa no espaço de fase do que a estrutura encontrada nos tipos de atratores já citados. De fato, de acordo com a definição de dimensão vista na Seção 2.2.1, há casos em que o atrator de um sistema contínuo no tempo é um conjunto de dimensão inteira, como por exemplo dimensão zero (caso do ponto fixo), assim como existem atratores cujo valor para sua dimensão não é um número inteiro. Na terminologia de Mandelbrot, tais objetos geométricos são chamados fractais. E, quando um atrator é fractal, ele é chamado de *atrator estranho* (Ott, 2002).

De acordo com Ott (2002), um atrator estranho é aquele que possui estrutura fractal no espaço de fase, enquanto o *atrator caótico* é aquele em que as órbitas não periódicas e limitadas apresentam dependência das condições iniciais. Ou seja, o termo caos descreve a dinâmica do atrator, enquanto o termo ‘estranho’ refere-se à geometria do atrator. E, embora para a maioria dos casos envolvendo equações diferenciais atratores estranhos sejam também caóticos, há casos em que atratores caóticos não são estranhos.

O estudo de atratores ligados a sistemas caóticos pode também ser o estudo de suas propriedades estatísticas, o que permite tanto a descrição geométrica (via dimensão fractal) quanto a caracterização dinâmica (via expoentes de Lyapunov, entropia) dos atratores.

2.2.3 Bifurcação e Rotas para o Caos

Poincaré usou o termo *bifurcação* para descrever a divisão dos estados de um sistema dinâmico. Segundo Hilborn (1994), bifurcação significa uma divisão em dois, uma ruptura em fragmentos, uma mudança. Em uma bifurcação o que ocorre é a mudança da natureza qualitativa das soluções. O espaço de fase, como consequência, muda qualitativamente; novos pontos estacionários podem aparecer e outros anteriormente estáveis podem se tornar instáveis e vice-versa.

Um sistema dinâmico que descreve um sistema físico real depende de um ou mais *parâmetros*, de modo que o primeiro sistema pode ser descrito como sendo uma função de seus parâmetros. Faure e Korn (2001) afirmam que “sistemas físicos (assim como os dados modelados) podem passar por transições entre várias dinâmicas conforme alguns de seus parâmetros básicos são variados”. Neste contexto, as mudanças qualitativas no comportamento decorrentes das variações dos parâmetros são conhecidas como bifurcações (Ott, 2002). De fato, como poderá ser visto nos exemplos a seguir, em um sistema dinâmico, é comum a alteração da dinâmica do sistema observado no momento à medida que um ou mais parâmetros variam, de modo que pequenas alterações nos parâmetros do sistema podem levá-lo a se comportar de forma totalmente diferente. Os valores dos parâmetros para os quais ocorre a mudança são chamados de *pontos de bifurcação* ou *pontos críticos*. No espaço de fase, tais pontos de bifurcação são os marcos para a ocorrência das mudanças de comportamento do sistema.

É possível verificar o comportamento geral do sistema ao se traçar suas órbitas como sendo uma função de seus parâmetros. O diagrama gerado é chamado *diagrama de bifurcação*. Entre outros comportamentos dinâmicos, pode ser facilmente observado neste diagrama o fenômeno da *bifurcação de duplicação de período*. Este fenômeno, em particular, é um cenário de surgimento de órbitas estáveis cujo período é o dobro da anterior que, por sua vez, se torna instável. Mais formalmente, apresentando-se um período 2^{k-1} , um novo período 2^k surge a cada valor $\mu = \mu_k$ do parâmetro, à proporção que o valor de μ vai sendo aumentado. A seqüência de valores μ_k aproxima-se de um valor crítico μ_∞ , que, por sua vez, deve marcar o limite entre o comportamento regular e o caótico. Essa seqüência de bifurcações conduzindo ao regime caótico é conhecida como a *rota de duplicação de período para o caos*, como ilustrado na Figura 2.7.

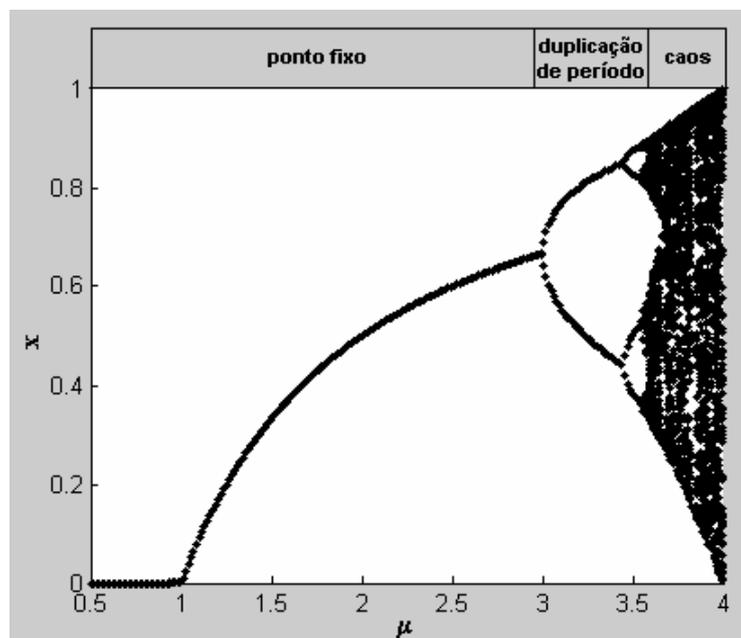


Figura 2.7. Diagrama de bifurcação mostrando as mudanças qualitativas na resposta de um sistema dinâmico devido à variação do parâmetro μ no mapa logístico $x_{k+1} = \mu x_k (1 - x_k)$. O sistema converge para zero, para um valor constante, para valores periódicos depois de bifurcações até atingir o caos dependendo do valor de μ .

Como pode ser notado na Figura 2.7, após ocorrer uma série de bifurcações, o sistema tornou-se caótico. O aparecimento de caos em sistemas dinâmicos está usualmente ligado às ocorrências de bifurcações de algum tipo. No entanto, um fato enfatizado por Araújo (1988) é que embora a bifurcação possa simplesmente coexistir com caos, e geralmente uma bifurcação precede uma situação caótica, não é necessária sua presença para que o caos ocorra.

Além da bifurcação do tipo duplicação de período, pelo fato do tema bifurcação ser bastante extenso e complexo, serão aqui apenas citadas algumas das bifurcações mais comuns: bifurcação de Hopf, associada a dois autovalores complexos que cruzam o eixo imaginário e a bifurcação simples, associada à divisão de um único estado estacionário estável em três estados estacionários. Existe ainda uma série de outras bifurcações, tais como bifurcação sela-nó, bifurcação transcritical, bifurcação de Neimark, cada uma com suas características e importâncias específicas. Mais detalhes sobre bifurcações podem ser encontrados em Hale e Koçak (1991).

Ferrara e Prado (1995) consideram como os três principais cenários de transição ordem-caos: o cenário de Ruelle-Taken, via quasi-periodicidade; o cenário de Geigenbauj, via duplicação de período; e o cenário de Pomeau-Manneville, via intermitência.

Existem ainda outras rotas possíveis para o caos (Ferrara e Prado, 1995). Uma dessas outras rotas possíveis para o caos é denominada *crise* (Ott, 2002). Segundo Ott, mudanças repentinas nos atratores caóticos geradas pela variação paramétrica do sistema são muito comuns. Tais mudanças

têm como causa a colisão do atrator caótico com uma órbita periódica instável. Primeiramente estudado por Grebogi, Ott e Yorke (1982), o fenômeno denominado crise foi por eles definido como “uma mudança abrupta e descontínua no atrator caótico, conforme um ou mais parâmetros são variados”. A crise é caracterizada pela destruição do atrator caótico ou pela sua expansão para um intervalo maior.

De acordo com Grebogi, Ott e Yorke (1983), transições entre ordem e caos são reversíveis, pelo menos dentro do domínio caótico; ou seja, em zonas críticas do parâmetro de controle (fronteiras das regiões de periodicidade dentro do caos), o caos pode ser criado ou destruído, dependendo do aumento ou diminuição do valor do parâmetro de controle.

2.2.4 Medidas de Caos

Apesar da sensibilidade às condições iniciais apresentada por sistemas caóticos, existem, segundo Faure e Korn (2001) métodos e medidas para reconhecimento e caracterização de sistemas caóticos que são invariantes, o que significa que “seus resultados não dependem do ponto inicial da trajetória no atrator, e nem das unidades usadas para definir as coordenadas no espaço de fase”.

Essas invariantes são possíveis graças à suposição de que os atratores caóticos satisfazem as condições da hipótese ergódica (de que as trajetórias estão sempre visitando as mesmas regiões próximas ao atrator). Na presente sub-seção será apresentada uma das medidas de caos considerada invariante: o expoente de Lyapunov.

Os *expoentes de Lyapunov* são utilizados para medir a taxa de divergência das trajetórias e assim poder quantificar a dependência sensível às condições iniciais. Segundo Ott (2002), considerando-se duas condições iniciais próximas $\mathbf{x}_1(0)$ e $\mathbf{x}_2(0) = \mathbf{x}_1(0) + \Delta(0)$, e admitindo-se que elas evoluam ao longo do tempo de modo a produzir as órbitas $\mathbf{x}_1(t)$ e $\mathbf{x}_2(t)$ (Figura 2.8), no instante de tempo t , a separação entre as duas órbitas é dada por $\Delta(t) = \mathbf{x}_2(t) - \mathbf{x}_1(t)$. Se em $|\Delta(0)| \rightarrow 0$ as órbitas permanecem limitadas e, para t grande, a diferença entre as soluções $|\Delta(t)|$ crescem exponencialmente para a orientação típica do vetor $\Delta(0)$ (ou seja, $|\Delta(t)|/|\Delta(0)| \sim \exp(ht)$, $h > 0$), então o sistema apresenta dependência sensível às condições iniciais e é, portanto, um sistema caótico.

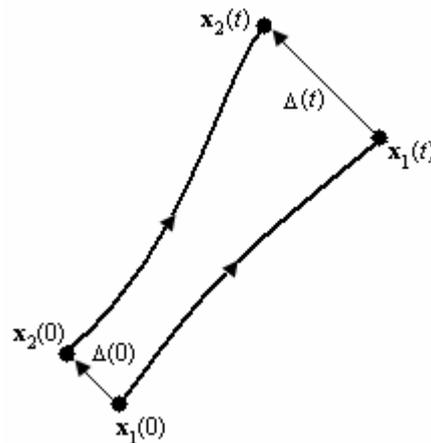


Figura 2.8. Ilustração da divergência de duas órbitas vizinhas. A dinâmica caótica leva duas curvas inicialmente próximas a distanciarem-se.

Para uma definição formal de expoentes de Lyapunov (Ott, 2002), seja considerado o caso do mapa \mathbf{M} . Seja \mathbf{x}_0 a condição inicial e \mathbf{x}_n ($n=0,1,2,\dots$) a órbita correspondente. Considerando-se um deslocamento infinitesimal de \mathbf{x}_0 na direção do vetor tangente \mathbf{y}_0 , a evolução do vetor tangente, dada pela Equação 2.8, determina a evolução do deslocamento infinitesimal da órbita a partir da órbita não perturbada \mathbf{x}_n .

$$\mathbf{y}_{n+1} = \mathbf{DM}(\mathbf{x}_n) * \mathbf{y}_n \quad (2.8)$$

em que \mathbf{DM} é a matriz Jacobiana de \mathbf{M} em \mathbf{x}_n .

Em particular, $\mathbf{y}_n / |\mathbf{y}_n|$ fornece a direção do deslocamento infinitesimal da órbita a partir de \mathbf{x}_n , e $|\mathbf{y}_n| / |\mathbf{y}_0|$ é o fator pelo qual o deslocamento infinitesimal cresce (se $|\mathbf{y}_n| > |\mathbf{y}_0|$) ou decresce (se $|\mathbf{y}_n| < |\mathbf{y}_0|$).

Da Equação 2.8 obtém-se:

$$\mathbf{y}_n = \mathbf{DM}^n(\mathbf{x}_0) * \mathbf{y}_0 \quad (2.9)$$

em que

$$\mathbf{DM}^n(\mathbf{x}_0) = \mathbf{DM}(\mathbf{x}_{n-1}) \cdot \mathbf{DM}(\mathbf{x}_{n-2}) \cdot \dots \cdot \mathbf{DM}(\mathbf{x}_0) \quad (2.10)$$

O expoente de Lyapunov, para a condição inicial \mathbf{x}_0 e para a orientação inicial do deslocamento inicial dada por $\mathbf{u}_0 = \mathbf{y}_0 / |\mathbf{y}_0|$, é definido como:

$$\begin{aligned}\lambda = h(\mathbf{x}_0, \mathbf{u}_0) &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left(\frac{|\mathbf{y}_n|}{|\mathbf{y}_0|} \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left| \mathbf{DM}^n(\mathbf{x}_0) \cdot \mathbf{u}_0 \right|\end{aligned}\quad (2.11)$$

Se a dimensão do mapa é N , então existirão N ou menos expoentes de Lyapunov (λ) distintos.

Não são muitos os casos em que se pode calcular os expoentes de Lyapunov analiticamente, pois o cálculo analítico de λ depende das particularidades do mapa em questão. Considerando-se o mapa $\mathbf{x}_{n+1} = \mathbf{F}(\mathbf{x}_n)$, tal cálculo só é factível quando é possível obter-se expressões analíticas simples para $\left(d\mathbf{F}^N / d\mathbf{x} \right)_{\mathbf{x}_0}$ ou para o produto $\prod_{i=1}^N \mathbf{F}'(\mathbf{x}_i)$. Em todos os outros casos deve-se recorrer ao cálculo numérico, sempre lembrando que existem diversas maneiras de se fazer a estimativa numérica dos expoentes de Lyapunov, quase todas variantes de um método geral, válido para fluxos e mapas, sugerido por Shimada e Nagashima (1979).

Diante do cálculo de λ é possível então caracterizar o comportamento do sistema, dando uma imagem das propriedades de um sistema que o levam à estabilidade ou à instabilidade. Se λ é negativo, significa que há um efeito de contração no espaço de fase de um atrator, os pontos próximos se aproximam, as trajetórias próximas convergem e a evolução não é caótica. Se λ é positivo, significa distensão, os pontos próximos se distanciam, trajetórias próximas divergem, e a evolução será caótica, e, portanto, sensível às condições iniciais. A Tabela 2.1 resume esta relação entre comportamento do sistema e o valor do expoente de Lyapunov.

Tabela 2.1. Relação entre comportamento de um sistema e os valores do expoentes de Lyapunov.

Comportamento	Expoente de Lyapunov
Não Caótico	$\lambda_i \leq 0$: quando não existe nenhum λ positivo num sistema dinâmico é porque a convergência do sistema não é sensível à condição inicial. Portanto, pontos inicialmente próximos continuam nessa condição ao longo da trajetória.
Caótico	$\lambda_i > 0$: quando existe algum λ positivo num sistema dinâmico, significa que a convergência do sistema é sensível à condição inicial, logo pontos inicialmente próximos podem separar-se exponencialmente ao longo da trajetória.

Resumidamente, pode-se dizer que os expoentes de Lyapunov fornecem informações valiosas a respeito do comportamento do sistema, valendo-se do fato de que o expoente de Lyapunov é: positivo nas regiões caóticas; negativo nas regiões periódicas; e nulo nos pontos de bifurcação.

Como exemplo, será apresentado o cálculo de expoentes de Lyapunov no mapa logístico. Em mapas unidimensionais, o cálculo do expoente de Lyapunov é imediato. Basta aplicar a equação:

$$\lambda_{x_0} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |F'(x_i)| \quad (2.12)$$

em que:

$x_i = F^i(x_0)$: resultado da i -ésima iteração do mapa F a partir de x_0 ;

$$F'(x_i) = (dF/dx)_{x_i}$$

Vale observar que, no limite ($n \rightarrow \infty$), o expoente de Lyapunov não depende mais da condição inicial x_0 .

Aplicando-se a equação acima no mapa logístico citada anteriormente, em que $x_{k+1} = \mu x_k (1 - x_k)$, o cálculo do expoente de Lyapunov será dado por:

$$\lambda \approx \frac{1}{n} \sum_{i=1}^n \ln |\mu - 2\mu x_i| \quad (2.13)$$

A Tabela 2.2 ilustra como o número chamado expoente de Lyapunov (λ) foi utilizado para determinar o comportamento do mapa logístico para diferentes valores de μ .

Tabela 2.2. Resultado do cálculo do expoente de Lyapunov para diversos valores paramétricos do mapa logístico e seus respectivos comportamentos. Nos pontos de bifurcação os valores dos expoentes de Lyapunov são próximos de zero, mas não exatamente zero, como consequência das aproximações computacionais. Deve também ser observada a correspondência entre o diagrama de bifurcação da Figura 2.7, que mostra as mudanças qualitativas no mapa logístico para determinado valor de μ , e o valor do expoente de Lyapunov apresentado nesta tabela.

μ	λ	Comportamento
1	$\lambda = -0.0872\dots$	Inicia-se um ponto fixo estável
3.001	$\lambda = -0.0173\dots$	Inicia-se um ciclo estável de período 2
3.4497	$\lambda = -0.0223\dots$	Inicia-se um ciclo estável de período 4
3.5699	$\lambda = 0.0197\dots$	Inicia-se o caos

A Tabela 2.3 procura resumir os diferentes tipos de comportamento apresentados por uma trajetória, baseando-se nos expoentes de Lyapunov.

Tabela 2.3. Resumo da relação entre comportamento e valor dos expoentes de Lyapunov.

Comportamento	Fluxo do Conjunto Atrator	Expoentes de Lyapunov
Ponto de equilíbrio	Ponto	$0 > \lambda_1 \geq \dots \geq \lambda_n$
Periódico	Curva fechada	$\lambda_1 = 0$ $0 > \lambda_2 \geq \dots \geq \lambda_n$
Período-k	Toróide	$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ $0 > \lambda_{k+1} \geq \dots \geq \lambda_n$
Caótico	Estranho	Ao menos um $\lambda_k > 0, k=1, \dots, n$

Após a compreensão dos conceitos básicos para a análise do comportamento caótico, será dada, a seguir, uma definição mais formal do fenômeno caótico, a qual garante condição suficiente para a existência do caos.

Definição 2.1

Sistema caótico é aquele que possui pelo menos um expoente de Lyapunov positivo.

Esta definição de caos é uma das mais utilizadas atualmente, e é possível pelo fato de a sensibilidade às condições iniciais poder ser quantificada. Como foi visto anteriormente, a característica essencial do caos é a sensibilidade às condições iniciais. Uma vez que esta dependência está associada a um fluxo expansivo, um atrator estranho terá pelo menos um expoente de Lyapunov positivo. Portanto, é possível definir um sistema caótico baseando-se no valor dos expoentes de Lyapunov (Mahla, 1991).

2.3 Papel do Caos em Redes Neurais Artificiais

Conforme citado anteriormente, o caos é uma teoria matemática recente. Semelhantemente, as redes neurais artificiais também constituem um tópico relativamente novo na literatura científica. Como consequência, o uso da teoria do caos no estudo e desenvolvimento de sistemas neurais artificiais é ainda mais recente.

Limitações da natureza estática das redes neurais artificiais convencionais têm inspirado a investigação de neurônios com dinâmicas mais complexas, tais como os neurônios caóticos e os neurônios *spiking*¹. Em redes neurais artificiais criadas com tais neurônios, o sistema não apresenta apenas convergência para pontos fixos, mas também para ciclos limites e atratores caóticos. Um grande número de pesquisadores acredita que o comportamento dinâmico caótico desempenha um papel importante tanto em neurônios reais quanto em redes neurais artificiais e tem procurado modelar redes neurais artificiais com dinâmicas caóticas através de equações diferenciais determinísticas.

De fato, é possível verificar que as redes neurais artificiais com dinâmicas convencionais frequentemente apresentam problemas do tipo: (a) convergência para atratores indesejáveis, (b) convergência muito lenta para os atratores, e (c) falhas na reprodução de atividades cerebrais (Potapov e Ali, 2001). Enquanto isso, a idéia de que as redes neurais artificiais podem trabalhar em regimes caóticos tem sido proposta em muitos artigos. Uma boa revisão destas abordagens pode ser encontrada em Tsuda (1992) e Freeman (1992).

Considerando-se que o caos em redes neurais artificiais é um tópico de recente interesse no meio acadêmico, o objetivo desta seção é apresentar alguns dos possíveis papéis das dinâmicas caóticas nesses tipos de redes.

O fato de o sistema caótico comportar-se de modo imprevisível pode ser utilizado na busca de uma nova maneira de solucionar problemas. Enquanto em arquiteturas de redes neurais artificiais convencionais os comportamentos costumam ser previsíveis, ou seja, para uma dada entrada a rede apresenta a mesma saída, em redes neurais caóticas o caos contribui para uma exploração de possibilidades de resultados. Redes neurais artificiais capazes de exploração durante a fase de aprendizagem têm sido pesquisadas. Por exemplo, têm sido feitas tentativas de se usar sinais caóticos e algoritmos para minimização aleatória durante a fase de aprendizagem de perceptron multicamadas (Potapov e Ali, 2001).

O fenômeno caótico também tem sido explorado na tentativa de melhorar o desempenho de sistemas neurais artificiais. Há trabalhos apontando a habilidade de redes neurais caóticas no reconhecimento de novos padrões, e também a habilidade de reconhecimento mais rápido dos padrões. Tan e Ali (1998), por exemplo, introduziram caos no modelo de Gardner (1988) ao reduzir o número de conexões entre os neurônios da rede. Os resultados mostraram que a rede neural caótica tornou-se mais eficaz na tarefa de reconhecer padrões. Os autores mostraram que os padrões podem ser reconhecidos no modelo caótico criado através do controle de caos realizado na história da evolução dos estados da rede. Utilizando o controle de caos e o critério de sincronização

¹ Redes de neurônios *spiking* possuem entradas e saídas temporalmente codificadas (Maass, 1997).

do tempo, a informação escondida na história da evolução foi explorada e o processo de reconhecimento de padrão realizado com eficiência.

Talvez uma das melhores indicações de que o caos pode ser utilizado na prática em sistemas neurais artificiais esteja no desempenho de um sistema artificial caótico projetado para reconhecer opticamente quatro tipos diferentes de peças industriais e determinar se elas eram defeituosas. Quando este sistema neural caótico foi comparado com as implementações feitas para o mesmo tipo de tarefa com o sistema neural convencional, os resultados mostraram que a performance do sistema neural caótico teve desempenho significativamente superior em identificar corretamente as peças aceitáveis e inaceitáveis (Yao *et al.*, 1991).

Ainda, experimentos têm sido realizados na tentativa de se utilizar a instabilidade local do fenômeno caótico como uma ferramenta para o processo de busca pela memória. Esta idéia envolve usar a trajetória de excursão entre imagens geradas pela rede caótica para ligar os dados de entrada com as imagens armazenadas. A instabilidade local também representa uma propriedade útil do caos para evitar trajetórias que conduzam o sistema a locais indesejados no espaço de fase (Potapov e Ali, 2001).

Um outro importante papel do caos em redes neurais artificiais que merece destaque encontra-se relacionado com a própria finalidade dessas redes. Sistemas neurais artificiais têm como objetivo modelar algumas das características do sistema neural biológico, a fim de capturar e explorar aspectos do processamento da informação. Grande parte da ênfase nas pesquisas com redes neurais artificiais tem sido tentar uma simulação mais cuidadosa das atividades cerebrais, tanto em nível microscópico (neurônios), quanto em nível macroscópico (atividade cerebral global). Freeman (1991) chega a afirmar que “o caos deve ser a propriedade essencial que torna o cérebro diferente de uma máquina com inteligência artificial”.

O próprio Freeman, que descobriu o comportamento caótico em traços de eletroencefalogramas de bulbos olfatórios, tem conduzido uma grande pesquisa do papel do caos em sistemas neurais artificiais. Desde 1988 ele vem trabalhando em uma simulação computacional do córtex olfatório. O modelo de rede proposto por Freeman, o qual utiliza apenas oito neurônios artificiais caóticos simples, é capaz de replicar muitas das características encontradas pelo pesquisador na reprodução biológica.

Pistolesi (1996) argumenta que a presença de caos tem sido revelada em diversas atividades cerebrais do sistema nervoso central correlacionadas com atividades mentais envolvendo processos mnemônicos. Para Pistolesi, redes neurais artificiais apresentando comportamentos dinâmicos complexos seriam eficazes para simular artificialmente os processos cognitivos.

Muitas vantagens da inclusão de caos em sistemas neurais artificiais têm sido demonstradas (Raffone e Leeuwen, 2003, Kozma e Freeman, 2000, He *et al.*, 1999, Yao *et al.*, 1991, Sandler,

1990). No entanto, o uso do caos em sistemas artificiais é uma área ainda em expansão e previsões encontradas na literatura indicam que a teoria do caos fará parte do desenvolvimento futuro dos sistemas neurais artificiais.

2.4 Controle de Caos

Apesar do reconhecimento da existência de dinâmicas caóticas em sistemas neurais biológicos (Matsumoto *et al.*, 1987) e do desejo de se alcançar maior plausibilidade fisiológica para os neurônios artificiais através da inclusão de comportamentos caóticos nestes neurônios, na prática, segundo Ott *et al.* (1990a), também é freqüente a situação em que se deseja que o caos seja evitado em sistemas físicos e/ou que o desempenho do sistema seja melhorado ou alterado de alguma maneira. Ott *et al.* (1990a) mostram que existe uma grande flexibilidade inerente em situações nas quais o comportamento dinâmico encontra-se em um atrator caótico. Em particular, este autor prova que pequenas alterações no sistema, quando cuidadosamente escolhidas, tornam possível a criação de uma variedade de atratores periódicos e também possibilitam a escolha do atrator periódico desejado como o comportamento a ser alcançado pelo sistema.

Esta seção tem como objetivo a introdução do termo controle de caos e a apresentação dos principais tipos de controle de caos, a fim de que se possa justificar a seleção de um tipo de controle para as redes caóticas propostas nesta tese.

2.4.1 O Controle de Caos

De acordo com Andrievskii e Fradkov (2003), o termo *controle de caos* é usado para denotar a área de estudos que se estende entre a teoria de controle e a teoria de sistemas dinâmicos determinísticos com comportamento caótico.

Desde a década de 1990, problemas envolvendo controle de caos atraem a atenção de pesquisadores em diversas áreas, tais como tecnologias a lasers, engenharia mecânica e química, telecomunicações. Estatísticas de publicações no *Science Citation Index* constam um pequeno número de artigos publicados na área em 1990, enquanto que entre 1997-2001, apenas as revistas acadêmicas publicaram, anualmente, cerca de quatrocentos artigos (Andrievskii e Fradkov, 2003).

O desenvolvimento do controle de caos foi desencadeado essencialmente a partir do artigo de Ott *et al.* (1990a), realizado na Universidade de Maryland e publicado pela revista *Physical Reviews Letters*, em 1990. Foi neste artigo que o termo “controle de caos” foi introduzido (Fradkov e Evans, 2005). Um ponto-chave no artigo de Ott *et al.* (1990) foi a demonstração que uma mudança significativa no comportamento de um sistema caótico pode ser feita através de uma simples

correção de seus parâmetros. Tal observação abriu possibilidade para mudanças no comportamento natural dos sistemas sem a interferência em suas propriedades inerentes.

Ott (2002), também declara que o controle de caos é útil quando se deseja utilizar o conhecimento de sistemas caóticos para se alcançar alguma meta específica desejada. Segundo Ott, *controle* pode ser geralmente entendido como controle por realimentação, ou seja, o sistema deve possuir alguma variável de controle que pode ser variada como uma função do tempo, e, então, decide-se como fazer essa variação com base no conhecimento, talvez limitado, da história do sistema e/ou do estado atual deste sistema.

2.4.2 Como Controlar o Caos

O controle de um sistema é sempre realizado com algum objetivo programado. De acordo com Ott (2002), são três as possíveis metas que recebem atenção no contexto de sistemas caóticos:

Meta 1: dado um sistema caótico operando regularmente, o objetivo é melhorar seu desempenho.

Meta 2: dado um sistema caótico que está em algum estado inicial, o objetivo é trazê-lo rapidamente para algum outro estado específico desejado.

Meta 3: dado um sistema caótico operando regularmente, o objetivo é evitar que este sistema evolua para alguma região indesejável do espaço.

Para a utilização da Meta 1, a propriedade do caos da existência de uma órbita de estrutura complexa é o fator mais relevante a ser considerado. Para a utilização das Metas 2 e 3, a propriedade do caos da dependência exponencialmente sensível às condições iniciais é o fator mais relevante a ser considerado. No entanto, nestas três metas de maior repercussão no contexto de sistemas caóticos, talvez o fator de maior interesse no problema de controle de caos é que estas metas de controle podem, em princípio, ser alcançadas com estratégias simples de controle. A seguir serão introduzidas breves descrições de como se atingir cada uma das três metas citadas acima, objetivando a avaliação de qual meta de controle é a mais adequada para uma futura proposição de estratégia de controle para a família C-BAM.

2.4.2.1 Controlando um Processo Caótico para Melhorar seu Desempenho

(Meta 1)

O ponto chave para este problema de controle é a presença de uma órbita de estrutura complexa para o sistema caótico. Isto significa que existem muitas órbitas topologicamente distintas dentro do conjunto caótico invariante. Uma outra característica deste tipo de sistema é que existem muitas órbitas periódicas instáveis diferentes ‘embutidas’ em seu atrator caótico. Para que

sejam consideradas apenas pequenas perturbações no controle de um sistema deste tipo, não poderiam ser criadas novas órbitas com diferentes propriedades além daquelas já existentes. Sendo assim, a idéia é explorar as órbitas periódicas instáveis já existentes na ausência do controle. A abordagem para esta exploração, segundo Ott, Gregori e Yorke (1990a, 1990b) envolve os seguintes passos:

1- Examinar as dinâmicas não controláveis do sistema e selecionar algumas órbitas periódicas instáveis de baixo período embutidas no atrator caótico;

2- Examinar as órbitas selecionadas no passo (1), comparar o desempenho de cada uma delas e escolher aquela que apresentar o melhor desempenho;

3- Formular um algoritmo de controle que estabilize a órbita periódica instável selecionada em alguma vizinhança.

Utilizando-se esta abordagem, espera-se que, conforme as órbitas não controláveis transitam sobre o atrator, elas eventualmente se aproximarão da órbita periódica instável selecionada e, quando isto acontecer, ocorrerão pequenos deslocamentos para colocá-las dentro da órbita periódica instável desejada. Depois disto, se as órbitas são novamente deslocadas da órbita periódica instável desejada (por algum ruído, por exemplo), elas podem ainda ser re-inseridas na órbita desejada. Segundo Ott, Gregori e Yorke (1990a, 1990b), esta estratégia geral tem tido uma grande aplicabilidade em controle de sistemas caóticos. Algumas questões ligadas a esta estratégia geral são: (i) escolha do algoritmo de controle referido no passo (3) seleção da abordagem para exploração de órbitas periódicas instáveis dado acima, (ii) a implementação de coordenadas de atraso, (iii) efeito do ruído no sistema, (iv) possibilidade de uso de técnicas para redução do tempo de espera para as órbitas entrarem na vizinhança onde serão estabilizadas.

Aqui foi apenas apresentada a idéia geral de como funciona uma técnica de controle para a estabilização de órbitas periódicas instáveis embutidas em atratores caóticos. Métodos formais para esta estabilização podem ser encontrados na literatura, por exemplo, em Ott (2002) e em Ott, Grebogi e Yorke (1990a, 1990b).

2.4.2.2 Controlando um Processo Caótico para um Estado Desejado (Meta 2)

O segundo tipo de controle mencionado por Ott (2002) tem como principal característica direcionar rapidamente a órbita de um sistema caótico para algum local desejado no espaço de fase. Este método é chamado alvejar (*targeting*). O conceito básico deste método, encontrado em Kostelich *et al.* (1993) e Shinbrot *et al.* (1990, 1992), origina-se da idéia de que órbitas caóticas são exponencialmente sensíveis a pequenas perturbações. Assim, uma pequena perturbação nas órbitas tem um grande efeito em um período de tempo relativamente curto. Se as perturbações são ruídos desconhecidos, a órbita caótica torna-se imprevisível. No entanto, se o ruído é pequeno e possibilita

a aplicação de perturbações de controle, então é possível esperar que se consiga direcionar a órbita para um estado desejado com estratégias mais simples de controle.

Um exemplo das *diversas* estratégias de controle possíveis para o método alvejar encontra-se em Shinbrot *et al.* (1990) e é chamado avançar-retroagir (*forward-backward*). Neste método, o controle é ativado apenas inicialmente, e nenhum outro tipo de controle é subsequente aplicado (situação que precisaria ser modificada na presença de ruído, e de erros na observação do estado do sistema, por exemplo). No desenvolvimento do método, o sistema encontra-se em um ponto inicial, chamado ‘fonte’, sendo que o objetivo é atingir uma região chamada ‘região alvo’. Através de pequenas variações de um parâmetro de controle escalar p é possível perturbar o ponto fonte para um conjunto de localizações no espaço de estado do sistema. Tal conjunto de localizações pode ser imaginado como um pequeno segmento de curva iniciando-se no ponto fonte. Cada ponto neste segmento de curva corresponde a um valor particular de p . O suposto conhecimento do sistema é então utilizado para se evoluir adiante no tempo, dentro do segmento de curva, em uma quantidade Q . Simultaneamente, regride-se da região alvo de acordo com Q . Começa a ocorrer a seguinte situação: conforme Q aumenta, tanto a imagem para frente do segmento de controle inicial, quanto a imagem para trás da região alvo expandem-se exponencialmente em extensão. O segmento de controle expande-se em direção à variedade (*manifold*) instável do conjunto caótico, enquanto a pré-imagem da região alvo expande-se em direção à variedade estável. Eventualmente, conforme Q cresce, a imagem para frente do segmento de controle e a pré-imagem da região alvo irão se interceptar. Tal ponto de intersecção evolui para frente em direção à região alvo e para trás em direção ao segmento de controle inicial, especificando uma órbita conectando as duas regiões. Seguindo para trás do ponto de intersecção, é possível determinar um ponto no segmento de controle inicial. Este ponto determina um valor para o parâmetro de controle inicial $p = p_{ini}$ que produz uma órbita que caminha em direção à região alvo. Encontrado p_{ini} , esta seria a perturbação de controle ($p = p_{ini}$) aplicada ao sistema. Idealmente, a órbita evoluiria para alcançar a região alvo novamente no instante de tempo $2Q$. No entanto, pequenos ruídos, erros de modelagem, ou imprecisões ao determinar ou gerar p_{ini} , fariam que divergências exponenciais de trajetória ocorressem. Neste caso, o erro seria eliminado ao se aplicar repetidamente o método avançar-retroagir conforme as órbitas evoluíssem em direção à região alvo. Shinbrot *et al.* (1990) apresentam a ilustração de como o procedimento computacional aqui descrito se aplica em um exemplo específico.

Para sistemas de dimensões mais altas, o método avançar-retroagir torna-se de difícil aplicação, uma vez que é necessário trabalhar no tempo com grandes dimensões. Uma alternativa computacionalmente mais apropriada para situações de dimensões maiores é formulada por Kostelich *et al.* (1993). A idéia básica destes últimos autores é a de pré-computar uma árvore de

caminhos de órbitas controláveis que conduzem ao alvo. Seja o ponto alvo chamado x_T e suas P pré-imagens $x_{T-1}, x_{T-2}, \dots, x_0$. Estes pontos, os ‘caminhos raízes’, formam o tronco da árvore, a partir dos quais brotam ramos (‘caminhos secundários’), os quais, por sua vez, também possuem seus próprios ramos. Caminhando à frente no tempo, estes ramos divergem exponencialmente do caminho raiz. Este procedimento torna possível construir e armazenar uma rede de caminhos que permeiam o conjunto caótico e conduzem o sistema para o alvo. Neste momento é dito que o sistema inicia-se em um ponto específico. Este ponto é iterado até que se torne próximo a um dos pontos da árvore. Aí então se aplica o controle com o objetivo de colocar tal ponto em uma variedade estável do ramo vizinho mais próximo do caminho para o alvo.

2.4.2.3 Controlando um Processo Caótico para Prevenir uma Região Indesejável no Espaço de Fase (Meta 3)

Das três metas apresentadas por Ott (2002) como sendo as de maior atenção no contexto de sistemas caóticos, a terceira meta é a menos estudada, e a respeito dela será introduzida apenas uma breve discussão.

O interesse na meta 3, que é a de prevenir uma região indesejável no espaço de fase, deve-se principalmente ao propósito de se evitar a ocorrência de algum evento catastrófico quando se conhece previamente que isto acontecerá, quando órbitas caóticas do sistema atingirem uma região particular do espaço de fase. Um exemplo de aplicação desta meta é mencionado no trabalho de In *et al.* (1997), no qual ele considera um combustor térmico, dispositivo cujo propósito é queimar uma mistura de combustível e ar para produzir na saída de sua câmara de combustão um fluxo de gasolina quente. Foi descoberto que, conforme a proporção entre combustível e ar, o combustor pulsa de modo caótico. Com um decréscimo desta proporção entre combustível e ar, o atrator caótico passa pelo fenômeno caótico chamado crise, e como consequência, apaga-se a chama na câmara de combustão e o dispositivo pára de operar. Em seu trabalho, In *et al.* (1997) discutem como a órbita no estado de caos transitório pode ser controlada por pequenas perturbações para evitar que ela chegue na região do espaço de estado em que ocorre a crise. Se isto é feito, o movimento caótico é sustentado e o combustor térmico mantém-se operando.

Outros exemplos de controle envolvendo a Meta 3 incluem a prevenção de um navio virar em um mar agitado (Ding *et al.*, 1994) e uma possível intervenção estratégica para prevenir ataques epiléticos (Yang *et al.*, 1995).

2.5 Resumo

Para que a teoria do caos e os comportamentos dinâmicos nela presentes sejam entendidos, é necessário familiaridade com o uso de conceitos e de técnicas vindos, em sua maioria, da matemática, física e estatística. Este capítulo procurou fornecer condições para o entendimento e identificação do fenômeno caótico. Também foram analisadas as três metas de controle de caos que recebem atenção no contexto de sistemas caóticos, visando posterior proposição de métodos de controle de caos para a família de memórias associativas bidirecionais caóticas.

Foi tratado como o caos determinístico se comporta em sua evolução a longo prazo, ou seja, como o comportamento do sistema caótico se modifica ao longo do tempo. Há duas importantes aplicações práticas no estudo de caos: (i) as previsões do sistema a longo prazo não devem ser consideradas, pois são muito imprecisas e (ii) o comportamento complexo pode ter causas simples. Foi também verificado que, apesar de à primeira vista o comportamento caótico parecer irregular, é possível reconhecer ordem dentro do caos.

Dentre as importantes características que podem identificar o caos, encontra-se a sensibilidade às condições iniciais, a qual pode ser medida através do cálculo dos expoentes de Lyapunov.

A teoria do caos, incluindo sensibilidade às condições iniciais, imprevisibilidade a longo prazo, expoentes de Lyapunov, entre outros aspectos, é um tópico interdisciplinar, estudado por especialistas de várias áreas. Foi mostrado que o caos ocorre tanto em sistemas naturais quanto em sistemas artificiais e que há evidências de que o caos pode contribuir positivamente no desempenho das redes neurais artificiais.

No presente trabalho, o fenômeno caótico é analisado no domínio das redes neurais artificiais, particularmente em redes de memória associativa. Os principais pontos da teoria do caos a serem diretamente empregados na tese envolvem a bifurcação de duplicação de período, crise, caos determinístico e o cálculo dos expoentes de Lyapunov, uma vez que tais tópicos farão parte dos comportamentos emergentes das redes de memória associativas caóticas. O capítulo a seguir descreve os primeiros modelos de neurônios artificiais capazes de reproduzir fenômenos caóticos, os quais serão utilizados como base para a construção da família de modelos de memória associativa bidirecional caótica (Capítulo 4).

Capítulo 3

Caos, Modelos Auto-associativos Caóticos e Fenômeno Saltar Passos: Fundamentos

O Capítulo 2 apresentou importantes fenômenos característicos do caos presentes em uma ampla variedade de áreas da ciência e em diversas aplicações. Adachi e Aihara (1997) relatam que as dinâmicas caóticas existem tanto em neurônios reais quanto em neurônios de redes neurais artificiais. Conseqüentemente, a ocorrência de caos em redes neurais artificiais pode ser utilizada para modelar uma série de efeitos e fenômenos.

Os objetivos deste capítulo englobam: (a) descrever um modelo de neurônio artificial capaz de reproduzir o fenômeno caótico encontrado em experimentos com animais (modelo neurônio caótico) (Aihara, Takabe e Toyoda, 1990); (b) apresentar um modelo de memória auto-associativa construído com esse neurônio caótico e mostrar alguns comportamentos resultantes desse modelo; (c) conceituar controle de caos e descrever o método de controle por pinagem (*pinning control*) para modelos de memória auto-associativas caóticas; (d) apresentar os conceitos do fenômeno psicológico chamado saltar passos (*step skipping*) (Blessing e Anderson, 1996), o qual será modelado no Capítulo 5 com base nas características comportamentais de uma família de memórias hetero-associativas bidirecionais com neurônios caóticos que será introduzida no próximo capítulo deste documento.

A Seção 3.1 descreve a criação do modelo neurônio caótico, capaz de gerar comportamentos não periódicos em redes neurais artificiais, e também a criação da rede neural caótica, gerada com os neurônios do modelo neurônio caótico. Inclui, ainda, a aplicação do modelo neurônio caótico em um modelo de memória auto-associativa e os resultados de tal aplicação. A apresentação do método de controle por pinagem proposto para controlar a dinâmica do modelo de memória auto-associativa caótica é feita na Seção 3.2. A Seção 3.3 descreve o fenômeno saltar passos, que será

modelado pelas redes criadas neste trabalho no Capítulo 4. Um resumo final destaca os conceitos deste capítulo que são mais relevantes para o presente trabalho.

3.1 Caos em Modelos Auto-associativos

Esta seção comenta a motivação encontrada para que neurônios artificiais capazes de produzir caos, e não apenas respostas periódicas, fossem construídos; descreve um breve histórico das tentativas de se modelar redes neurais artificiais caóticas até a construção do modelo neurônio caótico; apresenta a criação da rede neural caótica, construída com o modelo neurônio caótico; mostra a construção de uma rede neural de memória auto-associativa caótica e o comportamento diferente resultante da inclusão do modelo neurônio caótico neste tipo de rede, e termina apresentando um método de controle de caos capaz de estabilizar a rede de memória auto-associativa caótica em um padrão desejado.

3.1.1 Modelo Neurônio Caótico

Uma rede neural artificial é composta por elementos chamados neurônios artificiais, os quais têm como objetivo modelar, em diferentes níveis de abstração, os neurônios biológicos reais, ou seja, modelar a maneira pela qual o cérebro desempenha uma tarefa particular ou uma função de interesse. Portanto, um neurônio é a unidade de processamento de informação fundamental para a operação de uma rede neural artificial (Haykin, 1999). Um modelo comum de neurônio artificial é aquele em que um simples elemento limiar transforma um somatório ponderado de entradas em uma saída, através de uma função de saída não linear com um limiar. Entretanto, do ponto de vista da neurofisiologia, há críticas envolvendo a função das redes neurais artificiais de modelar neurônios reais, uma vez que estes últimos são considerados muito mais complexos do que os simples elementos limiares (Aihara, Takabe e Toyoda, 1990).

Uma das razões para estas críticas deve-se à constatação da falta de uma característica típica de neurônios biológicos em neurônios artificiais comuns, que é a presença de comportamentos caóticos. Por exemplo, foi possível mostrar tanto através de experimentos realizados com axônios de lulas gigantes, quanto numericamente, através das equações Hodgkin-Huxley (Cole, Guttman e Bezanilla, 1970), que as respostas dos neurônios a estímulos periódicos nem sempre são periódicas e que as respostas aparentemente não periódicas que surgiam poderiam ser entendidas como sendo casos de caos determinístico.

Aihara, Takabe e Toyoda (1990) apresentam um resumo histórico das tentativas de uma modelagem mais adequada das dinâmicas de neurônios biológicos. Esta história remonta a antigos modelos importantes, tais como o neurônio de McCulloch-Pitts (1943) e o de Caianiello (1961). A

equação do modelo de neurônio de Caianiello, a qual inclui o modelo de McCulloch-Pitts como um caso especial, é descrita por:

$$x_i(t+1) = u \left(\sum_{j=1}^m \sum_{\sigma=0}^t w_{ij}^{(\sigma)} x_j(t-\sigma) - \theta_i \right) \quad (3.1)$$

em que $x_i(t+1)$ representa a saída do i -ésimo neurônio no tempo discreto $t+1$; $u(\cdot)$ é a função degrau, de modo que $u(y) = 1$ para $y \geq 0$ ou $u(y) = 0$ para $y < 0$; m é o número de neurônios da rede neural artificial; $w_{ij}^{(\sigma)}$ (para $i \neq j$) é a conexão ponderada com a qual o j -ésimo neurônio afeta o i -ésimo neurônio após $\sigma+1$ unidades de tempo; $w_{ii}^{(\sigma)}$ é o coeficiente de memória de refração relativa com a qual o i -ésimo neurônio retém influências nele próprio após $\sigma+1$ unidades de tempo; e θ_i é o limiar para a ativação tudo-ou-nada do i -ésimo neurônio.

Nagumo e Sato (1972) analisaram respostas de um neurônio com uma única entrada baseando-se numa modificação do modelo de Caianiello. Estes autores afirmaram que a influência da refração decresce exponencialmente com o tempo, ou seja, $w_{ii}^{(\sigma)} = -\rho k^\sigma$, em que k assume valores entre 0 e 1 e ρ é um parâmetro positivo. A Equação 3.2 descreve o modelo de Nagumo e Sato:

$$x_i(t+1) = u \left(A_i(t) - \rho \sum_{\sigma=0}^t k^\sigma x_i(t-\sigma) - \theta_i \right) \quad (3.2)$$

em que $A_i(t)$ é a intensidade da entrada do i -ésimo neurônio no tempo discreto t e k é o fator de decaimento da refração.

A Equação 3.2 pode ser simplificada ao se definir uma nova variável $y_i(t+1)$, descrita na Equação 3.3, representando o estado interno do neurônio i :

$$y_i(t+1) = A_i(t) - \rho \sum_{\sigma=0}^t k^\sigma x_i(t-\sigma) - \theta_i \quad (3.3)$$

Deste modo, a Equação 3.2 pode ser descrita como:

$$y_i(t+1) = ky_i(t) - \rho u\{y_i(t)\} + \Theta_i(t) \quad (3.4)$$

$$x_i(t+1) = u\{y_i(t+1)\} \quad (3.5)$$

em que:

$$\Theta_i(t) = A_i(t) - kA_i(t-1) - \theta_i[1-k] \quad (3.6)$$

No caso acima, enquanto o comportamento dinâmico do neurônio é calculado com a Equação 3.4, através do estado interno y , o valor da saída é obtido ao se transformar este estado interno y na saída x , de acordo com a Equação 3.5.

Respostas características das Equações 3.4 e 3.5 foram analisadas no trabalho de Aihara, Takabe e Toyoda (1990) através de diagramas de bifurcação, cálculo do expoente de Lyapunov e do cálculo da taxa média de ativação dos neurônios. Desta análise conclui-se que quase todas as soluções das Equações 3.4 e 3.5 são periódicas. Este resultado manteve a diferença entre o modelo e os experimentos executados com axônios de lulas gigantes, nos quais as respostas caóticas puderam ser facilmente observadas.

Tal diferença entre o modelo e o neurônio biológico indicava que novas modificações deveriam ser feitas nas Equações 3.1 e 3.2. Pôde ainda ser mostrado tanto experimentalmente, através de estudos com axônios de lulas gigantes, quanto numericamente, através das equações de Hodgkin-Huxley, que a curva estímulo-resposta de um neurônio não é algo descontínuo do tipo tudo-ou-nada como a função degrau u utilizada nas Equações 3.1 e 3.2, mas que a propriedade de estímulo-resposta deveria ser representada por uma função contínua.

Diante desta constatação, Aihara, Takabe e Toyoda (1990) fizeram a substituição da função degrau unitária u usada na Equação 3.2 por uma função de saída contínua:

$$x_i(t+1) = f_i \left(A_i(t) - \rho \sum_{\sigma=0}^t k^\sigma g_i \{x_i(t-\sigma)\} - \theta_i \right) \quad (3.7)$$

em que $x_i(t+1)$ é a saída do i -ésimo neurônio no tempo discreto $t+1$, a qual assume um valor analógico entre 0 e 1; f_i é uma função de saída contínua que pode ser, por exemplo, a função logística $f(y) = 1/(1+e^{-y/\varepsilon})$ com parâmetro de escalada (*steepness*) ε ; $A_i(t)$ é o estímulo externo no tempo t ; g_i é uma função que descreve a relação entre a saída analógica e a magnitude da refração; k é o parâmetro de decaimento da refração e θ_i é o limiar.

Como no caso do modelo de Nagumo e Sato, foi definida uma nova variável ($y_i(t+1)$) para o estado interno do neurônio:

$$y_i(t+1) = A_i(t) - \rho \sum_{\sigma=0}^t k^\sigma g_i \{x_i(t-\sigma)\} - \theta_i \quad (3.8)$$

A Equação 3.7 passa a ser equivalentemente descrita pelas Equações 3.9 e 3.10:

$$y_i(t+1) = k y_i(t) - \rho g_i \{f_i \{y_i(t)\}\} + \Theta_i(t) \quad (3.9)$$

$$x_i(t+1) = f_i \{y_i(t+1)\} \quad (3.10)$$

Na Equação 3.9, $\Theta_i(t) \equiv A_i(t) - kA_i(t-1) - \theta_i [1-k]$. Para fins de simplificação e para que se possa focalizar nos efeitos da função de saída contínua f_i , a função g_i escolhida por Aihara, Takabe e Toyoda (1990) foi a função $g_i(v) = v$. Também para simplificar foi selecionado $\Theta_i(t) = \Theta_i$, que é temporalmente constante.

Tal como no modelo anteriormente apresentado (modelo de Nagumo e Sato), as respostas características das Equações 3.9 e 3.10, sendo f_i a função logística, foram analisadas através de diagramas de bifurcação, do cálculo do expoente de Lyapunov e do cálculo da taxa média de ativação dos neurônios. Foi verificado (Aihara, Takabe e Toyoda, 1990) que as respostas destas novas equações reproduziam, de modo qualitativo, as seqüências alternadas de respostas periódicas-caóticas experimentalmente observadas com os axônios de lula gigante. Tais análises demonstraram ainda que o modelo de neurônio das Equações 3.9 e 3.10 possui respostas caóticas em amplas regiões de seu espaço paramétrico.

O novo modelo de neurônio com dinâmicas caóticas, o *modelo neurônio caótico*, foi então generalizado como o elemento constituinte de uma rede neural artificial chamada *rede neural caótica* (Aihara, Takabe e Toyoda, 1990). Esta rede neural caótica capaz de produzir caos determinístico em redes neurais artificiais é portanto composta por neurônios caóticos excitados por: (a) entradas de realimentação vindas de cada um dos neurônios caóticos, (b) entradas de auto realimentação e (c) entradas externas. A ilustração do neurônio caótico é mostrada na Figura 3.1.

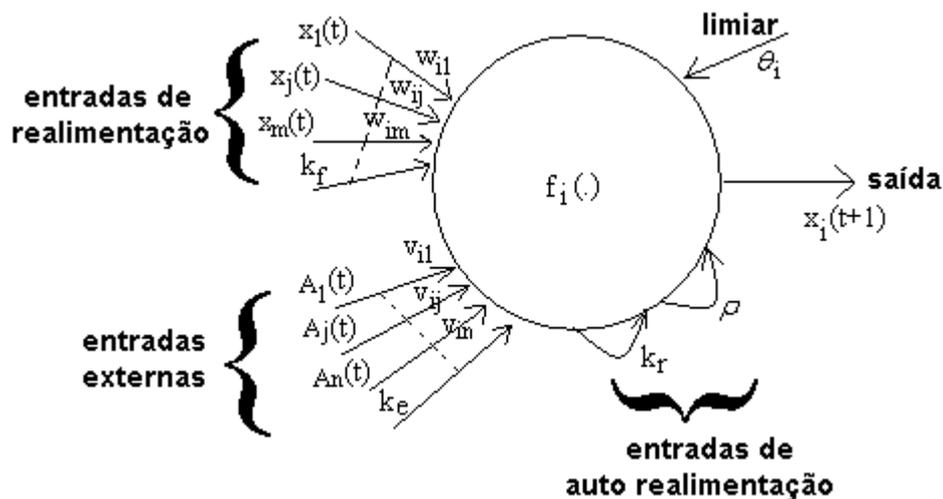


Figura 3.1. Modelo do neurônio caótico como um elemento da rede neural caótica.

A dinâmica do i -ésimo neurônio caótico em uma rede neural caótica composta de m neurônios caóticos pode ser modelada de acordo com a Equação 3.11.

$$x_i(t+1) = f_i \left[\sum_{j=1}^m w_{ij} \sum_{\sigma=0}^t k_f^\sigma x_j(t-\sigma) + \sum_{j=1}^n v_{ij} \sum_{\sigma=0}^t k_e^\sigma A_j(t-\sigma) - \rho \sum_{\sigma=0}^t k_r^\sigma g_i \{x_i(t-\sigma)\} - \theta_i \right] \quad (3.11)$$

Na Equação 3.11, $x_i(t+1)$ é a saída do i -ésimo neurônio caótico no tempo discreto $t+1$; f_i é a função de saída contínua do neurônio caótico; m é o número de neurônios caóticos na rede neural;

w_{ij} é o peso da conexão entre o j -ésimo neurônio caótico e o i -ésimo neurônio caótico; k_f é o parâmetro de decaimento para entradas realimentadas; n é o número de entradas externas aplicadas na rede, v_{ij} é o peso da conexão entre a j -ésima entrada externa e o i -ésimo neurônio caótico, $A_j(t-\sigma)$ é a intensidade da j -ésima entrada externa aplicada no tempo $t-\sigma$, k_e é o parâmetro de decaimento para entradas externas; ρ é um parâmetro positivo que pondera a magnitude da refração; k_r é o parâmetro de decaimento da refração; g_i é a função de refração do neurônio caótico e θ_i é o limiar do i -ésimo neurônio caótico. Para serem fisiologicamente plausíveis, os valores de k_e , k_f e k_r devem estar dentro do intervalo $[0,1]$.

Segundo Aihara, Takabe e Toyoda (1990), a Equação 3.11 pode ser re-escrita pelas Equações 3.12 a 3.15:

$$x_i(t+1) = f_i \{ \xi_i(t+1) + \eta_i(t+1) + \varsigma_i(t+1) \} \quad (3.12)$$

em que :

$$\xi_i(t+1) = k_e \xi_i(t) + \sum_{j=1}^n v_{ij} A_j(t) \quad (3.13)$$

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^m w_{ij} x_j(t) \quad (3.14)$$

$$\varsigma_i(t+1) = k_r \varsigma_i(t) - \rho g_i \{ x_i(t) \} - \Theta_i; \quad \Theta_i \equiv \theta_i [1 - k_r] \quad (3.15)$$

As Equações 3.13, 3.14 e 3.15 são os termos dos estados internos da rede para entradas externas (ξ), para entradas realimentadas vindas de outros neurônios caóticos da rede (η), e para refração (ς), respectivamente. Enquanto a dinâmica da rede é calculada através das Equações 3.13 a 3.15, o padrão de saída da rede é obtido pela Equação 3.12.

Foi possível observar que a modelagem dos neurônios caóticos feita por Aihara, Takabe e Toyoda (1990) é uma extensão da rede neural discreta no tempo convencional, a fim de que as dinâmicas caóticas sejam introduzidas nos modelos de redes neurais artificiais. Isto porque o modelo neurônio caótico das Equações 3.12-3.15 generaliza alguns dos modelos convencionais de neurônios discretos no tempo, tais como, por exemplo, os de McCulloch e Pitts (1943) (para $k_e = k_f = k_r = \rho = 0$ e $f_i =$ função discreta) e o de Rumelhart, Hinton e Williams (1986) (para $k_e = k_f = k_r = \rho = 0$ e $f_i =$ função contínua), por exemplo. Os neurônios caóticos se enquadram na situação em que $k_e \neq 0$, $k_f \neq 0$, $k_r \neq 0$, $\rho \neq 0$ e f_i é uma função contínua, tal como a função logística.

Outros pesquisadores tentaram modelar redes neurais artificiais com dinâmicas caóticas. Yao e Freeman (1990), Tani e Fujita (1992), por exemplo, tentaram tais modelos utilizando equações diferenciais determinísticas. Já Shiino e Fukai (1990), e Tsuda (1992), tentaram modelar redes neurais caóticas utilizando modelos estocásticos. No entanto, o modelo de rede neural caótica de Aihara, Takabe e Toyoda (1990) apresenta vantagens de tempo computacional e de memória para

análises numéricas, uma vez que a dinâmica complexa de seus neurônios é descrita por equações de diferenças determinísticas simples e que, dada a sua função de saída f_i ser contínua, possibilitam a análise da estabilidade orbital da rede através do cálculo dos expoentes de Lyapunov, os quais são comumente usados para analisar sistemas caóticos.

3.1.2 Aplicação do Modelo Neurônio Caótico em Memórias Auto-associativas

Em seu trabalho de 1997, Adachi e Aihara apresentaram a construção de uma rede neural auto-associativa simples, composta por cem neurônios caóticos interconectados através de uma matriz de pesos sinápticos auto-associativa convencional, com a finalidade de analisar as propriedades fundamentais desta rede. Para a aplicação do modelo de rede neural caótica à memória associativa, foi utilizada uma versão simplificada das Equações 3.12 a 3.15. A rede neural caótica para modelos auto-associativos é então modelada com apenas dois estados internos η_i e ς_i :

$$x_i(t+1) = f_i(\eta_i(t+1) + \varsigma_i(t+1)) \quad (3.16)$$

em que :

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^m w_{ij} x_j(t) \quad (3.17)$$

$$\varsigma_i(t+1) = k_r \varsigma_i(t) - \rho x_i(t) + a_i \quad (3.18)$$

No conjunto de equações acima, enquanto a dinâmica da rede é descrita pelas Equações 3.17 e 3.18, as saídas dos neurônios são obtidas pela Equação 3.16. Na Equação 3.18, a_i denota a soma do limiar com as entradas externas temporalmente constantes para o i -ésimo neurônio. A simplificação do modelo de Equações 3.16 a 3.18 ocorre ao se supor que os estímulos externos para a rede são temporalmente constantes, ou seja, o estado interno $\xi_i(t)$ da Equação 3.13 pode ser incluído na variável temporalmente constante a_i . Esta simplificação foi feita para facilitar análise das características do modelo em questão, uma vez que é mais fácil fazer isto trabalhando somente com as duas variáveis η_i e ς_i do que com os três estados internos variáveis η_i , ς_i e ξ_i . Ainda para simplificação, da Equação 3.15 para a Equação 3.18, foi escolhida como g_i a função $g_i(v) = v$.

As interconexões de realimentação são determinadas de acordo com a matriz auto-associativa simétrica de p padrões armazenados descrita na Equação 3.19:

$$w_{ij} = \frac{1}{p} \sum_{k=1}^p (2 X_i^k - 1)(2 X_j^k - 1) \quad (3.19)$$

em que X_i^k denota o i -ésimo componente do k -ésimo padrão armazenado, para $k=1, \dots, p$.

Para os testes da rede neural auto-associativa com neurônios caóticos, Adachi e Aihara (1997) selecionaram quatro padrões a serem armazenados. A Figura 3.2 ilustra \mathbf{X}^1 , \mathbf{X}^2 , \mathbf{X}^3 e \mathbf{X}^4 , os

\mathbf{X}^k ($k=1, \dots, 4$) padrões armazenados na rede de acordo com a Equação 3.19.

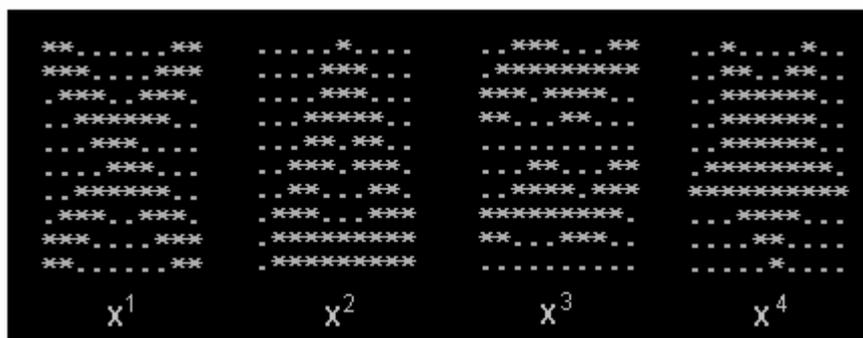


Figura 3.2. Padrões armazenados utilizados nos testes com a rede neural auto-associativa contendo neurônios caóticos.

Na Seção 3.1.1 foi visto que a rede neural caótica apresenta dinâmicas aparentemente não periódicas. O objetivo da inclusão do modelo desta seção foi o de analisar como estas dinâmicas não periódicas dos neurônios caóticos influenciam a rede de memória auto-associativa e esclarecer características do processo de recuperação da rede de memória auto-associativa caótica. Para que as diferenças entre a rede de memória auto-associativa convencional e a rede de memória auto-associativa caótica possam ser aqui destacadas, deve ser lembrado que a primeira se caracteriza por convergir sempre para um ponto fixo, o qual pode ser ou não ser um padrão armazenado. Com relação à rede de memória auto-associativa caótica, serão aqui tratadas suas características de recuperação utilizando-se de (i) frequência de recuperação e (ii) transição entre padrões armazenados, a fim de esclarecer a habilidade do processo de busca da memória.

Quando foi atribuído valor zero aos parâmetros da rede de memória auto-associativa caótica ($k_f = k_r = \rho = a = 0$), a rede apresentou uma dinâmica idêntica à rede auto-associativa convencional, estabilizando-se em um estado de equilíbrio. Conforme foram sendo introduzidos os componentes dos neurônios caóticos, ou seja, conforme os valores dos parâmetros dos neurônios caóticos tinham seus valores aumentados, o comportamento dinâmico da rede de memória auto-associativa caótica tornava-se mais complexo.

Alguns experimentos de Adachi e Aihara (1997) relacionados com a rede de memória auto-associativa caótica foram reproduzidos a fim de um melhor entendimento do processo dinâmico da nova rede criada. Para obtenção destes resultados, foi implementado em linguagem C, em um computador Pentium IV, o modelo de Adachi e Aihara (1997). Resultados equivalentes aos encontrados no artigo de Adachi e Aihara (1997) foram obtidos e serão aqui ilustrados para facilitar a descrição das características da dinâmica e do processo de recuperação da rede de memória auto-associativa caótica.

Nos experimentos reproduzidos, valores relativamente baixos para a realimentação inibitória

auto-recorrente, peculiar aos neurônios caóticos (k_i), provocaram um tipo de comportamento no qual, dado um padrão inicial, a rede mantinha-se ao redor de um estado de equilíbrio, ou recuperando exatamente o padrão mais próximo deste padrão inicial ou recuperando variações aproximadas dele. A Figura 3.3 demonstra a ocorrência deste comportamento dinâmico. O padrão inicial ($t=0$) é uma versão com ruído do padrão armazenado \mathbf{X}^3 da Figura 3.2. Nos instantes de tempo $t=[1,4]$ a rede recupera exatamente o padrão armazenado \mathbf{X}^3 . Em seguida, nos instantes $t=[5,15]$ a rede passa a apresentar como saída versões aproximadas de \mathbf{X}^3 . A versão exata do padrão armazenado é novamente recuperada entre os instantes de tempo $t=[16,19]$. Entre os instantes $t=[20,26]$ versões aproximadas de \mathbf{X}^3 são recuperadas e entre $t=[27,29]$ verifica-se a recuperação de \mathbf{X}^3 .

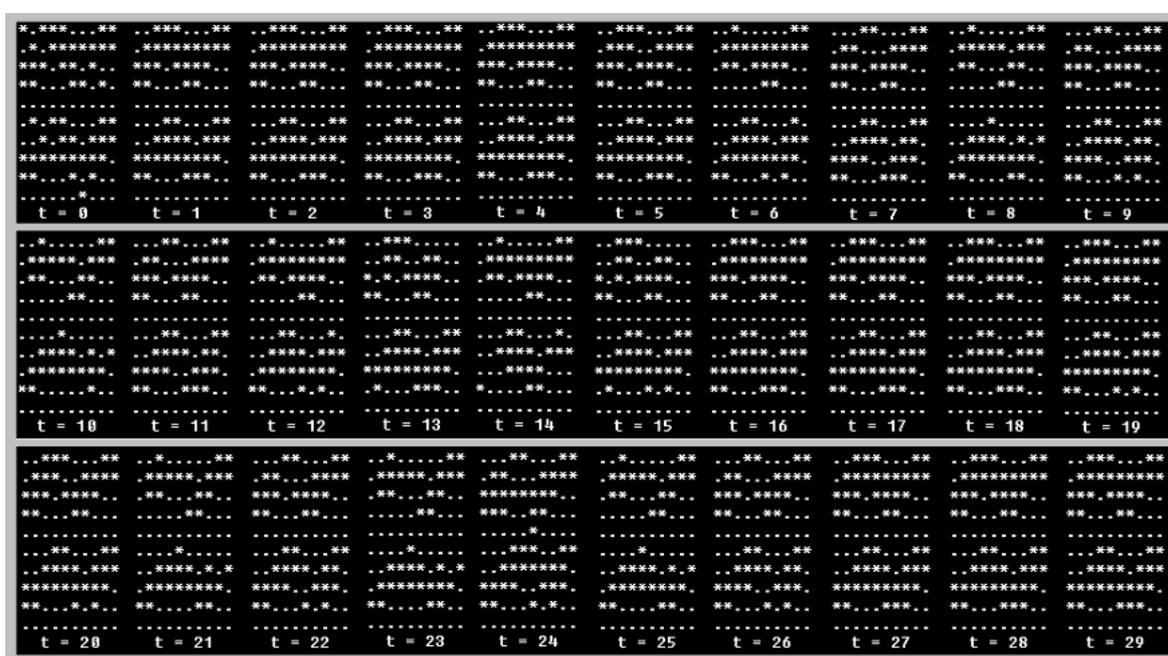


Figura 3.3. Exemplo de seqüência de padrões de saída da rede de memória auto-associativa caótica, com capacidade de refração relativamente fraca, em que $k_i=0.5$, $k_j=0.15$, $\rho=8.0$ e $a_i=2.0$ ($i=1,\dots,100$).

Quando os valores dos parâmetros da rede de memória auto-associativa caótica foram ajustados para valores mais altos, em especial para a capacidade de refração, os resultados mostraram que a rede não mais se mantém ao redor de um estado de equilíbrio. A frequência de recuperação do padrão de saída da rede modifica-se. A rede passa a mostrar uma dinâmica associativa interessante e complexa, na qual os padrões armazenados são recuperados seqüencialmente, não necessariamente na ordem original de armazenamento e freqüentemente intercalados com a recuperação de padrões espúrios. Além disto, tal dinâmica parece ser não periódica. A Figura 3.4 ilustra um exemplo de seqüência de padrões de saída com este tipo de

dinâmica. A Tabela 3.1 aponta, entre os instantes de tempo $t=[0,170]$, a recuperação de padrões armazenados e a recuperação de padrões inversos dos padrões armazenados.

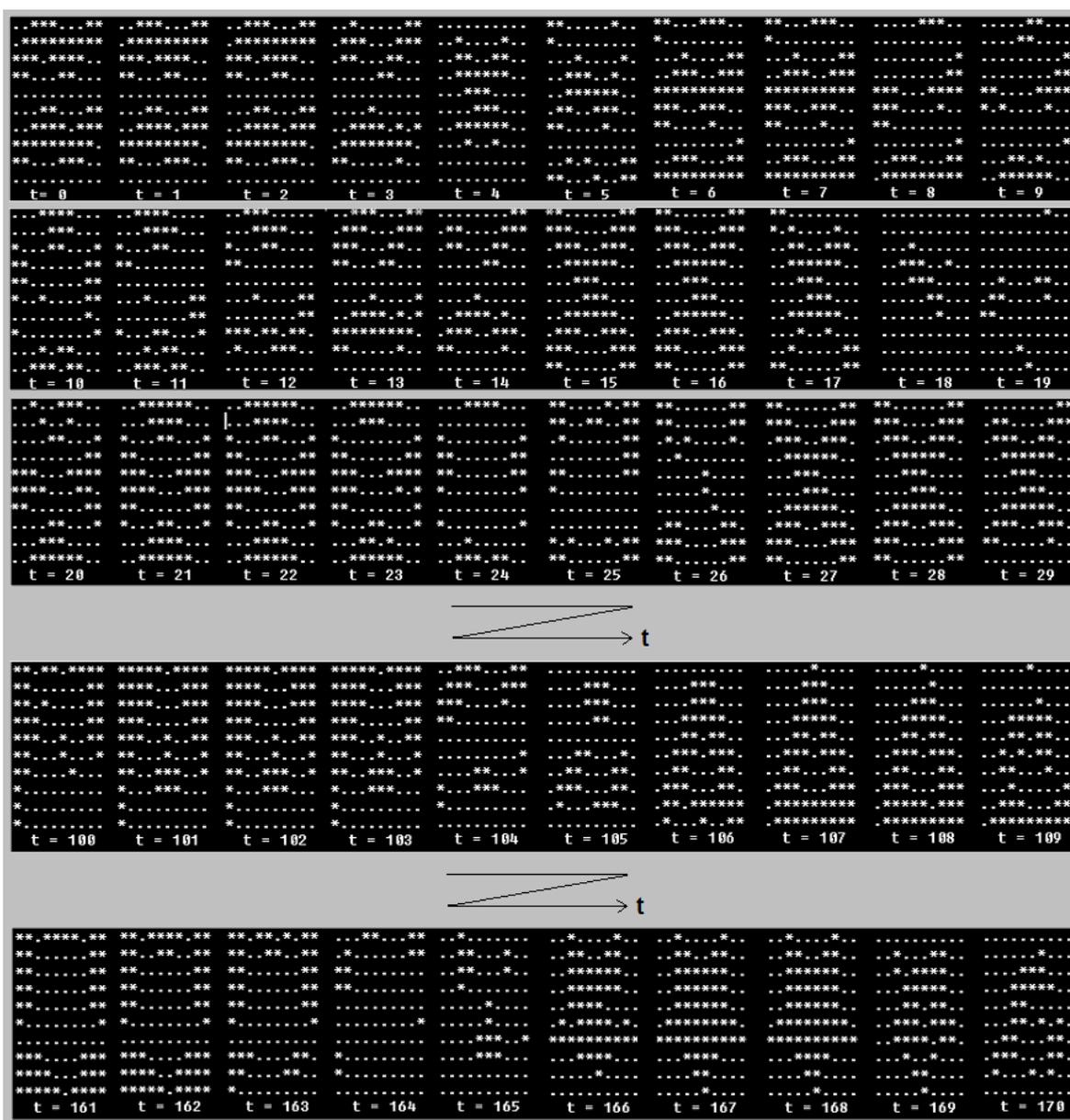


Figura 3.4. Exemplo de seqüência de padrões de saída da rede de memória auto-associativa caótica, com capacidade de refração alta, em que $k_i=0.8$, $k_f=0.2$, $\rho=8.0$ e $a_i=2.0$ ($i=1,\dots,100$). O padrão inicial selecionado é o padrão armazenado \mathbf{X}^3 , da Figura 3.2.

Tabela 3.1. Instantes de tempo de recuperação de padrões armazenados e seus inversos para apresentação de X^3 .

Padrões recuperados	Instantes de tempo de recuperação	Padrões recuperados	Instantes de tempo De recuperação
X^1	15,16,27,28	Inverso de X^1	21,22,34
X^2	107	Inverso de X^2	101,102,103
X^3	1,2,54,55	Inverso de X^3	6,7,60,61,122,134
X^4	167,168	Inverso de X^4	162

3.2 Controle de Caos para Modelos Auto-associativos

Diante do fato exposto por Ott (2002) de que o controle de um sistema é sempre realizado com algum objetivo programado e que as três metas que mais recebem atenção no contexto de sistemas caóticos são as metas apresentadas na Seção 2.4 do Capítulo 2, o primeiro passo para a escolha de uma estratégia de controle para a dinâmica caótica das redes de memória associativas seria selecionar qual das metas apresentadas é a mais adequada para a situação em que se deseja que estas redes alcancem um comportamento específico.

Foi visto na Seção 3.1.2 que o processo de recuperação de padrões armazenados pela rede de memória auto-associativa caótica não pode ser estabilizado em um padrão armazenado específico por causa da dinâmica caótica desta rede. Foi verificado por experimentação que a rede de memória auto-associativa caótica, mediante a escolha apropriada de seus valores de parâmetros, é capaz de transitar por todos os padrões armazenados, sem, contudo, estabilizar-se em nenhum destes padrões. Nesta circunstância seria interessante a busca por uma estratégia de controle que permitisse a estabilização da rede de memória auto-associativa caótica em um padrão armazenado desejado.

A idéia de estabilizar a rede de memória auto-associativa caótica em um padrão armazenado poderia enquadrar-se nas Metas 1 e 2 mencionadas nas Seções 2.4.2.1 e 2.4.2.2, respectivamente. Por exemplo, no caso de um controle objetivando a Meta 1, poderia ser proposta a seleção de uma órbita periódica instável na vizinhança do padrão armazenado desejado, considerando-se algum critério de otimalidade. No caso de um controle objetivando a Meta 2, poderia ser proposta uma perturbação no controle que levaria a rede ao padrão armazenado desejado.

Através da análise de alguns métodos de controle de caos encontrados na literatura que se enquadram nas Metas 1 (Ott *et al.*, 1990a, So *et al.*, 1997) e 2 (Shinbrot *et al.*, 1992, Kostelich *et al.*, 1993) descritas na seção anterior, foi possível notar que, em termos práticos, seria muito mais simples encontrar uma maneira de direcionar a órbita de um sistema caótico para algum local

desejado no espaço de fase (principal característica da Meta 2) do que explorar as órbitas periódicas instáveis presentes na estrutura complexa de um sistema caótico (um ponto chave para a utilização da Meta 1). Portanto, o presente trabalho adota como objetivo de controle a Meta 2, que é a de controlar um processo caótico trazendo-o rapidamente para um estado específico desejado.

Existem várias estratégias de controle para rapidamente direcionar a órbita de um sistema caótico em algum local desejado no espaço de fase do sistema. Além das estratégias já mencionadas, o método avançar-retroagir e o método de Kostelich, o controle por pinagem (Hu e Qu, 1994) mostrou-se apropriado ao controle de caos desejado para a rede de memória auto-associativa caótica pelo fato de ser um método simples, robusto e eficaz.

A idéia por trás do método de controle por pinagem é, de fato, muito simples: colocar alguns controles locais (*pinagem*) no espaço de estado da rede. Os “pinos” de realimentação são utilizados para controlar o caos do sistema ao estabilizar um certo estado instável. Quando a distribuição dos pinos torna-se suficientemente densa, o estado instável pode ser estabilizado. O controle por pinagem foi proposto para controlar o caos de um sistema treliça de mapa acoplado (*coupled map lattice*) em uma dimensão, cuja equação do modelo é descrita da seguinte maneira:

$$x_{i+1}(l) = (1 - \varepsilon)f[x_i(l)] + \frac{1}{2}\varepsilon\{f[x_i(l-1)] + f[x_i(l+1)]\} \quad (3.20)$$

em que $l=1,2,\dots,L$ são os locais da rede (*lattice sites*), L é o tamanho do sistema, ε é a intensidade de acoplamento (*coupling strength*) e $f(x) = ax(l-x)$. A dinâmica deste modelo permite a ocorrência de caos (Hu e Qu, 1994). Para controlar este sistema, foram utilizados pinos definidos da seguinte maneira:

$$x_{i+1}(l) = (1 - \varepsilon)f[x_i(l)] + \frac{1}{2}\varepsilon\{f[x_i(l-1)] + f[x_i(l+1)]\} + \sum_{k=0}^{L/I} \delta(l - Ik - 1)g_i \quad (3.21)$$

$$g_i = (1 - \varepsilon)p_i(l)x_i(l)[x_i(l) - \tilde{x}_i(l)] + \frac{1}{2}\varepsilon\{p_i(l-1)x_i(l-1)[x_i(l-1) - \tilde{x}_i(l-1)] + p_i(l+1)x_i(l+1)[x_i(l+1) - \tilde{x}_i(l+1)]\} \quad (3.22)$$

em que I é a distância entre dois pinos vizinhos, $\tilde{x}_i(l)$ é o estado em que se quer estabilizar o modelo, $p_i(l)$ é a intensidade da realimentação adicionada ao l -ésimo lado do mapa e $\delta(j)=1$ para $j=0$ ou $\delta(j)=0$, caso contrário. Foi verificado, por experimentação, que as estruturas instáveis da treliça de mapa acoplado podem ser estabilizadas para um certo intervalo de p . Conforme I aumenta, o intervalo de p decresce para um controle eficaz. No caso de um valor de I maior do que I_m , que é definido como sendo a densidade mínima de pinos para estabilizar o modelo, a densidade de pinos fica tão baixa que o modelo não consegue se estabilizar no estado desejado. No entanto,

conforme a intensidade de pinagem p aumenta, o caos do sistema torna-se cada vez mais eficazmente controlado.

Em 2003, He *et al.* propuseram um método de controle por pinagem para a rede neural caótica de Adachi e Aihara (1997). Esta rede neural caótica apresenta dinâmica espaço-temporal mais complexa que a treliça de mapa acoplado além de terem conexões neurais mais complicadas, levando a proposição de um controle por pinagem distinto daquele utilizado na treliça (Equações 3.21 e 3.22). A rede neural caótica com o controle por pinagem é descrita da seguinte maneira:

$$x_i(t+1) = f_i(\eta_i(t+1) + \zeta_i(t+1)) \quad (3.23)$$

em que :

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^m w_{ij} \left\{ x_j(t) + \sum_{l=0}^{m/I} \delta(j-l-1) K u_j(t) \right\} \quad (3.24)$$

$$u_j(t) = x_j(t) - (1 - \tilde{x}_j) \quad (3.25)$$

$$\zeta_i(t+1) = k_r \zeta_i(t) - \rho x_i(t) + a_i \quad (3.26)$$

em que $\delta(\cdot)$ é uma função pulso, a qual é usada para distinguir os neurônios ‘pinados’ (*pinned*) daqueles que não são; u_j é o sinal de controle, o qual é construído ao se comparar a última saída e a saída esperada (\tilde{x}_j); I é a distância de pinagem (*pinning distance*), que é a distância entre os dois neurônios pinados mais próximos, e K é a intensidade do controle.

Pode ser notado que a Equação 3.25 utiliza o inverso do padrão armazenado desejado ($1 - \tilde{x}_j$) como controle alvo. Isto ocorre porque He *et al.* (2003) observaram, por simulação, que ao adicionar o controle por pinagem no sistema utilizando-se o próprio controle alvo, a rede se estabilizava no inverso deste controle alvo. No entanto, simulações mostraram que, ao se utilizar o inverso do padrão armazenado desejado como controle alvo, a rede se estabilizava corretamente no padrão desejado.

He *et al.* (2003) mostraram que a dinâmica da rede neural caótica pode ser estabilizada em um padrão armazenado desejado se a intensidade do controle for maior do que seu limiar e se a densidade de pinagem (*pinning density*) não for muito pequena. He *et al.* (2003) apresentam uma ilustração na qual, a partir de uma versão ruidosa do padrão \mathbf{X}^4 da Figura 3.2, a rede de memória auto-associativa caótica com controle de caos (Equações 3.23 a 3.26), para os valores de parâmetros $k_r=0.95$, $k_f=0.2$, $\rho=10.0$, $a_i=2.0$ ($i=1,2,\dots,100$), $I=2$, $K=20.0$, e para $\tilde{x}_j = \mathbf{X}^4$, estabiliza-se no padrão armazenado desejado (\mathbf{X}^4).

3.3 Saltar Passos Transforma um Noviço em um Especialista

Uma das muitas características interessantes da dinâmica caótica é que o caos tem diversas funções potencialmente úteis (Aihara, 2002). Dentre as várias aplicações possíveis, o presente trabalho escolheu um processo cognitivo particular. Esta seção tem como objetivo o entendimento deste processo cognitivo a ser simulado por redes neurais caóticas propostas no Capítulo 4.

3.3.1 Caracterização do Saltar Passos

De acordo com Koedinger e Anderson (1990), à medida que as pessoas resolvem o mesmo tipo de problema repetidas vezes, elas se tornam mais rápidas com a repetição e freqüentemente mudam o processo utilizado para resolvê-lo. Tal reorganização pode permitir que alguns passos originais do problema sejam saltados, de modo que os problemas passem a ser resolvidos de forma mais fácil, rápida e eficiente. Em 1996, Blessing e Anderson investigaram o fenômeno de resolver problemas de modo mais eficiente conhecido como **saltar passos** (*step skipping*).

O salto de passos é freqüentemente entendido como sendo um processo composicional, no qual uma pessoa que costumava usar n passos para realizar uma tarefa ou resolver um problema, passa a realizá-la em m passos, em que $n > m$. Intuitivamente, se a pessoa utiliza um menor número de passos para completar uma tarefa, significa que ela é mais eficiente em seu desempenho e gasta menos tempo nessa atividade.

Pessoas que aplicam explicitamente cada operador ensinado para resolver um tipo de tarefa são chamadas de *noviças*, enquanto aquelas pessoas que conseguem resolver o mesmo tipo de tarefa usando um menor número de operadores são chamadas de *especialistas* nesse domínio particular. Segundo Blessing e Anderson (1996), o fenômeno de saltar passos não ocorre em tarefas de qualquer tipo de domínio. Por exemplo, a contagem de objetos ou o cálculo de adição não se enquadram no fenômeno, pois pessoas contando ou somando não podem pular os objetos a serem contados ou somados. Além disso, passos essenciais à realização da tarefa também não podem ser saltados. Por exemplo, quando o passo a ser saltado corresponde ao resultado desejado, este passo não pode ser saltado. Quando a tarefa é realizada em um único passo, sem que existam passos intermediários, também não há como ocorrer o salto de passos.

Os tipos de tarefas mais fáceis de ilustrar o fenômeno de saltar passos são aquelas tipicamente formais, tais como a álgebra e a física. Em tais domínios, os passos são caracterizados pela aplicação de uma regra, de modo que, dado o estado atual do problema a regra é aplicada e um próximo estado no caminho da solução do problema é alcançado. Nestes domínios, os passos a serem saltados são os passos intermediários entre o estado atual do problema e a sua solução. Tal efeito pode ser aplicado ao se reconhecer qualquer tipo de padrão que esteja contido no problema e

que possa ser diretamente alcançado, sem que outros possíveis padrões intermediários necessitem ser antes atingidos.

Para exemplificar o saltar passos, Blessing e Anderson (1996) descrevem o seguinte caso de estudo: quando se pede a um estudante, um aprendiz de álgebra, que a equação $-x - A = B$ seja resolvida, ele provavelmente não estará apto a saltar imediatamente da equação apresentada para a equação final, ou seja, para a solução do problema. Ele deverá agir como um noviço e resolver a equação passo-a-passo, conforme ilustrado na Tabela 3.2.

Tabela 3.2. Passos utilizados na resolução de uma equação algébrica.

Equação dada $\Rightarrow -x - A = B$		
Passo	Ação a ser realizada	Equação Resultante
Passo 1	Adicionar + A em ambos os lados da equação	$-x - A + A = B + A$
Passo 2	Realizar a operação $-A + A$	$-x = B + A$
Passo 3	Multiplicar ambos os lados da equação por -1	$x = -B - A$

Pessoas com experiência em resolver este tipo de equação, ao receberem a tarefa de solucionar $-x - A = B$, provavelmente dão a resposta $x = -B - A$ imediatamente, em um único passo, saltando dois passos (o passo 1 e o passo 2) incluídos no procedimento do noviço. Tais pessoas são chamadas especialistas neste tipo de tarefa.

Apesar do fenômeno de saltar não ser possível para todos os tipos de tarefas, os efeitos do saltar passos puderam ser verificados e testados em outros tipos de tarefas distintas de operações algébricas, inclusive em domínios menos formais. Os seguintes experimentos, por exemplo, também mostraram evidências da ocorrência de saltar passos: tarefas envolvendo circuitos eletrônicos (Hindes, Patterson e Pfeffer, 2001), tarefas envolvendo aquisição de habilidades cognitivas (tais como xadrez e composição musical) (Wenger, 1999), tarefas envolvendo desenhos de figuras humanas (Blessing e Anderson, 1996). Neste último caso, por exemplo, desenhistas iniciantes (os noviços) costumam aprender primeiro a fazer um desenho oval para gerar a cabeça e, a seguir, são feitas mais algumas dessas figuras geométricas para as outras partes do corpo. São necessárias ainda marcas de apoio para partes como ombros e quadris. À medida que o desenhista se torna mais especializado, ele começa a pular algumas figuras ovais e algumas marcas de apoio. Este tipo de saltar passos coincide com um entendimento mais intuitivo de como realizar tarefas (no caso, de desenhar figuras humanas).

3.3.2 Explicações do Saltar Passos

Esta seção apresenta descrições de alguns dos mecanismos de aprendizagem que têm sido propostos para se saltar passos na solução de um problema. Foi visto que, freqüentemente, os noviços aplicam explicitamente, em um domínio, cada operador necessário na solução do problema, enquanto os especialistas saltam passos durante sua solução particular do problema. Blessing e Anderson (1996) entendem que o procedimento de solução que os especialistas usam costuma ser organizado de modo diferente daquele dos noviços.

Diversas teorias de aquisição de habilidades tentam responder por mudanças no processo de resolução de problemas, incluindo as mudanças ocorridas no fenômeno de saltar passos. Especificamente para o fenômeno saltar passos, apesar das abordagens teóricas que o explicam diferirem em seus detalhes, elas costumam envolver a mesma estrutura de resolução de problemas (Newell e Simon, 1972), a qual geralmente faz uso de operadores e regras de produção. Um operador é uma regra que pode ser aplicada para transformar o estado atual do problema em um outro estado. Operadores como regras de produção que são declarações do tipo “se-então” atuando no estado atual do problema.

Um bom número dessas explicações do saltar passos correspondem à afirmação de que pular passos envolve substituir operadores de estruturas menores por operadores de estruturas maiores. São três as explicações dadas que envolvem este tipo de substituição:

Composição de regras

1. Criação de pedaços (*chunking*)
2. Compilação de regras por analogia

Uma outra visão de explicação de saltar passos é dada pela teoria de Logan (1988, 1990), a qual se difere das 3 anteriores por afirmar que o desempenho especialista não é verdadeiramente caracterizado pela aplicação de operadores, mas sim pela recuperação de um exemplo específico.

Através de breves descrições, serão apresentadas, a seguir, as idéias principais de cada uma das explicações do saltar passos citadas acima.

- **Composição de Regras**

Um ponto de vista de como ocorre o saltar passos, defendido por Anderson (1983) e Lewis, (1978), é através de operadores colapsantes (*collapsing operators*), que seguem-se um ao outro na solução de um problema (chamados de operadores adjacentes), dentro de um único operador. Tal organização ou encadeamento de operadores é chamado de composição de regras.

Estes operadores de composição supostamente produzem o comportamento de saltar passos por não utilizarem os passos intermediários que o operador único produz. Neste tipo de explicação

do saltar passos enquadra-se o exemplo de solucionar a equação algébrica $-x - A = B$ citado na Seção 3.3.1. Neste exemplo, quando um iniciante no assunto soluciona a equação, ele provavelmente resolve o problema em três passos, ou seja, aplica três operadores na solução: adicionar $+A$ em ambos os lados da equação, simplificar o termo $-A + A$ no lado esquerdo da equação, e multiplicar toda a equação por -1 , a fim de alcançar a resposta $x = -B - A$. De acordo com a teoria da composição de regras, o operador de composição une os dois primeiros passos desta solução algébrica em um único operador, de modo a produzir diretamente $-x = B + A$, e, a seguir, une este novo passo de operação criado com o último passo da solução original, criando assim um único operador capaz de produzir o resultado final.

A teoria da composição de regras é uma idéia atrativa porque fornece uma avaliação bem concisa de como o saltar passos ocorre. No entanto, em 1981, Lewis publicou um trabalho no qual lança dúvidas se o saltar passos pode ser meramente devido à composição de passos adjacentes. Nesse trabalho ele investiga o que chama de operadores poderosos (*powerful operators*). Tais operadores fazem o trabalho de uma série de outros operadores, assim como a composição de operadores na álgebra dada acima, mas sua crítica é que os operadores poderosos podem combinar passos não adjacentes, condição que não ocorre na composição de regras padrão. Por exemplo, um noviço provavelmente resolveria a equação $x + 2(x+1) = 4$ primeiro multiplicando 2 por x e por 1 e, então, adicionando x com $2x$. É provável que um especialista responderia imediatamente $3x = 2$, com o $2x$ sendo multiplicado e adicionado ao x , antes que o 2 fosse isoladamente multiplicado por x e por 1. Assim, passos não adjacentes também podem ser combinados para produzirem um único passo.

- **Criação de Pedacos (*Chunking*)**

A teoria de cognição de Newell (1990), chamada *SOAR*, tem um mecanismo que tenta fornecer uma resposta psicologicamente plausível para todo tipo de aquisição de habilidade. Este mecanismo, chamado criação de pedacos (*chunking*), opera quando ocorre um impasse no processo de solução de um problema. O impasse ocorre quando o sistema *SOAR* alcança um estado no qual:

- nenhum operador se aplica;
- não existe conhecimento suficiente para se aplicar um operador.

Quando ocorre o impasse, o *SOAR* estabelece uma submeta para resolvê-lo, criando assim um novo espaço de problema para resolver a submeta. Resolvida a submeta, o *SOAR* cria um novo operador, um naco ou pedaço (*chunk*), usando os resultados obtidos após a solução do impasse.

O pedaço é construído a partir dos resultados do impasse e é encadeado para trás através de elementos anteriores ao impasse e que geraram estes resultados. O operador que representa o naco é, portanto, construído ao se usar elementos originais que ocorrem antes do impasse como a

condição no lado de produção da regra e ao se usar os resultados após a solução do impasse lado da ação da regra.

Os pedaços podem ser pensados como um operador de composição, uma vez que os operadores usados para resolver a submeta poderiam ser representados pelo naco resultante. Sendo assim, a proposta de Newell é que uma pessoa adquira habilidade através do processo chamado criação de pedaços, no qual as respostas são adquiridas em nacos de vários tamanhos. Sob tal esquema, um noviço constrói número maior de pedaços, enquanto um especialista necessitaria de número menor para solucionar dado problema. Por exemplo, um especialista poderia criar um único naco, na forma de uma produção, capaz de produzir a solução $x = -B - A$ a partir da equação $-x - A = B$.

Newell ressalta que dentro do SOAR está não apenas o processo pelo qual um sistema resolve dado problema, mas o produto do processo também está presente. Por causa disto, o SOAR evita o problema dos operadores não adjacentes que ocorre na teoria da composição de regras.

- **Compilação de Regras por Analogia**

A teoria denominada controle adaptativo do pensamento – teoria racional (*adaptive control of thought – rational theory – ACT-R*) de Anderson (1993) explica que qualquer habilidade procedimental nova é adquirida por analogia a exemplos. Dentro da estrutura *ACT-R*, habilidades cognitivas são realizadas através de regras de produção. O único meio pelo qual novas regras de produção podem entrar no sistema é através de analogia a um exemplo já existente na memória declarativa. Sendo assim, quando uma dada tarefa deve ser feita e ainda não existe regra de produção aplicável a ela, o *ACT-R* tenta localizar um exemplo declarativo adquirido previamente e que exemplifique a aplicação de um operador correto, similar ao objetivo atual. Caso um exemplo similar o suficiente seja encontrado, o mecanismo de analogia cria uma regra de produção que gerará a ação correta.

De acordo com esta teoria, noviços resolvem problemas por referência a exemplos, enquanto especialistas aplicam operador. Saltar passos neste mecanismo de analogia seria então formar analogias com exemplos que já incorporaram a aplicação de múltiplos operadores.

- **Recuperação de Exemplos**

Na teoria de Logan (1988, 1990), o desempenho da habilidade é geralmente caracterizado pela recuperação de uma memória específica, e não pelo uso de um procedimento. Nesta teoria, o desempenho de uma tarefa é medido pela taxa entre (i) executar o operador apropriado e (ii) recuperar um exemplo específico. Quando iniciantes na execução de uma tarefa, as pessoas geralmente aplicam todos os operadores necessários. Eventualmente uma associação

suficientemente forte será criada entre um estímulo específico e sua resposta apropriada, e então pode começar a acontecer a recuperação direta da resposta.

Em comparação à teoria de Anderson (1993), Logan faz uma predição oposta: a de que o desempenho de novatos é marcado pela aplicação de operadores, enquanto a habilidade de desempenho de especialistas é feita pela recuperação de exemplos passados. Nesta teoria, o saltar passos ocorreria quando fosse possível a recuperação direta da memória na solução de um problema ou na execução de uma tarefa. Se feita a suposição de que o traço de memória recuperado da solução do problema pode conter múltiplos passos desta solução, então a pessoa encarregada de resolver o problema pode produzir apenas o último passo contido no caso, eliminando passos intermediários e assim permitindo que a pessoa salte passos intermediários.

As teorias de Anderson (*ACT-R*) e de Logan não são necessariamente mutuamente excludentes. Existem evidências que apóiam ambas. Por exemplo, em um estudo feito por Anderson e Fincham em 1994, alguns participantes receberam apenas exemplos de aplicação de regras. Os participantes foram capazes de extrair regras corretas destes exemplos e de aplicá-las aos problemas que lhe foram apresentados. Mesmo na ausência de exemplos repetidos, os participantes apresentaram um alto índice de aprendizagem, ao contrário do que a versão forte da teoria de Logan previa. Entretanto, de acordo com Blessing e Anderson (1996), estudos posteriores não publicados de Anderson e Fincham revelaram que a aprendizagem torna-se mais rápida quando se têm exemplos repetidos. Carlson e Lundy, em 1992, também testaram a habilidade de pessoas aprenderem regras. O método deles variava a consistência de ambos os exemplo e a seqüência da operação. Os resultados apontaram benefícios de velocidade de aprendizagem tanto no caso dos dados apresentados serem consistentes (ou seja, no caso de aparecerem exemplos repetidos) quanto no caso de a ordem das operações ser consistente.

3.3.3 Discussões sobre o Saltar Passos

Ao se verificar a ocorrência de mudança de organização no procedimento de solução de um problema entre novatos e especialistas dentro do fenômeno saltar passos, algumas questões, tais como (i) o que ocorre durante a mudança do número de passos na solução do problema? (ii) um especialista apenas salta passos? (iii) sob quais circunstâncias inicia-se fenômeno de saltar passos?, merecem ser melhor examinadas. Pode-se também questionar se, ao saltar passos na solução de um problema, as pessoas estão realmente pulando passos mentalmente ou se estão apenas pulando passos fisicamente. Ou seja, as pessoas passam por cada passo do problema mentalmente e apenas executam fisicamente alguns dos passos ou de fato conseguem mentalmente ir direto do primeiro passo para o último? Na teoria de aquisição de habilidades estes modos de saltar passos são

referidos como saltar passos declaradamente (*overtly skipping steps*) ou saltar passos disfarçadamente (*covertly skipping steps*).

Em seu artigo de 1996, Blessing e Anderson procuram responder algumas destas questões através de dois experimentos, nos quais as pessoas aprendem as regras da tarefa ou do problema e, com estas regras, devem solucioná-lo. Os procedimentos de solução dos problemas foram monitorados e protocolos verbais foram tomados. As ações dos participantes, seus conhecimentos e uso das regras foram examinados entre dois pontos: no início da tarefa e no final. No artigo citado, o desempenho de saltar passos foi analisado dentro de domínios particulares, similares à álgebra.

O primeiro experimento teve como objetivo explorar a questão de quando e como as pessoas começam a saltar passos quando estão resolvendo um problema. Para isto, os participantes foram instruídos nas regras da tarefa e conduzidos a resolver os problemas, mas não foram forçados a fazer todos os passos explicitamente. Foi criada uma situação na qual se poderia estudar como participantes aprenderiam a pular passos naturalmente. No início da execução da tarefa, os participantes a executaram de acordo com as regras dadas, mas logo passaram a não se referir de volta às regras.

A análise das perguntas que o primeiro experimento procura examinar apontaram para as seguintes respostas:

- **Como ocorreu o saltar passos?** No início do experimento, os participantes moviam os operadores do lado esquerdo da equação algébrica dada para o lado direito e depois mexiam no operador em frente à variável x . No final do experimento, os participantes primeiro consideravam o operador em frente à variável x e então faziam as operações de transposição. Isto representou a principal mudança estratégica. Outra diferença encontrada entre o início e o final das operações foi a constatação de que os mesmos participantes, no final, estavam usando atalhos. No início eles apenas seguiam as regras dadas. No final, eles criavam alguns atalhos para ajudar na manipulação das regras.
- **Quando começou o saltar passos?** Logo que os participantes começaram a solucionar os problemas, eles começaram a saltar passos. Muitos começaram a saltar passos após dois ou três problemas resolvidos. A princípio, quando os participantes começaram a saltar passos, eles estavam meramente fazendo exatamente todas as mesmas operações em suas cabeças e estavam apenas declaradamente saltando passos. Mais tarde, conforme os participantes tornaram-se mais habilidosos na solução do problema, os resultados sugeriram que os participantes estavam realmente mudando a regra base e, portanto, também estavam saltando passos mentalmente.

Quando se começa a saltar passos é, de fato, altamente plausível que se está passando por cada passo do problema mentalmente e executando fisicamente apenas alguns dos passos, pois não houve ainda chance de se estudar o modelo de saltar passos antes de iniciar a resolução dos problemas. Como estes participantes não tem experiência com os resultados das combinações de passos, fica difícil imaginar que eles inicialmente já consigam regras para produzir os resultados diretamente.

O segundo experimento realizado por Blessing e Anderson (1996) teve como objetivo verificar se as pessoas podem iniciar o salto de passos em qualquer ponto do experimento, inclusive na primeira tentativa de solução do problema. Isso significa investigar se o saltar passos não requer que primeiro se execute os saltos de passos da solução do problema declaradamente para depois executá-los mentalmente. Para este experimento, os participantes são informados quando podem começar a saltar passos. Ao exercer mais controle sobre o comportamento de saltar passos dos participantes, uma melhor análise pode ser feita da diferença entre executar os passos explicitamente e saltar passos.

Um grupo de participantes foi instruído a saltar todos os passos intermediários imediatamente, de modo que a primeira regra escolhida daria a solução do problema. Outros dois grupos começaram a executar os passos do problema explicitamente e, em diferentes pontos do experimento, foram informados de que deveriam começar a saltar todos os passos intermediários. Conforme esperado, o grupo do primeiro experimento, aqueles que tinham começado a saltar passos de imediato, tiveram dificuldade para começar a desempenhar a tarefa. No entanto, embora no início os participantes que começaram a saltar passos imediatamente demorassem a resolver cada passo do experimento, o maior número de passos executados pelos outros participantes fizeram com que aqueles que não saltaram passos desde o início acabassem sendo ainda mais lentos para completar o experimento. Além disso, os participantes que resolveram os problemas com maior número de passos acabaram cometendo mais erros.

Blessing e Anderson (1996) concluíram que, quando examinados conjuntamente, os dois experimentos citados acima permitem uma avaliação do processo de saltar passos. Segundo os autores, os resultados do primeiro experimento fornecem uma visão qualitativa de como as pessoas começam a saltar passos quando aprendem a resolver sozinhas os problemas. A princípio, os participantes executaram todos os passos intermediários explicitamente. Apenas depois de os participantes começarem a aplicar cada regra do experimento ao menos duas vezes é que eles começaram a saltar passos declaradamente. Entretanto, o segundo experimento indicou que saltar passos declaradamente foi uma escolha arbitrária dos participantes, e eles foram capazes de saltar passos imediatamente. Tal experimento também indicou que a decisão de saltar passos acarretava

em uma redução substancial no tempo total de resolução do problema, devido à redução do tempo de execução.

No Capítulo 5, Seção 5.4, o efeito de saltar passos será modelado por redes de memória hetero-associativas caóticas, de modo que os passos intermediários na realização de uma tarefa poderão ser eliminados. Será utilizada a abordagem denominada composição de regras e o efeito do saltar passos será realizado em operações algébricas. Será verificado que, conforme os passos intermediários são eliminados, a rede passa a comportar-se como um especialista na resolução de uma tarefa e passa a alcançar o estado final desejado em um intervalo de tempo menor.

3.4 Resumo

Um modelo de neurônio simples com dinâmica caótica (o modelo neurônio caótico) e um modelo de rede neural artificial composto por tais neurônios caóticos (a rede neural caótica) foram propostos por Aihara, Takabe e Toyoda (1990).

O modelo neurônio caótico é obtido ao se modificar e estender os modelos de neurônios discretos no tempo de Caianiello (1961) e de Nagumo e Sato (1972). Este modelo neurônio caótico, proposto por Aihara, Takabe e Toyoda (1990) é um modelo simples de neurônio individual capaz de qualitativamente reproduzir seqüências alternadas de respostas periódicas-caóticas experimentalmente observadas em axônio de lulas gigantes. Estes autores mostraram uma estrutura de rede neural caótica que tornou possível introduzir funções do caos determinístico em redes neurais artificiais sempre que necessário, porque o modelo de rede neural caótica inclui alguns modelos convencionais de redes neurais artificiais como casos especiais.

Uma rede auto-associativa é construída com os modelos neurônio caóticos interconectados através de uma matriz auto-associativa convencional. Foram analisadas as características de recuperação da rede de memória auto-associativa caótica. Foi mostrado que, enquanto nas redes de memória associativas convencionais a sua propriedade de recuperação tem sido avaliada principalmente por seu estado de equilíbrio estável, ou seja, avalia-se se o estado de equilíbrio estável corresponde a um padrão armazenado, na rede de memória auto-associativa caótica é gerado um comportamento curioso e peculiar, no qual a rede recupera um padrão armazenado e escapa dele após alguns instantes de tempo. Em sua dinâmica, a rede nunca se estabiliza em um ponto de equilíbrio, mas move-se não periodicamente na vizinhança dos padrões armazenados. Isto é interessante porque o processo de busca da memória é mais dinâmico do que o da rede associativa convencional. No entanto, foi visto que há situações nas quais se deseja evitar ou suspender a dinâmica caótica. He *et al.* (2003) propuseram o método de controle por pinagem para

a rede de memória auto-associativa caótica de Adachi e Aihara (1997), o qual foi capaz de estabilizar tal rede em um padrão armazenado desejado.

Na análise da dinâmica encontrada para a rede de memória auto-associativa caótica foi destacada apenas a existência do caos determinístico. Pela riqueza do processo de busca da memória, é provável que outros comportamentos dinâmicos, tais como bifurcações e crise pudessem também estar ocorrendo neste modelo de rede.

O fenômeno chamado saltar passos foi adicionado neste capítulo com o objetivo de apresentar o contexto do efeito que será trabalhado e modelado nas redes de memória hetero-associativa caóticas criadas nesta tese. Será visto no Capítulo 5 como a dinâmica das redes de memória hetero-associativas caóticas permitem que o fenômeno de saltar passos seja por elas modelado. A explicação do saltar passos utilizada para esta modelagem será a composição de regras e as tarefas selecionadas envolverão operações algébricas.

O capítulo seguinte propõe uma extensão do modelo de rede de memória auto-associativa caótica para o modelo de rede de memória hetero-associativa caótica.

Capítulo 4

Família de Modelos de Memória Associativa Bidirecional Caótica (Família C-BAM)

Capítulos anteriores descrevem os fenômenos presentes na teoria do caos (Capítulo 2), descrevem a ocorrência e o controle da dinâmica caótica em modelos de memória auto-associativas e contêm a descrição do efeito saltar passos na transformação do novíço em especialista (Capítulo 3). O presente capítulo tem como objetivo descrever a criação da família de modelos na qual os fenômenos caóticos, a recuperação de memórias inacessíveis e o processo cognitivo para salto de passos serão aplicados.

Na primeira parte do capítulo é apresentado o paradigma original selecionado para este trabalho (memória associativa bidirecional - BAM), suas características (tais como tipo de arquitetura, a natureza das associações entre padrões) e as equações que o descrevem. A segunda parte do capítulo descreve a família de modelos de memória associativa bidirecional caótica (família C-BAM), construída através da inclusão do neurônio caótico nas redes BAM originais.

Convém destacar que paradigmas do tipo “memórias associativas” armazenam e recuperam padrões ou pares de padrões. A importância das memórias associativas está relacionada à maneira como o cérebro humano aprende e armazena informações. A associação é uma característica proeminente da memória humana, de modo que a memória associativa pode ser reconhecida como uma das funções primárias do cérebro. Facilmente uma pessoa, ao sentir um cheiro conhecido de perfume, associa-o a alguém que tem o costume de usá-lo. É comum em programas de TV apresentarem jogos em que, ao se mostrar em uma foto apenas parte do corpo de uma pessoa famosa, o jogador descobrir quem está na foto.

Qualitativamente, uma memória associativa pode ser descrita como um sistema de memória em que, dada uma informação indicando um ponto de partida (uma entrada de dados), tal memória disponibiliza informações anteriormente armazenadas e associa uma resposta à informação apresentada como entrada. Portanto, uma propriedade fundamental da memória associativa é que ela associa um padrão de saída de uma atividade neural a um padrão de entrada desta atividade.

Em termos formais, Araújo (1994) apresenta memória associativa como sendo uma transformação de $T: R^n \rightarrow R^m$, geralmente não linear, a qual mapeia vetores \mathbf{x}_k do espaço R^n em vetores \mathbf{y}_k do espaço R^m . Ao ser apresentado um vetor $\mathbf{x}_k(0)$ como entrada de uma memória associativa, este vetor deve ser suficientemente similar ao vetor desejado \mathbf{x}_k para que ocorra a recuperação do vetor \mathbf{y}_k associado. Caso o conjunto de vetores \mathbf{x}_k for idêntico ao conjunto de vetor \mathbf{y}_k , essa memória associativa é chamada de auto-associativa. No caso de \mathbf{x}_k e \mathbf{y}_k serem conjuntos distintos, tal memória é classificada como sendo hetero-associativa.

A escolha da família derivada da rede de memória associativa bidirecional (BAM) (Kosko, 1987) como paradigma para este trabalho, visando estudar comportamentos complexos que emergem em redes neurais com a inclusão de neurônios caóticos (Aihara, Takabe e Toyoda, 1990) e posterior aplicação destes comportamentos no processo de saltar passos na aquisição de conhecimento, foi feita devido às seguintes propriedades da BAM: fácil implementação, capacidade de lidar com pares de padrões, baixa capacidade de memória na BAM (característica melhorada na BAM com atraso e na BAM exponencial), existência de padrões memorizados não acessíveis para recuperação, convergência comprovada e a ocorrência de padrões espúrios.

Na literatura podem ser encontradas diversas redes neurais relacionadas com a memória associativa bidirecional (BAM), como por exemplo a DCBAM (*Discrete Chainable Bidirectional Associative Memory*) (Zhou e Quek, 1996), a BAM Assimétrica (Xu, Leung e He, 1994), a BAM com lógica fuzzy (Belohlavek, 2000), entre outras. Para o estudo proposto neste trabalho foram selecionados três tipos de modelos BAM: a BAM original (Kosko, 1987), a BAM com atraso (Zhang e Yang, 2001) e a BAM exponencial (Jeng, Yeh e Chiveh, 1990).

A primeira seção deste capítulo descreve sucintamente a família original das redes tipo BAM selecionadas. A segunda seção tem o propósito de descrever como o neurônio caótico foi utilizado para criar a família C-BAM, mostrar a diferença entre as famílias BAM e C-BAM e apresentar as equações que descrevem a dinâmica das redes da família C-BAM. No final, um resumo evidencia os principais resultados do capítulo.

4.1 Modelos Originais de Memória Associativa Bidirecional

Esta seção descreve a classe original de redes BAM selecionadas para o presente trabalho. São três os modelos selecionados: BAM (Seção 4.1.1), BAM com atraso (Seção 4.1.2) e BAM exponencial (Seção 4.1.3).

4.1.1 Modelo Memória Associativa Bidirecional Discreto (BAM)

O modelo chamado Memória Associativa Bidirecional (BAM), proposto por Kosko (1987, 1988), é uma das versões mais simples de rede hetero-associativa com recorrência. Sua estrutura é a de uma rede neural artificial de duas camadas não lineares com realimentação. Como pode ser visto na Figura 4.1, os neurônios de uma camada são completa e simetricamente conectados com os neurônios de outra camada, sem a presença de conexões laterais.

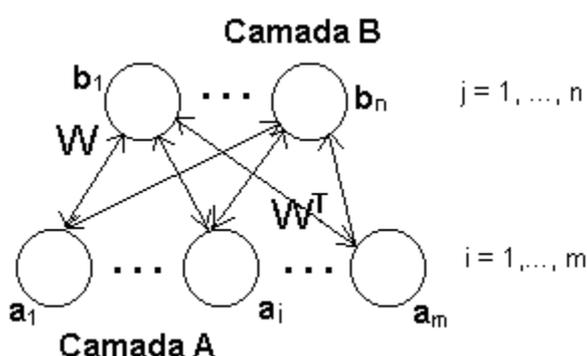


Figura 4.1. Topologia do modelo BAM.

O modelo BAM comporta-se como uma memória hetero-associativa de conteúdo endereçável, cujo objetivo original constitui-se em armazenar e recuperar p pares de vetores binários $(\mathbf{A}^1, \mathbf{B}^1), \dots, (\mathbf{A}^p, \mathbf{B}^p)$, em que $\mathbf{A} \in \{0,1\}^m$ e $\mathbf{B} \in \{0,1\}^n$.

A maneira de se armazenar as associações desejadas em BAM é pela construção de uma matriz de correlação, a qual distribui a informação de $(\mathbf{A}^k, \mathbf{B}^k)$ em uma estrutura matricial. Dessa forma, os p pares de vetores são armazenados na rede pela matriz correlação de pesos $\mathbf{W}_{n \times m}$, construída utilizando-se o modelo de associação linear da Equação 4.1:

$$\mathbf{W} = \sum_{k=1}^p \mathbf{A}^k (\mathbf{B}^k)^T \quad (4.1)$$

As buscas associativas por pares ou associações armazenadas $(\mathbf{A}^k, \mathbf{B}^k)$ são feitas por meio de um fluxo de informações de duas direções: para frente (*forward*) e para trás (*backward*). A informação passada através de \mathbf{W} fornece uma direção de fluxo, enquanto \mathbf{W}^T fornece a outra

direção. O adjetivo bidirecional do modelo BAM deriva-se deste tipo de fluxo. Tal bidirecionalidade implica em uma relação entre \mathbf{A}^k e \mathbf{B}^k , de modo que, quando \mathbf{A}^k for apresentada à rede, seja possível obter-se \mathbf{B}^k , sendo também possível o procedimento inverso.

A recuperação de associações armazenadas na BAM é feita por um procedimento de realimentação não linear. O par $(\mathbf{A}^k, \mathbf{B}^k)$ define o estado da BAM. Cada neurônio a_i na camada A e cada neurônio b_j na camada B calcula a soma de entradas que vêm dos neurônios da outra camada da rede e, então, após examinar se o valor calculado é superior, igual ou inferior a um valor de limiar, seu estado é atualizado. Um neurônio em particular será ativado (+1) se a soma de entrada for superior ao limiar, será desativado (0 ou -1) se a soma for inferior e irá manter seu estado atual se a soma for igual ao valor do limiar. Em 1987, Kosko assumiu valor zero como o limiar para todos os neurônios.

A soma de entrada para um neurônio a_i e para um neurônio b_j é, respectivamente:

$$BW_i^T = \sum_j b_j w_{ij} \quad (4.2)$$

$$AW_j = \sum_i a_i w_{ji} \quad (4.3)$$

A não linearidade é introduzida na BAM pela função limiar (Equações 4.4 e 4.5) para os neurônios a_i e b_j :

$$a_i = \begin{cases} 1, & \text{se } BW_i^T > 0 \\ 0, & \text{se } BW_i^T < 0 \end{cases} \quad (4.4)$$

$$b_j = \begin{cases} 1, & \text{se } AW_j > 0 \\ 0, & \text{se } AW_j < 0 \end{cases} \quad (4.5)$$

Segundo Kosko (1987), o esquema aditivo da Equação 4.1 implica que, ao se usar apenas vetores binários, \mathbf{W} não conterá sinapses inibitórias, de modo que as somas das entradas BW_i^T e AW_j nunca serão negativas. Como vetores bipolares não produzem este problema, Kosko (1988) afirma que a codificação bipolar geralmente é melhor do que a codificação binária. Para atender a esta argumentação de Kosko, que pode ser encontrada em detalhes no seu artigo de 1988, antes de se montar \mathbf{W} , os padrões de memória binários serão transformados em padrões bipolares:

$$\mathbf{W} = \sum_{k=1}^p \mathbf{X}^k (\mathbf{Y}^k)^T \quad \text{em que} \quad \begin{cases} \mathbf{X}^k = 2\mathbf{A}^k - 1 \\ \mathbf{Y}^k = 2\mathbf{B}^k - 1 \end{cases} \quad (4.6)$$

Formalmente, são estas as equações de recuperação da rede BAM original discreta:

$$x_i(t+1) = S\left(\sum_{j=1}^n w_{ij} y_j(t)\right) \quad i=1, \dots, m \quad (4.7)$$

$$y_j(t+1) = S\left(\sum_{i=1}^m w_{ji} x_i(t)\right) \quad j=1, \dots, n \quad (4.8)$$

em que:

$$S(a) = \begin{cases} 1, & \text{se } a > 0 \\ -1, & \text{se } a < 0 \\ S(a), & \text{se } a = 0 \end{cases} \quad (4.9)$$

Descrita desta maneira, a rede BAM permite que, por exemplo, dado um padrão \mathbf{X} , obtenha-se um padrão \mathbf{Y} e, dado um padrão \mathbf{Y} , obtenha-se um padrão \mathbf{X} , numa seqüência de recuperação dada por:

$$\begin{array}{l} \mathbf{X} \rightarrow \mathbf{W} \rightarrow \mathbf{Y}' \\ \mathbf{X}' \leftarrow \mathbf{W}^T \leftarrow \mathbf{Y}' \\ \mathbf{X}'' \rightarrow \mathbf{W} \rightarrow \mathbf{Y}'' \\ \mathbf{X}'' \leftarrow \mathbf{W}^T \leftarrow \mathbf{Y}'' \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{X}^f \rightarrow \mathbf{W} \rightarrow \mathbf{Y}^f \\ \mathbf{X}^f \leftarrow \mathbf{W}^T \leftarrow \mathbf{Y}^f \end{array}$$

Após o padrão \mathbf{X} ser apresentado à rede, os neurônios começam a realizar suas mudanças de estado até que um estado bidirecionalmente estável seja encontrado. No exemplo acima, $(\mathbf{X}^f, \mathbf{Y}^f)$ é o estado final estável para o qual a rede evoluiu.

Idealmente, o par $(\mathbf{X}^f, \mathbf{Y}^f)$ seria a resposta desejada para o padrão de entrada fornecido à rede. No entanto, na prática, o ponto fixo $(\mathbf{X}^f, \mathbf{Y}^f)$ pode coincidir com o par desejado ou pode ser um outro par (atrator espúrio). Tal atrator espúrio pode ser um outro par $(\mathbf{X}^f, \mathbf{Y}^f)$ armazenado pela rede, pode ser o par inverso do par desejado ou ainda, um novo par diferente daqueles armazenados.

Para que a rede sempre convirja para um estado final, deve ser mostrado que a matriz \mathbf{W} é uma matriz bidirecionalmente estável. Kosko (1988) prova que “toda matriz real arbitrária $\mathbf{W}_{n \times m}$ é bidirecionalmente estável” e, portanto, um estado estável é alcançado por qualquer matriz \mathbf{W} . O autor da BAM também demonstra que o estado estável alcançado corresponde a um mínimo local da função de energia E (Kosko, 1988):

$$E(\mathbf{X}^r, \mathbf{Y}^r) = -1/2 \mathbf{X}^r \mathbf{W} (\mathbf{Y}^r)^T - 1/2 \mathbf{Y}^r \mathbf{W}^T (\mathbf{X}^r)^T = -\mathbf{X}^r \mathbf{W} (\mathbf{Y}^r)^T \quad (4.10)$$

Com relação à capacidade de armazenamento (C) da BAM, que constitui um importante fator relacionado com o desempenho da rede, Kosko (1992) afirma que “a rede tende a exceder sua

capacidade de memória conforme o número p de padrões aproxima-se do número n ou m , ou do $\min(n, m)$ de neurônios da rede”. No entanto, esta estimativa apresenta-se muito otimista, pois a avaliação de Kosko foi feita para uma rede em que o número de unidades ativadas é aproximadamente o mesmo número de unidades desativadas em cada camada BAM. McEliece *et al.* (1987) avaliaram a capacidade de memória da rede de Hopfield e concluíram que o número de vetores bipolares escolhidos aleatoriamente que uma rede de Hopfield de n unidades pode armazenar, a fim de que a maioria das memórias originais sejam apropriadamente recuperadas, é dado pela equação $C = n / (2 \log_2 n)$. As análises de McEliece *et al.* (1987) podem ser adequadamente estendidas para a rede BAM. Sendo assim, esta última equação também é adequada para expressar o número máximo de pares de padrões que a rede BAM pode memorizar, considerando-se n a dimensão da menor camada da rede BAM.

4.1.2 Modelo Memória Associativa Bidirecional Discreto com Atraso

Faz parte dos estudos da família BAM o modelo de memória associativa bidirecional com atraso (BAM com atraso). A BAM com atraso difere-se da BAM original pelo fato de considerar atrasos de tempo em suas entradas. O atraso de tempo ($\sigma > 0$) representa a velocidade finita da transmissão de sinal nos axônios.

O estudo dos atrasos é importante porque os mesmos encontram-se presentes tanto em sistemas biológicos quanto em sistemas neurais artificiais. Na biologia, muitos atrasos são motivos para variações do crescimento do organismo, por exemplo. Em se tratando de sistemas artificiais, quando são estudados atrasos, é natural que se pergunte quais os efeitos deles nas dinâmicas e em outras propriedades da rede neural artificial. Um efeito importante dos atrasos diz respeito à estabilidade e às propriedades de convergência.

Baldi e Atiya (1994) sugerem a existência de dois lados na inclusão de atrasos em um modelo de rede. O lado ruim inclui a tendência em se introduzir algum grau de instabilidade no sistema e maior lentidão no processamento da informação. O outro lado é a parte boa da inclusão de atrasos, que seriam as evidências de possibilidade de se alcançar, através dos atrasos, efeitos dinâmicos particulares.

Liao e Yu (1998), além de concordarem com o fato de que o atraso de tempo no sinal de transmissão entre neurônios pode influenciar a estabilidade da rede ao criar fenômenos oscilatórios ou instáveis, adicionam comentários de que o atraso de tempo pode ser usado para produzir comportamentos dinâmicos tais como a recuperação de seqüências associativas e também pode ser usado para produzir um ruído na dinâmica das redes capaz de melhorar o processo de recuperação das memórias armazenadas.

Há, na literatura, a concordância de que a introdução de atrasos em redes neurais artificiais do tipo memória associativa pode: (a) ter grande influência nas propriedades de estabilidade e convergência dessas redes, (b) possibilitar a incorporação de mais realismo biológico aos modelos e (c) desempenhar papéis úteis em computações neurais (por exemplo, a construção de osciladores similares, mas com frequências diferentes).

De acordo com Baldi e Atiya (1994), a inclusão de atrasos pode gerar diferentes fenômenos quando diferentes modelos de neurônios são usados. Por exemplo, em se tratando em casos de algoritmos de aprendizagem de ponto fixo, os atrasos afetam as dinâmicas de ativação e podem mudar o ponto fixo associado a um dado estado inicial. Já dentro de arquiteturas oscilatórias, os efeitos de atrasos parecem ser principalmente quantitativos, apresentando também mudanças nos pontos de bifurcação. A inclusão da BAM com atraso na família BAM considerada neste trabalho visa verificar o comportamento dinâmico deste modelo quando lhe são adicionados componentes caóticos.

Atrasos de tempo podem ser facilmente introduzidos em modelos de memória hetero-associativa. As Equações 4.11 e 4.12 descrevem a dinâmica do modelo BAM com atrasos de tempo. Nestas equações, a função S é dada pela Equação 4.9, e o atraso de tempo $\sigma > 0$ representa a velocidade finita da transmissão de sinal nos axônios.

$$x_i(t+1) = S\left(\sum_{j=1}^n w_{ij} y_j(t-\sigma)\right) \quad i = 1, \dots, m \quad (4.11)$$

$$y_j(t+1) = S\left(\sum_{i=1}^m w_{ji} x_i(t-\sigma)\right) \quad j = 1, \dots, n \quad (4.12)$$

4.1.3 Modelo Memória Associativa Bidirecional Discreto Exponencial

O modelo de memória associativa bidirecional exponencial (eBAM), proposto por Jeng, Yeh e Chiveh em 1990, difere-se da BAM original pelo fato de apresentar uma não linearidade exponencial. O modelo de rede eBAM emprega um esquema de fluxo de informação para melhorar, de modo exponencial, a similaridade entre o padrão de entrada da rede e o padrão armazenado mais próximo dele. Tal esquema proporciona melhoria na capacidade de correção de erros de recuperação e proporciona uma capacidade de armazenamento escalada exponencialmente com o número de bits. Ela também converge para um ponto estável onde a energia é um mínimo local.

Supondo que se queira armazenar p pares associados de padrões $(\mathbf{X}^k, \mathbf{Y}^k)$, $k=1, \dots, p$, em que $\mathbf{X}^k \in \{-1, 1\}^m$ e $\mathbf{Y}^k \in \{-1, 1\}^n$, o processo de recuperação bidirecional da eBAM obedece o ciclo

dado pelas Equações 4.13 e 4.14. Nestas equações, $\langle \mathbf{P}, \mathbf{Q} \rangle$ denota o produto interno dos vetores \mathbf{P} e \mathbf{Q} , β é um número maior do que um e a função sinal, $\text{sgn}(a)$, tem valor 1, se $a \geq 0$ ou -1 , se $a < 0$.

$$\mathbf{Y} = f(\mathbf{X}) = \text{sgn} \left(\sum_{k=1}^p \mathbf{Y}^k \beta \langle \mathbf{X}^k, \mathbf{X} \rangle \right) \quad (4.13)$$

$$\mathbf{X} = g(\mathbf{Y}) = \text{sgn} \left(\sum_{k=1}^p \mathbf{X}^k \beta \langle \mathbf{Y}^k, \mathbf{Y} \rangle \right) \quad (4.14)$$

Com $\beta > 1$, a não linearidade exponencial do modelo aumenta a similaridade entre a entrada dada à rede e o par armazenado selecionado e reduzirá a interferência dos outros termos da equação. No entanto, um valor β muito alto não é uma boa escolha, pois, nesta circunstância, a rede polariza demais esse valor e perde sua capacidade de generalizar adequadamente. Assim, o termo $\langle \mathbf{Y}^k, \mathbf{Y} \rangle$ procura pelo padrão armazenado (\mathbf{Y}^k) mais próximo do padrão atual (\mathbf{Y}) para ponderá-lo. Desse modo, o padrão mais forte deve prevalecer em relação aos outros.

Finalmente, pode-se dizer que, através do uso de um esquema exponencial no fluxo de informação, os autores da eBAM mostraram, por meio de simulações computacionais, que ocorreu uma melhoria exponencial na similaridade entre o padrão de entrada e seu padrão armazenado mais próximo, uma melhoria na capacidade de armazenamento e na capacidade de correção de erros em relação a BAM original, e que a propriedade de convergência foi mantida.

4.2 Modelos de Memória Associativa Bidirecional Caótica (C-BAM)

Esta seção apresenta a diferença entre o neurônio original da rede BAM e neurônio caótico (Aihara, Takabe e Toyoda, 1990) utilizado para a construção da família C-BAM, assim como a análise do papel dos parâmetros caóticos na nova família de rede criada. Em seguida, são apresentadas as equações que descrevem a dinâmica de cada rede da família C-BAM.

4.2.1 Diferenças entre BAM e C-BAM

Para a criação da família de modelos de memória bidirecional caótica (família C-BAM), o neurônio original das redes BAM foi substituído pelo neurônio caótico de Adachi e Aihara (1997). Na família C-BAM, a fase de aprendizagem e a construção da matriz de correlação permaneceram a mesma do modelo original. Deste modo, as redes C-BAM não possuem nenhum tipo de aprendizagem adicional em relação à BAM.

As diferenças entre as redes BAM e C-BAM são enfatizadas nas Figuras 4.2 e 4.3. A Figura 4.2 mostra um esboço das entradas e saídas do neurônio original da BAM, enquanto a Figura 4.3 faz o mesmo para a C-BAM.

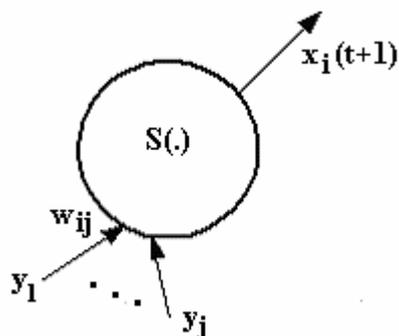


Figura 4.2. Representação do i -ésimo neurônio da rede BAM original.

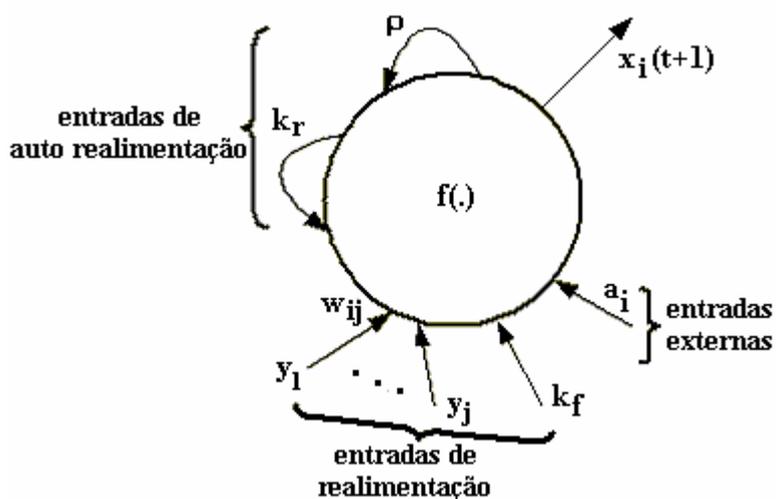


Figura 4.3. Representação do i -ésimo neurônio da rede C-BAM.

Através da Figura 4.3, pode-se notar que as entradas de auto realimentação, ponderadas pelos parâmetros k_r e ρ , e as entradas de realimentação, ponderadas por k_f , são adicionadas aos neurônios da rede BAM. As entradas externas (a_i) podem fazer parte do neurônio BAM.

Da teoria do neurônio caótico de Adachi e Aihara (1997), deve ser lembrado que, ao aumentar o valor de k_r para um valor maior que zero, passa a ser introduzida nos neurônios a capacidade de refração, que é a realimentação inibitória auto-recorrente peculiar a neurônios caóticos. A influência da refração é devida a um passado ativo do neurônio e costuma decrescer exponencialmente com o tempo. Através dessa capacidade de refração, a ativação de um neurônio retém influências nele próprio após um certo tempo. A Figura 4.4 esboça como ocorre a influência

do tempo na força de refração da rede ($k_r^{(t_{\text{atual}} - t_{\text{anterior}})}$). Deve ser lembrado que, por questão de plausibilidade biológica, o parâmetro k_r deve estar dentro do domínio $0 \leq k_r \leq 1$.

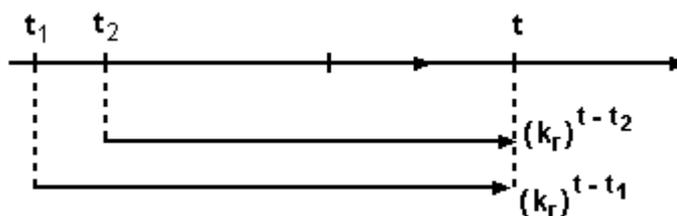


Figura 4.4. Influência do tempo na força de refração da rede C-BAM.

A Figura 4.5 ilustra como o parâmetro de decaimento para as entradas realimentadas (k_f) e o parâmetro de decaimento da capacidade de refração (k_r) atuam na dinâmica da rede caótica. Pode ser observado que a capacidade de refração provoca o distanciamento de um estado de equilíbrio (ponto fixo), de modo que a rede tende a não permanecer ao redor deste estado de equilíbrio. Quando a refração acumulada é grande, ou seja, quando o valor de k_r é suficientemente alto, a rede não mais permanece em um ponto fixo. Enquanto isto, uma vez que os pesos sinápticos da rede caótica são os mesmos da rede convencional, as entradas de realimentação vindas dos neurônios através das conexões recorrentes têm o efeito de fazer a rede convergir para um ponto fixo, tendendo a provocar um comportamento semelhante ao da rede original, não caótica.

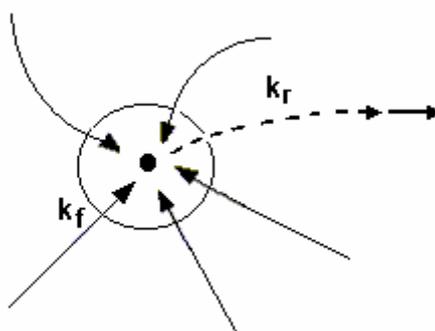


Figura 4.5. Representação da neurodinâmica caótica da rede C-BAM. O ponto central representa um ponto fixo. O parâmetro k_f representa o efeito de estabilização devido às entradas de realimentação da rede. O parâmetro k_r representa o efeito de desestabilização devido à refração ou à inibição auto-recorrente.

As diferenças entre o neurônio BAM e o neurônio C-BAM, portanto, consistem dos seguintes fatores: inclusão dos quatro parâmetros (k_r , k_f , ρ e a) que constituem o neurônio caótico e escolha da função de saída $f(\cdot)$. A dinâmica da C-BAM passa a ser determinada por estes fatores, os quais caracterizam a auto realimentação e a não linearidade capazes de gerar o comportamento caótico.

4.2.2 C-BAM

O modelo C-BAM é uma rede de duas camadas (X,Y), não linear e com auto realimentação. A arquitetura é ilustrada na Figura 4.6. Nesta ilustração, a atividade na unidade de entrada i é definida por $x_i(t)$, enquanto a atividade na unidade de saída j é definida por $y_j(t)$. As entradas externas são consideradas constantes, ou seja, $a_i = a$.

A dinâmica da rede é descrita por dois conjuntos de equações. As equações que descrevem a camada de entrada são:

$$x_i(t+1) = f(\eta_i(t+1) + \zeta_i(t+1)), \quad (4.15)$$

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^n w_{ij} y_j(t) \quad (4.16)$$

$$\zeta_i(t+1) = k_r \zeta_i(t) - \rho x_i(t) + a \quad (4.17)$$

A camada de saída é descrita por:

$$y_j(t+1) = f(\eta_j(t+1) + \zeta_j(t+1)), \quad (4.18)$$

$$\eta_j(t+1) = k_f \eta_j(t) + \sum_{i=1}^m w_{ji} x_i(t) \quad (4.19)$$

$$\zeta_j(t+1) = k_r \zeta_j(t) - \rho y_j(t) + a \quad (4.20)$$

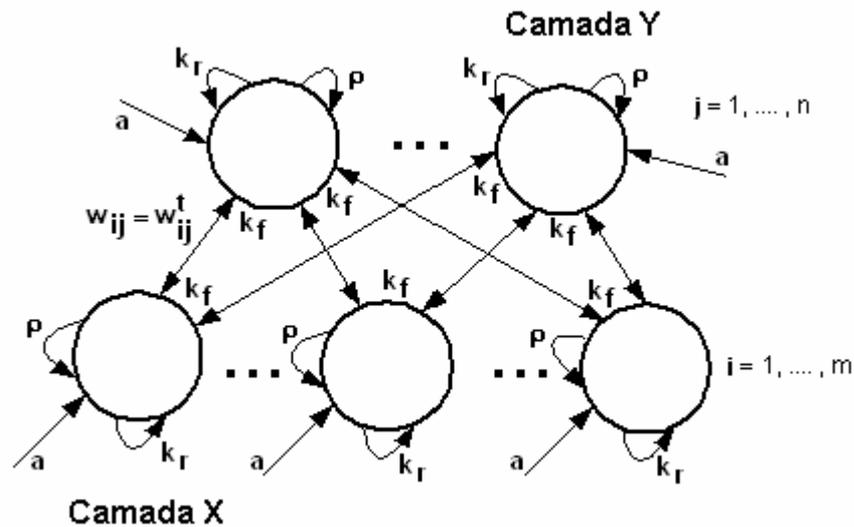


Figura 4.6. Arquitetura C-BAM. A atividade na unidade de entrada i é determinada por $x_i(t)$, e a atividade na unidade de saída j é determinada por $y_j(t)$.

Na figura que ilustra a arquitetura C-BAM (Figura 4.6), pode ser notado que as entradas de auto realimentação, ponderadas por k_r e ρ , e as entradas de realimentação, ponderadas por k_f , foram adicionadas na C-BAM. A matriz de conexão dos pesos é descrita na Equação 4.21. A função $f(\cdot)$ escolhida foi a função logística, descrita na Equação 4.22.

$$\mathbf{W} = \sum_{k=1}^p \mathbf{X}^k (\mathbf{Y}^k)^T \quad (4.21)$$

em que $(\mathbf{X}^k, \mathbf{Y}^k)$ são os vetores dos pares de padrões armazenados, para $k=1, \dots, p$.

$$f(v) = \frac{1}{\{1 + \exp(-v/\varepsilon)\}} \quad (4.22)$$

em que $\varepsilon = 0.02$.

Em aplicações práticas com a família C-BAM, notou-se que os valores de $f(\cdot)$ são sempre aproximadamente iguais a 0 ou 1, para as redes que lidam com padrões binários (C-BAM e C-BAM com atraso) e são aproximadamente iguais a -1 e 1 no caso da C-eBAM, que lida com padrões bipolares. Sendo assim, como os pares de padrões armazenados na família C-BAM são vetores binários ou bipolares, em todos os experimentos com a família C-BAM considerou-se a aproximação apresentada na Equação 4.23. A escolha apropriada da curva da função exponencial $f(v)$, através do parâmetro ε , possibilita o ajuste feito nas saídas das redes C-BAM.

$$h(f(v)) = \begin{cases} 1, & \text{se } f(v) > 0.98 \\ 0 \text{ ou } -1, & \text{se } f(v) < 0.02 \end{cases} \quad (4.23)$$

4.2.3 C-BAM com atraso

O modelo C-BAM com atraso é uma rede de duas camadas (X, Y) , não linear, com auto realimentação e que tem um atraso de tempo na transmissão de sinal entre os neurônios. A dinâmica desta rede é descrita pelo conjunto de equações a seguir:

$$x_i(t+1) = f(\eta_i(t+1) + \zeta_i(t+1)), \quad (4.24)$$

$$\eta_i(t+1) = k_f \eta_i(t-\sigma) + \sum_{j=1}^n w_{ij} y_j(t-\sigma) \quad (4.25)$$

$$\zeta_i(t+1) = k_r \zeta_i(t-\sigma) - \rho x_i(t-\sigma) + a \quad (4.26)$$

$$y_j(t+1) = f(\eta_j(t+1) + \zeta_j(t+1)), \quad (4.27)$$

$$\eta_j(t+1) = k_f \eta_j(t-\sigma) + \sum_{i=1}^m w_{ji} x_i(t-\sigma) \quad (4.28)$$

$$\zeta_j(t+1) = k_r \zeta_j(t-\sigma) - \rho y_j(t-\sigma) + a \quad (4.29)$$

4.2.4 C-eBAM

O modelo C-eBAM é uma rede de duas camadas (X, Y), não linear, com auto realimentação e que usa uma estratégia exponencial visando melhorias na similaridade entre a entrada fornecida à rede e o par armazenado selecionado como resposta (saída da rede). A dinâmica desta rede é descrita por dois conjuntos de equações, uma descrevendo a camada de entrada X e outra descrevendo a camada de saída Y :

$$x_i(t+1) = f(\eta_i(t+1) + \zeta_i(t+1)), \quad (4.30)$$

$$\eta_i(t+1) = k_f \eta_i(t) + \left\{ \sum_{k=1}^p X_i^k \beta \langle \mathbf{Y}^k, \mathbf{Y}(t) \rangle \right\} \quad (4.31)$$

$$\zeta_i(t+1) = k_r \zeta_i(t) - \rho x_i(t) + a \quad (4.32)$$

$$y_j(t+1) = f(\eta_j(t+1) + \zeta_j(t+1)) \quad (4.33)$$

$$\eta_j(t+1) = k_f \eta_j(t) + \left\{ \sum_{k=1}^p Y_j^k \beta \langle \mathbf{X}^k, \mathbf{X}(t) \rangle \right\} \quad (4.34)$$

$$\zeta_j(t+1) = k_r \zeta_j(t) - \rho y_j(t) + a \quad (4.35)$$

4.3 Resumo

Os estudos de Aihara, Takabe e Toyoda (1990) propuseram um modelo neurônio caótico simples e um modelo de rede neural artificial composto por esses neurônios caóticos apresentando respostas não periódicas. Os trabalhos destes autores (Aihara, Takabe e Toyoda, 1990) incluem também a aplicação do modelo neurônio caótico em redes de memória auto-associativa. Os estudos embasaram a introdução de neurônios caóticos em uma família de modelos de memórias hetero-associativas bidirecionais, visando analisar o comportamento destes modelos.

Com o objetivo de descrever a criação da família de modelos de memória hetero-associativa bidirecional caótica (família C-BAM), construída através da inclusão do neurônio caótico nas redes BAM originais, este capítulo iniciou-se com a apresentação das características dos modelos

originais de redes de memória hetero-associativas bidirecionais selecionadas para o presente trabalho. Em seguida o capítulo destacou as diferenças entre a família BAM original e a família C-BAM construída. Finalmente, foram mostradas as equações que descrevem a dinâmica dos novos modelos.

Após a descrição dos fenômenos caóticos, o entendimento de como o caos é introduzido em redes neurais artificiais, a criação de uma família de redes BAM capaz de gerar caos e a descrição do fenômeno psicológico envolvendo pular passos na resolução de uma tarefa, as dinâmicas emergentes na família C-BAM criada e suas aplicações serão apresentadas no próximo capítulo.

Capítulo 5

Experimentos com a Família de Modelos de Memória Associativa Bidirecional Caótica

Criada a família de modelos de Memória Associativa Bidirecional Caótica (família C-BAM), realizaram-se, então, diversos experimentos com o propósito de verificar o comportamento dinâmico dessa família de modelos. O objetivo deste capítulo é, portanto, apresentar e discutir os experimentos e as análises realizadas com a família C-BAM.

A princípio procurou-se investigar a importância de cada parâmetro dos neurônios da família C-BAM, visando a estimar o efeito destes sobre os comportamentos dinâmicos dos modelos C-BAM. A Seção 5.1.1 apresenta como foi feita esta avaliação paramétrica, utilizando-se do *planejamento fatorial de experimentos* (Montgomery, 1997). Por meio da variação dos quatro parâmetros (k_r , k_f , ρ e a) da família C-BAM e de acordo com os resultados obtidos nos experimentos da Seção 5.1.1, foi possível verificar a presença de diferentes tipos de comportamentos dinâmicos emergentes na família C-BAM. A Seção 5.1.2 apresenta quais são estes comportamentos dinâmicos emergentes e os experimentos realizados para sua obtenção. Para a comprovação dos comportamentos caóticos dos modelos C-BAM sugeridos pelos resultados dos experimentos da Seção 5.1.2 é mostrado o cálculo dos expoentes de Lyapunov para as redes C-BAM na Seção 5.2.

Nos experimentos realizados com a família C-BAM foi também observado o processo de recuperação das redes. A Seção 5.3 mostra como o fenômeno caótico permite que a família C-BAM acesse padrões memorizados que os modelos originais destas redes não podiam recuperar. As características de recuperação das redes C-BAM sugeriram que a nova maneira de estas acessarem os padrões armazenados poderia ser usada para modelar a habilidade de se saltar passos na solução

de uma tarefa. A Seção 5.4 descreve como as C-BAM são capazes de modelar o fenômeno de saltar passos. A Seção 5.5 finaliza com um resumo dos experimentos e resultados alcançados neste capítulo.

5.1 Comportamentos Dinâmicos da Família C-BAM

Esta seção apresenta (a) os resultados obtidos com a técnica de planejamento de experimentos na avaliação dos parâmetros de controle da família C-BAM (Seção 5.1.1) e (b) os comportamentos dinâmicos emergentes, obtidos com variação paramétrica, para esta família de redes (Seção 5.1.2).

5.1.1 Planejamento de Experimentos

Foi visto no Capítulo 4 que as diferenças entre os neurônios da BAM e os neurônios da C-BAM consistem basicamente da inclusão dos quatro parâmetros de realimentação e de entrada de estímulos (k_r , k_f , ρ e a) que constituem o neurônio caótico e a escolha da função de saída. Sendo assim, a presente seção tem como objetivo o estudo formal da importância de cada parâmetro dos neurônios da família C-BAM, visando a estimar o efeito deles sobre a dinâmica da rede. O *planejamento fatorial de experimentos* (Montgomery, 1997) foi a técnica de planejamento de experimentos escolhida para alcançar esta meta, pois permite que vários fatores de interesse em um experimento (no caso da C-BAM são quatro fatores) sejam avaliados e porque os experimentos fatoriais avaliam com metodologia as interações entre os fatores (Montgomery e Runger, 2003).

Inicialmente, experimentos empíricos foram realizados com a finalidade de observar os comportamentos apresentados pela família C-BAM. Tais experimentos sugeriram a existência de respostas periódicas, bifurcações, caos e crise. Os parâmetros da rede C-BAM foram, então, avaliados dentro de seus intervalos de definição de modo que o planejamento fatorial contribuísse para formalmente determinar a probabilidade de se alcançar um comportamento específico dentro de algumas faixas de valores de parâmetros.

Dois tipos de experimentos usando planejamento fatorial foram realizados. No primeiro conjunto de experimentos, os valores de parâmetros da C-BAM foram selecionados aleatoriamente, a fim de se obter a relevância de cada parâmetro sem a preocupação de alcançar um comportamento específico. No segundo conjunto de experimentos foram avaliados os valores de parâmetros que geravam comportamentos específicos (por exemplo, caos), a fim de estimar-se a relevância de cada parâmetro em um comportamento particular.

Sendo assim, o primeiro tipo de análise consistiu em planejamentos fatoriais para verificar a influência dos parâmetros no *comportamento geral* da C-BAM. Após alguns experimentos, a média

dos resultados obtidos com as tabelas ANOVA (Montgomery e Runger, 2003) sugeriu que k_r e suas interações com os outros parâmetros constituem a maior fonte de influência sobre o comportamento do sistema (freqüentemente acima de 50%), sendo seguido em grau de importância pelo parâmetro ρ , que costuma afetar o comportamento do sistema em média de 20% do total. Os parâmetros k_f e a têm influências similares, em torno de 10%. Finalmente, as interações entre os parâmetros, sem considerar k_r , alcançaram as menores porcentagens de influência no sistema.

Na segunda parte da análise, experimentos resultantes em *comportamentos periódicos*, quando comparados com os experimentos do primeiro tipo de análise, sugeriram um decréscimo na relevância de parâmetros individuais ao custo de um acréscimo nas interações entre eles. Nestes casos, em média, k_r explica em torno de 18% da variação total dos comportamentos periódicos, sendo seguido pela influência das interações entre outros parâmetros (em torno de 10% para cada interação). A influência de ρ e k_f ficaram em torno de 7%.

A relevância de k_r em *comportamentos caóticos* (em torno de 40%) é mais alta do que em comportamentos periódicos. Além disto, ρ usualmente afeta este tipo de comportamento em torno de 10%. A influência de k_f e a variaram muito em cada planejamento fatorial, ficando entre 0.1% a 9% nos experimentos realizados.

Dado que o planejamento fatorial de experimentos é uma técnica facilmente encontrada na literatura, a teoria e a aplicação dessa técnica de planejamento de experimentos na rede C-BAM estão apresentados no Apêndice A. Maiores detalhes podem ser consultados em Montgomery (1997), Montgomery e Runger (2003) e Jain (1991).

Um experimento fatorial com quatro fatores será utilizado para ilustrar os resultados acima. Neste experimento, cada fator teve dois níveis considerados: $k_r = 0.8$ e 0.9 (fator A), $k_f = 0.2$ e 0.4 (fator B), $\rho = 7$ e 8 (fator C) e $a = 2$ e 3 (fator D). O comportamento caótico foi encontrado para as combinações paramétricas desse experimento fatorial. O experimento foi realizado com duas réplicas. A análise de variância é mostrada na Tabela 5.1.

De acordo com Montgomery (1997), o número de níveis de cada fator foi o que determinou o número de graus de liberdade na tabela ANOVA (Tabela 5.1). Além disso, o baixo valor do erro (erro = 0.001) significa que as réplicas para cada fonte de variação são praticamente idênticas.

Considerando o nível de significância $\alpha=0.05$ na função de distribuição F, foi obtido (Tabela 5.1) que $F_{\text{fator}} > F_{0.05,1,16} = 4.49$ (Spiegel, 1978). Conseqüentemente, a hipótese de os efeitos de qualquer fator e suas interações serem nulas foi rejeitada com um nível de significância 0.05. Isto indica que todos os parâmetros e suas interações têm um efeito estatisticamente significativo no comportamento do sistema.

Tabela 5.1. Resultados da tabela ANOVA (Montgomery e Runger, 2003) para o experimento fatorial com quatro fatores.

Fonte de Variação	Soma Quadrática	Graus de Liberdade	Média Quadrática	F ₀
k_r	328,974	1	328,974	4,4E+08
k_f	0,907	1	0,907	1,2E+06
ρ	54,429	1	54,429	7,3E+07
a	6,901	1	6,901	9,2E+06
$k_r * k_f$	6,331	1	6,331	8,4E+06
$k_r * \rho$	25,601	1	25,601	3,4E+07
$k_r * a$	28,144	1	28,144	3,8E+07
$k_f * \rho$	0,571	1	0,571	7,6E+05
$k_f * a$	12,655	1	12,655	1,7E+07
$\rho * a$	7,106	1	7,106	9,5E+06
$k_r * k_f * \rho$	164,512	1	164,512	2,24E+08
$k_r * k_f * a$	42,550	1	42,550	5,7E+07
$k_r * \rho * a$	0,091	1	0,091	1,2E+05
$k_f * \rho * a$	0,577	1	0,577	7,7E+05
$k_r * k_f * \rho * a$	8,617	1	8,617	1,1E+07
Erro	0,001	16	0	
Total	687,967	31		

Outra maneira de analisar os resultados na Tabela 5.1 é através do procedimento que divide a variância total da resposta em componentes significativos (Jain, 1991). A porcentagem de variação explicada por cada fator sustenta a decisão de se o fator considerado tem ou não um impacto relevante na resposta. Sendo assim, os resultados da tabela ANOVA (Tabela 5.1) também indicam que o fator k_r explica aproximadamente 47.82% (328,974/687,967) da variação total do sistema, de modo que a escolha de k_r tem um efeito significativamente alto no comportamento dinâmico da rede C-BAM. Influências altas no comportamento da C-BAM também foram encontradas para a interação $k_r * k_f * \rho$ ($\approx 23.91\%$). Influências médias foram encontradas para ρ ($\approx 7.91\%$) e para as seguintes interações: $k_r * k_f * a$ ($\approx 6.18\%$), $k_r * a$ ($\approx 4.09\%$) e $k_r * \rho$ ($\approx 3.72\%$). Baixas influências foram encontradas para a ($\approx 1\%$), k_f ($\approx 0.13\%$) e para as outras interações.

5.1.2 Comportamentos Emergentes na Família C-BAM

Com o objetivo de ilustrar o comportamento da família C-BAM para os diferentes parâmetros (k_r , k_f , ρ e a) variando de acordo com os resultados da Seção 5.1.1, alguns experimentos foram selecionados para descrição. Estes experimentos puderam também exemplificar os comportamentos dinâmicos emergentes da família C-BAM. As variações paramétricas da C-BAM incluíram a seleção de valores capazes de: (i) gerar comportamento semelhante à rede BAM original; (ii) produzir respostas periódicas; (iii) produzir bifurcações; (iv) gerar caos; (v) resultar em crise; (vi) mostrar diversidade de recuperação de padrões armazenados e, finalmente, (vii) exemplificar o acesso a memórias anteriormente inacessíveis. Todos estes comportamentos foram observados nos três membros da família C-BAM.

Para os experimentos com a família C-BAM foram criadas redes com quinze neurônios caóticos na camada X e dez neurônios caóticos na camada Y. A Figura 5.1 mostra os quatro pares de padrões armazenados nos pesos sinápticos com a matriz de correlação da Equação 4.21 os quais são usados nos experimentos a seguir. Os padrões de entrada de quinze neurônios caóticos são apresentados na forma de um vetor de quinze bits, enquanto os padrões de saída de dez neurônios caóticos são apresentados na forma de um vetor de dez bits. Os quadrados pretos simbolizam os bits de valor 1. Os quadrados brancos simbolizam os bits de valor 0 no caso dos padrões binários utilizados na C-BAM e na C-BAM com atraso, e bits de valor -1 no caso dos padrões bipolares utilizados na C-eBAM.

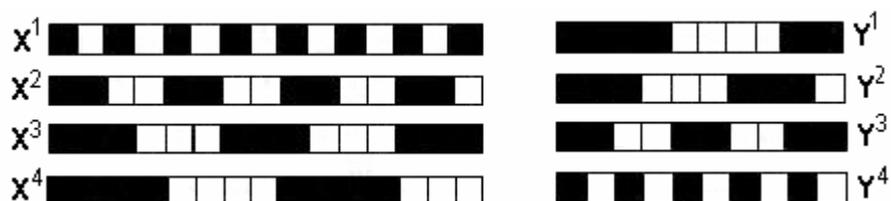


Figura 5.1. Pares de padrões armazenados (X^k, Y^k), $k=1, \dots, 4$, usados nos experimentos com a família C-BAM.

Em um primeiro experimento, todos os parâmetros da família C-BAM receberam valor zero. O comportamento da família C-BAM tornou-se idêntico ao da família BAM original, no qual a rede estabiliza-se em um estado de equilíbrio estável. Este comportamento sugere que a modelagem dos neurônios caóticos em redes de memória hetero-associativa é uma extensão natural da rede de memória hetero-associativa convencional. A Tabela 5.2 mostra que a família C-BAM converge para pontos fixos a partir de padrões iniciais específicos.

Tabela 5.2. Relação entre entradas da família C-BAM (\mathbf{X}) e suas saídas (\mathbf{Y}) para $k_r = k_f = \rho = a = 0$.

<i>Padrão Inicial</i>	\mathbf{X}^1	\mathbf{X}^2	\mathbf{X}^3	\mathbf{X}^4
<i>Saída</i>				
C-BAM	\mathbf{Y}^1	\mathbf{Y}^2	\mathbf{Y}^3	\mathbf{Y}^4
C-BAM with delay ($\sigma = 5$)	\mathbf{Y}^1	\mathbf{Y}^2	\mathbf{Y}^3	\mathbf{Y}^4
C-eBAM ($\beta = 1.3$)	\mathbf{Y}^1	\mathbf{Y}^2	\mathbf{Y}^3	\mathbf{Y}^4

O segundo tipo de experimento ilustra a influência dos parâmetros que constituem o neurônio caótico no comportamento dinâmico da família C-BAM. A escolha dos parâmetros para gerar o comportamento dinâmico desejado obedeceu às probabilidades estatísticas informadas na Seção 5.1.1. Logo, ao se desejar a ocorrência de comportamentos caóticos, uma atenção especial foi dada na escolha dos parâmetros k_r (optando-se por valores relativamente altos) e ρ .

Nos experimentos com a rede C-BAM com atraso, foram selecionados atrasos de valores $\sigma = 3$ e $\sigma = 5$ pelo fato de serem valores intermediários, que permitem a presença de diferentes efeitos dinâmicos na rede C-BAM sem, no entanto, introduzir um alto grau de instabilidade nesta rede. O valor de $\beta = 1.3$ foi selecionado para a rede C-eBAM pelo fato de ele permitir um aumento na similaridade entre a entrada dada à rede e o par armazenado selecionado em seu processo de recuperação, sem provocar perdas na capacidade de generalização da C-eBAM.

As Figuras 5.2, 5.3 e 5.4 ilustram as distâncias Hamming entre os padrões de saída e os padrões armazenados. Estas distâncias assumem valores inteiros de zero (quando ocorre a recuperação exata do padrão armazenado) a dez (quando ocorre a recuperação do inverso do padrão armazenado). Tais experimentos ilustram que, após a família C-BAM ter sido treinada com os padrões desejados (construção das matrizes de correlação), sem que haja qualquer tipo de treinamento complementar, os parâmetros caóticos (ou seja, os parâmetros inseridos no neurônio caótico), por si só, podem gerar uma grande diversidade de padrões recuperados, envolvendo complexas transições entre os padrões memorizados. Estes resultados confirmam que a família C-BAM é capaz de acessar todos os padrões armazenados, sem convergir para nenhum padrão em particular. Os resultados encontrados estimularam a investigação sobre novas dinâmicas emergentes a partir da variação dos parâmetros caóticos da família C-BAM.

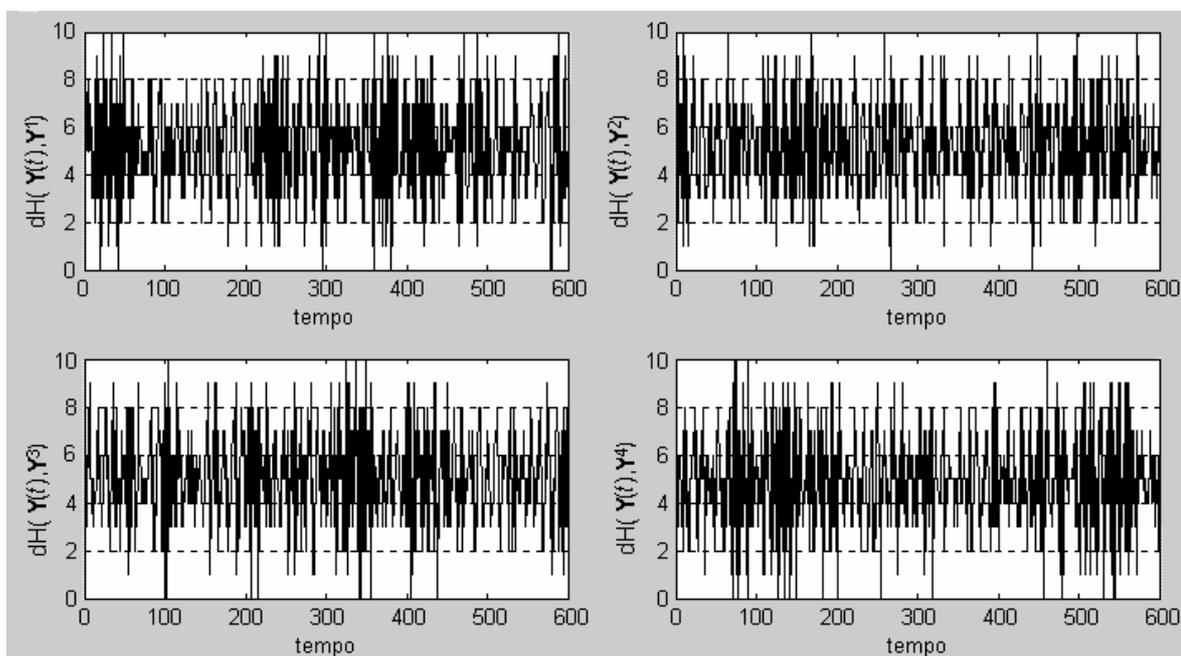


Figura 5.2. Distância Hamming (dH) para C-BAM, em que $k_r=0.85$, $k_f=0.2$, $\rho=8.0$ e $a=2.0$. Para o cálculo da distância Hamming entre o padrão de saída e o padrão armazenado ($dH(\mathbf{Y}(t), \mathbf{Y}^k)$) foi considerada a diferença entre os bits da sequência numérica da saída da rede no tempo t ($\mathbf{Y}(t)$) e o padrão \mathbf{Y}^k armazenado ($k=1, \dots, 4$). A cada bit diferente entre $\mathbf{Y}(t)$ e \mathbf{Y}^k (neurônio a neurônio da rede), foi adicionado o valor um na dH entre eles.

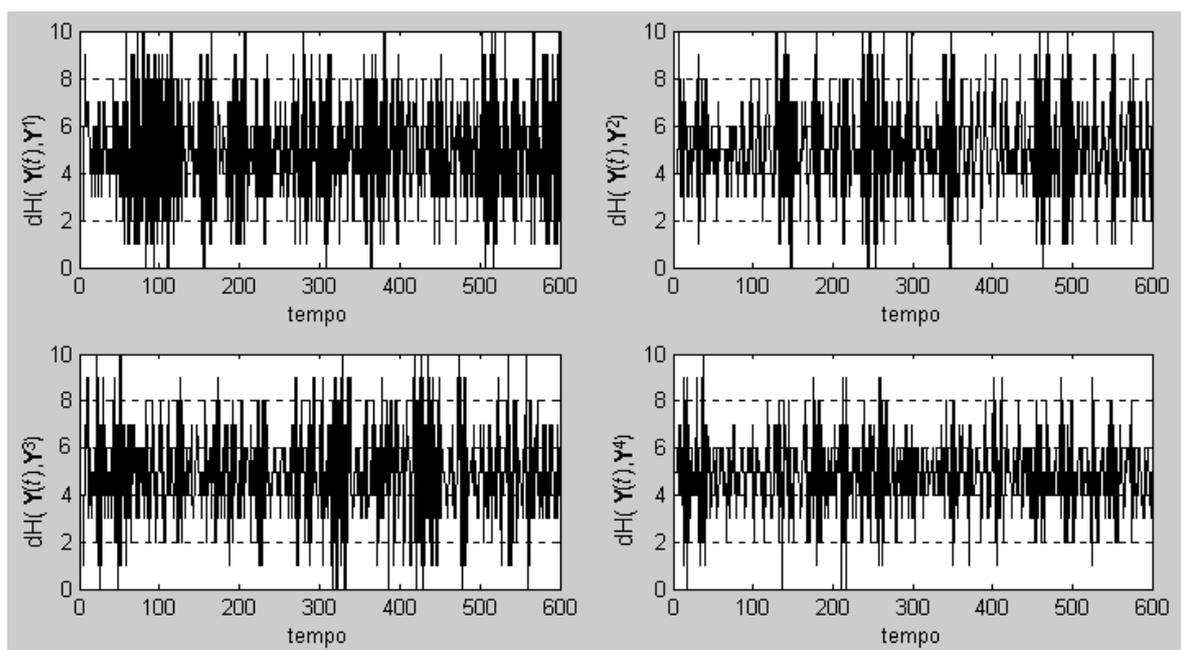


Figura 5.3. Distância Hamming para C-BAM com atraso, em que $k_r=0.8$, $k_f=0.15$, $\rho=8.0$, $a=2.0$ e $\sigma=5$.

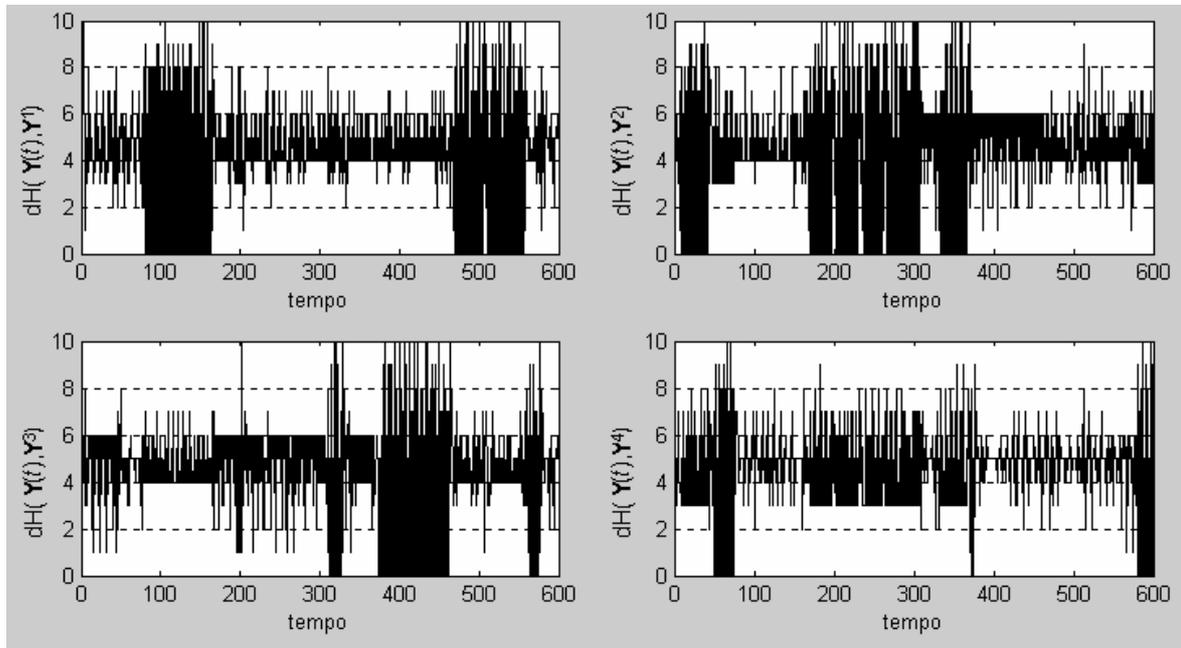


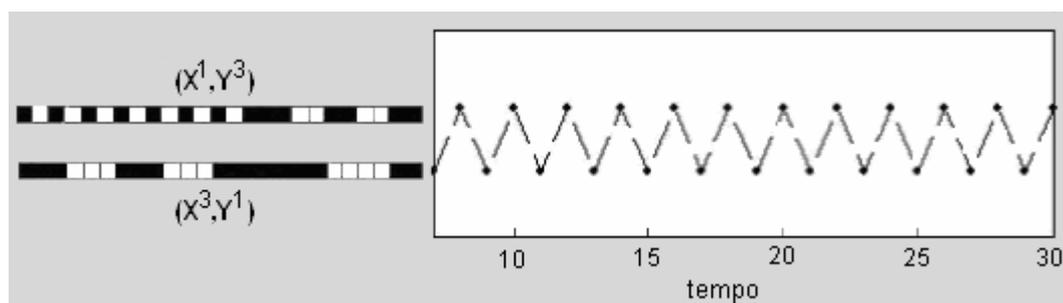
Figura 5.4. Distância Hamming para C-eBAM, em que $k_r=0.9$, $k_f=0.2$, $\rho=10.0$, $a=2.0$ e $\beta=1.3$.

A Tabela 5.3 utiliza a rede C-BAM para ilustrar a influência de k_r na dinâmica da rede. Os parâmetros $k_f=0.15$, $\rho=8.0$ e $a=2.0$ permaneceram inalterados, enquanto k_r foi variado no intervalo de 0 a 1.0. Pode ser verificado que cada novo domínio de k_r começou a apresentar período duas vezes maior que o domínio anterior (por exemplo, 2, 4, 8, 16) até chegar na ocorrência do caos. O resultado deste comportamento é uma **bifurcação** conhecida como *duplicação de período*, e, conforme visto no Capítulo 2, tal bifurcação corresponde a um processo no qual o aumento no valor do parâmetro de controle do sistema gera uma seqüência de atratores periódicos de período dobrado que, no final do processo, conduz ao comportamento caótico.

Dois casos particulares de comportamento descritos na Tabela 5.3 são mostrados nas Figuras 5.5 e 5.6. A Figura 5.5 mostra a ocorrência de **atrator periódico** de período 2 dentro da bifurcação do tipo duplicação de período. Este comportamento consiste na recuperação de dois pares que se mantêm repetindo na mesma ordem. Pode ser notado na Figura 5.5 a ocorrência de repetição periódica de pares em cada dois passos de tempo ($(\mathbf{X}, \mathbf{Y})_t = (\mathbf{X}, \mathbf{Y})_{t+2}$). A resposta do sistema forma um ciclo para frente e para trás entre os pares de padrões $(\mathbf{X}^3, \mathbf{Y}^1)$ e $(\mathbf{X}^1, \mathbf{Y}^3)$.

Tabela 5.3. Comportamento da C-BAM para k_r variando de 0 a 1 e $k_f = 0.15$, $\rho = 8.0$, $a = 2.0$.

<i>Comportamento Geral</i> \Rightarrow bifurcação do tipo duplicação de período \rightarrow caos		
k_r	Início do Comportamento	Comportamento
0	$t = 4$	período 2
0.1	$t = 3$	período 2
0.2	$t = 7$	período 2
0.3	$t = 7$	período 2
0.4	$t = 16$	período 4
0.5	$t = 7$	período 4
0.55	$t = 23$	período 8
0.6	$t = 31$	período 16
0.7	-	caos
0.8	-	caos
0.9	-	caos
1.0	-	caos

**Figura 5.5.** Sequência de pares recuperados na C-BAM: (X^3, Y^1) e (X^1, Y^3) . Período 2 ocorre do instante $t = 7$ em diante (até $t = 30$ nesta figura). Valores de parâmetros utilizados: $k_r = 0.2$, $k_f = 0.15$, $\rho = 8.0$, e $a = 2.0$.

A Figura 5.6 apresenta um tipo de dinâmica diferente e mais complexa do que a dinâmica periódica anteriormente apresentada na Figura 5.5. Na atual situação, o caos determinístico é o comportamento resultante. Tal dinâmica pode ser entendida como uma evolução imprevisível a longo prazo, a qual ocorre em sistemas dinâmicos determinísticos e não lineares por causa da sensibilidade às condições iniciais. A recuperação irregular de pares (armazenados ou espúrios) é mostrada, na Figura 5.6, dentro do intervalo $t = 192$ a $t = 211$.

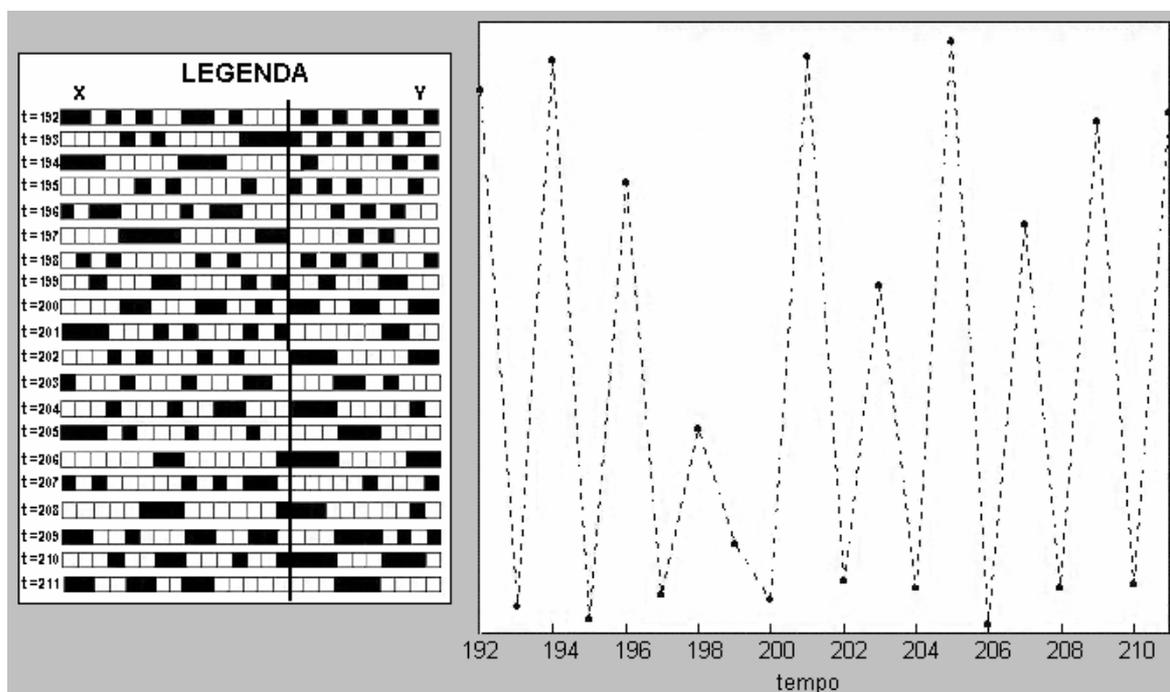


Figura 5.6. Seqüência de pares recuperados na C-BAM. À esquerda, na legenda, é mostrada a seqüência binária de pares de padrões recuperados pela rede (retângulos pretos representam valor 1, enquanto os brancos representam valor 0). À direita da figura, cada par binário recuperado é transformado em seu número decimal equivalente, representado pelos pontos. Pode ser notada a ocorrência de comportamento não periódico dentro do intervalo $t=192$ a $t=211$. Valores de parâmetros utilizados: $k_r = 0.8$, $k_f=0.15$, $\rho=8.0$, e $a=2.0$.

A mudança qualitativa no comportamento do sistema, produzindo bifurcação do tipo duplicação de período, também foi observada nas outras duas redes da família C-BAM: a C-BAM com atraso e a C-eBAM. Por exemplo, fixando $k_f=0.15$, $\rho=8.0$, $a=2.0$ e $\sigma=3$, a rede C-BAM com atraso teve suas respostas periódicas duplicadas até alcançar comportamento caótico, conforme o valor de k_r era aumentado dentro da faixa 0.3 a 0.7. A bifurcação de duplicação de período foi encontrado para a C-eBAM para o conjunto de valores de parâmetros $k_f=0.15$, $\rho=4.0$, $a=2.0$ e $\beta=1.3$ fixo e k_r variando dentro da faixa 0.5 a 0.9.

O comportamento conhecido como **crise**, no qual mudanças qualitativas na dinâmica caótica ocorrem conforme um parâmetro do sistema é variado, de modo que o comportamento caótico mantém-se intercalando com o comportamento periódico, foi também observado na família C-BAM. A Tabela 5.4 mostra um exemplo desta ocorrência na C-BAM com atraso para o conjunto de parâmetros: $k_r=0.65$, $\rho=9.0$, $a=2.0$, $\sigma=5$, e k_f variando de 0 a 1.0. O experimento ilustrado na Tabela 5.4 sugere que, durante a variação do parâmetro k_f , o atrator caótico é destruído para dar lugar à órbita periódica, o que caracteriza a ocorrência de crise de fronteira (*boundary crisis*).

Embora os experimentos realizados tenham sugerido a ocorrência de crise de fronteira, há ainda a possibilidade da ocorrência de um outro fenômeno caótico, chamado *intermitência*, que se caracteriza por estouros de comportamento caótico intercalados com intervalos irregulares de comportamentos aparentemente periódicos. É importante ressaltar que, para a confirmação de crise ou de intermitência, seria necessária uma visualização de como os atratores evoluem no espaço. Pelo fato de a família C-BAM possuir vários neurônios em cada uma de suas camadas, ainda não foi possível encontrar uma forma precisa de visualizar os fenômenos emergentes nas redes C-BAM. O aperfeiçoamento da visualização de atratores é uma das propostas para futuras pesquisas envolvendo a família C-BAM (Capítulo 7).

Tabela 5.4. Crise na rede C-BAM com atraso, para k_f variando e $k_r = 0.65$, $\rho=9.0$, $a=2.0$ e $\sigma=5$ fixos.

<i>Comportamento Geral</i> \Rightarrow crise		
k_f	Início do Comportamento	Comportamento
0	t = 55	período 16
0.1	t = 22	período 8
0.2	-	caos
0.3	-	caos
0.4	-	caos
0.5	-	caos
0.6	t = 58	período 33
0.7	t = 19	período 20
0.8	-	caos
0.9	t = 23	período 26
1.0	t = 8	período 2

Ainda na Tabela 5.4, é interessante destacar a presença de períodos ímpares (tal como o período 33) e de outros períodos que não são potências de 2 (tais como os períodos 20 e 26). Estes períodos diferem daqueles encontrados na Tabela 5.3, em que há bifurcação de duplicação de período. Tais ocorrências podem significar comportamentos caóticos diferentes dos tratados neste trabalho, e poderão ser analisados futuramente.

A intercalação de ordem e caos (vista na Tabela 5.4), foi também observada para as redes C-BAM e C-eBAM. Um exemplo de conjunto paramétrico em que este comportamento ocorre para a C-BAM é $k_r=0.7$, $\rho=8.0$, $a=2.0$, enquanto k_f varia entre 0 e 1.0. Na C-eBAM: $k_r=0.7$, $\rho=10.0$, $a=2.0$, $\beta=1.3$ fixos e k_f variando entre 0 e 1.0.

Foi também verificada a recuperação de pares de padrões exatamente do modo como eles foram treinados. Na Figura 5.7, através das distâncias Hamming, é mostrada a recuperação dos pares associados treinados na C-BAM com atraso: $(\mathbf{X}^1, \mathbf{Y}^1)$, $(\mathbf{X}^2, \mathbf{Y}^2)$, $(\mathbf{X}^3, \mathbf{Y}^3)$ e $(\mathbf{X}^4, \mathbf{Y}^4)$, para a entrada inicial $\mathbf{X}(t=0)=\mathbf{X}^1$.

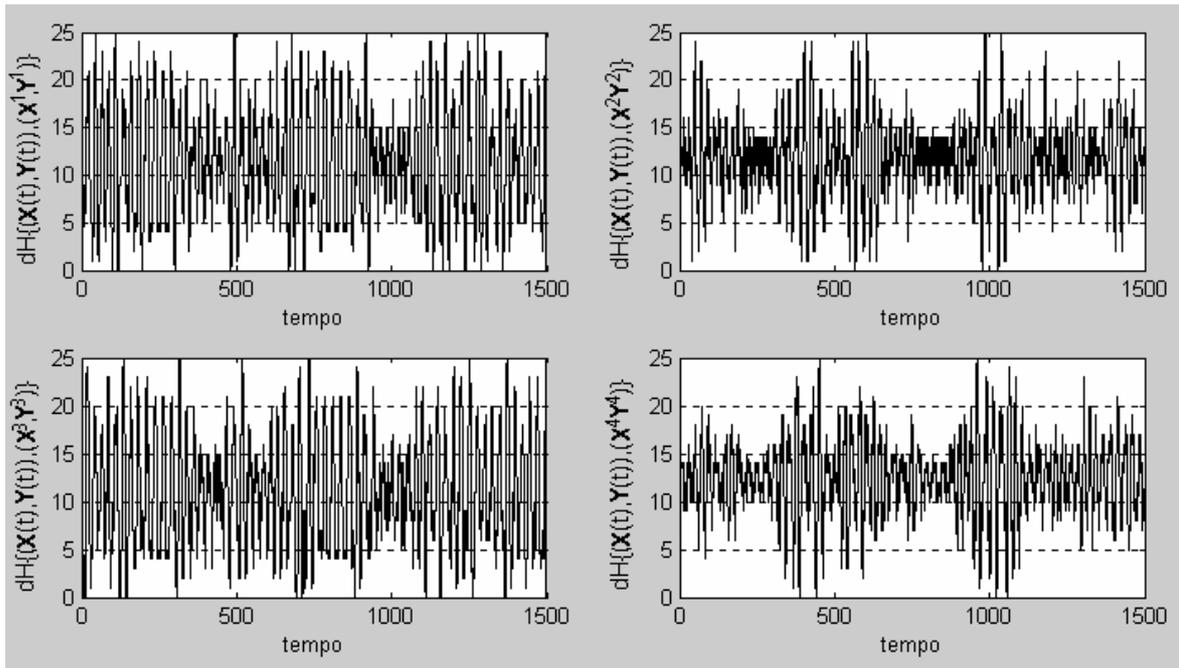


Figura 5.7. Recuperação dos pares associados treinados na C-BAM com atraso para os parâmetros $k_r = 0.92$, $k_f = 0.46$, $\rho = 4.0$, $a=2.0$ e $\sigma=3$. Neste caso, para o cálculo da distância Hamming, $dH\{(\mathbf{X}(t), \mathbf{Y}(t)), (\mathbf{X}^k, \mathbf{Y}^k)\}$ foi considerada a diferença entre os bits da seqüência numérica binária do par associado encontrado no tempo t $(\mathbf{X}(t), \mathbf{Y}(t))$ e do par associado armazenado $(\mathbf{X}^k, \mathbf{Y}^k)$, $k=1, \dots, 4$. A cada bit diferente entre $(\mathbf{X}(t), \mathbf{Y}(t))$ e $(\mathbf{X}^k, \mathbf{Y}^k)$, foi adicionado o valor um na distância Hamming entre eles.

Outro comportamento interessante encontrado na investigação da dinâmica família C-BAM é a capacidade de recuperação de memórias não acessíveis. Este comportamento será apresentado em uma seção exclusiva, a Seção 5.3.

A Figura 5.8 ilustra quando o comportamento caótico emerge em cada rede da família C-BAM de acordo com o aumento do parâmetro k_r . Para esta ilustração, os outros parâmetros foram mantidos fixos em $k_f=0.3$, $\rho=9.0$, $a=2.0$, $\sigma=5$ (para C-BAM com atraso) e $\beta=1.3$ (para C-eBAM). As barras pretas indicam comportamento periódico, enquanto as barras horizontais indicam comportamento caótico. As mudanças de dinâmica nas redes da família C-BAM podem ser consideradas como qualitativamente idênticas, devido à faixa aproximada de ocorrência de caos entre elas.

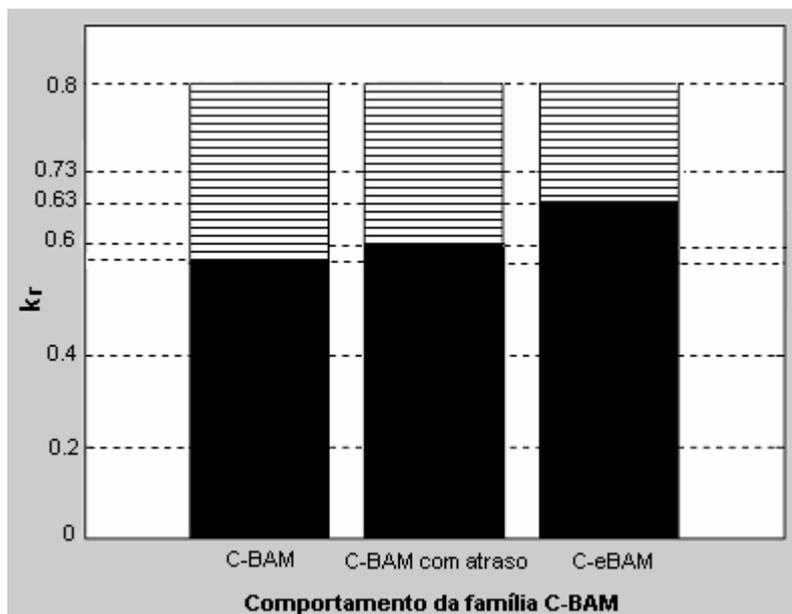


Figura 5.8. Limites entre comportamentos periódicos e caóticos de cada rede da família C-BAM, de acordo com o aumento do parâmetro k_r . Os outros parâmetros foram mantidos fixos em $k_f=0.3$, $\rho=9.0$, $a=2.0$, $\sigma=5$ (para C-BAM com atraso) e $\beta=1.3$ (para C-eBAM). As barras pretas indicam comportamento periódico, enquanto as barras horizontais indicam comportamento caótico.

5.2 Expoentes de Lyapunov

Como foi visto na seção anterior, vários fenômenos interessantes relacionados com a introdução do caos nas redes da família C-BAM foram sugeridos por simulações empíricas. A partir dos resultados procurou-se obter meios de tornar matematicamente precisa a presença de caos nestes modelos C-BAM e assegurar condição suficiente para esta formalização. Para sistemas discretos, o fenômeno caótico é usualmente definido pela positividade do maior expoente de Lyapunov (Parker e Chua, 1989).

Embora seja conhecido que diferentes comunidades apresentem definições diferentes para caos (Ikeguchi, Hasegawa e Aihara, 1999), todas as definições encontradas compartilham algumas características em comum, em particular a dependência sensível às condições iniciais. Uma maneira eficiente de descrever a sensibilidade às condições iniciais é através do cálculo dos expoentes de Lyapunov, pois constituem um método conveniente para computação e análise dos resultados (Ikeguchi, Hasegawa e Aihara, 1999).

Em uma breve recordação do conceito apresentado no Capítulo 2, o expoente de Lyapunov é um número que descreve a dinâmica da evolução da trajetória. Ele é usado para medir a taxa média de convergência ou divergência de duas trajetórias vizinhas no espaço de fase. Seus valores podem

ser números negativos, zero ou positivos. Valores negativos significam que as duas trajetórias seguem próximas uma da outra. Expoentes de Lyapunov positivos, ao contrário, resultam de trajetórias divergentes e aparecem apenas dentro do domínio caótico. Pelo fato do expoente de Lyapunov positivo medir a dependência das condições iniciais mostrando a taxa média pela qual dois pontos inicialmente próximos afastam-se com o passar do tempo, ele é um dos mais importantes indicadores de caos de um sistema. Um expoente de Lyapunov com valor zero corresponde a um espaço entre as trajetórias que permanece constante ou aumenta em um taxa menor do que a taxa exponencial (Ott, 2002).

Portanto, nesta seção, a família C-BAM é analisada e os expoentes de Lyapunov destas redes são calculados a fim de confirmar a existência da dinâmica caótica nesta família de modelos. A rede C-BAM, da família C-BAM, será utilizada para demonstração passo-a-passo de como se obteve a avaliação dos expoentes de Lyapunov.

O expoente de Lyapunov foi computado ao se utilizar a matriz Jacobiana do mapa $(2m+2n)$ -dimensional descrito pela Equação 5.1, dada a seguir, e o processo de ortonormalização de Gram-Schmidt (Anton e Rorres, 2001). Seja o estado $(2m+2n)$ -dimensional definido como na Equação 5.1:

$$\mathbf{z}(t) = (\eta_{x1}(t), \eta_{x2}(t), \dots, \eta_{xm}(t), \zeta_{x1}(t), \zeta_{x2}(t), \dots, \zeta_{xm}(t), \\ \eta_{y1}(t), \eta_{y2}(t), \dots, \eta_{yn}(t), \zeta_{y1}(t), \zeta_{y2}(t), \dots, \zeta_{yn}(t)) \quad (5.1)$$

A rede C-BAM pode ser interpretada como um sistema dinâmico $(2m+2n)$ -dimensional descrito pela Equação 5.2:

$$\mathbf{z}(t+1) = \mathbf{F}(\mathbf{z}(t)) \quad (5.2)$$

em que $\mathbf{z} \in \mathfrak{R}^{2m+2n}$ e $\mathbf{F}: \mathfrak{R}^{2m+2n} \rightarrow \mathfrak{R}^{2m+2n}$ é um mapa não linear no espaço de estado $(2m+2n)$ -dimensional, definido pelo conjunto de Equações 4.15 a 4.20.

Considerando o desvio infinitesimal $\delta\mathbf{z}(t)$, a partir de $\mathbf{z}(t)$, na Equação 5.2 e aplicando a linearização, obtém-se que:

$$\mathbf{z}(t+1) + \delta\mathbf{z}(t+1) = \mathbf{F}\{\mathbf{z}(t) + \delta\mathbf{z}(t)\} \quad (5.3)$$

Expandindo a Equação 5.3 por série de Taylor e descartando os termos de ordem maiores, o mapa linear da Equação 5.4 é obtido:

$$\delta\mathbf{z}(t+1) = \mathbf{DF}\{\mathbf{z}(t)\} \delta\mathbf{z}(t) \quad (5.4)$$

em que \mathbf{DF} é a matriz Jacobiana de \mathbf{F} em $\mathbf{z}(t)$. Ou seja, se $\mathbf{z}(t)$ é dada pela Equação 5.1 e $\mathbf{F}(\mathbf{z}(t))$ pela Equação 5.5 então $\mathbf{DF}\{\mathbf{z}(t)\}$ é descrita pela Equação 5.6.

$$\mathbf{F}(\mathbf{z}(t)) = \begin{pmatrix} F_1(\mathbf{z}(t)) \\ F_2(\mathbf{z}(t)) \\ \vdots \\ F_{2m+2n}(\mathbf{z}(t)) \end{pmatrix} \quad (5.5)$$

$$\mathbf{DF}(\mathbf{z}) = \begin{pmatrix} \frac{\partial F_1}{\partial \eta_{x1}} & \frac{\partial F_1}{\partial \eta_{x2}} & \cdots & \frac{\partial F_1}{\partial \zeta_{xm}} & \frac{\partial F_1}{\partial \eta_{y1}} & \cdots & \frac{\partial F_1}{\partial \zeta_{yn}} \\ \frac{\partial F_2}{\partial \eta_{x1}} & \frac{\partial F_2}{\partial \eta_{x2}} & \cdots & \frac{\partial F_2}{\partial \zeta_{xm}} & \frac{\partial F_2}{\partial \eta_{y1}} & \cdots & \frac{\partial F_2}{\partial \zeta_{yn}} \\ \vdots & \vdots & & \vdots & & & \\ \frac{\partial F_{2m}}{\partial \eta_{x1}} & \frac{\partial F_{2m}}{\partial \eta_{x2}} & \cdots & \frac{\partial F_{2m}}{\partial \zeta_{xm}} & \frac{\partial F_{2m}}{\partial \eta_{y1}} & \cdots & \frac{\partial F_{2m}}{\partial \zeta_{yn}} \\ \frac{\partial F_{2m+1}}{\partial \eta_{x1}} & \frac{\partial F_{2m+1}}{\partial \eta_{x2}} & \cdots & \frac{\partial F_{2m+1}}{\partial \zeta_{xm}} & \frac{\partial F_{2m+1}}{\partial \eta_{y1}} & \cdots & \frac{\partial F_{2m+1}}{\partial \zeta_{yn}} \\ \vdots & \vdots & & \vdots & & & \\ \frac{\partial F_{2m+2n}}{\partial \eta_{x1}} & \frac{\partial F_{2m+2n}}{\partial \eta_{x2}} & \cdots & \frac{\partial F_{2m+2n}}{\partial \zeta_{xm}} & \frac{\partial F_{2m+2n}}{\partial \eta_{y1}} & \cdots & \frac{\partial F_{2m+2n}}{\partial \zeta_{yn}} \end{pmatrix} \quad (5.6)$$

Para calcular o expoente de Lyapunov, bases ortonormais no espaço tangente $(2m+2n)$ -dimensional, $u_i(t)$, $(i=1,2,\dots, 2m+2n)$, são dadas como sendo $\delta\mathbf{z}(t)$ da Equação 5.4. Com o mapa linear \mathbf{DF} da Equação 5.4, as bases $u_i(t)$ são transformadas de acordo com a Equação 5.7.

$$e_i(t+1) = \mathbf{DF} u_i(t) \quad (5.7)$$

Uma vez que as bases $e_i(t+1)$ não são mais mutuamente ortogonais, o processo de ortonormalização de Gram-Schmidt é aplicado para formar novas bases ortonormais (Shimada e Nagashima, 1979) $q_i(t+1)$. Começando o processo tomando $q_1(t+1) = e_1(t+1)$, os demais vetores ortonormais são obtidos de acordo com a Equação 5.8.

$$\begin{aligned} q_i(t+1) &= e_i(t+1) - \left[\begin{aligned} &\langle e_i(t+1), u_1(t+1) \rangle u_1(t+1) + \\ &\langle e_i(t+1), u_2(t+1) \rangle u_2(t+1) + \cdots \\ &+ \langle e_i(t+1), u_{i-1}(t+1) \rangle u_{i-1}(t+1) \end{aligned} \right] \\ &= e_i(t+1) - \sum_{j=1}^{i-1} \langle e_i(t+1), u_j(t+1) \rangle u_j(t+1) \end{aligned} \quad (5.8)$$

em que \langle , \rangle denota o produto interno canônico (Lima, 2001) e $u_j(t+1)$ é dado pela Equação 5.9.

$$u_j(t+1) = \frac{q_j(t+1)}{\|q_j(t+1)\|} \quad (5.9)$$

em que $\| \cdot \|$ denota a norma euclidiana.

Repetindo este procedimento, os valores dos expoentes de Lyapunov são calculados de acordo com a Equação 5.10, para $i=1,2, \dots, (2m+2n)$.

$$\lambda_i = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \log |q_i(t)| \quad (5.10)$$

De acordo com a teoria (Ott, 2002), se pelo menos o maior expoente de Lyapunov, λ_1 , tiver valor positivo, isso significará que a rede C-BAM apresenta instabilidade orbital, que é uma das características fundamentais do caos determinístico.

5.2.1 Expoentes de Lyapunov na Família C-BAM

Conforme citado anteriormente, para os experimentos do presente capítulo foram construídas redes da família C-BAM com quinze neurônios caóticos ($m=15$) na camada X e dez neurônios caóticos ($n=10$) na camada Y. Sendo assim, as funções coordenadas da função F da Equação 5.5 relativas a camada X da C-BAM incluem $F_1(\mathbf{z}(t))$ a $F_{2m}(\mathbf{z}(t))$, sendo $F_1(\mathbf{z}(t))$ a $F_m(\mathbf{z}(t))$ relativas a η_{xi} e $F_{m+1}(\mathbf{z}(t))$ a $F_{2m}(\mathbf{z}(t))$ relativas a ς_{xi} . O conjunto de Equações 5.11 mostra as coordenadas de F relativas a camada X.

$$\begin{aligned} F_1(\mathbf{z}(t)) &= \eta_{x1}(t+1) = k_f \eta_{x1}(t) + \sum_{j=1}^n w_{1j} y_j(t) = k_f \eta_{x1}(t) + \sum_{j=1}^n (w_{1j} * f_y[\eta_{yj}(t) + \varsigma_{yj}(t)]) \\ F_2(\mathbf{z}(t)) &= \eta_{x2}(t+1) = k_f \eta_{x2}(t) + \sum_{j=1}^n w_{2j} y_j(t) = k_f \eta_{x2}(t) + \sum_{j=1}^n (w_{2j} * f_y[\eta_{yj}(t) + \varsigma_{yj}(t)]) \\ &\vdots \\ F_m(\mathbf{z}(t)) &= \eta_{xm}(t+1) = k_f \eta_{xm}(t) + \sum_{j=1}^n w_{mj} y_j(t) = k_f \eta_{xm}(t) + \sum_{j=1}^n (w_{mj} * f_y[\eta_{yj}(t) + \varsigma_{yj}(t)]) \\ F_{m+1}(\mathbf{z}(t)) &= \varsigma_{x1}(t+1) = k_r \varsigma_{x1}(t) - \rho x_1(t) + a = k_r \varsigma_{x1}(t) - \rho \{f_x[\eta_{x1}(t) + \varsigma_{x1}(t)]\} + a \\ &\vdots \\ F_{2m}(\mathbf{z}(t)) &= \varsigma_{xm}(t+1) = k_r \varsigma_{xm}(t) - \rho x_m(t) + a = k_r \varsigma_{xm}(t) - \rho \{f_x[\eta_{xm}(t) + \varsigma_{xm}(t)]\} + a \end{aligned} \quad (5.11)$$

As funções coordenadas da função F da Equação 5.5 relativas a camada Y da C-BAM incluem $F_{2m+1}(\mathbf{z}(t))$ a $F_{2m+2n}(\mathbf{z}(t))$, sendo $F_{2m+1}(\mathbf{z}(t))$ a $F_{2m+n}(\mathbf{z}(t))$ relativas a η_{yj} e $F_{2m+n+1}(\mathbf{z}(t))$ a $F_{2m+2n}(\mathbf{z}(t))$ relativas a ς_{yj} . O conjunto de equações 5.12 mostra as coordenadas de F relativas a camada Y.

$$\begin{aligned}
F_{2m+1}(\mathbf{z}(t)) &= \eta_{y1}(t+1) = k_f \eta_{y1}(t) + \sum_{i=1}^m w_{1i} x_i(t) = k_f \eta_{y1}(t) + \sum_{i=1}^m (w_{1i} * f_x[\eta_{xi}(t) + \varsigma_{xi}(t)]) \\
&\vdots \\
F_{2m+n}(\mathbf{z}(t)) &= \eta_{yn}(t+1) = k_f \eta_{yn}(t) + \sum_{i=1}^m w_{mi} x_i(t) = k_f \eta_{yn}(t) + \sum_{i=1}^m (w_{mi} * f_x[\eta_{xi}(t) + \varsigma_{xi}(t)]) \\
F_{2m+n+1}(\mathbf{z}(t)) &= \varsigma_{y1}(t+1) = k_r \varsigma_{y1}(t) - \rho x_1(t) + a = k_r \varsigma_{y1}(t) - \rho \{h_x[\eta_{y1}(t) + \varsigma_{y1}(t)]\} + a \\
&\vdots \\
F_{2m+2n}(\mathbf{z}(t)) &= \varsigma_{yn}(t+1) = k_r \varsigma_{yn}(t) - \rho x_n(t) + a = k_r \varsigma_{yn}(t) - \rho \{h_x[\eta_{yn}(t) + \varsigma_{yn}(t)]\} + a
\end{aligned} \tag{5.12}$$

Para exemplificar como foi realizado o cálculo das derivadas parciais do mapa linear DF da Equação 5.5, foi selecionada a função coordenada $F_1(\mathbf{z}(t))$.

$$\frac{\partial F_1(\mathbf{z}(t))}{\partial z_1} = \frac{\partial F_1(\mathbf{z}(t))}{\partial \eta_{x1}} = \frac{\partial(k_f \eta_{x1}(t))}{\partial \eta_{x1}(t)} + \frac{\partial \left[\sum_{j=1}^n (w_{1j} * f_y[\eta_{yj}(t) + \varsigma_{yj}(t)]) \right]}{\partial \eta_{x1}(t)} \tag{5.13}$$

Dividindo a Equação 5.13 em duas partes: a parte (A) para a primeira parcela da soma e a parte (B) para a segunda parcela. A solução da parte (A) é facilmente calculada:

$$\frac{\partial(k_f \eta_{x1}(t))}{\partial \eta_{x1}(t)} = k_f \tag{5.14}$$

Para solucionar a parte (B) da Equação 5.13, é preciso calcular a derivada parcial de $f_y(v)$, sendo $v = g(\eta, \varsigma) = \eta + \varsigma$. Ou seja, dada $f_y(g(\eta, \varsigma))$, deseja-se saber calcular sua derivada parcial com relação a η e a ς . Estas duas derivações são calculadas da seguinte forma:

- 1- derivada parcial com relação a η : $Df_y(g(\eta, \varsigma)) * D_1g(\eta, \varsigma)$
- 2- derivada parcial com relação a ς : $Df_y(g(\eta, \varsigma)) * D_2g(\eta, \varsigma)$

Como w_{1j} é um número, então a parte (B) da Equação 5.13 pode ser re-escrita de acordo com a Equação 5.14:

$$\frac{\sum_{j=1}^n w_{1j} \partial(f_y[\eta_{yj}(t) + \varsigma_{yj}(t)])}{\partial \eta_{x1}(t)} \tag{5.15}$$

Abrindo o somatório da Equação 5.15 em n parcelas é obtida a Equação 5.16:

$$w_{11} * f_y[\eta_{y1}(t) + \varsigma_{y1}(t)] + w_{12} * f_y[\eta_{y2}(t) + \varsigma_{y2}(t)] + \dots + w_{1n} * f_y[\eta_{yn}(t) + \varsigma_{yn}(t)] \tag{5.16}$$

Derivando a Equação 5.16 com relação a η_{x1} :

$$w_{11} * Df_y[\eta_{y1}(t) + \varsigma_{y1}(t)] * D_1[\eta_{y1}(t) + \varsigma_{y1}(t)] + \dots + w_{1n} * Df_y[\eta_{yn}(t) + \varsigma_{yn}(t)] * D_1[\eta_{yn}(t) + \varsigma_{yn}(t)] \tag{5.17}$$

Como $D_l[\eta_{yj}(t) + \zeta_{yj}(t)]$ tem valor zero para $j=1, \dots, n$, uma vez que esta é uma derivada de η_y e ζ_y por η_x e ζ_x , a Equação 5.17 também terá valor zero.

Portanto, juntando a solução das partes (A) e (B), o resultado da Equação 5.13 pode ser visto na Equação 5.18.

$$\frac{\partial F_1(\mathbf{z}(t))}{\partial z_1} = \frac{\partial F_1(\mathbf{z}(t))}{\partial \eta_{x1}} = k_f + 0 = k_f \quad (5.18)$$

As derivadas de $F_1(\mathbf{z}(t))$ com relação a η_{xi} , para $i=2, \dots, m$, e com relação a ζ_{xi} , para $i=1, \dots, m$, serão nulas, pois tanto a parte (A) quanto a parte (B) destas derivadas parciais terão valor zero.

Para exemplificar o cálculo das derivadas de $F_1(\mathbf{z}(t))$ com relação a η_{yj} , para $j=1, \dots, n$, e com relação a ζ_{yj} , para $j=1, \dots, n$, será realizado o cálculo de $F_1(\mathbf{z}(t))$ com relação a η_{y1} .

$$\begin{aligned} \frac{\partial F_1(\mathbf{z}(t))}{\partial \eta_{y1}(t)} &= \frac{\partial(k_f \eta_{x1}(t))}{\partial \eta_{y1}(t)} + \frac{\left[\sum_{j=1}^n (w_{1j} * f_y[\eta_{yj}(t) + \zeta_{yj}(t)]) \right]}{\partial \eta_{y1}(t)} \\ &= 0 + [w_{11} * Df_y(\eta_{y1}(t) + \zeta_{y1}(t)) * 1] \\ &= w_{11} * Df_y[\eta_{y1}(t) + \zeta_{y1}(t)] \end{aligned} \quad (5.19)$$

Como a função f da rede C-BAM é a função logística, então:

$$f(v) = \frac{1}{\{1 + e^{(-v/\varepsilon)}\}} \Rightarrow \frac{\partial f(v)}{\partial v} = \frac{e^{(-v/\varepsilon)}}{\varepsilon \{1 + e^{-v/\varepsilon}\}^2} \quad (5.20)$$

Substituindo a Equação 5.20 na Equação 5.19, a solução da derivada $F_1(\mathbf{z}(t))$ com relação a η_{y1} fica:

$$\frac{\partial F_1(\mathbf{z}(t))}{\partial \eta_{y1}} = w_{11} * \left(\frac{e^{\{-[\eta_{y1}(t) + \zeta_{y1}(t)]/\varepsilon\}}}{\varepsilon * [1 + e^{\{-[\eta_{y1}(t) + \zeta_{y1}(t)]/\varepsilon\}}]^2} \right) \quad (5.21)$$

Realizando de modo semelhante os cálculos feitos para encontrar as derivadas da função coordenada $F_1(\mathbf{z}(t))$ do mapa linear DF , as derivadas parciais das outras funções coordenadas podem ser facilmente encontradas. A Tabela 5.5 apresenta as derivadas parciais da matriz Jacobiana DF de F em $\mathbf{z}(t)$, de acordo com a Equação 5.6, aplicada à rede C-BAM. A fim de que a tabela não ficasse sobrecarregada de equações, apenas as funções coordenadas com mudanças significativas nos cálculos das derivadas foram explicitadas na tabela. As demais derivadas podem ser facilmente deduzidas a partir das próprias equações inseridas na tabela.

Tabela 5.5. Derivadas parciais da matriz Jacobiana DF de F em $z(t)$.

	η_{x1}	η_{x2}	...	η_{xm}	ζ_{x1}	...	ζ_{xm}	η_{y1}	...	η_{yn}	ζ_{y1}	...	ζ_{yn}
F_1	k_f	0	0	0	0	0	0	$w_{11}^* \text{Df}[\eta_{y1} + \zeta_{y1}]$		$w_{1n}^* \text{Df}[\eta_{yn} + \zeta_{yn}]$	$w_{11}^* \text{Df}[\eta_{y1} + \zeta_{y1}]$		$w_{1n}^* \text{Df}[\eta_{yn} + \zeta_{yn}]$
F_2	0	k_f	0	0	0	0	0	$w_{21}^* \text{Df}[\eta_{y1} + \zeta_{y1}]$		$w_{2n}^* \text{Df}[\eta_{yn} + \zeta_{yn}]$	$w_{21}^* \text{Df}[\eta_{y1} + \zeta_{y1}]$		$w_{2n}^* \text{Df}[\eta_{yn} + \zeta_{yn}]$
....													
F_m	0	0	0	k_f	0	0	0	$w_{m1}^* \text{Df}[\eta_{y1} + \zeta_{y1}]$		$w_{mn}^* \text{Df}[\eta_{yn} + \zeta_{yn}]$	$w_{m1}^* \text{Df}[\eta_{y1} + \zeta_{y1}]$		$w_{mn}^* \text{Df}[\eta_{yn} + \zeta_{yn}]$
F_{m+1}	$-\rho^* \text{Df}[\eta_{x1} + \zeta_{x1}]$	0	0	0	$k_r - \rho^* \text{Df}[\eta_{x1} + \zeta_{x1}]$	0	0	0		0	0	0	0
....													
F_{2m}	0	0	0	$-\rho^* \text{Df}[\eta_{xm} + \zeta_{xm}]$	0	0	$k_r - \rho^* \text{Df}[\eta_{xm} + \zeta_{xm}]$	0		0	0	0	0
F_{2m+1}	$w_{11}^* \text{Df}[\eta_{x1} + \zeta_{x1}]$	$w_{12}^* \text{Df}[\eta_{x2} + \zeta_{x2}]$		$w_{1m}^* \text{Df}[\eta_{xm} + \zeta_{xm}]$	$w_{11}^* \text{Df}[\eta_{x1} + \zeta_{x1}]$		$w_{1m}^* \text{Df}[\eta_{xm} + \zeta_{xm}]$	k_f		0	0	0	0
....													
F_{2m+n}	$w_{n1}^* \text{Df}[\eta_{x1} + \zeta_{x1}]$	$w_{n2}^* \text{Df}[\eta_{x2} + \zeta_{x2}]$		$w_{nm}^* \text{Df}[\eta_{xm} + \zeta_{xm}]$	$w_{n1}^* \text{Df}[\eta_{x1} + \zeta_{x1}]$		$w_{nm}^* \text{Df}[\eta_{xm} + \zeta_{xm}]$	0		k_f	0	0	0
F_{2m+n+1}	0	0	0	0	0	0	0	$-\rho^* \text{Df}[\eta_{y1} + \zeta_{y1}]$		0	$k_r - (\rho^* \text{Df}[\eta_{y1} + \zeta_{y1}])$		0
....													
F_{2m+2n}	0	0	0	0	0	0	0	0		$-\rho^* \text{Df}[\eta_{yn} + \zeta_{yn}]$	0		$k_r - (\rho^* \text{Df}[\eta_{yn} + \zeta_{yn}])$

Usando o procedimento descrito pelas Equações 5.7 a 5.10, assumindo $u_i(t) = 0$ como sendo vetores da matriz Identidade $I_{2n \times 2n}$ e $N=1000$, foram encontrados os valores apresentados na Tabela 5.6 para os expoentes de Lyapunov da rede C-BAM, em que $k_r=0.9$, $k_f=0.3$, $\rho=10.0$, e $a=2.0$. Como pode ser visto na Tabela 5.6, o maior expoente de Lyapunov computado é 0.057. A existência do expoente de Lyapunov positivo comprova que a rede C-BAM apresenta resposta caótica.

Tabela 5.6. Expoentes de Lyapunov da rede C-BAM com os parâmetros $k_f = 0.2$, $k_r = 0.6$, $\rho = 8.0$ e $a=2.0$.

λ_1	0.057	λ_{18}	-0.137	λ_{35}	-0.205
λ_2	-0.020	λ_{19}	-0.140	λ_{36}	-0.210
λ_3	-0.045	λ_{20}	-0.147	λ_{37}	-0.212
λ_4	-0.077	λ_{21}	-0.147	λ_{38}	-0.215
λ_5	-0.080	λ_{22}	-0.152	λ_{39}	-0.217
λ_6	-0.080	λ_{23}	-0.155	λ_{40}	-0.220
λ_7	-0.090	λ_{24}	-0.157	λ_{41}	-0.225
λ_8	-0.092	λ_{25}	-0.157	λ_{42}	-0.230
λ_9	-0.097	λ_{26}	-0.160	λ_{43}	-0.237
λ_{10}	-0.100	λ_{27}	-0.160	λ_{44}	-0.240
λ_{11}	-0.102	λ_{28}	-0.162	λ_{45}	-0.240
λ_{12}	-0.110	λ_{29}	-0.165	λ_{46}	-0.240
λ_{13}	-0.110	λ_{30}	-0.170	λ_{47}	-0.240
λ_{14}	-0.117	λ_{31}	-0.172	λ_{48}	-0.252
λ_{15}	-0.120	λ_{32}	-0.172	λ_{49}	-0.252
λ_{16}	-0.125	λ_{33}	-0.177	λ_{50}	-0.260
λ_{17}	-0.132	λ_{34}	-0.197		

De acordo com Li e Chen (2004), um sistema dinâmico é considerado caótico se o maior expoente de Lyapunov for positivo. No entanto, pode haver a ocorrência de um número qualquer de expoentes de Lyapunov positivos, quase tantos quanto a dimensão do próprio espaço de fase do sistema. Se o sistema dinâmico apresenta mais do que dois expoentes de Lyapunov positivos, então o sistema é considerado hiper caótico (Li e Chen (2004)). A Tabela 5.7 apresenta uma situação ($k_r=0.9$, $k_f=0.3$, $\rho=10.0$, e $a=2.0$) em que ocorrem seis expoentes de Lyapunov para a rede C-BAM.

Os expoentes de Lyapunov foram também calculados para as outras redes da família C-BAM: a BAM com atraso e a C-eBAM, com cálculos similares aos detalhadamente apresentados nesta Seção para a rede C-BAM.

A Tabela 5.8 mostra os valores dos expoentes de Lyapunov para a rede C-BAM com atraso quando a ela são atribuídos valores $k_f=0.1$, $k_r=0.7$, $\rho=8.0$, $a=2.0$ e $\sigma=3$. O valor positivo do maior expoente de Lyapunov (0.045) confirma a presença de caos neste modelo C-BAM.

Tabela 5.7. Expoentes de Lyapunov da rede C-BAM com os parâmetros $k_f = 0.3$, $k_r = 0.9$, $\rho = 10.0$ e $a = 2.0$.

λ_1	0.165	λ_{18}	-0.087	λ_{35}	-0.182
λ_2	0.150	λ_{19}	-0.090	λ_{36}	-0.185
λ_3	0.132	λ_{20}	-0.095	λ_{37}	-0.190
λ_4	0.112	λ_{21}	-0.102	λ_{38}	-0.195
λ_5	0.070	λ_{22}	-0.115	λ_{39}	-0.200
λ_6	0.045	λ_{23}	-0.145	λ_{40}	-0.207
λ_7	-0.010	λ_{24}	-0.152	λ_{41}	-0.217
λ_8	-0.050	λ_{25}	-0.165	λ_{42}	-0.232
λ_9	-0.057	λ_{26}	-0.170	λ_{43}	-0.245
λ_{10}	-0.065	λ_{27}	-0.170	λ_{44}	-0.250
λ_{11}	-0.065	λ_{28}	-0.172	λ_{45}	-0.252
λ_{12}	-0.067	λ_{29}	-0.172	λ_{46}	-0.252
λ_{13}	-0.070	λ_{30}	-0.172	λ_{47}	-0.262
λ_{14}	-0.072	λ_{31}	-0.175	λ_{48}	-0.262
λ_{15}	-0.072	λ_{32}	-0.177	λ_{49}	-0.267
λ_{16}	-0.080	λ_{33}	-0.177	λ_{50}	-0.280
λ_{17}	-0.085	λ_{34}	-0.180		

Tabela 5.8. Expoentes de Lyapunov da rede C-BAM com atraso com os parâmetros $k_f = 0.1$, $k_r = 0.7$, $\rho = 8.0$, $a = 2.0$ e $\sigma = 3$.

λ_1	0.045	λ_{18}	-0.092	λ_{35}	-0.197
λ_2	-0.002	λ_{19}	-0.150	λ_{36}	-0.200
λ_3	-0.045	λ_{20}	-0.152	λ_{37}	-0.202
λ_4	-0.045	λ_{21}	-0.157	λ_{38}	-0.207
λ_5	-0.055	λ_{22}	-0.160	λ_{39}	-0.207
λ_6	-0.060	λ_{23}	-0.162	λ_{40}	-0.210
λ_7	-0.060	λ_{24}	-0.172	λ_{41}	-0.225
λ_8	-0.070	λ_{25}	-0.172	λ_{42}	-0.230
λ_9	-0.072	λ_{26}	-0.172	λ_{43}	-0.235
λ_{10}	-0.075	λ_{27}	-0.180	λ_{44}	-0.250
λ_{11}	-0.077	λ_{28}	-0.180	λ_{45}	-0.252
λ_{12}	-0.080	λ_{29}	-0.180	λ_{46}	-0.255
λ_{13}	-0.080	λ_{30}	-0.182	λ_{47}	-0.255
λ_{14}	-0.080	λ_{31}	-0.182	λ_{48}	-0.255
λ_{15}	-0.085	λ_{32}	-0.187	λ_{49}	-0.260
λ_{16}	-0.087	λ_{33}	-0.187	λ_{50}	-0.267
λ_{17}	-0.090	λ_{34}	-0.190		

Os valores dos expoentes de Lyapunov para a rede C-eBAM assumindo os valores de parâmetros $k_f = 0.25$, $k_r = 0.7$, $\rho = 9.0$, $a = 3.0$ e $\beta = 1.3$ podem ser vistos na Tabela 5.9. Neste caso, o maior expoente de Lyapunov vale 0.075 e comprova a existência de caos na C-eBAM.

Tabela 5.9. Expoentes de Lyapunov da rede C-eBAM com os parâmetros $k_f=0.25$, $k_r=0.7$, $\rho=9.0$, $a=3.0$ e $\beta=1.3$.

λ_1	0.075	λ_{18}	-0.475	λ_{35}	-0.645
λ_2	-0.255	λ_{19}	-0.500	λ_{36}	-0.655
λ_3	-0.285	λ_{20}	-0.520	λ_{37}	-0.695
λ_4	-0.290	λ_{21}	-0.525	λ_{38}	-0.700
λ_5	-0.295	λ_{22}	-0.525	λ_{39}	-0.730
λ_6	-0.330	λ_{23}	-0.545	λ_{40}	-0.775
λ_7	-0.350	λ_{24}	-0.550	λ_{41}	-0.785
λ_8	-0.355	λ_{25}	-0.550	λ_{42}	-0.785
λ_9	-0.365	λ_{26}	-0.555	λ_{43}	-0.800
λ_{10}	-0.365	λ_{27}	-0.560	λ_{44}	-0.860
λ_{11}	-0.375	λ_{28}	-0.580	λ_{45}	-0.875
λ_{12}	-0.410	λ_{29}	-0.595	λ_{46}	-0.905
λ_{13}	-0.410	λ_{30}	-0.605	λ_{47}	-0.905
λ_{14}	-0.415	λ_{31}	-0.610	λ_{48}	-0.955
λ_{15}	-0.425	λ_{32}	-0.630	λ_{49}	-0.960
λ_{16}	-0.450	λ_{33}	-0.630	λ_{50}	-1.015
λ_{17}	-0.470	λ_{34}	-0.635		

5.3 Recuperação de Memórias Não Acessíveis

O papel fundamental de uma memória associativa é armazenar e recuperar padrões. Estas operações envolvem o treinamento de rede e a recuperação de um padrão memorizado em resposta à apresentação de uma versão ruidosa ou distorcida de um padrão chave inicial (padrão anteriormente apresentado à rede).

É um requisito de memórias hetero-associativas armazenar um conjunto arbitrário de padrões de entrada que esteja pareado com outro conjunto arbitrário de padrões de saída. Um padrão chave \mathbf{X}^k ($k = 1, 2, \dots, p$, em que p é o número de padrões armazenados na rede) age como se fosse um estímulo que não apenas determina a posição de armazenamento do padrão \mathbf{Y}^k memorizado, mas também funciona como um parâmetro para a recuperação de \mathbf{Y}^k . Para uma recuperação perfeita, deveria ser encontrado $\mathbf{Y} = \mathbf{Y}^j$, em que \mathbf{Y}^j é o padrão armazenado associado ao padrão chave \mathbf{X}^j . Quando $\mathbf{Y} \neq \mathbf{Y}^j$, para $\mathbf{X} = \mathbf{X}^j$, é dito que houve um erro na recuperação do padrão (Haykin, 1999).

O número de padrões armazenados nas memórias hetero-associativas fornece uma medida direta da capacidade de armazenamento da rede. É desejável que a capacidade de armazenamento de um modelo neural de memória associativa seja a maior possível para permitir que uma grande parte dos padrões armazenados seja recuperada corretamente.

A rede BAM, paradigma selecionado para o presente trabalho, comporta-se como uma memória hetero-associativa de conteúdo endereçável, armazenando e recuperando pares de

padrões. Foi provado (Kosko, 1988) que a BAM converge para algum ponto estável. Entretanto, a capacidade de armazenamento da BAM é pequena (limitada pelo número de neurônios da rede) e freqüentemente a BAM converge para atratores espúrios. É importante esclarecer que, na BAM, alguns pares armazenados que deveriam formar mínimos locais, não o fazem.

Novos modelos de memórias hetero-associativas foram propostas com o objetivo de melhorar o desempenho da BAM, entre eles as redes da família BAM consideradas neste trabalho: a C-BAM com atraso e a C-eBAM. A BAM exponencial (e-BAM) procura melhorar o desempenho da BAM utilizando-se das vantagens da estratégia exponencial: a não linearidade exponencial desta memória de conteúdo endereçável tem uma capacidade de armazenamento exponencialmente escalonada com o número de bits e tem maior capacidade de correção de erros (Wang e Don, 1995) (Wilson e Hancock, 2001). Xiaofeng e Jueban (1998) sugeriram que o atraso de tempo na BAM produz um ruído eficaz na dinâmica da rede, fazendo com que o processo de recuperação de memórias seja melhorado.

Apesar de a e-BAM e de a BAM com atraso apresentarem alternativas para melhorar o desempenho da BAM, ambas continuam apresentando convergência para padrões $\mathbf{Y} \neq \mathbf{Y}^j$ para $\mathbf{X} = \mathbf{X}^j$. Como conseqüência, existem padrões memorizados \mathbf{Y}^j que não podem ser recuperados nem através de seu par associado \mathbf{X}^j nem por outro padrão chave \mathbf{X}^k . Nesta tese, tais padrões são chamados de padrões inacessíveis.

Nos experimentos realizados com a família C-BAM foi também observado o processo de recuperação das redes. A presente seção mostra como o fenômeno caótico permite que a C-BAM, a C-BAM com atraso e a C-eBAM acessem padrões memorizados que os modelos originais destas redes (ou seja, os modelos não caóticos destas redes hetero-associativas) não podiam recuperar. Em outras palavras, a seção mostra como ocorre a recuperação de padrões memorizados inacessíveis ou padrões memorizados escondidos na família C-BAM. Deste modo, novos experimentos empíricos foram realizados com o objetivo de mostrar como o fenômeno caótico modifica a sensibilidade e a acessibilidade da família de redes C-BAM e permitem que elas recuperem padrões memorizados anteriormente inacessíveis.

A Figura 5.9 ilustra os quatro pares hetero-associativos armazenados utilizados nos experimentos desta seção. Os pares de vetores $(\mathbf{X}^k, \mathbf{Y}^k)$, $k=1, \dots, 4$, apresentam as seguintes dimensões: $\mathbf{X} \in \{0,1\}^m$, $m=1, \dots, 15$ e $\mathbf{Y} \in \{0,1\}^n$, $n=1, \dots, 10$. Os quadrados pretos simbolizam bits de valor um, enquanto os quadrados brancos simbolizam bits de valor zero.

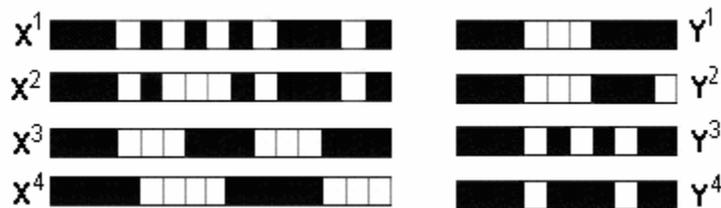


Figura 5.9. Os pares de vetores $(\mathbf{X}^k, \mathbf{Y}^k)$, $k=1, \dots, 4$, usados nos experimentos com os modelos C-BAM.

Conforme visto anteriormente, o comportamento das redes C-BAM torna-se idêntico ao da rede BAM original quando todos os parâmetros (k_r , k_f , ρ e a) das redes C-BAM recebem valor zero. Nesta situação, as redes C-BAM convergem para um ponto fixo (período um) a partir de um padrão de entrada específico. É importante destacar que, apesar de as redes C-BAM e as redes BAM apresentarem o mesmo comportamento, tais redes são diferentes por causa da função da ativação, que é $S(\cdot)$ para a BAM e $f(\cdot)$ para a C-BAM. Esta diferença pode ser visualizada nas Figuras 4.2 e 4.3.

A Tabela 5.10 mostra a ocorrência de erros no processo de recuperação, porque $\mathbf{Y} \neq \mathbf{Y}^j$, para $\mathbf{X} = \mathbf{X}^j$, são encontrados (Haykin, 1999). Pode ser notado na Tabela 5.10 que a eBAM, melhor, a C-eBAM com valores de parâmetros zerados, não converge para \mathbf{Y}^3 dado o padrão inicial de entrada \mathbf{X}^3 . Também, tanto a BAM quanto a BAM com atraso não convergem para \mathbf{Y}^2 , \mathbf{Y}^3 e \mathbf{Y}^4 dados os padrões de entrada \mathbf{X}^2 , \mathbf{X}^3 e \mathbf{X}^4 , respectivamente. Isto significa que há padrões memorizados que não podem ser recuperados ou ainda, são inacessíveis às redes da família C-BAM.

Tabela 5.10. Relação entre as entradas (\mathbf{X}) e as saídas (\mathbf{Y}) das redes da família C-BAM para $k_r = k_f = \rho = a = 0$.

Saída \ Entrada	\mathbf{X}^1	\mathbf{X}^2	\mathbf{X}^3	\mathbf{X}^4
C-BAM	\mathbf{Y}^1	\mathbf{Y}^1	\mathbf{Y}^1	\mathbf{Y}^1
C-BAM with delay ($\sigma = 5$)	\mathbf{Y}^1	\mathbf{Y}^1	\mathbf{Y}^1	\mathbf{Y}^1
C-eBAM ($\beta = 1.3$)	\mathbf{Y}^1	\mathbf{Y}^2	\mathbf{Y}^1	\mathbf{Y}^4

Outros conjuntos de valores de parâmetros para a família C-BAM mostraram que a inclusão do neurônio caótico nestes modelos resultou em uma nova maneira de acessibilidade da memória. Foi visto na Seção 5.1 (Figuras 5.2, 5.3 e 5.4) que a refração acumulada inerente aos neurônios caóticos possibilita que os modelos C-BAM escapem de qualquer ponto fixo e passem a transitar entre todos os padrões armazenados. Uma vez que ocorre a transição entre todos os padrões armazenados (para alguns valores de parâmetros), então os modelos C-BAM permitem, em princípio, o acesso a todos os padrões, inclusive daqueles inacessíveis nos modelos originais, para qualquer padrão inicial de entrada. Como consequência, as redes BAM caóticas construídas no Capítulo 4 eliminam a ocorrência de padrões inacessíveis apresentada nos modelos BAM originais. Por outro lado, as redes BAM caóticas nem sempre convergem para pontos fixos.

As Tabelas 5.11, 5.12 e 5.13 mostram a acessibilidade a todos os padrões armazenados das redes C-BAM. Por exemplo, o padrão \mathbf{Y}^3 , que é um padrão não acessível à rede eBAM (ver Tabela 5.10), torna-se acessível à C-eBAM (Tabela 5.11). A Tabela 5.11 ilustra a frequência de recuperação dos padrões armazenados para cada padrão \mathbf{X}^k ($k=1,\dots,4$) inicial de entrada, em um total de 3000 iterações. Pode ser verificado que o padrão memorizado \mathbf{Y}^3 é recuperado em 123, 112, 123 e 108 iterações para os padrões iniciais de entrada \mathbf{X}^1 , \mathbf{X}^2 , \mathbf{X}^3 e \mathbf{X}^4 , respectivamente.

Similarmente, os padrões \mathbf{Y}^2 , \mathbf{Y}^3 e \mathbf{Y}^4 , antes não acessíveis para a BAM e para a BAM com atraso (ver Tabela 5.10), tornam-se acessíveis para a C-BAM e para a C-BAM com atraso, conforme ilustrado nas Tabelas 5.12 e 5.13, respectivamente.

Tabela 5.11. Frequência de recuperação dos padrões armazenados da Figura 5.10 na C-eBAM, para $k_r=0.9$, $k_f=0.45$, $\rho=8.0$, $a=2.0$ e $\beta=1.3$.

Saída \ Padrão Inicial	\mathbf{X}^1	\mathbf{X}^2	\mathbf{X}^3	\mathbf{X}^4
\mathbf{Y}^1	1113	1121	1127	1126
\mathbf{Y}^2	425	414	400	418
\mathbf{Y}^3	123	112	123	108
\mathbf{Y}^4	92	120	95	75

Tabela 5.12. Frequência de recuperação dos padrões armazenados da Figura 5.10 na C-BAM, para $k_r=0.95$, $k_f=0.45$, $\rho=7.0$ e $a=3.0$.

Saída \ Padrão Inicial	X^1	X^2	X^3	X^4
Y^1	710	710	714	712
Y^2	289	289	269	262
Y^3	202	202	201	200
Y^4	20	20	20	18

Tabela 5.13. Frequência de recuperação dos padrões armazenados da Figura 5.10 na C-BAM com atraso, para $k_r=0.9$, $k_f=0.45$, $\rho=8.0$, $a=2.0$ e $\sigma=5$.

Saída \ Padrão Inicial	X^1	X^2	X^3	X^4
Y^1	1538	1532	1531	1531
Y^2	424	429	425	420
Y^3	213	214	212	209
Y^4	219	223	218	210

5.4 Modelagem do Fenômeno Saltar Passos

Os experimentos apresentados nas seções anteriores deste capítulo mostraram que a dinâmica caótica modifica a acessibilidade das memórias armazenadas das redes C-BAM: ao invés de estas redes recuperarem um único par de padrões, tal como ocorre com as redes BAM originais, as versões caóticas da BAM produzem uma grande diversidade de padrões recuperados, em cujas versões emergem diversas seqüências de recuperações, compostas de padrões treinados e não treinados. Por esta característica de recuperação, os experimentos com a C-BAM sugeriram que a

nova maneira de acessar os padrões armazenados pode ser usada para modelar a habilidade de um indivíduo de se saltar passos na solução de uma tarefa.

Portanto, o objetivo desta seção é mostrar como a transição entre os padrões armazenados da C-BAM pode modelar a tarefa de resolver uma equação algébrica com a habilidade de saltar passos. O fenômeno a ser modelado, ou seja, o fenômeno de saltar passos, está descrito no Capítulo 3, na Seção 3.3.

5.4.1 Representação

Com base no trabalho de Blessing e Anderson (1996), optou-se por escolher como a dinâmica da C-BAM pode modelar o saltar passos através da memorização de estágios de operações algébricas. Assim, cada padrão representa um passo na trajetória em direção à solução do problema algébrico.

Duas equações foram selecionadas como estudo de casos. A primeira equação selecionada é aquela utilizada como exemplo do saltar passos no artigo original de Blessing e Anderson (1996). A Tabela 5.14 mostra os quatro passos que foram propostos pelos autores para a solução desta equação feita por um noviço.

A segunda equação, $-A / B - X = -C$, foi selecionada como um estudo de caso porque sua solução gera mais operações passo-a-passo e dá a chance de saltar mais passos intermediários. Seis passos foram propostos pelo autor deste documento para a solução da equação feita por um noviço. Estes passos estão mostrados na Tabela 5.15.

Em ambas as equações, cada passo da resolução representa um padrão a ser armazenado e a C-BAM pode ser vista como um operador algébrico no qual o experimento tem um número de padrões a ser memorizado. No caso da primeira equação de estudo são quatro os padrões memorizados: \mathbf{P}^1 ($-X-A=B$), \mathbf{P}^2 ($-X-A+A=B+A$), \mathbf{P}^3 ($-X=B+A$) e \mathbf{P}^4 ($X=-B-A$). Na segunda equação de estudo são seis os padrões memorizados: \mathbf{P}^1 ($-A/B-X=-C$), \mathbf{P}^2 ($-A/B-X+A=-C+A$), \mathbf{P}^3 ($/B-X=-C+A$), \mathbf{P}^4 ($/B-X*B=-C+A*B$), \mathbf{P}^5 ($-X=-C+A*B$) e \mathbf{P}^6 ($X=+C-A*B$). Assim, dado \mathbf{P}^k , a C-BAM teria que produzir a seqüência de \mathbf{P}^k a $\mathbf{P}^{\text{final}}$.

Tabela 5.14. Passos utilizados na modelagem da resolução da equação algébrica exemplificada em Blessing e Anderson (1996) pela C-BAM.

Equação dada $\Rightarrow -X - A = B$		
Passo	Ação a ser realizada	Equação Resultante
Passo 1	Escrever a equação	$-X - A = B$
Passo 2	Adicionar +A em ambos os lados da equação	$-X - A + A = B + A$
Passo 3	Realizar as operações com o +A adicionado	$-X = B + A$
Passo 4	Multiplicar a equação por -1	$X = -B - A$

Para o caso da primeira equação algébrica a ser modelada, a C-BAM tem 470 neurônios caóticos na camada de entrada X e 470 neurônios caóticos na camada de saída Y. A Figura 5.10 mostra cada um dos quatro padrões na forma de uma matriz 10 por 47, na qual cada neurônio representa um bit do padrão. Quadrados pretos significam bits de valor um, enquanto quadrados brancos significam bits de valor zero.

No caso da segunda equação a ser modelada, a C-BAM tem 650 neurônios caóticos na camada de entrada X e 650 neurônios caóticos na camada de saída Y. A Figura 5.11 mostra cada padrão na forma de uma matriz 10 por 65, em que cada neurônio representa um bit do padrão.

Tabela 5.15. Passos utilizados na modelagem da resolução da equação algébrica $-A/B - X = -C$.

Equação dada $\Rightarrow -A/B - X = -C$		
Passos	Ação a ser realizada	Equação Resultante
Passo 1	Escrever a equação	$-A/B - X = -C$
Passo 2	Adicionar um $+A$ em ambos os lados da equação	$-A/B - X + A = -C + A$
Passo 3	Realizar a operação com o $-A + A$ cancelado	$/B - X = -C + A$
Passo 4	Multiplicar ambos os lados da equação por $*B$	$/B - x * B = -C + A * B$
Passo 5	Realizar a operação com o $/B * B$ cancelado	$-x = -C + A * B$
Passo 6	Multiplicar a equação por -1	$x = C - A * B$



Figura 5.10. Os quatro padrões utilizados na modelagem do saltar passos da equação algébrica $-X - A = B$.



Figura 5.11. Os seis padrões armazenados utilizados na resolução da equação algébrica $-A/B-X=-C$ para mostrar a modelagem do fenômeno saltar passos.

5.4.2 Testes

A modelagem da resolução da equação algébrica $-A/B-X=-C$ pode gerar mais operações passo-a-passo e possibilitar maior chance de saltar passos intermediários. Portanto, a apresentação da modelagem do saltar passos pela C-BAM será iniciada com esta equação. Em seguida, será também ilustrada a modelagem da equação $-X-A = B$, de modo a confirmar a ocorrência do fenômeno saltar passos na C-BAM em um outro conjunto de padrões.

Para os experimentos com a equação $-A/B-X=-C$, a matriz de correlação foi construída de acordo com a Equação 4.21, para os pares (P^1, P^2) , (P^2, P^3) , (P^3, P^4) , (P^4, P^5) , (P^5, P^6) . A solução da tarefa executada no modo do noviço consta de se apresentar o padrão P^k e fazer com que a rede recupere P^{k+1} , $k= 1, \dots, 5$, após um período transitório.

Inicializando a rede C-BAM de 650 neurônios caóticos com todos os seus parâmetros valendo zero, o comportamento da rede caótica foi idêntico ao da rede BAM original. A Figura 5.12 mostra a C-BAM convergindo para o ponto fixo (P^1, P^2) , a partir do padrão inicial P^1 . A partir dos padrões iniciais P^2 e P^3 , a rede C-BAM com os parâmetros zerados convergiu para os pares associados armazenados (P^2, P^3) e (P^3, P^4) , respectivamente. No entanto, para os padrões iniciais P^4

e \mathbf{P}^5 , a C-BAM não conseguiu recuperar os pares $(\mathbf{P}^4, \mathbf{P}^5)$ e $(\mathbf{P}^5, \mathbf{P}^6)$. Ao invés disto, a C-BAM convergiu para pares espúrios.

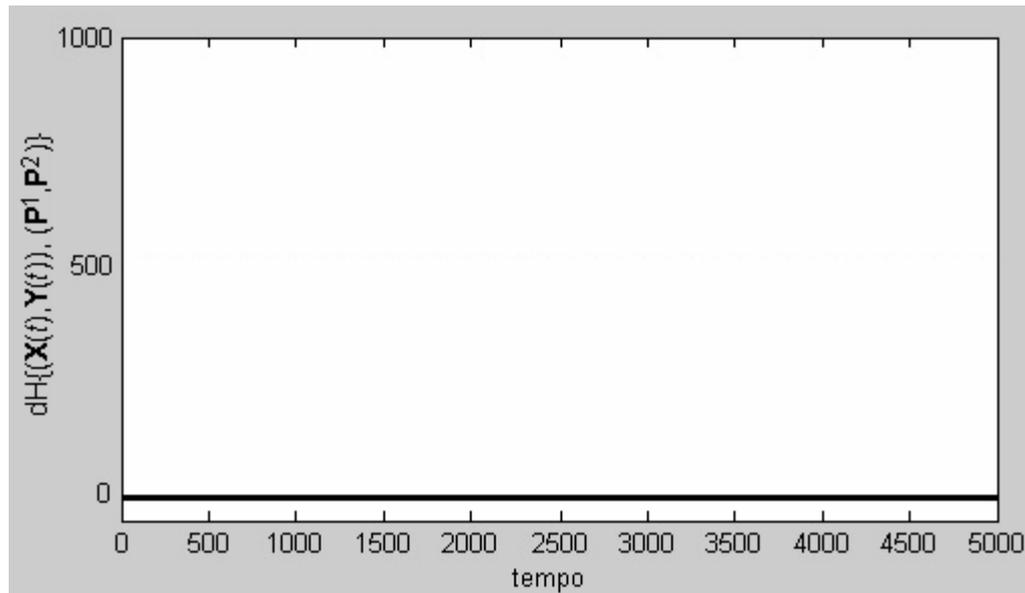


Figura 5.12. Distância Hamming entre os pares recuperados e o par $(\mathbf{P}^1, \mathbf{P}^2)$ para $k_r = k_f = \rho = a = 0$ e padrão inicial \mathbf{P}^1 . Período um (ponto fixo) ocorreu do instante $t=1$ em diante.

Atribuindo outros valores de parâmetros para C-BAM, esta rede apresentou comportamentos diferentes. A Figura 5.13 mostra o comportamento da C-BAM quando $k_r = 0.995$, $k_f = 0.25$, $\rho = 8.0$ e $a = 3.0$. Nesta situação, o modelo C-BAM apresenta diversidade de recuperação para os padrões armazenados. Quando o padrão inicial \mathbf{P}^1 foi apresentado à rede, a distância Hamming entre os padrões de saída e os padrões armazenados variou entre zero (recuperação dos pares armazenados) e 1300 (recuperação do inverso dos pares armazenados). Com este resultado da Figura 5.13 é possível verificar o acesso a todos os padrões armazenados, sem que ocorra a convergência para nenhum padrão em particular e, devido a esta propriedade, foi possível verificar a recuperação de pares inacessíveis na rede BAM original. Em resumo, os resultados das Figuras 5.12 e 5.13 mostram a ocorrência dos comportamentos verificados nas Seções 5.1 e 5.3 para um outro conjunto de padrões, o da Figura 5.11.

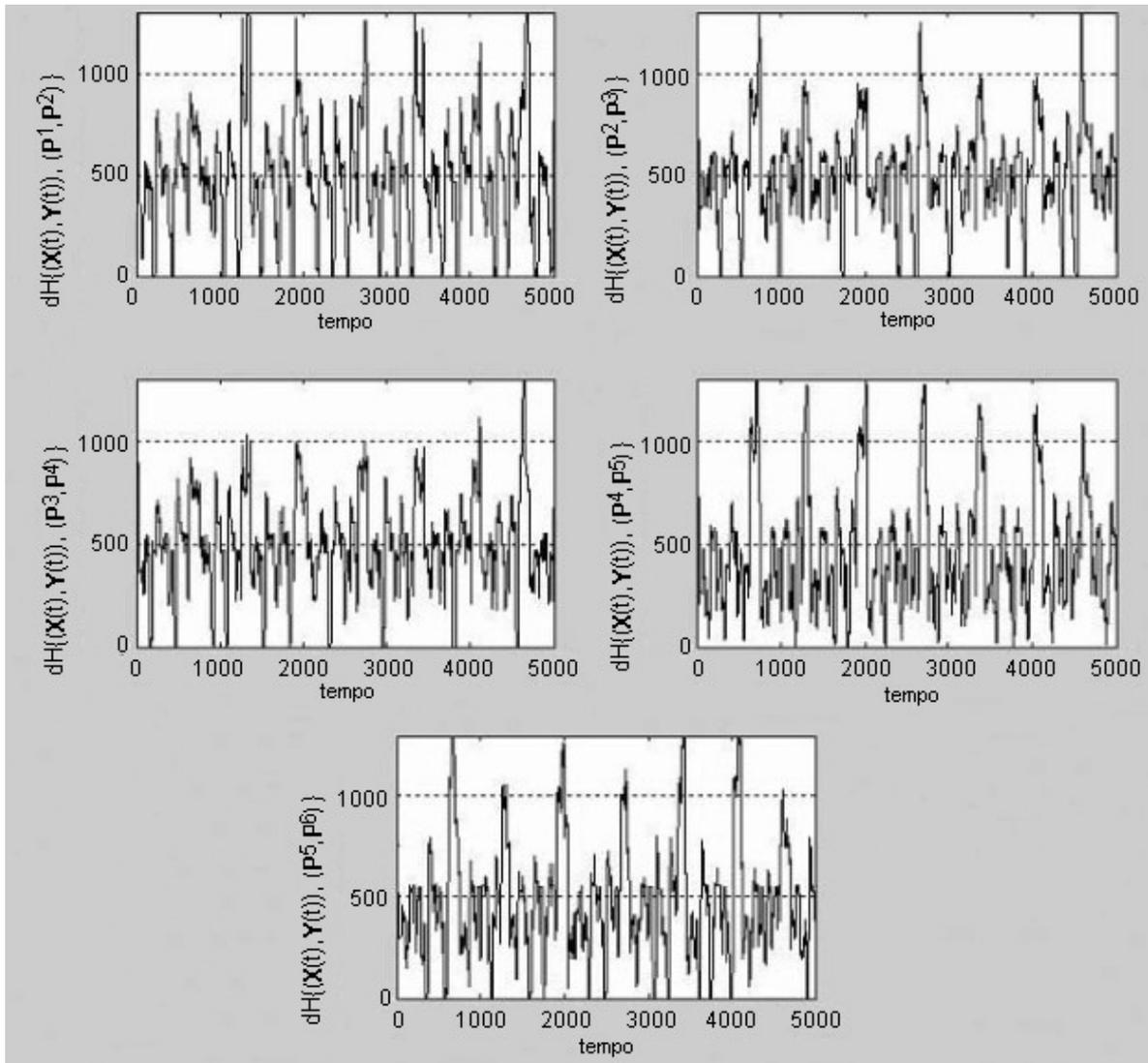


Figura 5.13. Distância Hamming entre os pares recuperados $(\mathbf{X}(t), \mathbf{Y}(t))$ e os pares associados armazenados $(\mathbf{P}^k, \mathbf{P}^{k+1})$, $k=1, \dots, 5$, para $k_r = 0.995$, $k_f = 0.25$, $\rho = 8.0$ e $a = 3.0$. Comportamento não periódico foi encontrado.

Outros experimentos com a C-BAM para os padrões da Figura 5.11 sugeriram a existência de uma trajetória de recuperação, criada pela dinâmica caótica. Foi verificado que a rede recebe o padrão inicial (representado por \mathbf{P}^1) para produzir o padrão final (\mathbf{P}^6), sem receber nenhuma entrada suplementar, nem necessariamente passar por todos os padrões da seqüência original de padrões armazenados (saltos de um padrão armazenado para outro podem acontecer).

A redução da trajetória de recuperação da seqüência de padrões armazenados pode representar um fenômeno psicológico chamado saltar passos, no qual um ou mais passos na solução de um problema pode ser saltado.

Quando o caminho para a recuperação de pares armazenados obedece a seqüência original de armazenamento ($\mathbf{P}^1 \rightarrow \mathbf{P}^2 \rightarrow \mathbf{P}^3 \rightarrow \mathbf{P}^4 \rightarrow \mathbf{P}^5 \rightarrow \mathbf{P}^6$), isto pode significar que a rede C-BAM

comporta-se como um noviço, uma vez que a rede transita todos os possíveis pares de padrões armazenados para alcançar o último padrão armazenado. A Figura 5.14 ilustra esta situação. Nesta figura, os pontos pretos na linha do par $(\mathbf{P}^k, \mathbf{P}^{k+1})$ indicam a ocorrência do par nos instantes de tempo especificados na coordenada horizontal. Note que o par $(\mathbf{P}^5, \mathbf{P}^6)$ foi recuperado na iteração número 2467 (ou $t=2467$), quando todos os demais pares já haviam sido recuperados.

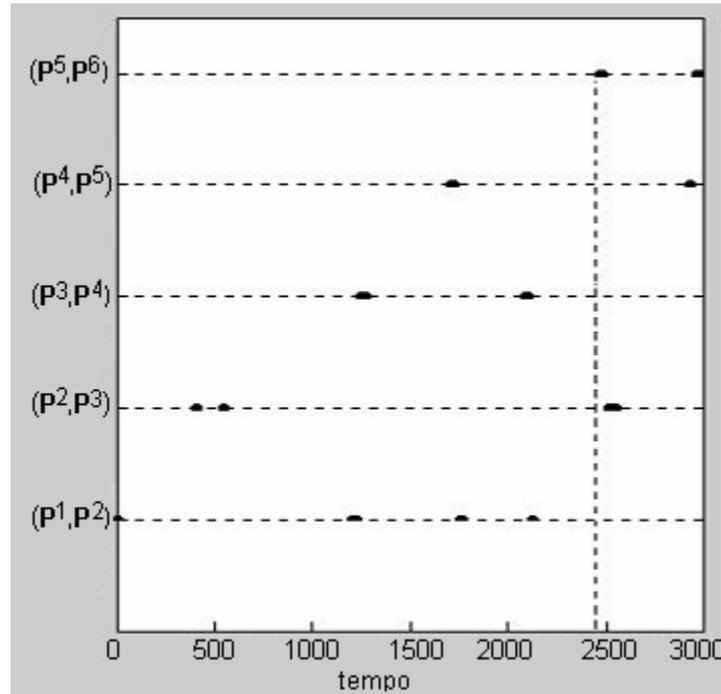


Figura 5.14. Seqüência de recuperação dos pares de padrões armazenados da Figura 5.11 para $k_r = 0.9975$, $k_f = 0.35$, $\rho = 9.0$ e $a = 4.0$.

Quando o caminho para a recuperação de pares armazenados de \mathbf{P}^1 a \mathbf{P}^6 elimina um ou mais padrões armazenados, isto pode significar que a C-BAM salta passos intermediários para executar esse trajeto. Neste caso, a rede opera com a habilidade de um especialista na solução da tarefa.

Variando apenas o parâmetro k_r e permanecendo com k_f , ρ e a inalterados, foi possível verificar que, muito freqüentemente, a C-BAM pula passos ao produzir trajetórias menores do que a trajetória completa de recuperação de \mathbf{P}^1 a \mathbf{P}^6 . As Figuras 5.15, 5.16 e 5.17 ilustram esse comportamento.

A Figura 5.15 mostra uma trajetória na qual a rede inicia a recuperação com \mathbf{P}^1 , gera padrões intermediários (treinados ou espúrios) e segue a seqüência correta de armazenamento sem, no entanto, passar pelo par armazenado $(\mathbf{P}^3, \mathbf{P}^4)$. A seqüência de recuperação dos pares treinados para $k_r = 0.9974$, $k_f = 0.35$, $\rho = 9.0$ e $a = 4.0$ é $(\mathbf{P}^1, \mathbf{P}^2) \rightarrow (\mathbf{P}^2, \mathbf{P}^3) \rightarrow (\mathbf{P}^4, \mathbf{P}^5) \rightarrow (\mathbf{P}^5, \mathbf{P}^6)$. O primeiro instante de tempo em que ocorre a recuperação de $(\mathbf{P}^5, \mathbf{P}^6)$ é em $t=1665$.

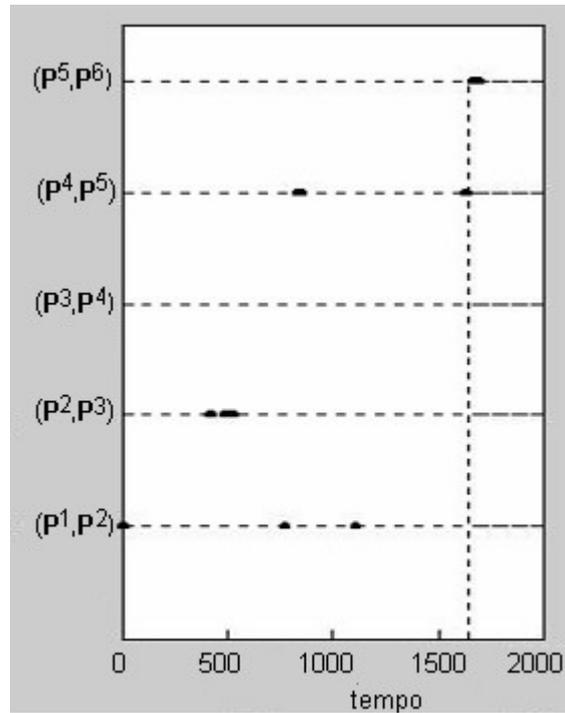


Figura 5.15. Seqüência de recuperação dos pares de padrões armazenados para $k_r = 0.9974$, $k_f = 0.35$, $\rho = 9.0$ e $a = 4.0$.

A Figura 5.16 mostra uma trajetória na qual a rede inicia a recuperação com \mathbf{P}^1 , e alcança \mathbf{P}^6 sem passar pelos pares armazenados $(\mathbf{P}^2, \mathbf{P}^3)$ e $(\mathbf{P}^3, \mathbf{P}^4)$. Isso significa que a C-BAM salta alguns passos da seqüência original de padrões associados memorizada. A seqüência recuperada para $k_r = 0.9973$, $k_f = 0.35$, $\rho = 9.0$ e $a = 4.0$ é $(\mathbf{P}^1, \mathbf{P}^2) \rightarrow (\mathbf{P}^4, \mathbf{P}^5) \rightarrow (\mathbf{P}^5, \mathbf{P}^6)$. O primeiro instante de tempo em que ocorre a recuperação de $(\mathbf{P}^5, \mathbf{P}^6)$ é em $t = 1268$.

A Figura 5.17 mostra uma trajetória na qual a rede inicia a recuperação com \mathbf{P}^1 e alcança \mathbf{P}^6 sem passar por nenhum par armazenado intermediário entre $(\mathbf{P}^1, \mathbf{P}^2)$ e $(\mathbf{P}^5, \mathbf{P}^6)$. Isso significa que a C-BAM salta os passos intermediários da seqüência original de padrões associados memorizada. A seqüência recuperada para $k_r = 0.9972$, $k_f = 0.35$, $\rho = 9.0$ e $a = 4.0$ é $(\mathbf{P}^1, \mathbf{P}^2) \rightarrow (\mathbf{P}^4, \mathbf{P}^5) \rightarrow (\mathbf{P}^5, \mathbf{P}^6)$, em que $(\mathbf{P}^5, \mathbf{P}^6)$ aparece em $t = 488$.

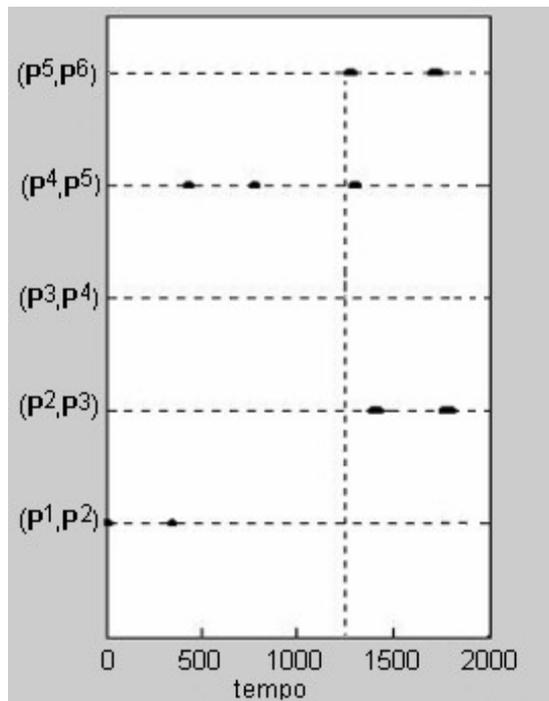


Figura 5.16. Seqüência de recuperação dos pares de padrões armazenados para $k_r=0.9973$, $k_f=0.35$, $\rho=9.0$ e $a=4.0$.

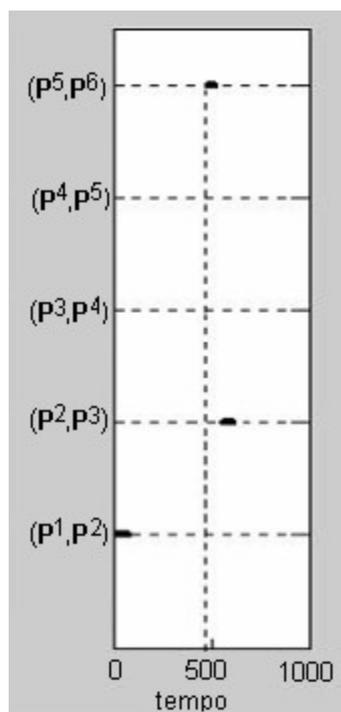


Figura 5.17. Seqüência de recuperação dos pares de padrões armazenados para $k_r=0.9972$, $k_f=0.35$, $\rho=9.0$ e $a=4.0$.

A modelagem do saltar passos foi também realizada para o exemplo de equação algébrica dado por Blessing e Anderson (1996): $-x-A=B$. Variando apenas o parâmetro k_r e permanecendo com $k_f=0.7$, $\rho=5.0$ e $a=3.0$ inalterados, foi verificada a ocorrência do saltar passos na modelagem da solução de $-x-A=B$, com a C-BAM saltando passos ao produzir trajetórias menores do que a trajetória completa de recuperação de \mathbf{P}^1 a \mathbf{P}^4 (Figura 5.10). As Figuras 5.18 e 5.19 ilustram este fenômeno.

A Figura 5.18 mostra uma trajetória na qual a rede inicia a recuperação com \mathbf{P}^1 e gera os padrões intermediários seguindo a seqüência correta de armazenamento. A seqüência de recuperação dos pares armazenados para $k_r=0.998$ é $(\mathbf{P}^1, \mathbf{P}^2) \rightarrow (\mathbf{P}^2, \mathbf{P}^3) \rightarrow (\mathbf{P}^3, \mathbf{P}^4)$. O primeiro instante de tempo em que ocorre a recuperação de $(\mathbf{P}^3, \mathbf{P}^4)$ é em $t=1302$.

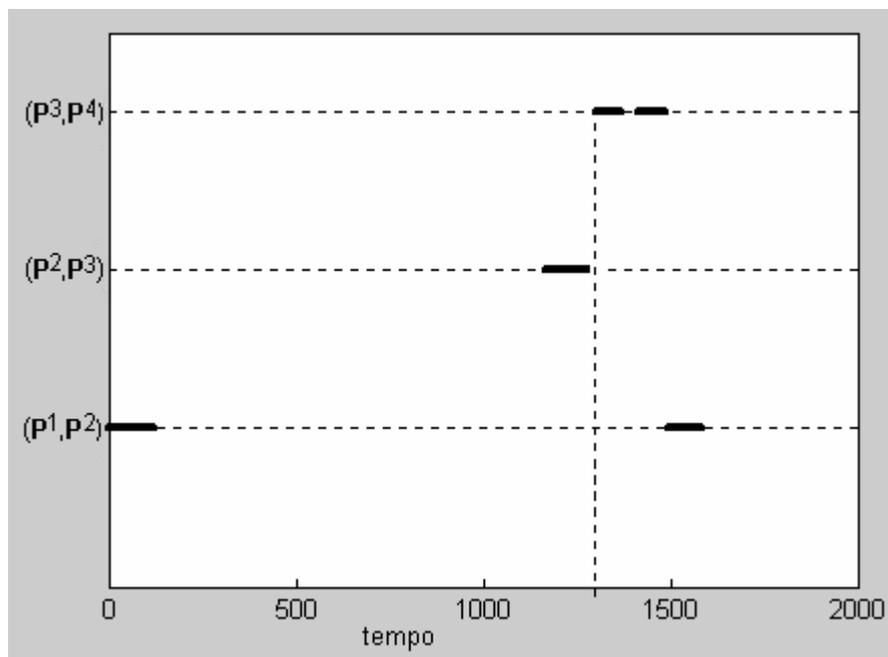


Figura 5.18. Seqüência de recuperação dos pares de padrões armazenados para $k_r = 0.998$, $k_f = 0.7$, $\rho = 5.0$ e $a = 3.0$.

A Figura 5.19 mostra uma trajetória na qual a rede inicia a recuperação com \mathbf{P}^1 e alcança \mathbf{P}^4 sem passar por nenhum par armazenado intermediário. Nesta situação, para $k_r=0.999$, a C-BAM salta o passo intermediário $(\mathbf{P}^2, \mathbf{P}^3)$ da seqüência original de padrões associados memorizada para alcançar imediatamente o passo final $(\mathbf{P}^3, \mathbf{P}^4)$ no tempo $t=832$.

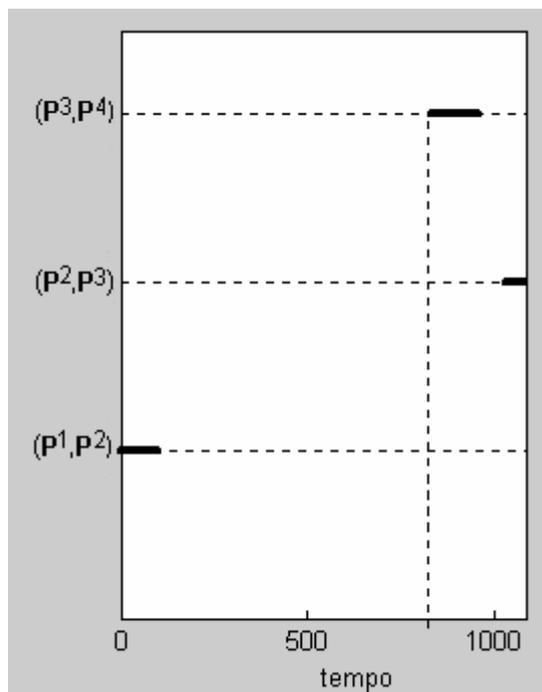


Figura 5.19. Seqüência de recuperação dos pares de padrões armazenados da Figura 5.10 para $k_r = 0.999$, $k_f = 0.7$, $\rho = 5.0$ e $a = 3.0$.

As Figuras 5.14 a 5.17 e 5.18 a 5.19 mostram que existem situações nas quais a habilidade de saltar passos está diretamente relacionada com o aumento ou o decréscimo do parâmetro k_r . Outros experimentos realizados apontaram casos em que a simples variação seqüencial de k_r não é suficiente para gerar os saltos obtidos na recuperação dos padrões armazenados. Isto ocorre porque a dinâmica da família C-BAM é muito complexa. Nestas figuras pode também ser notada a repetição de pares de padrões nas trajetórias entre (P^1, P^2) e (P^3, P^4) e entre (P^1, P^2) e (P^5, P^6) . Além disso, após a ocorrência do último padrão armazenado, costumam aparecer padrões que foram saltados na seqüência original de armazenamento. No entanto, o interesse no momento é o de analisar somente a capacidade da C-BAM de saltar passos intermediários até se alcançar o padrão final, desprezando a ocorrência de novas recuperações após o objetivo ser atingido.

Deve ser enfatizado que, no processo de modelar a habilidade de um noviço ou de um especialista na solução de um problema algébrico, a C-BAM não desenvolve uma percepção da equação matemática e, portanto, não “entende” os cálculos matemáticos ilustrados nos passos P^1 a P^{final} . Quando a rede salta um ou mais passos na solução da equação algébrica, isso significa que os parâmetros caóticos afetaram o modo de acessibilidade da memória e permitiram a recuperação dos padrões armazenados sem a ocorrência de alguns passos intermediários da seqüência original treinada, produzindo trajetórias menores do que a trajetória completa, na qual se passa por todos os

padrões treinados. O objetivo é mostrar um novo processo de acessibilidade de memória aplicado a um efeito psicológico, que é o fenômeno saltar passos.

5.5 Resumo

Os experimentos realizados com a família C-BAM mostraram comportamentos dinâmicos ricos e complexos para esta família de redes. Caos, bifurcações, crises, transição entre todos os padrões associados armazenados e recuperação de memórias antes não acessíveis na família BAM original foram os comportamentos destacados com os experimentos realizados para a família C-BAM. A Tabela 5.16 apresenta exemplos de variações paramétricas para as redes da família C-BAM em que ocorre cada um desses comportamentos citados.

Tabela 5.16. Resumo com os exemplos de comportamentos interessantes alcançados nos experimentos com a família C-BAM.

Modelo	Comportamento	Conjunto Paramétrico Utilizado				
		k_r	k_f	ρ	A	
<i>C-BAM</i>	período	0.3	0.15	8.0	2.0	
	caos	0.7	0.15	8.0	2.0	
	bifurcação	0.3 a 0.7	0.15	8.0	2.0	
	crise	0.7	0 a 1.0	8.0	2.0	
	transição entre pares armazenados	0.85	0.2	8.0	2.0	
	recuperação de memórias não acessíveis	0.95	0.45	7.0	3.0	
<i>C-BAM com atraso</i>	período	$\sigma = 3$	0.3	0.15	8.0	2.0
	caos		0.7	0.15	8.0	2.0
	bifurcação		0.3 a 0.7	0.15	8.0	2.0
	crise	$\sigma = 5$	0.65	0 a 1.0	9.0	2.0
	transição entre pares armazenados		0.8	0.15	8.0	2.0
	recuperação de memórias não acessíveis		0.9	0.45	8.0	2.0
<i>C-eBAM</i> ($\beta = 1.3$)	período	0.5	0.15	4.0	2.0	
	caos	0.9	0.15	4.0	2.0	
	bifurcação	0.5 a 0.9	0.15	4.0	2.0	
	crise	0.7	0 a 1.0	10.0	2.0	
	transição entre pares armazenados	0.9	0.2	10.0	2.0	
	recuperação de memórias não acessíveis	0.9	0.45	8.0	2.0	

Com os resultados obtidos através dos experimentos para verificação de comportamentos dinâmicos emergentes na família C-BAM, foi notado que as características de recuperação das C-BAM possibilitavam modelar o fenômeno psicológico do saltar passos. Por meio de mudanças

paramétricas no parâmetro k_r (enquanto os outros parâmetros permaneciam fixos), foi possível modelar a habilidade de novíços e de especialistas na resolução de um problema. Quando há a ocorrência do saltar passos (C-BAM representando a habilidade de um especialista), o desempenho, em termos de rapidez de recuperação do último padrão armazenado, é melhorado. A Tabela 5.17 resume as seqüências de recuperação apresentadas pelos conjuntos de padrões das Figuras 5.10 e 5.11.

Tabela 5.17. C-BAM para k_r variando e $k_f=0.70$, $\rho=5.0$ e $a=3.0$. (P^1 a P^4); $k_f=0.35$, $\rho=9.0$ e $a=4.0$ (P^1 a P^6).

Conjunto de Padrões	Habilidade	k_r	Sequencia de padrões recuperada	Instante de Recuperação de (P^5, P^6)
P^1 a P^4 (Figura 5.10)	Noviço	0.998	$(P^1, P^2) \rightarrow (P^2, P^3) \rightarrow (P^3, P^4)$	$t = 1302$
	Especialista	0.999	$(P^1, P^2) \rightarrow (P^3, P^4)$	$t = 832$
P^1 a P^6 (Figura 5.11)	Noviço	0.9975	$(P^1, P^2) \rightarrow (P^2, P^3) \rightarrow (P^3, P^4) \rightarrow (P^4, P^5) \rightarrow (P^5, P^6)$	$t = 2467$
		0.9974	$(P^1, P^2) \rightarrow (P^2, P^3) \rightarrow (P^4, P^5) \rightarrow (P^5, P^6)$	$t = 1665$
	Especialista	0.9973	$(P^1, P^2) \rightarrow (P^4, P^5) \rightarrow (P^5, P^6)$	$t = 1268$
		0.9972	$(P^1, P^2) \rightarrow (P^5, P^6)$	$t = 488$

Pode ser notado nos experimentos apresentados neste capítulo que o processo de recuperação de padrões armazenados pela família de modelos de memória associativa bidirecional caótica não se estabiliza em um padrão armazenado específico por causa da dinâmica caótica dessa família de redes. De fato, foi verificado, por experimentação, que as redes de memória associativa bidirecionais caóticas, mediante uma escolha apropriada de valores de parâmetros, são capazes de transitar por todos os padrões armazenados, sem, contudo, se estabilizarem em nenhum desses padrões. O próximo capítulo apresenta a proposição de estratégias de controle de caos para a família C-BAM considerando-se dois focos: (a) foco na recuperação de memórias inacessíveis e (b) foco no fenômeno saltar passos.

Capítulo 6

Controle de Caos na Família C-BAM

No capítulo anterior o presente trabalho concentrou-se no estudo das dinâmicas da família C-BAM e na aplicabilidade dessas dinâmicas em seu estado natural, tal como elas emergem das redes C-BAM. No entanto, durante os estudos com as redes C-BAM, percebeu-se que seria muito interessante a inclusão de um tipo de estratégia de controle de caos capaz de lidar com o novo modo de acesso a memórias armazenadas gerado por estas redes, de modo a não deixar esta acessibilidade à memória da família C-BAM como um processo pseudo-aleatório, isto é, decorrente apenas da evolução da dinâmica caótica. Assim, através do controle de caos, a dinâmica das redes associativas bidirecionais caóticas realmente concretiza a acessibilidade à memória, convergindo para o padrão anteriormente não recuperável. Portanto, o objetivo deste capítulo é propor e implementar estratégias de controle de caos para o novo modo de acessibilidade a memórias armazenadas gerado pela família C-BAM. Dentro deste objetivo, o controle de caos foi proposto considerando-se dois casos distintos: (a) recuperação de memórias inacessíveis e (b) fenômeno saltar passos.

Vale enfatizar que o novo modo de acessibilidade a memórias armazenadas, emergente na família C-BAM, é uma consequência direta da dinâmica caótica destas redes e significa que, ao invés de as redes C-BAMs recuperarem um único par de padrão, como as redes BAMs originais fazem, elas apresentam complexas transições entre os padrões memorizados, para alguns conjuntos de valores de parâmetros. A grande diversidade de padrões recuperados permite que um padrão antes inacessível ou não recuperável nas redes BAM originais, definido como um padrão \mathbf{Y}^j , $k=1, \dots, j, \dots, p$, que não é recuperado a partir de nenhum padrão de entrada \mathbf{X}^k , possa ser recuperado pela família C-BAM. A nova acessibilidade a memórias também permite que a rede caótica excursionasse por todos os padrões \mathbf{P}^k ($k=1, \dots, p$) do fenômeno saltar passos e, assim, modele a

habilidade de saltar passos por um indivíduo na solução de uma tarefa, enquanto que, na BAM original, a entrada \mathbf{P}^k é capaz de gerar apenas um próximo passo \mathbf{P}^{k+1} na solução da tarefa.

A Seção 6.1 apresenta o controle de caos aplicado na recuperação de memórias inacessíveis. Através do controle de caos, realizado com base no método de controle por pinagem (*pinning control method*), a família C-BAM não apenas excursiona passando pelo padrão inacessível como uma consequência direta de sua dinâmica caótica, mas as redes C-BAM tornam-se aptas a convergir para memórias acessíveis ou inacessíveis desejadas, completando com sucesso o processo de recuperação.

A Seção 6.2 apresenta o controle de caos aplicado ao fenômeno saltar passos. Pelo fato de a modelagem deste fenômeno ter sido proposta apenas para a rede C-BAM, o controle de caos também será realizado apenas para esta rede. Duas estratégias de controle de caos são propostas para o fenômeno saltar passos: o método de controle por pinagem e o método de controle adaptativo. Tais estratégias permitem que a rede C-BAM, utilizada para a modelagem do saltar passos, estabilize-se no padrão final, o qual representa a solução da operação algébrica memorizada pela rede, diminuindo o número de estados intermediários necessários para chegar ao estado final desejado.

6.1 Controle de Caos para a Recuperação de Memórias Inacessíveis

Na Seção 5.3 mostrou-se como as dinâmicas caóticas permitem à família C-BAM acessar padrões armazenados que os modelos originais não podem fazê-lo, ou seja, mostrou-se como a própria dinâmica caótica da família C-BAM conduz ao acesso a memórias inacessíveis pela dinâmica não-caótica. Embora os resultados impliquem na possibilidade de se evitar, na família C-BAM, o problema de memórias inacessíveis como uma consequência direta da evolução caótica deste modelo de redes, o próprio movimento caótico destas redes impossibilita que o processo de busca de memória seja estabilizado em um padrão desejado. A rede pode recuperar os padrões armazenados, mas eles aparecem não periodicamente, uma vez que a rede está em caos. Esta situação impossibilita a recuperação do padrão inacessível desejado, isto é, a saída passa por ele mas não permanece nele. Assim, é fundamental deter o estado da rede na situação desejada através de um mecanismo de controle.

A presente seção tem como objetivo propor uma estratégia de controle de caos que permite à família C-BAM suprimir sua dinâmica caótica e estabilizar-se em uma memória inacessível desejada utilizando-se apenas da informação de entrada, sem qualquer informação prévia a respeito da saída desejada (que será uma memória acessível ou inacessível). Utilizando o método de controle de caos que será aqui proposto, será possível verificar que as redes C-BAM podem, de

modo eficaz, realizar a hetero-associação correta de memórias antes inacessíveis, ou seja, as redes C-BAM podem realizar a recuperação dos pares associados tal como foram armazenados em sua aprendizagem.

Com base no método de controle por pinagem de He *et al.* (2003), visto na Seção 3.2, é proposta uma estratégia de controle de caos para a família C-BAM que utiliza apenas a informação de entrada como referência e que permite hetero-associações bem sucedidas para todos os padrões armazenados, incluindo as memórias inacessíveis dos modelos originais.

6.1.1 Controle de Caos para a Rede C-BAM

Com o objetivo de controlar a dinâmica da rede C-BAM a partir do conhecimento apenas da informação de entrada da rede, foi aplicado o controle de caos somente na Camada Y desta rede. Conforme ilustrado na Figura 6.1, o controle de realimentação é aplicado em um estado interno (η) da rede C-BAM.

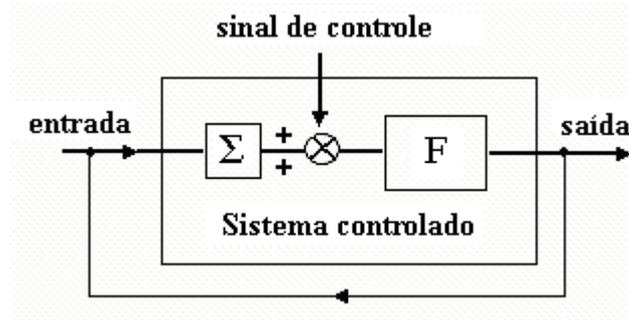


Figura 6.1. Esquema do método de controle aplicada na camada Y da rede C-BAM.

Assim sendo, a rede C-BAM com controle, capaz de estabilizar-se na saída desejada, é descrita da seguinte maneira:

$$x_i(t+1) = f(\eta_i(t+1) + \zeta_i(t+1)), \quad (6.1)$$

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^n w_{ij} y_j(t) \quad (6.2)$$

$$\zeta_i(t+1) = k_r \zeta_i(t) - \rho x_i(t) + a \quad (6.3)$$

$$y_j(t+1) = f(\eta_j(t+1) + \zeta_j(t+1)) \quad (6.4)$$

$$\eta_j(t+1) = k_f \eta_j(t) + \sum_{i=1}^m w_{ji} \{x_i(t) + Ku_i(t)\} \quad (6.5)$$

$$u_i(t) = x_i(t) - (1 - \tilde{x}_i) \quad (6.6)$$

$$\zeta_j(t+1) = k_r \zeta_j(t) - \rho y_j(t) + a \quad (6.7)$$

em que u_i é o sinal de controle, o qual é construído ao se comparar a saída real da camada X ($x_i(t)$) e a saída desejada para esta mesma camada (\tilde{x}_i); e K é a intensidade do controle aplicado ao u_j .

Simulações computacionais comprovam que esta estratégia de controle de caos proposta para a rede C-BAM é capaz de estabilizar esta rede em um padrão armazenado específico, e de permitir a ocorrência de hetero-associações corretas para padrões armazenados antes inacessíveis. Por exemplo, foi visto na Tabela 5.10 que a memória armazenada \mathbf{Y}^3 é uma memória inacessível para a rede BAM convencional. A Figura 6.2 mostra que a C-BAM sem caos ($k_r = k_f = \rho = a = 0$), tal como a BAM original, converge para o ponto fixo \mathbf{Y}^1 desde o instante $t=1$.

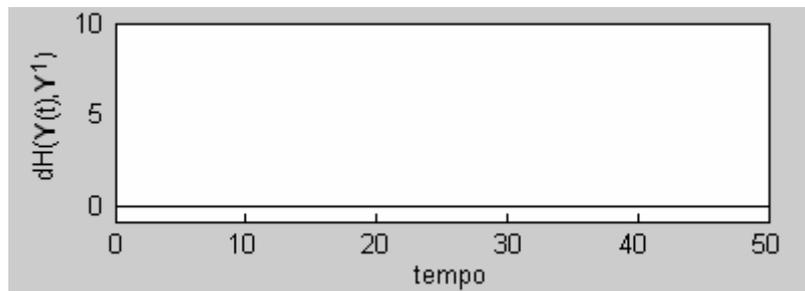


Figura 6.2. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado \mathbf{Y}^1 , no intervalo de tempo $[0, 50]$, para $k_r = k_f = \rho = a = 0$ e $\mathbf{X}(t=0) = \mathbf{X}^3$. Pode ser aqui verificado que, a partir do padrão inicial \mathbf{X}^3 , a C-BAM sem caos não consegue recuperar o par associado \mathbf{Y}^3 . Ao invés de \mathbf{Y}^3 a rede recupera \mathbf{Y}^1 .

Considerando-se \mathbf{X}^3 como o estado inicial da rede e fazendo-se $k_r = 0.95$, $k_f = 0.25$, $\rho = 10.0$ e $a = 4.0$, verificou-se que a rede C-BAM sem sinal de controle está em caos e os padrões armazenados (Figura 5.9) são ativados em meio à dinâmica caótica. A Figura 6.3 mostra a distância Hamming entre a saída da rede C-BAM sem a estratégia de controle e o padrão alvo (\mathbf{Y}^3) no intervalo de tempo $[0, 3000]$. Através da Figura 6.3 pode-se facilmente observar que, embora a C-BAM sem a estratégia de controle atinja o padrão desejado antes inacessível, \mathbf{Y}^3 , a partir de seu par associado \mathbf{X}^3 , em alguns instantes de tempo (quando a distância Hamming entre a saída da rede e \mathbf{Y}^3 é zero), esta rede não se estabiliza no padrão \mathbf{Y}^3 .

A Figura 6.4 ilustra a distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado \mathbf{Y}^3 quando a estratégia de controle de caos é adicionada a C-BAM. Esta Figura 6.4 mostra que, quando o controle proposto nas Equações 6.1 a 6.7 é adicionado ao sistema, a rede C-BAM com controle de caos estabiliza-se no padrão desejado \mathbf{Y}^3 , uma vez que a distância Hamming entre a saída $\mathbf{Y}(t)$ e o padrão \mathbf{Y}^3 é zero a partir de $t=10$. Com a inclusão da estratégia de controle de caos na rede C-BAM, esta rede deixa de simplesmente passar pelo padrão antes inacessível, para rapidamente estabilizar-se neste padrão. O par associado ($\mathbf{X}^3, \mathbf{Y}^3$) armazenado agora é corretamente recuperado.

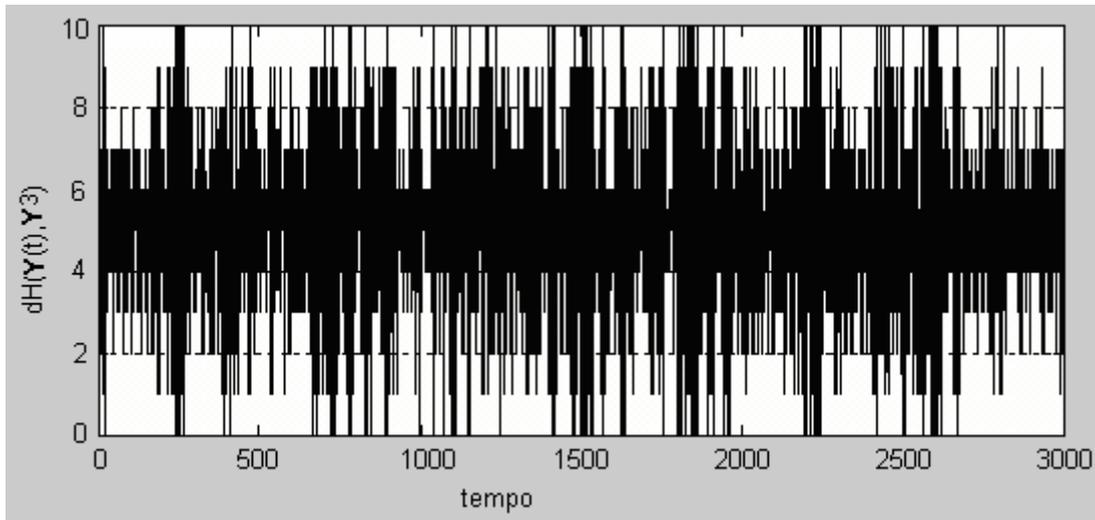


Figura 6.3. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado \mathbf{Y}^3 , de $t = 0$ a 3000, para $k_r = 0.95$, $k_f = 0.25$, $\rho = 10$, $a = 4.0$ e $\mathbf{X}(t=0) = \mathbf{X}^3$.

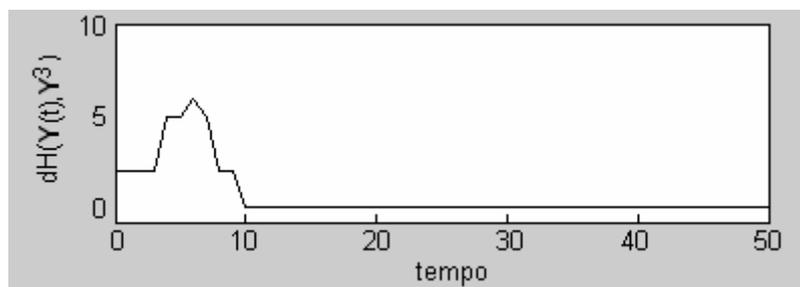


Figura 6.4. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado \mathbf{Y}^3 , de $t = 0$ a 50, para $k_r = 0.95$, $k_f = 0.25$, $\rho = 10.0$, $a = 4.0$, $K = 20$ e $\mathbf{X}(t=0) = \mathbf{X}^3$.

Deve ser notado que a hetero-associação correta (\mathbf{X}^3 , \mathbf{Y}^3) é obtida sem qualquer informação sobre a saída desejada (\mathbf{Y}^3). Enquanto a rede BAM convencional recupera o par (\mathbf{X}^3 , \mathbf{Y}^1) quando \mathbf{X}^3 é o padrão inicial de entrada da rede, na rede C-BAM com controle, o par recuperado para o padrão inicial \mathbf{X}^3 é o par associado armazenado \mathbf{Y}^3 , sendo que o controle é realizado apenas com a informação do próprio padrão \mathbf{X}^3 , que é o par associado correspondente a \mathbf{Y}^3 .

Embora a Seção 6.1 dedique atenção especial ao caso de memórias inacessíveis, é importante destacar que o controle de caos, feito pelo método de controle de pinagem, também é eficaz para controlar as memórias acessíveis de toda a família C-BAM. A seguir será mostrado um exemplo de controle de memória acessível.

A Tabela 5.10 mostra que a C-BAM sem caos faz a hetero-associação correta do par associado (\mathbf{X}^1 , \mathbf{Y}^1). No entanto, em sua dinâmica caótica, a C-BAM com os valores de parâmetros $k_r = 0.95$, $k_f = 0.45$, $\rho = 10.0$ e $a = 2.0$, passa apenas a excursionar pelo padrão \mathbf{Y}^1 , sem convergir para este padrão (Figura 6.5).

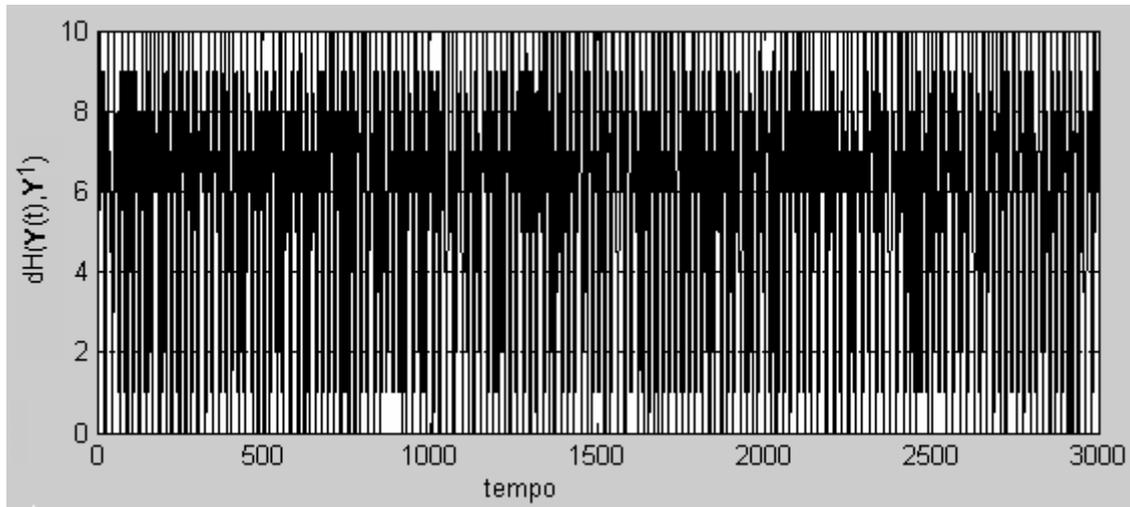


Figura 6.5. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado \mathbf{Y}^1 , de $t = 0$ a 3000, para $k_r=0.95$, $k_f=0.45$, $\rho=10.0$, $a=2.0$ e $\mathbf{X}(t=0)=\mathbf{X}^1$.

A aplicação do método de controle por pinagem, na situação de caos ilustrada na Figura 6.5, permite à C-BAM convergir para o padrão \mathbf{Y}^1 e fazer a hetero-associação correta ($\mathbf{X}^1, \mathbf{Y}^1$), desde o instante inicial, para o padrão acessível \mathbf{X}^1 (Figura 6.6).

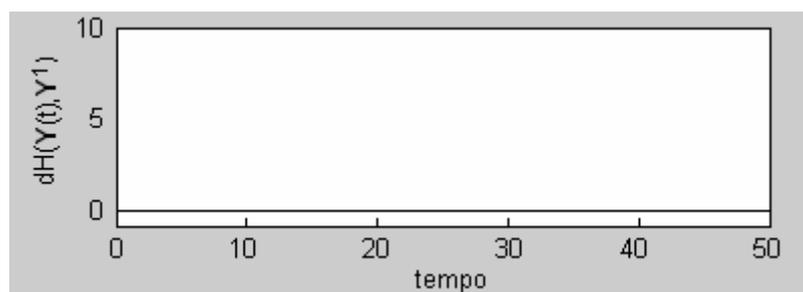


Figura 6.6. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado \mathbf{Y}^1 , de $t = 0$ a 50, para $k_r=0.95$, $k_f=0.45$, $\rho=10.0$, $a=2.0$, $K=20$ e $\mathbf{X}(t=0)=\mathbf{X}^1$.

6.1.2 Controle de Caos para a Rede C-BAM com atraso

Também foi proposta uma estratégia de controle de caos para a rede C-BAM com atraso que segue o esquema da Figura 6.1. O controle de caos proposto para a C-BAM com atraso, o qual utiliza apenas a informação de entrada como referência, levou-a a convergência a um padrão específico, atingindo hetero-associações corretas para os padrões armazenados inacessíveis no modelo convencional da BAM com atraso. A fim de controlar a dinâmica da rede C-BAM com atraso a partir do conhecimento da informação de entrada, o controle de caos é aplicado somente na

camada Y deste modelo. Sendo assim, a C-BAM com atraso será estabilizada na saída desejada ao se substituir as Equações 4.24 a 4.29, da C-BAM com atraso original, pelas seguintes equações:

$$x_i(t+1) = f(\eta_i(t+1) + \zeta_i(t+1)), \quad (6.8)$$

$$\eta_i(t+1) = k_f \eta_i(t-\sigma) + \sum_{j=1}^n w_{ij} y_j(t-\sigma) \quad (6.9)$$

$$\zeta_i(t+1) = k_r \zeta_i(t-\sigma) - \rho x_i(t-\sigma) + a \quad (6.10)$$

$$y_j(t+1) = f(\eta_j(t+1) + \zeta_j(t+1)) \quad (6.11)$$

$$\eta_j(t+1) = k_f \eta_j(t-\sigma) + \sum_{i=1}^m w_{ji} \{x_i(t-\sigma) + Ku_i(t-\sigma)\} \quad (6.12)$$

$$u_i(t-\sigma) = x_i(t-\sigma) - (1 - \tilde{x}_i) \quad (6.13)$$

$$\zeta_j(t+1) = k_r \zeta_j(t-\sigma) - \rho y_j(t-\sigma) + a \quad (6.14)$$

Simulações computacionais foram realizadas utilizando-se o conjunto paramétrico $k_r=0.96$, $k_f=0.45$, $\rho=9.0$, $a=2.0$, $\sigma=5$, e o padrão armazenado \mathbf{X}^3 como estado inicial da rede. A Figura 6.7 ilustra a distância Hamming entre as saídas da rede C-BAM com atraso ($\mathbf{Y}(t)$) e a saída desejada \mathbf{Y}^3 , antes do controle de caos, durante o intervalo de tempo $[0,3000]$. Através desta figura pode ser verificado que, em sua dinâmica original, a C-BAM com atraso apenas passa pelo padrão alvo \mathbf{Y}^3 .

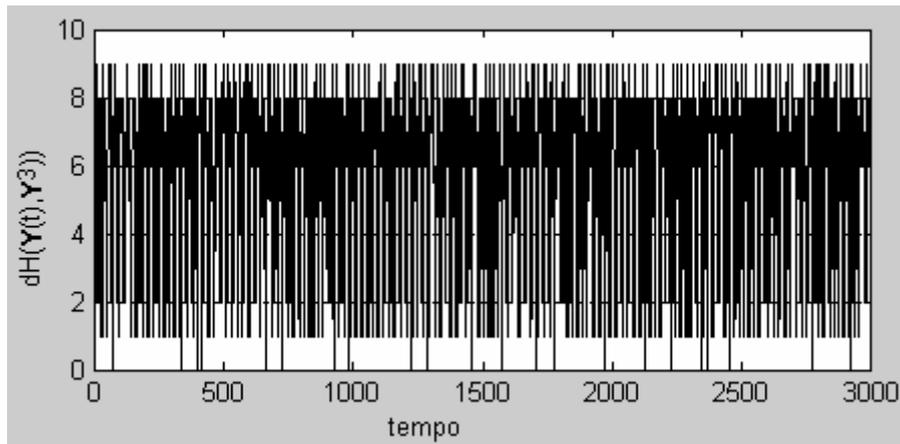


Figura 6.7. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM com atraso e o padrão armazenado \mathbf{Y}^3 , de $t=0$ a 3000, para $k_r=0.96$, $k_f=0.45$, $\rho=9.0$, $a=2.0$, $\sigma=5$ e $\mathbf{X}(t=0)=\mathbf{X}^3$.

Quando a estratégia de controle de caos é adicionada à C-BAM com atraso de acordo com as Equações 6.8 a 6.14, é possível verificar que esta rede, de fato, pode ser estabilizada em um padrão armazenado específico e a hetero-associação pode ocorrer de forma apropriada. Por exemplo, calculando-se a distância Hamming para a C-BAM com atraso para $k_r=0.96$, $k_f=0.45$, $\rho=9.0$, $a=2.0$, $\sigma=5$, $K=10$ e \mathbf{X}^3 como padrão inicial, foi obtido que a rede se estabiliza no padrão armazenado \mathbf{Y}^3 ,

um exemplo de memória inacessível para a BAM com atraso convencional (Tabela 5.10), a partir de $t=16$. A Figura 6.8 ilustra esta ocorrência.

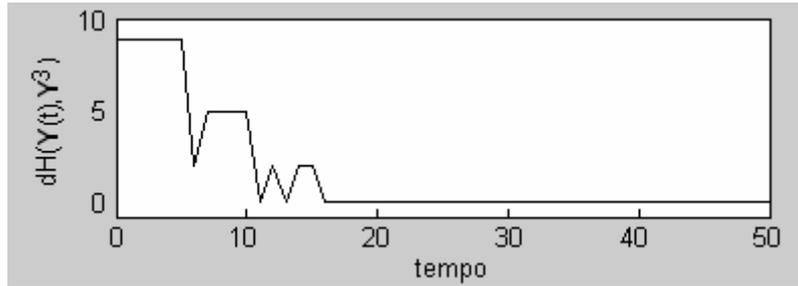


Figura 6.8. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM com atraso e o padrão armazenado \mathbf{Y}^3 , de $t = 0$ a 50, para $k_r = 0.96$, $k_f = 0.45$, $\rho=9.0$, $a=2.0$, $\sigma=5$, $K=10$ e $\mathbf{X}(t=0)=\mathbf{X}^3$.

6.1.3 Controle de Caos para a Rede C-eBAM

A estratégia de controle de caos esboçada na Figura 6.1 também foi aplicada à rede C-eBAM. A fim de controlar a dinâmica da C-eBAM apenas com o conhecimento da informação de entrada, o controle de caos foi aplicado apenas a sua camada X. Deste modo, a C-eBAM com controle, capaz de estabilizar-se em um padrão de saída desejado, é descrita da seguinte maneira:

$$x_i(t+1) = f(\eta_i(t+1) + \zeta_i(t+1)) \quad (6.15)$$

$$\eta_i(t+1) = k_f \eta_i(t) + \left\{ \sum_{k=1}^p X_i^k \left(\beta \left\langle \mathbf{Y}^k, \mathbf{Y}(t) \right\rangle * K u_i(t) \right) \right\} \quad (6.16)$$

$$u_i(t) = x_i(t) - \tilde{x}_i(t) \quad (6.17)$$

$$\zeta_i(t+1) = k_r \zeta_i(t) - \rho x_i(t) + a \quad (6.18)$$

$$y_j(t+1) = f(\eta_j(t+1) + \zeta_j(t+1)) \quad (6.19)$$

$$\eta_j(t+1) = k_f \eta_j(t) + \left\{ \sum_{k=1}^p Y_j^k \beta \left\langle \mathbf{X}^k, \mathbf{X}(t) \right\rangle \right\} \quad (6.20)$$

$$\zeta_j(t+1) = k_r \zeta_j(t) - \rho y_j(t) + a \quad (6.21)$$

Antes de adicionar o controle de caos à C-eBAM, esta rede é capaz de transitar pelos padrões armazenados, sem, contudo, estabilizar-se em um padrão desejado. A Figura 6.9 ilustra esta situação através da distância Hamming entre a saída da rede na camada Y e o padrão armazenado

\mathbf{Y}^3 , que é uma memória inacessível para a eBAM convencional (Tabela 5.10), quando $k_r=0.9$, $k_f=0.45$, $\rho=8.0$, $a=2.0$, e $\beta=1.3$, e \mathbf{X}^3 é o estado inicial da rede.

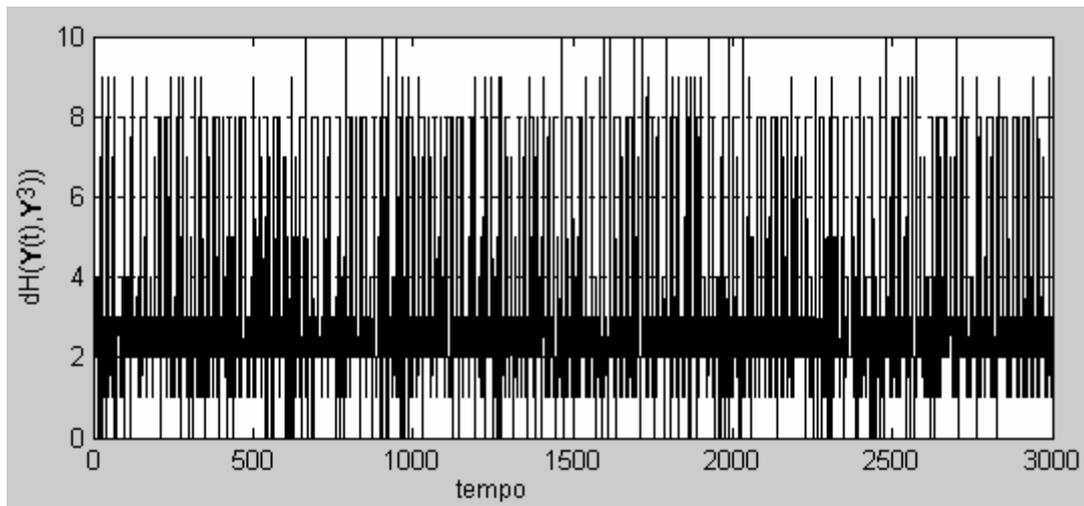


Figura 6.9. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-eBAM e o padrão armazenado \mathbf{Y}^3 , de $t=0$ a 3000, para $k_r=0.9$, $k_f=0.45$, $\rho=8.0$, $a=2.0$, $\beta=1.3$ e $\mathbf{X}(t=0)=\mathbf{X}^3$.

Os resultados obtidos através de simulações computacionais envolvendo a rede C-eBAM com controle comprovaram que o modelo proposto é capaz de estabilizar-se em um padrão armazenado desejado. Com isso, a hetero-associação envolvendo padrões antes inacessíveis pode acontecer de forma correta. A distância Hamming entre a saída da rede na camada Y e o padrão armazenado \mathbf{Y}^3 , para o mesmo conjunto paramétrico e o padrão inicial utilizado na Figura 6.9, e com $K=10$, mostra que a C-eBAM com controle estabiliza-se rapidamente ($t=6$) no padrão \mathbf{Y}^3 (Figura 6.10).

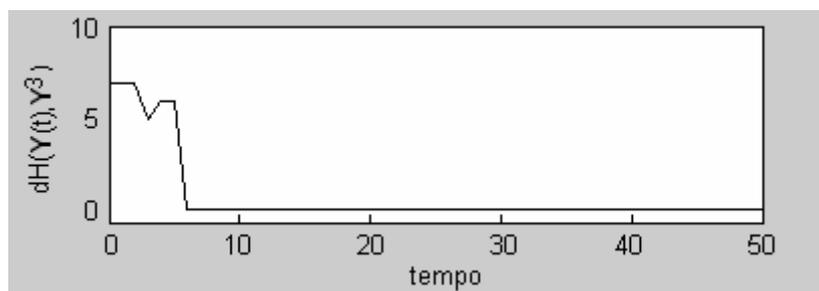


Figura 6.10. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-eBAM e o padrão armazenado \mathbf{Y}^3 , de $t=0$ a 50, para $k_r=0.9$, $k_f=0.45$, $\rho=a=0$, $\beta=1.3$, $K=10$ e $\mathbf{X}(t=0)=\mathbf{X}^3$.

6.1.4 Discussão da Estratégia de Controle de Caos Proposta para a Recuperação de Memórias Inacessíveis na Família C-BAM

A estratégia de controle de caos proposta nesta seção está baseada no método de controle por pinagem (Hu e Qu, 1994), (He *et al.*, 2003). No método de controle por pinagem original apenas uma parte dos neurônios da rede são controlados e o efeito de controle é estendido para toda a rede neural. No caso do exemplo de recuperação de memórias inacessíveis pela família C-BAM, apresentado na Seção 5.3, o conjunto de padrões (Figura 5.9) utilizado é pequeno e os pares de padrões são muito similares. Como consequência, a idéia de aplicar o controle apenas em uma parte dos neurônios da rede, tal como no método de controle por pinagem original, não foi bem sucedida. No entanto, quando o controle foi adicionado a todos os neurônios de uma determinada camada das redes C-BAM, estas redes foram estabilizadas no padrão desejado. Deste modo, a estratégia de controle de caos com foco na recuperação de memórias inacessíveis proposta neste trabalho, que engloba conjunto de padrões pequenos e similares, considera simplesmente a intensidade de controle (K) e o sinal de controle (u) do método de controle por pinagem original, ambos aplicados em todos os neurônios de uma camada da rede, como parâmetros para tornarem a família C-BAM controlável.

O valor da intensidade de controle (K) é importante para a ocorrência do controle de caos. Para uma distância de pinagem fixa, o controle alvo pode ser alcançado apenas quando a intensidade de controle é maior do que um limiar. Se a intensidade de controle é menor do que o limiar, o sistema permanece em caos. Os experimentos simulados mostraram que um valor de $K \geq 10$ pode, de fato, controlar as redes da família C-BAM para o caso das memórias inacessíveis em questão.

É importante lembrar que os métodos de controle por pinagem encontrados na literatura estão focados em redes de uma única camada. Os modelos C-BAM propostos neste trabalho constituem-se redes de duas camadas, portanto, a arquitetura da família C-BAM é diferente da rede controlada por Hu *et al.* (2003), na qual qualquer neurônio da rede está ligado a todos os outros neurônios da rede. A idéia da estratégia de controle aqui proposta é a de estabilizar as redes C-BAM em um padrão desejado ao aplicar controle de caos apenas em uma das camadas do modelo. A escolha da camada na qual será aplicado o controle, em cada modelo específico da família C-BAM, é determinada pela informação de entrada, ou seja, o controle de caos será aplicado na camada que lida com o padrão de entrada da rede, sem que se inclua no controle qualquer informação prévia a respeito do padrão de saída desejado.

Quando se projetou o controle de caos para a família C-BAM, cada modelo desta família recebeu a informação de entrada em uma camada específica. Desta forma, em C-BAM e C-BAM com atraso, o sinal de controle (u) foi aplicado na camada Y, enquanto o sinal de controle foi

utilizado na camada X para a C-eBAM. Nas equações da família C-BAM com controle é possível notar que o controle é específico para o modelo e, ao mesmo tempo, é compatível com o modelo original.

A estratégia de controle de caos proposta para a família C-BAM utiliza apenas a entrada associada à saída desejada como referência para se alcançar a estabilidade das redes neurais. Por outro lado, no método de controle por pinagem proposto He *et al.* (2003), a resposta final é empregada como referência, ou seja, o estado de referência é o próprio padrão onde se quer que a rede seja estabilizada. Além disto, as simulações de He *et al.* (2003) mostram que a rede estabiliza no próprio padrão alvo (\mathbf{X}), se o estado de todo neurônio do padrão alvo for substituído por $(1 - \tilde{x}_i)$ na variável de controle. Foi determinado por experimentação que isto também é verdade para a C-BAM e para a C-BAM com atraso, que tratam com a saída da rede. No entanto, no caso da C-eBAM, em que o controle é feito em cima do próprio conjunto de padrões armazenados, os experimentos mostraram que esta substituição não é necessária.

6.2 Controle de Caos para o Fenômeno Saltar Passos

Foi visto, na Seção 5.4, como a transição entre os padrões armazenados da C-BAM pode modelar o fenômeno saltar passos através da memorização de estágios de operações algébricas. Nesta modelagem, o objetivo é estudar a trajetória da recuperação dos padrões armazenados pela C-BAM. A Seção 5.4.2 mostra que a C-BAM é capaz de, a partir de um padrão inicial \mathbf{P}^1 , representando o passo inicial na solução de uma operação algébrica, saltar padrões da seqüência original de armazenamento e recuperar o padrão final que representa a solução da operação algébrica. No entanto, a rede C-BAM parte do padrão inicial \mathbf{P}^1 e, após alcançar o padrão final na solução da equação algébrica, ela apenas passa por este último padrão sem, contudo, efetivamente convergir para ele. Tal como ocorre no caso da recuperação de memórias inacessíveis pela família C-BAM, a rede C-BAM é capaz de recuperar o padrão que representa a solução de uma equação algébrica, mas, a partir deste ponto, outros padrões armazenados e padrões espúrios podem reaparecer não periodicamente, uma vez que a rede está em caos.

As Tabelas 6.1 e 6.2 mostram a recuperação de padrões da rede C-BAM sem caos, ou seja, a C-BAM com os valores de parâmetros iguais a zero, para os padrões algébricos das Figuras 5.10 e 5.11. Essas tabelas mostram que a rede C-BAM sem caos (BAM) não foi capaz de saltar passos. Dado \mathbf{P}^k como padrão inicial, a BAM converge para \mathbf{P}^{k+1} ou para um padrão espúrio.

Tabela 6.1. Relação entre as entradas e as saídas das redes da família C-BAM para $k_r = k_f = \rho = a = 0$ considerando-se os padrões algébricos da Figura 5.11.

Entrada \ Saída	P^1	P^2	P^3	P^4	P^5
C-BAM	P^2	P^3	P^4	padrão espúrio	padrão espúrio

Tabela 6.2. Relação entre as entradas e as saídas das redes da família C-BAM para $k_r = k_f = \rho = a = 0$ considerando-se os padrões algébricos da Figura 5.10.

Entrada \ Saída	P^1	P^2	P^3
C-BAM	P^2	P^3	P^4

As Figuras 6.11 e 6.12 mostram a rede C-BAM funcionando caoticamente, ou seja, com valores de parâmetros que conduzem ao caos. Pode ser verificado que, dado o estado inicial P^1 , a rede é capaz de saltar passos e alcançar, respectivamente, P^4 e P^6 , os estados finais de conjuntos de padrões algébricos (Figuras 5.10 e 5.11). Conforme visto na Seção 5.4, com o saltar passos há uma diminuição do número de estados intermediários para atingir o estado final. Enquanto a recuperação de P^4 , para a C-BAM modelando a habilidade de um noviço com os padrões da Figura 5.10, ocorre no instante tempo $t=1302$; com o saltar passos (Figura 6.11), a recuperação de (P^3, P^4) se dá em $t=832$. No caso da C-BAM modelando a habilidade de um noviço com os padrões da Figura 5.11, P^6 é recuperado no instante $t=2467$, sendo que, com o saltar passos (Figura 6.12), recupera-se P^6 em $t=488$. Apesar de, com a C-BAM funcionando caoticamente, ser possível modelar esta diminuição do número de estados intermediários para se atingir o estado final, a rede C-BAM não converge para os estados finais desejados, escapando deles logo após alcançá-los.

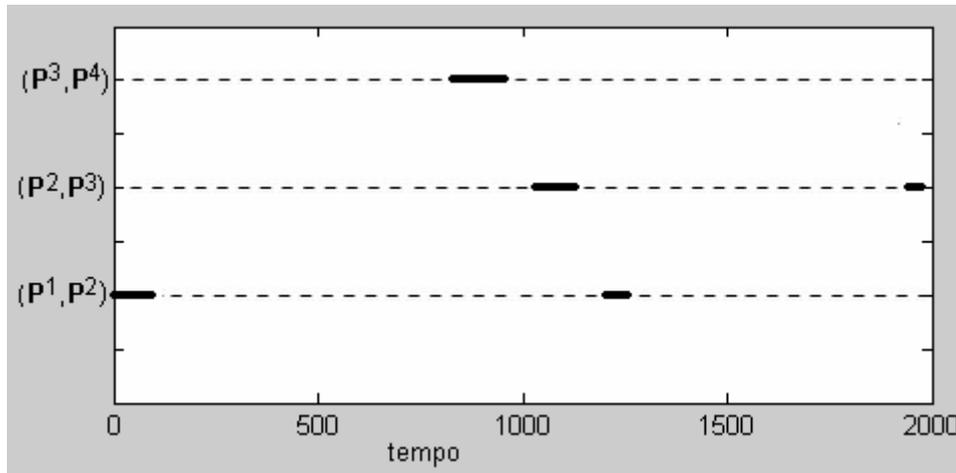


Figura 6.11. Seqüência de recuperação dos pares de padrões armazenados na C-BAM para o conjunto de padrões da Figura 5.10 e $k_r=0.999$, $k_f=0.7$, $\rho=5.0$, $a=3.0$. Pode ser notado que, após recuperar o par associado contendo o padrão final \mathbf{P}^4 , no instante $t=832$, a rede continua transitando por outros pares armazenados, sem se estabilizar no padrão desejado.

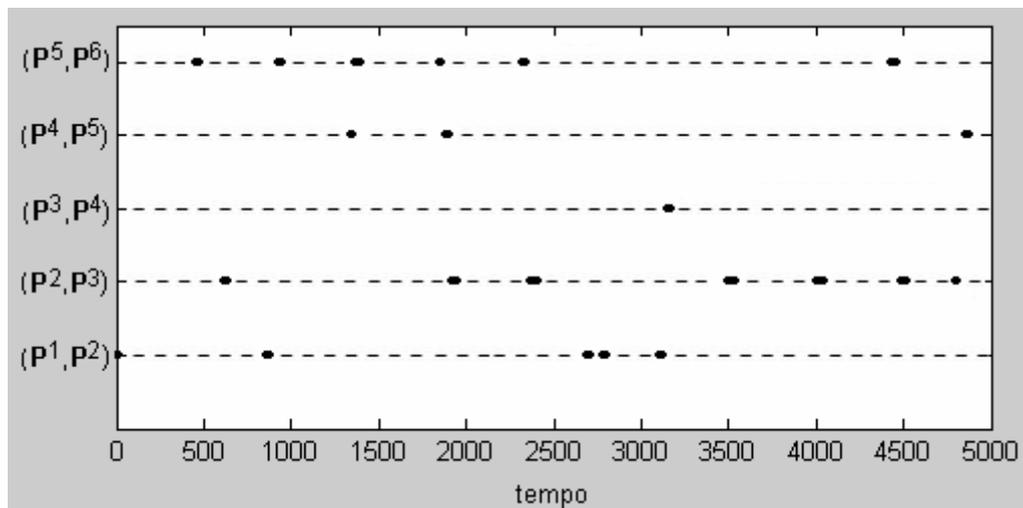


Figura 6.12. Seqüência de recuperação dos pares de padrões armazenados na C-BAM para o conjunto de padrões da Figura 5.11 e $k_r=0.9972$, $k_f=0.35$, $\rho=9.0$, $a=4.0$. Pode ser notado que, após recuperar o par associado contendo o padrão final \mathbf{P}^6 , no instante $t=488$, a rede continua transitando por outros pares armazenados, sem se estabilizar no padrão desejado.

A dinâmica caótica permite a recuperação de todos os padrões armazenados (passos na resolução da equação da solução), mas impossibilita que o processo de recuperação da C-BAM se estabilize no padrão final desejado, levando a não distinção entre a recuperação do padrão final (a solução da equação algébrica) e a recuperação de um outro padrão qualquer. Sendo assim, decidiu-se incluir, neste trabalho, uma estratégia de controle de caos na rede C-BAM, com foco no fenômeno saltar passos, que possibilite a esta rede suprimir sua dinâmica caótica e estabilizar-se no

padrão final da seqüência de estados. Tal estratégia de controle de caos permite caracterização fidedigna da recuperação da memória desejada, pois a rede passa a alcançar e a estabilizar-se nesta determinada memória. Portanto, o objetivo para a inclusão de um modelo caótico controlável para o fenômeno saltar passos é o de, a partir de do padrão inicial \mathbf{P}^1 , que representa o passo ou estado inicial da solução do problema, alcançar-se o padrão final e estabilizar-se neste padrão.

Serão abordadas duas estratégias de controle de caos para a rede C-BAM utilizada no fenômeno saltar passos: uma baseada no método de controle por pinagem (Hu e Qu, 1994), (He *et al.*, 2003), e outra baseada na técnica de controle adaptativo (Astrom e Wittenmark, 1995), (Andrievskii e Fradkov, 2003), (Fradkov e Evans, 2005).

6.2.1 Método de Controle por Pinagem

Uma das estratégias de controle de caos proposta na modelagem C-BAM para o fenômeno saltar passos está baseada no método de controle por pinagem (Hu e Qu, 1994), (He *et al.*, 2003). Ao contrário do caso da recuperação de memórias inacessíveis, a dimensionalidade dos padrões a serem armazenados pela rede C-BAM na modelagem do saltar passos é alta, ou seja, cada padrão do conjunto de padrões a ser armazenado é composto por um número alto de bits. Como consequência, a idéia original do método de controle por pinagem, que é a de aplicar o controle de caos apenas para uma parte dos neurônios da rede, de modo que o efeito de controle seja estendido para toda a rede a partir de alguns neurônios controláveis, foi retomada para o controle de caos na modelagem do saltar passos através da rede C-BAM.

Como o objetivo do controle com foco na modelagem saltar passos é o de estabilizar a rede no último padrão armazenado da rede, que representa a solução de uma determinada equação algébrica, a estratégia de controle de caos aqui proposta para a C-BAM utiliza a informação de saída da rede na realização da comparação entre a saída real e a saída desejada (padrão final). Pode ser notado que, diferentemente da estratégia de controle de caos utilizada na família C-BAM com o objetivo de as redes estabilizarem-se em uma memória inacessível, que não utiliza qualquer informação prévia a respeito da saída desejada, na estratégia de controle de caos para o fenômeno saltar passos, a própria saída desejada é utilizada como alvo do controle. Pode ser observado que, no caso da C-BAM, os neurônios da camada X estão conectados com os neurônios da camada Y, de modo que é a camada X desta rede quem efetua a comparação entre a saída real da rede e a saída desejada. Portanto, é na camada X que o controle de caos com o objetivo de estabilizar a C-BAM em uma saída desejada vai ser incluído.

As equações para a rede C-BAM com controle, capaz de estabilizar-se na saída desejada, é descrita da seguinte maneira:

$$x_i(t+1) = f_i(\eta_i(t+1) + \zeta_i(t+1)) \quad (6.22)$$

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^n w_{ij} \left\{ y_j(t) + \sum_{l=0}^{n/I} \delta(j-l-1) K u_j(t) \right\} \quad (6.23)$$

$$u_j(t) = y_j(t) - (1 - \tilde{y}_j) \quad (6.24)$$

$$\zeta_i(t+1) = k_r \zeta_i(t) - \rho x_i(t) + a \quad (6.25)$$

$$y_j(t+1) = f(\eta_j(t+1) + \zeta_j(t+1)) \quad (6.26)$$

$$\eta_j(t+1) = k_f \eta_j(t) + \sum_{i=1}^m w_{ji} x_i(t) \quad (6.27)$$

$$\zeta_j(t+1) = k_r \zeta_j(t) - \rho y_j(t) + a \quad (6.28)$$

Quando o método de controle por pinagem é adicionado a C-BAM, conforme as Equações 6.22 a 6.28, a distância Hamming entre a saída da camada ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado desejado, que é o padrão final no processo de memorização de estágios de operações algébricas, é mostrada na Figura 6.13. Nesta simulação da Figura 6.13 foi considerado o conjunto de padrões da Figura 5.10 para $k_r=0.9$, $k_f=0.7$, $\rho=5$, $a=3$, $I=2$, $K=3$ e padrão inicial \mathbf{P}^1 . Pode ser notado que a rede C-BAM converge rapidamente ($t=5$) para o padrão final desejado (\mathbf{P}^4).

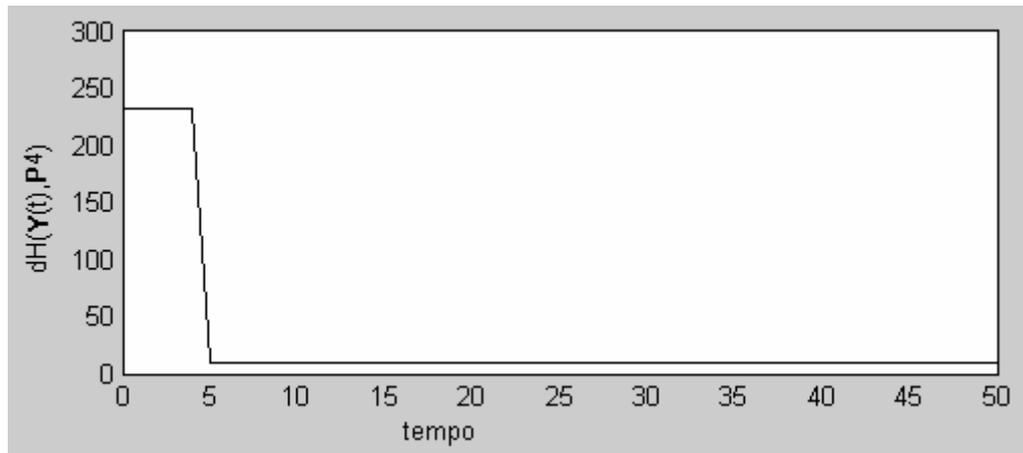


Figura 6.13. Distância Hamming entre a saída da camada Y ($\mathbf{Y}(t)$) da rede C-BAM e o padrão armazenado \mathbf{P}^4 , de $t=0$ a 50, para $k_r=0.9$, $k_f=0.7$, $\rho=5$, $a=3$, $I=2$, $K=3$ e padrão inicial \mathbf{P}^1 .

Os experimentos realizados sugerem que a proposição do método de controle por pinagem para a família de memórias associativas bidirecionais caóticas permite que a C-BAM seja estabilizada em um padrão armazenado. Entretanto, no método de controle por pinagem, dois parâmetros: a distância de pinagem (I) e a intensidade do controle (K), precisam ser determinados *a priori*. Nas Tabelas 6.3 e 6.4 são listados os valores mínimos da intensidade de controle K para cada distância de pinagem I , de modo que os conjuntos de padrões das Figuras 5.10 e 5.11, respectivamente, sejam apropriadamente controlados pelo método de controle por pinagem. Sem

uma escolha apropriada dos valores destes dois parâmetros, as redes C-BAM não convergem para o padrão desejado.

Tabela 6.3. Correlação entre a distância de pinagem (I) e a intensidade mínima de controle (K) para C-BAM com os padrões da Figura 5.10 e $k_r=0.9$, $k_f=0.7$, $\rho=5$, $a=3$, padrão inicial \mathbf{P}^1 .

Distância de Pinagem I	2	3	4	5	6
Intensidade Mínima de Controle K_{min}	3	4	5	7	6
Tempo Inicial de Convergência	5	8	8	8	8

Tabela 6.4. Correlação entre a distância de pinagem (I) e a intensidade mínima de controle (K) para C-BAM com os padrões da Figura 5.11 e $k_r=0.9$, $k_f=0.35$, $\rho=9$, $a=4$, padrão inicial \mathbf{P}^1 .

Distância de Pinagem I	2	3	4	5	6
Intensidade Mínima de Controle K_{min}	2	2	3	3	4
Tempo Inicial de Convergência	11	17	13	27	14

6.2.2 Método de Controle Adaptativo

Procurando melhorar as limitações do método de controle por pinagem, foi proposta uma outra estratégia de controle de caos para a modelagem C-BAM do fenômeno saltar passos, a qual utiliza o método de controle adaptativo. De acordo com Fradkov e Evans (2005), muitas publicações consideram a possibilidade de se aplicar os métodos de adaptação para o controle de processos caóticos pelo fato de, em muitas aplicações físicas, os parâmetros da planta controlada serem desconhecidos e também porque as informações a respeito da estrutura do modelo (por exemplo, a dimensionalidade das equações do modelo ou as suas características não lineares) costumam ser incompletas. A escolha do método de controle adaptativo para o controle de caos da C-BAM é justificada por fatores como: (i) presença de variações na dinâmica do processo, (ii) engenharia eficiente e (iii) facilidade de uso do método.

Intuitivamente, controle adaptativo é caracterizado por modificar seu comportamento em resposta a mudanças na dinâmica do processo e a natureza das perturbações (Astrom e Wittenmark, 1995).

Existem muitas tentativas de se definir formalmente controle adaptativo. Em 1961, um simpósio na área seguiu a seguinte definição: “Um sistema adaptativo é qualquer sistema físico que tenha sido projetado com um ponto de vista adaptativo”. Em 1973, um comitê do IEEE fez uma nova tentativa de definição, na qual foi proposta um vocabulário baseado em noções de sistema de controle auto-organizado (SOC), parâmetro-adaptativo SOC, desempenho-adaptativo SOC e sistema de controle aprendiz (*learning control system*). No entanto, tais esforços em se definir formalmente um controle adaptativo não foram amplamente aceitos. Uma definição significativa para controle adaptativo, a qual faria possível olhar para um controlador de hardware e software e decidir se ele é ou não adaptativo, ainda estava faltando. Apenas o que parecia ser um consenso era o fato de que um sistema de realimentação de ganho constante não é um sistema adaptativo. Astrom e Wittenmark (1995), em uma atitude pragmática, consideraram que “um controle adaptativo é um controle com parâmetros ajustáveis e um mecanismo para ajustar os parâmetros”. Esta definição será utilizada nesta tese.

6.2.2.1 Modelos da Planta Controlada

Sejam considerados sistemas contínuos no tempo, com parâmetros concentrados descritos no espaço de estado por equações diferenciais. De acordo com Andrievskii e Fradkov (2003), Fradkov e Evans (2005), os modelos destes sistemas são representados na forma paramétrica como:

$$\dot{\mathbf{z}} = \mathbf{F}(\mathbf{z}, \boldsymbol{\theta}, \mathbf{u}), \quad \mathbf{q} = \mathbf{h}(\mathbf{z}) \quad (6.29)$$

em que \mathbf{z} é o vetor n -dimensional de variáveis de estado; $\dot{\mathbf{z}} = dz/dt$ significa a derivada no tempo de \mathbf{z} ; $\boldsymbol{\theta}$ é o vetor de parâmetros desconhecidos, e \mathbf{u} é o vetor m -dimensional de entradas (variáveis de controle). Usualmente assume-se que a função vetor $\mathbf{F}(\mathbf{z}, \mathbf{u})$ é continuamente diferenciável. O modelo precisa também incluir uma descrição da medida do processo (*measurement process*). O vetor l -dimensional de variáveis de saída (\mathbf{q}) é definido, por exemplo, como na Equação 6.29. Se as saídas não são definidas explicitamente, assume-se que todas as variáveis de estado são disponíveis para medição, ou seja, $\mathbf{q} = \mathbf{z}$. Baseada na Equação 6.29, a representação paramétrica do controlador é:

$$\mathbf{u} = \mathbf{U}(\mathbf{z}, \mathbf{u}, \boldsymbol{\xi}) \quad (6.30)$$

em que $\xi = \Phi(\theta)$ para algum mapeamento $\Phi(\cdot)$. Os processos obtidos ao se medir o estado do sistema $\{z(t)\}$ ou sua saída $\{q(t)\}$ são utilizados para determinar estimativas $\hat{\theta}(t)$ dos parâmetros desconhecidos do sistema ($\theta(t)$), ou para ajustar diretamente os parâmetros do controlador $\hat{\xi}(t)$. Para a adaptação dos parâmetros, uma variedade de métodos existentes na literatura de controle (Fradkov e Evans, 2005) pode ser usada, tais como o gradiente, gradiente rápido (*speed-gradient*), mínimos quadrados (*least squares*), mínimos quadrados sobrecarregados (*weighted least squares*), entre outros. O controlador da Equação 6.30 é usualmente projetado para ser utilizado em abordagens modelo-referência ou linearização com realimentação (*feedback linearization*).

Muitos autores (Fradkov e Evans, 2005) consideram modelos espaço de estado discretos no tempo:

$$\mathbf{z}_{k+1} = \mathbf{F}_d(\mathbf{z}_k, \mathbf{u}_k), \quad \mathbf{q} = \mathbf{h}(\mathbf{z}) \quad (6.31)$$

em que $\mathbf{z}_k \in \mathfrak{R}^n$, $\mathbf{u}_k \in \mathfrak{R}^m$, $\mathbf{q}_k \in \mathfrak{R}^l$, representam o valor dos vetores de estado, de entrada e de saída no k -ésimo estágio do processo. O modelo da Equação 6.31 é determinado nos mapas \mathbf{F}_d , \mathbf{h} .

Astrom e Wittenmark (1995), afirmam que um sistema de controle adaptativo possui uma estrutura característica, que é formada por dois laços (*loops*). Um laço é o da realimentação normal, com o controle e o processo, sendo que o outro laço, o dito externo, é de ajuste paramétrico. Um diagrama de bloco de um sistema adaptativo é mostrado na Figura 6.14.

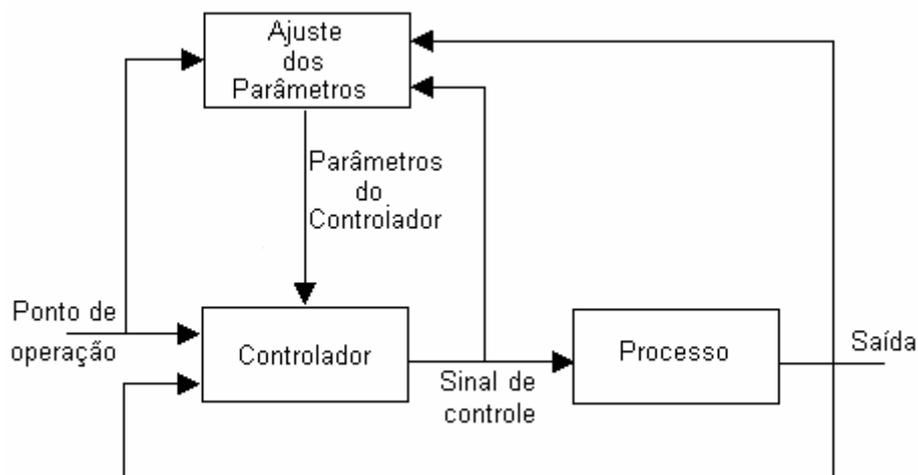


Figura 6.14. Diagrama de bloco de um sistema adaptativo.

A construção de um controle adaptativo contém os seguintes passos:

- Caracterizar o comportamento desejado do sistema de laço fechado;
- Determinar uma lei de controle apropriada para os parâmetros ajustáveis;

- Encontrar um mecanismo para ajustar os parâmetros;
- Implementar a lei de controle.

O passo chave na construção do controle adaptativo é a escolha do mecanismo de ajuste dos parâmetros. Astrom e Wittenmark (1995) descrevem quatro maneiras de se fazer o mecanismo de ajuste dos parâmetros em esquemas adaptativos: agendamento do ganho (*gain scheduling*), controle adaptativo modelo-referência (*model-reference adaptive control*), regulador auto-ajustável (*self-tuning regulators*) e controle dual (*dual control*).

Considerando-se que o objetivo de um controle na rede C-BAM capaz de modelar o fenômeno saltar passos é o de suprimir a dinâmica caótica desta rede, de modo que ela se estabilize em um padrão desejado, o mecanismo de ajuste dos parâmetros selecionado para o problema em questão foi o controle adaptativo modelo-referência (MRAS), pelo fato de este modelo lidar diretamente com a saída da rede, informando como a saída do processo deve, idealmente, responder a um sinal de comando.

6.2.2.2 Sistema Adaptativo Modelo-Referência (MRAS)

O sistema adaptativo modelo-referência (MRAS) foi originalmente proposto para resolver um problema no qual o desempenho desejado é expresso em termos de um modelo referência, o qual dá a resposta desejada para um sinal de comando. Um diagrama de bloco do sistema é mostrado na Figura 6.15.

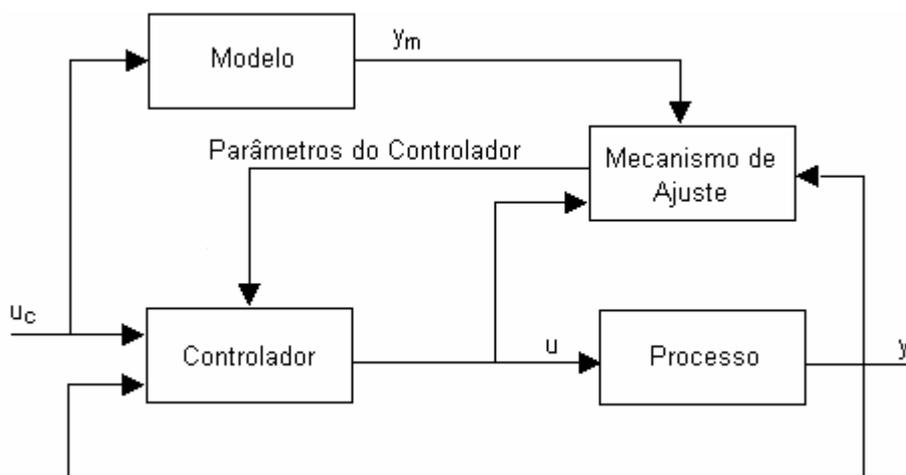


Figura 6.15. Diagrama de bloco de um sistema adaptativo modelo-referência (MRAS).

O MRAS pode ser pensado como consistindo de dois laços. O laço interno é um laço de realimentação comum, envolvendo o processo e o controlador. No laço interno, u_c denota o sinal de

comando e u o sinal de controle. O laço externo ajusta os parâmetros do controlador de tal maneira que o erro, definido como a diferença entre a saída do processo (y) e a saída do modelo (y_m), seja minimizada. Portanto, no MRAS, o comportamento desejado para o sistema é especificado por um modelo, e os parâmetros do controlador são ajustados com base na discrepância entre o comportamento do modelo e do processo.

O mecanismo para ajustar os parâmetros do controlador em um MRAS pode ser obtido de duas maneiras: usando o método gradiente ou aplicando a teoria da estabilidade (Astrom e Wittenmark, 1995). Sistemas adaptativos modelo-referência foram originalmente derivados para sistemas determinísticos contínuos no tempo. Entretanto, sua extensão para sistemas de tempo-discreto pode ser facilmente obtida.

6.2.2.3 A Regra MIT

A Regra MIT é a abordagem original para o controle adaptativo modelo-referência. O nome é derivado do fato de que ele foi desenvolvido pelo Laboratório de Instrumentação (atualmente o Laboratório Negociador – Draper Laboratory), no MIT.

A regra MIT (Figura 6.16) é utilizada em um sistema realimentado no qual o controlador tem um parâmetro ajustável, θ . A resposta desejada para o laço fechado interno é especificada por um modelo, cuja saída é y_m . Seja $e = y - y_m$ o erro entre a saída do laço fechado (y) e a saída do modelo (y_m). É possível ajustar o parâmetro θ de tal maneira que a função perda seja minimizada:

$$J(\theta) = \frac{1}{2}e^2 \quad (6.32)$$

Para diminuir J , mudam-se os parâmetros na direção do gradiente negativo de J :

$$\frac{d\theta}{dt} = -\gamma \frac{\partial J}{\partial \theta} = -\gamma e \frac{\partial e}{\partial \theta} \quad (6.33)$$

A Equação 6.33 expressa a Regra MIT. A derivada parcial $\partial e / \partial \theta$, chamada derivada de sensibilidade do sistema, informa como o erro é influenciado pelo parâmetro ajustável e o parâmetro γ determina a taxa de adaptação. Assumindo-se que as mudanças paramétricas são mais lentas do que as outras variáveis no sistema, então a derivada $\partial e / \partial \theta$ pode ser avaliada supondo-se θ constante.

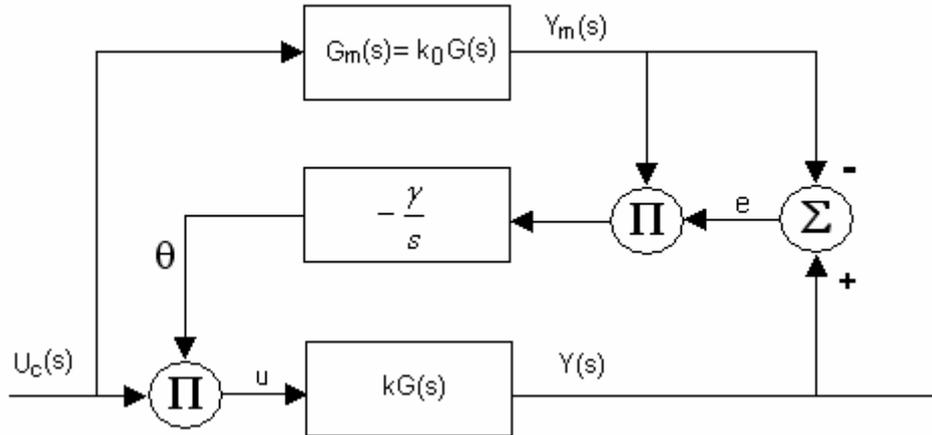


Figura 6.16. Diagrama de bloco do MRAS com um ajuste paramétrico baseado na Regra MIT.

Tendo analisado a Regra MIT, o resumo do sistema adaptativo modelo-referência (MRAS) pode ser descrito como se segue:

Seja o desempenho desejado especificado por um modelo referência com função transferência $kG(s)$, e seja $G_m(s)=k_0 G(s)$, em que k_0 é uma constante, a função transferência do controlador da planta e u_c o sinal de comando. O MRAS tenta mudar os parâmetros de controle de modo que o erro

$$e = y - y_m = kG(p)\theta u_c - k_0 G(p)u_c, \quad \text{em que } p=d/dt \quad (6.34)$$

vá para zero. Logo

$$\frac{\partial e}{\partial \theta} = kG(p)u_c \quad (6.35)$$

Na Figura 6.16 tem-se que $y_m = k_0 G(p)u_c$. Portanto

$$G(p) = \frac{y_m}{k_0 u_c} \quad (6.36)$$

Substituindo-se a Equação 6.36 na Equação 6.35, tem-se que

$$\frac{\partial e}{\partial \theta} = \frac{k}{k_0} y_m \quad (6.37)$$

Finalmente, a Equação 6.33 pode ser re-escrita como

$$\frac{d\theta}{dt} = -\gamma e \left(\frac{k}{k_0} y_m \right) = -\beta y_m e, \quad \text{em que } \beta = \gamma \frac{k}{k_0} \quad (6.38)$$

A escolha do ganho de adaptação é crítica e depende dos níveis dos sinais. Os valores iniciais dos parâmetros devem ser tais que o sistema laço fechado seja estável. Evidências empíricas indicam que o sistema é estável para pequenos ganhos de adaptação, enquanto altos ganhos conduzem à instabilidade.

6.2.2.4 Controle de Caos da Rede C-BAM com o Esquema Adaptativo MRAS

Um método de controle adaptativo no qual não se escolhe dois parâmetros desconhecidos e não se utilize padrão inverso do desejado como padrão alvo é proposto para o controle da rede C-BAM capaz de modelar o fenômeno saltar passos. O esquema de controle proposto possui um controlador adaptativo do tipo modelo referência (MRAS), e o mecanismo utilizado para ajustar os parâmetros do MRAS foi a Regra MIT (Figura 6.16).

No caso de controlar a rede C-BAM com o esquema MRAS baseado na Regra MIT, assume-se que o processo possui uma função de transferência $G(s)$ e que o desempenho desejado seja especificado por um modelo referência com função transferência $G_m(s)$ ($k_0 = k = 1$). Com o controlador $u = \theta u_c$, em que u é o sinal de controle e u_c é o sinal de comando, a função de transferência a partir do sinal de comando para a saída torna-se $\theta G(s)$. Esta função de transferência é igual a $G_m(s)$ se o parâmetro θ for escolhido apropriadamente. Utilizando-se a Regra MIT para obter um método para ajustar o parâmetro θ , o erro é dado por:

$$e = y - y_m = G(p)\theta u_c - G(p) u_c \quad (6.39)$$

em que u_c é o sinal de comando, y_m é a saída do modelo, y é a saída do processo, θ é o parâmetro ajustável, e $p = d/dt$ é o operador diferencial. A derivada de sensibilidade é dada por:

$$\frac{\partial e}{\partial \theta} = G(p)u_c = y_m \quad (6.40)$$

Ajustando θ de modo que a função perda $J(\theta) = 1/2e^2$ seja minimizada, a Regra MIT produz a seguinte lei de adaptação:

$$\frac{d\theta}{dt} = -\gamma \frac{\partial J}{\partial \theta} = -\gamma e \frac{\partial e}{\partial \theta} = -\gamma e y_m \quad (6.41)$$

Discretizando a idéia acima, o controlador adaptativo discreto com o diagrama de bloco do MRAS e regra MIT usando a transformada-z é mostrado na Figura 6.17.

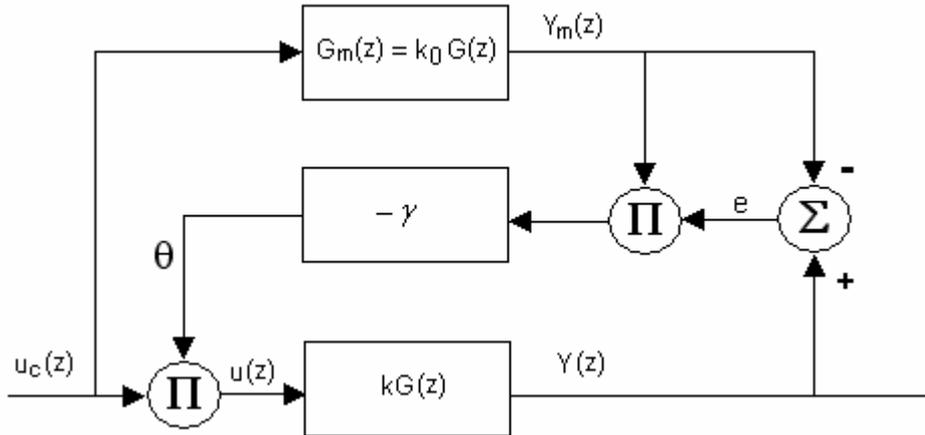


Figura 6.17. Controlador adaptativo discreto com diagrama de bloco usando transformada-z.

A dinâmica da rede C-BAM caótica com um esquema de controle adaptativo do tipo MRAS que utiliza a Regra MIT (da Equação 6.41) é dada pelo seguinte conjunto de equações:

$$x_i(t+1) = f(\eta_i(t+1) + \varsigma_i(t+1)) \quad (6.42)$$

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^n w_{ij} y_j(t) \quad (6.43)$$

$$\varsigma_i(t+1) = k_r \varsigma_i(t) - \rho x_i(t) + a \quad (6.44)$$

$$y_j(t+1) = f(\eta_j(t+1) + \varsigma_j(t+1)) \quad (6.45)$$

$$\eta_j(t+1) = k_f \eta_j(t) + \sum_{i=1}^m w_{ji} x_i(t) \theta_j(t) \quad (6.46)$$

$$\theta_j(t) = \theta_j(t-1) - \gamma \Delta \theta_j \quad (6.47)$$

$$\Delta \theta_j = (y_m)_j * e_j \quad (6.48)$$

$$e_j = y_j(t) - (y_m)_j \quad (6.49)$$

$$\varsigma_j(t+1) = k_r \varsigma_j(t) - \rho y_j(t) + a \quad (6.50)$$

Os resultados dos experimentos sugerem a eficácia do esquema de controle (Equações 6.42 a 6.50) para a rede C-BAM. Seja o conjunto de quatro padrões da Figura 5.10, utilizado na modelagem do saltar passos da equação algébrica $-X-A = B$, em uma representação bipolar, na qual os quadrados pretos significam bits de valor 1, enquanto quadrados brancos significam bits de valor -1. Foi visto na Figura 6.11 que, após a rede C-BAM recuperar o padrão final (\mathbf{P}^4), a rede continua transitando por outros pares armazenados, sem se estabilizar no padrão desejado. As

Figuras 6.18 a 6.20 apresentam simulações do MRAS para a rede C-BAM, de acordo com as equações dadas acima. Nestas figuras estão ilustradas as distâncias Hamming entre a saída do processo (y , representada pela linha contínua) e a saída do modelo (variável y_m , representada pela linha pontilhada).

As Figuras 6.18 a 6.20 também mostram que a rede converge rapidamente para o padrão desejado com valores ainda pequenos de ganho de adaptação. A convergência ocorre quando a saída do processo torna-se idêntica à saída do modelo.

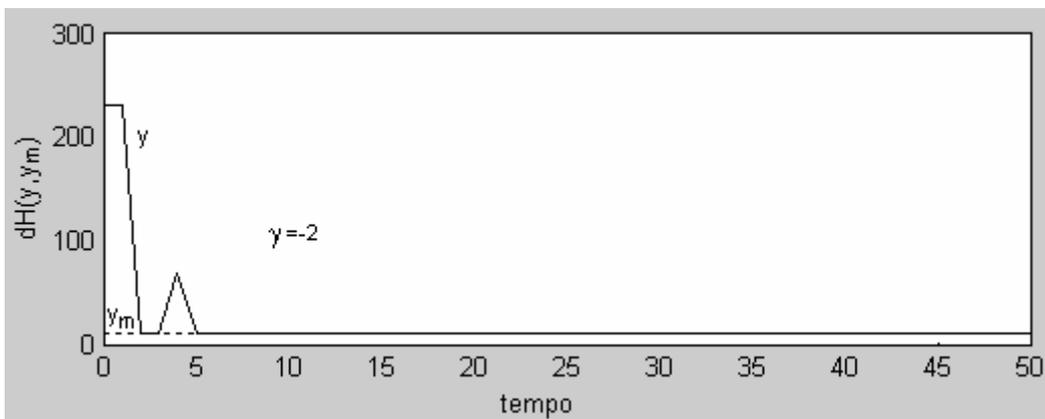


Figura 6.18. Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo (linha contínua) e a saída do modelo (linha pontilhada) é ilustrada para $k_f=0.7$, $k_r=0.9$, $\rho=5$, $a=3$, ganho de adaptação $\gamma = -2$, padrão inicial \mathbf{P}^1 e $y_m = \mathbf{P}^4$.

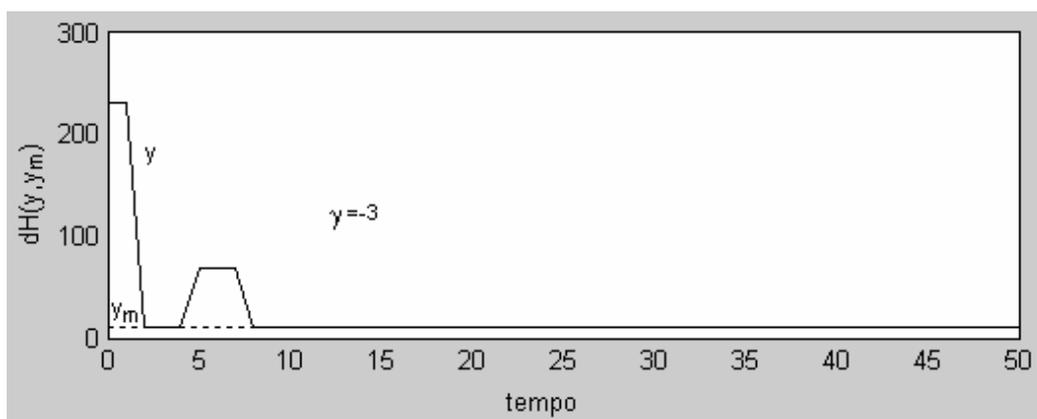


Figura 6.19. Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo (linha contínua) e a saída do modelo (linha pontilhada) é ilustrada para $k_f=0.7$, $k_r=0.9$, $\rho=5$, $a=3$, ganho de adaptação $\gamma = -3$, padrão inicial \mathbf{P}^1 e $y_m = \mathbf{P}^4$.

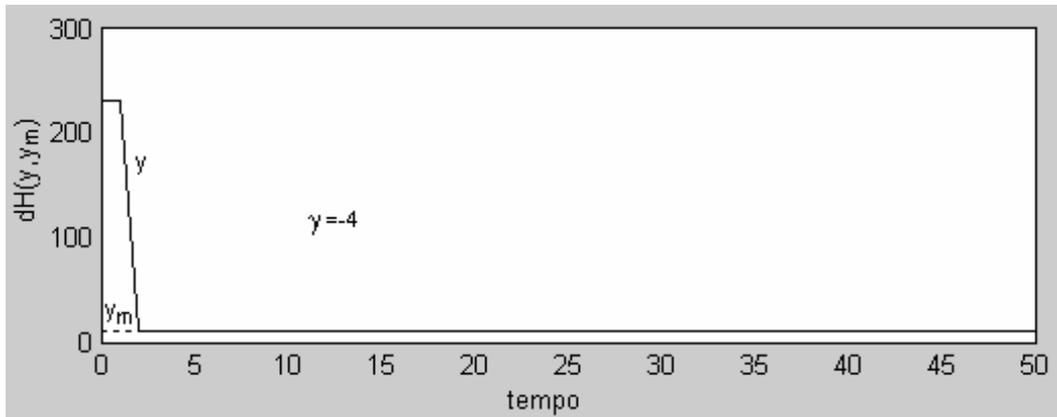


Figura 6.20. Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo e a saída do modelo é ilustrada para $k_f=0.7$, $k_r=0.9$, $\rho=5$, $a=3$, $\gamma=-4$, padrão inicial \mathbf{P}^1 e $y_m = \mathbf{P}^4$. Foi verificado que a dinâmica da rede não se modifica ao utilizar-se $\gamma > -4$.

Mesmo com um rápido processo de convergência pode ser notado que a taxa de convergência depende do ganho de adaptação. Torna-se, portanto, importante conhecer um valor razoável para este parâmetro. O exemplo da Figura 6.21 mostra claramente que a escolha do ganho de adaptação é crucial, e que a dinâmica da rede mantém-se imprevisível para valores inadequados de γ .

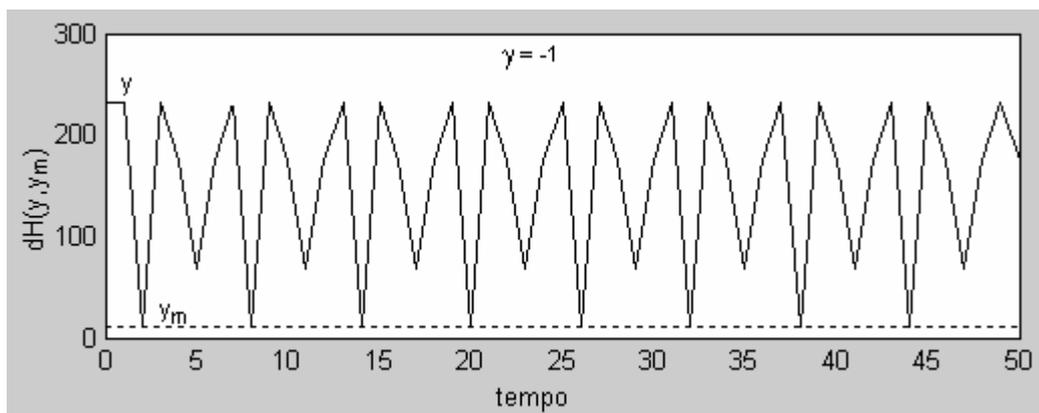


Figura 6.21. Simulação de um MRAS para a rede C-BAM. A distância Hamming entre a saída do processo e a saída do modelo é ilustrada para $k_f=0.7$, $k_r=0.9$, $\rho=5$, $a=3$, ganho de adaptação $\gamma = -1$, padrão inicial \mathbf{P}^1 e $y_m = \mathbf{P}^4$.

Experimentos com o conjunto de seis padrões da Figura 5.11, utilizado na modelagem do saltar passos da equação algébrica $-A/B-X=-C$, em uma representação bipolar, também sugeriram que a rede C-BAM com controle adaptativo é capaz de estabilizar-se no padrão final (\mathbf{P}^6). Foi verificado, por simulação computacional, que para $k_f=0.35$, $k_r=0.9$, $\rho=9$, $a=4$ e $y_m = \mathbf{P}^6$, a rede se estabiliza no padrão final desejado com $y \geq -3$. Na Figura 6.22 estão ilustradas as distâncias Hamming entre a saída do processo (y) e a saída do modelo (y_m).

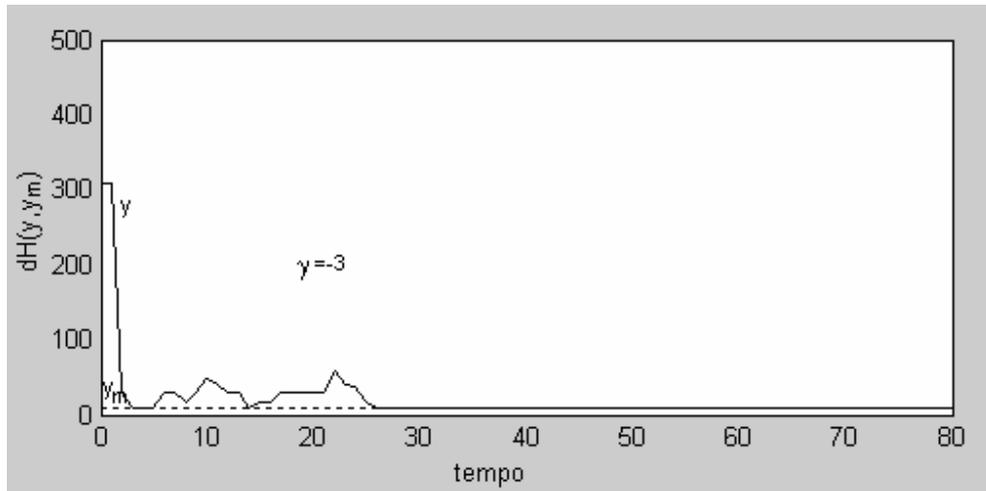


Figura 6.22. Simulação do MRAS para a rede C-BAM. A distância Hamming entre a saída do processo e a saída do modelo é ilustrada para $k_f=0.35$, $k_r=0.9$, $\rho=9$, $a=4$, $y_m=\mathbf{P}^6$ e ganho de adaptação $\gamma=-3$.

6.2.2.5 Discussão da Estratégia de Controle de Caos Proposta para a Modelagem do Fenômeno Saltar Passos

Duas estratégias de controle de caos foram propostas para a rede C-BAM com foco na modelagem do fenômeno saltar passos: uma utilizando o método de controle por pinagem e outra utilizando o método de controle adaptativo. Os resultados dos experimentos sugerem a eficácia destes dois esquemas de controle. Ambos são práticos, eficazes e apresentam rápida convergência. Também em ambos o padrão desejado é previamente conhecido e apresentado ao modelo.

Uma das limitações encontradas no método de controle por pinagem para controlar a rede C-BAM é a necessidade da determinação apropriada, *a priori*, de dois parâmetros: a distância de pinagem (I) e a intensidade do controle (K). Sem uma escolha apropriada destes dois parâmetros, as redes C-BAM não convergem para o padrão desejado. Verificou-se por experimentação que, ao utilizar o padrão desejado como referência, a C-BAM converge para o inverso deste padrão. Assim, o padrão referência no método de controle por pinagem precisa ser o inverso do padrão desejado. Além disso, para padrões armazenados de grande extensão, como é o caso dos conjuntos de padrões aqui utilizados no fenômeno saltar passos, o período de tempo gasto na computação do método de controle por pinagem é considerável, pois muitos cálculos são necessários (Equações 6.23 e 6.24).

Na tentativa de reduzir as limitações encontradas no método de controle por pinagem, um método de controle adaptativo foi proposto para controlar a rede C-BAM. No método de controle adaptativo, apenas o parâmetro γ , que determina a taxa de adaptação, precisa ser escolhido. Em

relação ao tempo de cálculo computacional (tempo processamento do modelo), o método de controle adaptativo é bem mais eficiente. Os experimentos mostraram que a utilização do padrão desejado em si como padrão referência é suficiente para que a rede convirja corretamente, não sendo mais necessária a utilização do padrão inverso nas equações da rede. Embora o método de controle adaptativo apresente algumas vantagens sobre o método de controle por pinagem, esse controle adaptativo também apresenta uma limitação. Foi visto que, no método de controle por pinagem, cada erro u_j é calculado na camada X. Conseqüentemente, os erros são adicionados aos próprios neurônios y_j (Equações 6.23 e 6.24). Já no método de controle adaptativo, cada erro e_j é calculado na camada Y, estando, portanto, relacionado com o neurônio x_i (Equações 6.46 a 6.49). Como o controle é feito variável a variável, para que o cálculo deste controle seja possível, é necessário que o tamanho i da camada X seja do mesmo tamanho j da camada Y. No caso da modelagem do fenômeno saltar passos todos os padrões armazenados apresentam mesma dimensão, uma vez que constituem uma seqüência de estágios de operações algébricas. Sendo assim, o tamanho da camada X é igual ao tamanho da camada Y (ou seja, $i=j$), de modo que o método de controle adaptativo proposto para a C-BAM pode ser utilizado para controlar a modelagem do fenômeno saltar passos.

6.3 Resumo

Embora as propriedades das dinâmicas caóticas da família C-BAM e a aplicabilidade dessas dinâmicas em seu estado natural tenham se mostrado ricas, procurou-se uma forma de lidar com a com o novo modo de acesso a memórias armazenadas gerado pelas redes C-BAM, de modo a não deixar que as saídas da família C-BAM permaneçam vagueando em torno de todos os padrões armazenados, as quais mudam continuamente e não podem ser estabilizadas em um dos padrões armazenados.

O controle de caos foi proposto considerando-se dois focos: um na recuperação de memórias inacessíveis e outro na modelagem do fenômeno saltar passos. Através da proposição do método de controle por pinagem para a família C-BAM, as redes C-BAM tornam-se aptas a se estabilizar nas memórias inacessíveis desejadas, de modo a permitir a hetero-associação correta de memórias antes inacessíveis.

A proposição do controle de caos no contexto da modelagem do fenômeno saltar passos foi realizada por meio de duas estratégias, uma baseada no método de controle por pinagem e outra baseada em um método de controle adaptativo. Os experimentos sugerem que ambas permitem que a rede C-BAM se estabilize no padrão final representando a solução da operação algébrica memorizada pela rede, de forma prática e eficaz, desde alguns parâmetros sejam escolhidos de

forma apropriada. Em casos de conjuntos de padrões de mesmo tamanho, o método de controle adaptativo apresenta-se mais eficiente, pois apenas um parâmetro precisa ser determinado a priori, o tempo de cálculo computacional é menor e o próprio padrão desejado é utilizado como padrão referência.

Embora neste capítulo o controle de caos para o caso do fenômeno saltar passos esteja associado a uma tarefa específica, a solução de equações algébricas, deve ser enfatizado que este é apenas um caso de estudo. O controle de caos, tanto através do método de controle por pinagem quanto através do método de controle adaptativo, serve para detectar e convergir para algum estado da uma seqüência de passos, quando ele é visitado.

Capítulo 7

Conclusões e Trabalhos Futuros

O objetivo deste capítulo é apresentar as conclusões derivadas da modelagem, implementação e testes realizados com a família C-BAM. Inicialmente, na Seção 7.1, são discutidos os principais resultados alcançados e são apontadas limitações encontradas na família C-BAM. Na Seção 7.2 são enfatizadas as principais contribuições desta tese. A Seção 7.3 encerra o capítulo, sugerindo propostas para futuras pesquisas com a família C-BAM, as quais foram planejadas a partir de limitações observadas e de tópicos interessantes ainda não desenvolvidos. Evidências e explicações detalhadas das conclusões aqui apresentadas podem ser encontradas no conteúdo dos capítulos anteriores.

7.1 Desenvolvimento e Limitação da Família C-BAM

Nos últimos anos a literatura tem registrado um número crescente de redes neurais artificiais que apresentam fenômenos caóticos. Entre essas redes encontram-se os neurônios artificiais caóticos originalmente inseridos nos modelos auto-associativos (Adachi e Aihara, 1997), os quais ilustram respostas não periódicas previamente detectadas em neurônios biológicos através de experimentos com animais (Aihara e Matsumoto, 1987).

Em 1996, Osana *et al.* publicaram um trabalho inicial para tornar caótico o modelo BAM. Os autores introduziram alguns neurônios caóticos em parte da BAM, representando uma informação contextual. O objetivo do trabalho de Osana *et al.* é permitir à rede BAM lidar com uma-paramuitas associações de padrões, tais como $\{(\underline{\mathbf{X}}^1, \mathbf{Y}^1), (\underline{\mathbf{X}}^1, \mathbf{Y}^2), (\underline{\mathbf{X}}^1, \mathbf{Y}^3), (\underline{\mathbf{X}}^2, \mathbf{Y}^4)\}$, o que não é possível no caso da BAM convencional.

A contribuição inicial desta tese é estender o trabalho de Aihara e colegas (Aihara *et al.* 1990, Adachi e Aihara, 1997, Aihara, 2002) ao introduzir uma família de modelos de memórias

associativas bidirecionais caóticas na qual todos os neurônios originais da BAM, BAM com atraso e BAM exponencial são substituídos por neurônios caóticos. Os resultados após a substituição sugeriram que a modelagem dos neurônios caóticos em redes de memória hetero-associativa é uma extensão natural da rede de memória hetero-associativa convencional ($k_r = k_f = \rho = a = 0$) e que, mesmo em modelos totalmente caóticos, o comportamento das redes C-BAM pode ser idêntico ao dos modelos BAM originais (ao usar valores baixos para k_r , k_f , ρ e a).

A influência de cada parâmetro do neurônio caótico sobre as C-BAMs foi estimada através do planejamento de experimentos (Montgomery, 1997), de modo a sugerir quais parâmetros possuem maior influência sobre o comportamento dinâmico das redes de memória associativa bidirecionais caóticas. Embora a análise realizada através do planejamento de experimentos não tenha permitido criar faixas precisas de valores de parâmetros para ocorrência de comportamentos caóticos específicos (bifurcações, crises), ela permitiu comparar e relacionar valores de parâmetros para a ocorrência de caos. Os experimentos sugeriram que todos os parâmetros do neurônio caótico afetam o comportamento da família C-BAM, mas cada um deles apresenta um grau diferente de influência. O parâmetro k_r parece ser o mais influente. De fato, o planejamento fatorial revela que k_r está sempre envolvido no comportamento caótico.

Seguindo a identificação dos parâmetros de acordo com os resultados do planejamento de experimentos, foi possível ilustrar como as modificações paramétricas geram diferentes tipos de comportamento dinâmico. Enquanto o modelo caótico apresenta pontos fixos, comportamentos periódicos, bifurcação, caos e crise, o modelo original sempre converge para um ponto fixo (período 1). Os experimentos foram realizados para toda a família C-BAM e para mais de um conjunto de padrões armazenados. Resultados idênticos foram encontrados para todos os modelos da família C-BAM, de modo a descaracterizar a emergência da diversidade de comportamentos dinâmicos como uma descoberta isolada. A existência de um dos comportamentos dinâmicos sugeridos pelos experimentos empíricos, o caos, foi provada para toda a família C-BAM, através do cálculo dos expoentes de Lyapunov. Além disso, a rica dinâmica encontrada nos experimentos realizados com a família C-BAM incluiu a ocorrência de períodos ímpares e de outros períodos que não são potências de dois, algumas vezes em meio ao caos, em intervalos irregulares. Tais ocorrências sugerem a presença de comportamentos caóticos diferentes dos até então tratados neste trabalho, e poderão ser analisados futuramente.

A dinâmica caótica apresentada pela família C-BAM modificou o comportamento da acessibilidade à memória: ao invés das redes C-BAMs recuperarem um único par de padrão, como as redes BAMs faziam, foi detectada, na versão caótica, uma grande diversidade de padrões recuperados, incluindo a ocorrência de complexas transições entre os padrões memorizados para alguns conjuntos de valores de parâmetros. As complexas transições entre alguns ou entre todos os

padrões armazenados não são compulsórias: os padrões armazenados podem aparecer exatamente conforme eles foram treinados ($(\mathbf{X}^k, \mathbf{Y}^k)$, $k=1, \dots, p$) ou podem aparecer ocasionalmente, sem uma seqüência de recuperação específica e entre atratores espúrios. Além disso, algumas vezes a família C-BAM recupera recombinações de padrões armazenados (por exemplo, $(\mathbf{X}^1, \mathbf{Y}^3)$) ou então combinações de padrões treinados com padrões não treinados (por exemplo $(\mathbf{X}^1, \mathbf{Y}^j)$, em que $j \neq k=1, \dots, p$). Foi verificado, por experimentação, que o modo como os padrões são recuperados depende da escolha apropriada dos valores de parâmetros.

A grande diversidade de padrões recuperados, envolvendo excursões complexas sobre as memórias armazenadas como consequência direta do comportamento caótico da família C-BAM, permitiu que fossem modeladas duas habilidades presentes em organismos inteligentes naturais: a habilidade de recuperar memórias inacessíveis pelas redes BAM originais e a habilidade de saltar passos na realização de uma tarefa. Para a implementação e testes com esta última habilidade foi selecionada apenas uma rede da família C-BAM, que é a rede C-BAM.

No processo de recuperação de memórias antes inacessíveis, foi verificado que, para alguns valores de parâmetros, a emergência de várias seqüências de recuperações (com diferentes tamanhos), composta por padrões treinados e não treinados, possibilita a todas as redes da família C-BAM passarem por padrões armazenados que as redes originais não conseguiam recuperar, eliminando, portanto, a ocorrência de memórias inacessíveis nas C-BAMs.

Na modelagem do saltar passos foi verificado que, enquanto na BAM original a entrada \mathbf{P}^k é capaz de gerar apenas um próximo passo \mathbf{P}^{k+1} na solução de uma tarefa, a nova acessibilidade a memórias gerada pelas C-BAMs permite que a rede caótica excursione por todos ou alguns \mathbf{P}^k ($k=1, \dots, p$) e, assim, modele a habilidade de saltar passos por um indivíduo na solução de uma tarefa. Excursões pela memória que seguem a seqüência original dos padrões armazenados, sem saltar qualquer padrão desta seqüência, são consideradas como tendo o procedimento de um noviço na habilidade para executar uma tarefa, enquanto as excursões que pulam alguns padrões armazenados são classificadas como tendo o procedimento de um especialista. Além disso, são ilustrados casos nos quais a recuperação de um noviço pode ser transformada em uma recuperação de especialista através de variação paramétrica.

Os experimentos com o fenômeno saltar passos sugeriram que, por meio de mudanças paramétricas em k_r (enquanto os outros parâmetros permaneciam fixos), é possível encontrar trajetórias de excursão de memória começando do primeiro padrão armazenado e alcançando o último padrão, com seqüências de vários tamanhos. De fato, as Figuras 5.14 a 5.19 da Seção 5.4 mostram que há situações em que a habilidade para saltar passos está diretamente relacionada com o aumento ou diminuição do parâmetro k_r . No entanto, outros experimentos apontaram casos em que o simples aumento seqüencial de k_r não é suficiente para gerar os saltos obtidos na recuperação

dos padrões armazenados. Isso ocorre porque a dinâmica das C-BAMs é muito rica, incluindo diversos comportamentos dinâmicos.

Quando a excursão por padrões armazenados da C-BAM não obedece à seqüência originalmente armazenada, ou seja, quando a rede caótica comporta-se como um especialista na solução de problemas, ao reduzir o número de estados entre o padrão inicial e o padrão final, verifica-se uma diminuição no tamanho da trajetória entre esses dois padrões. Isso significa que a trajetória do especialista feita pela C-BAM modela o uso de atalhos (em número de passos) realizados pelos participantes dos experimentos de Blessing e Anderson (1996), o que acarreta em uma redução substancial no número de estados intermediários para resolução do problema.

Ainda na modelagem do saltar passos pela C-BAM foram verificadas ocorrências de repetições de pares de padrões entre a trajetória dos pares de padrões iniciais e finais. Por exemplo, na Figura 5.15 a seqüência da trajetória entre os pares de padrões iniciais e finais deu-se exatamente na forma $(P^1, P^2) \rightarrow (P^2, P^3) \rightarrow (P^1, P^2) \rightarrow (P^4, P^5) \rightarrow (P^1, P^2) \rightarrow (P^4, P^5) \rightarrow (P^5, P^6)$, ou seja, durante a trajetória de $(P^1, P^2) \rightarrow (P^2, P^3) \rightarrow (P^4, P^5) \rightarrow (P^5, P^6)$, a C-BAM repetiu a recuperação de pares (P^1, P^2) . Também pode ser verificado que, após a ocorrência do último par de padrões, pares de padrões que foram saltados na seqüência original podem aparecer. Por exemplo, na Figura 5.16, após a C-BAM alcançar (P^5, P^6) através da seqüência $(P^1, P^2) \rightarrow (P^4, P^5) \rightarrow (P^5, P^6)$, a rede recuperou o par saltado (P^2, P^3) . Foram também observadas operações de transposição na modelagem do saltar passos, nas quais a C-BAM troca a ordem dos passos intermediários entre os passos finais e iniciais. Tais ocorrências não foram melhor analisadas pelo fato de o interesse na modelagem do saltar passos estar na capacidade de a C-BAM passar pela seqüência de pares armazenados.

Com base nas evidências e nos experimentos de Blessing e Anderson (1996), deve ser enfatizado que, no processo de modelar a habilidade de um noviço ou de um especialista na solução de um problema algébrico, a rede C-BAM não desenvolveu uma percepção mais profunda das equações matemáticas, e o objetivo não foi fazê-la “entender” os cálculos matemáticos. Ao invés disso, quando a C-BAM salta passos na solução de uma equação algébrica significa que os parâmetros caóticos afetam a maneira de a C-BAM acessar os padrões armazenados, o que permite que a recuperação desses padrões seja feita desprezando-se alguns passos intermediários da seqüência originalmente treinada, gerando, com isso, trajetórias menores do que as completas.

Apesar de a dinâmica caótica da família C-BAM ter se mostrado muito rica, percebeu-se que seria interessante apresentar um mecanismo capaz de estabilizar a saída das redes C-BAMs em um padrão armazenado específico, ou seja, no padrão desejado. Sendo assim, através da proposição e da implementação de métodos de controle de caos: (a) a família C-BAM concretiza sua acessibilidade a memórias inacessíveis à família BAM ao convergir para padrões anteriormente não

recuperáveis, (b) a rede C-BAM, no fenômeno saltar passos, estabiliza-se no padrão final após alcançá-lo na seqüência de recuperação de padrões. No desenvolvimento desta tese, dois métodos de controle foram propostos e implementados: o método de controle por pinagem e o método de controle adaptativo.

O método de controle por pinagem foi utilizado nas duas situações de controle de caos: (a) na convergência da família C-BAM em uma memória inacessível, (b) na convergência da rede C-BAM em um padrão final, representando a solução da operação algébrica memorizada pela rede durante o fenômeno saltar passos. No caso das memórias inacessíveis, o objetivo do método de controle por pinagem foi extinguir o caos ao aplicar o controle de caos apenas na camada que opera sobre a informação de entrada, evitando qualquer uso de informações prévias sobre a saída desejada. Então, na C-BAM e na C-BAM com atraso, o sinal de controle (u_j) foi codificado na camada Y, enquanto o sinal de controle foi introduzido na camada X na rede C-eBAM. Como o conjunto de padrões do caso de memórias inacessíveis possui um número pequeno de bits, as redes da família C-BAM convergiram para um padrão desejado quando a intensidade de controle (K) e o sinal de controle (u_j) foram ambos aplicados em todos os neurônios da camada da rede contendo a regra de controle. Além disso, foi usualmente determinado por experimentação que, na C-eBAM, o sinal de controle não emprega o padrão complementar do estímulo desejado como controle alvo, enquanto a C-BAM e a C-BAM com atraso empregam o padrão complementar. Ao utilizar o método de controle por pinagem na modelagem do saltar passos, o objetivo foi aplicar o controle de caos apenas na camada que opera diretamente sobre o padrão desejado (o último padrão armazenado), uma vez que esta é a informação supostamente disponível. Sendo assim, a regra de controle para a C-BAM está incluída na camada em que a informação sobre a saída desejada está presente: a camada X. Neste caso, como o conjunto de padrões armazenados é formado por um grande número de bits, então a regra de controle de caos é aplicada em apenas alguns neurônios da rede, tal como na idéia original do método de controle por pinagem.

Nas duas situações de controle de caos realizadas pelo método de pinagem, os resultados sugeriram a efetividade desta estratégia de controle. O método de controle por pinagem proposto para a família C-BAM é facilmente implementado e apresenta rápida convergência para a saída desejada. No entanto, uma limitação na proposição deste método de controle é a necessidade da escolha aleatória, *a priori*, de dois parâmetros: a distância de pinagem (I) e a intensidade de controle (K). Estes parâmetros são interdependentes: para uma dada distância de pinagem o controle é realizado apenas quando a intensidade de controle é maior do que um limiar.

O método de controle adaptativo foi proposto na tentativa de reduzir limitações encontradas no método de controle por pinagem. Em sua proposta inicial, o método de controle adaptativo foi utilizado apenas para o controle de caos da C-BAM, no caso do fenômeno saltar passos.

Em comparação ao método de controle por pinagem, o método de controle adaptativo apresenta algumas vantagens: (a) apenas o parâmetro γ , que determina a taxa de adaptação, é previamente escolhido, (b) o tempo de cálculo computacional é menor para padrões armazenados de grande extensão, como é o caso dos conjuntos de padrões utilizados no fenômeno saltar passos, e (c) o próprio padrão desejado é utilizado como padrão referência, não há necessidade de se utilizar o padrão inverso. É importante lembrar que no caso de conjunto de padrões com muitos bits, o período de tempo gasto na computação do método de controle por pinagem é considerável, pois muitos cálculos são necessários.

No entanto, a versão do método de controle adaptativo proposta neste trabalho também apresenta uma limitação. O controle é realizado variável a variável e como cada erro e_j é calculado na camada Y da C-BAM, e_j encontra-se relacionado com o neurônio x_i . Sendo assim, para que o cálculo deste controle seja possível, é necessário que o tamanho i da camada X seja do mesmo tamanho j da camada Y.

Na modelagem do fenômeno saltar passos, pelo fato de os padrões da camada X e da camada Y serem todos de mesmo tamanho, uma vez que constituem uma seqüência de estágios de operações algébricas, tem-se que $i=j$, de modo que o método de controle adaptativo proposto para a C-BAM pode ser utilizado de modo prático, conveniente e eficiente no controle da modelagem do fenômeno saltar passos.

7.2 Resumo das Contribuições Científicas da Tese

As principais contribuições desta tese estão abaixo relacionadas:

Criação de uma família de redes hetero-associativas bidirecionais caóticas, a família C-BAM. No Capítulo 4 desta tese estão apresentadas as redes que fazem parte da família C-BAM criada: a rede C-BAM, a rede C-BAM com atraso e a rede C-BAM exponencial (C-eBAM). O novo paradigma, das redes de memórias associativas bidirecionais caóticas, representa uma extensão das redes de memória auto-associativas propostas por Adachi e Aihara (1997). Tal extensão significa uma contribuição para maior aprofundamento nos estudos das influências dos fenômenos caóticos em memórias associativas;

- (i). **Análise de comportamentos dinâmicos emergentes apresentados pelas redes da família C-BAM.** A fim de possibilitar a análise e o entendimento da dinâmica hetero-associativa das redes C-BAM de maneira simples, foram compostas redes C-BAM com 25 neurônios caóticos (15 na camada X e 10 na camada Y) e analisadas suas dinâmicas para grupos de padrões armazenados. Como resultado, foram relatados diversos comportamentos e propriedades interessantes, descobertos por

simulações, tais como o caos determinístico, bifurcação, crise, transição e recuperação de todos os padrões armazenados. A análise da ocorrência de bifurcação e crise nos experimentos realizados constitui uma ampliação de investigação comportamental nas redes associativas caóticas, uma vez que, nos artigos englobando memória auto-associativa caótica, apenas o caos determinístico foi destacado. Além disso, foi incluído, nas análises dos comportamentos, o planejamento fatorial de experimentos como uma ferramenta capaz de auxiliar na escolha da configuração paramétrica das redes C-BAM. Através do planejamento fatorial, a escolha dos valores dos parâmetros para se obter um determinado comportamento tornou-se mais objetiva e formal, não deixando que os caminhos possíveis para um determinado comportamento fossem escolhidos apenas por experimentação. O Capítulo 5, Seção 5.1, apresenta a análise do comportamento dinâmico da família C-BAM;

- (ii). **Demonstração da existência de caos nas redes da família C-BAM.** A sugestão da presença de caos na família C-BAM pelos experimentos com ela realizados, foi demonstrada através do cálculo dos expoentes de Lyapunov. Resultados sobre este tópico estão apresentados no Capítulo 5, Seção 5.2;
- (iii). **Estudo da acessibilidade de memórias treinadas e armazenadas pelas redes da família C-BAM.** Em contraste com as redes BAM convencionais, as quais se estabilizam em um estado de equilíbrio estável, que pode ser um padrão armazenado ou um padrão espúrio, as redes C-BAM mostraram grande diversidade de padrões recuperados, envolvendo complexas transições entre os padrões memorizados e sem convergir para nenhum padrão em particular. Resultados sobre este tópico estão apresentados no Capítulo 5, Seção 5.1;
- (iv). **Mostrar a aplicabilidade da dinâmica caótica no acesso a memórias não recuperáveis pelas redes BAM convencionais.** Conjuntos de padrões armazenados que apresentavam memórias não recuperáveis para as redes BAM convencionais foram analisados em experimentos com a família C-BAM. Os resultados mostraram que a família C-BAM é capaz de recuperar memórias antes não recuperáveis, através da dinâmica caótica presente nas redes desta família, que permitem que as redes escapem de mínimos locais indesejáveis e mostrem uma nova sensibilidade de acesso a memórias. Resultados sobre este tópico estão apresentados no Capítulo 5, Seção 5.3;
- (v). **Proposição da modelagem do fenômeno psicológico chamado saltar passos a partir da aplicabilidade da dinâmica caótica das redes C-BAM.** Esta proposição

se inclui no objetivo de utilizar o conhecimento adquirido das dinâmicas da C-BAM para alcançar uma meta desejada. Foi possível mostrar que a nova maneira de a C-BAM acessar os padrões armazenados pode ser usada para modelar a habilidade de um indivíduo de saltar passos na solução de uma tarefa. Desse modo, os resultados dessa modelagem podem ser verificados no Capítulo 5, Seção 5.4.

- (vi). **Proposição de uma estratégia de controle de caos para lidar com a dinâmica caótica das redes C-BAM.** O Capítulo 6 apresentou como o método de controle por pinagem permite à família C-BAM convergir para memórias desejadas e, assim, completar com sucesso seu processo de recuperação, que inclui a recuperação de memórias inacessíveis às redes BAM convencionais (Seção 6.1). Ainda no Capítulo 6, Seção 6.2, foram propostos dois métodos de controle, o método de controle por pinagem e o método de controle adaptativo, os quais permitiram à rede C-BAM estabilizar-se no último estado de uma seqüência, quando este estado é visitado.

Analisando-se as contribuições enumeradas acima, acredita-se que a meta geral desta tese, que foi a de contribuir para um maior aprofundamento no estudo das influências dos fenômenos caóticos em modelos de memórias associativas, tenha sido alcançada, pelo menos em parte. A consolidação, a longo prazo, das contribuições vindas da criação e dos estudos realizados com os modelos de memórias associativas bidirecionais caóticas depende de mais estudos teóricos e de um elenco maior de testes e provas analíticas, em adição aos apresentados neste documento. A seção seguinte relaciona algumas das próximas etapas planejadas e visualizadas para um maior embasamento, aprimoramento e diversificação de aplicações da família C-BAM.

7.3 Propostas para Novos Trabalhos

Esta seção sugere algumas opções para futuras pesquisas com a família C-BAM, as quais foram planejadas a partir de limitações observadas e da visualização de tópicos interessantes ainda não desenvolvidos. As principais extensões deste trabalho devem incluir:

- (i). **Universalização dos comportamentos e resultados alcançados para memórias hetero-associativas a outros tipos de redes neurais que modelem memórias associativas.** O fenômeno caótico, originalmente modelado para redes auto-associativas foi estendido com sucesso a memórias hetero-associativas. Como consequência, todos os comportamentos e resultados obtidos neste trabalho devem ser extensivos a outros tipos de redes neurais artificiais que também modelam memórias associativas.

- (ii). **Estudo de novos tipos de comportamentos dinâmicos emergindo da família C-BAM.** Os experimentos computacionais realizados com a família C-BAM indicaram a presença de diversos períodos, tanto pares quanto ímpares, conduzindo a dinâmicas caóticas e intercalando-as. Nesta exibição de comportamentos dinâmicos complexos podem ser pesquisados outros fenômenos caóticos além dos considerados e ilustrados neste trabalho, tais como intermitência e outros tipos de bifurcações e crises. Deve também ser investigada uma forma de aperfeiçoar a visualização dos fenômenos emergentes na família C-BAM.
- (iii). **Aplicação do fenômeno saltar passos em outras redes da família C-BAM e em novos domínios.** Neste trabalho, o fenômeno saltar passos foi modelado apenas para a rede C-BAM e aplicado no caso de solução de operações algébricas. Para consolidar esta aplicabilidade, novos experimentos podem incluir a C-BAM com atraso e a C-eBAM, envolver tarefas de diferentes domínios e experimentar outras explicações do saltar passos, que não seja a composição de regras.
- (iv). **Análise qualitativa e quantitativa do fenômeno saltar passos quando todos os parâmetros caóticos variam.** O objetivo é expandir as análises do saltar passos para casos em que todos os parâmetros caóticos variam e não apenas o parâmetro k_r . Nesta análise poderia ser incluída a investigação da ocorrência de (a) repetições de pares de padrões entre os passos iniciais e finais, (b) operações de transposição, (c) recombinações de pares armazenados.
- (v). **Realização de provas analíticas dos diferentes fenômenos caóticos emergentes na família C-BAM.** Expandir a formalização da existência do caos para os diferentes fenômenos caóticos encontrados na família C-BAM.
- (vi). **Expansão do método de controle adaptativo para o caso de camadas de tamanhos diferentes.** Aprimorar o método de controle adaptativo de modo que o atual controle variável a variável não exija que as camadas X e Y das redes C-BAMs tenham o mesmo tamanho. Tal expansão permitirá que novos contextos de modelagem possam ser modelados pelas redes de memória associativas bidirecionais caóticas.
- (vii). **Extensão do método de controle adaptativo às outras redes da família C-BAM.** A princípio proposto apenas para a rede C-BAM, o método de controle adaptativo pode ser estendido à C-BAM com atraso e à C-eBAM.
- (viii). **Desenvolvimento da prova analítica da estabilidade dos métodos de controle de caos.** As propostas de métodos de controle de caos para a família C-BAM (método de controle por pinagem e método de controle adaptativo) tiveram um enfoque

experimental. Uma próxima etapa incluiria a prova analítica da estabilidade dos métodos de controle de caos aqui propostos.

- (ix). **Embasamento das modelagens realizadas com a família C-BAM dentro de um contexto neurofisiológico.** Este trabalho mostrou como o caos modifica o mecanismo de acessibilidade à memória e permite que a família C-BAM recupere padrões inacessíveis pelo processo de associação convencional. Futuras investigações devem incluir o estudo da relação entre este efeito e a memória de seres humanos. Isto possibilitaria incorporar aplicações biológicas ao problema de acessibilidade da memória.

Referências Bibliográficas

- Adachi, M. e Aihara, K. (1997). Associative dynamics in a chaotic neural network. *Neural Networks*, vol. 10, nº 1, pp.:83-98.
- Aihara, K. e Matsumoto, G. (1983). Two stable steady states in the Hodgkin-Huxley axons. *Biophysical Journal*, vol. 41, nº 1, pp.:87-89.
- Aihara, K. e Matsumoto, G. (1986). Chaotic oscillations and bifurcations in squid giant axons. In *Chaos*, A. V. Holden, Ed. Manchester, U.K.: Manchester Univ. Press, pp.: 257-269.
- Aihara, K., Takabe, T. e Toyoda, M. (1990). Chaotic Neural Networks. *Physics Letters A*, vol. 144, nº 6-7, pp.:333-340.
- Aihara, K. (2002). Chaos engineering and its application to parallel distributed processing with chaotic neural networks. *Proceedings of the IEEE*, vol. 90, nº 5, pp.:919-930.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. e Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology: learning, memory and cognition*, vol. 20, nº 6, pp.:1322-1340.
- Andrievskii, B.R. e Fradkov, A.L. (2003). Control of chaos: Methods and applications. I. Methods, *Automation and Remote Control*, vol. 64, nº 5, pp.:673-713.
- Anton, H. e Rorres, C. (2001). *Álgebra linear com aplicações*. 8ª ed. Porto Alegre: Bookman.

- Araújo, A.F.R. (1988). Caos em sistemas de controle com modelos discretos. *Dissertação (Mestrado)*, Faculdade de Engenharia Elétrica, UNICAMP, Campinas, SP.
- Araújo, A.F.R. (1994). Memory, Emotions, and Neural Networks: Associative Learning and Memory Recall Influenced by Affective Evaluation and Task Difficulty. *Ph.D. dissertation*, School of Cognitive and Computing Sciences, University of Sussex, Brighton, UK.
- Astrom, K.J. e Wittenmark, B. (1995). *Adaptive Control*. 2nd edition. Addison-Wesley Publishing Company, USA.
- Babloyantz, A. e Kaczmarek, L.K. (1977). Spatiotemporal patterns in epileptic seizures. *Biological Cybernetics*, vol.26, n° 4, pp.:199-208.
- Babloyantz, A. (1991). Probing dynamics of the cerebral cortex. In *Complexity, Chaos and Biological Evolution*, Mosekild E. e Mosekilde L. (Eds.), Plenum press, New York.
- Baldi, P. e Atiya, A. F. (1994). How delays affect neural dynamics and learning. *IEEE Transactions on Neural Networks*, vol.5, n° 4, pp.:612-617.
- Basar, E. (1980). *EEG-Brain Dynamics*. Elsevier North-Holland, Amsterdam.
- Basar, E., Basar-Eroglu, C. e Roschke. J. (1988). Do coherent patterns of the strange attractor EEG reflect deterministic sensory-cognitive states of the brain? In *From Chemical to Biological Organization*, Markus M., Muller S.C. e Nicolis G. (Eds.), Springer, Berlin, pp.:297-306.
- Basar, E., Basar-Eroglu, C., Roschke. J. e Schultz J. (1989). Chaos and alpha-preparation in brain function. In *Models of Brain Function*, Cotterill R.M.J. (Ed.), Cambridge: Cambridge Univ. Press, U.K., pp.:365-395.
- Belohlavek, R. (2000). Fuzzy logical bidirectional associative memory. *Information Sciences*, vol. 128, n° 1-2, pp.:91-103.
- Blessing, S. B. e Anderson, J. R. (1996). How People Learn to Skip Steps. *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 22, n° 3, pp.:576-598.
- Caianiello, E. R. (1961). Outline of a theory of thought processes and thinking machines. *Journal of Theoretical Biology*, vol. 1, n° 2, pp.:204-235.
- Carlson, R. A, e Lundy, D. H. (1992). Consistency and restructuring in learning cognitive procedural sequences. *Journal of Experimental Psychology: learning, memory and cognition*, vol. 18, n° 1, pp.:127-141.

- Chua, L. O., Wu, C. W., Huang, A. e Zhong, G.–Q. (1993). A universal circuit for studying and generating chaos – Part I: Routes to chaos”, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 40, n° 10, pp.:732-744.
- Cole, K.S., Guttman, R. e Bezanilla, F. (1970). Nerve Membrane Excitation without Threshold. *Proc. National Academy of Sciences (USA)*, vol. 65, n° 4, pp.:884-891.
- Destexhe, A., Sepulchre, J.A. e Babiloyantz, A. (1988). A comparative study of the experimental quantification of deterministic chaos. *Physics Letters A*, vol.132, n° 2-3, pp.:101-106.
- Ding, M., Grebogi, C. e Ott, E. (1994). Controlling chaos in a temporally irregular environment. *Physica D:Nonlinear Phenomena*, vol. 74, n° 3-4, pp.:386-394.
- Eisenber, J., Freeman, W.J. e Burke, B. (1989). Hardware architecture of a neural network model simulating pattern recognition by the olfactory bulb. *Neural Networks*, vol.2, n° 4, pp.:315-325.
- Faure, P. e Korn, H. (2001). Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *Life Sciences*, 324, pp.:773-793.
- Ferrara, N. F. e Prado, C. P. C. (1995). *Caos: uma introdução*. São Paulo: Editora Edgard Blucher Ltda.
- Ford, J. (1989). What is chaos, that we should be mindful of it? In *The New Physics*, P. Davies (Ed.), Cambridge: Cambridge University Press, pp.:348-372.
- Fradkov, A.L.e Evans, R.J. (2005). Control of chaos: Methods and applications in engineering, *Annual Reviews in Control*, vol. 29, n° 1, pp.:33-56.
- Freeman, W.J. (1987). Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biological Cybernetics*, vol.56, n° 2-3, pp.:139-150.
- Freeman, W. J. e Van Dijk, B. W. (1987). Spatial patterns of visual cortical fast EEG during conditioned reflex in a rhesus monkey. *Brain Research*, vol. 422, n° 2, pp.:267-276.
- Freeman, W. J., Yao, Y. e Burke, B. (1988). Central pattern generating and recognizing in olfactory bulb.: A correlation learning rule. *Neural Networks*, vol. 1, n° 4, pp.:277-288.
- Freeman, W. J. (1991). The physiology of perception. *Scientific American*, vol. 264, n° 2, pp.:78-85.

- Freeman, W. J. (1992). Tutorial on neurobiology: from single neuron to brain chaos. *International Journal of Bifurcation and Chaos*, vol. 2, nº 3, pp.:451-482.
- Gardner, E. (1988). The space of interactions in neural network models. *Journal of Physics A: Mathematical and General*, vol. 21, nº 1, pp.:257-270.
- Gleick, J. (1989). *Caos: a criação de uma nova ciência*. Rio de Janeiro: Editora Campus.
- Grebogi, C., Ott, E. e Yorke, J.A. (1982). Chaotic attractors in crisis. *Physical Review Letters*, vol. 48, nº 22, pp.:1507-1510.
- Grebogi, C., Ott, E. e Yorke, J.A. (1983). Crises, sudden changes in chaotic attractors, and transient chaos. *Physica D: Nonlinear Phenomena*, vol. 7, nº 1-3, pp.:181-200.
- Hale, J. e Koçak, H. (1991). *Dynamics and bifurcations*. Springer-Verlag, New York.
- Haykin, S. (1999). *Neural networks: a comprehensive foundation*. 2nd edition. New Jersey.
- He, Z., Zhang, Y.F. e Yang, L.X. (1999). The study of chaotic neural network and its applications in associative memory. *Neural Processing Letters*, vol. 9, nº 2, pp.:163-175.
- He, G., Cao, Z., Zhuo, P. e Ogura, H. (2003). Controlling chaos in a chaotic neural network. *Neural Networks*, vol. 16, nº 8, pp.:1195-1200.
- Hilborn, R. C. (1994). *Chaos and nonlinear dynamics: an introduction for scientists and engineers*. Oxford USA Trade.
- Hindes, P. J., Patterson, M. e Pfeffer, J. (2001). Bothered by abstraction: the effect of expertise on knowledge transfer and subsequent novice performance. *Journal of Applied Psychology*, vol. 86, nº 6, pp.: 1232-1243.
- Hodgkin, A. L. e Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, vol. 117, nº 4, pp.:500-544.
- Hu, G. e Qu, Z. (1994). Controlling spatiotemporal chaos in coupled map lattice systems. *Physical Review Letters*, vol. 72, nº 1, pp.:68-71.

- Ikeguchi, T., Hasegawa, M. e Aihara, K. (1999). An analysis on associative dynamics of chaotic neural networks by global and local Lyapunov exponents. *Proceeding of the 6th International Conference on Neural Information Processing (ICONIP 1999)*, vol. 3, pp.:1017-1022.
- In, V., Spano, M.L., Neff, J.D., Ditto, W.L., Daw, S., Edwards, K.D. e Nguyen, K. (1997). Maintenance of chaos in a computational model of a thermal pulse combustor. *Chaos*, vol. 7, n° 4, pp.:605-613.
- Jain, R. (1991). *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*. USA: John Wiley & Sons, Inc.
- Jeng, Y.-J., Yeh, C.-C. e Chiveh, T. D. (1990). Exponential bidirectional associative memories. *Electronics Letters*, vol. 26, n° 11, pp.:717-718.
- Koedinger, K. R. e Anderson, J. R. (1990). Abstract planning and perceptual chunks: elements of expertise in geometry. *Cognitive Science*, vol. 14, n° 4, pp.:511-550.
- Komuro, M. e Aihara, K. (2001). Hierarchical structure among invariant subspaces of chaotic neural networks. *Japan Journal of Industrial and Applied Mathematics*, vol. 18, n° 2, pp.:335-357.
- Kosko, B. (1987). Adaptive bidirectional associative memories. *Applied Optics*, vol. 26, n° 23, pp.:4947-4960.
- Kosko, B. (1988). Bidirectional associative memories. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 18, n° 1, pp.:49-60.
- Kosko, B. (1992). *Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence*. New Jersey, USA: Prentice Hall.
- Kostelich, E.J., Gregobi, C., Ott, E. e Yorke, J.A. (1993). Higher-dimensional targeting. *Physical Review E*, vol. 47, n° 1, pp.:305-310.
- Kozma, R. e Freeman, W.J. (2000). Encoding and recall of noisy data as chaotic spatio-temporal memory patterns in the style of the brains. *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Como, Italy, vol. 5, pp.:33-38.
- Lewis, C. H. (1978). Production system models of practice effects. *Ph.D. dissertation*, Department of Psychology, University of Michigan, Ann Arbor.

- Lewis, C. H. (1981). Skill in algebra. In *Cognitive skills and their acquisition*, J. R. Anderson (Ed.), Hillsdale, NJ: Erlbaum.
- Li, C. e Chen, G. (2004). Estimating the Lyapunov exponents of discrete systems. *Chaos*, vol. 14, n° 2, pp.:343-346.
- Liao, X.F. e Yu, J.B. (1998). Qualitative analysis of bi-directional associative memory with time delay. *International Journal of Circuit Theory and Applications*, vol. 26, n° 3, pp.:219-229.
- Lima, E. L. (2001). *Álgebra Linear*. 5ª ed. Rio de Janeiro: IMPA.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, vol. 95, n° 4, pp.:492-527.
- Logan, G. D. (1990). Repetition priming and automaticity: common underlying mechanisms? *Cognitive Psychology*, vol. 22, n° 1, pp.:1-35.
- Maass, W. (1997). Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, vol. 10, n° 9, pp.:1659-1671.
- Mahla, A. A. I. (1991). Caos em sistemas de controle de motores C.C. pela técnica PLL-DUAL. *Dissertação (Mestrado)*, Faculdade de Engenharia Elétrica, UNICAMP, Campinas, SP.
- Matsumoto, G., Takahashi, N. e Hanyu, Y. (1987). In *Chaos in biological systems*. Degn, H., Holden, A.V. e Olsen L.F. (Eds.), Plenum, New York, pp.:143-156.
- McCulloch, W. S. e Pitts, W. H. (1943). A logical calculus of the ideas immanent in neural nets. *Bulletin of Mathematical Biophysics*, vol. 5, pp.:115-133.
- McEliece, R.J., Posner, E.D., Rodemich, E.R. e Venkatesh, S.S. (1987). The capacity of the Hopfield associative memory. *IEEE Transactions on Information Theory*, vol. IT-33, n° 4, pp.:461-482.
- Montgomery, D. C. (1997). *Design and analysis of experiments*. 4th ed. Wiley, New York.
- Montgomery, D.C. e Runger, G.C. (2003). *Estatística aplicada e probabilidade para engenheiros*. 2ª edição. Rio de Janeiro: LTC Editora.
- Nagumo, J. e Sato, S. (1972). On a response characteristic of a mathematical neuron model, *Biological Cybernetics*, vol. 10, n° 3, pp.:155-164.
- Newell, A. e Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Ott, E., Grebogi, C. e Yorke, J.A. (1990a). Controlling chaos. *Physical Review Letters*, vol. 64, n° 11, pp.:1196- 1199.
- Ott, E., Grebogi, C. e Yorke, J.A. (1990b). Controlling chaotic dynamical systems. In *Chaos*, D.K. Campbell (Ed.), American Institute of Physics, New York.
- Ott, E. (2002). *Chaos in dynamical systems*. 2nd edition. Cambridge University Press, UK.
- Parker, T. S. e Chua, L. O. (1989). *Practical Numerical Algorithms for Chaotic Systems*. New York: Springer-Verlag.
- Pistolese, G. (1996). The role of chaos in memory: the concepts of short-term memory and long-term memory systems revisited. *Cybernetica* , vol. XXXIX, n° 1, pp.:31-47.
- Potapov, A.B. e Ali, M.K. (2001). Nonlinear dynamics and chaos in information processing neural networks. *Differential Equations and Dynamical Systems*, vol.9, n° 3-4, pp.:259-319.
- Raffone, A. e Leeuwen, C. (2003). Dynamic synchronization and chaos in an associative neural network with multiple active memories, *Chaos*, vol. 13, n° 3, pp.:1090-1104.
- Rosenbloom, P. S. e Newell, A. (1986). The chunking of goal hierarchies: A generalized model of practice. In Michalski, R.S., Carbonell, J.G., e Mitchell, T.M., editors, *Machine learning: An artificial intelligence approach*, vol. 2, pp.:247-288. Morgan Kaufmann Publishers, Inc., Los Altos, California, 1986.
- Rumelhart, D. E., Hinton, G. E. e Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, vol. 323, n° 6088, pp.:533-536.
- Sandler, Y.M. (1990). Model of neural networks with selective memorization and chaotic behavior. *Physics Letters A*, vol. 144, n° 8-9, pp.:462-466.
- Shaw, R. (1984). *The Dripping Faucet as a Model Chaotic System*. Santa Cruz, California: Aerial Press.
- Shiino, M. e Fukai, T. (1990). Chaotic image retrieval in Markovian asymmetric neural networks with sign-constrained synaptic couplings. *Journal of Physical Society of Japan*, vol. 59, n° 5, pp.:1529-1532.

- Shimada, I. e Nagashima, T. (1979). A numerical approach to ergodic problem of dissipative dynamical systems. *Progress of Theoretical Physics*, vol. 61, n° 6, pp.:1605-1616.
- Shinbrot T., Ott, E., Grebogi, C. e Yorke, J.A. (1990). Using chaos to direct trajectories to targets. *Physical Review Letters*, vol. 65, n° 26, pp.:3215-3218.
- Shinbrot, T., Ditto, W., Grebogi, C., Ott, E., Spano, M. e Yorke, J.A. (1992). Using the sensitive dependence of chaos (the “butterfly effect”) to direct trajectories in an experimental chaotic system. *Physical Review Letters*, vol. 68, n° 19, pp.:2863-2866.
- So, P., Ott, E., Sauer, T e Gluckman, B.J. (1997). Extracting unstable periodic orbits from chaotic time series. *Physical Review E*, vol. 55, n° 5, pp.:5398-5417.
- Spiegel, M. R. (1978). *Probabilidade e estatística*. São Paulo: McGraw-Hill do Brasil.
- Stewart, I. (1991). *Será que Deus joga dados?: a nova matemática do caos*. Jorge Zahar Editora.
- Tan, Z. e Ali, M. K. (1998). Pattern recognition in a neural network with chaos. *Physical Review E*, vol.58, n° 3, pp.:3649-3653.
- Tani, J. e Fujita, M. (1992). Coupling of memory search and mental rotation by a nonequilibrium dynamics neural network. *IEICE Transactions on Fundamentals*, vol. E75-A, n° 5, pp.:578-585.
- Tsuda, I. (1992). Dynamic link of memory – Chaotic memory map in nonequilibrium neural networks. *Neural Networks*, vol. 5, n° 2, pp.:313-326.
- Wang, X. (1991). Period-doublings to chaos in a simple neural network: an analytical proof. *Complex Systems*, vol.5, n° 4, pp.:425-444.
- Wang, W.J. e Lee, D.L. (1993). A modified bidirectional decoding strategy based on the BAM structure. *IEEE Transactions on Neural Network*, vol. 4, n° 4, pp.:710-717.
- Wang, C. C. e Don, H. S. (1995). An analysis of high-capacity discrete exponential BAM. *IEEE Transactions on Neural Networks*, vol. 6, n° 2, pp.:492-496.
- Wenger, M. J. (1999). On the Whats and Hows of Retrieval in the Acquisition of a Simple Skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 25, n° 5, pp.:1137-1160.
- Williams, G. P. (1997). *Chaos theory tamed* . Washington, DC: Joseph Henri Press.

- Wilson, R. C. e Hancock, E. R. (2001). Storage Capacity of the exponential correlation associative memory. *Neural Processing Letters*, vol. 13, n° 1, pp.:71-80.
- Xiaofeng, L. e Juebang, Y. (1998). Qualitative analysis of bi-directional associative memory with time delay. *International Journal of Circuit Theory and Applications*, vol. 26, n° 3, pp.:219-229.
- Xu, Z.B., Leung, Y. e He, X.W. (1994). Asymmetric bidirectional associative memories. *IEEE Transactions on Systems Man and Cybernetics*, vol. 24, n° 10, pp.:1558-1564.
- Yang, W., Ding, M., Mandell, A., e Ott, E. (1995). Preserving chaos: Control strategies to preserve complex dynamics with potential relevance to biological disorders. *Physical Review E*, vol. 51, n° 1, pp.:102-110.
- Yao, Y. e Freeman, W. J. (1990). Model of biological pattern recognition with spatially chaotic dynamics. *Neural Networks*, vol. 3, n° 2, pp.:153-170.
- Yao, Y., Freeman, W. J., Burke, B., e Yang, Q. (1991). Pattern recognition by a distributed neural network: An industrial application. *Neural Networks*, vol. 4, n° 1, pp.:103-121.
- Zhang, J. e Yang, Y. (2001). Global stability analysis of bidirectional associative memory neural networks with time delay. *International Journal of Circuit Theory and Applications*, vol. 29, n° 2, pp.:185-196.
- Zhou, R.W. e Quek, C. (1996). DCBAM: a discrete chainable bidirectional associative memory. *Pattern Recognition Letters*, vol. 17, n° 9, pp.:985-999.

Apêndice A

Experimento Fatorial

O objetivo deste Apêndice é oferecer mais detalhes da aplicação da metodologia utilizada nesta tese para formalizar a importância de cada parâmetro da família C-BAM.

As técnicas de planejamento de experimentos, baseadas em métodos estatísticos, permitem, entre outras numerosas aplicações, que se determine qual subconjunto de fatores do sistema tem maior influência no desempenho do processo. Fatores, no contexto de planejamento de experimentos, são definidos como as variáveis que afetam a resposta do sistema e que podem assumir valores alternativos.

Nesta tese, as técnicas de *planejamento de experimentos com vários fatores* foram aplicadas com o objetivo de se determinar quais parâmetros da família C-BAM são mais importantes numa dada circunstância e também quais deles são críticos para a obtenção de determinado comportamento.

A.1. Experimentos Fatoriais

No caso de haver vários fatores potencialmente importantes em um experimento, um planejamento fatorial de experimentos deve ser utilizado. Um experimento fatorial é aquele em que os fatores são variados em conjunto. Segundo Montgomery e Runger (2003), “por **um planejamento fatorial**, queremos dizer que, em cada tentativa completa ou réplica do experimento, todas as combinações possíveis dos níveis dos fatores são investigadas”.

O efeito de um fator é definido como a variação na resposta, produzida pela mudança do nível desse fator. Por exemplo, ao ser considerado um experimento fatorial com dois fatores A e B, cada um com dois níveis (A_{baixo} , A_{alto} , B_{baixo} , B_{alto}), o efeito principal do fator A é a diferença entre a resposta média no nível alto de A e a resposta média no nível baixo de A. Quando ocorre de a

diferença na resposta entre os níveis de um fator não ser a mesma em todos os níveis dos outros fatores, significa que existe uma *interação* entre fatores. Quando essa interação é significativa, ela pode ‘mascarar’ os efeitos individuais dos fatores. Conseqüentemente, o conhecimento da interação AB pode ser mais útil do que o conhecimento do efeito individual de cada fator.

Para Montgomery e Runger (2003), os experimentos fatoriais são a única maneira de descobrir interações entre as variáveis. Segundo esses autores, uma alternativa ao planejamento fatorial utilizada na prática é a mudança de um fator de cada vez, ao invés de variá-los simultaneamente. No entanto, esta alternativa é ineficiente, uma vez que não consegue detectar a interação entre os fatores. Ela necessita de mais experimentos do que um planejamento fatorial e não dá garantia de produzir resultados corretos.

Dentro da metodologia de experimentos fatoriais, cada experimento envolve uma seqüência de atividades:

Conjectura – a hipótese original que motiva o experimento;

Experimento – o teste feito para investigar a conjectura;

Análise – a análise estatística dos dados do experimento;

Conclusão – o que se aprendeu acerca da conjectura original do experimento.

Freqüentemente, o experimento conduzirá a uma conjectura revisada e a um novo experimento e assim por diante.

Como experimentos fatoriais gerais, que envolvem mais de dois fatores, requerem uma aritmética tediosa e computacionalmente confusa (Montgomery e Runger, 2003), para a ilustração do uso da metodologia de planejamento de experimentos na rede C-BAM será apresentado um experimento envolvendo apenas dois fatores (*experimento fatorial com dois fatores*).

O experimento fatorial com dois fatores enquadra-se no tipo mais simples desse tipo de experimento, justamente por envolver somente dois fatores. No entanto, entendida a aplicação dessa metodologia, sua ampliação para experimentos que envolvam mais de dois fatores torna-se fácil. No caso dos dois fatores, A e B, há a níveis do fator A e b níveis do fator B. Como conjectura, serão testadas as hipóteses de nenhum efeito para o fator A, nenhum efeito para o fator B e nenhum efeito para o AB. O experimento deve ter n réplicas, cada réplica contendo combinações ab . Em uma tabela de arranjo dos dados, a observação feita na ij -ésima célula para a k -ésima réplica é denotada por y_{ijk} .

A Tabela A.1 exemplifica como deve ser feito o arranjo dos dados para um planejamento fatorial com dois fatores:

Tabela A.1. Arranjo dos Dados para um Planejamento de Experimentos com Dois Fatores.

		Fator B					
		1	2	...	b	Totais	Médias
Fator A	1	$y_{111}, y_{112}, \dots,$ y_{11n}	$y_{121}, y_{122}, \dots,$ y_{12n}		$y_{1b1}, y_{1b2}, \dots,$ y_{1bn}	$y_{1..}$	$\bar{y}_{1..}$
	2	$y_{211}, y_{212}, \dots,$ y_{21n}	$y_{221}, y_{222}, \dots,$ y_{22n}		$y_{2b1}, y_{2b2}, \dots,$ y_{2bn}	$y_{2..}$	$\bar{y}_{2..}$
	.						
	.						
	.						
	a	$y_{a11}, y_{a12}, \dots,$ y_{a1n}	$y_{a21}, y_{a22}, \dots,$ y_{a2n}		$y_{ab1}, y_{ab2}, \dots,$ y_{abn}	$y_{a..}$	$\bar{y}_{a..}$
Totais	$y_{.1.}$	$y_{.2.}$		$y_{.b.}$	$y_{...}$		
Médias	$\bar{y}_{.1.}$	$\bar{y}_{.2.}$		$\bar{y}_{.b.}$		$\bar{y}_{...}$	

Com os dados da Tabela A.1 à disposição, é possível fazer a *análise da variância*, a qual poderá ser usada para testar hipóteses sobre os efeitos principais dos fatores A e B e da interação AB.

Alguns valores devem ser calculados para que se obtenha essa análise de variância. As representações do total das observações feitas no i -ésimo nível do fator A, $y_{i..}$, e suas médias correspondentes são dadas pelas Equações A.1 e A.2.

$$y_{i..} = \sum_{j=1}^b \sum_{k=1}^n y_{ijk} \quad (\text{A.1}) \quad \bar{y}_{i..} = \frac{y_{i..}}{bn} \quad (\text{A.2}) \quad i = 1, 2, \dots, a$$

As representações do total das observações feitas no i -ésimo nível do fator B, $y_{.j.}$, e suas médias correspondentes são dadas pelas Equações A.3 e A.4.

$$y_{.j.} = \sum_{i=1}^a \sum_{k=1}^n y_{ijk} \quad (\text{A.3}) \quad \bar{y}_{.j.} = \frac{y_{.j.}}{an} \quad (\text{A.4}) \quad j = 1, 2, \dots, b$$

As representações do total das observações feitas na ij -ésima célula da Tabela A.1, $y_{ij.}$, e suas médias correspondentes são dadas pelas equações A.5 e A.6.

$$y_{ij.} = \sum_{k=1}^n y_{ijk} \quad (\text{A.5}) \quad \bar{y}_{ij.} = \frac{y_{ij.}}{n} \quad (\text{A.6}) \quad i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b$$

As representações do total global de todas as observações, $y_{...}$, e suas médias correspondentes são dadas pelas Equações A.7 e A.8.

$$y_{...} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk} \quad (\text{A.7}) \quad \bar{y}_{...} = \frac{y_{...}}{abn} \quad (\text{A.8})$$

Conforme dito anteriormente, a análise da variância testará três hipóteses: a de nenhum efeito principal do fator A, a de nenhum efeito principal do fator B, e de nenhum efeito de interação AB. Para isso, tal procedimento de análise fará a decomposição da variabilidade total nos dados em partes componentes, comparando, então, os vários elementos nessa decomposição.

A *variabilidade total* é medida pela soma quadrática total das observações:

$$SQ_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2 \quad (\text{A.9})$$

Sendo que a decomposição da soma quadrática (ou seja, a decomposição da variabilidade total, a qual testa as hipóteses acima) é dada por:

$$\begin{aligned} SQ_T &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2 = bn \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 + an \sum_{j=1}^b (\bar{y}_{.j.} - \bar{y}_{...})^2 + \\ &+ n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij.})^2 \end{aligned} \quad (\text{A.10})$$

A equação acima é chamada de a identidade da soma quadrática para a análise de variância com dois fatores, e pode ser escrita simbolicamente como:

$$SQ_T = SQ_A + SQ_B + SQ_{AB} + SQ_E \quad (\text{A.11})$$

Nesta Equação A.11, SQ_A representa a soma quadrática devido ao fator A, SQ_B , a soma quadrática devido ao fator B, SQ_{AB} representa a soma quadrática devido à interação entre A e B, e SQ_E representa a soma quadrática devido ao erro.

As Equações A.12 a A.16 fornecem os cálculos para as **somas quadráticas** em uma análise de variância com dois fatores:

$$SQ_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{y_{...}^2}{abn} \quad (\text{A.12})$$

$$SQ_A = \sum_{i=1}^a \frac{y_{i..}^2}{bn} - \frac{y_{...}^2}{abn} \quad (\text{A.13})$$

$$SQ_B = \sum_{j=1}^b \frac{y_{.j.}^2}{an} - \frac{y_{...}^2}{abn} \quad (\text{A.14})$$

$$SQ_{AB} = \sum_{i=1}^a \sum_{j=1}^b \frac{y_{ij.}^2}{n} - \frac{y_{...}^2}{abn} - SQ_A - SQ_B \quad (\text{A.15})$$

$$SQ_E = SQ_T - SQ_{AB} - SQ_A - SQ_B \quad (\text{A.16})$$

Os efeitos principais de A, B, e o efeito de interação possuem, respectivamente, $(a-1)$, $(b-1)$ e $(a-1)(b-1)$ **graus de liberdade**. No total, existem $(abn-1)$ graus de liberdade. Dentro de cada uma das células ab da Tabela A.1, há $n-1$ graus de liberdade entre as n réplicas, podendo haver alguma diferença devido ao erro aleatório. Conseqüentemente, há $(ab)(n-1)$ graus de liberdade para o erro. Portanto, os graus de liberdade são divididos de acordo com:

$$abn-1 = (a-1) + (b-1) + (a-1)(b-1) + (ab)(n-1) \quad (\text{A.17})$$

Em geral, uma **média quadrática** é calculada dividindo-se uma soma quadrática por seu número de graus de liberdade. Sendo assim, as Equações A.18 a A.21 fornecem as médias quadráticas do planejamento de experimentos em questão.

$$MQ_A = \frac{SQ_A}{a-1} \quad (\text{A.18}) \qquad MQ_B = \frac{SQ_B}{b-1} \quad (\text{A.20})$$

$$MQ_{AB} = \frac{SQ_{AB}}{(a-1)(b-1)} \quad (\text{A.19}) \qquad MQ_E = \frac{SQ_E}{ab(n-1)} \quad (\text{A.21})$$

Finalmente, para testar o fato de os efeitos dos fatores nas linhas serem iguais a zero, é utilizada a razão: $F_0 = MQ_A/MQ_E$, que terá distribuição f , com $(a-1)$ e $ab(n-1)$ graus de liberdade se a hipótese nula for verdadeira. A hipótese nula será rejeitada com um nível de significância μ se $f_0 > f_{\mu, (a-1), ab(n-1)}$. Testes similares devem ser feitos para as outras hipóteses nulas: $F_0 = MQ_B/MQ_E$ e $F_0 = MQ_{AB}/MQ_E$.

A.2. Aplicação de Experimentos Fatoriais

Antes do início das aplicações do experimento fatorial às redes C-BAM, foram necessárias adequações de certos procedimentos exigidos no teste.

Cada réplica do experimento fatorial deve ser apresentada em uma representação numérica decimal. No entanto, as saídas das redes C-BAMs, que constituem as réplicas do planejamento de experimentos, possuem representação numérica binária. Sendo assim, antes da aplicação do planejamento de experimentos deve ser realizada uma conversão entre bases binárias e decimais.

A conversão de um número na base binária para uma base decimal foi realizada seguindo-se a ordem de grandeza da posição do número. Deste modo, dado um número binário, para obter este valor na base decimal, foi feita uma multiplicação de seus dígitos, da direita para esquerda, por potências crescentes de 2. O início dessa potência é obtido dividindo-se o número de bits do par associado por 2. Seja chamado de n o número inteiro resultante da divisão. Então, o bit mais à direita do número binário é multiplicado por 2^{-n} . Por exemplo, se o número de bits do par associado da rede caótica for 25, então $n=12$ e o bit mais à direita será multiplicado por 2^{-12} . Em seguida, da

direita para esquerda, os outros bits continuarão a ser multiplicados por 2^{-11} , 2^{-10} , ..., 2^0 , 2^1 , 2^2 , ..., 2^{12} . Finalmente, os resultados da multiplicação dos dígitos pelas potências devem ser somados. A soma total será o número decimal equivalente ao número binário. Para ilustração desta conversão, seja tomada uma saída da rede C-BAM, o par $(\mathbf{X}^k, \mathbf{Y}^k) = (1\ 1\ 1\ -1\ -1\ -1\ 1\ 1\ 1\ -1\ -1\ -1\ 1\ 1\ 1, 1\ 1\ -1\ -1\ 1\ -1\ -1\ 1\ 1)$. O processo de conversão passo-a-passo é apresentado na Tabela A.2.

Tabela A.2. Método utilizado para converter um par associado da C-BAM (binário) em um número decimal, a fim de que ele seja utilizado como réplica no planejamento de experimentos.

1	1	1	0	0	0	1	1	1	0	0	0	1	1	1
2^{12}	2^{11}	2^{10}	2^9	2^8	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	2^{-1}	2^{-2}
1x	1x	1x	0x	0x	0x	1x	1x	1x	0x	0x	0x	1x	1x	1x
4096	2048	1024	512	256	128	64	32	16	8	4	2	1	0.5	0.25
4096	2048	1024	0	0	0	64	32	16	0	0	0	1	0.5	0.25
4096 + 2048 + 1024 + 64 + 32 + 16 + 1 + 0.5 + 0.25 = 7281.25														

1	1	0	0	1	1	0	0	1	1
2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}	2^{-9}	2^{-10}	2^{-11}	2^{-12}
1x	1x	0x	0x	1x	1x	0x	0x	1x	1x
0.125	0.0625	0.03125	0.015625	0.00781	0.00391	0.001953	0.000976	0.000488	0.000244
0.125 + 0.0625 + 0.00781 + 0.00391 + 0.000488 +						TOTAL = 7281.25000			
0.000244 = 0.199952						0.199952 +			

						7281.449952			

Como pode ser visto na Tabela A.2, o par $(\mathbf{X}^k, \mathbf{Y}^k)$ equivale ao número decimal 7281.449952. Pelo fato de o conjunto de padrões utilizados nos experimentos desta tese constarem de um grande número de bits, o número decimal resultante tende a ser um valor muito alto. Para evitar altos valores no planejamento de experimentos, foi feita uma divisão por 1000 em todos os números decimais resultantes da conversão numérica.

Além deste, outros ajustes foram necessários para que os experimentos fatoriais fossem devidamente aplicados. Um segundo ajuste envolveu o número de réplicas, ou seja, o número de saídas que a rede fornece para dado nível de fator A *versus* dado nível de fator B. Como as redes C-BAMs são determinísticas, um único valor de saída é gerado para dado conjunto de valores de parâmetros com mesma condição inicial. No entanto, para que o planejamento de experimentos com vários fatores possa testar suas hipóteses a respeito dos efeitos dos fatores e suas interações de forma correta, são necessárias pelo menos duas réplicas para cada nível de fator A *versus* nível de

fator B. Para contornar tal situação, a partir da réplica exata resultante da saída da rede caótica, uma segunda réplica foi gerada aleatoriamente, somando-se ou subtraindo-se valores entre [0,0.1] da réplica original. Por exemplo, sendo 260,985 o valor de uma réplica original, uma segunda réplica, diante do sorteio aleatório do número -0.07 , teria o valor 260,915.

Um último ajuste dos dados a serem utilizados no planejamento fatorial envolveu o número de bits dos números binários das réplicas. Como as redes C-BAMs apresentam diferentes comportamentos emergentes, tais como períodos, bifurcação, caos, crise; as réplicas tiveram que se adequar aos diferentes comportamentos dinâmicos. Por exemplo, se para dado nível de fator A *versus* dado nível de fator B fosse encontrado um período 2 equivalente a $\{(X^1, Y^1) + (X^4, Y^4)\}$ e, para outro nível de fator A *versus* nível de fator B fosse encontrado um período 4 equivalente a $\{(X^3, Y^1) + (X^1, Y^1) + (X^3, Y^3) + (X^1, Y^1)\}$, não seria possível trabalhar com uma réplica contendo 50 bits binários e a outra contendo 100 bits binários (considerando-se o caso em que cada par associado tem 25 bits binários). Para haver equivalência no número de bits das réplicas, foi definido que os bits de uma réplica teriam o tamanho dos bits da réplica de maior período. Sendo assim, se o maior período encontrado nos a níveis do fator A *versus* os b níveis do fator B foi o período 16, então cada réplica foi computada como $\{(16/\text{período}) * (\text{soma dos pares})\} / 16$ (Tabela A.3).

Tabela A.3. Formatação do tamanho das réplicas de um experimento fatorial para o caso de período máximo 16.

Período 1	Período 2	Período 3	Período 8	Período 16
$\{(16/1) * (\text{par1})\} / 16$ = par1	$\{(16/2) * (\text{par1} + \text{par2})\} / 16$ = $\{8 * (\text{par1} + \text{par2})\} / 16$ = $(\text{par1} + \text{par2}) / 2$	$\{(16/3) * (\text{par1} + \text{par2} + \text{par3})\} / 16$		$\{(16/8) * (\text{par1} + \dots + \text{par8})\} / 16$ = $\{2 * (\text{par1} + \dots + \text{par8})\} / 16$ = $(\text{par1} + \dots + \text{par8}) / 8$	$\{(16/16) * (\text{par1} + \dots + \text{par16})\} / 16$ = $\{1 * (\text{par1} + \dots + \text{par16})\} / 16$ = $(\text{par1} + \dots + \text{par16}) / 16$

Na ocorrência de caos, o tamanho da réplica seguiu o período máximo encontrado no conjunto de níveis de fatores considerado. No caso de período máximo 16, escolheu-se uma seqüência de 16 pares dentro do caos e o valor da réplica foi calculado sobre eles.

A Tabela A.4 mostra o arranjo dos dados para um planejamento fatorial com dois fatores feito para a rede C-BAM. Neste experimento, tanto o fator A (parâmetro ρ) quanto o fator B (parâmetro k_r) tiveram três níveis considerados ($\rho = 2, 4$ e 8 ; $k_r = 0.3, 0.6$, e 0.7). O número de réplicas do experimento foi $n=2$.

Tabela A.4. Dados para um experimento fatorial com 2 fatores ($\rho = 2, 4$ e 8 ; $k_r = 0.3, 0.6$, e 0.7) realizado com a rede C-BAM. Os parâmetros k_f e a foram mantidos fixos.

		Fator B = k_r				
		0.3	0.6	0.7	Totais	Médias
Fator A = ρ	2	Período 2 6.37 6.3703	Período 2 6.37 6.368	Caos 5.81 5.816	37.12	6.19
	4	Período 2 6.37 6.375	Período 4 3.71 3.712	Período 8 3.38 3.379	26,91	4.49
	8	Período 4 3.96 3.967	Período 1 0.0 0.03	Período 1 0.0 0.05	8.06	1.34
	Totais	33.39	20.21	18.5	72.09	
	Médias	5.56	3.37	3.08		12.02

A partir deste ponto, basta realizar os cálculos das Equações A.1 a A.21 para obter a tabela ANOVA para o experimento fatorial com dois fatores.

Os cálculos para a obtenção da tabela ANOVA em experimentos fatoriais com mais de três fatores são muitos extensos e requerem uma aritmética tediosa. No entanto, atualmente há programas estatísticos que geram automaticamente a tabela ANOVA a partir dos dados do experimento (uma Tabela A.1 para 4 fatores). Nesta tese foi utilizado o programa *Minitab* para a geração das tabelas ANOVA para o experimento fatorial com quatro fatores (k_r, k_f, ρ e a).

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)